



ЕВГЕНИЙ
КУНИН

ВЫДАЮЩАЯСЯ КНИГА МИРОВОГО ЭКСПЕРТА
В ОБЛАСТИ КОМПЬЮТЕРНОЙ
И ЭВОЛЮЦИОННОЙ БИОЛОГИИ.

ЛОГИКА
СЛУЧАЯ

О ПРИРОДЕ
И ПРОИСХОЖДЕНИИ
БИОЛОГИЧЕСКОЙ
ЭВОЛЮЦИИ

EUGENE V. KOONIN

В этой амбициозной книге Евгений Кунин освещает переплетение случайного и закономерного, лежащих в основе самой сути жизни. В попытке достичь более глубокого понимания взаимного влияния случайности и необходимости, двигающих вперед биологическую эволюцию, Кунин сводит воедино новые данные и концепции, намечая при этом дорогу, ведущую за пределы синтетической теории эволюции. Он интерпретирует эволюцию как стохастический процесс, основанный на заранее непредвиденных обстоятельствах, ограниченный необходимостью поддержки клеточной организации и направляемый процессом адаптации. Для поддержки своих выводов он объединяет между собой множество концептуальных идей: сравнительную геномику, проливающую свет на предковые формы; новое понимание шаблонов, способов и непредсказуемости процесса эволюции; достижения в изучении экспрессии генов, распространенности белков и других фенотипических молекулярных характеристик; применение методов статистической физики для изучения генов и геномов и новый взгляд на вероятность самопроизвольного появления жизни, порождаемый современной космологией. Логика случая демонстрирует, что то понимание эволюции, которое было выработано наукой XX века, является устаревшим и неполным, и обрисовывает фундаментально новый подход – вызывающий, иногда противоречивый, но всегда основанный на твердых научных знаниях.

- [Евгений Кунин](#)
 -
 - [Предисловие автора к русскому переводу](#)
 -
 - [Введение. На пути к новому синтезу эволюционной биологии \[1\]](#)
 - [Глава 1. Основы эволюции: Дарвин и синтетическая теория эволюции](#)
 -
 - [Дарвин и первая синтетическая теория: величие замысла, ограничения и проблемы](#)
 - [Генетика и «черный день» дарвинизма](#)
 - [Популяционная генетика, теорема Фишера, адаптивные ландшафты, генетический дрейф и «эволюционная тяга»](#)
 - [Положительный и очищающий \(отрицательный\) отбор: классификация форм отбора](#)
 - [Синтетическая теория эволюции](#)
 - [Краткий обзор главы](#)
 - [Рекомендуемая дополнительная литература](#)
 - [Глава 2. От синтетической теории эволюции к эволюционной геномике: различные механизмы и пути эволюции](#)
 -
 - [Репликация цифровых носителей информации: центральный принцип биологии и необходимое и достаточное условие эволюции](#)
 - [Молекулярная эволюция и молекулярная филогенетика](#)
 - [Нейтральная теория молекулярной эволюции](#)
 - [Измерение естественного отбора сравнением последовательностей ДНК](#)
 - [Эгоистичные гены, мусорная ДНК и мобильные элементы](#)

- [Эволюция путем дубликации генов и геномов: ортологи и паралоги](#)
- [Прерывистое равновесие и несостоятельность градуализма](#)
- [Пандативы, экзаптация, эволюция как ремесленник и ошибочность панглоссианской парадигмы эволюции](#)
- [Эволюция в мире микробов и вирусов и трехдомное древо жизни](#)
- [Вирусы и рождение эволюционной геномики](#)
- [Эндосимбиоз](#)
- [Канализация и устойчивость в эволюции](#)
- [Краткий обзор и перспектива](#)
- [Рекомендуемая дополнительная литература](#)
- [Глава 3. Сравнительная геномика: эволюционирующие геномные ландшафты](#)
 - [Важность перехода к геномике](#)
 - [Эволюция геномных ландшафтов](#)
 - [Поразительное разнообразие геномов](#)
 - [Древние гены составляют в геноме большинство и имеют отчетливую эволюционную судьбу](#)
 - [Мультидомные белки и сложность связей ортологов](#)
 - [Контраст между эволюционной пластичностью генома и стабильностью индивидуальных генов](#)
 - [Геномные ландшафты: распределение эволюционных ограничений по разным классам сайтов в геноме](#)
 - [Вселенная генов](#)
 - [Минимальные наборы генов, замещение неортологичных генов \(ЗНОГ\) и ускользящее незаменимое ядро жизни](#)
 - [Единицы эволюции и фрактальная структура генетической вселенной](#)
 - [Элементарные события геномной эволюции](#)
 - [Краткий обзор главы](#)
 - [Рекомендуемая дополнительная литература](#)
- [Глава 4. Геномика, системная биология и универсалии эволюции: эволюция генома как феномен статистической физики](#)
 - [Взаимосвязь между эволюционными и фенотипическими параметрами, универсалии эволюции генов, белков и геномов и физическая модель эволюционного процесса](#)
 - [Почти нейтральные сети и белковая эволюция](#)
 - [Геномная эволюция путем дубликации генов, модель рождения и смерти гена и универсальное распределение численности паралогичных семейств](#)
 - [Структура и эволюция сетей: всеобщность степенного закона и стоящие за ним фундаментальные процессы](#)
 - [Разбиение генома по биологическим функциям: универсальный степенной закон](#)
 - [Стохастичность, нейтральность и отбор в эволюции](#)
 - [Краткий обзор и перспектива: о природе эволюционного процесса](#)
 - [Рекомендуемая дополнительная литература \[46\]](#)
- [Глава 5. Сетевая геномика мира прокариот: вертикальные и горизонтальные потоки генов, мобиломы и динамика пангеномов](#)
 -

- [Размер и общая организация бактериальных и архейных геномов](#)
- [Пространство-время прокариот и его эволюция](#)
 - [Фрактальное пространство-время генома, пангеномы и кластеризация прокариот](#)
 - [Эволюционная динамика архитектуры генома прокариот: опероны, суперопероны и сети соседствующих генов](#)
 - [Регуляция экспрессии генов и передачи сигналов у бактерий и архей: от базовой схемы оперона к сверхоперонам, регулонам и сложным сетям](#)
- [Горизонтальный перенос генов – определяющий процесс в эволюции прокариот](#)
 - [Повсеместное распространение ГПГ в мире прокариот](#)
- [Мобилом прокариот](#)
- [Незаменимость ГПГ для эволюции прокариот](#)
- [Горизонтальный перенос генов, универсальные законы геномики и хорошо перемешанный резервуар прокариотических генов](#)
- [Характерные геномные профили бактерий и архей с различными стилями жизни и неизоморфное отображение генного и функционального пространств](#)
- [Археи и бактерии в свете сравнительной геномики: как же быть с прокариотами?](#)
- [Краткий обзор и перспектива](#)
- [Рекомендуемая дополнительная литература](#)
- [Глава 6. Филогенетический лес и поиск неуловимого древа жизни в век геномики](#)
 - [Очень краткая история древа жизни](#)
 - [Фундаментальные единицы эволюции и присущая им древовидная природа](#)
 - [Лес жизни и почти универсальные филогенетические деревья](#)
 - [В глубь леса жизни: Большой взрыв или сжатый кладогенез?](#)
 - [Разделение эволюции прокариот на древовидный и сетевидный компоненты](#)
 - [Древовидная эволюция или неслучайный горизонтальный перенос генов?](#)
 - [Краткий обзор и перспектива](#)
 - [Рекомендуемая дополнительная литература](#)
- [Глава 7. Происхождение эукариот: эндосимбиоз, удивительная история интронов и исключительная важность единичных событий в эволюции](#)
 - [Эукариотическая клетка, ее внутренняя архитектура и пропасть между прокариотической и эукариотической клеточной организацией](#)
 - [Эндосимбиоз, митохондрии, гидрогеносомы и пластиды](#)
 - [Супергруппы эукариот и корень эукариотического эволюционного древа](#)
 - [Сложный, «растущий» LECA и темные века эволюции эукариот](#)
 - [Реконструкция LECA](#)
 - [Стволовая фаза: темные века эволюции эукариот](#)
 - [Корни эукариот среди архей и бактерий](#)
 - [Поиск архейного и бактериального «родителей» эукариот](#)
 - [Происхождение ключевых функциональных систем эукариотической клетки](#)
 - [Эукариогенез: источники уникальной эукариотической клеточной организации](#)
 - [Сценарий симбиогенеза в сравнении с архезойным сценарием](#)
 - [Симбиотический сценарий эукариогенеза: происхождение ключевых эукариотических инноваций, инициированных эндосимбиозом](#)
 - [Удивительная история эукариотических интронов](#)
 - [Три домена жизни: за пределами дерева Вёзе](#)

- [Краткий обзор главы](#)
- [Рекомендуемая дополнительная литература](#)
- [Глава 8. Неадаптивная нулевая гипотеза эволюции генома и истоки биологической сложности](#)
 - [Эволюционная энтропия и сложность](#)
 - [Эффективный размер популяции как общая мера эволюционных ограничений: неадаптивная теория эволюции генома](#)
 - [Генная архитектура эукариот: наглядная демонстрация неадаптивной теории эволюции генома](#)
 - [От мусора к функциональности: важность ослабленного очищающего отбора для эволюции сложности](#)
 - [Оптимизация генома в качестве основного пути эволюции и сложность как геномный синдром](#)
 - [За пределами нулевой гипотезы: ограничения популяционно-генетического взгляда на эволюцию генома](#)
 - [Дарвиновский глаз, нередуцируемая сложность, экзаптация и конструктивная нейтральная эволюция](#)
 - [Краткий обзор и перспектива: неадаптивная эволюционная парадигма и переоценка концепции эволюционного успеха](#)
 - [Рекомендуемая дополнительная литература](#)
- [Глава 9. Ламарковский, дарвиновский и райтовский режимы эволюции, эволюция эволюционируемости, надежность биологических систем и созидательная роль шума в эволюции](#)
 - [Драма ламаркизма](#)
 - [Ламарковский, дарвиновский и райтовский режимы эволюции и критерии для обнаружения ламарковского наследования](#)
 - [Ламарковские и квазиламарковские явления в эволюции](#)
 - [Системы антивирусного иммунитета CRISPR-Cas у прокариот: демонстрация аутентичного механизма по Ламарку](#)
 - [Другие \(квази\)ламарковские системы, функционирующие по принципу CRISPR](#)
 - [Горизонтальный перенос генов: важная ламарковская составляющая](#)
 - [Стресс-индуцированный мутагенез и активизация мобильных элементов: квазиламарковский феномен](#)
 - [Континуум дарвиновских и ламарковских механизмов эволюции](#)
 - [Точность передачи информации в биологических системах и ее \(не\)адаптивная эволюция](#)
 - [Шум в биологических системах и его созидательная роль в эволюции](#)
 - [Эволюция эволюционируемости, надежность биологических систем и реализуемость эволюционного предвидения](#)
 - [Краткий обзор и перспектива](#)
 - [Рекомендуемая дополнительная литература](#)
- [Глава 10. Мир вирусов и его эволюция](#)
 - [Необыкновенное разнообразие и повсеместное распространение вирусов](#)
 - [Что такое вирус?](#)
 - [Разнообразие стратегий репликации-экспрессии среди вирусов](#)

- [Диапазон сложности геномов, их функционального содержания и разнообразие геномной архитектуры вирусов](#)
- [Метагеномика вирусов, экспериментальная вирусология, агенты переноса генов и повсеместное распространение вирусов](#)
- [Эволюция вирусов: полифилия, монофилия и гены-сигнатуры](#)
 - [Краткая естественная история вирусных генов](#)
 - [Гены – вирусные сигнатуры: сигналы из древнего мира вирусов](#)
- [Конкурирующие концепции происхождения и эволюции вирусов](#)
- [Непрерывность мира вирусов и связи с миром клеточных организмов](#)
- [Фундаментальная неизбежность паразитов](#)
- [Вечная гонка вооружений между хозяином и паразитом и эволюция систем защиты и взлома защиты](#)
- [Краткий обзор и перспектива](#)
- [Рекомендуемая дополнительная литература](#)
- [Глава 11. Последний универсальный общий предок, происхождение клеток и первичный резервуар генов](#)
 - [Сравнительно-геномная реконструкция генного репертуара LUCA](#)
 - [Неклеточный компартиментализированный LUCA\(S\): сообщество разнородных репликаторов и лаборатория ранней эволюции](#)
 - [Краткий обзор и перспектива](#)
 - [Рекомендуемая дополнительная литература](#)
- [Глава 12. Происхождение жизни. Возникновение трансляции, репликации, метаболизма и мембран: биологический, геохимический и космологический подходы](#)
 - [Происхождение репликации и трансляции и мир РНК](#)
 - [Цикл Дарвина – Эйгена](#)
 - [Изучение эволюции белковых доменов дает аргументы в пользу сложного мира РНК: взгляд сверху вниз](#)
 - [Рибозимы и мир РНК](#)
 - [Природа и происхождение генетического кода](#)
 - [Происхождение трансляции: ключевые идеи и модели \[129\]](#)
 - [Скептический обзор моделей эволюции репликации и трансляции](#)
 - [Происхождение жизни с точки зрения химии и геохимии](#)
 - [Радикальная альтернатива: космология вечной инфляции, переход от случайности к биологической эволюции в истории жизни и переоценка роли крайне редких событий в эволюции \[132\]](#)
 - [Краткий обзор и перспектива](#)
 - [Рекомендуемая дополнительная литература](#)
 - [Дополнение](#)
- [Глава 13. Постсовременное состояние эволюционной биологии](#)
 - [Плюрализм паттернов и процессов эволюции: смена концепций отбора, вариации и древа жизни](#)
 - [Роль и статус отбора](#)

- [Меняющиеся концепции вариации и конец градуализма](#)
 - [От древа жизни к паутине жизни](#)
 - [Неожиданная значимость простых физических и математических моделей для понимания эволюции: биологическая эволюция как предмет статистической физики](#)
 - [Воспроизводимость эволюции: детерминизм и стохастика эволюционного процесса](#)
 - [Сложное и неоднозначное соответствие генома и фенотипа](#)
 - [Заря экспериментальной эволюционной биологии](#)
 - [Дивный новый мир вирусов и прокариот](#)
 - [Империи и домены жизни](#)
 - [Парадокс биологической сложности, обманчивость прогресса и значение неадаптивных храповиков](#)
 - [Клетки и организмы как устройства для самовоспроизведения генов](#)
 - [Так почему же организмы столь сложно устроены?](#)
 - [Величайшая загадка происхождения жизни](#)
 - [Нужна ли и полезна ли новая теория биологической эволюции? Появится ли постсовременная синтетическая теория?](#)
 - [Рекомендуемая дополнительная литература](#)
 - [Дополнение](#)
 - [Приложение I. Философия постмодерна, метанарративы; природа и цели научных исследований](#)
 -
 - [Философия постмодерна, \(не\)доверие к метанарративам и \(не\)осуществимость синтеза](#)
 - [Вопросы «почему?» и семантические ловушки: что на самом деле мы говорим об эволюции?](#)
 - [Природа и цели науки: зачем вообще изучать эволюцию?](#)
 - [Приложение II. Эволюция космоса и жизни: вечная инфляция, теория «мира многих миров», антропный отбор и грубая оценка вероятности возникновения жизни \[150\]](#)
 - [Краткое введение в инфляционную космологию для неспециалистов](#)
 - [Вероятность случайного возникновения различных революционных систем в Н-области: грубая прикидка верхних пределов](#)
 - [Литература](#)
 - [Благодарности](#)
 - [Об авторе](#)
- [notes](#)
 - [1](#)
 - [2](#)
 - [3](#)
 - [4](#)
 - [5](#)
 - [6](#)
 - [7](#)
 - [8](#)
 - [9](#)
 - [10](#)

- [11](#)
- [12](#)
- [13](#)
- [14](#)
- [15](#)
- [16](#)
- [17](#)
- [18](#)
- [19](#)
- [20](#)
- [21](#)
- [22](#)
- [23](#)
- [24](#)
- [25](#)
- [26](#)
- [27](#)
- [28](#)
- [29](#)
- [30](#)
- [31](#)
- [32](#)
- [33](#)
- [34](#)
- [35](#)
- [36](#)
- [37](#)
- [38](#)
- [39](#)
- [40](#)
- [41](#)
- [42](#)
- [43](#)
- [44](#)
- [45](#)
- [46](#)
- [47](#)
- [48](#)
- [49](#)
- [50](#)
- [51](#)
- [52](#)
- [53](#)
- [54](#)
- [55](#)
- [56](#)
- [57](#)

- [58](#)
- [59](#)
- [60](#)
- [61](#)
- [62](#)
- [63](#)
- [64](#)
- [65](#)
- [66](#)
- [67](#)
- [68](#)
- [69](#)
- [70](#)
- [71](#)
- [72](#)
- [73](#)
- [74](#)
- [75](#)
- [76](#)
- [77](#)
- [78](#)
- [79](#)
- [80](#)
- [81](#)
- [82](#)
- [83](#)
- [84](#)
- [85](#)
- [86](#)
- [87](#)
- [88](#)
- [89](#)
- [90](#)
- [91](#)
- [92](#)
- [93](#)
- [94](#)
- [95](#)
- [96](#)
- [97](#)
- [98](#)
- [99](#)
- [100](#)
- [101](#)
- [102](#)
- [103](#)
- [104](#)

- [105](#)
- [106](#)
- [107](#)
- [108](#)
- [109](#)
- [110](#)
- [111](#)
- [112](#)
- [113](#)
- [114](#)
- [115](#)
- [116](#)
- [117](#)
- [118](#)
- [119](#)
- [120](#)
- [121](#)
- [122](#)
- [123](#)
- [124](#)
- [125](#)
- [126](#)
- [127](#)
- [128](#)
- [129](#)
- [130](#)
- [131](#)
- [132](#)
- [133](#)
- [134](#)
- [135](#)
- [136](#)
- [137](#)
- [138](#)
- [139](#)
- [140](#)
- [141](#)
- [142](#)
- [143](#)
- [144](#)
- [145](#)
- [146](#)
- [147](#)
- [148](#)
- [149](#)
- [150](#)
- [151](#)



Евгений Кунин

**Логика случая. О природе и происхождении
биологической эволюции**

«The Logic of Chance. The Nature and Origin of Biological Evolution»

Copyright © 2012 by Pearson Education, Inc.

© Перевод, издание на русском языке, ЗАО «Издательство Центрполиграф», 2014

© Художественное оформление, ЗАО «Издательство Центрполиграф», 2014

Все права защищены. Никакая часть электронной версии этой книги не может быть воспроизведена в какой бы то ни было форме и какими бы то ни было средствами, включая размещение в сети Интернет и в корпоративных сетях, для частного и публичного использования без письменного разрешения владельца авторских прав.

©Электронная версия книги подготовлена компанией ЛитРес (www.litres.ru)

Предисловие автора к русскому переводу

Сообщение о том, что группа энтузиастов, самоорганизовавшаяся через LiveJournal, начала работу над переводом этой книги, было для автора полной неожиданностью, конечно же приятной. В XXI веке вопрос о необходимости перевода научной литературы с английского на какие-либо другие языки, мягко говоря, неоднозначен. Научные тексты теперь публикуются по-английски, и умение их читать на этом языке – элементарное требование профессиональной пригодности. Научно-популярная литература – дело, конечно, совершенно иное. Эта книга не популярная, но и не типичная специализированная монография. В идеале этот текст рассчитан на широкие круги ученых разных специальностей, включая аспирантов и студентов старших курсов. Было бы, конечно, прекрасно, если бы вся эта читательская аудитория могла свободно прочесть оригинал, однако пока что это вряд ли реалистично. Самым же главным аргументом в пользу перевода стал для автора сам факт, что немалый коллектив переводчиков собрался в считанные дни. В этой ситуации автор считал своей почетной обязанностью прочесть и отредактировать весь текст перевода, конечно следя в первую очередь за фактической точностью.

Оригинал этой книги был опубликован осенью 2011 года, за два года до русского издания. Биологические исследования в наше время прогрессируют в беспрецедентном темпе, и за эти годы, естественно, накопилось множество важных новых результатов и было опубликовано немало серьезных статей, проливающих свет на фундаментальные проблемы эволюционной биологии, обсуждаемые в книге. Разумеется, новые соображения, только частично опубликованные, появились и у автора. Более того, многие читатели, включая переводчиков, и сам автор при редактировании перевода отметили неточности и неясности в изложении (к счастью, насколько автору известно, ни одна из них не может считаться серьезной ошибкой). Учесть все это в русском переводе было невозможно, но автор сделал попытку отразить наиболее важные уточнения и некоторые самые интересные научные новости в примечаниях к русскому изданию. Таких новых примечаний в итоге оказалось куда больше, чем ожидалось в начале работы над редактированием перевода (а могло быть и еще больше – автор высказывался только тогда, когда уж совсем не мог молчать). Автора это очень радует, поскольку наглядно иллюстрирует скорость прогресса современной эволюционной биологии. Несколько примечаний относятся скорее к переводу, поясняя те места в тексте, где английскую игру слов не удалось точно передать по-русски. Разумеется, эти примечания не могут претендовать на то, чтобы сделать книгу «вторым изданием», это именно перевод, но все же автор надеется, что эти небольшие дополнения повышают его ценность.

С точки зрения автора, основные идеи книги пока выдерживают проверку временем (пусть коротким в астрономическом исчислении, но не пренебрежимым, учитывая поразительную скорость накопления новых данных); во всяком случае, потребности что-либо радикально пересмотреть до сих пор не возникло. Более того, автору представляется, что прошедшее время только усилило потребность в концептуальном обобщении информации о разнообразии организмов и их геномов и об эволюционных процессах. Новый эволюционный синтез на основе данных геномики и системной биологии кажется важным и актуальным, как никогда раньше. Без такого обобщения как-либо осмыслить море наблюдений становится просто невозможным.

Конечно, важно подчеркнуть, что эта книга ни в коем случае не может претендовать на роль такого нового синтеза. Это всего лишь некий эскиз, попытка угадать контуры будущего здания. Даже оставляя в стороне принципиальную открытость науки и считая, что какие-то

этапы завершенности и подведения итогов в ней действительно существуют, по мнению автора, завершение нового синтеза эволюционной биологии – дело как минимум двух научных поколений. Слишком много еще остается неясного, и слишком много надо сделать, чтобы уложить гигантские массивы данных, производимые геномикой и системной биологией, в рамки стройных и обоснованных теорий и концепций. Пожалуй, главной задачей этой книги и было выявить те области эволюционной биологии, где традиционные представления не работают, наметить возможные пути к решениям и только в некоторых случаях предложить сами решения, конечно же предварительные. Насколько все это удалось, судить читателям.

Благодарности учителям, сотрудникам и многочисленным коллегам, с которыми довелось обсуждать рассматриваемые в книге проблемы, приведены в конце книги. Здесь же приятный долг автора – выразить искреннюю благодарность Георгию Юрьевичу Любарскому за идею коллективного перевода и его организацию, всем переводчикам и редакторам издательства за работу над русским вариантом и персонально одному из переводчиков, Валерию Анисимову, за ценные комментарии, в значительной степени учтенные в авторских примечаниях к переводу.

Моим родителям

Введение. На пути к новому синтезу эволюционной биологии [1]

Название настоящей работы связано с четырьмя замечательными книгами: романом Пола Остера «Музыка случая» (Auster, 1991), знаменитым трактатом Жака Моно по молекулярной биологии, эволюции и философии «Случай и необходимость» (Monod, 1972), книгой Франсуа Жакоба «Логика жизни» (Jacob, 1993) и, конечно, «Происхождением видов» Чарльза Дарвина (Darwin, 1859). Каждая из этих книг в своем роде затрагивает одну и ту же всеохватную тему: взаимосвязь произвола и порядка, случайности и необходимости в жизни и эволюции.

Лишь после того, как эта работа была завершена и находилась уже на последней стадии редактирования, я узнал о книге Джона Венна, выдающегося логика и философа из Кембриджа, который в 1866 году опубликовал труд «Логика случая: эссе об основах и структуре теории вероятности» (Venn, 1866). В этой работе Венн вводит частотную интерпретацию вероятности, остающуюся основой теории вероятности и статистики по сей день. Более всего Джон Венн известен, естественно, вездесущими диаграммами, им изобретенными. Я смущен тем, что не знал о работе Венна, когда начал эту книгу. С другой стороны, мне трудно представить более достойного предшественника.

Основным толчком к написанию этой книги было мое убеждение в том, что сейчас, через 150 лет после Дарвина и 40 лет после Моно, мы собрали достаточно данных и идей, чтобы выработать более глубокое и, вероятно, более удовлетворительное толкование принципиально важной взаимосвязи между случаем и необходимостью. Мой главный тезис состоит в том, что ограниченная различными факторами случайность лежит в самой основе всей истории жизни.

К работе над этой книгой автора подтолкнуло множество событий. Самым непосредственным стимулом для того, чтобы описать возникающий новый взгляд на эволюцию, была революция в исследовании геномов, которая началась в последней декаде XX века и продолжается по сей день. Возможность сравнивать последовательности нуклеотидов в геномах тысяч организмов самых разнообразных видов качественно изменила ландшафт всей эволюционной биологии. Наши выводы о вымерших, предковых формах жизни – уже не те смутные догадки, какими они были раньше (по крайней мере для организмов, окаменелости которых не были обнаружены). Сравнение геномов выявляет разнообразные гены, сохраненные в основных группах ныне живущих существ (в некоторых случаях, даже во всех или большинстве из них), и таким образом приносит нам невообразимое прежде богатство достоверной информации о предковых формах. К примеру, не будет преувеличением заявить, что у нас есть достаточно полное понимание основного генетического состава последнего общего предка всех бактерий, который, вероятно, жил около 3,5 миллиарда лет назад. Более древние предки видятся менее ясно, но определенные черты расшифрованы даже для них. Геномная революция не просто позволила осуществить уверенную реконструкцию генных наборов древних форм жизни. Еще важнее то, что она буквально перевернула центральную метафору эволюционной биологии (и, возможно, всей биологии) – древо жизни (ДЖ), показав, что эволюционные траектории отдельных генов несовместимо разные. Вопрос о том, должно ли быть ДЖ возрождено и если так, то в каком виде, остается предметом ожесточенных споров, которые являются одной из важных тем этой книги.

Я рассматриваю падение ДЖ как «метареволюцию», крупнейшее изменение всей концептуальной структуры биологии. Явно рискуя вызвать гнев многих за связь с вредоносной культурной тенденцией, я тем не менее называю эту главную перемену переходом к

постмодернистскому биологическому взгляду на жизнь [2]. По существу, этот переход вскрывает множественность паттернов и процессов эволюции, центральную роль непредсказуемых событий в эволюции живых форм [«эволюция как халтура» (evolution as tinkering)] и, в особенности, крушение панадаптационизма как парадигмы эволюционной биологии. Несмотря на наше непоколебимое восхищение Дарвином, мы должны низвести викторианский взгляд на мир (включая его обновленные версии, процветающие в XX столетии) в почтенные музейные залы, где ему самое место, и исследовать последствия смены парадигмы.

У этого переворота в эволюционной биологии есть еще один план. Сравнительная геномика и эволюционная системная биология (например, сравнительное изучение экспрессии генов, концентрации белка и других молекулярных характеристик фенотипа) выявили несколько общих закономерностей, которые проступают во всех клеточных формах жизни от бактерий до млекопитающих. Существование таких универсальных закономерностей подсказывает, что сравнительно простые молекулярные модели, сходные с теми, что используются в статистической физике, могут объяснить важные аспекты биологической эволюции; некоторые подобные модели, обладающие значительной предсказательной силой, уже существуют. Пресловутая «зависть к физикам», которая, кажется, беспокоит многих биологов (включая меня), может быть утолена недавними и предстоящими теоретическими изысканиями. Взаимодополняющие отношения между всеобщими тенденциями и непредсказуемостью конкретных результатов эволюции являются центральными для биологической эволюции и текущей революции в эволюционной биологии – и это еще одна ключевая тема настоящей книги.

Еще одна причина появления наброска новой синтетической эволюционной теории, который предлагается в этой книге, специфическая, в какой-то мере личная. Я получил высшее образование и окончил аспирантуру в Московском государственном университете (еще во времена СССР), в области молекулярной вирусологии. Моя кандидатская работа включала экспериментальное изучение репродукции полиовируса и родственных вирусов, крошечный геном которых представлен молекулой РНК. Я никогда не умел как следует работать руками, да и место и время были не лучшими для экспериментов, потому что даже простейшие реагенты было сложно достать. Сразу по завершении моей кандидатской мы с моим коллегой Александром Евгеньевичем Горбаленей принялись за иное направление в исследованиях, которое в то время казалось многим совершенно ненаучным. Это было «разглядывание последовательностей» – попытки предсказать функции белков, закодированных в крошечных геномах вирусов (это были единственные полные геномы, доступные в то время), исходя из последовательности их кирпичиков-аминокислот. Сегодня кто угодно может легко провести такой анализ, используя удобные программные средства, которые можно бесплатно загрузить из Интернета; естественно, осмысленная интерпретация результата все равно потребует обдумывания и навыка (здесь с тех пор ничего особенно не изменилось). В 1985 году, однако, практически не было ни компьютеров, ни программ. И все же с помощью наших коллег-программистов нам удалось разработать несколько довольно полезных программ (мы тогда набивали их на перфокартах). Львиная доля анализа производилась вручную (или, точнее, на глаз). Вопреки всем трудностям и невзирая на некоторые упущенные возможности, наши усилия в последующие пять лет были довольно успешны. Мы смогли превратить функциональные карты тех самых крошечных геномов из большей частью неисследованных территорий в весьма насыщенные геномные карты биологических функций. Большинство предсказаний было впоследствии подтверждено на опыте, хотя некоторые из них до сих пор еще в работе: лабораторные эксперименты занимают куда больше времени, чем компьютерный анализ. Уверен, что нашему успеху послужило раннее осознание очень простого, но

удивительно мощного основного принципа эволюционной биологии: если явственно различимый мотив в последовательности белка сохраняется в течение долгой эволюции, то он функционально важен, и чем он консервативнее, тем важнее функция. Этот принцип, в сущности вытекающий из простого здравого смысла, но конечно же строго следующий из молекулярной эволюционной теории, прекрасно служил нашим целям и, уверен, сделал из меня эволюционного биолога до конца моих дней. Я склонен перефразировать известное изречение великого эволюционного генетика Феодосия Добржанского: «Ничто в биологии не имеет смысла, кроме как в свете эволюции» (Dobzhansky, 1973) – еще более прямым образом: *биология есть эволюция*.

В те ранние дни эволюционной геномики мы с Сашей часто говорили о возможности того, что наши любимые РНК-вирусы являются прямыми потомками древнейших форм жизни. В конце концов, это маленькие и простые генетические системы, использующие только один вид нуклеиновой кислоты, и репликация у них напрямую связана с экспрессией через трансляцию геномной РНК. Конечно, это были вечерние разговоры, вовсе не связанные с нашими дневными попытками картирования функциональных доменов вирусных белков. Сегодня, 25 лет спустя, когда изучены сотни различных геномов вирусов и хозяев, идея того, что вирусы (или сходные с вирусами генетические элементы) могли быть главными на ранних стадиях эволюции жизни, из туманных предположений выросла в концепцию, совместимую с огромным массивом экспериментальных данных. По моему мнению, это наиболее многообещающая линия размышлений и анализа в исследованиях ранних стадий эволюции жизни.

Таковы различные концептуальные линии, которые неожиданно для меня сошлись в растущем осознании того, что наше понимание эволюции, а с ней и самой природы биологии навсегда отошло от взглядов, преобладавших в XX веке, которые на сегодня выглядят скорее наивными и довольно догматичными. В определенный момент желание сплести эти линии в подобие связной картины стало непреодолимым, и отсюда появилась эта книга.

Некоторые стимулы для написания этой книги пришли вовсе не из биологии, а из поразительных достижений современной космологии. Эти открытия не только подняли космологию до уровня настоящей физики, но и полностью перевернули наши представления о мире, и особенно о природе случайности и необходимости. Когда дело доходит до границ биологии, таких как проблема происхождения жизни, этот новый взгляд на мир невозможно не принимать во внимание. Физики и космологи все чаще ставят вопрос, почему в мире существует что-то, а не ничто, – не только как философскую, но и как физическую проблему, и исследуют возможные ответы в форме определенных физических моделей. Трудно не задаться тем же вопросом о биологическом мире, причем на более чем одном уровне: почему существует жизнь, а не просто растворы ионов и маленьких молекул? И коли жизнь существует, почему есть пальмы и бабочки, кошки и летучие мыши, а не только бактерии? Уверен, что эти вопросы могут быть поставлены прямым научным образом, и мне кажется, на них уже появляются правдоподобные, пусть и предварительные, ответы.

Последние достижения в физике высоких энергий и космологии послужили вдохновением для этой книги не только в прямом научном смысле. Многие ведущие теоретические физики и космологи оказались одаренными писателями популярных и научно-популярных книг (что заставляет задуматься о связи между абстрактным мышлением на высочайшем уровне и литературным талантом), которые передают эмоциональный подъем, возникающий в связи с новейшими открытиями о строении Вселенной, с восхитительной ясностью, изяществом и пылом. Современная волна такой литературы, совпадающая с революцией в космологии, началась с классической «Краткой истории времени» Стивена Хокинга (Hawking, 1988). С тех пор появились десятки различных прекрасных книг. Одна из них, сильнее прочих изменившая

мой собственный взгляд на мир, – великолепная короткая книга Александра Виленкина «Мир многих миров» (Vilenkin, 2007), но не менее важны были и работы Стивена Вайнберга (Weinberg, 1994), Алана Гута (Guth, 1998a), Леонарда Зюскинда (Susskind, 2006b), Шона Кэрролла (Carroll, 2010) и Ли Смолина (в спорной книге о «космическом естественном отборе»; Smolin, 2010). Эти книги гораздо больше, чем просто великолепные популяризации: каждая из них пытается представить связный, общий взгляд как на фундаментальную природу мира, так и на состояние науки, которая ее исследует. Каждая из этих картин мира уникальна, но во многих аспектах они идут бок о бок и дополняют друг друга. Каждая из них основана на строгой науке, но содержит и элементы экстраполяции и предположения, широких обобщений и, несомненно, противоречий. Чем больше я читал эти книги и размышлял о значении возникающего нового мировоззрения, тем сильнее мне хотелось сделать что-то подобное и в моей собственной области, молекулярной биологии. В какой-то момент, читая книгу Виленкина, я осознал, что, возможно, существует прямая и принципиально важная взаимосвязь между новыми взглядами на вероятность и случай, диктуемыми современной космологией, и происхождением жизни – вернее, происхождением биологической эволюции. Огромная роль случая в возникновении жизни на Земле, присутствующая в этой линии размышления, безусловно, неординарна и непременно многих смутит, но я чувствовал, что она не может быть оставлена без внимания, если мы хотим серьезно подойти к проблеме происхождения жизни.

Эта книга – мой собственный подход к описанию текущего состояния эволюционной биологии с позиций сравнительной геномики и системной биологии; следовательно, она неизбежно включает в себя не только установленные факты и подтвержденные теоретические модели, но и догадки и предположения. В этой книге я пытаюсь провести границу между фактами и догадками настолько четко, насколько возможно. Я хотел написать книгу в стиле вышеупомянутых превосходных научно-популярных книг по физике, но изложение заупрявилось и отказалось быть написанным таким образом. В результате текст получился гораздо более научным, чем это задумывалось поначалу, хотя он большей частью не слишком специализирован и описывает совсем немного методов, притом в весьма упрощенной манере. Одна важная оговорка: хоть книга и посвящена различным аспектам эволюции, она остается сборником глав по выбранным темам и ни в коей мере не претендует быть всеохватывающим трудом. Многие важные и популярные темы, такие как происхождение многоклеточных организмов или эволюция развития животных, совершенно осознанно не затронуты. Насколько возможно, я пытался придерживаться лейтмотива книги: взаимодействия между случаем и упорядоченными процессами. Еще один щекотливый момент связан со ссылками на литературу: попытайтесь я включить пусть не все, но хотя бы основные источники, библиография составила бы много тысяч ссылок. Я отказался от попытки это сделать с самого начала, и таким образом список литературы в конце книги является лишь небольшой выборкой относящихся к теме работ, и их отбор частично субъективен. Приношу мои искренние извинения коллегам, чья важная работа осталась неупомянутой.

Невзирая на все эти предостережения, я надеюсь, что обобщения и идеи, представленные здесь, будут интересны многим моим коллегам-ученым и студентам – не только биологам, но и физикам, химикам, геологам и всем интересующимся эволюцией и происхождением жизни.

Глава 1. Основы эволюции: Дарвин и синтетическая теория эволюции

В этой и следующей главах дается краткое описание современного состояния эволюционной биологии, какой она была до 1995 года, когда возникло новое направление науки – сравнительная геномика. Мягко говоря, это сложная задача – спрессовать полтора века исследований в области эволюции в две кратких главы. Тем не менее я полагаю, что мы можем начать с прямого вопроса «Какой же итог всех этих десятилетий научной работы?». Мы можем коротко и осмысленно сформулировать выводы синтетической теории до возникновения геномики, пусть и опуская большую часть подробностей.

В этих двух главах я попытался объединить историю и логику, однако мне, естественно, не удалось избежать некоторого произвола. В этой главе я прослежу развитие основ эволюционной биологии от «Происхождения видов...» Чарльза Дарвина до сформулированной в 1950-х годах синтетической теории эволюции (СТЭ). Во второй главе речь пойдет об идеях и открытиях, которые оказали влияние на понимание эволюции после окончательного оформления синтетической теории эволюции и до революции в геномике 1990-х.

Дарвин и первая синтетическая теория: величие замысла, ограничения и проблемы

Довольно странно думать о том, что мы только что отметили 150-летие со дня первой публикации «Происхождения видов...» (Darwin, 1859) и 200-летний юбилей самого Дарвина. Учитывая, какой глубокий и неизгладимый след оставило «Происхождение...» в науке, философии и человеческой мысли в целом (далеко за пределами только биологии), кажется, что 150 лет прошли очень быстро.

Что же такого исключительного и важного в том изменении миропонимания, которое вызвал труд Дарвина? Дарвин не открыл эволюцию (как иногда заявляют или чаще подразумевают, особенно в массовом сознании и публичных обсуждениях). Многие ученые до него, включая светил науки своего времени, были убеждены, что организмы изменяются во времени и эти изменения не случайны. Если не считать великих (и в некоторой степени легендарных) древнегреческих философов Эмпедокла, Парменида и Гераклита и их индийских современников, обсуждавших поразительные пророческие идеи (хоть и странным для нас образом совмещенные с мифологией) о процессах изменения в природе, у Дарвина было много предшественников в XVIII и начале XIX века. В последующих изданиях «Происхождения...» Дарвин с присущими ему непредвзятостью и великодушием признал их вклад. Его дед, Эразм Дарвин, и знаменитый французский ботаник и зоолог Жан-Батист Ламарк (Lamarck, 1809) написали толстые фолианты об эволюции ^[3]. Ламарк даже предложил ясное объяснение действия механизма, который, как он считал, закрепляет эволюционные изменения. Более того, знаменитый учитель и друг Дарвина, великий геолог Чарльз Лайель, писал о «борьбе за выживание», в которой всегда выигрывает более плодовитый. И конечно, общеизвестно, что в то же самое время молодой современник Дарвина, Альфред Рассел Уоллес, предложил в целом идентичную концепцию эволюции и ее механизма.

Однако, несмотря на достижения всех эволюционистов более раннего периода, именно Дарвин в «Происхождении...» заложил основу современной биологии и навсегда изменил научное представление о мире. Что же определило уникальность и исключительную значимость работы Дарвина? Рассматривая его достижение спустя 150 лет, мы можем выделить три крупных обобщения:

- Дарвин представил свой взгляд на эволюцию исключительно с позиции натуралиста и рационалиста, не привлекая к объяснению никакие телеологические силы или стремление к совершенствованию (или прямо указывая на некоего создателя), как обычно поступали теоретики того времени.
- Дарвин предложил конкретный, прямой и доходчивый механизм эволюции, представляющий собой взаимодействие между наследственной изменчивостью и естественным отбором, в целом описываемое как выживание наиболее приспособленных.
- Дарвин смело расширил идеи эволюции на всю историю земной жизни, которая, как он полагал, может быть представлена величественным древом (знаменитая единственная иллюстрация в «Происхождении...»), и даже утверждал, что все существующие формы жизни происходят от единого общего предка.

Общая и обладающая огромной предсказательной силой модель эволюции, предложенная Дарвином, явилась резким контрастом к эволюционным идеям его предшественников, особенно

Ламарка и Лайеля, которые рассматривали преимущественно или даже исключительно внутривидовые эволюционные изменения. Четвертое значительное достижение Дарвина связано не столько с научным содержанием его работы, сколько с формой ее изложения. Главным образом в связи с вполне понятной срочностью, вызванной соперничеством с Уоллесом, Дарвин представил свой труд в виде небольшой и легко читаемой даже для неспециалиста книги, которая, несмотря на это, содержала скрупулезно и тщательно собранные доводы. Благодаря этим принципиальным достижениям, Дарвин не просто опубликовал очередную книгу об эволюции, но полностью изменил лицо науки. Сразу же после публикации «Происхождения...» большинство биологов и даже просто образованная часть общества признали эту работу как заслуживающее доверие естественно-научное объяснение возникновения многообразия форм жизни, и это послужило динамичной основой для дальнейших теоретических построений [4].

Рассматривая труд Дарвина с более отвлеченной позиции, которая является основной в этой книге, необходимо особо отметить, что Дарвин, похоже, первым обнаружил определяющее взаимодействие между случаем и направленностью (неизбежностью) в эволюции. В соответствии с идеей Дарвина, изменчивость почти полностью случайна, в то время как отбор является направленным и создает сложность. В этом Дарвин полностью противоположен Ламарку, который, в сущности, изгнал случайность из своей картины мира. В данной книге мы будем периодически возвращаться к этому ключевому конфликту мировоззрений.

Конечно, надо отдать должное предшественникам Дарвина – геологам и эволюционным биологам, однако Дарвин, несомненно, был первым ученым, который включил возможность эволюционных изменений (и, косвенно, происхождение) всей Вселенной в сферу явлений природы, подлежащих рациональному изучению. Другими словами, Дарвин положил начало научному изучению стрелы времени – то есть асимметричных во времени, необратимых процессов. Таким образом, он подготовил почву не только для развития биологии, но также для создания современной физики. Я полагаю, что знаменитый физик Людвиг Больцман, основатель статистической термодинамики и автор современной концепции энтропии, имел все основания назвать Дарвина «великим физиком», что может показаться парадоксальным, учитывая, что Дарвин крайне мало знал физику и математику. Пожалуй, и наш современник, философ Дэниел Деннет не так уж преувеличил, утверждая, что дарвиновская идея естественного отбора – величайшая идея в истории человечества (Dennett, 1996).

Конечно, эволюционное учение Дарвина со времени публикации «Происхождения...» и по меньшей мере до конца XIX века сталкивалось с острыми проблемами, всерьез беспокоившими Дарвина и на тот момент казавшимися непреодолимыми большинству ученых. Во-первых, значительную трудность представляло собой определение возраста Земли, который во времена Дарвина был существенно занижен. Даже не принимая во внимание религиозные мифы о сотворении мира, наиболее точно возраст Земли, по мнению физиков XIX века (в частности, лорда Кельвина), оценивался в 100 миллионов лет. Такого промежутка времени было явно недостаточно для эволюции жизни в том виде, в каком ее представил Дарвин, то есть путем постепенного накопления небольших изменений. В целом, действительно, 100 миллионов лет очень мало для эволюции жизни в ее нынешнем многообразии, хотя никто в XIX веке не мог количественно оценить скорость дарвиновской эволюции. Эта проблема разрешилась спустя 20 лет после смерти Дарвина. После открытия радиоактивности в начале XX века ученые подсчитали, что охлаждение Земли от первичного раскаленного состояния заняло миллиарды лет, то есть примерно столько, сколько, по предположению Дарвина, требовалось для эволюции с помощью естественного отбора.

Во-вторых, еще больше вопросов вызывали механизмы наследственности и так называемый кошмар Дженкина. Так как во времена Дарвина еще не существовало теории дискретных

наследственных детерминант (кроме малоизвестных статей Менделя), то было неясно, каким образом полезное благоприобретение может сохраниться в поколениях и закрепиться в эволюционирующей популяции, не растворяясь и не теряясь. Очевидно, сам Дарвин не обратил внимания на эту проблему в своей теории, когда писал «Происхождение...», однако о ней сообщил Дарвину необычайно критичный читатель его работы, инженер Дженкин. Оглядываясь назад, трудно понять, почему Дарвин (или Дженкин, или Гексли) не принял в расчет решение, предлагаемое Менделем. Вместо этого Дарвин выдвинул куда более странное объяснение, так называемую теорию пангенеза, которую даже он сам, по-видимому, не принимал всерьез. Противоречие было устранено с рождением (или, вернее, повторным рождением) генетики, хотя поначалу ее значение для дарвинизма ^[5] было неочевидным (см. следующую главу).

Третья проблема, которую Дарвин полностью осознавал и блестяще исследовал, – это эволюция сложных структур (*органов*, по терминологии Дарвина), для работы которых необходимо соединение множества частей. Такие сложные органы представляли собой классическую головоломку для эволюционной биологии, которая в XX веке была выразительно названа неупрощаемой сложностью ^[6]. Конечно, сразу непонятно, как может происходить эволюция таких органов путем естественного отбора, если считать, что отдельные части органа или «частично укомплектованный» орган не функциональны. Дарвин решительно обратился к этой проблеме в одном из самых известных отрывков «Происхождения...», сценарии эволюции глаза. Он предложил логически безупречное, убедительное и неординарное решение: Дарвин предположил, что эволюция сложных органов идет через серию промежуточных стадий, каждая из которых частично выполняет функцию развивающегося сложного органа. Таким образом, эволюция глаза, по Дарвину, начинается с простого светочувствительного участка, через примитивные постепенно усложняющиеся структуры, подобные глазу, к полноценным, функциональным сложным глазам членистоногих и позвоночных. Необходимо отметить, что примитивные светочувствительные структуры, похожие на те, существование которых предположил Дарвин исходя из общих предположений, были впоследствии обнаружены, что по крайней мере частично подтверждает его сценарий и показывает, что в этом случае неупрощаемость сложной структуры иллюзорна. Однако, несмотря на убедительность схемы, предложенной Дарвином, к ней следует относиться трезво, как к частично подтверждаемому, но все же гипотетическому сценарию эволюции одного конкретного органа. Предположение Дарвина показало одну из возможных траекторий эволюции сложной структуры, но не решило главную проблему в целом. Эволюция сложных структур на разных уровнях является центральным вопросом биологии, поэтому мы будем возвращаться к нему много раз в этой книге.

Четвертый сложный вопрос дарвинизма является и самым глубоким. Эта главная проблема имеет непосредственное отношение к названию книги Дарвина и к подразумеваемой основной ее теме, то есть к происхождению видов, и, в общем смысле, к крупным эволюционным событиям, которые в настоящее время носят собирательное название макроэволюция. В значительном отрыве от названия книги те неоспоримые примеры эволюции, которые представил Дарвин, относились к возникновению новых внутривидовых различий, а не новых видов, не говоря уже о новых таксонах более высокого уровня. Эта проблема сохранялась долго после смерти Дарвина и существует даже сейчас, хотя частично она была решена сначала прогрессом палеонтологии, затем развитием теории видообразования при поддержке биогеографических данных, а затем, наиболее убедительно, сравнительной геномикой (см. гл. 2 и 3). К чести Дарвина и в отличие от критиков эволюции по сей день, он твердо стоял на своем перед лицом всех трудностей, благодаря своей непоколебимой вере в то, что, несмотря на возможные пробелы в его теории, ей нет никакой разумной альтернативы. Единственным

слабым местом Дарвина оказалось включение неправдоподобной модели пангенеза в последующие издания «Происхождения...» как заплатки для маскировки кошмара Дженкина.

Генетика и «черный день» дарвинизма

Существует легенда, что Дарвин прочитал работу Менделя, но не нашел ее интересной (возможно, из-за ограниченного знания немецкого языка). Сложно предположить, насколько изменилась бы история биологии, если бы Дарвин использовал идеи Менделя, которые теперь нам кажутся предельно простыми. Однако этого не произошло.

Еще удивительнее, что сам Мендель, очевидно хорошо знакомый с «Происхождением...» [7], не рассматривал свое открытие в контексте теории Дарвина. Ожидать установления этой жизненно важной связи пришлось не только до возрождения генетики на заре XX века, но также до появления популяционной генетики в 1920-х годах. Повторное открытие механизма наследования и рождение генетики дало мощный толчок развитию дарвинизма, так как выявление дискретных носителей наследственности устраняло кошмар Дженкина. В связи с этим совершенно парадоксален тот факт, что первой реакцией большинства биологов на открытие генов было мнение, что генетика опровергает теорию Дарвина, хотя при этом никто из серьезных ученых не отвергал реальность эволюции. Основной причиной кажущейся несовместимости дарвинизма и генетики было то, что основатели генетики, в частности Хуго де Фриз, наиболее плодотворный ученый из трех биологов, переоткрывших законы Менделя, рассматривали мутации генов как прерывистые, скачкообразные наследственные изменения, противоречащие постепенной эволюции в теории Дарвина. Мутации с малым фенотипическим эффектом считались неотъемлемой чертой дарвинизма, в полном соответствии с «Происхождением...». Поэтому де Фриз полагал, что его теория мутаций «антидарвинистская». Таким образом, столетний юбилей Дарвина, а также 50-летие публикации «Происхождения...» в 1909 году были далеко не триумфальными, даже на фоне резкого роста генетических исследований и введения термина «ген» Вильгельмом Йогансеном в том же году.

Популяционная генетика, теорема Фишера, адаптивные ландшафты, генетический дрейф и «эволюционная тяга»

Основы крайне важного синтеза дарвинизма и генетики были заложены в конце 1920-х – начале 1930-х годов тремя выдающимися генетиками-теоретиками – Рональдом Фишером, Сьюэлом Райтом и Дж. Б. С. Холдейном. Основываясь на точных математических и статистических расчетах, они создали идеализированную модель эволюции в биологической популяции. Вероятно, великий ученый-статистик Фишер первым обратил внимание, что генетика никоим образом не противоречит дарвинизму, а, напротив, предоставляет естественный и твердый фундамент для теории дарвиновской эволюции. Фишер обобщил свои выводы в исторической работе 1930 года «Генетическая теория естественного отбора» (Fisher, 1930), пожалуй, втором по значимости для эволюционной биологии труде после дарвиновского «Происхождения...» [8]. Это стало началом блистательного возрождения дарвинизма, позже получившего название современный синтез (термин, используемый преимущественно в США), или неodarвинизм (в британской и европейских традициях) [9].

Нет ни надобности, ни практической возможности излагать здесь основы популяционной генетики [10]. Можно, однако, лаконично представить некоторые обобщения, имеющие отношение к остальной части обсуждения современной эволюционной биологии. Пусть и поверхностное, но такое резюме здесь будет существенно. По сути, основатели популяционной генетики осознали простой факт, что эволюция не действует на изолированные организмы или абстрактные виды, а направлена на конкретные группы скрещивающихся особей, называемые популяциями. Размер и структура эволюционирующей популяции в большой степени определяют направление и результат эволюции. В частности, Фишер сформулировал и доказал фундаментальную теорему естественного отбора (известную как теорема Фишера), в которой утверждается, что интенсивность отбора (и, следовательно, скорость эволюции путем отбора) пропорциональна величине генетической дисперсии по приспособленности эволюционирующей популяции, которая, в свою очередь, пропорциональна эффективному размеру популяции.

В табл. 1–1 собраны основные определения и уравнения, описывающие эффекты мутаций и давления отбора на устранение или закрепление мутантных аллелей в зависимости от эффективного размера популяции. Качественная суть этих уравнений в том, что при одинаковой скорости мутаций в популяции большего эффективного размера отбор более интенсивный. В таких популяциях даже мутации с небольшим положительным коэффициентом отбора («слегка» благоприятные мутации) закрепляются быстро. С другой стороны, мутации даже с очень маленьким отрицательным коэффициентом селекции («слегка» вредные мутации) быстро устраняются. Данный эффект был строго сформулирован в теореме Фишера.

Таблица 1–1. **Фундаментальное соотношение, описывающее роль отбора и генетический дрейф в эволюции популяции**

Почти нейтральная эволюция, управляемая дрейфом

$$1/Ne \gg |s|$$

Эволюция, управляемая отбором

$$1/Ne \ll |s|$$

Комбинированный режим, при сравнимом вкладе дрейфа и отбора

$$1/Ne \approx |s|$$

Ne : эффективный размер популяции (как правило, значительно меньше, чем общее количество особей в популяции, так как не все особи могут дать жизнеспособное потомство)

s : коэффициент отбора, или влияние мутации на приспособленность:

$$s = F_A - F_a$$

F_A, F_a : величина приспособленности каждого из двух аллелей гена

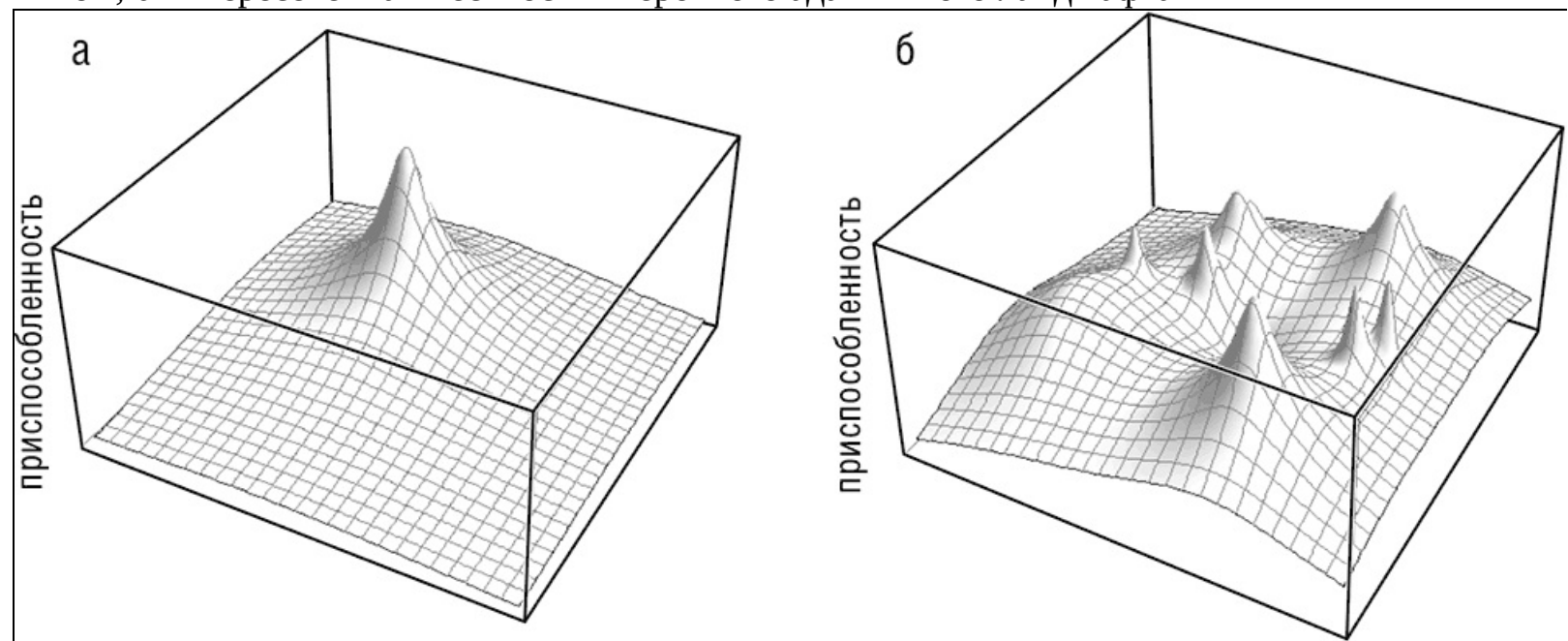
$s > 0$: полезная мутация

$s < 0$: вредная мутация

Из теоремы Фишера следует, что при эволюции, направляемой только естественным отбором, *средняя приспособленность популяции не может уменьшаться* (если, конечно, популяция собирается выжить). Пожалуй, наилучшим образом это можно представить с помощью образа «адаптивного ландшафта», который впервые был предложен другим отцом-основателем популяционной генетики, Сьюэлом Райтом. Райт создал этот чрезвычайно удачный образ в ответ на просьбу своего научного руководителя представить результаты математического анализа отбора в приемлемой для биологов форме. Благодаря своей простоте и изяществу это представление адаптивной эволюции сохраняет свою ценность по сей день и стимулировало многочисленные исследования, в результате которых появились более сложные и менее интуитивно понятные адаптивные ландшафты, в том числе и многомерные (Gavrilets, 2004) [\[11\]](#). В соответствии с теоремой Фишера популяция, эволюция которой идет только за счет отбора (строго говоря, популяция бесконечного размера – такие популяции, естественно, не существуют, но являются удобной абстракцией, часто используемой в популяционной генетике), никогда не будет двигаться вниз по адаптивному ландшафту (см. рис. 1–1). Легко представить, что адаптивный ландшафт, как и обычный ландшафт, может иметь самую различную форму. При определенных обстоятельствах ландшафт может быть очень гладким, с единственным пиком, соответствующим глобальному адаптивному максимуму (иногда такой ландшафт образно называют «гора Фудзияма» (см. рис. 1–1 а)). Реальный ландшафт, однако, неровный и содержит многочисленные пики различной высоты, разделенные долинами (см. рис. 1–1 б). Формально, согласно теореме Фишера (и в целом, в соответствии с теорией Дарвина), популяция, эволюционирующая с помощью отбора, может только подниматься вверх и, таким образом, достигнуть только локального пика, даже если его высота значительно меньше, чем высота глобального пика (см. рис. 1–1 а). Теория Дарвина и СТЭ утверждают, что движение популяции через долины запрещено, так как неизбежно подразумевает фазу спуска. Однако развитие популяционной генетики и ее применение к эволюционным процессам

изменило эту упорядоченную картину, привнеся в нее понятие «дрейфа генов», ключевую идею эволюционной биологии, которую также предложил Райт.

Рис. 1–1. Адаптивные ландшафты: а– «гора Фудзияма» с единственным (глобальным) пиком; б– «пересеченная местность» неровного адаптивного ландшафта



Как подчеркивалось ранее, Дарвин признавал важную роль случайности в эволюции, но эта роль была ограничена только одной частью эволюционного процесса: появлением изменений (в современной терминологии – мутаций). В остальном эволюция рассматривалась как строго детерминистский процесс, где отбором закрепляются выгодные мутации, а все прочие мутации устраняются без какого-либо вреда для дальнейшего существования популяции. Однако при рассмотрении популяции в динамике картина значительно меняется. Основатели количественной популяционной генетики отразили в простых формулах зависимость интенсивности отбора от размера популяции и частоты мутаций (см. табл. 1–1 и рис. 1–2). Отбор эффективен в большой популяции, и мутация, несущая незначительное преимущество, почти наверняка закрепится (в популяции бесконечного размера закрепляется мутация с бесконечно малым положительным коэффициентом отбора). Райт понял, что в малой популяции, особенно при низкой частоте мутаций, эволюционный процесс идет по-другому. В такой популяции решающую роль играет дрейф генов, с помощью которого случайным образом часто закрепляются нейтральные и даже вредные (но, конечно, не летальные) мутации. Очевидно, с помощью генетического дрейфа эволюционирующая популяция может избежать однонаправленного подъема по адаптивному ландшафту и может спускаться (см. рис. 1–2) [12]. Преимущественно это выражается в движении вниз и последующем вымирании, однако если долина, отделяющая один локальный пик от другого, возможно даже более высокого, достаточно узкая, становится возможным переход через нее и последующее восхождение на более высокую вершину (см. рис. 1–2). Введение понятия генетического дрейфа в изучение эволюции является центральным в моем рассказе. Это новый уровень проявления случая. Хотя Дарвин и его ближайшие последователи видели роль случая в появлении наследуемых изменений (мутаций), дрейф вводит случайность на следующей стадии, то есть при закреплении этих изменений, забирая у отбора часть ответственности. В этой книге я исследую, насколько значимой может быть роль дрейфа в различных ситуациях в ходе эволюции.

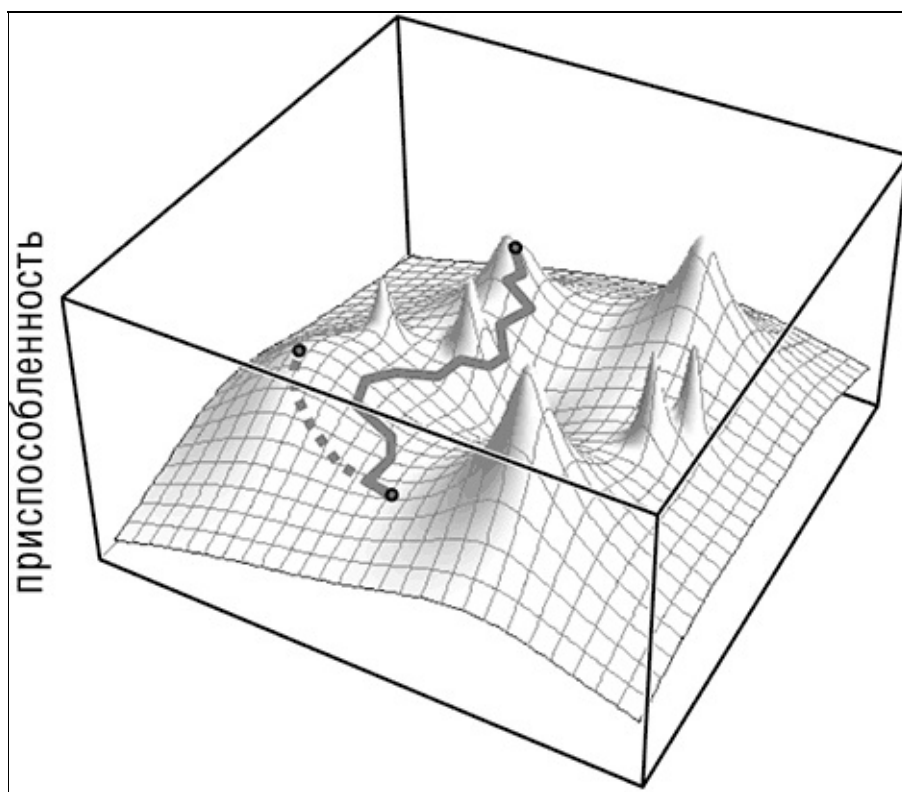


Рис. 1–2. Эволюционные траектории на неровном адаптивном ландшафте. Пунктирной линией обозначается эволюционная траектория при высоком значении эффективного размера популяции. Сплошной линией обозначается эволюционная траектория при низком значении эффективного размера популяции.

Джон Мейнард Смит и, позднее, Джон Гиллеспи разработали теорию и компьютерные модели для демонстрации существования особого режима нейтральной эволюции, который слабо зависит от эффективного размера популяции и актуален даже в популяции бесконечного размера с сильным отбором. Этот способ нейтрального закрепления мутаций стал известен как «генетическая тяга» и относится к ситуациям, в которых одна или несколько нейтральных или даже умеренно вредных мутаций распространяются в популяции и в конечном итоге закрепляются, будучи связанными с полезной мутацией. Иными словами, нейтральные или вредные аллели «двигаются в одной повозке» вместе с полезным аллелем (Barton, 2000). Похоже, что некоторые данные и модели популяционной генетики свидетельствуют, что «движение в одной повозке» даже важнее для эволюции популяции с половым размножением, чем дрейф. Очевидно, что эффект «езды в одной повозке» обусловлен совокупным воздействием естественного отбора и нейтральной изменчивости в различных участках генома и, в отличие от дрейфа, может происходить даже в популяции бесконечно большого эффективного размера (Gillespie, 2000).

За счет эффекта «движения в одной повозке» даже в больших популяциях могут закрепляться умеренно вредные мутации, что, соответственно, дает этой популяции возможность пересекать долины адаптивного ландшафта.

Положительный и очищающий (отрицательный) отбор: классификация форм отбора

Дарвин думал о естественном отборе в первую очередь с точки зрения закрепления благоприятных изменений. Он понимал, что эволюцией отсеиваются вредные изменения, но не интерпретировал эту ликвидацию в одной плоскости с естественным отбором. С развитием СТЭ понятие отбора было расширено за счет включения «очищающего» (отрицательного) отбора, который в некоторых фазах эволюции оказывается более распространенным (на самом деле на порядок более распространенным), чем «дарвиновский» положительный отбор. По сути, очищающий отбор – это просто элиминация неприспособленных особей. Тем не менее выделение этого процесса в особую форму отбора представляется оправданным и важным, потому что оно подчеркивает ключевую роль элиминации в формировании (сдерживании) биологического разнообразия на всех уровнях. Проще говоря, изменение допускается, только если оно не наносит существенного вреда никаким из выживающих особей. Интересен и открыт вопрос, до какой степени эти ограничения фактически сужают пространство, доступное для эволюции, и я коснусь этого вопроса позднее (см., в частности, гл. 3, 8 и 9).

Тонкая, но важная разница существует между очищающим отбором и *стабилизирующим отбором*, который является еще одной из форм отбора, которая действует на распределение частот отличительных признаков. Таким образом, выделяются следующие формы отбора: стабилизирующий отбор, основанный в первую очередь на очищающем отборе, движущий отбор, обусловленный положительным (дарвиновским) отбором, и более экзотические режимы дизруптивного и балансирующего отбора, которые являются результатом сочетания многочисленных ограничений (см. рис. 1–3).

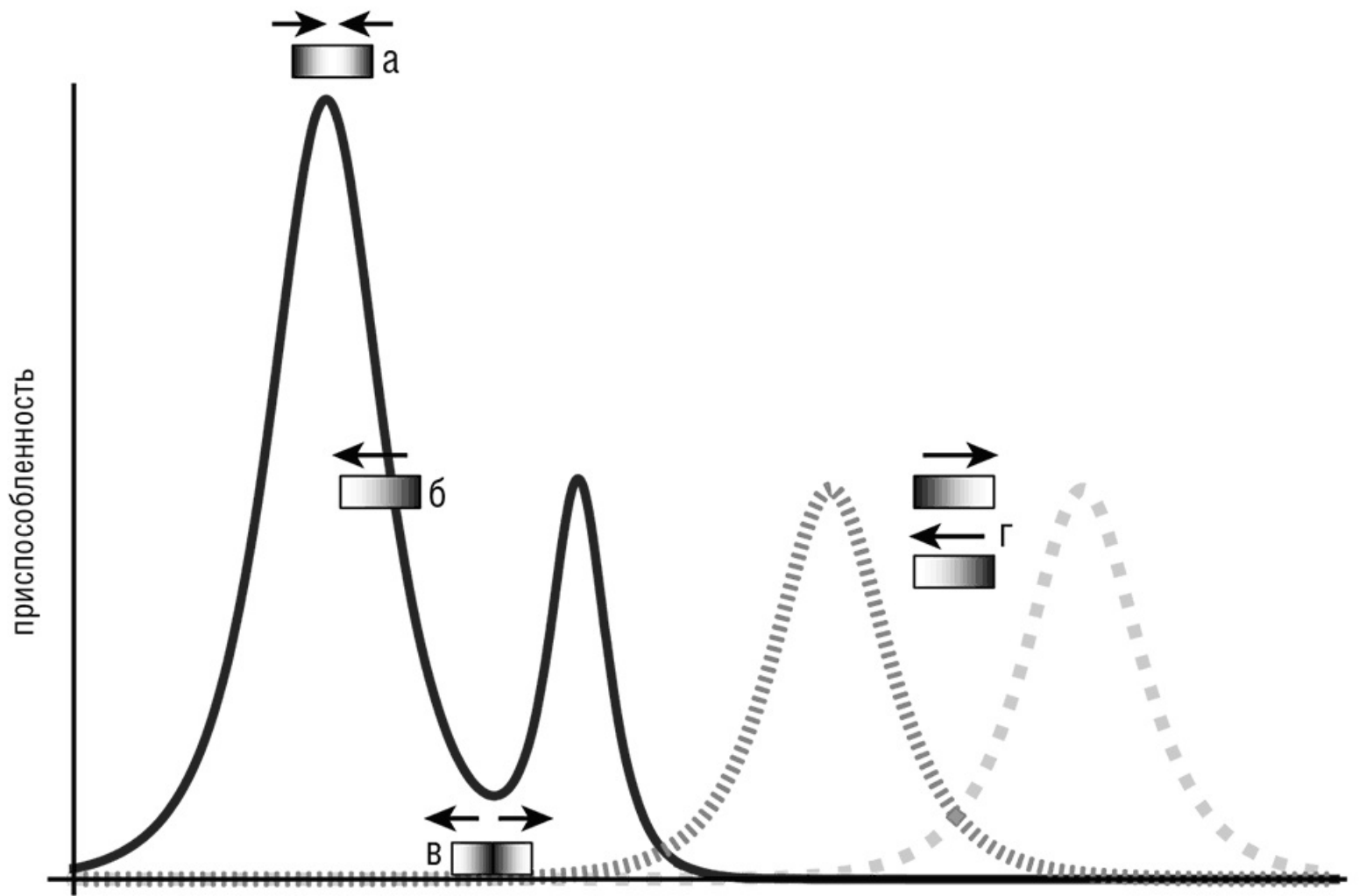


Рис. 1–3. Четыре различные формы отбора в эволюционирующей популяции: а — стабилизирующий отбор (адаптивный ландшафт представлен сплошной линией); б — движущий отбор (адаптивный ландшафт представлен сплошной линией); в — дизруптивный отбор (адаптивный ландшафт представлен сплошной линией); г — балансирующий отбор (адаптивный ландшафт периодически меняется, переключаясь между двумя пунктирными линиями)

Синтетическая теория эволюции

Объединение дарвиновской теории эволюции и генетики, состоявшееся в основополагающих исследованиях Фишера, Райта и Холдейна, подготовило почву для рождения синтетической теории эволюционной биологии. Само название идет от одноименной книги, опубликованной Джулианом Хаксли в 1942 году (Huxley, 2010), однако концептуальная структура СТЭ полностью сформировалась только в 1959 году в ходе мероприятий, посвященных 100-летию юбилею «Происхождения...». Новая синтетическая теория стала результатом работы многих выдающихся ученых. Можно утверждать, что главными архитекторами СТЭ были экспериментальный генетик Феодосий Добржанский, зоолог Эрнст Майр и палеонтолог Джордж Гейлорд Симпсон. Экспериментальные и полевые работы Добржанского с плодовой мушкой *Drosophila melanogaster* принесли насущные фактические свидетельства в поддержку теории популяционной генетики и стали первой крупномасштабной экспериментальной проверкой идеи естественного отбора. Книга Добржанского «Генетика и происхождение видов» (Dobzhansky, 1951) явилась основным программным документом СТЭ, в котором он сузил понятие эволюции до «изменения частоты аллеля в генетическом пуле». Знаменита также крылатая фраза Добржанского о том, что «ничто в биологии не имеет смысла, кроме как в свете эволюции» ^[13] (см. больше о «смысле» в прил. I). Заслугой Эрнста Майра, как никакого другого ученого, является серьезная, крайне влиятельная попытка теоретического решения принципиальной проблемы, поставленной Дарвином, – происхождения видов. Майр сформулировал так называемую биологическую концепцию вида, согласно которой видообразование происходит, когда две популяции (размножающиеся половым путем) изолированы друг от друга достаточно долго, чтобы обеспечить необратимую генетическую несовместимость (Mayr, 1963).

Симпсон реконструировал наиболее полную (на тот момент) картину эволюции жизни на основании палеонтологической летописи (Simpson, 1983). Замечательно, что Симпсон осознал стазис (отсутствие существенных изменений) в эволюции большинства видов и резкую смену доминантных видов. Он ввел понятие квантовой эволюции, которое предвосхитило теорию прерывистого равновесия, предложенную Стивеном Джемсом Гулдом и Нильсом Эддриджем (см. гл. 2).

Консолидация СТЭ в 1950-х годах была довольно странным процессом, сопровождавшимся странным «затвердеванием» (выражение Гулда) основных идей Дарвина (Gould, 2002). Так, доктрина СТЭ фактически отбросила идею Райта о случайном дрейфе генов и его эволюционной важности и стала бескомпромиссно панадапционистской. Более того, сам Симпсон отказался от идеи квантовой эволюции, так что градуализм продолжал оставаться одним из неоспоримых столпов СТЭ. Такое «затвердевание» сделало СТЭ относительно узкой, в некотором смысле даже догматичной, системой.

Чтобы продолжить обсуждение эволюции эволюционной биологии и ее преобразование в век геномики, представляется необходимым кратко резюмировать основные принципы эволюции, впервые сформулированные Дарвином, затем усовершенствованные первым поколением биологов-эволюционистов и, наконец, кодифицированные в СТЭ. Мы будем возвращаться к каждому из этих ключевых моментов на протяжении всей книги.

1. Ненаправленное случайное изменение – это главный процесс, обеспечивающий материал для эволюции. Дарвин впервые показал, что случайность является основным фактором в истории жизни, и это, несомненно, было одной из его наиболее важных идей. Дарвин также признавал роль направленной, ламарковской изменчивости и в последующих изданиях

«Происхождения...» склонялся даже к более весомой роли этого механизма эволюции. Однако СТЭ твердо настаивает на том, что случайные мутации являются единственным источником эволюционно значимой изменчивости.

2. Действие эволюции заключается в фиксации редких выгодных изменений и элиминации вредных изменений. Согласно Дарвину и СТЭ, в этом состоит процесс естественного отбора, который, наряду со случайной изменчивостью, является основной движущей силой эволюции. Естественный отбор, очевидно сходный и навеянный «невидимой рукой» рынка, которая, по теории Адама Смита, управляет экономикой, был первым из когда-либо предложенных механизмов эволюции, который был прост и правдоподобен и не требовал изначально мистического подхода. Таким образом, это вторая ключевая идея Дарвина. Сьюэл Райт подчеркивал, что случайность может играть вспомогательную роль не только в возникновении, но также и в закреплении изменений в ходе эволюции с помощью дрейфа генов, в результате которого случайно сохраняются нейтральные или умеренно вредные изменения. Согласно теории популяционной генетики, дрейф генов особенно значим в небольших популяциях, проходящих через «бутылочное горлышко». «Генетическая тяга», или «езда в одной повозке», – это другая форма случайного закрепления невыгодных мутаций. Однако СТЭ в ее догматизированной форме фактически отрицает стохастические процессы в эволюции, кроме возникновения изменений, и придерживается полностью адапционистского (панадапционистского) взгляда на эволюцию. Такая модель неизбежно приводит к концепции «прогресса», постепенного улучшения «органов» в ходе эволюции. Дарвин поддерживал эту идею как основное направление развития, несмотря на четкое понимание, что организмы все еще далеки от совершенства в плане адаптивности, как можно прекрасно увидеть на примере рудиментарных органов, и несмотря на свое резко отрицательное отношение к любым формам ламарковского внутреннего стремления к совершенству. СТЭ уходит от прогресса как антропоморфной идеи, но тем не менее поддерживает общую концепцию эволюции от простых форм к сложным.

3. Полезные изменения, закрепляемые естественным отбором, бесконечно малы (в современной терминологии, эволюционно значимые мутации обладают бесконечно малым влиянием на приспособленность), поэтому эволюция происходит путем постепенного накопления этих слабых изменений. Дарвин был убежден, что в основе его теории лежит *строгий градуализм*: «Естественный отбор действует только путем сохранения и кумулирования малых наследственных модификаций, каждая из которых выгодна для сохраняемого существа... Если бы возможно было показать, что существует сложный орган, который не мог образоваться путем многочисленных последовательных слабых модификаций, моя теория потерпела бы полное крушение» («Происхождение видов...», гл. 6 [цит. по: Дарвин Ч. Сочинения. Т. 3 / Пер. с англ. К. А. Тимирязева, С. Л. Соболя. М.: Изд-во АН СССР, 1939]). Даже некоторые современники Дарвина полагали, что это излишняя, искусственная строгость теории. В частности, хорошо известны заблаговременные возражения со стороны Томаса Гексли. Еще до публикации «Происхождения...» Гексли написал Дарвину: «Вы взяли на себя ненужный груз безоговорочного следования принципу *Natura non facit saltum*» (<http://aleph0.clarku.edu/huxley/>). Несмотря на эти своевременные предостережения и даже на идею Симпсона о квантовом характере эволюции, СТЭ бескомпромиссно настаивает на градуализме.

4. Униформизм (термин был заимствован Дарвином из геологии Лайеля) – это один из аспектов классической эволюционной биологии, который связан, но в то же время отличается от принципиального градуализма. Это убеждение, что эволюционные процессы не изменялись по существу на всем протяжении истории жизни.

5. Следующий ключевой принцип логически связан с градуализмом и униформизмом: *макроэволюция* (происхождение видов и высших таксонов) управляется теми же механизмами, что и *микроэволюция* (эволюция внутри вида). Главным апологетом данного принципа был Добржанский, определивший эволюцию как изменение частоты аллелей в популяциях. Дарвин не пользовался терминами *микроэволюция* и *макроэволюция*; тем не менее достаточность внутривидовых процессов для объяснения происхождения видов и, в более широком плане, всей эволюции жизни можно считать центральной аксиомой Дарвина (или, возможно, фундаментальной теоремой, такой, однако, для которой у Дарвина не было даже и намека на доказательство). Представляется разумным говорить о данном принципе как об «универсальном униформизме»: *эволюционные процессы одинаковы не только на протяжении всей истории жизни, но и на разных уровнях эволюционных изменений, включая крупные преобразования*. Загадка взаимосвязи между микроэволюцией и макроэволюцией является в некотором смысле осью эволюционной биологии, поэтому мы будем постоянно возвращаться к ней в этой книге.

6. Эволюцию жизни можно адекватно представить в виде «огромного дерева», что и подчеркнуто единственной иллюстрацией в «Происхождении...» (в гл. 4). Дарвин представил древо жизни только как общую идею и не пытался исследовать фактический порядок ветвления. Древо было заселено реальными формами жизни, насколько они были известны в то время, одним из главных последователей Дарвина, знаменитым немецким биологом Эрнстом Геккелем. Основатели СТЭ не проявляли большого интереса к древу жизни, но они, несомненно, включали его в теорию как описание эволюции животных и растений, убедительно поддержанное палеонтологической летописью в XX веке. Однако микробы, определяющее значение которых в глобальной экологии становилось все более очевидным, фактически остались за пределами эволюционной биологии.

7. Концепция единого древа жизни имеет следствие, которое заслуживает статуса отдельного принципа: существующее в настоящее время разнообразие форм жизни произошло от общего предка (или нескольких форм-предшественников, в соответствии с осторожной формулировкой Дарвина в главе 14 «Происхождения...», см. Darwin, 1859). Спустя много лет он был назван «последним универсальным клеточным предком» (*Last Universal Cellular Ancestor, LUCA* [\[14\]](#)). Для создателей СТЭ существование LUCA не вызывало сомнений, но они, по-видимому, не считали реалистичной или научно важной целью прояснение его природы.

Краткий обзор главы

В своей книге «Происхождение видов...» Чарльз Дарвин тщательно собрал доказательства изменений во времени, которые охватывают мир живых существ, и впервые предложил убедительный механизм эволюции: естественный отбор. Эволюция путем естественного отбора, безусловно, является одной из самых существенных концепций, когда-либо разработанных учеными, и даже была объявлена самой важной идеей в истории человечества (Dennett, 1996). В свете этого может показаться парадоксальным, что понятие естественного отбора нередко считают простой тавтологией. Если рассуждать с точки зрения выживания наиболее приспособленных, видно, что для этого взгляда есть основания. Однако, если рассматривать весь дарвиновский сценарий эволюции в целом, его решительно не тавтологические и не тривиальные аспекты становятся очевидными. В действительности Дарвин предложил механизм преобразования случайных изменений в отнюдь не случайные адаптации, вплоть до сложнейших приспособлений, исполняющих узкоспециализированные функции и тем самым повышающих приспособленность их носителей. Если рассматривать этот процесс в терминах физики и несколько вольно следовать идеям знаменитой книги Эрвина Шредингера, дарвиновская эволюция является машиной для создания отрицательной энтропии, другими словами, порядка из беспорядка. На мой взгляд, самым главным прозрением Дарвина было осознание того, что простой механизм, лишенный какого-либо телеологического содержания, вероятно, мог только благодаря случайным изменениям привести к появлению удивительного разнообразия форм жизни, каждая из которых в совершенстве приспособлена к условиям среды своего обитания. С этой точки зрения, «невидимая рука» естественного отбора кажется почти чудесно всемогущей, и нельзя не задаваться вопросом, действительно ли этого достаточно, чтобы объяснить историю жизни. Этот вопрос неоднократно использовался в качестве риторического приема креационистами всех мастей, но он также всерьез поднимался биологами-эволюционистами. В остальной части этой книги мы увидим, что разные ученые дают на него разные ответы, которые зависят от того, о каких именно ситуациях и этапах эволюции жизни идет речь.

Конечно, дарвинизм в его изначальном виде столкнулся с более значительными и непосредственными проблемами, чем вопрос о достаточности естественного отбора: Дарвин и его ранние последователи не имели представления о механизмах наследования и о том, будут ли когда-либо открыты механизмы, согласующиеся со сценарием Дарвина. В этом смысле здание теории Дарвина висело в воздухе. Повторное открытие законов генетики в начале XX века и последующее развитие теоретической и экспериментальной популяционной генетики обеспечило твердое основание для дарвиновской теории эволюции. Было показано, что, без сомнения, популяции эволюционируют посредством процесса, в котором дарвиновский естественный отбор играет важнейшую роль. СТЭ в эволюционной биологии явилась завершением работы Дарвина, последовательно объединив дарвинизм и генетику. По мере развития СТЭ заметно «окостенела», настаивая на градуализме, униформизме и, что наиболее важно, монополии естественного отбора как единственном пути эволюции. В соответствии с СТЭ все изменения, закрепляемые в ходе эволюции, являются адаптивными, по крайней мере изначальными. При всех своих выдающихся достоинствах СТЭ представляет собой довольно догматичную и удручающе незаконченную теорию. Назовем три наиболее бросающиеся в глаза проблемы: СТЭ совершенно бездоказательно распространяет механизмы и закономерности, принятые в микроэволюции, на макроэволюционные процессы; она ничего не говорит об эволюции микробов, являющихся наиболее широко распространенными и многообразными

формами жизни на Земле; и она даже не пытается обратиться к вопросу о происхождении жизни.

Рекомендуемая дополнительная литература

Futuyma, Douglas. (2009) *Evolution*, 2d edition. Sunderland, MA: Sinauer Associates.

Возможно, лучший из учебников по эволюционной биологии для студентов.

Gould, Stephen Jay. (2002) *The Structure of Evolutionary Theory*. Cambridge, MA: Harvard University Press.

Почти 1500-страничный том, очевидно, не для слабых духом, и не многие прочитают его полностью. Тем не менее по крайней мере первая часть ценна своим четким и точным описанием истории эволюционной биологии и острой критикой СТЭ.

Hartl, Daniel L., and Andrew G. Clark. (2006) *Principles of Population Genetics*, 4th edition. Sunderland, MA: Sinauer Associates.

Превосходный, достаточно серьезный, но доступный для понимания учебник по популяционной генетике.

Mayr, Ernst. (2002) *What Evolution Is*. New York: Basic Books.

Упрощенное, но ясное и полезное представление классической эволюционной биологии одним из основателей СТЭ.

Schroedinger, Erwin. (1992) *What Is Life?: With «Mind and Matter» and «Autobiographical Sketches»*. Cambridge, MA: Cambridge University Press. (Перевод: Шредингер Э. Что такое жизнь? Физический аспект живой клетки / Пер. с англ. 3-е изд. Ижевск: РХД, 2002.)

Первое издание этой блестящей книги вышло в 1944 году на основе лекций Шредингера (одного из основателей квантовой механики), прочитанных им в Эдинбурге, где он жил во время Второй мировой войны. Эта очевидно устаревшая, однако удивительно доходчивая, пророческая книга все еще важна для обсуждения роли энтропии и информации в биологии.

Глава 2. От синтетической теории эволюции к эволюционной геномике: различные механизмы и пути эволюции

В этой главе мы продолжим обсуждение эволюционной биологии в период до появления геномики. Многие из обсуждаемых направлений развития не являлись предшественниками синтетической теории эволюции (СТЭ). На самом деле они возникали параллельно с развитием СТЭ, но были отвергнуты «каноном» вследствие «ужесточения» СТЭ. Достижения, которые обсуждаются в этой главе, относятся к интервалу между 1930 (публикация книги Рональда Фишера, которая ознаменовала вторую, зрелую стадию развития эволюционной биологии) и 1995 годами (первые сравнения полных геномов клеточных форм жизни). Моя цель здесь – вкратце обрисовать сложную сеть эволюционных идей, теорий и наблюдений, которые дополнили достаточно жесткую структуру СТЭ и стали пусковой площадкой для нового, «геномного» подхода к изучению эволюции.

Репликация цифровых носителей информации: центральный принцип биологии и необходимое и достаточное условие эволюции

Модель структуры ДНК, представленная Джеймсом Уотсоном и Фрэнсисом Криком (очевидно, основанная на рентгеновских структурах, полученных Розалинд Франклин и другими), несомненно, является одним из главных открытий не только биологии XX века, но и всей истории биологии (Watson and Crick, 1953b). Однако этот прорыв не всегда упоминается в связи с принципами биологической эволюции.

С моей точки зрения, структура ДНК и модель ее репликации, которую Уотсон и Крик описали в своей второй классической статье как непосредственное следствие структуры (Watson and Crick, 1953a), являются важнейшим фундаментальным открытием в изучении эволюции со времени публикации «Происхождения видов...». По сути, Уотсон и Крик вывели из структуры ДНК биологическое воплощение общего принципа цифрового хранения, кодирования и передачи информации. Система биологической передачи информации, которую выявили их исследования, может рассматриваться как расширение принципа машины Тьюринга, сначала через правила комплементарности нуклеотидных оснований (в процессах репликации и транскрипции), а затем в процессе трансляции, через генетический код (см. рис. 2–1). По сути, пусть и не в историческом смысле, эти открытия вытеснили концепцию Дарвина, в том смысле, что *вся дарвиновская схема эволюции является прямым следствием механизма репликации ДНК*. Для всех известных форм жизни биологическая передача цифровой информации влечет за собой исполнение следующих простых фундаментальных принципов [\[15\]](#).

Генетический материал любого организма состоит из линейной последовательности символов, четырех оснований нуклеиновых кислот, которая, прямо или косвенно, кодирует всю информацию, необходимую для построения организма [\[16\]](#).

Репликация генетического материала, являющегося механической основой наследственности, осуществляется на основе принципа однозначного комплементарного соответствия между А и Т(У), и G и С. (Так называемые правила Чаргафа, по имени их первооткрывателя, австрийского, а затем американского химика Эрвина Чаргафа [\[17\]](#).)

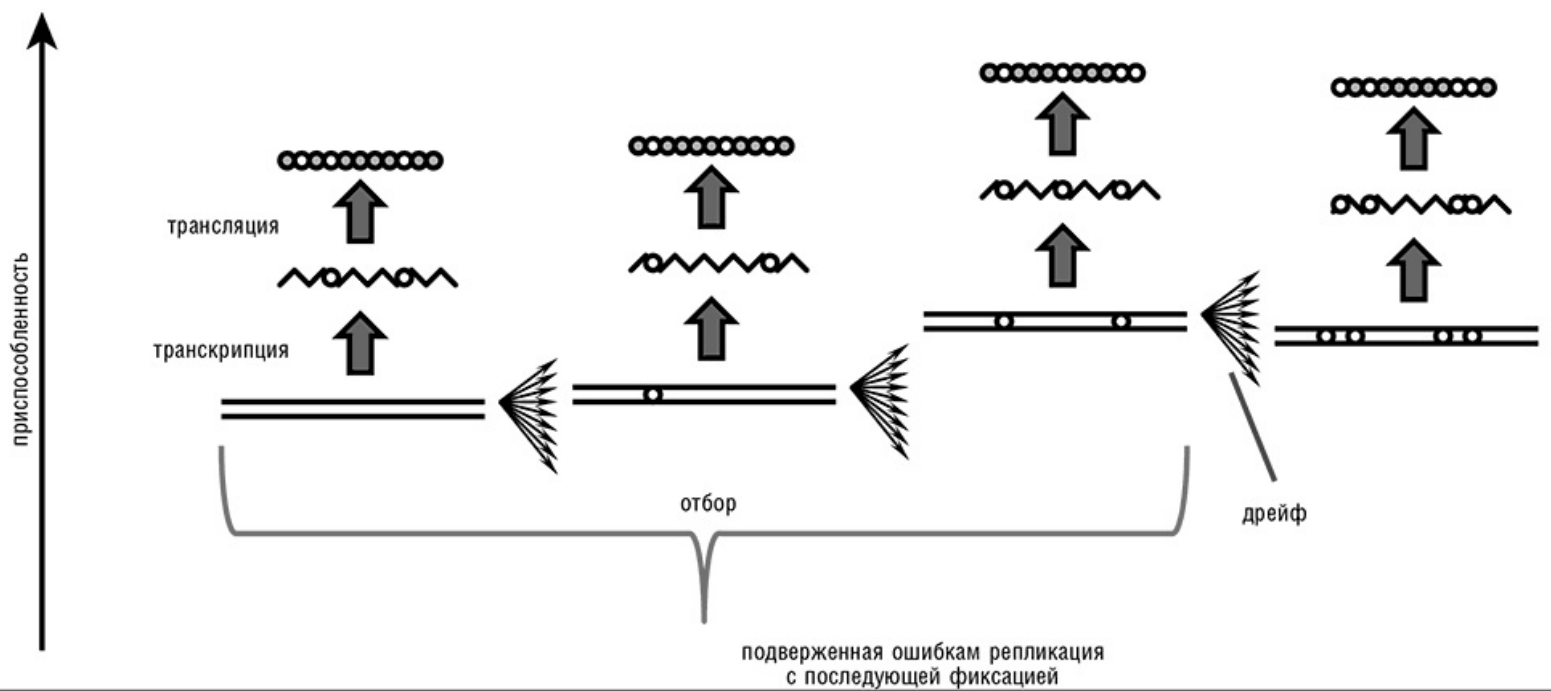


Рис. 2–1. Передача информации в биологических системах и превращение естественного отбора и генетического дрейфа в эпифеномены репликации. Белые круги на схеме обозначают изменения относительно оригинальной последовательности.

Уотсон и Крик описали эти ключевые принципы генетических систем в двух своих статьях, вышедших в 1953 году. Дальнейшие исследования добавили два очень важных аспекта:

1. Принцип комплементарности используется не только во время репликации, но и во время транскрипции ДНК во все виды РНК и во время трансляции мРНК в белок с помощью адапторных молекул тРНК.

2. Те же самые принципы цифровой репликации и декодирования применимы и для генетических систем, в которых генетический материал отличается от двойной спирали ДНК, изначально описанной Уотсоном и Криком, и состоит из РНК или односпиральной ДНК (например, у многих вирусов).

Теория информации твердо стоит на том, что передача информации абсолютно без ошибок невозможна в принципе. В реальности вероятность ошибки в любом конечном сообщении может быть сведена к минимуму, но любое снижение уровня ошибок при передаче информации возможно только за счет затраты энергии. Эта связь непосредственно следует из законов термодинамики. Центральный принцип эволюции может быть сформулирован следующим образом:

Репликация цифровых носителей информации неизбежно подвержена ошибкам, что влечет за собой эволюцию этих носителей путем естественного отбора и случайного дрейфа генов при условии, что уровень ошибок репликации ниже катастрофического порога, имеющего порядок величины от одной до десяти ошибок на геном за один цикл репликации [18].

Назовем это обобщение принципом подверженной ошибкам репликации (ПОР) [19]. Этот принцип становится самоочевидным, как только мы осознаем существование и основной механизм репликации. Он был впервые описан математически в теории Манфреда Эйгена (Eigen, 1971), который также ввел понятие концепции порога ошибки (Biebricher and Eigen, 2005) – эта теория и ее применение будут рассматриваться далее в главе 12. ПОР основывается на следующих двух предположениях, которые могут показаться очевидными, но заслуживают тем не менее особого упоминания:

1. Ошибки репликации наследуются (проходят через циклы репликации).

2. Существует обратная связь между генотипом и фенотипом: некоторые ошибки репликации влияют на эффективность и точность репликации как отрицательно, так и положительно.

Это и отличает биологические репликаторы с их «неограниченной наследуемостью» от репликаторов с «ограниченной наследуемостью», таких как кристаллы или ряд химических циклов, которые реплицируются, но не передают накопленные дефекты последующим поколениям (Szathmary, 2000). Упрощенно говоря, разница в том, что в нуклеиновых кислотах замена одного нуклеотида на другой влияет только на передаваемую информацию, а не на физические или химические свойства носителя информации (во всяком случае, не значительно), как в случае небиологических систем.

В принципе должен существовать и нижний порог частоты ошибок репликации. Очевидно, что если математическое ожидание числа ошибок на цикл репликации стремится к нулю, то разнообразие становится недостаточным для эволюционного процесса. Однако необходимо заметить, что уровень ошибок репликации нуклеиновых кислот, определяемый эмпирически, в отсутствие сложных корректирующих механизмов (как, например, в РНК вирусов), не намного ниже верхнего, катастрофического порога. Таким образом, необходимым условием эволюции является достаточно низкий (а не достаточно высокий) уровень ошибок репликации [\[20\]](#). Вопрос о том, в какой степени оптимизируется фактическая точность передачи информации в биологической системе (другими словами, эволюционирует ли эволюция), весьма сложен, интересен и широко обсуждается. Мы рассмотрим его подробнее в главе 9.

Несмотря на то что вся естественно эволюционировавшая жизнь основана на репликации нуклеиновых кислот, ПОР не зависит от физической природы репликаторов, как можно видеть на примере эволюции компьютерных вирусов и различных компьютерных моделей эволюции искусственной жизни (Lenski et al., 2003). Тем не менее вопрос о том, необходим ли цифровой код для эволюции или эволюция может происходить и в аналоговых системах, весьма интересен и до сих пор остается открытым.

В главе 1 был затронут вопрос о псевдодарвиновской природе естественного отбора. По сути, ПОР действительно в значительной мере тривиализирует естественный отбор и генетический дрейф (два фундаментальных принципа эволюции), как бы отбирая у них статус независимых феноменов и низводя их до эпифеноменов ПОР. Это ни в коей мере не умаляет достижения Дарвина, Райта и других выдающихся эволюционных биологов и не уменьшает важность концепций естественного отбора и генетического дрейфа для описания эволюционных процессов на абстрактном уровне. Тем не менее открытие репликации с контролируемым уровнем ошибок обнаруживает более фундаментальные принципы, которые лежат в основе классических положений эволюционной биологии.

Молекулярная эволюция и молекулярная филогенетика

Традиционные филогенетические исследования, наполненные содержанием дарвиновскую концепцию древа жизни, были основаны на сравнении морфологических черт организмов, таких как структура скелета животных или строение цветков растений (Futuyma, 2005). Эволюционные биологи не осознавали, что сравнивать следует реальную молекулярную базу эволюции, которая подвержена действию естественного отбора, то есть гены, просто потому, что они практически ничего не знали о химической основе этой составляющей и о способе, которым гены кодируют фенотип организма. Более того, согласно парадигме панадаптационизма в эволюционной биологии, гены, на каком бы молекулярном механизме они ни основывались, должны существенно различаться в неродственных организмах, если учесть фенотипические различия между этими организмами, как подчеркивалось, в частности, Эрнстом Майром, одним из главных строителей СТЭ.

Идея того, что последовательность оснований ДНК может использоваться для эволюционной реконструкции, была, вероятно, впервые высказана в печати, пусть и мимоходом, еще Криком (в той же самой основополагающей статье, в которой он сформулировал адапторную гипотезу синтеза белка, – Crick, 1958). Эмиль Цукеркандль и Лайнус Полинг развили принципы и описали первое фактическое использование молекулярного эволюционного анализа несколько лет спустя. Они напрямую опровергли гипотезу Майра, показав, что последовательности аминокислот нескольких белков, которые в то время были известны для нескольких видов, такие как цитохром с глобины, оказались чрезвычайно консервативными даже у дальнородственных животных (Zuckerkindl and Pauling, 1965). Цукеркандль и Полинг также предложили концепцию молекулярных часов: они предсказали, что скорость эволюции определенной последовательности белка будет неизменна (с учетом возможных флуктуаций) в течение длительных временных интервалов в отсутствие функциональных изменений. Здесь необходимо отметить, что то, что последовательности генов, кодирующих «один и тот же белок» (то есть белки с одинаковым действием и сходными свойствами), в различных организмах оказались очень сходными – и, более того, степень схожести этих последовательностей отрицательно коррелировала с филогенетическим расстоянием между данными организмами, – можно рассматривать как наилучшее и исчерпывающее доказательство реальности эволюции.

В течение последующих нескольких лет, в основном благодаря работам Маргарет Дэйхофф и ее коллег, консервативность кодирующих белки последовательностей была продемонстрирована на примерах самых различных форм жизни, от бактерий до млекопитающих (Dayhoff et al., 1983). Учитывая открытие консервативности белок-кодирующих последовательностей и гипотезу молекулярных часов, оказалось естественным перейти к конструированию филогенетических деревьев на основе (не)похожести этих последовательностей, что должно было показать примерное время возникновения расхождений в последовательностях генов (белков) от общего предка. И в самом деле, скоро были изобретены несколько методов измерения расстояний в молекулярной филогенетике, а также введен принцип наибольшей экономии (см. табл. 2–1). Последующее тестирование гипотезы молекулярных часов на все растущей базе последовательностей ДНК показало, что для большинства генов эти часы идут не с одинаковой скоростью; напротив, эти данные оказались значительно диспергированы, то есть отклонения в скорости эволюции значительно превышали среднее отклонение, которое могло быть предсказано распределением Пуассона (Bromham and Penny, 2003). Такая избыточная дисперсия молекулярных часов приводит к особенности

молекулярного филогенеза, известной как притяжения длинных ветвей (ПДВ), существенно искажающей результаты молекулярного филогенетического анализа (см. табл. 2–1). Молекулярная филогенетика, таким образом, превратилась в сложную ветвь прикладной математики и статистики в основном для того, чтобы справляться с эффектами ПДВ и им подобными (Felsenstein, 2004). Но, несмотря на все артефакты, молекулярная филогенетика остается краеугольным камнем современной эволюционной биологии и использует в первую очередь методы наибольшего правдоподобия (см. табл. 2–1).

Таблица 2–1. Краткое описание филогенетических методов

Методы, основанные на секвенировании

Требуют многократного сравнения гомологичных нуклеотидов или белковых последовательностей.

Дистанционно-матричные методы

Все эти методы используют матрицы межвидовых расстояний $\langle d_{ij} \rangle$ (i, j соответствуют видам), рассчитанные на основе сравнений последовательностей с корректировкой на множественные замещения. *Ультраматрические методы (простая иерархическая кластеризация)*. Достоверны только для стабильных молекулярных часов. В принципе не считаются приемлемыми филогенетическими методами, но могут быть использованы для целей классификации или генерации предварительных филогенетических деревьев.

Метод объединения ближайших соседей (neighbor-joining). Более сложный метод восходящей кластеризации, основанный на минимальном эволюционном критерии (кратчайшая суммарная длина ветвей дерева). Чувствителен к ПДВ и гораздо менее точен, чем метод наибольшего правдоподобия, однако высокоэффективен для вычислений и быстр. Не используется для исчерпывающего филогенетического анализа, но для анализа большого количества последовательностей может быть единственным практически применимым методом.

Метод наименьших квадратов, метод Фитча. Метод измерения расстояний, основанный на минимизации разностей между расстояниями на филогенетическом дереве и в соответствующей матрице расстояний. По точности и эффективности примерно равен *методу объединения ближайших соседей*. Считается неподходящим для исчерпывающего филогенетического анализа, но используется для построения предварительных филогенетических деревьев для метода наибольшего правдоподобия.

Принцип наибольшей экономии (maximum parsimony)

Не использует матрицы расстояний, вместо этого работает с наборами состояний признаков. Состояниями признаков, в частности, могут быть нуклеотиды или аминокислоты в определенных позициях множественных выравниваний. Принцип наибольшей экономии (НЭ), основанный на принципе наименьшего действия в физике, определяет как наиболее вероятный тот эволюционный сценарий (филогенетическое дерево), который включает в себя наименьшее количество событий (переходов состояний в наборе признаков).

Существует множество алгоритмов, вычисляющих деревья, наиболее соответствующие принципу НЭ и использующие значимые и незначимые признаки. Принцип НЭ часто ставится под сомнение, поскольку существуют деревья, лишь

слегка отличающиеся от наиболее экономичного варианта, но имеющие совершенно иную топологию. Метод высоко чувствителен к ПДВ.

Метод наибольшего правдоподобия (maximum likelihood)

Аналогично методу НЭ, в методе наибольшего правдоподобия (НП) оцениваются переходы между состояниями признаков и выбираются деревья, набравшие наибольший вес. В отличие от метода НЭ, метод НП является параметрическим статистическим подходом, который использует детальную модель эволюции признака для оценки вероятности данных на основе имеющегося эволюционного дерева. Дерево, которое имеет наибольшую вероятность возникновения наблюдаемых данных, признается наиболее вероятным. Метод НП зачастую производит деревья, аналогичные тем, которые получаются методом НЭ, но теоретически он предпочтительнее, будучи (в отличие от НЭ) статистически более достоверным (то есть при наличии достаточного количества данных гарантирует получение наиболее правдоподобного дерева). На практике метод НП часто превосходит метод НЭ. Методы НП чрезвычайно затратны с вычислительной точки зрения и непрактичны при работе с большими наборами данных. Таким образом, методы НП зачастую используются для оптимизации предварительных деревьев, полученных методом объединения ближайших соседей и методом Фитча. Для тех же филогенетических исследований, где точность построения дерева важнее скорости, следует выбирать методы НП. Более того, недавние алгоритмические достижения более чем на порядок ускорили построение филогенетических деревьев методами НП без серьезных потерь точности (Price et al., 2010).

Байесовский подход

Подобно методу НП, этот подход использует функцию правдоподобия, но прибегает к теореме Байеса с целью связать апостериорную вероятность дерева с правдоподобием данных и априорную вероятность дерева с эволюционной моделью. В отличие от методов НЭ и НП, которые выводят наилучшее дерево или набор деревьев, методы байесовского вывода выбирают деревья пропорционально их правдоподобию и определяют представительный набор деревьев. Метод хорошо работает для относительно небольших объемов данных, но непрактичен для больших.

Проверка точности филогенетических методов и достоверности деревьев

Модельные деревья

Филогенетические методы постоянно проверяются на искусственно смоделированных данных, для которых известна точная история эволюции. Методы сравниваются по критерию точности реконструкции топологии для искусственно построенных деревьев. Как правило, различные методы НП и байесовские методы превосходят все остальные для небольших наборов данных. Наилучшие результаты показывают итерационные методы, которые используют исходное дерево, построенное по методу НП, чтобы выравнять данные, перестраивать дерево и повторять так до сходимости.

Бутстреппинг

Наиболее часто используемый тест на надежность топологии филогенетического дерева, при котором рассматриваются выборки данных (колонки выравнивания) и дерево оценивается по большому числу выборок. Процент выборок (то есть

репликаций), в которых реконструируется данный узел дерева, называют уровнем поддержки. Статистика бутстреппинга еще не полностью разработана, поэтому пороговые значения для «достаточно высокого» уровня поддержки определяются путем моделирования или эмпирического анализа и могут варьировать в зависимости от целей конкретного исследования (например, значения более 90 процентов, или более 70 процентов; поддержка ниже 50 процентов обычно не считается надежной).

Статистические критерии проверки филогенетических гипотез (топологий деревьев)

Для сравнения правдоподобия различных топологий деревьев, выводимых из одного и того же набора данных, разработаны статистические критерии, основанные на различных моделях правдоподобия (самые известные – критерий Кишино – Хасегавы и приблизительно несмещенный критерий).

Когда исследователь интересуется филогенетическим сродством конкретного таксона, соответствующая ветвь переносится в различные положения в дереве, при сохранении топологии остальных ветвей, и правдоподобие каждого из полученных деревьев сравнивается при помощи статистических критериев с правдоподобием исходного дерева, полученного методом НП. Разновидность этого критерия применяется к деревьям с ограничениями, используемым для проверки филогенетических гипотез, таких как монофилия определенной группы (например, архей) в определенном наборе данных. В этом случае сравнивается правдоподобие дерева с ограничениями (монофилия в данном примере) с правдоподобием исходного НП-дерева.

Часто встречающиеся аномалии филогенетического анализа

Ни один филогенетический метод не застрахован от аномалий, которые часто оказывают заметное влияние на топологию дерева. Двумя основными классами филогенетических аномалий являются *гомоплазия притяжение длинных ветвей (ПДВ)*. Гомоплазия включает в себя параллельные, сходящиеся и обратные мутации, которые филогенетически не информативны и неверно истолковываются филогенетическими методами. ПДВ называется чрезвычайно распространенный случай, когда длинные ветви (быстро эволюционирующие линии) в дереве кластеризуются вместе только потому, что ни одна из них не проявляет сродства к другим группам, а не потому, что они на самом деле образуют монофилетическую группу. Филогенетики также иногда говорят о притяжении коротких ветвей, то есть ошибочной кластеризации коротких ветвей дерева. Разработка новых методов филогенетического анализа в большой степени побуждается необходимостью преодолеть эти аномалии, сохраняя притом вычислительный метод приемлемым с практической точки зрения.

Общие производные признаки

Важным подходом филогенетического анализа, дополняющим традиционные молекулярные филогенетические методы, является анализ общих производных признаков (так называемых синапоморфий), которые могут быть использованы для разграничения монофилетических групп (клад). Синапоморфии суть признаки, объединяющие всех членов монофилетической группы и исключаящие все другие виды. В принципе одна достоверная синапоморфия может определять кладу. Однако это верно только в отсутствие гомоплазии, которую невозможно исключить для

большинства признаков. Предполагаемые синапоморфии выбираются таким образом, чтобы свести вероятность гомоплазии к минимуму, например уникальные вставки в консервативных генах, в частности вставки мобильных элементов, мутации, которые требуют нескольких нуклеотидных замен, и слияния генов. В филогеномике идет активный поиск подобных редких геномных изменений. Одних синапоморфий часто недостаточно для несомненных филогенетических выводов, но они предоставляют дополнительные свидетельства для филогений, основанных на геномных последовательностях.

Деревья, не основанные на геномных последовательностях

Филогенетические методы пригодны не только для выравнивания гомологичных последовательностей, но и для анализа дистанционных матриц, полученных полногеномным сравнением любого числа других признаков (таких как содержание общих генов или оперонная организация). Например, в случае содержания общих генов расстояние между двумя геномами определяется как $D_{ij} = n_{ij} / n_i$, где n_{ij} – число генов, общих для двух геномов, а n_i – полное число генов в меньшем геноме. Геномные деревья, полученные этим методом, обычно не являются надежными филогениями из-за обширной гомоплазии. Однако эти деревья могут быть информативными для сравнения образа жизни организмов.

Нейтральная теория молекулярной эволюции

Вероятно, важнейшим прорывом в эволюционной биологии после СТЭ стала нейтральная теория молекулярной эволюции. Как правило, ее связывают с именем Мото Кимуры (Kimura, 1983), хотя Джукс и Кинг одновременно и независимо развивали аналогичные идеи. Вначале нейтральная теория развивалась как логическое продолжение популяционно-генетических идей Райта, основанных на важности генетического дрейфа в эволюции. Согласно нейтральной теории, значительное большинство всех фиксируемых в процессе эволюции мутаций являются относительно нейтральными; таким образом, фиксация возникает на основе случайного дрейфа. Следствием этой теории, неоднократно подчеркиваемым Кимурой, является то, что геномная последовательность эволюционирует равномерно, как по часам (в подтверждение исходной гипотезы молекулярных часов Цукеркандля и Полинга), при этом полезные мутации, подверженные естественному отбору, настолько редки, что ими можно с успехом пренебречь в целях количественного описания эволюционного процесса. Естественно, нейтральная теория отнюдь не подразумевает, что естественный отбор не важен для эволюции. На самом деле теория подчеркивает, что доминирующим способом отбора является не дарвиновский позитивный отбор на основе адаптивных мутаций, а отсекающий (очищающий) отбор, который удаляет вредные мутации, в то же время допуская фиксацию нейтральных мутаций путем генетического дрейфа.

Последующие исследования довели эту теорию до более реалистичной формы: чтобы зафиксироваться, мутация должна быть не в буквальном смысле нейтральной, а всего лишь достаточно мало вредной, чтобы избежать немедленного удаления отсекающим отбором. Современная теория «почти нейтральных» мутаций была разработана в первую очередь Томоко Отой (Ohta, 2002). То, какие мутации распознаются как вредные при вычищении отбором, в большой степени зависит от величины популяции: в небольших популяциях в ходе генного дрейфа могут зафиксироваться даже существенно вредные мутации, тогда как в больших популяциях даже малого негативного эффекта будет достаточно для удаления мутантной аллели (см. табл. 1–1).

Главной эмпирической проверкой теории (почти) нейтральных мутаций является измерение постоянства скорости эволюционного процесса в семействах генов. Несмотря на то что зачастую можно наблюдать значительную дисперсию молекулярных часов, такие измерения с уверенностью показывают, что доля нейтральных мутаций среди зафиксированных и в самом деле весьма существенна (Bromham and Penny, 2003; Novichkov et al., 2004). Теория почти нейтральных мутаций является значительным отступлением от селекционистской парадигмы СТЭ, поскольку однозначно утверждает, что большинство мутаций, зафиксированных в ходе эволюции, не подвержены дарвиновскому (позитивному) отбору. Хотя Дарвин и предвидел нейтралистскую парадигму, утверждая, что для целей классификации лучше всего подходят селективно нейтральные характеристики, однако он не развил эту прозорливую идею, и она, таким образом, не стала частью СТЭ.

Важно отметить, что в ходе последующего развития «нейтральной» теории Кимура, Ота и другие осознали, что те мутации, которые были почти нейтральными во время их фиксации, не были в то же время не важными для эволюции. Напротив, такие мутации составили резервуар вариаций (почти нейтральную сеть аллелей), который может использоваться естественным отбором в свете меняющихся условий среды, – феномен, важный как для микро-, так и для макроэволюции (Kimura, 1991). Эта идея стала ключевой для некоторых позднейших открытий в эволюционной теории, мы обсудим ее более детально позднее в этой книге (в частности, в гл.

Измерение естественного отбора сравнением последовательностей ДНК

Несмотря на всю свою важность, дарвиновский естественный отбор является концепцией, определенной в качественных терминах. В рамках же популяционной генетики и СТЭ отсекающий и положительный отбор оказались более конкретными и математически определенными. В описании СТЭ отбор скорее можно приравнять к силе в классической механике или потоку в классической термодинамике, то есть к феноменологически определяемому количеству. С появлением сравнения последовательностей ДНК стало возможно обнаруживать и измерять отбор в определенных механистических терминах, базируясь на подсчете различных типов замещений нуклеотидов. Для измерения отбора путем сравнения последовательностей используются две очень простые идеи (см. табл. 2–2). Эти два подхода имеют в своей основе очень много общего, поскольку оба определяют два класса сайтов, один из которых принимается в качестве фона нейтральной эволюции. Первый метод заключается в сравнении числа замен нуклеотидов в позициях, важных с точки зрения кодирования аминокислот (несинонимичные позиции), и в позициях, которые, из-за избыточности генетического кода, не имеют значения для кодирования белков (синонимичные замены). Если отношение скоростей несинонимичных и синонимичных замен (Ka/Ks , см. табл. 2–2) значительно ниже 1, то эволюция соответствующего гена в основном определяется отсекающим отбором, направленным на данную последовательность белка; напротив, в случае $Ka/Ks > 1$ эволюция определяется в основном положительным дарвиновским отбором (см. табл. 2–2). Второй, более точный подход использует так называемый критерий Макдональда – Крейтмана для измерения отбора, при котором соотношение Ka/Ks сравнивается для внутривидовых вариантов (полиморфизмы) и межвидовых вариантов (фиксированные мутации). Поскольку незафиксированные полиморфизмы в основном нейтральны, то межвидовое отношение Ka/Ks должно быть значительно меньше, чем Ka/Ks для полиморфизмов в случае отсекающего отбора, и значительно больше, чем значение для полиморфизмов в случае положительного отбора.

Таблица 2–2. Измерение отбора путем анализа последовательностей белок-кодирующих генов (Hurst, 2002; Li, 1997)

Белок-кодирующие последовательности состоят из двух видов сайтов:

- синонимичные, в которых замены не влияют на последовательность кодируемых аминокислот;
- несинонимичные, в которых замены ведут к заменам аминокислот.

Отношение Ka/Ks (где Ka – частота несинонимичных замен, Ks – синонимичных; обе вычисляются с коррекцией на множественные замены) является количественной мерой отбора, действующего на уровне белковых последовательностей.

$Ka/Ks = 1$ – нейтральная эволюция белковой последовательности (кодируемый белок не подвергается отбору).

Для большинства белок-кодирующих генов $Ka/Ks \ll 1$ – отсекающий отбор.

Для прокариот типично $Ka/Ks < 0,1$.

Для эукариот типично $Ka/Ks \approx 0,1–0,2$.

- $Ka/Ks > 1$ – положительный отбор; достаточно редко встречается для белок-кодирующих генов, но для некоторых категорий генов, несомненно, наличествует,

например для генов, участвующих в антипаразитической защите или в сперматогенезе, а также в вирусных белках, таких как гемагглютинин вируса гриппа.

- Для измерения Ka/ Ks для индивидуальных сайтов используют методы наибольшего правдоподобия; большинство белоккодирующих генов содержат несколько сайтов, подверженных положительному отбору.

- Использование Ka/ Ks для измерения уровня отбора предполагает нейтральность синонимичных сайтов.

- Однако Ka и Ks положительно коррелируют между собой – таким образом, отбор затрагивает и синонимичные сайты.

- Некодирующие сайты, такие как интронные последовательности, могут использоваться как фон нейтральной эволюции при измерении отбора на синонимичных сайтах (Ks/ Ki , где Ki – частота замен для интронных сайтов).

- Критерий Макдональда – Крейтмана (Aquadro, 1997; McDonald and Kreitman, 1991) широко используется для измерения отбора. Он сравнивает внутривидовые вариации (частота полиморфизма, P) с межвидовыми вариациями (дивергенция, D).

- $Dn/ Ds = Pn/ Ps$ – нейтральная эволюция белковой последовательности.

- $Dn/ Ds < Pn/ Ps$ – отсекающий отбор.

- $Dn/ Ds > Pn/ Ps$ – положительный отбор.

Появление таких количественных подходов к анализу отбора примечательно не только благодаря их технической применимости в изучении эволюции: они также являются признаком фундаментальных изменений в способах осмысления отбора биологами. Дарвиновская качественная идея, которая была выражена в абстрактной математической форме Фишером и впервые измерена с использованием генетических методов Добржанским и его учениками, теперь превратилась в прямо измеряемую статистическую характеристику ансамблей нуклеотидных сайтов. Такая трансформация концепции отбора сродни тому, как классическая термодинамика с ее абстрактными потоками превратилась в статистическую физику Больцмана и Гиббса (см. гл. 4).

Эгоистичные гены, мусорная ДНК и мобильные элементы

Хоть это и редко утверждается без обиняков, классическая генетика предполагает, что почти все части генома (все нуклеотиды, если употреблять более современные, молекулярные термины) имеют определенные функции. Это неявно выраженное утверждение также важно и с точки зрения СТЭ, с ее панадаптационистским подходом. Однако это понимание подверглось сомнению еще в 1960-х и 1970-х годах по мере накопления данных об отсутствии прямой связи между размером генома и фенотипической сложностью организма. Даже с использованием приблизительных методов, доступных в то время, становилось ясно, что организмы с примерно одинаковым уровнем фенотипической сложности зачастую имеют геномы, на порядок различающиеся по размеру (так называемый парадокс гаплоидной величины). Этот парадокс был концептуально разрешен с помощью двух связанных друг с другом фундаментальных идей: эгоистичных генов и мусорной ДНК [\[21\]](#). Концепция эгоистичных генов была предложена Ричардом Докинзом в одноименной книге, изданной в 1976 году (Dawkins, 2006). Резко отступив от организм-центричной парадигмы СТЭ, Докинз приходит к выводу, что естественный отбор может действовать не только на уровне организма в целом, но и на уровне индивидуального гена. Этот взгляд, поданный в умышленно провокационной манере, *представляет геномы и организмы, по сути, средствами размножения генов.*

Концепция эгоистичных генов породила множество важных выводов, и некоторые из них мы рассмотрим ниже в этой книге. Один из аспектов, имеющий непосредственное отношение к парадоксу гаплоидной величины, был всесторонне рассмотрен Фордом Дулиттлом и Кармен Сапиенцей (Doolittle and Sapienza, 1980), а также Лесли Оргелом и Фрэнсисом Криком (Orgel and Crick, 1980). Они предположили, что немалая или даже основная часть геномной ДНК (по крайней мере в сложных многоклеточных организмах) состоит из различных классов повторов, которые образуются в результате амплификации эгоистичных элементов – абсолютных паразитов, говоря хлестким языком Оргела и Крика. Другими словами, с точки зрения организма, большая часть геномной ДНК должна быть признана избыточной. Такой взгляд на геном в корне отличается от панселекционистской парадигмы, присущей СТЭ, в рамках которой большинство или даже все нуклеотиды в геноме подвержены влиянию отсекающего или положительного отбора, действующего на уровне организма.

Концептуально родственным важным открытием стало обнаружение транспозонов, или «прыгающих генов», сначала Барбарой Макклиток в 1940-х годах в растениях, а затем и в животных. Эти транспозоны стали затем известны как мобильные элементы (то есть генетические элементы, которые имеют тенденцию часто менять свое место в геноме; McClintock, 1984). Демонстрация вездесущности мобильных элементов привела к концепции высокодинамичных, постоянно меняющихся геномов задолго до рождения современной геномики [\[22\]](#).

Эволюция путем дубликации генов и геномов: ортологи и паралоги

СТЭ в полной мере унаследовала центральное положение дарвиновской теории, провозгласившей постепенные малые изменения единственно возможным материалом для эволюции. Однако эта концепция была поставлена под сомнение альтернативной концепцией эволюции дубликацией гена, разработанной Сусуму Оно в его классической книге 1970 года (Ono, 1970). Мысль о том, что дубликация частей хромосом может служить одной из движущих сил эволюции, восходит к основателям современной количественной генетики, в частности к Фишеру и Холдейну ^[23]. Однако Оно первым предположил, что дубликация генов является основой эволюции геномов и организмов, и первым подвел качественную теорию под это положение. Начав с цитогенетических свидетельств полногеномной дубликации (ПГД) в начале эволюции хордовых, Оно выдвинул гипотезу о том, что дубликация генов является важным, если не единственным, путем эволюции новых биологических функций. Согласно гипотезе Оно, дубликация гена высвобождает одну из копий от ограничений отсекающего отбора и, таким образом, эта копия получает потенциал развития новой функции (феномен, позднее названный неофункционализацией). Очевидно, что возникновение нового гена в результате дубликации, не говоря уже о дубликации геномного участка или ПГД, является огромным отличием от дарвиновских ничтожно малых изменений. Если такие крупные события в самом деле являются ключевыми для эволюции, то парадигма постепенных изменений в опасности. Позднейшие исследования дубликации генов, обсуждаемые далее в этой книге (см. гл. 8 и 9), привели к предположению о том, что неофункционализация вряд ли является основным путем эволюции дублицированных генов. Однако факт остается фактом: дубликация, как важнейший механизм эволюции, бросает вызов градуализму.

Примерно в то же время, когда была издана книга Оно об эволюции путем дубликации генов, Уолтер Фитч опубликовал весьма плодотворную статью, всю значимость которой стало возможным оценить лишь в свете более поздних достижений геномики. Фитч исследовал понятие гомологии (общего предка) генов и провел различие между двумя классами гомологичных генов: *ортологамии паралогами* (Fitch, 1970). Ортологи – это гены, которые эволюционировали вертикально от одного предкового гена, принадлежащего общему предку сравниваемых организмов, тогда как паралоги – гены, эволюционировавшие в результате дубликации. Понятия ортологии и паралогии очевидным образом тесно связаны между собой и зависят от конкретной топологии филогенетического дерева данного семейства генов, так что дубликация в определенном узле дерева порождает новый набор паралогов в поддереве-потомке (подробнее см. в гл. 3). Более того, концептуальное определение паралогии осложнено специфичными для каждой линии эволюции потерей и горизонтальным переносом генов (см. гл. 5 и 7). Тем не менее, если не принимать во внимание эти осложнения, классификация гомологов Фитча остается центральной для эволюционной геномики ^[24].

Прерывистое равновесие и несостоятельность градуализма

Недостача межвидовых переходных форм в палеонтологической летописи – постоянная тема эволюционной биологии. Дарвин осознавал эту проблему и традиционно считал ее (так же как и палеонтологи, следующие дарвиновским традициям) следствием драматической неполноты этой летописи. Однако обширное накопление палеонтологических данных в XX веке мало помогло (если не сказать совершенно не помогло) в решении этой проблемы, что привело к возникновению иной точки зрения, сначала с появлением концепции квантовой эволюции Джорджа Гэйлорда Симпсона, затем оформившейся в концепцию прерывистого равновесия Стивена Джея Гулда и Нильса Эддриджа (Eldredge and Gould, 1997; Gould, 2002). Гулд и Эддридж собрали обширную доказательную базу, свидетельствующую о том, что история большинства видов животных, отраженная в палеонтологической летописи, соответствует в основном состоянию покоя – то есть, фактически, отсутствия изменений. Состояние покоя (стасис) перемежается «внезапным» исчезновением видов, последовательно замещаемых новыми. Следствием такой модели является очень быстрое в сравнении с продолжительностью стасиса видообразование; возникновение новых видов в определенной области является следствием миграции из области видообразования; градуалистское видообразование (постепенная трансформация видов в новые) – довольно редкий процесс. Такая модель прерывистого равновесия кажется применимой и к эволюции высших таксонов и зачастую обобщается до несостоятельности градуализма в целом, хотя правомерность такого вывода часто подвергается критике.

Пандативы, экзаптация, эволюция как ремесленник и ошибочность панглоссианской парадигмы эволюции

Пусть и неявно, но принципу градуализма был брошен вызов гипотезой Оно об эволюции генов и геномов путем дубликации, а затем, в явной форме, концепцией прерывистого равновесия. Адапционистская программа эволюционной биологии подверглась решительной, сметающей все на своем пути атаке в статье 1979 года «Пандативы Святого Марка» Гулда и Левонтина (Gould and Lewontin, 1979), одной из самых необычных и влиятельных статей в истории биологии. Гулд и Левонтин саркастически описали адапционистскую картину мира как панглоссианскую парадигму, названную так в честь примечательного персонажа вольтеровского «Кандида», который утверждал, что «все к лучшему в этом мире» [пер. Ф. Сологуба] (даже катастрофы). Гулд и Левонтин подчеркивали, что вместо того, чтобы стряпать на скорую руку «сказки просто так» ^[25] о правдоподобных адаптациях, эволюционным биологам следовало бы искать объяснение наблюдаемых черт организации биологических организмов исходя из плюралистского подхода, который принимает во внимание не только отбор, но также и внутренние ограничения, случайный дрейф и другие факторы. Метафора пандатива означает, что многие функционально важные элементы биологической организации вовсе не эволюционировали как специальные устройства для выполнения определенных функций, но скорее являются продуктами неадаптивных архитектурных ограничений, подобно пандативам (spandrels), появляющимся в арках соборов и других зданий исключительно вследствие требований конструкции, и могут использоваться для различных целей, например для украшения собора (см. рис. 2–2). Процессу использования пандативов в биологии было дано специальное название *экзаптация*, и Гулд провозгласил его важным путем эволюции (Gould, 1997a). Концепция пандативов связана с почти нейтральной теорией, но в каком-то смысле идет дальше и подходит ближе к сути эволюционного мышления, показывая, что даже те фенотипические черты, которые выглядят как типичные адаптации, не обязательно эволюционировали под давлением естественного отбора.



Рис. 2–2. Один из пандативов базилики Святого Марка в Венеции. Фото Марии Шнитцмейер, Викисклад.

В более ранней статье по сходной тематике Франсуа Жакоб (один из первооткрывателей регуляции генов и автор нескольких других плодотворных идей в бактериальной генетике, см. гл. 5) ввел метафору мастера-самоучки. Отталкиваясь прежде всего от сравнительного анализа механизмов развития, Жакоб положил в основу своих рассуждений, что эволюция действует не как инженер или дизайнер, а скорее как ремесленник-самоучка, причем чрезвычайно зависимый от предыдущего опыта при решении стоящих перед ним проблем: «Сложно проводить аналогии между естественным отбором и какими-то аспектами человеческого поведения. Однако если очень хочется поиграть в сравнения, то можно сказать, что естественный отбор работает не как изобретатель или инженер. Он работает как дилетант – мастер на все руки, который не знает точно, что он собирается создать, и при этом использует все, что подвернется под руку, будь то обрывки ниток, куски дерева или старые коробки; короче, он действует как тот мастерской, который использует все, что есть в его распоряжении, чтобы сделать хоть что-то, лишь бы работало» (Jacob, 1977).

Ключевым выводом концепции ремесленника-самоучки становится то, что итоговый результат эволюции непредсказуем, или по крайней мере его невозможно предсказать, не зная в деталях всех предшествующих событий. Другими словами, если взять и «проиграть эволюционную пластинку заново» (любимая метафора Гулда) в некоем мысленном эксперименте, то результат будет отличен от того, что мы наблюдаем в реальности, возможно до неузнаваемости; мы вернемся к этому обсуждению позднее в этой книге (см. гл. 13).

Эволюция в мире микробов и вирусов и трехдоменное древо жизни

Вероятно, в ходе развития биологии наибольшее влияние на изменение представления об эволюции оказало распространение эволюционных исследований на мир микробов, а именно одноклеточных эукариот (протист), прокариот (бактерий и архей) и вирусов. Дарвиновское представление об эволюции и все достижения эволюционной биологии нескольких последующих десятилетий базировались исключительно на исследовании животных и растений, тогда как одноклеточные эукариоты (протисты) и бактерии (монеры) были сугубо номинально размещены у корня древа жизни Эрнстом Геккелем и его последователями. Хотя к 1950-м годам генетический анализ бактериофагов и бактерий продвинулся настолько, что стало очевидным, что эти формы жизни обладают эволюционирующими геномами, СТЭ не принимала во внимание эти открытия. То, что бактерии (не говоря уже о вирусах) эволюционируют по тому же самому принципу и с использованием тех же механизмов, что и животные и растения, отнюдь не очевидно, учитывая все их разительные биологические отличия от многоклеточных организмов, и в особенности из-за отсутствия у них типичного полового размножения и репродуктивной изоляции, ключевых для видообразования среди животных и растений.

Фактически прокариоты стали «видны» эволюционным биологам в 1977 году, после выхода революционной работы Вёзе и его коллег по филогенезу рРНК (Woese, 1987) [\[26\]](#). Рассмотренное в общем контексте, открытие Вёзе является эпохально важным и, возможно, даже заслуживает сравнения с открытием структуры ДНК. Вёзе установил, что в одной молекулярной структуре, а именно последовательности нуклеотидов рРНК, выявляется очевидная консервативность во всем диапазоне клеточных форм жизни. Кроме того, чрезвычайно информативным оказался и филогенетический анализ этой универсальной консервативной молекулы: он показал, что рРНК, в некотором приближении, эволюционирует с постоянной скоростью, то есть подчиняется модели молекулярных часов. Это привело к еще одному важному открытию, ставшему одним из символов эволюционной биологии конца XX века, – *трехдоменному древу жизни* (см. рис. 2–3; Woese et al., 1990). Три доменами являются бактерии, археи и эукариоты. Домен архей был открыт Джорджем Фоксом и Вёзе сравнительным анализом рРНК, когда в новой группе ничем, казалось бы, не примечательных «бактерий» обнаружили существенные отличия как от остальных бактерий, так и от более сложных эукариотических организмов. В дополнение к разграничению трех доменов, Вёзе и его коллеги использовали филогенетический анализ рРНК для идентификации нескольких основных ветвей архей и бактерий (Woese, 1987). Из этого следовало, что эволюция прокариот столь же доступна для изучения, как эволюция сложных эукариот, – концепция, чуждая микробиологам до работы Вёзе (Stanier and Van Niel, 1962). Благодаря достижениям Вёзе, его сотрудников и последователей появилась все усиливающаяся тенденция приравнивать филогенетическое древо рРНК, с его трехдоменной структурой, к древу жизни Дарвина и Геккеля (Rase, 2009a, 2006). В течение нескольких лет после публикации открытий Вёзе стало ясно, что топологически древо рРНК (по крайней мере, в своих основных чертах) конгруэнтно деревьям некоторых из самых консервативных белков, таких как рибосомные белки, факторы трансляции, субъединицы ДНК-зависимой РНК-полимеразы и мембранные АТФазы.

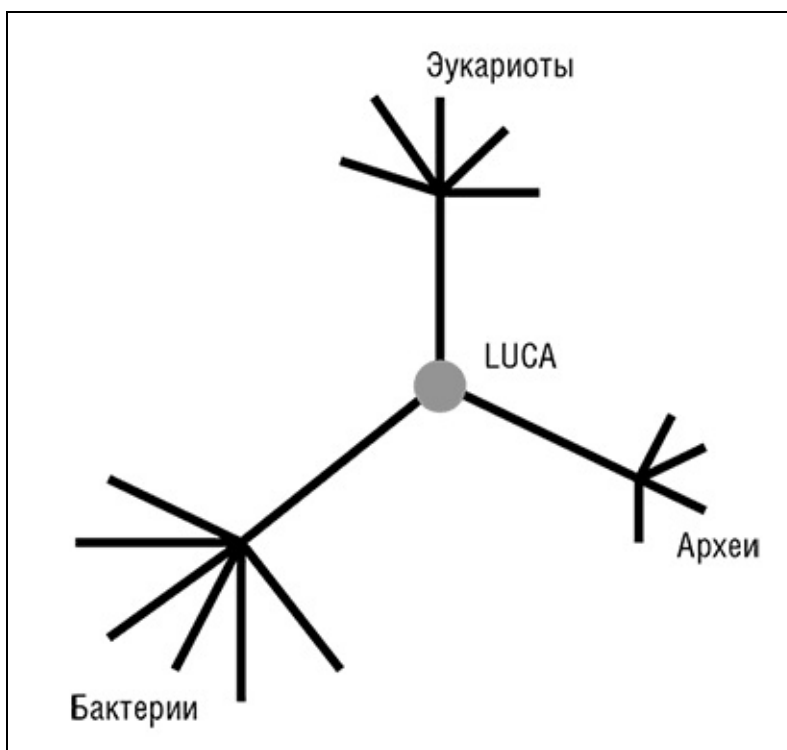


Рис. 2–3.Трехдоменное древо жизни Вёзе.

Две группы исследователей независимо друг от друга пришли к блестящей идее о том, как определить положение корня в эволюционном дереве, которое до этого было бескорневым (рис. 2–3). Для этой цели можно использовать древние паралоги, которые представлены в (почти) всех организмах и, таким образом, можно с уверенностью заключить, возникли в результате дупликации, предшествующей последнему общему предку всех живых организмов (LUCA). Когда дерево строится совместно для двух паралогичных множеств древних ортологов, положение корня между ними определено однозначно, и таким образом корень может быть выведен для каждого из множеств ортологов (см. рис. 2–4; Gogarten et al., 1989; Iwabe et al., 1989). Результаты анализа двух пар древних паралогов, факторов трансляции и субъединиц мембранных АТФаз были полностью совместимы и поместили корень на бактериальную ветвь, установив таким образом кладу архей-эукариотов (см. рис. 2–4). Тем не менее даже в догеномную эпоху было ясно, что не все деревья белок-кодирующих генов имеют ту же топологию, что и дерево рРНК; причины этих отличий оставались неясными и, как предполагалось, включали (за исключением артефактов метода) горизонтальный перенос генов (ГПГ. Smith et al., 1992). Эти расхождения оставались лишь интересным дополнением к трехдоменному ДЖ, но все резко изменилось с наступлением эры геномики.

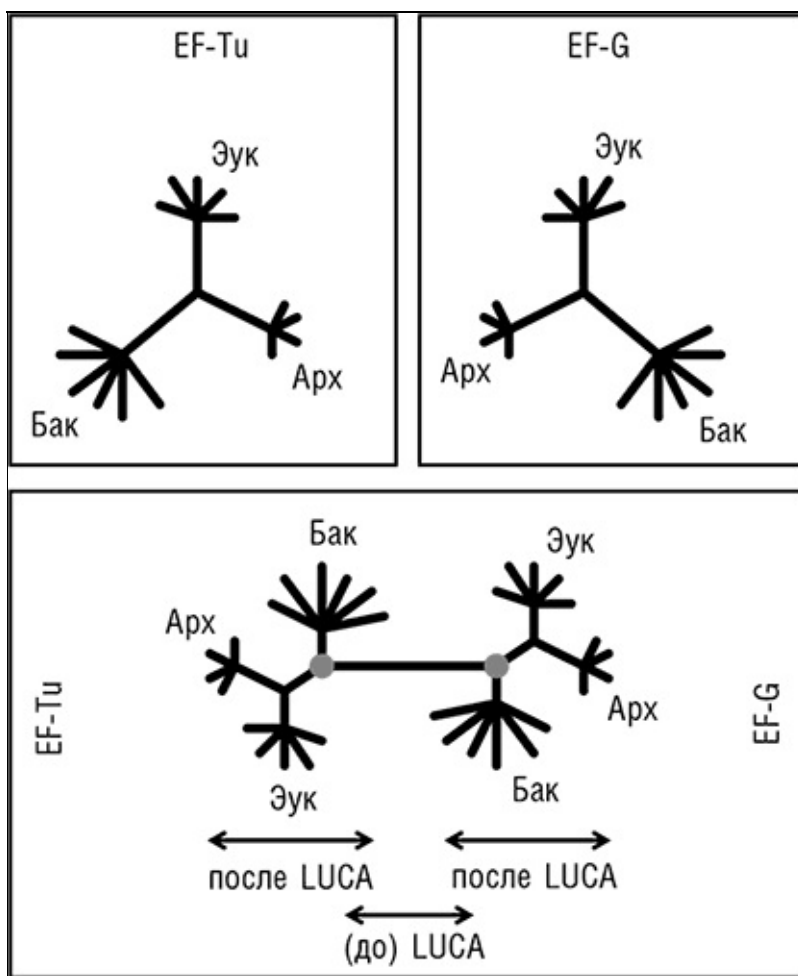


Рис. 2–4. Определение положения корня в трехдоменном древе жизни с помощью древних паралоогов. Схематически показаны филогенетические деревья двух широко распространенных факторов инициации трансляции EF-Tu и EF-G, реконструированные независимо (верхние диаграммы) и совместно (нижняя диаграмма). Кружками обозначено вычисленное положение корня в каждом из двух деревьев.

Вирусы и рождение эволюционной геномики

Эволюционная геномика родилась более чем за десять лет до исторического заявления о секвенировании первого бактериального генома. С меньшей помпой (но и не в неизвестности) было секвенировано несколько небольших (в рамках 4—100 Кб) различных вирусных геномов, и были разработаны принципы сравнения геномов, наряду с практическими вычислительными методами. Вирусы являются облигатными внутриклеточными паразитами, и вирусные геномы намного меньше и качественно отличаются от геномов клеточных форм жизни. Вирусы обычно не лишены некоторых классов генов, вездесущих и незаменимых в клеточных организмах, таких как гены компонентов системы трансляции и биогенеза мембран. Тем не менее вирусы следуют своей собственной «биологической стратегии» и кодируют все субъединицы вириона, а также по крайней мере некоторые белки, участвующие в репликации вирусного генома. (Одна из центральных тем этой книги – ключевая роль вирусов в эволюции биосферы, так что я возвращаюсь к ней достаточно подробно в гл. 10 [\[27\]](#).) Несмотря на быструю эволюцию последовательности генома, характерную для вирусов (в особенности вирусов с РНК-геномом), эти ранние сравнительные геномные исследования успешно выявили множества генов, консервативных в больших группах вирусов (Koonin and Dolja, 1993). Возможность структурного и функционального картирования всего генома определенной формы жизни была реализована в этих исследованиях впервые, и это стало краеугольным камнем эволюционной геномики. Кроме того, было сделано непредвиденное и важное обобщение: в то время как некоторые гены консервативны для удивительно обширного разнообразия вирусов, архитектура генома, структура вириона и биологические свойства вирусов демонстрируют гораздо большую пластичность (см. гл. 5 и 10, где этот вопрос обсуждается подробнее).

Гипотеза о том, что некоторые органеллы эукариотических клеток, в частности хлоропласты растений, произошли от бактерий, не намного моложе «Происхождения...» Дарвина: некоторые исследователи высказали эту идею в конце XIX века на основе микроскопического исследования клеток растений, показавшего заметное структурное сходство между хлоропластами и цианобактериями (известными тогда как сине-зеленые водоросли). Концепция симбиогенетической эволюции была последовательно представлена Константином Мережковским в начале XX века [\[28\]](#). Тем не менее в течение первых двух третей XX века гипотеза эндосимбиоза оставалась маргинальным теоретизированием. Такое восприятие изменилось вскоре после появления в 1967 году революционной статьи Линн Саган (Маргулис), где она обобщила данные о сходстве органелл и бактерий, и в особенности о совершенно неожиданно открытых незадолго до того геномах и системах трансляции органелл. Саган сделала вывод, что не только хлоропласты, но и митохондрии произошли от эндосимбиотических бактерий (Sagan, 1967). Последующие исследования, и в особенности филогенетический анализ как генов, содержащихся в митохондриальном геноме, так и генов, кодирующих белки, которые функционируют в митохондриях и, видимо, были перенесены из митохондриального в ядерный геном, превратили гипотезу эндосимбиоза в устоявшуюся концепцию с чрезвычайно прочными эмпирическими основаниями (Lang et al., 1999). Кроме того, эти филогенетические исследования убедительно продемонстрировали происхождение митохондрий от определенной группы бактерий, α -протеобактерий. Фундаментальная роль в эволюции, которая отводится уникальным (или крайне редким) событиям, таким как эндосимбиоз, не совместима ни с градуализмом, ни с униформизмом, и является одной из основных тем в остальной части этой книги, в частности в главах 7 и 12.

Канализация и устойчивость в эволюции

Выдающийся эволюционный генетик Конрад Уоддингтон выдвинул неортодоксальную идею канализации развития, которая является частью его общей концепции эпигенетического ландшафта [\[29\]](#). Эпигенетический ландшафт – это отображение решений, принимаемых развивающимся эмбрионом, так что развитие происходит за счет движения вдоль долин, по которым проходят группы сходных траекторий. Таким образом, относительно небольшие возмущения, вызванные либо факторами окружающей среды, либо мутациями, не влияют на развитие, то есть биологические системы существенно устойчивы. Согласно концепции Уоддингтона, эта устойчивость является эволюционировавшим, адаптивным свойством биологических систем. Внешнее давление может нарушить канализацию и обнаружить скрытую изменчивость, увеличивая тем самым эволюционный потенциал популяции (Waddington and Robertson, 1966). Во времена Уоддингтона эти идеи были за пределами главного русла эволюционной биологии, но в новой концепции эволюции надежность и эволюционный потенциал занимают центральное место, как обсуждается в главе 9.

Краткий обзор и перспектива

Вскоре после того, как была создана СТЭ, в эволюционной биологии произошли разительные перемены: эволюцию стало возможно проследить непосредственно к ее основе, эволюционирующему геному. На самом глубоком концептуальном уровне эволюция путем естественного отбора и дрейфа является неизбежным следствием подверженной ошибкам репликации генетической информации, кодируемой по цифровому принципу. Эволюция перестала быть несколько абстрактным процессом накопления мутаций, наблюдаемых лишь косвенно через их фенотипический эффект. Напротив, эволюция в настоящее время рассматривается как накопление конкретных изменений различного рода, больших и малых, выявляемых прямым сравнением все более доступных генных и геномных последовательностей. Наличие градиента дивергенции последовательностей от близкородственных к далеким видам само по себе является лучшим доказательством эволюции. Эта тенденция воплощается в теории (почти) нейтральной молекулярной эволюции и, на более практическом уровне, позволяет строить осмысленные филогенетические деревья. Молекулярная филогенетика достигла высшей точки с построением трехдоменного древа жизни, первоначально обнаруженного через филогении рРНК, а затем поддержанного филогениями многих белков. Анализ древних паралогов поместил корень на бактериальную ветвь трехдоменного ДЖ. Тем не менее первые выявленные расхождения между топологиями деревьев отдельных генов подсказали, что дерево рРНК не сможет рассказать всей истории эволюции жизни.

Сравнение первых секвенированных геномных последовательностей небольших вирусов положило начало эволюционной геномике. Стало ясно, что с помощью сравнительного анализа могут быть построены структурные и функциональные карты геномов, которые нельзя было охарактеризовать никаким иным способом, и что поразительная консервативность ключевых генов идет рука об руку с пластичностью архитектуры генома.

Одновременно с завершением развития СТЭ на заре молекулярной эволюции и молекулярной филогенетики эволюционная биология догеномной эпохи включала несколько концепций, таких как пандативы и канализация, выходящих за рамки неodarвинизма. В результате быстрый расцвет геномики в 1990-х годах происходил на фоне сложного, разнообразного ландшафта эволюционной теории и методологии.

Рекомендуемая дополнительная литература

Dawkins, Richard.(1976/2006) *The Selfish Gene*. Oxford: Oxford University Press. (Перевод: Докинз Р. Эгоистичный ген / Пер. с англ. Н. Фоминой. М.: Мир, 1993.)

Книга неоченимой концептуальной значимости, в которой впервые представлена плодотворная идея отбора на уровне индивидуальных генов.

Gould, S. J., and N. Eldredge.(1993) *Punctuated Equilibrium Comes of Age*. *Nature*366: 223–227. Краткий, но убедительный обзор данных в поддержку модели прерывистого равновесия.

Gould, S. J., and R. C. Lewontin.(1979). *The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme*. *Proceedings of the Royal Society of London B Biological Sciences*205: 581–598.

Вероятно, одна из самых выдающихся статей во всей истории биологии, замечательная как непримиримым отстаиванием плюралистического взгляда на эволюцию в противовес пандаптационизму, так и духом Возрождения, которым пропитан весь текст.

Jacob, F.(1977) *Evolution and Tinkering*. *Science*196: 1,161—1,166.

Великолепно написанная статья, которая остается незаменимой как опровержение идеи «естественный отбор как дизайнер» и обоснование определяющей роли исторического стечения обстоятельств в эволюции.

Kimura, Motoo.(1983) *The Neutral Theory of Molecular Evolution*. Cambridge, MA: Cambridge University Press. (Перевод: Кимура М. Молекулярная эволюция: теория нейтральности Пер. с англ. М.: Мир, 1985.)

Дефинитивное изложение теории нейтральной эволюции ее создателем, сочетающее скрупулезный популяционно-генетический анализ с совершенно прозрачными пояснениями для биологов. Несмотря на то что практические аспекты книги несколько устарели, она остается непревзойденной по ясности изложения и легко читаемой.

Koonin, E. V., and V. V. Dolja . (1993) *Evolution and Taxonomy of Positive-Strand RNA Viruses: Implications of Comparative Analysis of Amino Acid Sequences*. *Critical Reviews in Biochemistry and Molecular Biology*28: 375–430.

Исчерпывающий обзор ранних результатов сравнительной и эволюционной геномики обширного класса малых РНК-вирусов, одна из первых попыток глубокой эволюционной реконструкции. Основные выводы сохраняют свое значение, несмотря на последующий рост количества разнообразных секвенированных вирусных геномов.

Nei, Masatoshi, and Sudhir Kumar. (2000) *Molecular Evolution and Phylogenetics*. Oxford: Oxford University Press. (Перевод: Ней М., Кумар С. Молекулярная эволюция и филогенетика / Пер. с англ. Киев: КВЦ, 2004.)

Техническое, но вполне доступное описание широко используемых филогенетических методов, написанное автором метода объединения ближайших соседей и одним из авторов широко используемого филогенетического пакета MEGA.

Nielsen, R.(2009) *Adaptionism —30 Years After Gould and Lewontin*. *Evolution*63: 2,487—2,490.

Взгляд на «Пандативы Святого Марка» 30 лет спустя.

Ohno, Susumu.(1970) *Evolution by Gene Duplication*. Vienna: Springer.

Классическая книга, в которой описывается концепция дубликации генов как центральный путь развития эволюции.

Ohta, T., and J. H. Gillespie.(1996) *Development of Neutral and Nearly Neutral Theories*. *Theoretical Population Biology*49: 128–142.

Историческое и концептуальное освещение нейтральной и почти нейтральной теорий двумя выдающимися исследователями молекулярной эволюции.

Woese, C. R.(1987) Bacterial Evolution. *Microbiology Reviews*51: 221–271.

Блистательная, исчерпывающая работа, обобщающая ранние результаты филогенетики 16S-рРНК.

Woese, C. R., and N. Goldenfeld.(2009) How the Microbial World Saved Evolution from the Scylla of Molecular Biology and the Charybdis of the Modern Synthesis. *Microbiology and Molecular Biology Reviews*73: 14–21.

Эссе к 200-летию Дарвина, подчеркивающее исключительную важность геномики микроорганизмов для нового понимания эволюции.

Глава 3. Сравнительная геномика: эволюционирующие геномные ландшафты

Важность перехода к геномике

В догеномную эру были установлены фундаментальные принципы молекулярной эволюции и было сделано немало конкретных наблюдений, имеющих большое значение и повлиявших на основы эволюционной биологии (см. гл. 1 и 2). Но масштабные работы по расшифровке геномов, начавшиеся в середине 90-х и стремительно развивавшиеся в новом тысячелетии, качественно изменили всю эволюционную биологию. Важность обширной базы данных геномных последовательностей, имеющих различную степень расхождения, очевидна. Этот материал дает исследователям возможность изучать механизмы и отдельные события эволюции с необходимой статистической точностью и выявлять даже самые малозаметные эволюционные подвижки. Как бы то ни было, в эволюционной биологии получение разнообразных и полных геномных последовательностей чрезвычайно важно далеко не только ради накопления количества данных. Действительно, лишь полностью расшифрованный геном (в отличие от, скажем, расшифрованного лишь на 95 процентов) дает исследователю объективное и непредвзятое представление о геномном репертуаре той или иной формы жизни. Иными словами, исследователь может определить присутствие в организме тех или иных генов и, что одинаково важно, их отсутствие. Таким образом, сравнение полных геномов представляет собой единственный удовлетворяющий исследователя путь к реконструкции эволюции. Открывающаяся картина во многом отличается от всего, что можно было себе представить, оставаясь в рамках традиционной эволюционной биологии.

Если мы действительно стремимся «понять» эволюцию, принципиально важно исследовать геномные образцы как вглубь (для этого необходимы геномные последовательности множества близкородственных представителей одного и того же таксона), так и вширь (для этой цели нужны последовательности как можно большего числа различных таксонов – в идеале всех таксонов). Ко времени написания этих строк, в последние дни 2010 года, собрание секвенированных геномов состояло из нескольких тысяч геномов вирусов, более чем тысячи геномов бактерий и архей, а также приблизительно сотни геномов эукариот. Ко времени издания этой книги геномная база данных почти удвоится, а благодаря новому поколению методов секвенирования в предстоящие годы ее темпы роста должны еще более ускориться [\[30\]](#). Несмотря на то что не все основные таксоны должным образом охвачены, быстро пополняющееся собрание геномов все более отвечает потребностям исследований как в области микроэволюции, так и в области макроэволюции.

Успехи традиционной геномики дополняют и стремительно накапливающиеся в последнее время, обширные по объему данные по метагеномике – а именно всеобъемлющее (или, по меньшей мере, обширное) секвенирование нуклеиновых кислот форм жизни из разнообразных сред обитания. Хотя применяемые в настоящее время в метагеномике подходы обычно не обеспечивают полную расшифровку геномов, они предоставляют бесценную, объективную информацию о разнообразии жизни в различных средах.

В данной главе представлен обзор разнообразия и основных характеристик геномов. В последующих главах подробно исследуется влияние результатов сравнительных геномных исследований на развитие «постсовременной» синтетической теории эволюционной биологии.

Геном стал первым термином с окончанием «-ом» – и до сих пор является наиболее употребительным термином этой группы [31]. Как это всегда бывает в биологии, определить, что же такое геном, нелегко. Говоря просто, геном – это генетическая информация конкретного организма во всей ее полноте. Существование стабильного ядра унаследованной генетической информации (а более конкретно, генов) вытекает из самого факта существования надежной наследственности, а в терминах более фундаментальных – из принципа подверженной ошибкам репликации (ПОР, см. гл. 2). Однако связь между «генетической информацией во всей ее полноте» и «стабильным ядром» не так уж проста. Стоит, к примеру, задать на первый взгляд невинный вопрос: «Что есть геном кишечной палочки *Escherichia coli*?» – как тут же возникает целый ряд серьезных затруднений. А вопрос «Что такое геном человека?» вызывает свои, не менее сложные проблемы. Вернемся мы к этому обсуждению позднее (см. гл. 5), а сейчас рассмотрим многообразие геномов, расшифрованных за последние 15 лет.

Новая эра геномики наступила на исходе лета 1995 года. Тогда лаборатория Дж. Крейга Вентера опубликовала результаты секвенирования генома условно-патогенной бактерии гемофильного гриппа *Haemophilus influenzae* (Fleischmann et al., 1995). В процессе расшифровки геномной последовательности *H. influenzae* Вентер, Гамильтон Смит и их коллеги усовершенствовали так называемый «метод дробовика». Этот подход грубого деления генома на короткие произвольные участки с расшифровкой их по частям и последующим восстановлением полной геномной последовательности быстро превратил секвенирование длинных нуклеотидных цепочек в рутинное дело. В течение года были расшифрованы геномы некоторых других бактерий, первый геном археи (*Methanocaldococcus jannaschii*) и первый геном эукариота (пекарские дрожжи *Saccharomyces cerevisiae*) (Koonin et al., 1996). К 1999 году установился стабильный экспоненциальный рост коллекции секвенированных геномов (см. рис. 3–1).

В диапазоне от вирусов до животных геномы различаются по размеру на шесть порядков – от нескольких тысяч до нескольких миллиардов нуклеотидов; для клеточных организмов, исключая вирусы, ширина диапазона составляет четыре порядка (см. рис. 3–2). По количеству генов диапазон значительно уже и составляет всего около четырех порядков, от двух-трех генов у простейших вирусов до приблизительно 40 тысяч генов у некоторых животных. Если же исключить вирусы и паразитические (симбиотические) бактерии, диапазон по числу генов становится довольно узким, немногим более одного порядка (см. рис. 3–2; Koonin, 2009a; Lynch, 2007c). Кажется весьма удивительным, что млекопитающие или цветковые растения имеют всего примерно в десять раз больше (легко идентифицируемых) генов, чем какая-нибудь средняя свободно живущая бактерия, и лишь примерно в два раза больше, чем бактерия из разряда наиболее сложных (см. рис. 3–2). Далее в книге рассматриваются всевозможные объяснения этих явных ограничений по числу генов в геномах всех форм жизни (см. гл. 5, 7 и 10).

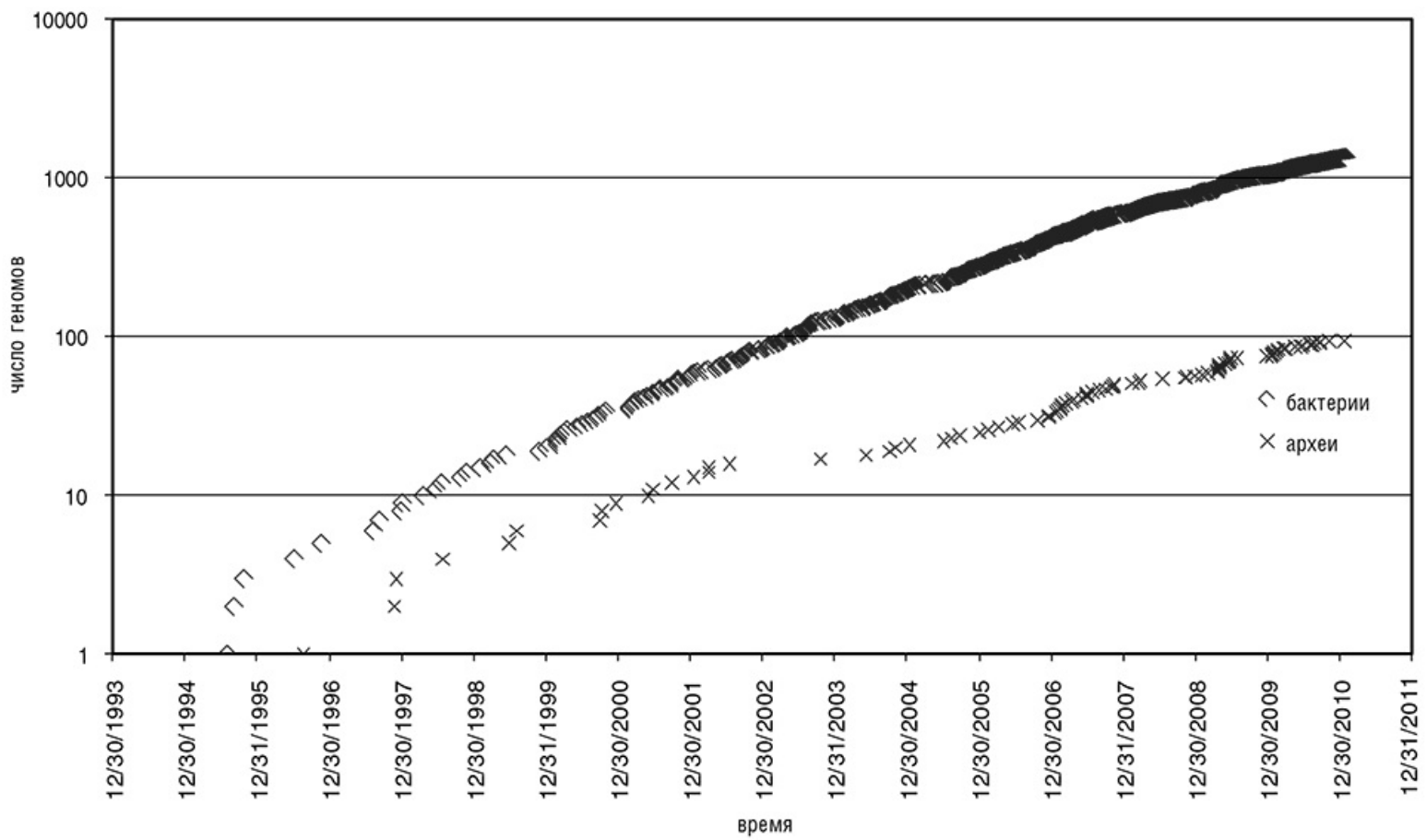


Рис. 3–1. Экспоненциальный рост коллекции секвенированных геномов. Данные с веб-сайта Национального центра биотехнологической информации (www.ncbi.nlm.nih.gov/genome/)

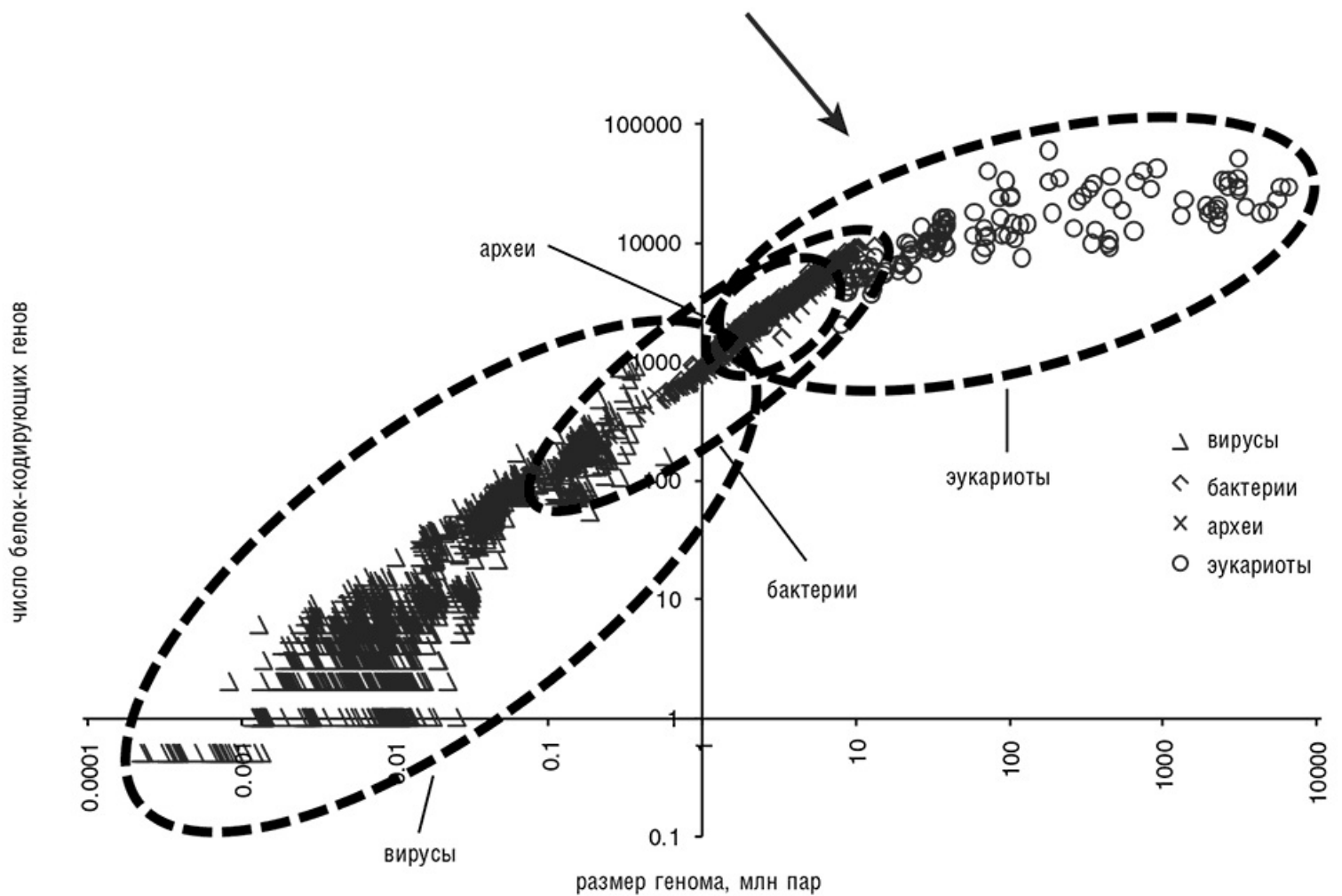


Рис. 3–2. Общий размер геномов и число генов у вирусов, бактерий, архей и эукариот. Данные с веб-сайта Национального центра биотехнологической информации. Представлено в двойном логарифмическом масштабе. Стрелка указывает на точку изменения наклона кривой, соответствующую переходу от «малых» к «большим» геномам.

Грубо говоря, геномы могут быть разделены на два четко выделенных класса (Koonin, 2009a). Граница, разделяющая эти классы, находится в точке изменения наклона кривой на графике, представленном на рис. 3–2.

1. Геномы со строгим соответствием между размером генома и числом генов. К ним относятся геномы всех вирусов и прокариот, имеющие огромную плотность генов от 0,5 до 2 генов на тысячу пар оснований и очень короткие участки между генами (10–15 процентов длины генома и даже меньше), состоящие главным образом из регуляторных элементов. Иногда, говоря о таких геномах, вспоминают ковер «от стены к стене» (wall to wall genomes) [32], так как они почти полностью состоят из легко определяемых генов. Геномы большинства одноклеточных эукариот демонстрируют несколько меньшую зависимость между размером генома и числом генов, чем геномы вирусов и прокариот, тем не менее они могут быть отнесены к этому же классу.

2. Геномы, у которых нет четкой взаимосвязи между размером генома и числом генов, в частности большие геномы многоклеточных и некоторых одноклеточных эукариот. Здесь в лучшем случае наблюдается слабая корреляция между общим размером генома и числом генов. Соответственно, доля генома, занимаемая межгенными участками (а также другими некодирующими последовательностями, такими как интроны), сильно варьирует. В некоторых

наиболее сложных геномах, в частности у млекопитающих, основную часть генома составляют именно некодирующие последовательности.

Вариабельность размеров генома и числа генов дополняется разнообразием в других измерениях – например, в физической организации и композиции нуклеотидов. При рассмотрении как вирусной, так и клеточной формы жизни геномы предстают во всевозможных формах нуклеиновых кислот (подробнее см. в гл. 10). Все геномы клеточных организмов состоят из двухцепочечных ДНК, однако количество геномных сегментов (хромосом) и их размеры, форма (кольцевая или линейная), а также пloidность (число наборов) широко разнятся. Азбучная истина гласит, что прокариоты имеют гаплоидные, простые кольцевые хромосомы, в то время как у геномов эукариот, сильно различающихся по пloidности, гены распределены между множеством линейных хромосом. И хотя такие геномные формы, по-видимому, действительно доминируют, на самом деле разнообразие геномов выходит далеко за рамки такого простого дихотомического разделения. В частности, у многих прокариот имеется несколько хромосом, в отдельных случаях – линейных. Более того, вопреки распространенному заблуждению, у прокариот большинство клеток не гаплоидные, то есть они содержат несколько копий генома.

Древние гены составляют в геноме большинство и имеют отчетливую эволюционную судьбу

Как уже упоминалось в предыдущей главе, Эрнст Майр, великий эволюционист XX века и один из основателей СТЭ, с уверенностью предсказывал исходя из больших фенотипических различий между организмами, что гены разных организмов, даже близкородственных, не будут иметь узнаваемого сходства. Ошибочность этого предсказания оказалась просто феерической, что само по себе делает его нетривиальным и ценным. Сравнение последовательностей даже в догеномный период выявило высокий консерватизм последовательностей у некоторых гомологичных белков и молекул некодирующих РНК по всему спектру жизни, от бактерий до млекопитающих (см. предыдущую главу). Более того, высокая степень сходства последовательностей существует у древних паралогов, которые, по-видимому, происходят от копий, ведущих свое происхождение от LUCA (Gogarten et al., 1989; Iwabe et al., 1989). Геномика позволяет перевести это общее понимание в количественное разбиение генов любого генома на классы эволюционной консервативности (см. рис. 3–3; Koopin and Wolf, 2008b).

Ключевое открытие сравнительной геномики состоит в том, что большинство генов в каждом геноме могут считаться высококонсервативными – они имеют легко обнаруживаемые гомологи в организмах, разделяемых сотнями миллионов лет эволюции (например, в случае генов человека, на уровне общего предка позвоночных; см. рис. 3–3; Wolf et al., 2009). Это открытие демонстрирует поразительную устойчивость последовательностей РНК и белков в процессе эволюции: *типичное время исчезновения сходства последовательностей у гомологичных генов сравнимо со временем существования жизни на Земле*. Помимо основополагающего значения, данный факт имеет огромные практические последствия: благодаря ему, прежде всего, сравнительная геномика становится крайне информативной и действенной.

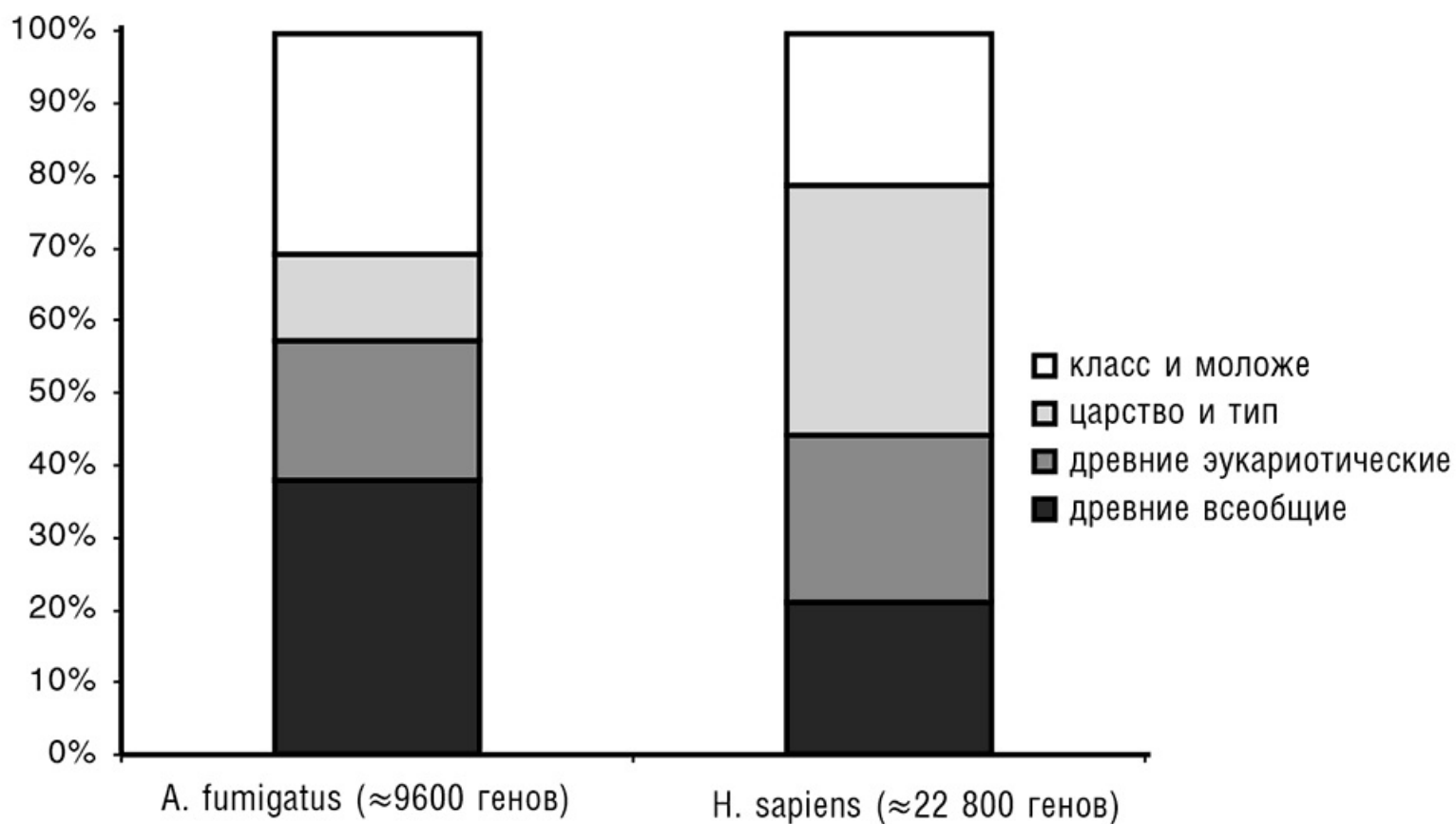


Рис. 3–3. Распределение генов по эволюционному возрасту. «Эволюционный возраст» соответствует самому старому таксономическому узлу, в котором могут быть определены гомологи для белка, производимого данным геном. В частности, для человека *древние всеобщие* означает «гомологи, обнаруживаемые у прокариот», *древние эукариотические* означает «гомологи, обнаруживаемые у прокариот вне супергруппы униконтов» (см. гл. 7), *царство и тип* означает «гомологи, обнаруживаемые у животных вне класса млекопитающих», а *класс и моложе* означает «вне класса млекопитающих гомологи не обнаружены» (данные по Wolf et al., 2009)

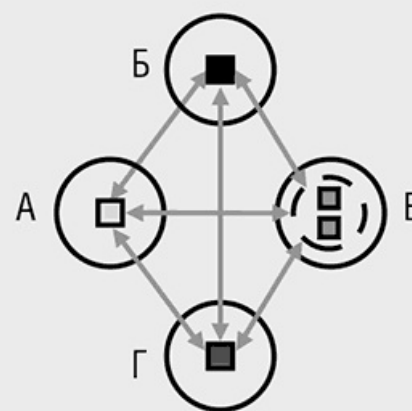
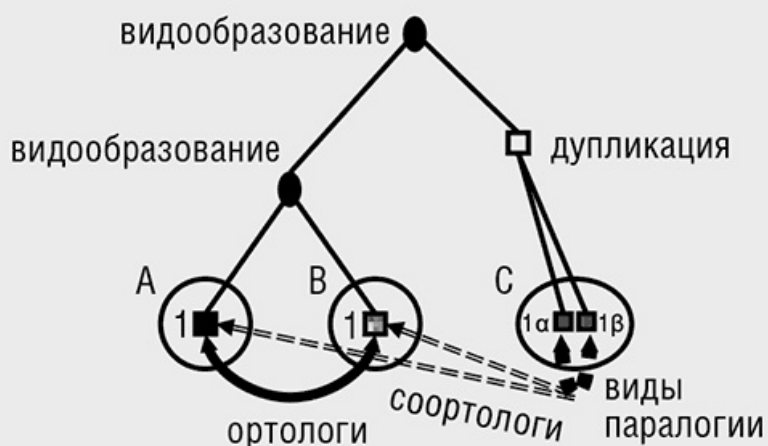
Структуру эволюционного процесса определяют не только консервативные последовательности. На протяжении чрезвычайно длительных эволюционных периодов не просто сохраняется сходство последовательностей РНК и белков, но и гены имеют свойство сохранять свою уникальность. Иными словами, большинство генов развиваются как *ортологичные линии*, с редкими случаями дупликации (Koonin, 2005). Устойчивость ортологии генов становится очевидной благодаря простой процедуре, широко применяемой в сравнительной геномике и позволяющей эффективно выявлять ортологичные наборы генов. При этом ортологи обнаруживаются как «наилучшие совпадения при двунаправленном сравнении» (bidirectional best hits): все закодированные в геноме белковые последовательности сравниваются со всеми белками, закодированными в другом геноме, а затем процедура повторяется в обратном направлении (Tatusov et al., 1997). Пары генов, дающие наилучшие совпадения (те, которые демонстрируют наибольшее сходство последовательностей) при обоих направлениях сравнения, считаются возможными ортологами; нетрудно применить эту процедуру к нескольким видам путем совмещения треугольников двунаправленных совпадений, имеющих общую сторону (см. табл. 3–1). Примечательно, что такой прямолинейный подход в большинстве случаев хорошо срабатывает: к примеру, порядка 70 процентов генов организмов,

разделенных приблизительно 100 миллионами лет эволюции, таких как люди и мыши, легко идентифицируются как ортологи при помощи описанной процедуры (Wolf et al., 2009). Если применить простую модификацию этого алгоритма и включить дубликации генов, характерных для одной линии наследования (дубликации, образовавшиеся после расхождения сравниваемых видов), такой подход позволяет идентифицировать наборы ортологов (известных как *кластеры ортологичных генов*, КОГ) во многих геномах, в том числе столь удаленных друг от друга, как археи и бактерии – представители двух доменов прокариот (см. гл. 5). Более точные и мощные способы обнаружения ортологов требуют подробного анализа филогенетических деревьев (см. табл. 3–1); впрочем, результаты такого анализа обычно близки к тем, что дают более простые методы, основанные только на сравнении последовательностей. Разумеется, для части генов история дубликаций и потерь настолько сложна, что обнаружить КОГ трудно, поэтому они становятся нечеткими кластерами с неопределенной внутренней структурой. По счастью, этих «трудных» генов в каждом геноме относительно немного.

Таблица 3–1. Классификация гомологичных связей генов: ортологи, паралоги и методы их определения.

Эволюционные связи генов:

- Гомология: гены, имеющие общее происхождение.
- Ортология: гомологичные гены, эволюционировавшие путем видообразования.
- Паралогия: гомологичные гены, эволюционировавшие путем дубликации.
- Ксенология: гомологичные гены, имитирующие ортологи, но образовавшиеся в результате горизонтального переноса гена из другой ветви.
- Паралогия, внутренняя и внешняя: паралогичные гены, возникшие в результате видоспецифической дубликации после (внутренняя) или до (внешняя) определенного события видообразования.
- Со-ортология: внутренне-паралогичные гены, совокупно ортологичные по отношению к генам другой ветви (из-за их общего происхождения в ходе видообразования).
- Ортологичная группа (КОГ): совокупность всех потомков данного предкового гена.



КОГ с (со)ортологами 4 видов (А, Б, В, Г) и схема вычисленных двусторонних совпадений

Изначально не вполне складная аббревиатура КОГ относилась к кластерам ортологичных групп (белков), чтобы обозначать соортологичные связи, вызванные дубликацией генов (см. табл. 3–1; Tatusov et al., 1997). Сейчас я предпочитаю расшифровывать КОГ просто как кластеры ортологичных генов, однако само по себе это сокращение остается чрезвычайно удобным для обозначения фундаментального свойства таких кластеров. Эта трехбуквенная аббревиатура широко используется в литературе, и я использую ее в данной книге в качестве сокращенного названия наборов ортологичных генов. Обычно каждый секвенированный геном более чем на 70 процентов состоит из генов, относящихся к КОГ (см. рис. 3–4). В эволюции генома, к которой мы обращаемся в этой книге неоднократно, эта величина представляется важной. Таким образом, существенное большинство генов в каждом геноме весьма консервативно, то есть представлено ортологами во многих далеко отстоящих друг от друга организмах.

Мультидоменные белки и сложность связей ортологов

В этой главе основной упор делается на рассмотрении взаимосвязи между стабильностью и изменчивостью в ходе эволюции. В настоящем разделе мы сосредоточим внимание на отдельных элементах белковой структуры, доменах и мультидоменной организации многих белков (Doolittle, 1995). Таким образом, мы заглянем по другую сторону генной эволюции, которая противостоит стабильности ортологичных линий, отмеченной ранее, и дополняет ее. Домен – центральное понятие в исследовании белков, и определение ему можно дать по меньшей мере на двух уровнях. По первому определению, домены представляют собой компактные элементы белковой структуры с характерными размерами около ста аминокислотных остатков. В этой главе нас интересуют родственные связи геномов, в частности ортология, поэтому необходимости рассматривать структурные элементы нет. Второе определение доменов относится к компактным единицам эволюции, которые могут охватывать один или несколько структурных элементов; здесь нас интересуют именно такие *эволюционные домены*.

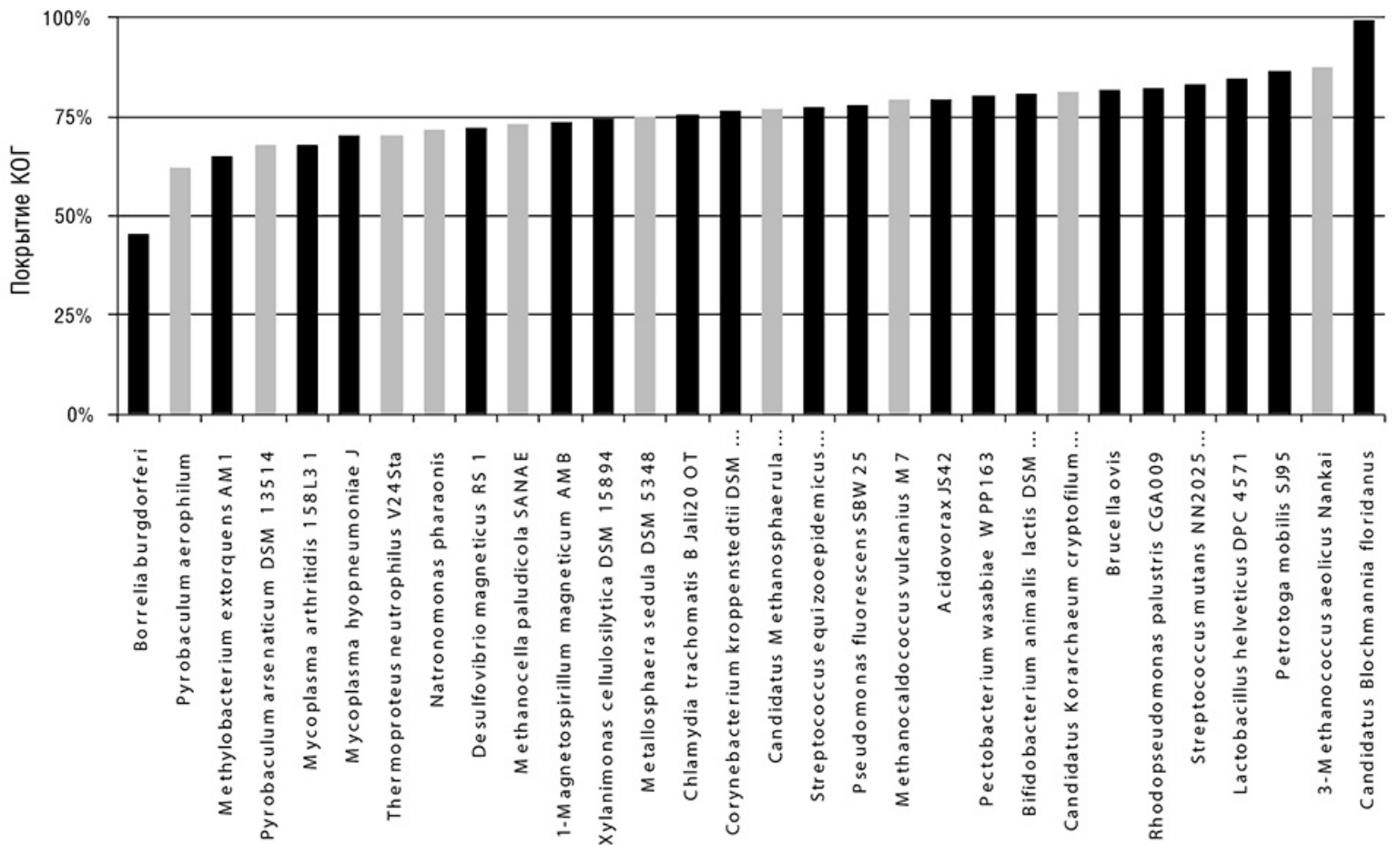


Рис. 3–4. Уровень покрытия КОГ в геномах архей и бактерий. Полные наборы белков в 20 отобранных геномах бактерий (показаны черным) и 10 геномах архей (показаны серым), отнесенные к КОГ (Tatusov et al., 2003). Применялся метод COGNITOR (Makarova et al., 2007b)

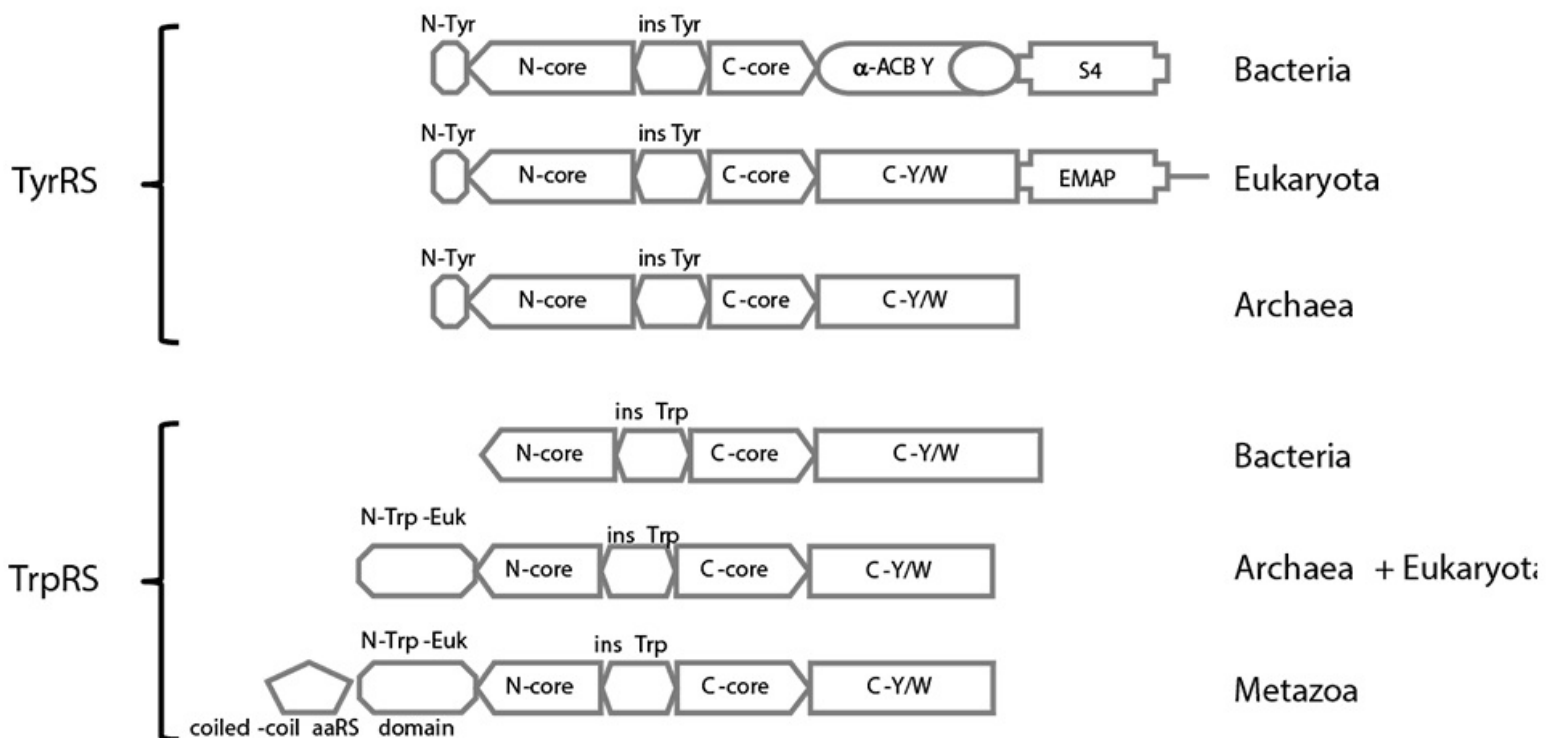


Рис. 3–5. Разнообразие мультидоменной архитектуры гомологичных белков. На схеме сравнивается доменная архитектура двух паралогичных наборов древних и функционально

незаменимых для всех организмов ортологичных белков: тирозил-тРНК синтетаза (TyrRS) и триптофанил-тРНК синтетаза (TrpRS). Каждый домен обозначен своей собственной геометрической формой (по Wolf et al., 1999a).

Мультидоменные белки обнаружены у всех форм жизни, но особенно характерны для сложных многоклеточных эукариот (Koonin et al., 2000a; Koonin et al., 2000b). Доменная архитектура этих белков демонстрирует различную степень эволюционной пластичности. Изменчивость особенно выражена у белковых архитектур, включающих так называемые «неразборчивые домены» (promiscuous domains), имеющие склонность к слиянию с разнообразными другими доменами (Basu et al., 2009). Разнообразная мультидоменная архитектура белков запутывает понятие ортологии. Считается, что в ходе долгой эволюции ортологичные гены сохраняют свою уникальность, в том числе функциональную (имеют одну и ту же эволюционную историю). Однако это правило нарушается в тех случаях, когда гены, казалось бы подпадающие под определение ортологии (см. табл. 3–1), меняют доменную архитектуру (см. рис. 3–5): в этих случаях лишь части соответствующих белков в разных организмах имеют одну и ту же эволюционную историю и выполняют одни и те же функции (хотя второе и не может быть гарантировано, поскольку взаимодействие доменов вполне может иметь существенные функциональные последствия).

Контраст между эволюционной пластичностью генома и стабильностью индивидуальных генов

Мы видели, что большинство генов в каждом геноме весьма консервативно: гомологи этих генов – чаще всего легко определяемые ортологи – обнаружены у организмов, эволюционно далеких друг от друга. Тем не менее эта поразительная эволюционная устойчивость генов – лишь одна сторона медали сравнительной геномики. Другая же, обратная сторона – это «текучесть» генного набора и архитектуры геномов всех форм жизни. Геномы прокариот особенно подвержены изменчивости. Наглядным примером этого является сравнение различных штаммов классической модели бактерий, лабораторного штамма K12 и нескольких патогенных штаммов кишечной палочки *Escherichia coli* (Perna et al., 2001). Последовательности ортологичных генов у этих бактерий почти одинаковы, однако некоторые патогенные штаммы имеют на 30 процентов больше генов, чем штамм K12, и генные наборы патогенных штаммов радикально различаются. Неизбежно возникает заключение, что «лишние» гены, формирующие так называемые островки патогенности, одними штаммами были приобретены, а другими утеряны (в гл. 5 мы еще вернемся к этой теме).

В более общем плане можно измерить дистанцию между геномами, сравнив, с одной стороны, последовательности консервативных генов-маркеров, таких как рРНК или рибосомных (р) белков, а с другой стороны, исследовав ту часть генов, что формирует легко узнаваемые пары совпадающих ортологов (см. табл. 3–1). В отличие от постепенного, относительно медленного изменения нуклеотидной последовательности генов, наблюдается резкое несовпадение генных наборов (см. рис. 3–6). Заметим, что нет никакого противоречия между этим наблюдением и выводом о том, что для значительного большинства генов в геноме бактерии или археи имеются ортологи в *некоторых* эволюционно удаленных от них организмах. Здесь слово *некоторых* ключевое, поскольку у многих генов в любом геноме разное эволюционное происхождение и разная история, и потому их ближайшие родственники могут быть обнаружены в разных таксонах (см. гл. 5). Дистанцию между геномами, определяемую как доля общих (ортологичных) генов, можно использовать для описания «геномной вселенной»,

рассматриваемой далее в этой главе, а также для построения особого рода дерева эволюции (см. гл. 5).

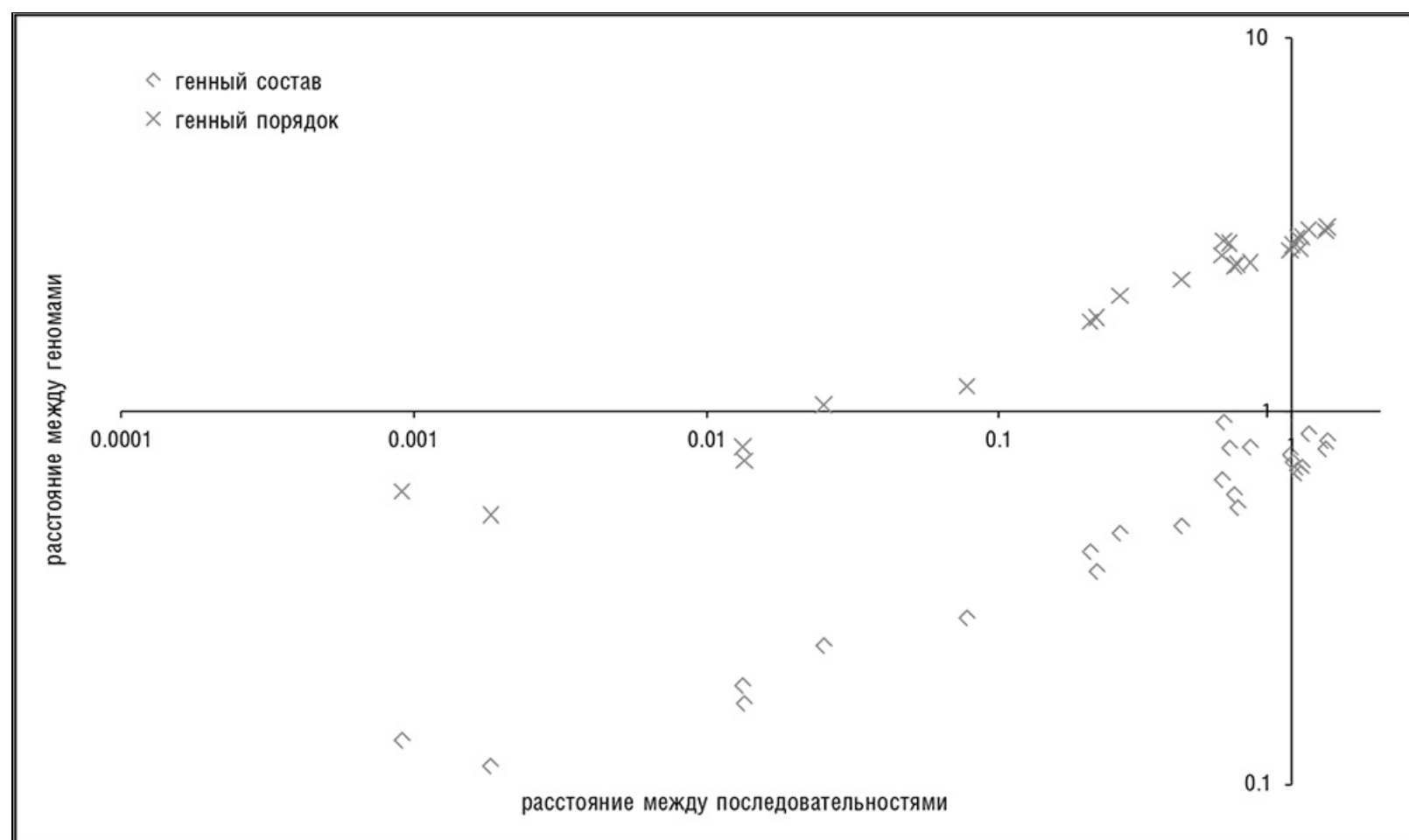


Рис. 3–6. Расхождение порядка генов и генных составов между бактериями по сравнению с расхождением высококонсервативных последовательностей белков. Были вычислены расстояния от кишечной палочки K12, штамм MG1655, до 24 других разнообразных протеобактерий. Расстояние между последовательностями: расстояние наибольшего подобия (maximum likelihood distance) для соединенных выравниваний рибосомных белков вычислено с использованием программы PROTDIST пакета программного обеспечения филогенетического анализа Phylip (Felsenstein, 1996). Расстояние между генными порядками: $-\ln(JCOG)$, где $JCOG$ – коэффициент подобия (коэффициент Жаккара) для набора КОГ в двух геномах. Расстояние между генными составами: $-\ln(JPAIR)$, где $JPAIR$ – коэффициент Жаккара для множества неупорядоченных пар соседних КОГ в двух геномах. График исполнен в двойных логарифмических координатах.

Геномная архитектура, то есть расположение генов в геноме, проявляет еще большую эволюционную нестабильность, чем генный состав геномов, что контрастирует с консервативностью генных последовательностей (Koonin, 2009a; Novichkov et al., 2009). За исключением организации малых групп функционально связанных генов в оперонах, порядок генов сравнительно слабо сохраняется даже у близкородственных организмов [33]. У прокариот сохранение порядка генов на большом протяжении генов не просматривается даже в некоторых группах геномов, которые сохраняют почти однозначное соответствие ортологичных генов и в среднем более 99 процентов идентичности последовательностей ортологичных белков (см. рис. 3–6). Эукариоты демонстрируют несколько большую сохранность порядка генов. Тем не менее даже в случае эукариот имеется мало общих элементов архитектуры генома между,

например, разными типами в царстве животных и вообще никаких между животными и грибами или животными и растениями.

Геномные ландшафты: распределение эволюционных ограничений по разным классам сайтов в геноме

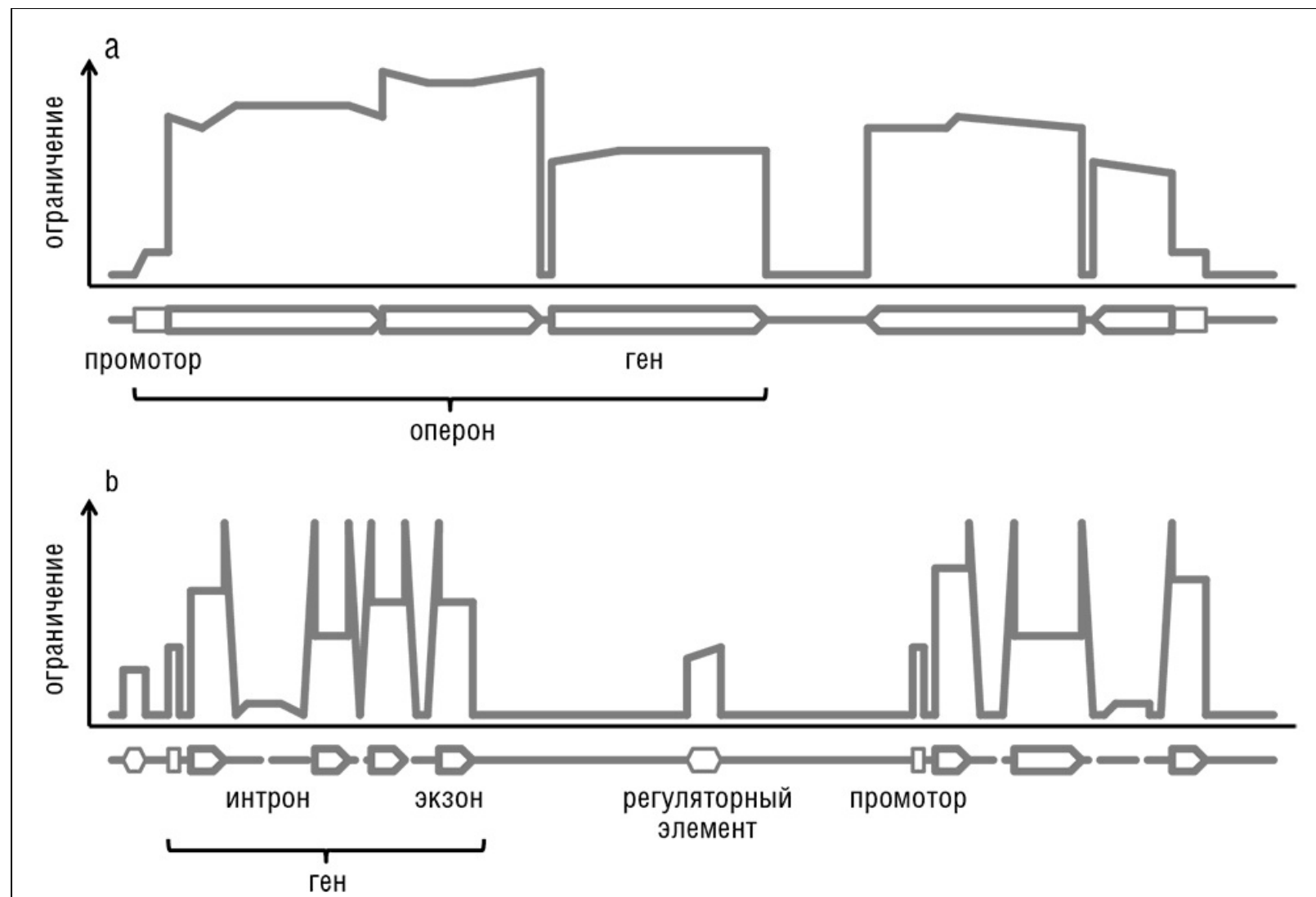


Рис. 3–7. Схематические геномные ландшафты. Распределение эволюционных ограничений по разным сайтам в геномах прокариот и эукариот вскрывает различные принципы геномной архитектуры: *а*– геном прокариот; *б*– геном эукариот.

Любой геном может быть представлен в виде геномного ландшафта, графика, напоминающего панораму города, где каждому нуклеотидному сайту присваивается высота, пропорциональная силе влияющих на него эволюционных ограничений. Ограничения имеет смысл рассматривать как меняющиеся в диапазоне от 0 (лишенная ограничений, нейтрально эволюционирующая, функционально не значимая позиция) до 1 (полностью ограниченная, функционально важная позиция, в которой изменения недопустимы, см. рис. 3–7; Koornin and Wolf, 2010b). Распределения ограничений по геному значительно отличаются у форм жизни с различными архитектурами генома. Эти отличия проявляются особенно ярко, если говорить о сравнении, с одной стороны, вирусов и прокариот с их геномами «стена к стене», в основном состоящими из генов, кодирующих белок или РНК, и, с другой стороны, многоклеточных эукариот, в геномах которых кодирующие нуклеотиды находятся в меньшинстве (см. рис. 3–7). В пересчете на один сайт, ограничения в компактных геномах, особенно у прокариот, на

несколько порядков сильнее, чем ограничения в больших геномах многоклеточных эукариот. Белок-кодирующие последовательности и последовательности, кодирующие структурные РНК, подвержены наиболее сильным ограничениям во всех геномах. Подавляющее большинство белок-кодирующих генов, особенно у прокариот, имеют низкие значения Ka/Ks , что указывает на сильное давление очищающего отбора на эти последовательности (см. рис. 3–8 и предыдущую главу). В то же время во всех группах организмов существует значительная положительная корреляция между Ka и Ks , указывающая, что даже синонимические сайты в белок-кодирующих генах ограничены примерно в пропорции к ограничениям на несинонимические сайты (Drummond and Wilke, 2008; см. также гл. 4). Учитывая, что прокариотические геномы почти полностью состоят из белоккодирующих генов со вкраплениями генов структурных РНК и коротких межгенных промежутков, в основном занятых разно образно ограниченными регуляторными регионами, эти компактные геномы содержат мало неограниченных сайтов. Заметным исключением являются псевдогены, редкие у большинства прокариот, но распространенные у некоторых паразитических бактерий, особенно растущих внутри эукариотических клеток, например *Rickettsia* и *Mycobacterium leprae* (Harrison and Gerstein, 2002). Геномы большинства вирусов еще более компактны, чем геномы прокариот, причем почти вся последовательность генома занята белок-кодирующими генами.

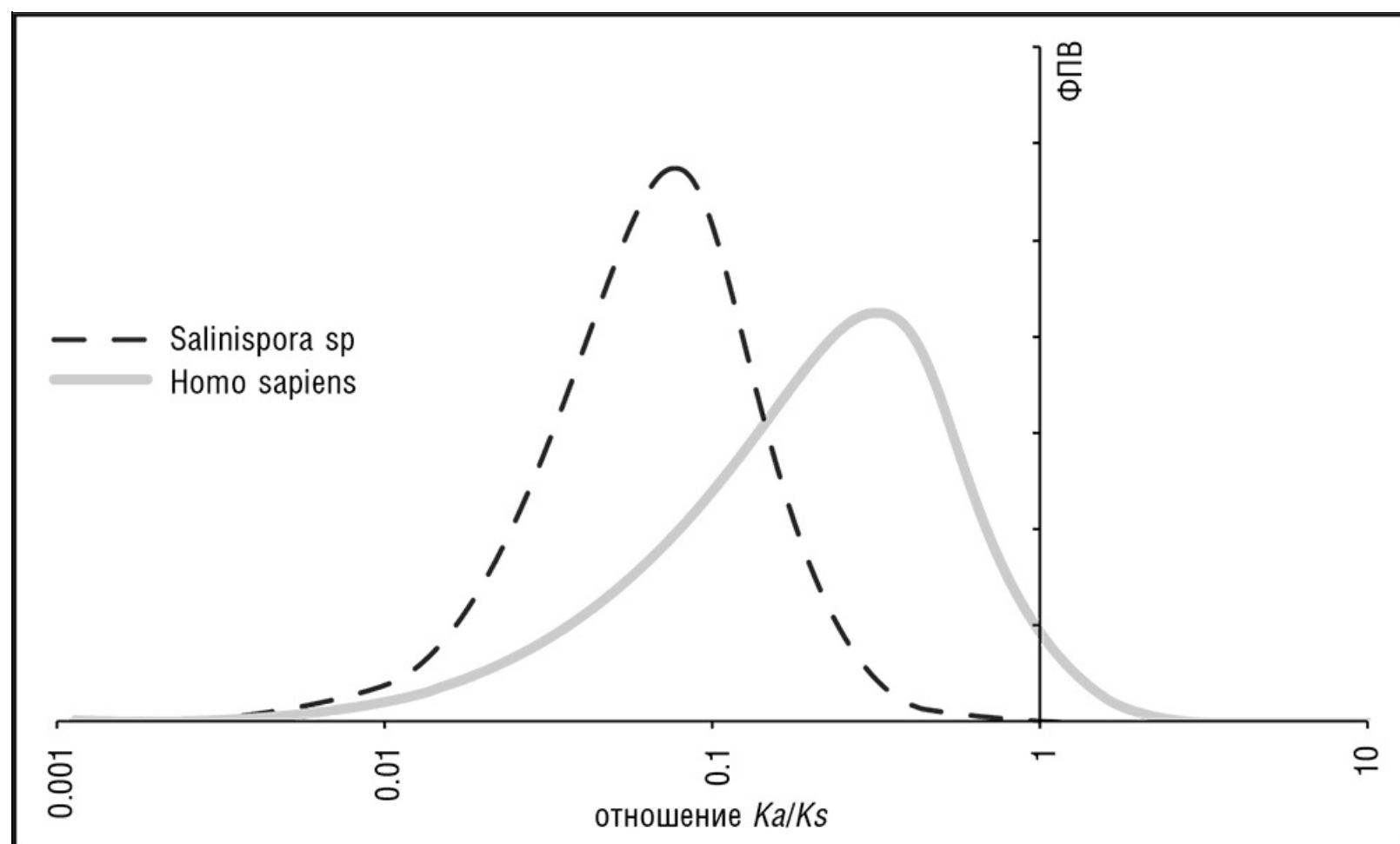


Рис. 3–8. Распределение отношения Ka/Ks в геномах прокариот и эукариот. *Salinispora sp.*: вычислено по ортологам в *S. arenicola* CNS-205 и *S. tropica* CNB-440 (актинобактерии). *Homo sapiens*: вычислено по ортологам в *H. sapiens* *Macaca mulatta* (приматы). Значения Ka и Ks оценены с использованием программного обеспечения PAML (Yang, 2007). График в логарифмических координатах по оси абсцисс; ФПВ обозначает функцию плотности вероятности.

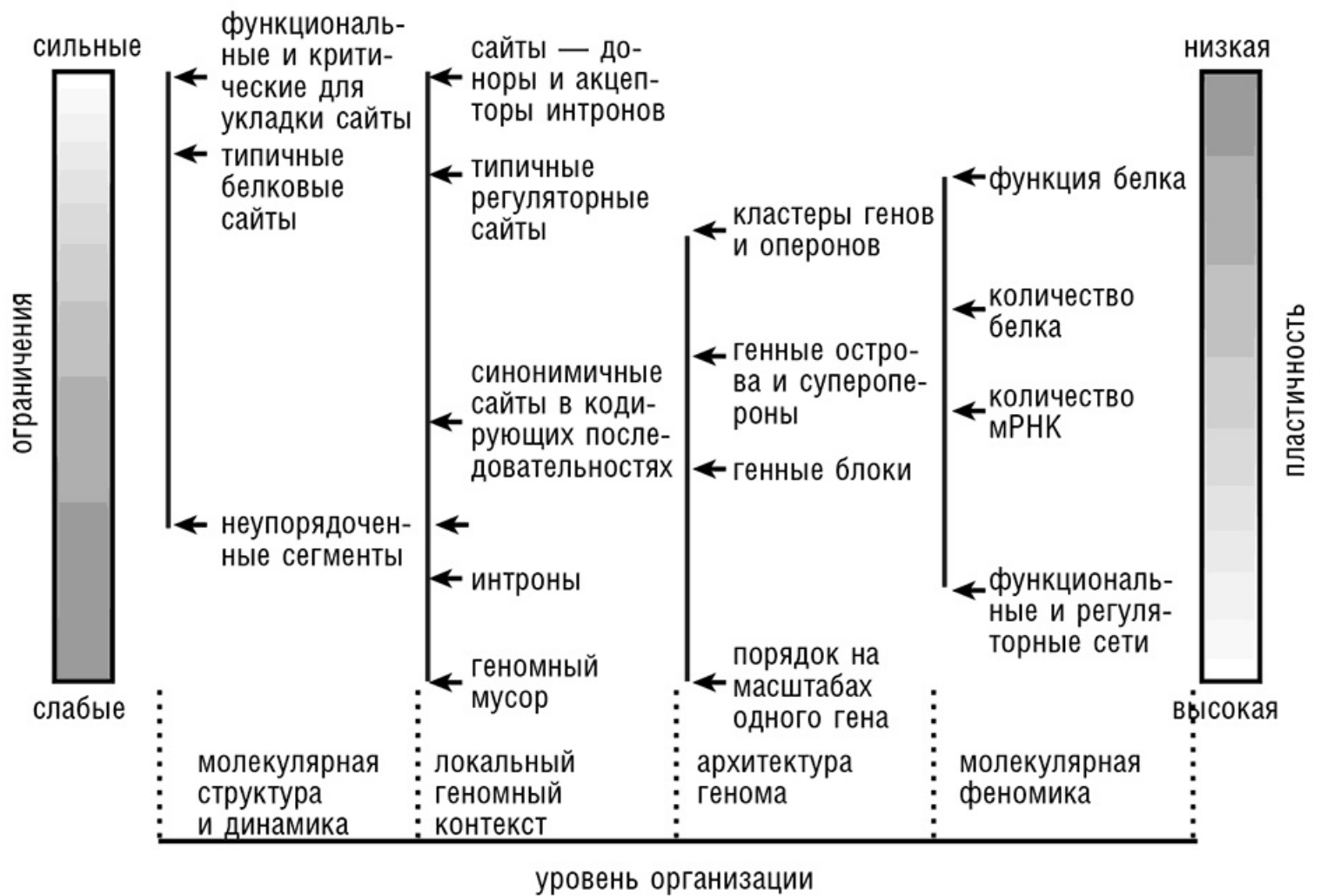


Рис. 3–9. Схематическая сводка эволюционных ограничений, действующих на различные классы геномных сайтов.

Одноклеточные эукариоты, напоминающие прокариот общей архитектурой генома, демонстрируют примерно одинаковые распределения эволюционных ограничений, хотя доля очевидно не подверженных ограничениям некодирующих последовательностей в их геномах несколько выше. Геномы многоклеточных эукариот (растений и особенно животных) являют собою разительную противоположность. Эти организмы имеют богатые интронами геномы с длинными межгенными промежутками; существенная, хотя и переменная часть этих некодирующих последовательностей, по-видимому, эволюционирует, не подвергаясь ограничениям. Доля нуклеотидов в геноме, подверженных эволюционным ограничениям, оценивается методами, основанными на критерии Макдональда – Крайтмана (см. табл. 2–2). Полученные оценки существенно отличаются даже между животными: у *Drosophila* около 70 процентов нуклеотидных сайтов в геноме, в том числе 65 процентов некодирующих участков, по всей видимости, подвержены отбору (в том числе положительному), а у млекопитающих эта доля оказывается в интервале 3–6 процентов (Koonin and Wolf, 2010b). Примечательно, однако, что абсолютное число подверженных отбору сайтов в столь разных по размеру геномах этих животных довольно близко. Напротив, в *Arabidopsis*, растении с геномом, размером и общей архитектурой сравнимыми с таковыми *Drosophila*, доля некодирующих подверженных ограничениям участков, по-видимому, существенно ниже.

Резюмируя существующее понимание ограничений, влияющих на различные классы и сайты во всем известном разнообразии геномов (см. рис. 3–9), отметим, что некоторые

фундаментальные простые выводы являются бесспорными. В частности, нет никаких сомнений, что несинонимичные сайты в белок-кодирующих последовательностях и последовательности, кодирующие структурные РНК, являются одними из наиболее сильно ограниченных во всех геномах и что характерное распределение ограничений (геномный ландшафт) сильно коррелирует с архитектурой генома (Koonin and Wolf, 2010b). Однако помимо этих основных принципов, и довольно неожиданно, оказывается, что эволюционные режимы сильно различаются даже для некоторых относительно близких таксонов, таких как членистоногие и позвоночные. Чтобы выработать всеобъемлющую картину эволюционных ограничений и давления, формирующих геном, требуется еще множество дополнительных исследований по различным организмам. В последующих главах рассматриваются различные проявления давления отбора, влияющие на разные части генома.

Интеграция результатов сравнительной геномики позволяет нам начать строить карту всей «вселенной генов». Глобальная эволюционная устойчивость генов, проявляющаяся прежде всего в сохранении белковых и РНК-последовательностей, стала очевидной в результате самых первых сравнений секвенированных прокариотических и эукариотических геномов: бактерии *Haemophilus influenzae* и *Mycoplasma genitalium*, археи *Methanocaldococcus jannaschii* и эукариотических дрожжей *Saccharomyces cerevisiae* (Tatusov et al., 1997). Ключевое обобщение сравнительной геномики состоит в том, что гены не просто сохраняются на различных интервалах эволюционного процесса, но и представляют собой дискретные единицы эволюции, а именно ортологичные эволюционные линии (см. табл. 3–1). В сегодняшней коллекции секвенированных геномов найдены ортологи в далеких таксонах для значительного большинства белок-кодирующих генов в каждом геноме. Ярким примером являются недавние результаты секвенирования генома примитивных животных: многочисленные гены *Trichoplax* губки связаны ортологичными отношениями с генами млекопитающих и птиц (Putnam et al., 2007; Srivastava et al., 2008; Srivastava et al., 2010). Один из выводов состоит в том, что характерная продолжительность жизни животного гена в этих линиях охватывает по меньшей мере сотни миллионов лет. Многие другие группы животных, такие как насекомые, утратили многочисленные гены (Koonin et al., 2004), так что судьба одного и того же гена в большинстве случаев отличается в разных линиях, в результате чего мы получаем «пятнистую» филетическую модель. (Как подчеркивается далее в этой главе, множество поистине универсальных генов чрезвычайно мало.) Судьбы конкретных генов в разных линиях зависят как от случайных факторов, так и от различий в давлении отбора (см. гл. 9). Результаты обширного сравнительного анализа геномов растений, грибов и прокариот полностью совместимы с этим выводом. Когда гены в геноме классифицируются по их относительному «возрасту» (то есть филогенетической глубине, на которой обнаруживаются гомологи), наблюдаемое расхождение подобно для удаленных друг от друга организмов, как показано на рис. 3–3 для генетических наборов человека и грибка *Aspergillus fumigatus* (Wolf et al., 2009), двух организмов, разделенных, по-видимому, миллиардами лет эволюции. Тем не менее распределения генных возрастов поразительно похожи: в каждом случае древних генов, для которых легко обнаруживаются гомологи в далеких таксонах, значительно больше, чем «молодых» генов. *Несмотря на частую потерю в отдельных эволюционных линиях, гены характеризуются чрезвычайной долговечностью, и многие из них, возможно, бессмертны* [34].

Как обсуждается далее в этой книге (гл. 5 и 7), пути передачи генетической информации у прокариот принципиально отличаются от таковых у эукариот. Тем не менее доли консервативных генов у них примерно равны. В настоящее время эта доля хорошо известна и очень близка у разнообразных бактерий и архей, почти как фундаментальная постоянная: для 70–80 процентов генов ортологи обнаруживаются в далеких организмах (Koonin and Wolf, 2008b; см. рис. 3–4).

Минимальные наборы генов, замещение неортологичных генов (ЗНОГ) и ускользящее незаменимое ядро жизни

Секвенирование геномов симбиотических и паразитических бактерий привело к соблазнительной идее, что их генетический репертуар может быть близок к «наименьшему возможному набору генов», то есть такому, который является необходимым и достаточным

для поддержания простой (прокариотической) клетки при самых благоприятных условиях, какие только могут существовать вне других клеток (Fraser et al., 1995; Mushegian and Koonin, 1996b). Последнее условие чрезвычайно важно, поскольку «наименьшим возможным» набор генов будет лишь в отношении к окружающей среде, в которой соответствующий организм существует (или мог бы существовать, в случае «концептуальных» геномов, полученных компьютерными методами). Однако, как только появились первые два полных бактериальных генома, вторым из которых был геном *Mycoplasma genitalium* [35], лишенной клеточной стенки паразитической бактерии с размером генома всего около 570 генов, возникла очевидная идея, что «истинный» наименьший набор можно естественным образом вывести, сравнивая геномы этих двух существенно различно специализированных бактериальных патогенов (Mushegian and Koonin, 1996b). Точнее, можно было бы ожидать, что ортологичные гены в двух организмах будут представлять собой набор основных биологических функций, которые необходимы для выживания клетки, независимо от уникального образа жизни каждого организма.

Сравнение геномных наборов *H. influenzae* и *M. genitalium* дало 240 пар ортологичных генов, охватывающих большую часть очевидно существенных клеточных функций. Тем не менее в этом консервативном наборе несколько важных функций явно отсутствовали. До сих пор мы не говорили о «настоящей биологии», о биологических функциях, ролях генов, но теперь мы должны начать думать биологически. Определение минимального набора основных биологических функций – задача непростая. Соблазнительно, конечно, попытаться «разобрать эволюцию по винтику»: идя от сравнительной геномики, определить минимальный набор основных генов, сохраняемых во всех клеточных формах жизни. Но этот подход упускает возможность, что разные организмы могли прийти к решению одной и той же принципиальной задачи независимыми путями. Мы увидим далее в этой главе, что такая гипотетическая возможность действительно отражает важный аспект биологической реальности. Таким образом, чтобы очертить минимальный набор клеточных функций, нам необходимо обратиться к логике биохимии и клеточной биологии. Знаний в этих областях несомненно достаточно, чтобы составить разумный каталог основных функций. Само собой, это знание несовершенно, поэтому на самом деле вычисление минимального набора генов требует многократного поочередного обращения к биологическому обоснованию и сравнительному геномному анализу. Мы с Аркадием Мушегианом предположили, что принципиально важные функции, отсутствующие среди 240 ортологов *H. influenzae* и *M. genitalium*, вероятно, исполняются неродственными или отдаленно родственными белками в этих двух бактериях. Мы привлекли определенные догадки, чтобы увеличить предполагаемый минимальный набор на 16 дополнительных генов *M. genitalium* (см. рис. 3-10). Это простое упражнение в получении минимального набора генов соединением сравнительной геномики и биологической логики оказалось достаточно успешным и, по-видимому, определило приближенный функциональный репертуар простейшей бактериальной клетки, способной к самостоятельному росту в наиболее благоприятных условиях. В самом деле, последующие эксперименты с нокаутом генов подтвердили, что большинство из генов, включенных в минимальный набор, необходимы для выживания бактерий и что гены из минимального набора присутствуют в большинстве (хотя и не обязательно во всех) вновь секвенированных бактериальных геномах (Delaye and Moya, 2010; Koonin, 2003).

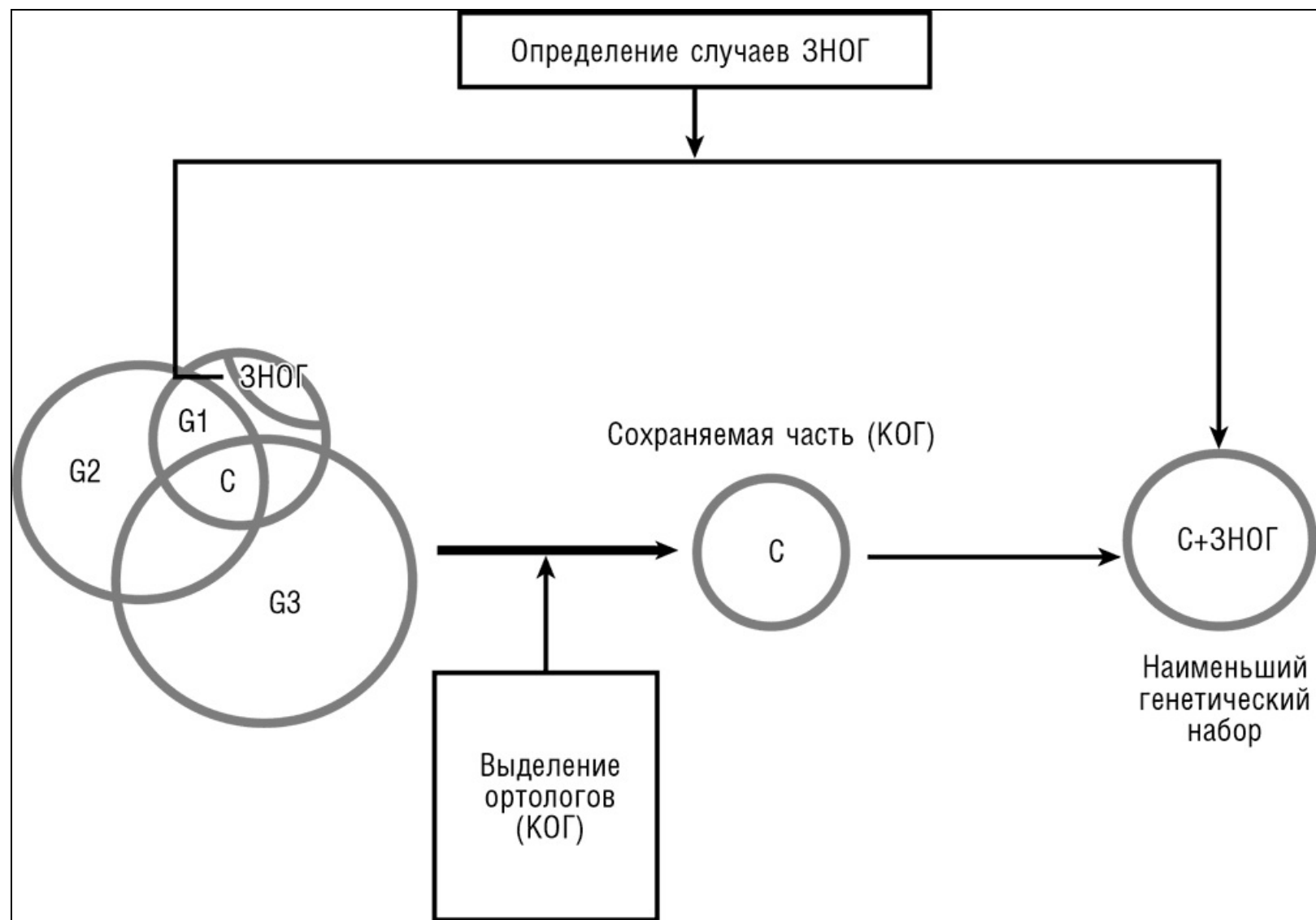


Рис. 3-10. Выделение минимального набора генов клеточной жизни методами сравнительной геномики. G1, G2, G3 – три сравниваемых генома; C – набор консервативных генов.

Наименьший набор генов



Необходимый набор *B. subtilis*

Все КОГ

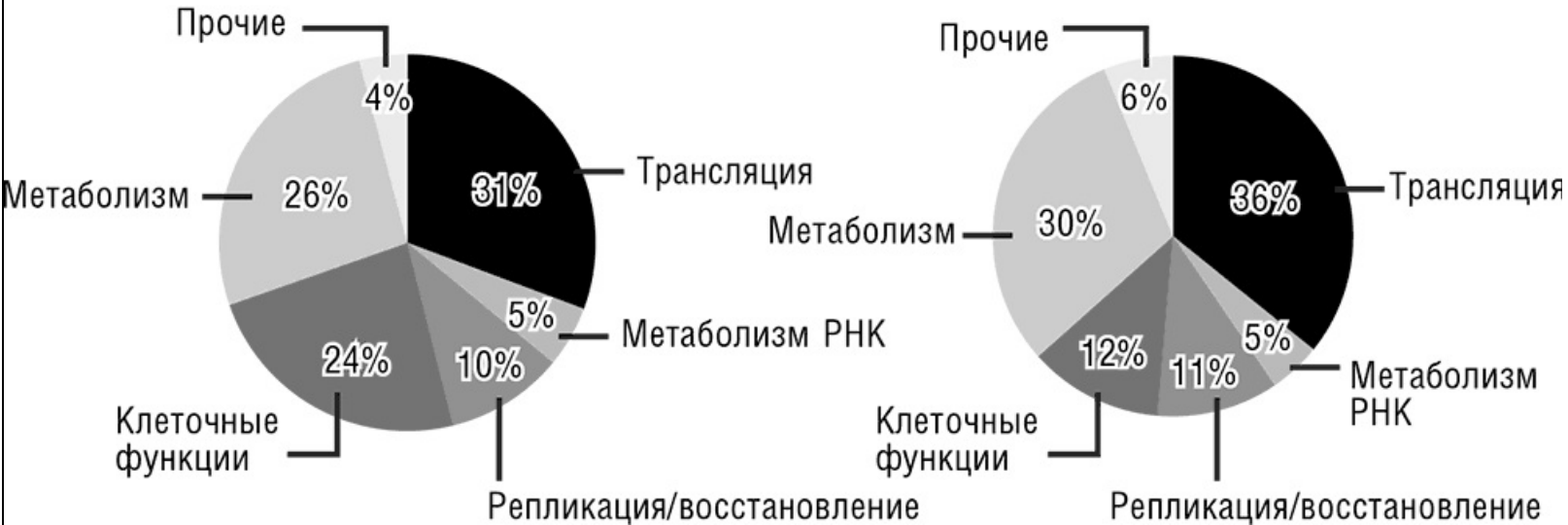


Рис. 3-11. Распределение числа генов по биологическим функциям в минимальном геномном наборе, полном наборе КОГ и среди экспериментально определенных незаменимых генов бактерии *Bacillus subtilis*. Данные по Koonin, 2003.

Поучительно провести теперь функциональную перепись минимального бактериального набора генов. В этом наборе преобладают гены, которые кодируют белки, участвующие в передаче информации в клетке (репликации, транскрипции и, прежде всего, трансляции). Метаболические ферменты и белки транспортной системы представлены куда более разреженно, что вполне ожидаемо для организма, растущего в самой богатой из возможных сред. Этим минимальный набор генов резко отличается от полного набора КОГ, но напоминает набор незаменимых бактериальных генов (нокаут которых убивает бактерию, см. рис. 3-11). Эта особая эволюционная устойчивость систем передачи клеточной информации является одним из центральных обобщений сравнительной геномики. Мы вернемся к этому вопросу позднее.

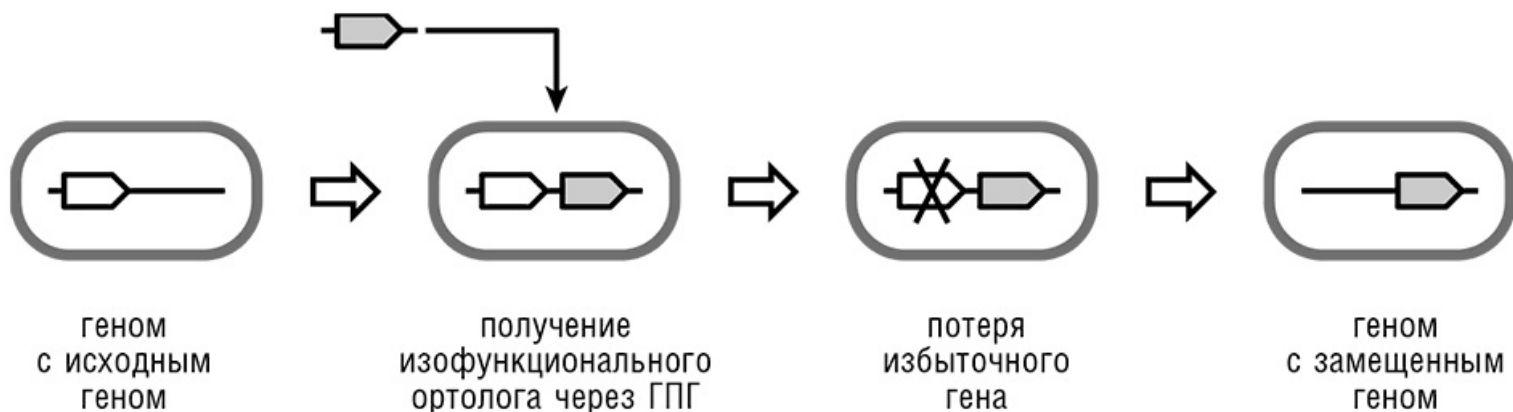


Рис. 3-12. Эволюционный сценарий замещения неортологичных генов.

По-видимому, самым весомым результатом поиска минимального набора генов было открытие, что в списке легко обнаружимых ортологов недостает нескольких важных функций. Этот результат был подтвержден и значительно усилен при сравнении бактериальных геномов с первым архейным (*Methanocaldococcus jannaschii*), когда в наборе консервативных необходимых функций выявилось несколько дополнительных зияющих пробелов. Эти выводы были обобщены в понятии *замещения неортологичных генов* (ЗНОГ), эволюционном сценарии, при котором неродственные или отдаленно родственные гены (иными словами, не ортологи) становятся ответственными за одни и те же необходимые функции в разных организмах (Koonin, 2003). Сам этот эволюционный сценарий легко представить (см. рис 3-12): чтобы произошло ЗНОГ, эволюционирующая линия приобретает альтернативный, функционально избыточный ген для некоторой незаменимой функции и таким образом проходит через промежуточное состояние, в котором присутствуют обе реализации данной функции (такая избыточность часто наблюдается у организмов с более сложными геномами), а затем теряет начальный ген (Koonin and Mushegian, 1996). С ростом коллекции секвенированных геномов обнаруживается все больше организмов, в которых оба варианта действительно представлены для разнообразных функций; таким образом, сценарий эволюции ЗНОГ, представленный на рис. 3-12, становится все более правдоподобным.

В табл. 3–2 описывается несколько примеров ключевых биологических функций, для которых два или несколько неродственных ферментов по-разному представлены в частично дополнительных, но обычно перекрывающихся группах эволюционных линий. Даже эти отдельные примеры показывают, что ЗНОГ происходит в самых различных функциональных системах и путях. В дальнейшем, с заметным увеличением числа секвенированных геномов, стало ясно, что ЗНОГ и утрата генов в отдельных линиях настолько широко распространены, что лишь малое число функций являются действительно мономорфными и вездесущими (то есть представлены ортологичными генами во всех организмах). Вместе с тем универсальное ядро жизни уменьшилось почти до исчезновения: все, что остается универсальным, – это около тридцати генов белков трансляции и три больших субъединицы РНК-полимеразы, а также примерно равное число генов структурных РНК (рРНК и тРНК).

Таблица 3–2. Примеры замещения неортологичных генов.

Биологическая функция	Гены/белки	Филетическое разнообразие формы 1	Филетическое разнообразие формы 2	Филетическое разнообразие формы 3
Репликация ДНК	ДНК-полимераза, примаза, геликаза	Археи, эукариоты, двухцепочечные ДНК-вирусы прокариот и эукариот	Бактерии	—
Биосинтез предшественников ДНК	Тимидилатсинтаза	Некоторые археи и бактерии, большинство эукариот	Некоторые археи и бактерии, меньшинство эукариот	—
Трансляция	Лизил-тРНК-синтаза	Большинство бактерий, меньшинство архей, подавляющее большинство эукариот	Большинство архей, меньшинство бактерий	—
Центральный метаболизм углеводов	Глюкозо-6-фосфат-изомераза	Большинство бактерий, некоторые археи, эукариоты	Меньшинство бактерий, некоторые археи	—
Центральный метаболизм углеводов	Фруктозо-бисфосфатаза	Примерно половина бактерий, некоторые археи, большинство эукариот	Примерно половина бактерий, малое меньшинство архей и эукариот	Большинство архей, малое меньшинство бактерий
Метаболизм жиров	Диацилглицеролкиназа	Большинство бактерий	Эукариоты, малое меньшинство бактерий	—
Кислотно-щелочной гомеостаз, реакция на окислительный стресс	Карбонатдегидратаза	Большинство бактерий, некоторые археи, некоторые эукариоты	Меньшинство бактерий, большинство архей, меньшинство эукариот	Меньшинство бактерий, большинство эукариот

Данные в основном по Omelchenko et al., 2010.

Даже при исключении паразитических бактерий перечень универсальных генов расширяется незначительно (Koonin, 2003). Таким образом, за исключением небольшого числа генов, участвующих в основных этапах передачи информации, *не существует универсального генетического ядра жизни, в связи с повсеместными ЗНОГ и потерей генов.* Концепция *небольшого, универсального набора функций*, необходимых для поддержания клетки, остается жизнеспособной, но, учитывая комбинаторику ЗНОГ, этот наименьший набор функциональных ниш может заполняться огромным разнообразием генных ансамблей.

Единицы эволюции и фрактальная структура генетической вселенной

Результаты сравнительной геномики приводят к ключевому обобщению, которое позволяет нам выполнять продуктивные эволюционные исследования: основные единицы эволюции могут быть довольно четко определены, и единицы эти – кластеры ортологичных генов, или эволюционные домены (КОГ), или, еще точнее, линии эволюционирующих ортологичных генов (доменов). Истории отдельных генов часто сложны (а во многих случаях даже чрезвычайно сложны) и включают в себя множественные утраты генов, дубликацию и горизонтальный перенос (ниже в настоящей книге мы обсудим эти явления подробнее, см. гл. 5 и 7). Предрасположенность генов к дубликации, утрате и переносу варьирует в широких пределах. Однако, невзирая на все эти осложнения, атомарное свойство наборов ортологичных генов твердо соблюдается: *КОГ суть естественные элементы генетической вселенной.*

Генетическая (геномная) вселенная (это только метафора, но удобная и, возможно, продуктивная) может быть представлена как развивающееся пространство-время, заполненное кластерами, состоящими из генов (то есть КОГ), или, точнее, эволюционирующими линиями ортологов, элементарными единицами эволюции. Ортология легче всего прослеживается между прокариотическими генами, так что здесь мы обсудим прокариотическую область геномной

вселенной. Тенденции среди эукариотов в принципе похожи, но осложнены распространенной мультидоменной организацией белков и обширной паралогией. В нашем геномном пространстве заметно характерное распределение КОГ по геномам, хорошо аппроксимируемое тремя экспонентами с разными показателями, которые делят генную популяцию на три класса (см. рис. 3-13, *a – в*; Koonin and Wolf, 2008b).

1. (Почти) универсальные гены, те, что представлены в (почти) всех геномах клеточных форм жизни, составляют лишь малую часть генетической вселенной: это ядро *клеточной жизни* состоит, самое большее, из приблизительно 70 генов. В каждом конкретном геноме доля этих «ядерных» генов составляет не более 10 процентов, если говорить о самых маленьких геномах клеточных форм жизни (паразитических бактерий, таких как *M. genitalium*), но обычно ближе к 1 проценту или менее от общего числа генов (см. рис. 3-14).

2. Умеренно консервативная генная *оболочка* состоит из КОГ, представленных в самых разнообразных геномах, но не в подавляющем их большинстве. Недавний анализ имеющихся прокариотических геномов дает число КОГ оболочки около 5000. Гены оболочки составляют большую часть числа генов в любом геноме (см. рис. 3-14).

3. Малоконсервативное «облако» состоит из КОГ, встречающихся в узких группах организмов, и «генов-сирот» – генов в открытых рамках считывания (ОРС), обнаруженных пока что в одном-единственном геноме, но гомологи которых обычно обнаруживаются во вновь появляющихся геномных данных. Гены «облака» составляют переменную долю в каждом геноме, обычно в интервале 10–30 процентов от общего числа генов (см. рис. 3-14).

Примечательно, что эта структура является самоподобной, или *фрактальной*: те же три компонента – крошечное ядро, сравнительно большая оболочка и огромное «облако» – проявляются на любом уровне, где бы ни рассекалось геновое пространство-время, от всего мира прокариот и до узких групп бактерий (см. рис. 3-14). Мы возвратимся к последствиям этой фрактальности геномного пространства-времени прокариот в главе 5. Заметим, однако, что эволюционная модель, которая объясняла бы наблюдаемую фрактальность, еще ожидает своей разработки [\[36\]](#).

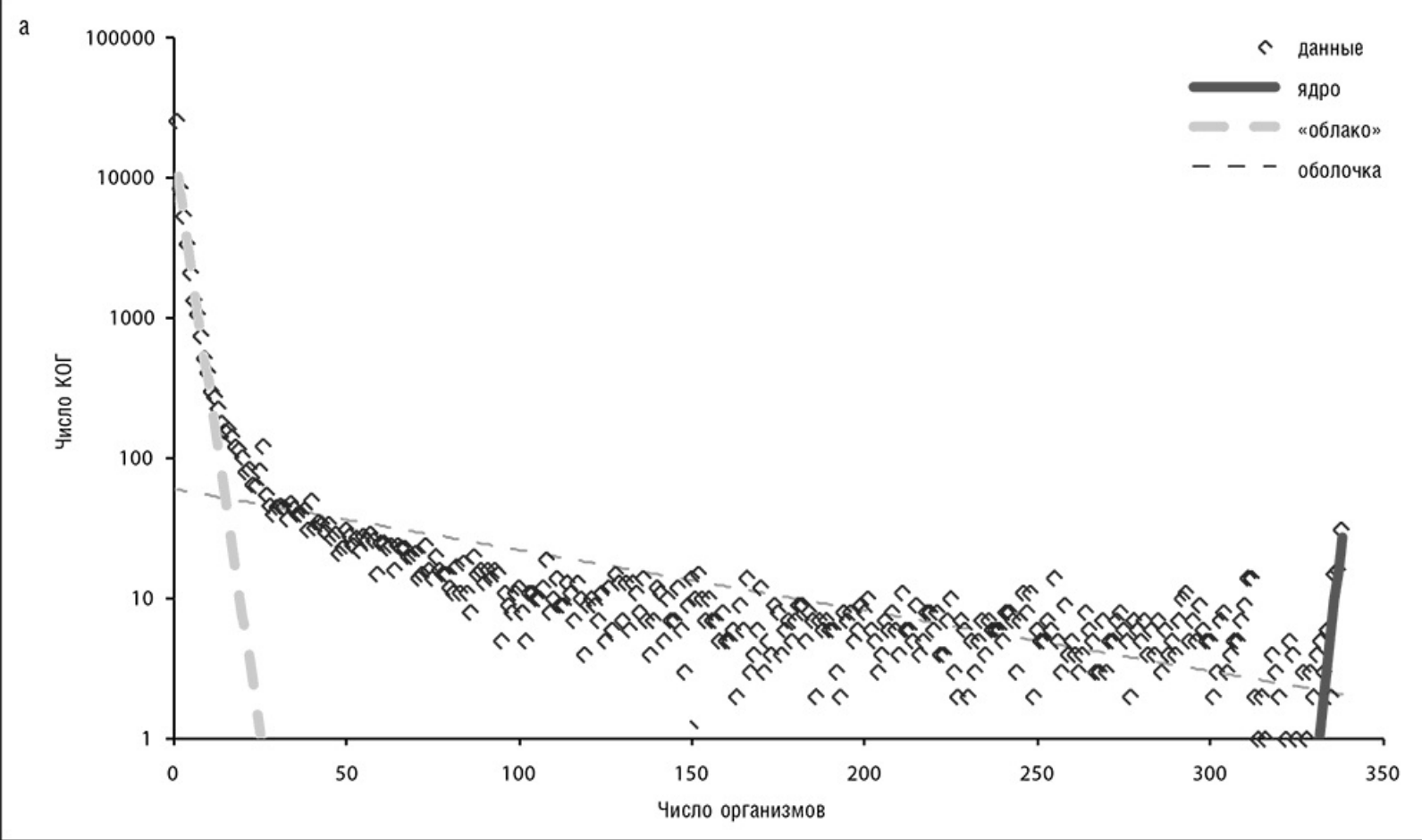


Рис. 3-13 а. Глубокий уровень: 338 прокариот из базы данных EggNOG (Jensen et al., 2008)

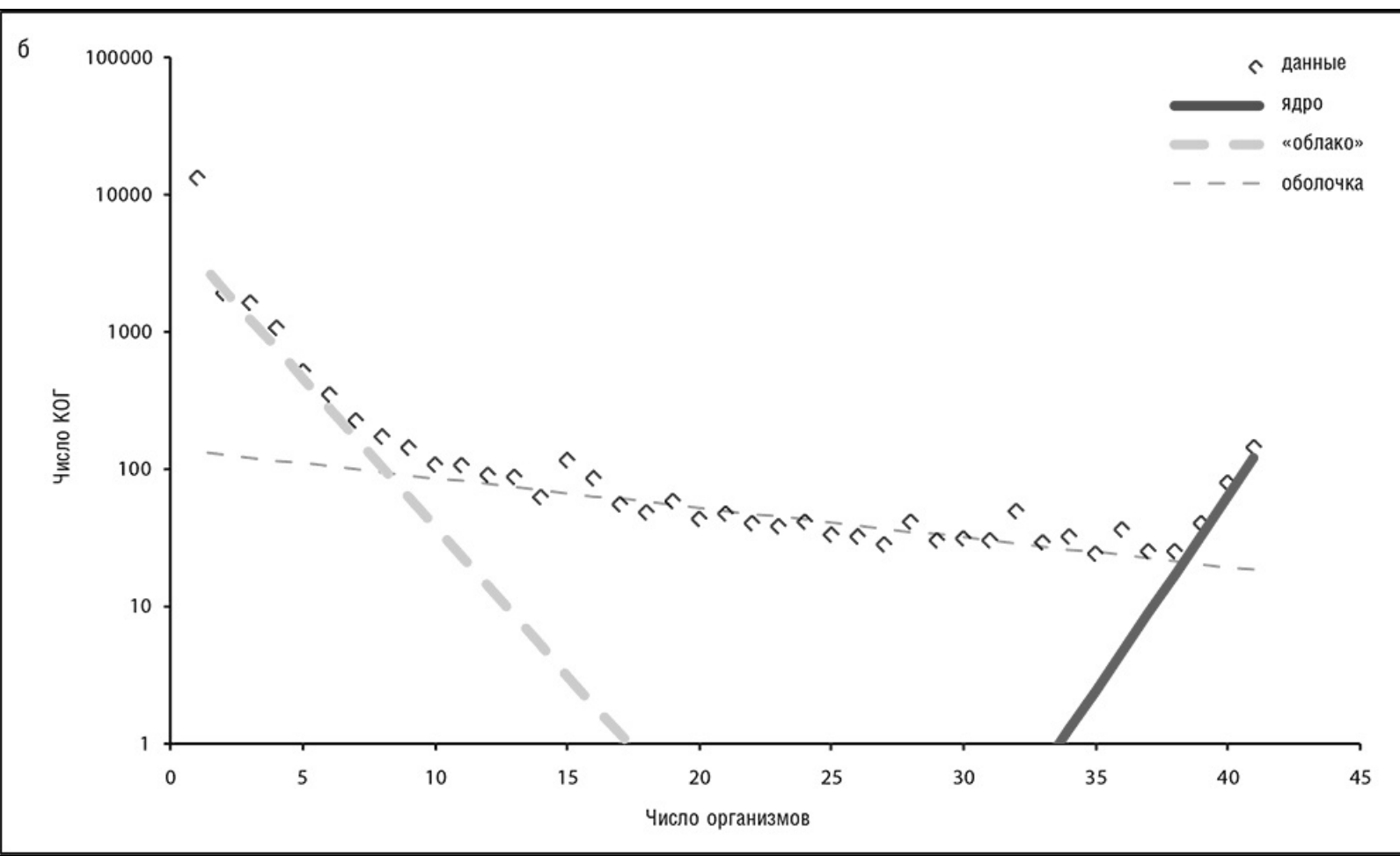


Рис. 3-13 б. Средний уровень: 41 архей из базы данных arCOG (Makarova et al., 2007b)

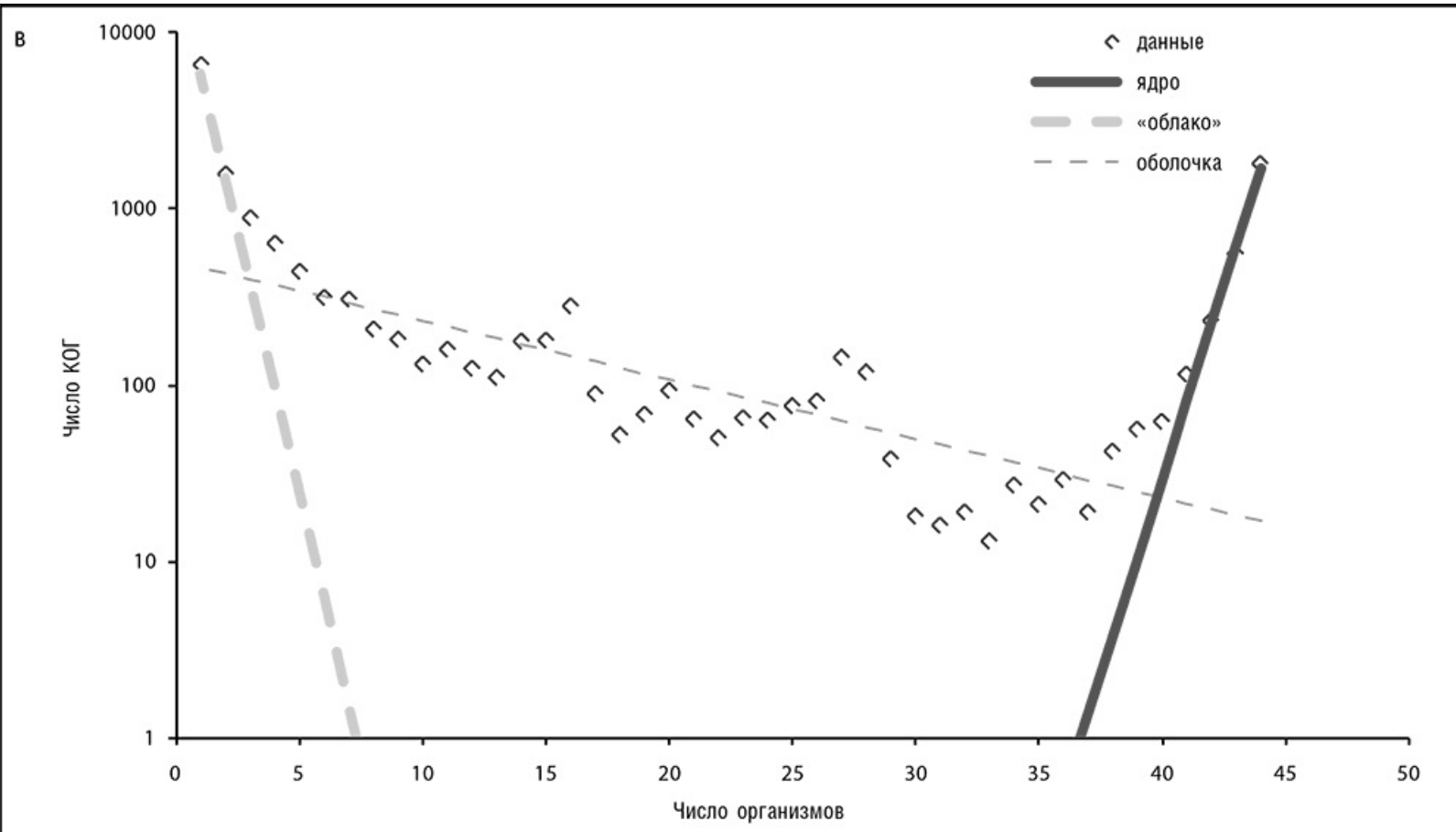


Рис. 3-13 в. Мелкий уровень: 44 вида *Escherichia*, *Shigella* и *Salmonella* из базы данных COG (Tatusov et al., 2003). Регрессия данных экспоненциальными функциями на всех трех рисунках (Koonin and Wolf, 2008b) изображена пунктирными и непрерывными линиями.

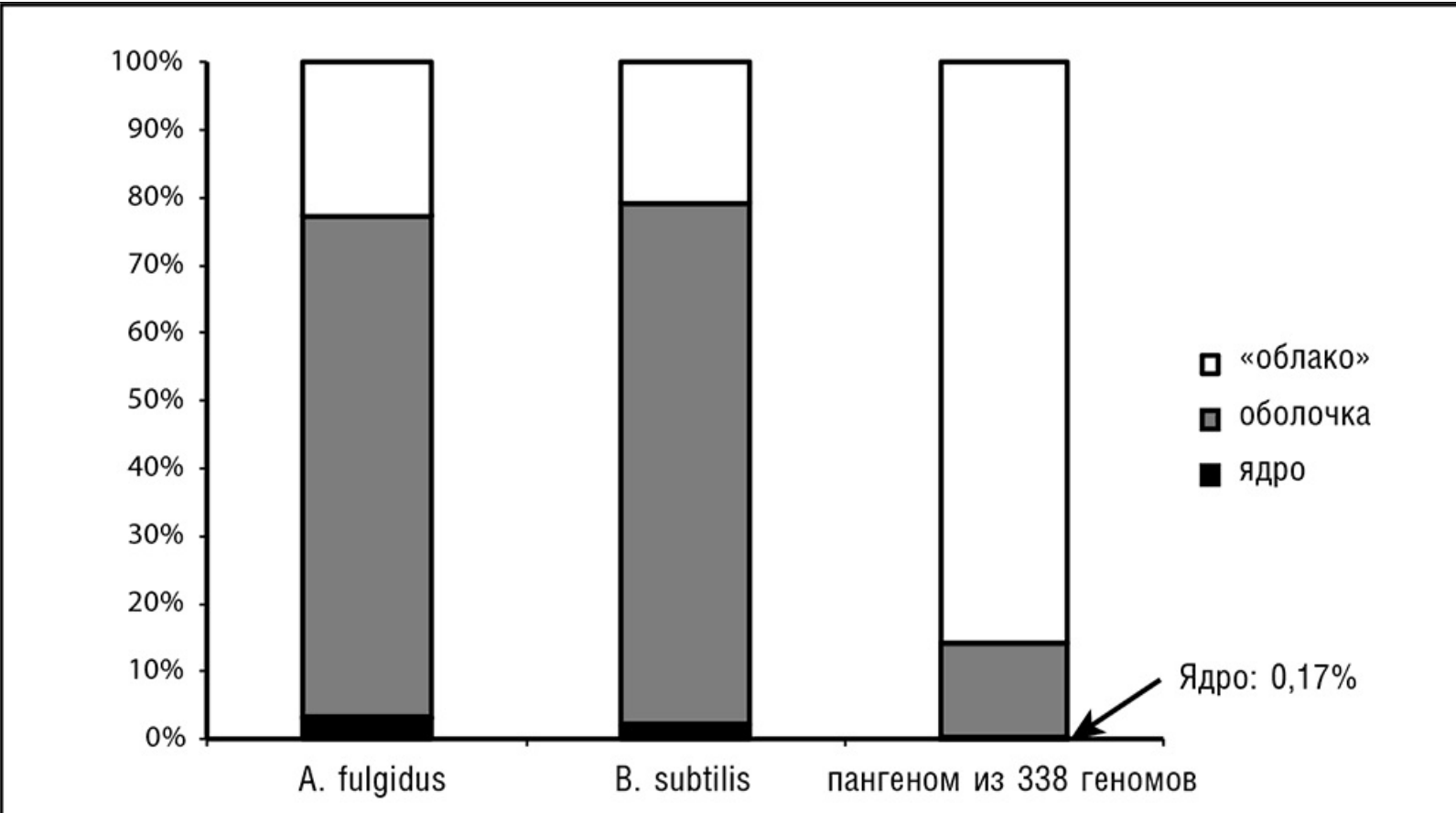


Рис. 3-14. Вклад ядра, оболочки и «облака» в состав индивидуальных геномов и геномной

вселенной как целого. Расчет произведен по данным из базы EggNOG (Jensen et al., 2008). *A. fulgidus*– архея *Archaeoglobus fulgidus*; *B. subtilis*– бактерия *Bacillus subtilis*.

Следует обратить внимание на кажущийся парадокс в распределении КОГ в геномном пространстве. Хотя в каждом геноме большинство генов относятся к оболочке, то есть являются общими с дальнородственными организмами, при рассмотрении всей геномной вселенной оказывается, что гены (или, вернее, КОГ) ядра и оболочки составляют лишь незначительное меньшинство (см. рис. 3-14). Вполне очевидно, что эта разница возникает потому, что КОГ оболочки представлены во многих геномах, в то время как КОГ «облака», особенно «гены-сироты», являются редкими или уникальными. С учетом этой характерной структуры вселенной генов эволюционные реконструкции неизбежно приводят к картине динамичной эволюции генома, где многочисленные гены (в основном из «облака» и, в меньшей степени, из оболочки) утрачиваются, а многие другие приобретаются путем ГПГ (в основном у прокариот), а также в результате многочисленных дупликаций, в первую очередь у эукариот (см. ниже в этой главе).

Элементарные события геномной эволюции

Теперь, определив единицы геномной эволюции и разработав идею организации вселенной генов, мы можем осмысленно дополнить эти понятия списком основных операций, элементарных событий эволюции генома, которые можно будет сравнить с элементарными событиями эволюции отдельных генов. Алфавиты элементарных событий довольно кратки и фактически подобны (изоморфны) на соответствующих уровнях (см. табл. 3–3). Однако относительный вклад и частота различных типов событий разнятся в эволюции генов и геномов самым коренным образом. Существенное различие между эволюцией отдельных генов и целых геномов заключается в особой важности и высокой частоте дупликации генов, в отличие от много более ограниченного вклада внутригеномных дупликаций. Далее, внутригеномные рекомбинации редко закрепляются в эволюции, за исключением близких геномов, а важнейшие механизмы перестройки генома, такие как инверсии и транслокации, не играют особой роли в эволюции отдельных генов. В итоге различия в относительном вкладе разнообразных элементарных механизмов (см. табл. 3–2) лежат в основе *значительно более динамичного характера эволюции геномов по сравнению с эволюцией отдельных генов.*

Таблица 3–3. Сравнение элементарных событий эволюции гена и генома.

Тип эволюционного события	Ген	Геном
Замена	Замена нуклеотида, один из ключевых процессов	Замена гена неортологом или ксенологом; важна, но относительно редко встречается
Делеция/потеря	Малые делеции почти столь же часты, как и замены; большие делеции встречаются реже, обратно пропорционально величине	Потеря гена путем делеции или инактивации широко распространена во многих эволюционных линиях

Тип эволюционного события	Ген	Геном
Вставка	Небольшие вставки обычны, хотя и менее часты, чем делеции	Приобретение генов посредством ГПГ — один из важнейших путей эволюции генома; вставки другого происхождения много более редки
Рекомбинация/ГПГ	Внутригенная рекомбинация сравнительно редка, за исключением гомологической рекомбинации в близкородственных генах	Важнейший путь эволюции генома, доминирующий у прокариот
Дупликация	Дупликация небольших участков обычна; большие дупликации встречаются реже, обратно пропорционально величине	Важнейший путь эволюции генома, доминирующий у эукариот

Сравнительная геномика раскрывает примечательный контраст между относительной эволюционной устойчивостью отдельных генов, многие из которых сохраняют значительное сходство на протяжении сотен миллионов или даже миллиардов лет эволюции, и пластичностью состава и архитектуры генома, которые изменяются на несколько порядков быстрее. Отсюда возникает характерное устройство вселенной генов, в котором сравнительно небольшое число плотных кластеров образуют ядро, гены которого представлены в большинстве геномов, в то время как большую часть пространства-времени занимает огромное количество все более разреженных «туманностей», состоящих из редких генов. Поразительно, что организация генетической вселенной явно фрактальна, то есть проявляется на всех масштабах эволюционных расстояний.

Атомарная сущность генов (или, точнее, КОГ, ортологичных эволюционных линий) лежит в основе всей исследовательской программы сравнительной геномики: сравнение геномов оказывается весьма информативным, несмотря на нетривиальные отношения между отдельными генами и геномами, обусловленные изменчивостью геномной архитектуры.

Геномные ландшафты различных форм жизни – распределение ограничений по геномным сайтам – разнообразны и сложны. Компактные геномы вирусов, прокариот и, в меньшей степени, одноклеточных эукариот в основном занимают «высокогорные плато», так что почти все сайты подвергаются существенным ограничениям. Геномные ландшафты многоклеточных эукариот состоят в основном из «долин» со слабыми ограничениями, разделенных редкими «гребнями» сильного отбора. Эти отличия отражают разные эволюционные режимы, которые мы обсудим в главе 8. Парадоксально, но именно «неэффективность» режима эволюции, характерного для многоклеточных эукариот, позволяет организационной сложности возникнуть. Этот парадокс должен заставить задуматься всех неравнодушных к идее эволюционного «прогресса». Мы вернемся к подробному обсуждению этого вопроса в главах 8 и 13.

Рекомендуемая дополнительная литература

Ellegren, H.(2008) Comparative Genomics and the Study of Evolution by Natural Selection. *Molecular Ecology*17: 4,586—4,596.

Обзор факторов отбора, важных для эволюции разных классов геномных последовательностей.

Koonin, E. V.(2005) Orthologs, Paralogs, and Evolutionary Genomics. *Annual Review of Genetics*39: 309–338.

Детальное обсуждение концепций ортологии и паралогии, а также определенных категорий эволюционных отношений между генами в этих широких классах.

Koonin, E. V.(2009) Evolution of Genome Architecture. *International Journal of Biochemistry and Cell Biology*41: 298–306.

Обзор разнообразия и эволюционных тенденций геномных архитектур различных форм клеточной жизни.

Koonin, E. V.(2003) Comparative Genomics, Minimal Gene-Sets, and the Last Universal Common Ancestor. *Nature Reviews Microbiology*1: 127–136.

Критическое обсуждение концепции минимального набора генов и ее применения к организмам, имеющим различный стиль жизни, а также сравнение минимальных наборов и реконструкция предковых геномов.

Koonin, E. V., and Y. I. Wolf.(2010) Constraints and Plasticity in Genome and Molecular-Phenome Evolution. *Nature Reviews Genetics*11: 487–498.

Попытка исчерпывающей полногеномной переписи эволюционных ограничений, действующих на различные классы последовательностей и сайтов в геномах.

Koonin, E. V., and Y. I. Wolf. (2008) Genomics of Bacteria and Archaea: The Emerging Dynamic View of the Prokaryotic World. *Nucleic Acids Research*36: 6,688—6,719.

Детальный обзор геномики прокариот с особым упором на динамику генома, в том числе ГПГ.

*Levitt, M.*Nature of the Protein Universe. (2009) *Proceedings of the National Academy of Sciences USA*106: 11,079—11,084.

Подробное исследование происхождения новизны в эволюции белка. Делается вывод о том, что ключевым механизмом порождения новизны явилось возникновение мультидоменной архитектуры.

Lynch, Michael.(2007) The Origins of Genome Architecture. Sunderland, MA: Sinauer Associates.

Принципиально важная книга по неадаптивной теории эволюции геномной сложности и ее различным последствиям (подробное обсуждение в гл. 8).

Wilkins, A. S.(1997) Canalization: A Molecular Genetic Perspective. *Bioessays*19: 257–262.

Переоценка Уоддингтоновой концепции канализации в контексте современной эволюционной биологии.

Глава 4. Геномика, системная биология и универсалии эволюции: эволюция генома как феномен статистической физики

В предыдущей главе была подчеркнута относительная стабильность отдельных генов, составляющая яркий контраст динамизму геномной эволюции. Если гены или домены принять за атомарные единицы геномной эволюции, тогда геномы можно рассмотреть как статистические ансамбли таких единиц. Мы можем продолжить эту очень упрощенную, но очевидно не бессмысленную и потенциально продуктивную физическую аналогию и рассмотреть геномы как структуры, подобные газу или жидкости, в которых силы межмолекулярного взаимодействия хоть и являются важными параметрами, но слабы по сравнению с внутримолекулярными взаимодействиями (лежащими в основе стабильности молекул), в отличие от твердых тел, в которых межмолекулярные взаимодействия сильны и имеют определяющее значение.

Из статистической физики известно, что поведение ансамбля слабовзаимодействующих частиц (молекул) следует простым и универсальным статистическим закономерностям, таким как распределение Больцмана для скоростей частиц. Аналогия между ансамблями генов (геномами) и ансамблями молекул (газами и жидкостями) наталкивает нас на поиск статистических закономерностей в функционировании и эволюции генома. Более того, размышляя таким образом, мы можем с некоторой степенью уверенности предположить, что эти статистические закономерности должны представлять собой математически простые, универсальные законы распределения значений определенных параметров, описывающих процесс эволюции. Мы убедимся в этой главе, что поиск таких эволюционных универсалий – дело далеко не безнадежное.

Перед обсуждением статистических свойств генных ансамблей необходимо обратить внимание на еще одно ведущее направление биологических исследований первой декады третьего тысячелетия, представляющее собой новую область науки, часто называемую, может быть не очень удачно, системной биологией. Системная биология провозглашает своей конечной целью построение моделей и понимание функционирования биологических систем во всей их сложности. Реальное положение дел на данном этапе становления этой области исследований заключается в том, что основное внимание направлено на агрегацию обширных данных специфического типа, таких как транскриптомы (совокупность всех экспрессируемых РНК клетки, ткани или организма), протеомы (совокупность всех экспрессированных белков), метаболомы (совокупность всех метаболитов) и другие «-омы» (Bruggeman and Westerhoff, 2007; Koopin and Wolf, 2008a). Все эти «-омы» описываются системной биологией с помощью количественных понятий, таких как концентрация белка или метаболита.

Так же как и генетику в ее первые годы, системную биологию многие ученые приняли за скучную «большую науку» и слишком хлопотливое занятие. (Подозреваю, что это отношение до сих пор преобладает.) Так же как и с генетикой, первый взгляд оказался, мягко говоря, недалеким. Наличие высококачественных данных по генной экспрессии, генетическим и белок-белковым взаимодействиям, локализации белка в клетке и других данных системного уровня в масштабе генома открыло новые измерения эволюционного анализа (иначе иногда называемого эволюционной системной биологией) и обеспечило его взаимопроникновение с эволюционной геномикой. В исследованиях системной биологии, в масштабах генома, уже были открыты нетривиальные связи между эволюцией генных последовательностей, генной

экспрессией, структурой белка и другими характеристиками генов и белков. Эти открытия в целом оказались совместимыми с точкой зрения на геном как на статистический ансамбль генов и дали возможность в новом свете рассмотреть селективную и нейтральную составляющие эволюции структуры и функций генома.

Взаимосвязь между эволюционными и фенотипическими параметрами, универсалии эволюции генов, белков и геномов и физическая модель эволюционного процесса

В предыдущей главе было показано, что белок-кодирующие гены (по крайней мере в отношении мутационных замен, приводящих к изменению аминокислот в кодируемом белке) принадлежат к числу наиболее консервативных последовательностей генома. Однако уже на раннем этапе исследований в молекулярной эволюции стало понятно, что скорости эволюционирования белок-кодирующих генов могут очень сильно различаться (Wilson et al., 1977). Этот широкий разброс значений в общем объясняли существованием широкого спектра функций белка, которые по-разному ограничивают скорость эволюции соответствующих генов. В самом деле, кажется само собой разумеющимся, что огромная роль ДНК-полимеразы, сложнейшего фермента, который катализирует встраивание комплементарных матрице нуклеотидов в растущую цепь ДНК, требует значительного ограничения на скорость эволюции для соответствующей ей генной последовательности, в то время как, например, для структурного белка, чья единственная задача состоит в поддержании целостности ядерного матрикса, такого сильного ограничения не требуется. Фундаментальное представление о том, что эволюция белок-кодирующих генов может сводиться не только к уникальным особенностям молекулярной структуры и функции белков, возникло уже на этом раннем этапе. В богатой идеями обзорной статье, опубликованной Аланом Вильсоном и коллегами в 1977 году, выдвигается гипотеза о том, что скорость эволюции генных последовательностей зависит как от уникальных функций кодируемого белка, так и от важности этого белка для выживания организма (Wilson et al., 1977). Однако в то время не было прямых способов изучения эволюционных ограничений, так что эти идеи, хоть и интригующие, тогда находились всецело в области умозрительных построений.

В начале третьего тысячелетия геномика и системная биология полностью преобразили область эволюционных исследований. Доступность множества данных по геномным последовательностям позволила проанализировать и сравнить распределения скоростей эволюции для полных наборов ортологичных генов в различных таксонах, а также изучить взаимосвязи скоростей эволюции ортологов в различных эволюционных линиях. Значения скоростей эволюции по несинонимичным сайтам в ортологичных генах могут различаться на три-четыре порядка, и это распределение значений гораздо шире, чем распределение скоростей по синонимичным сайтам. Замечательно, что формы графиков распределений по ортологичным белкам исключительно похожи, практически одинаковы для всех изученных клеточных форм жизни, от бактерий и архей до млекопитающих (см. рис. 4–2; Grishin et al., 2000; Wolf et al., 2009). Все эти распределения имеют так называемую логарифмически нормальную форму, то есть распределение логарифма эволюционной скорости близко к нормальному (распределению Гаусса, функция плотности вероятности которого имеет колоколообразную форму). В теории случайных процессов такая форма обычно представляет собой результат произведения многих независимых случайных величин. Универсальность функции распределения среди различных организмов, обладающих глубокими различиями в функциональной организации и сильно различающихся по размеру геномов, представляется неожиданной и может указывать на существование фундаментальных, простых объяснений, которые мы и обсудим в этой главе.

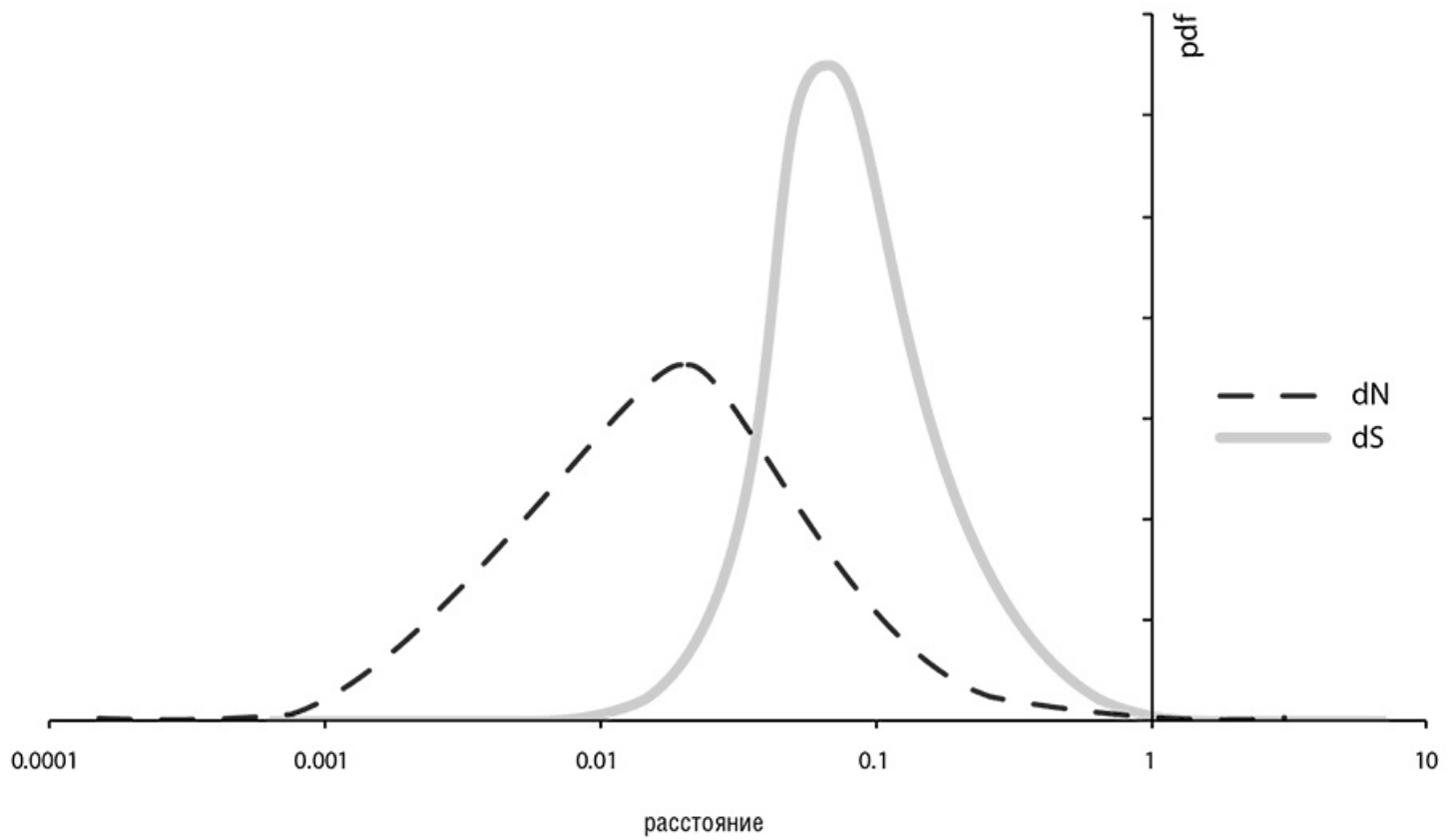


Рис. 4–1. Распределения скорости эволюции по несинонимичным и синонимичным сайтам в ортологичных генах человека и мыши: dN = скорость эволюции по несинонимичным сайтам; dS = по синонимичным; pdf = функция плотности вероятности. Данные из Wolf et al., 2009; для расчетов использовался пакет PALM (Yang, 2007)

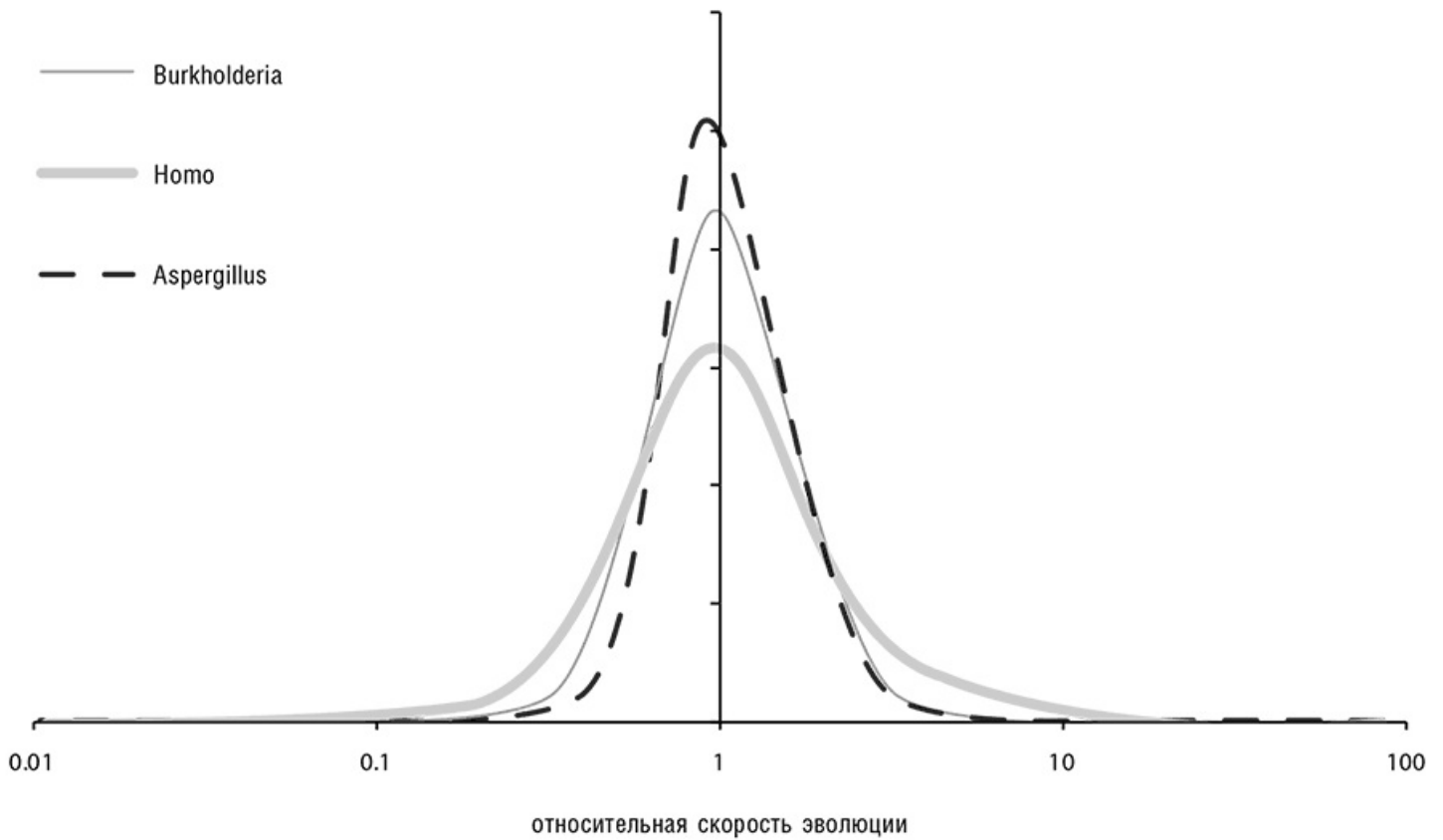


Рис. 4–2. Распределения скорости эволюции в наборах ортологичных генов бактерий и эукариот; *Burkholderia* = распределения для ортологов *Burkholderia cenosepacia* и *Burkholderia vietnamiensis* (протеобактерия); *Homo* = для ортологов человека и макаки-резус (приматы). *Aspergillus* = для ортологов *Aspergillus fumigatus* и *Neosartorya fischeri* (аскомицеты). Данные из Lobkovsky et al., 2010; для расчетов использовался пакет PALM (Yang, 2007)

Прогресс в системной биологии позволил измерить корреляции между скоростью эволюции и всеми возможными молекулярно-фенотипическими величинами, такими как уровень экспрессии, концентрации белков, белок-белковые взаимодействия, фенотипический эффект генной мутации и другими (Koonin and Wolf, 2006). Эти поиски корреляций стали практически самостоятельной областью исследований, цель которых, однако, состоит не в описании самих корреляций, а в построении физически осмысленной модели эволюции геномов и феномов. Было найдено много важных корреляций, что позволило увидеть существование некоторых закономерностей, несмотря на «зашумленность» молекулярно-фенотипических данных (особенно данных, полученных на ранних этапах исследований). На рис. 4–3 представлена простая и наглядная, хоть и неизбежно упрощенная общая картина результатов исследований (Wolf et al., 2006). Обобщение результатов показывает, что существуют два обширных класса переменных:

1. *Интенсивные, эволюционные переменные* – различные скорости геномных изменений, включая эволюцию последовательностей, потерю гена, перестройку генома и другие виды эволюционных процессов.

2. *Экстенсивные, фенотипические переменные* – скорость экспрессии, скорость трансляции, концентрация белка, частота взаимодействия с другими изучаемыми объектами.

Корреляции внутри каждого из двух классов обычно положительные, а корреляции между двумя классами – отрицательные (рис. 4–3). Эта закономерность предполагает модель «статуса

генов», в которой высокостатусные гены эволюционируют медленно, имеют высокий уровень экспрессии и взаимодействуют со многими другими генами. Гены с низким статусом меняются быстро и имеют низкий уровень экспрессии и меньшее число партнеров (рис. 4–4).

	эволюционные переменные	фенотипические переменные
	скорость скорость скорость ...	КОЛИ- ЧЕСТВО КОЛИ- ЧЕСТВО КОЛИ- ЧЕСТВО ...
скорость скорость скорость ...	положительная корреляция	отрицательная корреляция
КОЛИ- ЧЕСТВО КОЛИ- ЧЕСТВО КОЛИ- ЧЕСТВО ...	отрицательная корреляция	положительная корреляция

Рис. 4–3. Схематическая обобщенная картина корреляций эволюционных и молекулярно-фенотипических переменных.

Сильнейшая, универсальная связь между эволюционными и молекулярно-фенотипическими переменными состоит в отрицательной корреляции скорости эволюции белок-кодирующих генов и уровня экспрессии: *высокоэкспрессированные гены эволюционируют медленно*. Эта зависимость наблюдается у всех организмов, для которых есть данные по экспрессии генов (Drummond et al., 2006; Drummond and Wilke, 2008; Pal et al., 2001). Поскольку, как отмечено выше, существует положительная корреляция между *K_{ai}* и *K_s*, неудивительно, что скорости эволюции синонимических и несинонимических сайтов связаны с

уровнем экспрессии гена качественно одним и тем же образом. Более неожиданно то, что зависимость между экспрессией и скоростью эволюции соблюдается и для 3'-нетранслируемого участка (НТУ), хотя и не обнаружена для 5'-НТУ (Jordan et al., 2004). Эта универсальная отрицательная корреляция проявляется еще сильнее, если сравнивать скорость эволюции напрямую с экспериментально измеренными концентрациями белка (Schrimpf et al., 2009).

Открытие универсальной связи между экспрессией генов и их эволюцией стимулировало смелую попытку новой интерпретации, согласно которой эволюция белков определяется в большей степени принципами структуры и укладки белка, общими для всех организмов, чем его уникальными биологическими функциями. Было выдвинуто предположение, впервые – в работе Алана Драмонда и Клауса Вилке, о том, что главным фактором отбора в эволюции белка является его устойчивость к неправильной укладке. Согласно этой гипотезе, влияние мутации, как геномной, так и фенотипической (вызванной ошибками трансляции), на приспособленность организма в первую очередь рассматривается как следствие отрицательного эффекта от неправильной укладки белка, которая, помимо вызываемых ею энергетических издержек, может быть еще и токсичной для клетки (Drummond et al., 2005; Drummond and Wilke, 2008). Не углубляясь в детали, заметим, что эта интуитивно привлекательная модель может естественным образом объяснить отрицательную корреляцию между экспрессией генов и эволюцией генных последовательностей: очевидно, что негативный эффект от неправильной укладки должен быть выше для высокоэкспрессированных белков, чем для белков, производимых в небольших количествах. Другими словами, уровень экспрессии – это линза, которая увеличивает любое негативное влияние на приспособленность, связанное с данной последовательностью белков, и важнейшее из таких влияний обусловлено неправильной укладкой белка. Таким образом, гены высокоэкспрессируемых белков подвергаются бóльшим ограничениям, следствием чего является низкая скорость их эволюции. Эта гипотеза совместима с твердо установленным принципом предпочтительного выбора кодона (среди синонимичных кодонов чаще встречается оптимальный) ^[37] в высокоэкспрессируемых и высококонсервативных белок-кодирующих генах, а также с положительной корреляцией между *K_{ai}* и *K_s*. Согласно гипотезе эволюции, движимой ошибками укладки, эволюция синонимичных сайтов ограничена, по крайней мере частично, теми же самыми факторами, что и эволюция последовательности белка, поскольку выбор оптимального кодона, обеспечивающий более быструю и точную трансляцию, особенно важен для высокоэкспрессированных белков и для тех конкретных позиций, которые влияют на укладку белка. Таким же образом можно объяснить и эволюцию 3-НТУ – этот нетранслируемый участок используется для регуляции процесса трансляции.

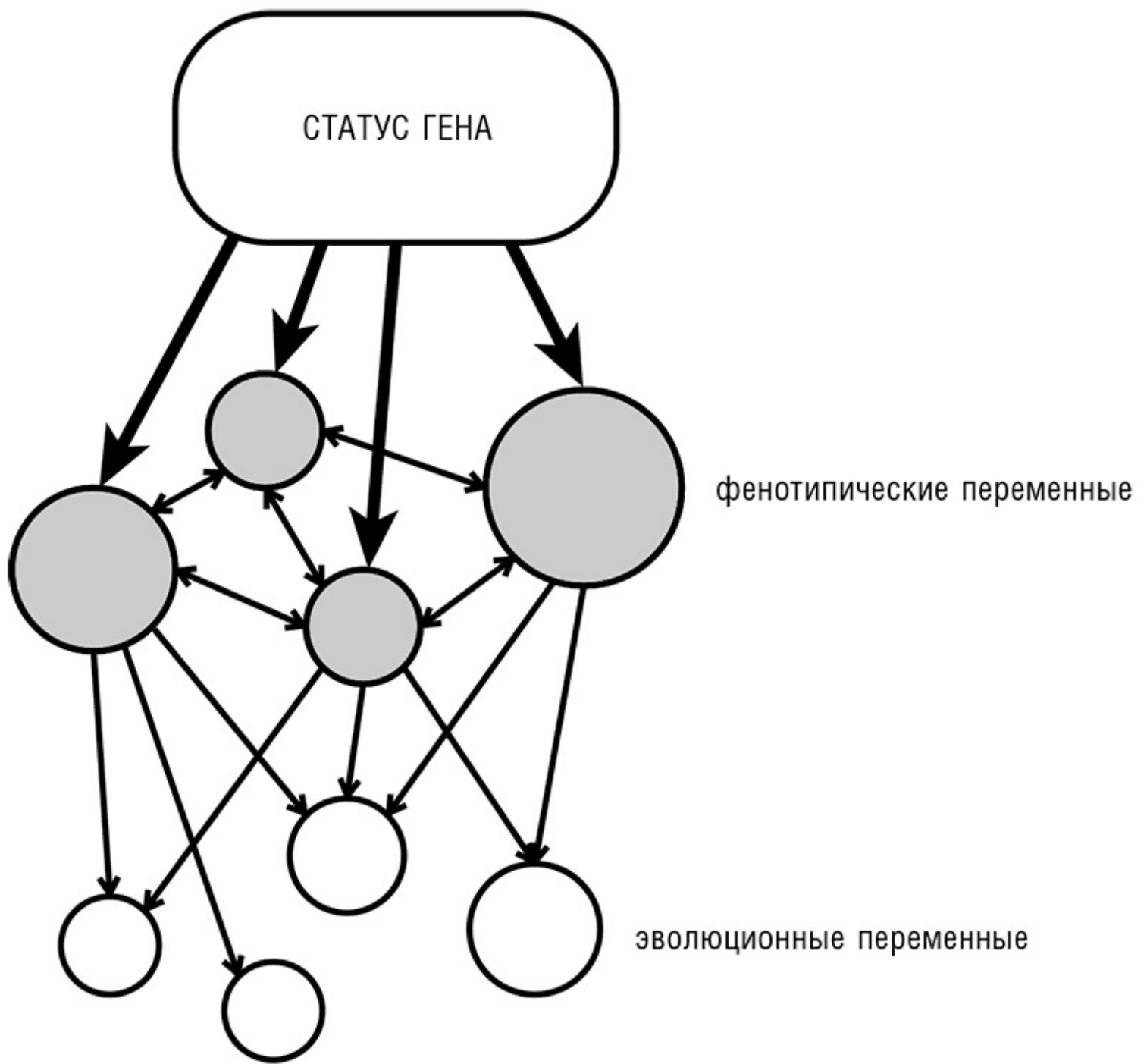


Рис. 4–4. Модель «статуса генов».

В работе, выполненной совместно с Александром Лобковским и Юрием Вульфом, мы задались вопросом, не удастся ли в данном случае убить двух зайцев, то есть возможно ли объяснить эволюцию белок-кодирующих генов и практически повсеместную отрицательную корреляцию между эволюционными скоростями и уровнем экспрессии в рамках одной простой модели (Lobkovsky et al., 2010). В анализе эволюции, ведомой ошибками укладки, проводимом для «безрешеточной» [38] модели укладки белка, были получены оценки эволюционных скоростей для гипотетического случая, в котором ошибки укладки белка являются единственным фактором, влияющим на приспособленность организма. Результаты анализа воспроизвели, и весьма точно, универсальное распределение эволюционных скоростей белков, а также зависимость между скоростью эволюции и экспрессией. Этот результат позволяет предположить, что универсальный закон распределения скоростей эволюции и в самом деле вытекает из фундаментальных физических принципов укладки белка.

Слабость или даже отсутствиен некоторых интуитивно ожидаемых корреляций между эволюционными и фенотипическими переменными кажутся не менее поразительными, чем обнаруженные корреляции. В самом деле, биологическая интуиция всей этой области знаний подсказывает, как было указано в начале этого раздела, что гены с большей «биологической значимостью» будут эволюционировать медленнее и будут теряться с меньшей вероятностью (Wilson et al., 1977). Общее понятие биологической важности можно конкретизировать измерением фенотипических эффектов от нокаута или других мутаций многих генов – желательно всех генов многих организмов. Можно предположить, что чем больше эффект от нокаута гена, тем медленнее этот ген будет эволюционировать, и гены, утрата которых вызывает летальный эффект, будут эволюционировать значительно медленнее генов, менее существенных для выживания. К настоящему времени проведено сравнение фенотипических эффектов нокаута генов и скоростей эволюционирования генов для множества модельных организмов и получен недвусмысленный и как будто парадоксальный результат: *связь между экспериментально измеренной биологической важностью гена и скоростью его эволюции очень слаба, если вообще существует* (Hurst and Smith, 1999; Jordan et al., 2002; Krylov et al., 2003; Wang and Zhang, 2009). Еще более удивительным кажется отсутствие сильной корреляции между скоростью утраты гена в течение эволюции, представляющей в некотором смысле временную меру биологической важности или существенности гена, и экспериментально определенным эффектом на приспособленность: только те наборы генов, которые вообще не утрачиваются на продолжительных отрезках эволюции, таких как вся эволюция эукариот, обогащены «важными» генами (Krylov et al., 2003; Wang and Zhang, 2009). Первые работы, показавшие почти полное отсутствие связи между скоростью эволюции и биологической значимостью, были основаны на простом измерении эффекта (присутствует/отсутствует) нокаута гена (соответственно, существенный/несущественный ген). Можно предполагать, что такие измерения слишком грубы и не дают осмысленной оценки биологической важности. Однако, например, в последних работах лаборатории Джорджа Занга было продемонстрировано почти полное отсутствие корреляции между скоростью эволюции и весьма точно измеренным влиянием на приспособленность пекарских дрожжей (*S. cerevisiae*) во множестве различных условий (Wang and Zhang, 2009), что уменьшает вероятность получения неадекватных результатов.

Чем же тогда объяснить неожиданно слабую связь между эволюцией и функцией гена? По отношению к эволюции последовательностей можно предположить, что скорость эволюции более зависит от внутренних характеристик гена (в особенности от структуры кодируемого белка), чем от его биологической значимости. Однако это объяснение неприменимо к случаям утраты гена. Наиболее осмысленным – хотя опять-таки противоречащим здравому смыслу – кажется следующее объяснение: *фенотипический эффект нокаута гена (и в целом – набора необходимых генов) не проявляется как консервативное свойство в эволюционном процессе и быстро меняется (в масштабах эволюционной шкалы), вероятно благодаря высокой скорости эволюции сетей взаимодействующих генов* [39]. Ясно, что это предположение можно проверить опытным путем, пусть и с помощью трудоемких экспериментов.

Почти нейтральные сети и белковая эволюция

В целом скорость эволюции гена определяется размером его почти нейтральной сети, то есть множества последовательностей, получаемых друг из друга в результате одношаговых мутаций (пусть и необязательно с помощью одной замены) и имеющих приспособленность примерно такую же, как и у наиболее приспособленной последовательности (Wagner, 2008a; Wolf et al., 2010). Чем больше нейтральная сеть, тем слабее ограничения для конкретного гена, тем быстрее он может эволюционировать (рис. 4–5).

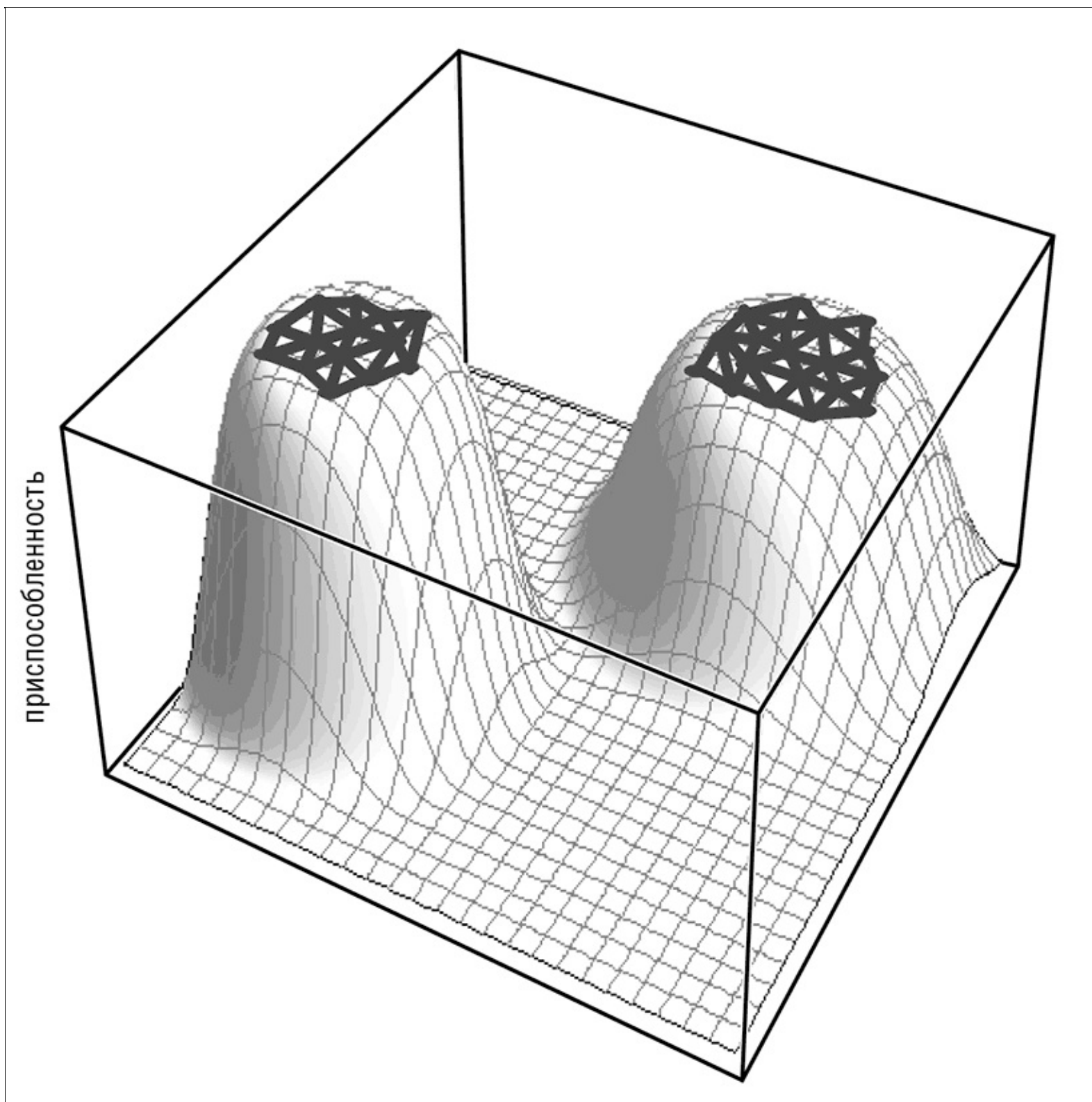


Рис. 4–5. Почти нейтральная сеть и белковая эволюция. Две почти нейтральные сети для двух вымышленных белков схематично представлены как две области, находящиеся на широких

вершинах пиков приспособленности.

В эволюции белка приспособленность отдельной последовательности в основном зависит от ее устойчивости к ошибкам укладки и от уровня экспрессии, а размер почти нейтральной сети зависит от высоты и формы пика, занимаемого этой последовательностью и ее соседями на ландшафте устойчивости (рис. 4–6). В этой модели высокоэкспрессированные белки, чьи исходные последовательности высоко устойчивы к ошибкам укладки, занимают высокие и крутые пики с небольшой областью высокой приспособленности (малые почти нейтральные сети) и, следовательно, подвержены сильному стабилизирующему отбору и медленно эволюционируют. И наоборот, белки с более низким уровнем экспрессии и меньшей устойчивостью находятся на более низких, пологих пиках, имеют более широкую область высокой приспособленности и, соответственно, подвержены более слабому отбору и имеют высокую скорость эволюции (рис. 4–6; Wolf et al., 2010).

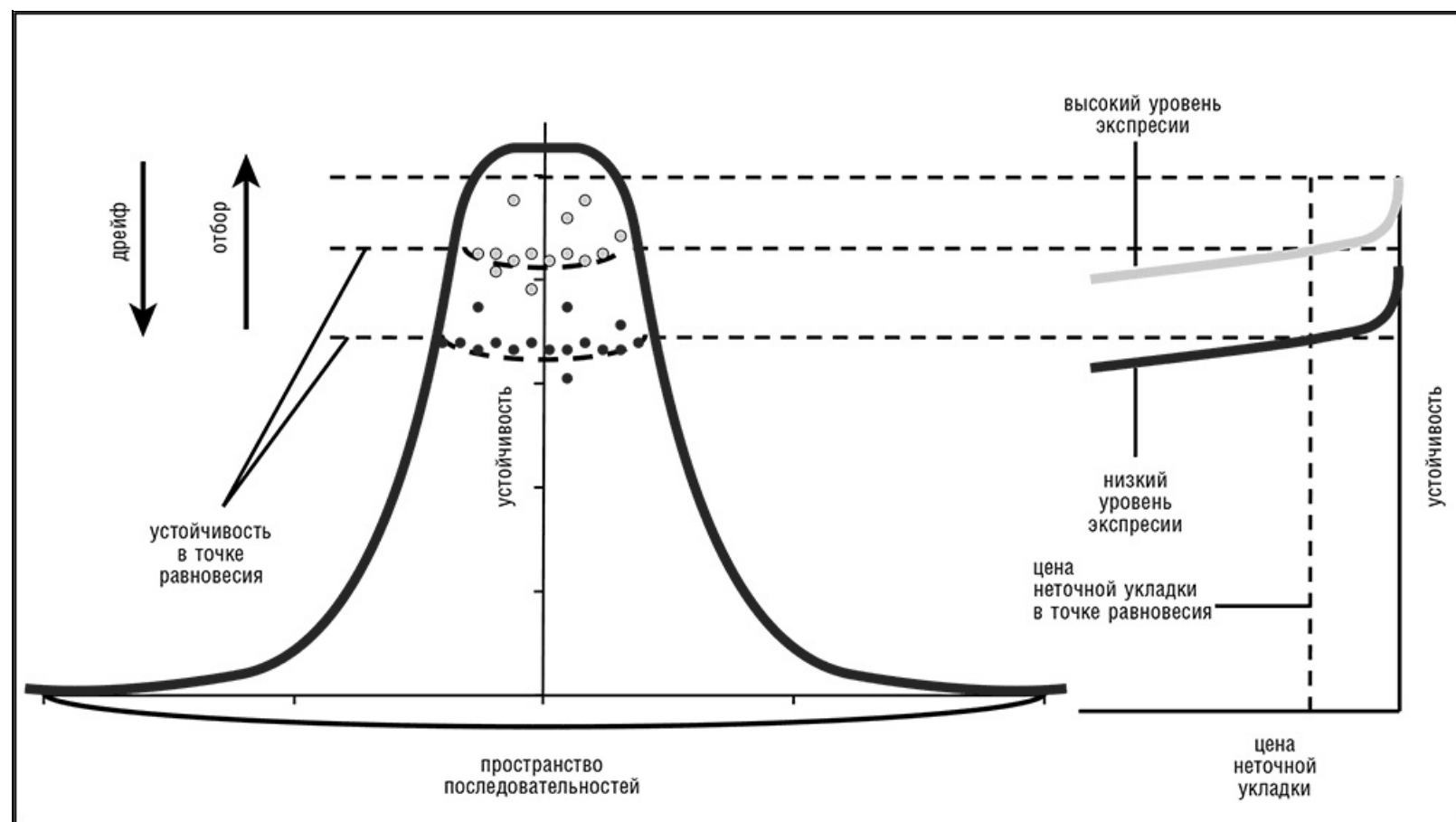


Рис. 4–6. Модель эволюции белка, управляемой издержками неточной укладки.

Геномная эволюция путем дупликации генов, модель рождения и смерти гена и универсальное распределение численности паралогичных семейств

Мы уже касались вопроса дупликации генов в предыдущей главе, в частности при построении списка важнейших механизмов геномной эволюции. Однако есть еще по крайней мере два хороших повода вернуться к этому механизму эволюции и обсудить его более детально. Во-первых, дупликация несомненно является одним из главных путей геномной эволюции для всех форм жизни и играет принципиальную роль в эволюции эукариот (см. гл. 8). Во-вторых, эволюция путем дупликации генов представляет собой формально простой процесс, для которого довольно легко построить хорошо работающие физические (или математические) модели, которые и рассматриваются в этой главе.

Представление о дупликации как об исключительно эффективном способе геномной эволюции лежит в основе современного эволюционного мышления. Упрощенно говоря, сущность этого представления состоит в том, что создание новых функциональных объектов (белков и РНК) путем модификации уже имеющихся (вспомните модель «эволюции как мастерского» Жакоба; Jacob, 1977) – это, очевидно, намного более простой путь, чем создание этих объектов с самого начала, с нуля (история этой идеи рассказывается в гл. 2). Как и другие представления геномики, эволюция путем геномной дупликации имеет строгие количественные подтверждения – большинство генов в геноме принадлежат семействам паралогов (за исключением очень небольших геномов, таких как у микоплазмы и других паразитических бактерий; Jordan et al., 2001). Более детальная реконструкция эволюции показывает, что дупликация проявляется, с различной интенсивностью, на всех этапах эволюции, таким образом, любой геном – это набор дупликаций самого различного возраста. Выбирая некоторую эволюционную линию, допустим, животные – хордовые – млекопитающие – приматы – и т. д., мы можем обнаружить в геноме (например, в нашем) все соответствующие классы дупликаций: дупликации, специфичные для животных, дупликации, специфичные для хордовых, специфичные для приматов и т. д. (Lespinet et al., 2002).

Распределение численности паралогичных семейств в любом геноме – еще одна универсальная статистическая закономерность, обнаруженная сравнительной геномикой (рис. 4–7). Распределения для всех геномов приблизительно описываются степенной функцией с отрицательным показателем степени: $y = ax^{-\gamma}$ (где y – положительное число, a – коэффициент; Koonin et al., 2002; Luscombe et al., 2002). Эти распределения, имеющие в двойных логарифмических координатах вид прямых линий, показывают, что большинство семейств по численности малы (включая семейства геномов с преобладанием синглетонных сайтов), и только немногие семейства включают в себя большое число паралогов.

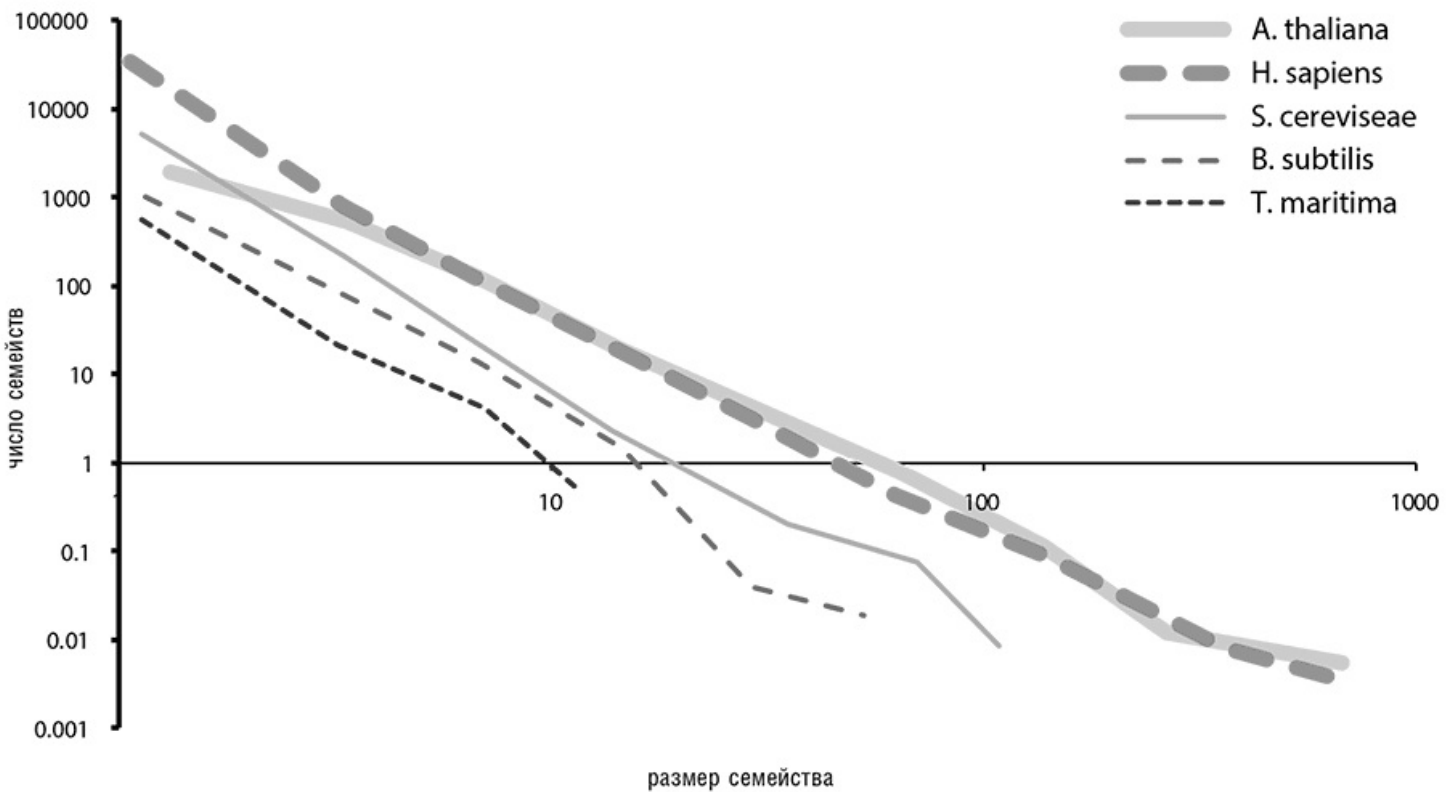


Рис. 4–7. Распределение размера семейств паралогичных генов для нескольких сильно отличающихся геномов. Показаны распределения для растения резуховидки Таля (*A. thaliana*), человека (*H. sapiens*), для пекарских дрожжей (*S. cerevisiae*), для сенной палочки (*B. subtilis*) и для бактерии *Thermotoga maritima*. Использованы данные из базы данных EggNog (Jensen et al., 2008)

Возникновение универсальной степенной закономерности распределения численности паралогичных семейств может быть описано с высокой точностью простой математической моделью эволюционного процесса (рис. 4–8). Эта модель основана на математической теории так называемых процессов рождения и смерти (один из видов марковских процессов) и для случая эволюции путем геной дупликации чаще называется моделью рождения, смерти и инновации (Karev et al., 2002). В рамках этой модели рождение – это такая геной дупликация, при которой появляется новый член паралогичного семейства, смерть – утрата гена, а инновация – это рождение нового семейства либо путем такой дупликации, которая вызывает быструю эволюции и тем самым как бы стирает «память» старого семейства, либо путем горизонтального переноса генов [40]. Наиболее интересный результат этого моделирования состоит в том, что эта модель эволюции путем геной дупликации воспроизводит рассмотренные нами распределения численности семейств паралогичных генов только при соблюдении вполне определенных условий: частоты рождения и смерти гена должны быть примерно равными и зависеть от численности семейства таким образом, чтобы большие семейства оказывались более динамичными, чем маленькие.

Стоит подчеркнуть, что динамика эволюции геного семейства описывается именно той стохастической моделью, которая используется в статистической физике. Однако, чтобы эта модель была совместима с полученными данными, необходимо соблюдение тонкого баланса между рождением, смертью и обновлением, и похоже, что этот баланс поддерживается

естественным отбором. Примечательно, что эта и подобные модели описывают с одинаковой точностью эволюцию геномов как прокариот, так и эукариот, несмотря на существенные различия между процессами, ведущими к образованию семейств паралогичных генов. Для эукариот важнейшим, если не единственным, процессом, лежащим в основе эволюции семейств, является «честная» геновая дупликация, а для прокариот количественно более важным является горизонтальный перенос генов (поэтому такие геновые семейства «псевдопаралогичны»; см. гл. 5 и 7). Тот факт, что рассмотренные здесь модели одинаково хорошо описывают биологически отличающиеся процессы эволюции генома, ведущие к сходным результатам, с одной стороны, подчеркивает универсальность этих моделей, а с другой – указывает на их ограниченную ценность для биолога.

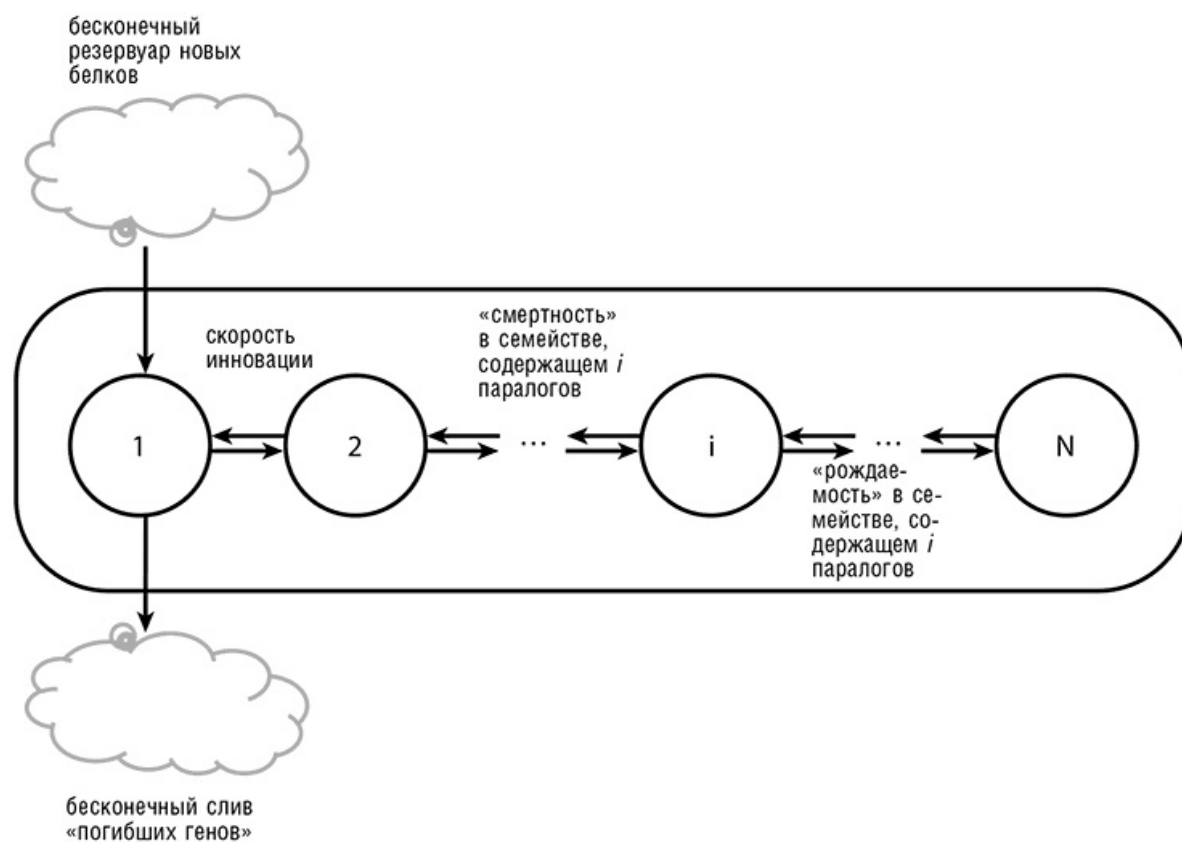


Рис. 4–8. Модель рождения, смерти и инновации в применении к эволюции геновых семейств. Под рождением подразумевается геновая дупликация или приобретение псевдопаралогичного гена путем горизонтального переноса с последующим расширением паралогичного семейства, смертью называется утрата гена (независимо от способа утраты), а инновацией считается приобретение нового гена, который становится родоначальником нового семейства (Karev et al., 2002)

Структура и эволюция сетей: всеобщность степенного закона и стоящие за ним фундаментальные процессы

Сеть (*network*) – популярнейшее понятие системной биологии, повсеместно пронизывающее современную культуру, не только в рамках биологии или науки в целом [41]. В самом деле, трудно придумать более естественный способ представления связей между многочисленными объектами, чем сеть (в математике рассматриваемую как ориентированный или неориентированный граф). В биологическом контексте узлами (или иначе – вершинами) сети часто представляют гены или белки, а ребрами (связями между узлами) обозначают их взаимодействия, которые могут быть физическими, генетическими или регуляторными (Barabasi and Oltvai, 2004). К настоящему времени разработано множество методов описания и сравнения структур (топологий) сетей (табл. 4–1). Наиболее часто для анализа используется понятие функции распределения степеней вершин, где под степенью вершины понимают число ребер, связывающих эту вершину с другими. Сравнение таких функций, выполненное для сетей различного типа, показало принципиальное отличие биологических сетей (а также многих небологических, включая Интернет) от случайных графов: случайные графы имеют колоколообразное распределение Пуассона, а для биологических сетей распределения описываются степенной функцией (табл. 4–1). Сети, имеющие степенные функции распределения степеней вершин, называют масштабнo-инвариантными сетями, так как графики их функций внешне не меняются при масштабировании (обратите внимание на прямую линию в двойных логарифмических координатах на табл. 4–1). Такие сети всегда содержат небольшое число вершин с высокими степенями, так называемых хабов (*hubs*), и большое число слабосвязанных вершин.

Таблица 4–1. Случайные и масштабнo-инвариантные сети.

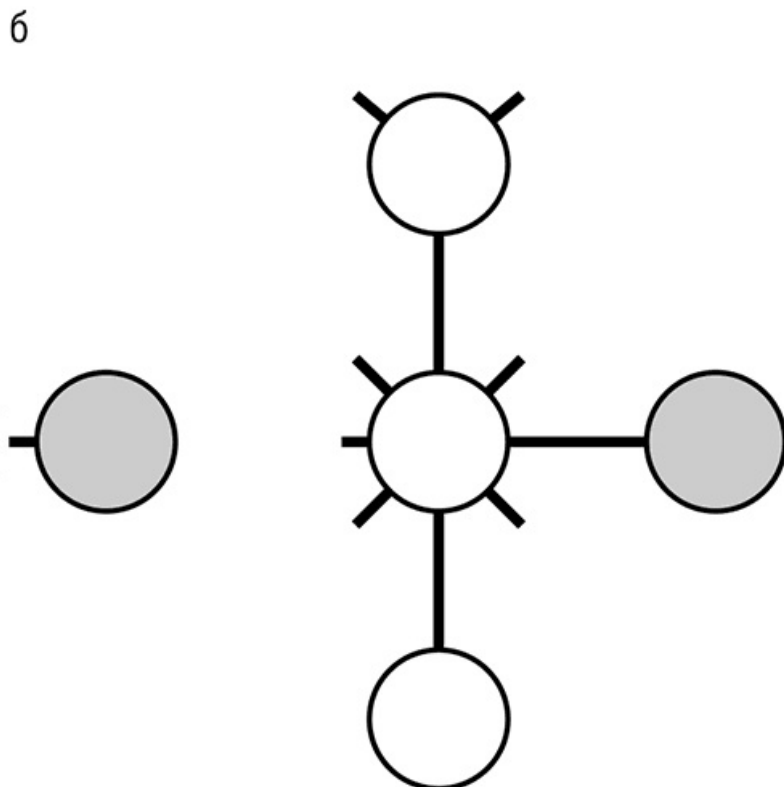
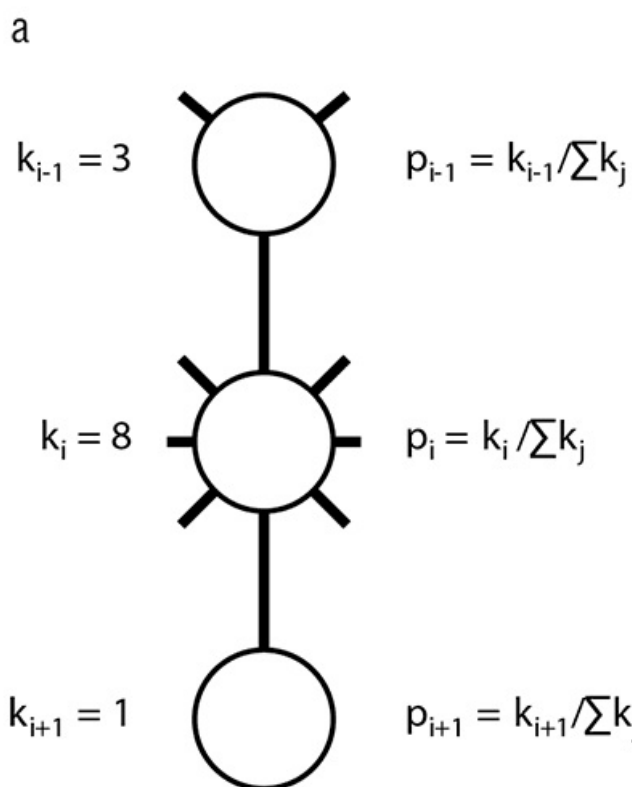
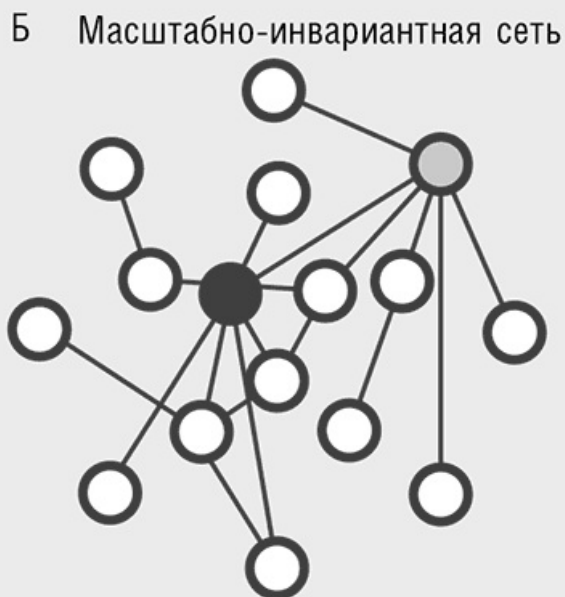
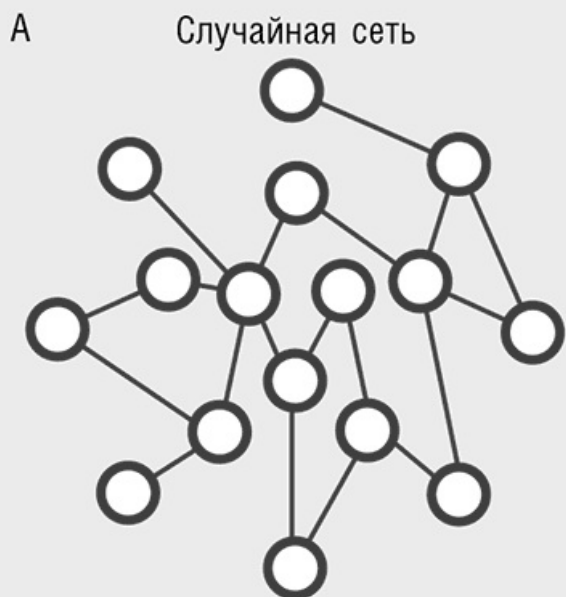


Рис. 4–9. Предпочтительное присоединение в эволюции биологических сетей: а– фрагмент

сети и новый добавляемый элемент; b – результат включения новой вершины в сеть. K_i = степень i -й вершины, p_i = вероятность присоединения новой вершины к вершине i (см. табл. 4–1)

Примечательно, что степенная функция распределения степеней вершин, по всей видимости, является неотъемлемым свойством эволюционирующих сетей (включая Интернет) и не обязательно имеет биологическое происхождение. Все типы биологических сетей, как описывающих физические взаимодействия между белками, так и отражающих взаимную регуляцию генов, несомненно, появились в результате эволюции и обладают указанным типом распределения (другими словами, являются масштабнo-инвариантными). Для объяснения универсального степенного закона распределения Барабаша с коллегами предложили принцип *предпочтительного присоединения* (preferential attachment) новых вершин, что на простом циничном языке означает, что в процессе эволюции сети «богатые делают еще богаче» (Barabasi, 2002). Предпочтительное присоединение представляет собой стохастический, неадаптивный процесс. В самом деле, когда создается новый сайт в Интернете и случайно связывается с другими сайтами, с большей вероятностью он окажется связанным с хабом, чем с изолированным сайтом, просто потому что очень многие различные пути в сети ведут к хабам (табл. 4–1). Этот режим эволюции по своей природе консервативен – сеть сохраняет свою структуру в процессе роста. Является ли предпочтительное присоединение главным принципом эволюции биологических сетей? По этому вопросу еще не достигнуто согласия. В случае если этот принцип существен для биологических сетей, должны обнаружиться некоторые специфические биологические механизмы, обеспечивающие его выполнение (рис. 4–9). Высокая интерактивность хабов, представленная «липкостью» некоторых белков, склонных к взаимодействиям, не обязательно функционально значимым, со многими другими белками, могла бы быть одним из таких механизмов. Еще более важный вклад в формирование сетей осуществляется посредством важнейшего механизма эволюции – геной дупликации. Когда ген удваивается, все имеющиеся его связи с другими генами также удваиваются, а потом начинают постепенно расходиться в процессе последующей эволюции. В простейшей модели эволюции (такой как сбалансированная модель рождения, смерти и обновления), если частота геной дупликации пропорциональна размеру семейства, структура сети (то есть распределение степеней вершин) будет сохраняться даже при отсутствии давления отбора (Koonin et al., 2002; Lynch, 2007a).

Разбиение генома по биологическим функциям: универсальный степенной закон

До сих пор в нашем обсуждении универсальных количественных закономерностей в геномной эволюции мы преднамеренно обходили стороной вопрос биологических функций. Конечно, это абстракция: геном ни в коем случае не сумма безликих «молекул», а ансамбль генов, каждый из которых кодирует определенную биологическую функцию [42]. Сначала может показаться неожиданным, что способ рассуждения, позаимствованный из статистической физики, может быть применен и к биологическим функциям. Для применения такого подхода необходимо разделить гены на большие функциональные классы, о которых можно думать как о разных типах «молекул» и которые пригодны для статистического анализа, если они включают достаточно много генов.

Как показывается в серии доскональных исследований Эрика Ван Нимвегена [43], различные функциональные классы генов по-разному соотносятся с общим числом генов в геноме (Molina and van Nimwegen, 2009; van Nimwegen, 2003). Не учитывая некоторые отклонения, для прокариот можно указать три основных показателя степени, описывающие эти соотношения: 0, 1 и 2. Генам белков, участвующих в информационных процессах (трансляции, транскрипции и репликации), соответствует показатель степени 0 – число таких генов достигает некоторого константного значения уже в минимальных геномах и в принципе не зависит от сложности генома. Число метаболических ферментов и транспортных белков примерно прямо пропорционально общему числу генов (показатель степени 1). Регуляторные гены и компоненты систем передачи сигналов показывают квадратичную зависимость (показатель степени равен 2; рис. 4-10). Показатели степени этих трех обширных классов остаются неизменными, с очень небольшими отклонениями, для всех групп прокариот, и это позволяет предположить, что разница в эволюционной динамике генов с различными функциями отражает какие-то фундаментальные законы эволюции клеточных организмов, или, другими словами, строгие и четко выраженные ограничения в функциональном устройстве геномов. Для генов эукариот обнаружены похожие, хотя и не такие явные, степенные соотношения, показатель степени для регуляторных генов эукариот значительно больше 1 (хотя и меньше 2). Имея в виду все вышесказанное, можно заключить, что эти соотношения представляют еще один набор универсалий геномной эволюции, которые становятся еще интереснее при рассмотрении их связи с функциональным устройством клетки.

Фундаментальные причины существования различных соотношений для различных функциональных классов генов еще не выяснены. Привлекательно простая модель эволюции метаболических сетей прокариот как «ящика с инструментами», предложенная Сергеем Масловым и коллегами, может быть первым шагом на пути объяснения квадратичной зависимости, характерной для регуляторных генов (Maslov et al., 2009). В этой модели ферменты, необходимые для утилизации новых метаболитов, добавляются вместе с соответствующими им регуляторами (в первую очередь посредством горизонтального переноса генов, гл. 5) во все более развитую сеть реакций. В результате усложнения сети, обеспечивающей все большее разнообразие ферментов промежуточных реакций, увеличивается отношение числа регуляторных генов к регулируемым. В какой-то момент, и его наступление можно точно предсказать, цена добавления новых регуляторов неизбежно станет слишком невыгодной («разрастающаяся бюрократия») и будет ограничивать рост сложности генома.

Гипотеза «бюрократического потолка» для верхней границы сложности генома выглядит особенно правдоподобно в свете почти полного отсутствия роста числа генов в геномах

позвоночных, особенно млекопитающих (и в наших геномах тоже), для которых связь между числом генов и размером генома очевидно нарушена (см. гл. 3 и 8). В принципе число генов могло бы быть напрямую ограничено ценой репликации ДНК, но для огромных геномов позвоночных этот фактор можно смело исключить как главное ограничение. Соответственно, цена регуляции, возможно совместно с ценой экспрессии, выглядит наиболее вероятным кандидатом на роль основного фактора, ограничивающего рост числа генов. Поэтому не случайно позвоночные (и в меньшей степени другие многоклеточные эукариоты) выработали новые, замысловатые способы увеличения сложности протеома, такие как широко распространенный альтернативный сплайсинг, альтернативная трансляция и сложная регуляция (в особенности обширный, все еще слабо изученный набор регуляторных РНК). Такие формы сложности не вызывают инфляционный рост числа белок-кодирующих генов и, таким образом, снижают по крайней мере некоторые издержки, особенно издержки трансляции (см. гл. 8).

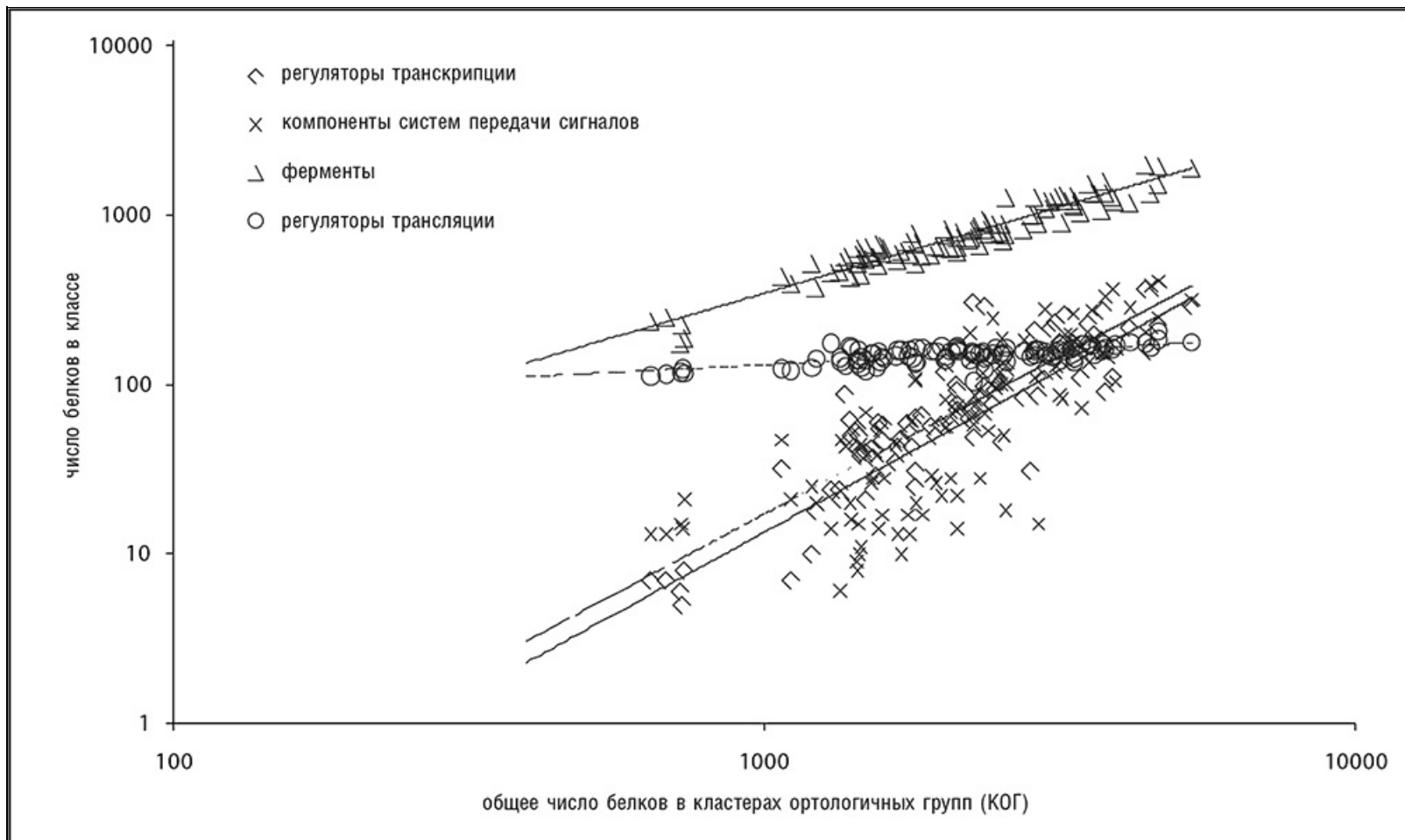


Рис. 4-10. Четко различающиеся соотношения между различными функциональными классами генов и размером генома (общим числом белок-кодирующих генов) прокариот. Использована информация из базы данных COG (Tatusov et al., 2003). График построен в двойных логарифмических координатах.

Универсальная зависимость численности функциональных классов генов от размера генома находится в обратной зависимости с ранее описанным степенным законом распределения численности семейств генов. Чем больше положительный показатель степени функции зависимости от размера генома для функционального класса генов (рис. 4-10), тем меньше отрицательный показатель степени функции распределения численности семейств в этом классе (рис. 4–7). Кажется интуитивно правдоподобным, что функциональные классы с сильной зависимостью от размера генома должны содержать много больших семейств паралогичных

генов. Обратная зависимость этих двух универсалий генома была выведена в рамках простой эволюционной модели, в которой используются правила пропорциональных вычислений для функциональной композиции генома, например «добавить два регулятора на каждый метаболический фермент» (Grilli et al., 2011). Предсказания этой модели подтверждаются эмпирическими данными для многих геномов бактерий и архей.

Стохастичность, нейтральность и отбор в эволюции

В предыдущих разделах этой главы мы ознакомились со многими количественными универсалиями, отражающими важнейшие аспекты эволюции и функционирования генома. Среди этих универсалий мы рассматривали вездесущий степенной закон распределения, который описывает как структуру всех биологических сетей, так и семейства паралогичных генов в разнообразных геномах, близкое к логарифмически нормальному распределение скоростей эволюции генов и универсальные корреляции, такие как отрицательная корреляция между генной экспрессией и скоростью эволюции. Какова природа этих универсалий? Отражают ли они какие-то глубокие свойства эволюции или это просто статистические эффекты, не имеющие отношения к пониманию биологических явлений? Здесь и далее в этой книге (гл. 13) будет отстаиваться точка зрения на эти универсалии как на нетривиальные, характерные и биологически значимые тенденции, хотя они отражают только одну из двух (а возможно, и большего числа) дополнительных (в смысле принципа дополненности Бора) составляющих эволюции жизни [44].

Во-первых, как это уже отмечалось и теперь совершенно очевидно, все эти универсалии зависимы от поведения совокупностей генов, фундаментальных единиц эволюции, рассматриваемых как статистические ансамбли. Таким образом, эти универсальные зависимости и распределения являются *эмергентными свойствами биологических систем*, то есть свойствами, проявляющимися в результате того, что эти системы состоят из многочисленных (достаточно многочисленных для проявления устойчивых статистических закономерностей) элементов (генов или белков, в зависимости от контекста), слабо взаимодействующих друг с другом (если сравнивать эти взаимодействия с теми, которые поддерживают целостность самих этих элементов).

Во-вторых, как мы уже видели, современный эволюционный анализ не останавливается на демонстрации существования универсальных понятий и законов. По крайней мере некоторые ключевые универсалии, такие как распределение эволюционных скоростей, отрицательные корреляции между скоростью эволюции и экспрессией и распределение численности паралогичных семейств, были теоретически выведены в рамках простых, но достаточно детализированных, формальных моделей эволюции. Способность простых моделей, в которых в качестве элементарных событий рассматриваются наиболее общие эволюционные процессы (такие как дупликация и утрата генов), объяснять геномные универсалии убеждает в том, что эти универсалии отражают существенные черты эволюции.

Третье, и, возможно, наиболее важное, замечание о новой парадигме понимания эволюции, которую мы пытаемся здесь обрисовать, состоит в том, что порождающие модели для общегеномных универсалий либо совсем не используют понятие отбора, либо используют только понятие очищающего (стабилизирующего) отбора. Эта форма отбора направлена на сохранение статуса-кво и наблюдается для укладки белковых молекул, для распределения численности генных семейств и для универсальной зависимости численности функциональных классов генов от общего числа генов (Koonin and Wolf, 2010b).

Аналогия между эволюционным процессом и статистической физикой не ограничена существованием универсальных зависимостей и распределений, некоторые из которых могут быть выведены в рамках простых моделей. Возможно также составить схему *детального соответствия ключевых параметров этих двух областей* (Barton and Coe, 2009; Sella and Hirsh, 2005). Такой параметр состояния (степень свободы), как положение частицы, в этой схеме является аналогом либо состояния сайта в нуклеотидной или белковой последовательности,

либо состояния гена в геноме (в зависимости от уровня моделирования эволюции), и тогда параметрам скорости эволюции для сайта или гена будет соответствовать скорость частицы. Более того, значение эффективной численности популяции будет очевидно аналогичным значению температуры в статистической физике, а приспособленность будет соответствовать свободной энергии.

Краткий обзор и перспектива: о природе эволюционного процесса

Результаты взаимопроникновения сравнительной геномики и системной биологии, обсуждаемые в данной главе, приводят нас к следующему ключевому обобщению.

Многие, чтобы не сказать все, общие закономерности геномной и молекулярно-фенотипической эволюций описываются стохастическими процессами, основанными на принципе подверженной ошибкам репликации и ограниченными очищающим отбором, который поддерживает существующую общую (но не специфическую) архитектуру генома и устройства клеток.

Это обобщение не следует понимать как исключение адаптации из числа важнейших эволюционных понятий. Разумеется, адаптация – это общее и неотъемлемое явление в эволюции всех форм жизни. Тем не менее становится все более ясным, что общие количественные характеристики геномной архитектуры, функционирования и эволюции в первую очередь определяются неадаптивными, стохастическими процессами. Адаптация только модулирует эти закономерности. Здесь становится очень соблазнительным провести вполне очевидную параллель с нейтральной теорией Кимуры. В ходе высокоуровневого анализа геномных и молекулярно-фенотипических параметров мы начинаем различать контуры «неонейтрализма» (см. также гл. 8).

Аналогия между эволюцией и стохастическими физическими процессами ни в коем случае не отрицает метафоры «эволюции как мастерского» Жакоба. Напротив, новые открытия в эволюционной геномике прекрасно вписываются в это представление об эволюции: естественный отбор (адаптивный компонент эволюции) представляет собой процесс «латания», не полную перестройку или создание нового объекта, а добавление к существующему новых частей из уже имеющихся подручных материалов. Таким образом, первичная форма отбора – это очищающий отбор, который поддерживает статус-кво. Это обобщение имеет довольно удивительное, но неизбежное следствие: большая часть наиболее значимых событий во всей истории жизни произошла в течение первых нескольких сотен миллионов лет существования жизни на Земле, до появления современного типа клеток. Этот период в истории жизни должен был качественно отличаться от всей остальной эволюции; есть основания считать, что *важнейшее достижение эволюции – это появление клетки*, все остальное уже не так важно. Мы будем обсуждать происхождение жизни с этой точки зрения в главах 11 и 12 и вернемся к обсуждению общей природы эволюции в главе 13.

Параллели между эволюционной биологией и статистической физикой оказались точными и фундаментальными до такой степени, что кажется вполне справедливым заключение о том, что это не аналогии, а *проявление общих статистических принципов (если не сказать законов) поведения больших ансамблей слабозаимодействующих объектов* [45]. Как в физике, так и в эволюционной биологии такие ансамбли (например, идеальный газ в физике и геном как сумма генов в биологии) являются идеализациями. В реальности отклонения от поведения, которое предсказывается простыми статистическими моделями, неизбежны и значимы. В эволюционной биологии такие отклонения, кроме всего прочего, вызываются различными взаимодействиями генов, что приводит к неожиданным эффектам, таким как отсутствие строгой корреляции между биологической значимостью гена и скоростью его эволюции. Тем не менее существенный эвристический потенциал прямого статистического подхода в объяснении по крайней мере некоторых фундаментальных свойств как физических, так и биологических процессов неоспорим.

Barabasi, A. L., and Z. N. Oltvai. (2004) Network Biology: Understanding the Cell's Functional Organization. // *Nature Reviews Genetics*5: 101–113.

Обзор свойств биологических сетей с акцентом на масштабной инвариантности.

Barton, N. H., and J. B. Coe. (2009) On the Application of Statistical Physics to Evolutionary Biology. *Journal of Theoretical Biology*259: 317–324.

Технически сложная, но важная работа по термодинамическому подходу в эволюционной биологии.

Drummond, D. A., and C. O. Wilke. (2009) The Evolutionary Consequences of Erroneous Protein Synthesis. *Nature Reviews Genetics*10: 715–724.

Критический обзор концепции эволюции, ограниченной ошибками трансляции и ошибками укладки белка.

Lobkovsky, A. E., Y. I. Wolf, and E. V. Koonin. (2010) Universal Distribution of Protein Evolution Rates As a Consequence of Protein Folding Physics. *Proceedings of the National Academy of Sciences USA*107: 2,983—2,988.

В этой работе эволюционная динамика выводится в рамках простой модели укладки белка и с хорошей точностью воспроизводится универсальное распределение эволюционных скоростей.

Koonin, E. V., and Y. I. Wolf. (2006) Evolutionary Systems Biology: Links Between Gene Evolution and Function // *Current Opinion in Biotechnology*17: 481–487.

Обзор корреляций между эволюционными и молекулярно-фенотипическими параметрами.

Koonin, E. V., Y. I. Wolf, and G. P. Karev. (2002) The Structure of the Protein Universe and Genome Evolution // *Nature*420: 218–223.

Обсуждение универсальных распределений и зависимостей с акцентом на роли стохастических процессов и принципе предпочтительного присоединения.

Molina, N., and E. van Nimwegen. (2009) Scaling Laws in Functional Genome Content Across Prokaryotic Clades and Lifestyles // *Trends in Genetics*25: 243–247.

Последние данные по универсальным степенным функциям для различных функциональных классов генов.

Sella, G., and A. E. Hirsh. (2005) The Application of Statistical Physics to Evolutionary Biology // *Proceedings of the National Academy of Sciences USA*102: 9,541—9,546.

Богатая идеями статья, детально описывающая формальную аналогию между статистической физикой и эволюционной динамикой.

Schroedinger, Erwin. (1944/1992) What Is Life?: With «Mind and Matter» and «Autobiographical Sketches». Cambridge, MA: Cambridge University Press. [Шредингер Э. Что такое жизнь? Физический аспект живой клетки / Пер. с англ. 3-е изд. Ижевск: РХД, 2002.

Небольшая классическая книга, которую можно порекомендовать еще раз как взгляд на биологию с физической точки зрения, не так уж и изменившийся со времен Шредингера.

Глава 5. Сетевая геномика мира прокариот: вертикальные и горизонтальные потоки генов, мобиломы и динамика пангеномов

Когда Дарвин писал об эволюции, он имел в виду животных и растения, по крайней мере он использовал эти сложные многоклеточные организмы во всех своих конкретных примерах. Одноклеточные формы жизни практически не упоминаются в «Происхождении видов...» или любой другой книге Дарвина. В любом случае, учитывая, что Дарвин серьезно обсуждал происхождение всех существующих ныне видов от одной или нескольких предковых форм (см. гл. 2 и 11), он должен был исходить из того, что эти предки были одноклеточными [\[47\]](#). Эрнст Геккель, плодовитый немецкий последователь Дарвина, поместил протист (одноклеточных эукариотов, часто называемых этим термином даже сейчас) и дробянок (ныне известных как прокариоты – бактерии и археи) в основании своего монументального древа жизни, первого из подобных деревьев, которое было населено реальными жизненными формами. Естественно, животные доминировали на дереве Геккеля, в то время как протисты и дробянки располагались на неопределенных позициях поблизости от корня.

Вездесущность и важность бактерий в биосфере постепенно становились очевидными параллельно с развитием эволюционной биологии, вначале благодаря полному драматизму исследованиям бактериальных патогенов, а позднее в результате достижений экологической микробиологии. Достаточно рано микробиологи показали, что бактерии в буквальном смысле являются основным действующим началом в биосфере. Подавляющее большинство живых клеток на нашей планете – это именно бактерии, они демонстрируют наибольшее биохимическое разнообразие среди всех организмов и являются главной геохимической силой. Однако, несмотря на биологическую важность и поразительное биохимическое и экологическое разнообразие микробов и огромный прогресс микробиологии в середине XX столетия (в качестве примеров можно упомянуть открытие антибиотиков и демонстрацию химической природы генетического материала бактерий), микробиология ничего не внесла в СТЭ и не была эволюционной дисциплиной на протяжении большей части этого столетия. Не то чтобы микробиологи совсем не думали об эволюции, но все их попытки расшифровать эволюционные взаимоотношения между бактериями, используя морфологию клеток, а также метаболические и фенотипические характеристики, и на основе этих признаков построить филогенетическую таксономию приводили к несовместимым и неприемлемым результатам. Весьма любопытно, что приблизительно в то время, когда происходила консолидация СТЭ, ведущие микробиологии того времени, включая Роджера Стейнира и Корнелиуса Ван Нейла, пришли к выводу, что, если какие-то эволюционные процессы и происходят в микромире, описать их и как-либо применить в сфере таксономии микробов и микробиологии вообще практически невозможно (Stanier and Van Niel, 1962; Van Niel, 1955).

Как отмечено в главе 3, все резко изменилось в 1977 году, когда Карл Вёзе с сотрудниками применил филогенетический анализ рРНК как основной метод изучения эволюции микробов и создания их таксономии (Woese, 1987). Возможности новой методологии были эффектно продемонстрированы открытием архей, по-видимому, первым крупным открытием в биологии, которое было сделано исключительно на основе анализа нуклеотидных последовательностей. За этим прорывом последовал период «бури и натиска» 1980-х и начала 1990-х годов, когда филогения рРНК была успешно применена для прояснения взаимоотношений среди многих групп прокариот. Среди молекулярных эволюционистов того времени превалировало мнение,

что в принципе эти методы позволят точно реконструировать эволюцию микробов.

Однако дивный новый мир микробной эволюции оказался недолговечным – эволюционная геномика вновь запутала картину самым неожиданным образом. Первый полный бактериальный геном был секвенирован в 1995 году, а первый геном археи – в 1996-м [48]. Вскоре после этого прорыва установился экспоненциальный темп секвенирования геномов со временем удвоения около 20 месяцев для бактерий и около 34 месяцев для архей (см. рис. 3–1). Сравнительный анализ сотен секвенированных бактериальных геномов и десятков геномов архей привел к важнейшему выводу: *микробы определенно эволюционируют, но их эволюция сильно отличается от той, что описана СТЭ* (Doolittle, 1999b; Woese and Goldenfeld, 2009). Ключевым стало осознание того, что геномы прокариот ведут себя не так, как если бы они были стабильными, точно наследуемыми носителями генетической информации организма (вида). Геномы микробов оказались чрезвычайно динамичными, неоднородными образованиями, которые относительно стабильны лишь на коротких интервалах времени, имеют свою характерную скорость распада и существуют в динамическом равновесии между различными формами жизни, которые отличаются по принципам геномной организации. В «мире прокариот» эти взаимосвязанные и постоянно взаимодействующие формы жизни включают не только бактерии и археи, но также различные плазмиды, вирусы и другие мобильные элементы. В этой новой, динамической парадигме прокариотической эволюции традиционная концепция видов с четко определенным, стабильным геномом теряет существенную, если не большую часть своей применимости (Doolittle and Zhaxybayeva, 2009). Становится осмысленнее говорить о сериях «пангеномов» на всех уровнях, от пангенома, например, *Escherichia coli* или любого другого «вида» бактерий или архей, до пангенома всех прокариот (Lapierre and Gogarten, 2009; Mira et al., 2010).

В главе 3 мы уже обсуждали важные аспекты структуры генетической вселенной прокариот. Она рассматривалась в основном как сложный статичный объект, то есть в терминах распределения различных существенных переменных. В этой главе мы также рассматриваем распределения, но в основном пытаемся встать на динамическую точку зрения и исследовать мир прокариот в терминах потоков генов и взаимодействия между репликонами.

Размер и общая организация бактериальных и архейных геномов

Несмотря на огромные различия в образе жизни, а также метаболической и геномной организации, бактериальные и архейные геномы демонстрируют легко различимые общие архитектурные принципы (см. обзор в гл. 3). Секвенированные бактериальные и архейные геномы охватывают два порядка величины по размерам от около 144 Кб для внутриклеточного симбионта *Hodgkinia cicadicola* до примерно 13 Мб для обитающей в почве бактерии *Sorangium cellulosum* (Koonin and Wolf, 2008b). Примечательно, что бактерии демонстрируют бимодальное распределение размеров генома ^[49] с пиком в районе примерно 5 Мб и дополнительным плато в районе примерно 2 Мб (см. рис. 5–1). Хотя существует много геномов промежуточного размера, это распределение предполагает существование двух в достаточной степени разделенных классов бактерий с «малым» и «большим» геномами. К этим наблюдениям нужно относиться с известной осторожностью, так как они могут быть артефактом, обусловленным предпочтительным секвенированием небольших геномов (в первую очередь бактериальных патогенов), но с ростом числа секвенированных геномов такое объяснение становится все менее удовлетворительным.

Археи демонстрируют более узкое, но также сложное распределение размеров генома от примерно 0,5 Мб у паразита/симбионта *Nanoarchaeum equitans* до примерно 5,5 Мб у *Methanosarcina barkeri*, с острым пиком в районе 2 Мб, который практически точно соответствует расположению плато бактериальных геномов малого размера, вторым небольшим пиком около 3 Мб и тяжелым хвостом, соответствующим геномам большего размера (см. рис. 5–1). При этом смещения в базе данных опять могут быть существенными, так как в настоящее время геномов архей секвенировано примерно на порядок меньше, чем геномов бактерий, так что пока может быть еще просто недостаточно данных для выявления истинной формы распределения размеров геномов. Однако более вероятно, что археи действительно являются менее разнородной группой, как будет обсуждаться далее в данном разделе.

Все очень маленькие (менее 1 Мб) геномы бактерий и архей принадлежат бактериям-паразитам и внутриклеточным симбионтам эукариот и единственной известной архее-паразиту (или симбионту) *Nanoarchaeum equitans*, которая живет за счет другой археи, *Ignicoccus hospitalis*. Таким образом, кажется все более вероятным, что минимальный размер генома свободно живущего прокариота, по крайней мере автотрофа, который не зависит от других форм жизни для добывания пищи, немного превышает 1 Мб. Текущий рекорд редукции генома среди свободно живущих клеток, около 1,3 Мб, принадлежит фотосинтезирующей морской альфа-протеобактерии *Pelagibacter ubique* (SAR11), которая также является наиболее распространенной из известных клеточных форм жизни на Земле (Giovannoni et al., 2005). (Связь между размером популяции и размером генома потенциально важна, мы вернемся к этому вопросу в гл. 8.)

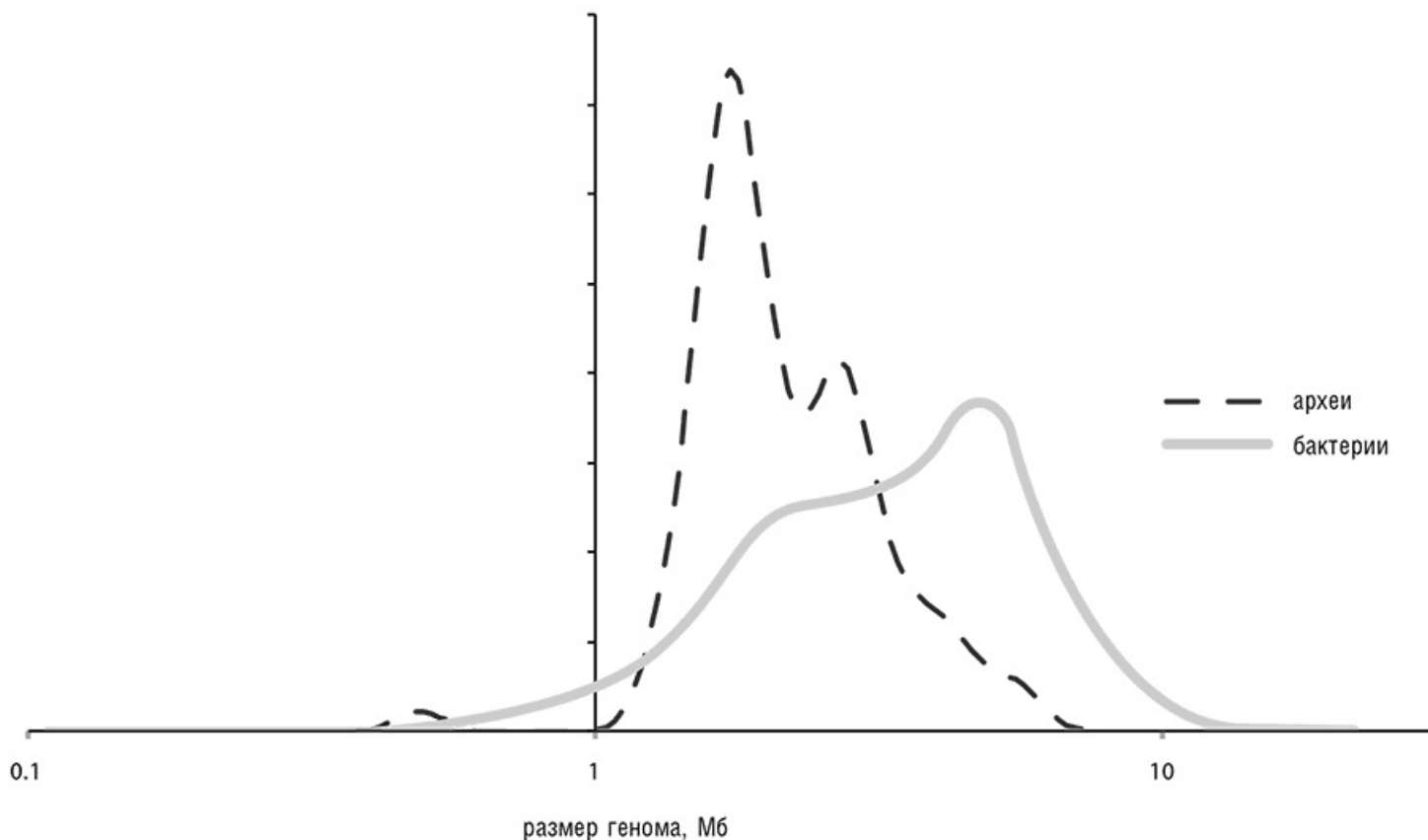


Рис. 5–1. Распределение размеров геномов среди бактерий и архей.

Как мы уже обсуждали в главе 3, бактериальные и архейные геномы характеризуются высокой плотностью белок-кодирующих генов, которые занимают большую часть ДНК. Бактериальные и архейные геномы демонстрируют одномодальное и довольно острое распределение плотности генов, большей частью 0,8–1,2 гена на Кб геномной ДНК (отсюда предельно простое эмпирическое правило: 1 ген на 1000 пар нуклеотидов). Архейное распределение по сравнению с бактериальным сдвинуто в сторону более высоких плотностей, таким образом, в среднем архейные геномы даже более компактны, чем бактериальные. Похоже, что как кодирующие, так и межгенные области у архей немного короче по сравнению с бактериями.

Таким образом, археи и бактерии весьма похожи в смысле характерных размеров и общей архитектуры геномов, но резко отличаются от эукариот, которые охватывают много больший интервал размеров генома, имеют белок-кодирующие гены, часто прерываемые интронами, и более длинные межгенные промежутки (см. гл. 8). Эти общие признаки бактерий и архей подтверждают концепцию «прокариотного принципа организации генома» (см. более подробно ниже).

Фрактальное пространство-время генома, пангеномы и кластеризация прокариот

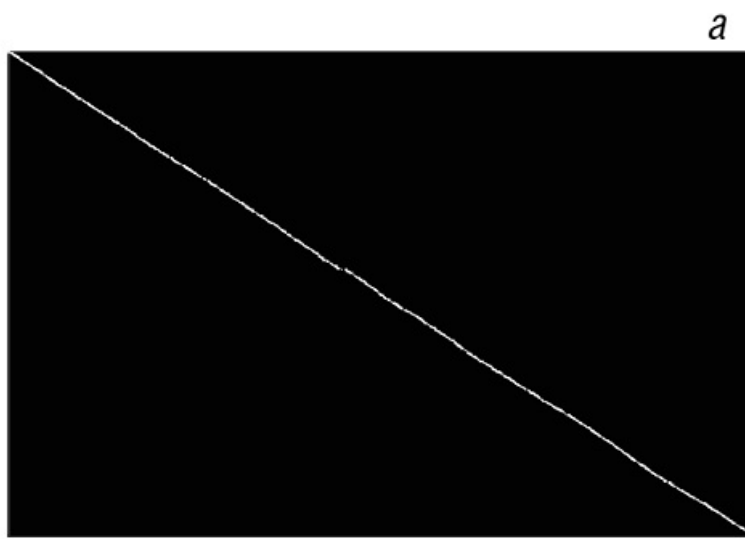
В главе 3 мы сосредоточились на трехкомпонентной структуре прокариотического геномного пространства, состоящего из ядра, оболочки и облака, и показали, что эта структура *фрактальна*. Одни и те же три компонента, а именно небольшое ядро, оболочка большего размера и огромное по сравнению с ними «облако», проявляются на любом уровне разбиения геномного пространства, от мира прокариот в целом до совсем небольших групп бактерий (см. рис. 3-14). Непосредственным следствием этой фрактальности является важность «пангеномов» – всей общности генов, представляющих геномы, принадлежащие к кластеру архей или бактерий на данном уровне. Читатель может (и должен) немедленно спросить, что определяет кластеры и откуда берутся уровни. Пока предположим, что дерево рРНК Карла Вёзе (см. рис. 2–3) разумно описывает организацию пространства-времени мира прокариот и является по крайней мере одним из источников для кластеризации. В главе 6 мы обсудим применимость и смысл концепции древа жизни более глубоко и покажем, что дерево рРНК, хотя ни в коем случае и не является полным представлением истории эволюции прокариот, тем не менее вполне осмысленно.

Огромное множество архейных и бактериальных генов кодируют белки, которые не имеют никакого измеримого сходства с какими-либо другими доступными последовательностями белков. Эти гены часто обозначают как одинокие рамки считывания (ОРС) ^[50] (Daubin and Ochman, 2004). Обычно в архейных и бактериальных геномах ОРС составляют 10–15 процентов от всех предсказанных генов. Многие ОРС – очень короткие, и некоторые из них могут быть не реальными генами, а результатом ошибочного предсказания при анализе генома (Ochman, 2002). Кроме того, высказывается предположение, что большинство ОРС, являющихся полноценными генами, произошли от генов бактериофагов и, соответственно, характеризуются высокой горизонтальной мобильностью, хотя в некоторых случаях они могут быть задействованы для клеточных функций и, соответственно, фиксируются в бактериальных и архейных геномах. Последние оценки, следующие из метагеномных исследований бактериофагов, предполагают, что разнообразие фаговых последовательностей очень велико и остается по большей части неизученным (Edwards and Rohwer, 2005). Таким образом, кажется привлекательной идея, что большая часть бактериальных и архейных ОРС произошла из этого огромного резервуара генов. В трехкомпонентной структуре вселенной прокариотических генов, с которой мы теперь знакомы, ОРС естественным образом объединяются с «облаком» редких генов, которые количественно доминируют в геномном пространстве, но не в индивидуальных геномах, как обсуждалось в главе 3.

Насколько велико все геномное пространство прокариот? Сколько генов в общей сложности оно содержит? Надежная экстраполяция расширения геномного пространства в результате продолжающегося секвенирования бактериальных и архейных геномов и достоверная оценка реального размера этого пространства трудноосуществимы. Однако с учетом большого разнообразия микробных виромов, которые являются основным резервуаром генов и их переносчиком (см. также гл. 10), наиболее вероятно, что число элементов прокариотического геномного пространства увеличится на порядки величины, в основном, если не исключительно, за счет расширения «облака» (Koonin and Wolf, 2008b; Lapierre and Gogarten, 2009).

Эволюционная динамика архитектуры генома прокариот: опероны, суперопероны и сети соседствующих генов

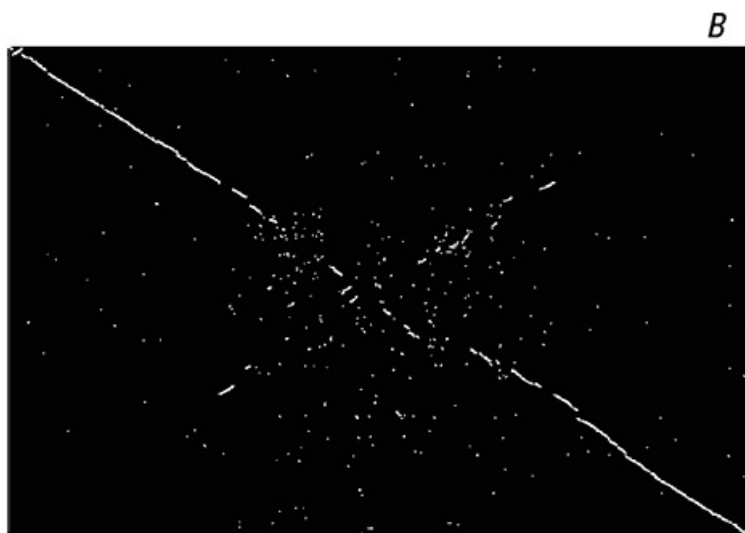
Как уже отмечалось в главе 3, практически сразу же, как только были опубликованы первые полные геномные последовательности, стало очевидным, что последовательность генов в бактериальных и архейных геномах относительно мало консервативна, она сохраняется существенно хуже, чем последовательность нуклеотидов в самих генах (см. рис. 3–6). Для того чтобы анализировать эволюцию последовательности генов, необходимо иметь надежный набор ортологичных генов в сравниваемых геномах (см. табл. 3–1). Как только такое множество ортологичных генов задано, становится достаточно просто оценить степень сохранения последовательности генов, например с помощью точечного графика (одно из самых ранних представлений степени сходства нуклеотидных и белковых последовательностей), в котором каждая точка представляет собой пару ортологов. Исследование этих графиков показывает быстрое расхождение порядка генов у прокариот таким образом, что даже между близкородственными организмами коллинеарность хромосом разрушена в нескольких точках (см. рис. 5–2 а), а умеренно разошедшиеся организмы показывают лишь несколько протяженных коллинеарных районов (см. рис. 5–2 б и 5–2 в). Для любой пары более отдаленных друг от друга организмов график выглядит как карта звездного неба (см. рис. 5–2 г). Разрушение синтении в процессе эволюции бактериальных и архейных геномов обычно явно бросается в глаза на графике, образуя картину в форме буквы X (см. рис. 5–2 б и 5–2 в). В свое время было сделано предположение, что такая картина возникает в результате симметричных хромосомных инверсий с центром в точке начала репликации (Eisen et al., 2000). Исходной причиной таких инверсий может быть высокая частота рекомбинаций в репликационных вилках, которые в кольцевых хромосомах бактерий и архей обычно располагаются с обеих сторон и на одинаковом расстоянии от точки начала репликации.



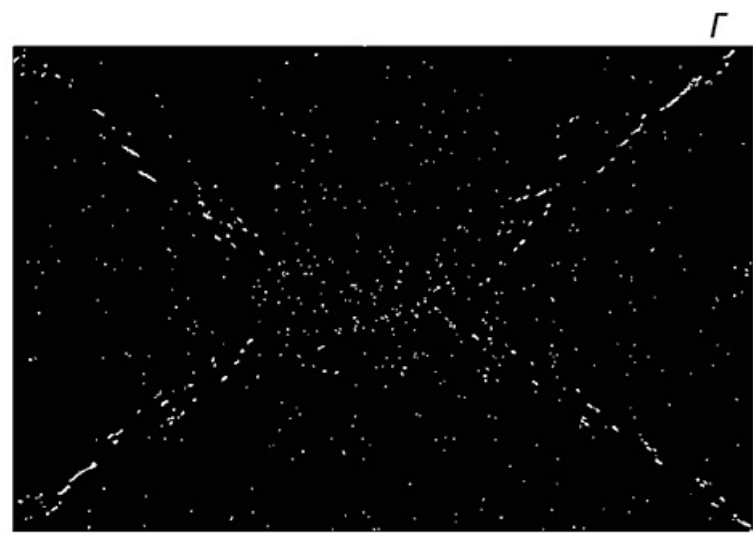
DY ~ 0; DN = 0.03



DY = 0.048; DN = 0.03



DY = 0.159; DN = 0.075



DY = 0.336; DN = 0.123

Рис. 5–2. Расхождение порядка следования генов между геномами бактерий: *a* – *Borrelia afzelii* PKo по сравнению с *Borrelia burgdorferi* B31; *б* – *Shewanella oneidensis* MR-1 по сравнению с *Shewanella sp.* ANA-3; *в* – *Pseudomonas fluorescens* PfO-1 по сравнению с *Pseudomonas fluorescens* Pf-5; *г* – *Pseudomonas fluorescens* Pf-5 по сравнению с *Pseudomonas syringae* pv. tomato str. DC3000. Каждая точка представляет пару ортологичных генов, идентифицированных с использованием метода наилучшего сходства при двунаправленном сравнении (см. табл. 3–1). Яркие точки показывают пары ортологичных генов, принадлежащих консервативным массивам генов; бледные точки показывают изолированные ортологи. DY – расстояние между сравниваемыми геномами в терминах порядка следования генов, как описано в Novichkov et al., 2009. DN – медианное расстояние между последовательностями несинонимических сайтов в белок-кодирующих генах.

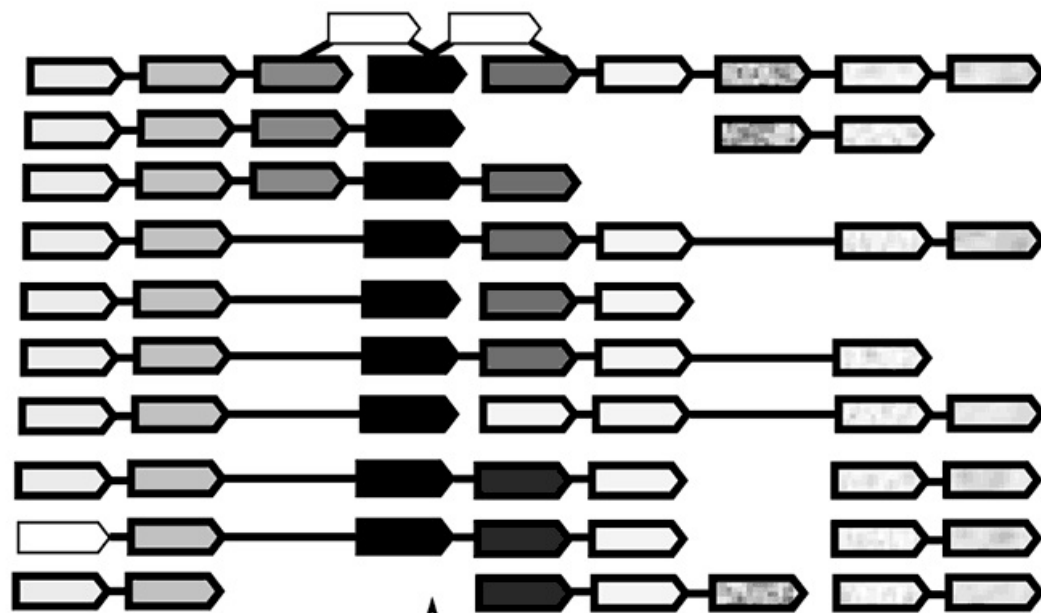
Одной из наиболее ранних фундаментальных концепций бактериальной генетики является оперон, то есть группа совместно транскрибируемых и регулируемых генов (Jacob and Monod, 1961). Гипотеза оперона – выдающийся концептуальный прорыв Франсуа Жакоба и Жака Моно. Хотя за 50 лет, прошедших с момента ее первой публикации, было открыто огромное количество вариаций простой схемы регуляции лактозного оперона Lac репрессором, оперон выдержал проверку сравнительной геномикой как главный организационный принцип бактериальных и архейных геномов. В процессе эволюции опероны сохраняются гораздо лучше,

чем протяженные синтении. Тем не менее сравнительный анализ порядка следования генов в бактериях и археях выявил небольшое количество оперонов, общих для широкого многообразия организмов. Как уже было отмечено ранее, высококонсервативные опероны, как правило, кодируют физически взаимодействующие белки, тенденция, легко объяснимая отбором, направленным против вредных эффектов дисбаланса между субъединицами сложных белковых комплексов. Наиболее эффективной иллюстрацией этой тенденции является рибосомный супероперон, включающий более 50 генов рибосомных белков, который встречается в различных комбинациях и локализациях во всех секвенированных архейных и бактериальных геномах. Анализ рибосомного супероперона и других частично сохраняющихся групп оперонов меньшего размера привел к идее *сверхоперона* (Lathe et al., 2000), или *консервативного окружения гена* (Rogozin et al., 2002), как некоего множества перекрывающихся, частично консервативных цепочек генов (известных или предсказанных оперонов; см. рис. 5–3). В дополнение к рибосомному супероперону, яркими примерами консервативного окружения являются предсказанная группа перекрывающихся оперонов, которая кодирует субъединицы экзосомного комплекса архей, и *cas*-гены, из которых состоит антивирусная система защиты (см. также гл. 9 и 10).

Большинство генов в каждом консервативном окружении кодируют белки, вовлеченные в один и тот же процесс или комплекс, но существуют и высококонсервативные участки, которые включают гены с функциями, как кажется, несвязанными. Яркий пример – частое присутствие гена енолазы в рибосомном окружении или генов субъединиц протеасомы в экзосомном окружении архей. Присутствие этих генов, на первый взгляд кажущихся неуместными в консервативном геномном окружении, может объясняться скрытой функциональной связью, плейотропией (множественностью функций соответствующих белков), или «геномным автостопом», когда оперон объединяет гены функционально не связанные, но экспрессируемые в одинаковых условиях (Rogozin et al., 2002).

Концепция геномного окружения воплощает в себе парадигму эволюции генома прокариот, если не эволюции геномов вообще, так как она ярко демонстрирует *баланс между частичным сохранением элементов ядра и огромной диверсификацией периферии* (см. рис. 5–3 а). Так же как для многих других объектов и их взаимоотношений в биологии, эти частично консервативные окружения могут быть естественным образом представлены в виде сети, в которой гены являются узлами, соседи соединены ребрами, а вес ребер пропорционален частоте встречаемости данной связи в геномах (см. рис. 5–3).

а



б

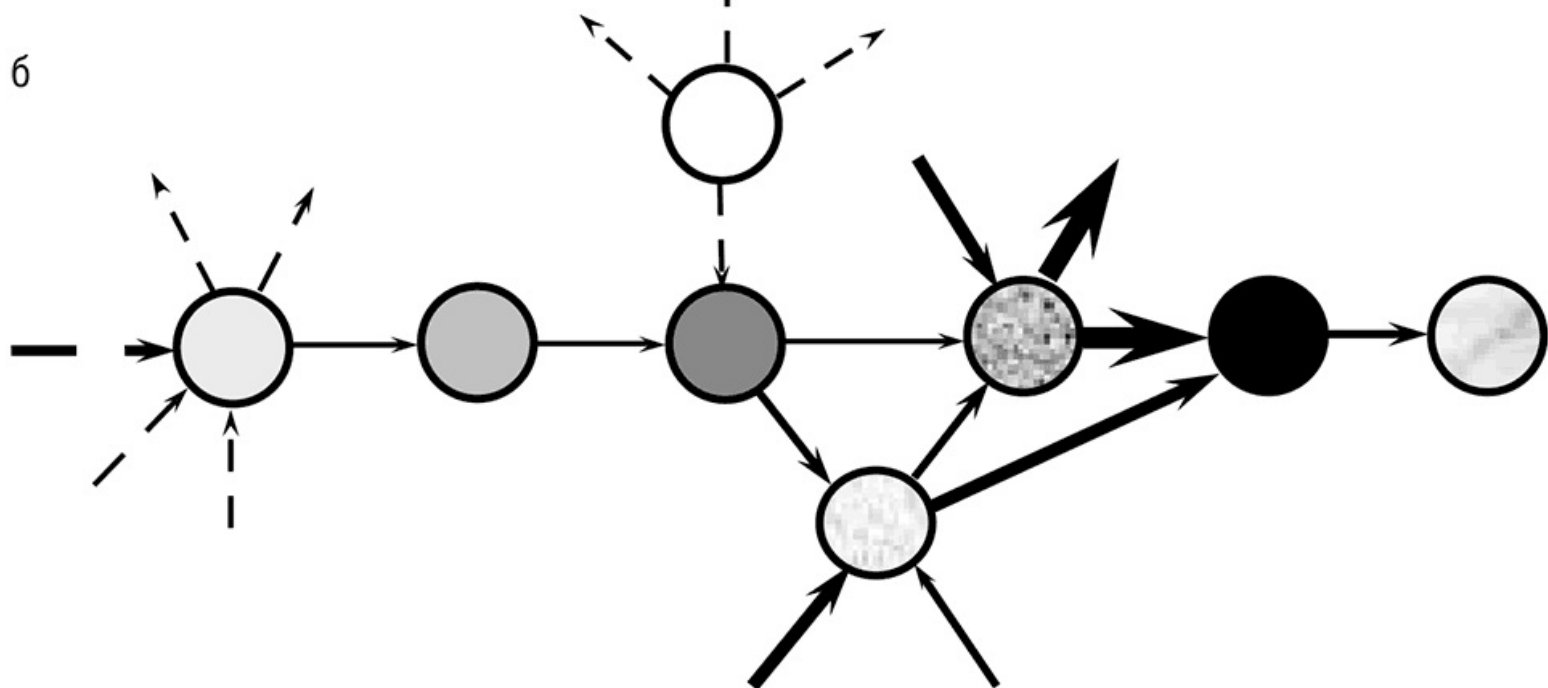


Рис. 5–3. Частично сохраняющееся окружение гена в геноме прокариот: а – перекрывающиеся, частично сохраняющиеся массивы генов. Гены показаны в форме стрелок с уникальной штриховкой или текстурой. Связывающие их жирные линии обозначают короткие межгенные промежутки, а тонкие линии показывают протяженные районы, разделяющие соответствующие гены. (Они содержат дополнительные гены и изображены без учета масштаба.) В случаях, в которых опероны не связаны, они могут располагаться в различных частях генома. На рисунке показаны реальные массивы генов, но названия конкретных геномов и генов не указаны, чтобы подчеркнуть общий характер геномной организации подобного типа. Данные по Rogozin et al., 2002; б – представление окружения гена в виде сети. Закрашенные окружности показывают гены, которые принадлежат к окружению, автоматически вычлененному с использованием алгоритма, описанного в Rogozin et al., 2002; показана только часть окружения. Незакрашенная окружность соответствует гену, который принадлежит окружению, но не был включен в него автоматической процедурой. Стрелки показывают связь между генами в оперонах (жирные стрелки соответствуют связям внутри окружения, а пунктирные стрелки – внешним связям). Толщина стрелок примерно пропорциональна числу

геномов, в которых представлена данная пара генов.

Большинство оперонов находится не в сложном окружении, включающем разнообразные связи, а представляет собой простую последовательность от двух до четырех генов, порядок которых может различаться. Идентичные или похожие в смысле организации генов опероны часто обнаруживаются в сильно различающихся организмах и в различных функциональных системах. Примечательны в данном случае многочисленные опероны транспорта метаболитов, которые состоят из расположенных в одинаковом порядке генов, кодирующих трансмембранные пермеазы, АТФазы и периплазматические субъединицы так называемых АВС-транспортеров (три субъединицы обозначаются соответственно А, В и С). Присутствие таких общих оперонов в разнообразных бактериях и археях было интерпретировано в рамках гипотезы *эгоистичного оперона* (Lawrence, 1999), которая постулирует, что оперон так хорошо сохраняется не из-за функциональной важности совместной регуляции входящих в него генов, а из-за «эгоистичности» этой компактной генетической единицы, которая склонна к горизонтальному распространению среди прокариот (ниже в этой главе мы еще вернемся к данной концепции при обсуждении горизонтального переноса генов).

Систематическое сравнение расположения ортологичных генов в архейных и бактериальных геномах выявило относительно небольшую долю сохраняющихся (предсказанных) оперонов и гораздо большую распространенность уникальных директонов (последовательностей генов, считываемых в одинаковом направлении и разделенных короткими межгенными участками; Wolf et al., 2001). Как было показано, возможно несколько неожиданно, директоны довольно точно предсказывают опероны: большинство директонов в действительности, по-видимому, являются оперонами (Salgado et al., 2000). Таким образом, архейные и бактериальные геномы сформированы на оперонных принципах с небольшим числом высококонсервативных оперонов и намного более многочисленными редкими и уникальными оперонами. С учетом этого обстоятельства модель консервации оперонов, по крайней мере качественно, напоминает распределение кластеров ортологичных генов, с его трехкомпонентной структурой (см. выше): *редкие гены и редкие опероны гораздо более многочисленны, чем повсеместно распространенные гены и опероны*.

Степень «оперонизации» генома у бактерий и архей широко варьирует: некоторые геномы, например как у гипертермофильной бактерии *Thermotoga maritima*, почти полностью состоят из (предсказанных) оперонов, в то время как другие, как у большинства цианобактерий, по-видимому, содержат очень немного оперонов. Остается неясным, что определяет распространенность оперонов в организме, хотя высказывались предположения, что степень «оперонизации» зависит от баланса между интенсивностью рекомбинации и горизонтального потока генов, а также факторов отбора, препятствующих разрушению оперонов.

Регуляция экспрессии генов и передачи сигналов у бактерий и архей: от базовой схемы оперона к сверхоперонам, регулонам и сложным сетям

Бактерии и археи обладают сложной и элегантной системой регуляции экспрессии генов. Сравнительная геномика драматически изменила существующие взгляды на принципы организации, распределение в природе и эволюцию этих регуляторных механизмов. Концепция оперона Жакоба и Моно, представленная в предыдущем разделе как основной принцип локальной архитектуры бактериальных и архейных геномов, также является концепцией регуляции экспрессии генов и передачи сигналов у прокариот. В модели Жакоба – Моно регулятор (репрессор лактозы в их оригинальной работе) является сенсором внеклеточных и

внутриклеточных сигналов (в данном случае концентрации лактозы), что влияет на структуру белка-регулятора и, опосредованно, на экспрессию оперона (в случае лактозного оперона репрессор, связывая лактозу, отсоединяется от регуляторной части оперона, делая тем самым возможной транскрипцию). В течение полувека, прошедших с момента фундаментального открытия Жакоба – Моно, было обнаружено множество вариаций этой темы, включая регуляторы, которые симметрично влияют на транскрипцию разных расположенных по соседству генов, и глобальные регуляторы, которые контролируют экспрессию многочисленных разрозненных генов и оперонов, в противоположность репрессору простого оперона в модели Жакоба – Моно. Наиболее заметными глобальными регуляторами являются белки – подавители катаболизма (CRP) и регулятор ответа на стресс (SOS) LexA. С учетом открытия этих и других глобальных регуляторов концепция оперона была усовершенствована понятием регулона – набора генов, экспрессия которых регулируется одним и тем же белком-регулятором. Сравнительный геномный анализ регулонов выявил их чрезвычайную эволюционную пластичность с существенными различиями между регулонами даже у близкородственных организмов (Lozada-Chavez et al., 2006). Глобальные регуляторы транскрипции, такие как LexA, широко распространены и высококонсервативны в различных бактериях, но состав генов в регулоне LexA является очень вариативным. Пластичность регулонов, наряду с изменчивостью архитектуры генома (см. выше), хорошо согласуется с идеей, что регуляция экспрессии генов и архитектура генома в эволюции архей и бактерий тесно взаимосвязаны. В резком контрасте с изменчивостью и пластичностью регулонов, регуляторы транскрипции у бактерий и архей демонстрируют примечательное единство архитектуры и структуры. Как правило, эти регуляторы содержат домен, связывающий небольшие молекулы-сенсоры и ДНК-связывающий домен. Подавляющее большинство ДНК-связывающих доменов являются вариациями одной и той же структурной темы, спираль – поворот – спираль. Более специфические, но тоже распространенные домены связывания с ДНК включают мотивы лента – спираль – спираль и цинковая лента (Aravind et al., 2005; Aravind and Koonin, 1999).

Более сложная схема передачи сигналов и регуляции экспрессии генов, которая процессирует сигналы, приходящие из окружающей среды, основана на так называемых двухкомпонентных системах (Casino et al., 2010). Двухкомпонентные системы состоят из мембранных гистидин-киназ и растворимых регуляторов ответа, между которыми сигнал передается путем переноса фосфата. Примечательно, что классические регуляторы транскрипции и гистидин-киназы содержат много общих сенсорных доменов. Это родство указывает на то, что регуляторы транскрипции (однокомпонентные системы) и двухкомпонентные системы образуют единую, интегрированную структуру передачи сигналов и регуляции экспрессии. Однокомпонентные системы, которые распространены практически повсеместно и, как правило, численно доминируют у бактерий и архей, предположительно являются наиболее древними устройствами передачи сигналов, в то время как двухкомпонентные системы, вероятно, являются произошедшей от них более сложной формой передачи сигнала, которая эволюционировала как механизм реагирования на стимулы, приходящие из окружающей среды (Ulrich et al., 2005).

Сравнительная геномика бактерий и архей внесла решающий вклад в открытие новых, до того неизвестных, но в действительности весьма распространенных систем передачи сигналов. В течение многих лет было известно, что широко распространенная форма глобальной регуляции у бактерий использует в качестве посредника цАМФ (циклический АМФ), при участии различных аденилатциклаз (яркий пример неортологичной замены генов), многочисленных белков, содержащих сенсоры цАМФ, такие как GAF-домен, а также белки катаболитной репрессии (CRP и FNR) и другие регуляторы транскрипции, которые тоже

содержат цАМФ-связывающие домены. Сравнительный анализ выявил многочисленные неклассифицированные белки, содержащие гомологичные сенсорные домены, которые типичны для цАМФ-зависимых регуляторов и двухкомпонентных систем, объединенные с одним или двумя новыми доменами, GGDEF и EAL (обозначенными так по соответствующим мотивам консервативных последовательностей аминокислот). Геномный контекст этих доменов и наблюдение, что домен GGDEF является отдаленным гомологом одного из семейств аденилатциклаз, привели к гипотезе, что эти белки являются компонентами новой системы (или систем) передачи сигналов. Впоследствии эти предсказанные системы были открыты после того, как было показано, что домен GGDEF обладает активностью ди-ГМФ-циклазы, в то время как EAL является ди-ГМФ-фосфодиэстеразой. Зависимая от ц-ди-ГМФ передача сигнала, существование которой даже не предполагалось в догеномную эру, начинает рассматриваться как главная регуляторная система бактерий и архей (Seshasayee et al., 2010).

Другая интересная тема дискуссий – широкое представительство у прокариот различных модулей сложных систем передачи сигналов, которые, как считалось ранее, характерны только для эукариот. В частности, сравнительный геномный анализ убедительно показал, что белковые серин-треонин-киназы и соответствующие фосфатазы широко распространены и диверсифицированы среди архей и бактерий и являются важным компонентом многогранной системы передачи сигналов у прокариот. Анализ большего количества бактериальных геномов неожиданно выявил гомологи белков, которые, как считалось ранее, имеются только у эукариот, где они вовлечены в известные пути передачи сигналов, такие как программируемая клеточная смерть (ПКС), или апоптоз. Эти белки включают протеазы из суперсемейства каспаз, семейство апоптозных АТФаз и семейство ГТФаз NACHT; все они вовлечены в различные формы ПКС растений и животных (Koonin and Aravind, 2002; Leipe et al., 2004). Как правило, эти белки обладают сложной мультидоменной модульной архитектурой, для которой характерно соединение каталитических доменов с разнообразными доменами, обеспечивающими специфичность белок-белковых взаимодействий. Эти предполагаемые сигнальные молекулы наиболее распространены в бактериях со сложными фазами развития, таких как цианобактерии, актинобактерии и миксобактерии, а также присутствуют у метаносарцин, единственной известной группы архей с относительно большими геномами и сложной морфологией. Детальное исследование функций этих белков еще предстоит, но есть предварительные признаки того, что у некоторых бактерий они могут быть вовлечены в ПКС (Bidle and Falkowski, 2004). Эти наблюдения показывают, что по крайней мере для некоторых из сложных сигнальных систем эукариот существуют аналоги и вероятные эволюционные предшественники среди бактерий. Мы еще вернемся к этим связям, когда будем обсуждать в главе 7 происхождение эукариот.

Наряду с вышеупомянутой приблизительно квадратичной зависимостью от размера генома сравнительный геномный анализ выявил огромную вариацию в сложности систем передачи сигналов среди бактерий и архей. Эта изменчивость, по-видимому, отражает разнообразие стилей жизни среди соответствующих организмов. Вариации в доле генов, ответственных за передачу сигналов, были количественно отражены в «бактериальном IQ», показателе, который пропорционален квадратному корню от числа белков передачи сигналов (учитывая квадратичное масштабирование) и обратно пропорционален общему количеству генов (Galperin, 2005). IQ отражает способность бактерий и архей отвечать на различные стимулы, приходящие из внешней среды. Соответственно, внутриклеточные симбионты (паразиты) имеют наименьшие значения IQ. Он лишь ненамного выше у организмов с компактными геномами, живущих в стабильной внешней среде, таких как морские цианобактерии, и существенно больше у организмов, живущих в сложной и переменчивой среде, даже у тех, которые обладают

сравнительно небольшими геномами.

Горизонтальный перенос генов – определяющий процесс в эволюции прокариот

Повсеместное распространение ГПГ в мире прокариот

Вездесущность и огромную важность горизонтального переноса генов (ГПГ) в эволюции архей и бактерий можно рассматривать как самую большую новость, выявленную с помощью сравнительного геномного анализа прокариот. Никакое другое открытие не было причиной такого большого количества споров и (порою желчных) дебатов, в которых сталкивались прямо противоположные точки зрения на ГПГ, от утверждений о его повсеместном распространении и всеобъемлющей роли в эволюции бактерий и архей до отрицания любого значимого вклада ГПГ в эволюцию (Gogarten and Townsend, 2005; Kurland et al., 2003; O'Malley and Boucher, 2005). Существование ГПГ, *переноса генов между неродственными организмами иным путем, нежели посредством вертикальной передачи реплицированной хромосомы в процессе деления клетки*, было осознано задолго до того, как был секвенирован первый геном (Syvanen, 1994). Более того, стало понятно, что ГПГ может происходить исключительно быстро и эффективно – во всяком случае, под давлением отбора, как в случае распространения устойчивости к антибиотикам в популяции патогенных бактерий. Однако, до того как появилась возможность сравнения множества полных геномных последовательностей, ГПГ по молчаливому соглашению рассматривался как маргинальный феномен, возможно важный для таких специфических областей, как эволюция сопротивляемости инфекциям, но по большей части не принимавшийся во внимание при изучении эволюции организмов. Как читатель, вероятно, помнит, сама важность вопроса о роли ГПГ в эволюции была осознана в связи с другим революционным открытием: демонстрацией Вёзе и соавторами того, что филогенетический анализ рРНК прокариот реально возможен и может быть потенциально использован для описания эволюции бактерий и недавно открытых архей. Для большинства биологов *трехдоменное эволюционное дерево рРНК, полученное Вёзе, стало синонимом гипотетического древа жизни (ДЖ)*, исходно постулированного Дарвином, а теперь реально полученного и готового для использования в качестве основы для картирования эволюционных событий всевозможного рода (Rase, 2006). Такова была парадигма, когда сравнительная геномика вызвала революцию, связанную с осознанием роли ГПГ.

Исторически и методологически проблема идентификации актов ГПГ и его влияния на эволюцию бактерий и архей резко различается для случаев (сравнительно) недавних и древних переносов, с одной стороны, и переносов между близкородственными и давно разошедшимися организмами, с другой стороны (Koonin et al., 2001a). Недавние случаи ГПГ, особенно между близкородственными организмами, широко распространены, бесспорны и легко обнаруживаются. Действительно, сравнение геномов бактериальных штаммов предоставляет отчетливые свидетельства большого количества актов ГПГ. Вероятно, наиболее характерным примером является открытие так называемых островов патогенности – генных кластеров, которые несут информацию, типичную для патогенов, подобную той, что содержится в генах, кодирующих различные токсины, компоненты секреторной системы третьего типа и другие подобные системы у бактерий-паразитов, а также похожие «симбиотические острова» у бактерий-симбионтов. Острова патогенности представляют собой протяженные районы генома размером до 100 Кб, которые обычно расположены недалеко от генов тРНК и содержат множество генетического материала профагов, откуда напрашивается предположение, что вставка этих островов в геном была осуществлена при посредничестве бактериофагов (Juhas et

al., 2009). Ставший классическим сравнительный геномный анализ энтерогеморрагического штамма O157:H7 и лабораторного штамма K12 бактерии *E. coli* показал, что патогенный штамм содержит 1,387 дополнительного гена, распределенного между несколькими специфичными для штамма кластерами (островами патогенности), сильно различающимися по размерам. Таким образом, до 30 процентов генов у патогенных штаммов, по-видимому, были приобретены посредством недавнего ГПГ (Perna et al., 2001). Последующий детальный анализ индивидуальных линий O157:H7 показал, что процесс ГПГ непрерывно продолжается, внося свой вклад в различие степени вирулентности этих штаммов (Zhang et al., 2007). Воздействие недавнего ГПГ определенно не ограничивается патогенетическими эффектами. Большинство недавних (случившихся, по оценкам, в течение последних 100 миллионов лет) добавлений в метаболическую сеть *E. coli* явно были вызваны ГПГ, часто включающим опероны, кодирующие два и более фермента или белка-переносчика одного и того же метаболического пути. Вклад дубликации генов в метаболические инновации оказался в количественном плане существенно менее важным.

Многочисленные исследования выявили фундаментальный вклад ГПГ в эволюцию конкретных функциональных систем прокариот. Возможно, наиболее впечатляющие результаты были получены для кластера генов фотосинтеза цианобактерий и других фотосинтезирующих бактерий (Raymond et al., 2002). Филогенетический анализ уверенно показывает, что эти кластеры представляют собой сложную мозаику генов, собранных посредством множественных актов ГПГ. Попросту говоря, кислородный фотосинтез, который сформировал нынешнюю атмосферу Земли, вероятно, появился благодаря ГПГ (Mulkiđjanian et al., 2006). К тому же большинство цианофагов несут один или более ген фотосинтеза, предположительно используя их для усиления фотосинтетической активности инфицированных клеток. Таким образом, бактериофаги фактически являются переносчиками генов фотосинтеза при ГПГ (Lindell et al., 2005).

Особенно важным представляется открытие агентов переноса генов (АПГ) в нескольких группах бактерий и архей. АПГ являются дефектными производными от хвостатых бактериофагов, которые упаковывают и переносят случайные фрагменты хромосомы (не являющиеся генами профагов, кодирующими капсид и аппарат упаковки) между прокариотами (Paul, 2008). В прямых экспериментах с сообществами морских микроорганизмов было показано, что АПГ переносят гены необычайно эффективно и малоизбирательно по отношению к реципиенту (McDaniel et al., 2010). Таким образом, как это ни поразительно, АПГ вполне правомочно можно рассматривать как *специализированные средства перемещения генов путем ГПГ, которые, вероятно, вносят важный вклад в потоки генов в мире прокариот*. Мы вернемся к роли вирусов и АПГ в ГПГ и эволюции геномов в целом в главе 10.

Кроме непосредственной экспериментальной проверки и сравнения геномов, недавний ГПГ определяется посредством анализа состава нуклеотидов, частот олигонуклеотидных последовательностей, частоты использования тех или иных кодонов и других «лингвистических» признаков нуклеотидных последовательностей, которые обнаруживают приобретение генов в горизонтальном направлении в виде композиционных аномалий данного генома. Однако перенесенные горизонтально последовательности относительно быстро изменяются, так как приобретенные гены «одомашниваются» в процессе эволюции, так что перенесенные гены в геноме-реципиенте скоро становятся «лингвистически» неразличимы (Ragan, 2001). Важно отметить, что молекулярные механизмы ГПГ между близкородственными организмами хорошо понятны (если не сказать полностью понятны) и включают конъюгацию, перенос бактериофагами (трансдукцию) и трансформацию (Bushman, 2001).

В отличие от четко установленных недавних случаев ГПГ, особенно между компактными

группами родственных организмов, обобщение явления ГПГ на большие эволюционные расстояния, особенно в отдаленном прошлом, их механизмы и влияние на эволюцию архей и бактерий остаются предметом жарких споров (Gogarten and Townsend, 2005; Kurland et al., 2003) [51]. Сравнительная геномика предоставила достаточно свидетельств вероятного ГПГ, включая перенос между очень отдаленными организмами, в частности археями и бактериями. Первым явным свидетельством массивного ГПГ между археями и бактериями было обнаружение того, что гипертермофильные бактерии (конкретно *Aquifex aeolicus* и *Thermotoga maritima*) содержат намного больше гомологов белков, характерных для архей, чем мезофильные бактерии, а также белки, имеющие гомологи как среди архей, так и среди бактерий, но с аминокислотными последовательностями, существенно больше похожими на архейные гомологи, чем на бактериальные (см. рис. 5–4) [52]. Сравнение с мезофильными бактериями показало, что доля «архейных» белков у бактерий-гипертермофилов была намного больше (причем с высокой статистической значимостью), чем у мезофилов (Aravind et al., 1998; Nelson et al., 1999). Впоследствии было продемонстрировано, что мезофильные археи с относительно большим геномом, *Methanosarcina* галобактерии, обладают намного большим количеством «бактериальных» генов, чем термофильные археи с небольшими геномами (см. рис. 5–4; Derrenmeier et al., 2002). Это обстоятельство позволяет грубо оценить долю генов, которые могли быть приобретены археями и бактериями в местах их совместного обитания за счет ГПГ между ними, по крайней мере в 20 процентов.

Несмотря на эти замечательные открытия, ГПГ между эволюционно далекими прокариотами остается предметом дебатов, и все имеющиеся к настоящему времени свидетельства часто (иногда жестко) критикуются (Kurland, 2005; Kurland et al., 2003). Таксономический анализ результатов сравнения последовательностей большого числа геномов дает серьезные аргументы в пользу предполагаемого ГПГ, особенно если учесть, что для прокариот с различными стилями жизни получаются сильно разнящиеся результаты (см. рис. 5–4). И все же эти свидетельства не «доказывают» ГПГ, и в свое время были предложены иные объяснения (не всегда правдоподобные), такие как конвергенция белковых последовательностей удаленных организмов, которые живут в похожих условиях среды обитания (например, гипертермофильные археи и бактерии). Так или иначе, в течение первой декады третьего тысячелетия многочисленные филогеномные исследования – анализ филогенетических деревьев для всех (или почти всех) генов прокариот, которые являются в достаточной степени консервативными и, таким образом, содержат достаточно филогенетической информации для надежных выводов, выявили обширный перенос генов между хорошо известными группами архей и бактерий, включающий даже перенос генов между царствами (Beiko and Hamilton, 2006; Puigbo et al., 2009; Sicheritz-Ponten and Andersson, 2001). Кроме того, эти исследования продемонстрировали существование «скоростных магистралей» (highways), предпочтительных путей для потока генов (Beiko et al., 2005). Широкие магистрали связывают, в частности, различные термофильные организмы (см. также гл. 6).

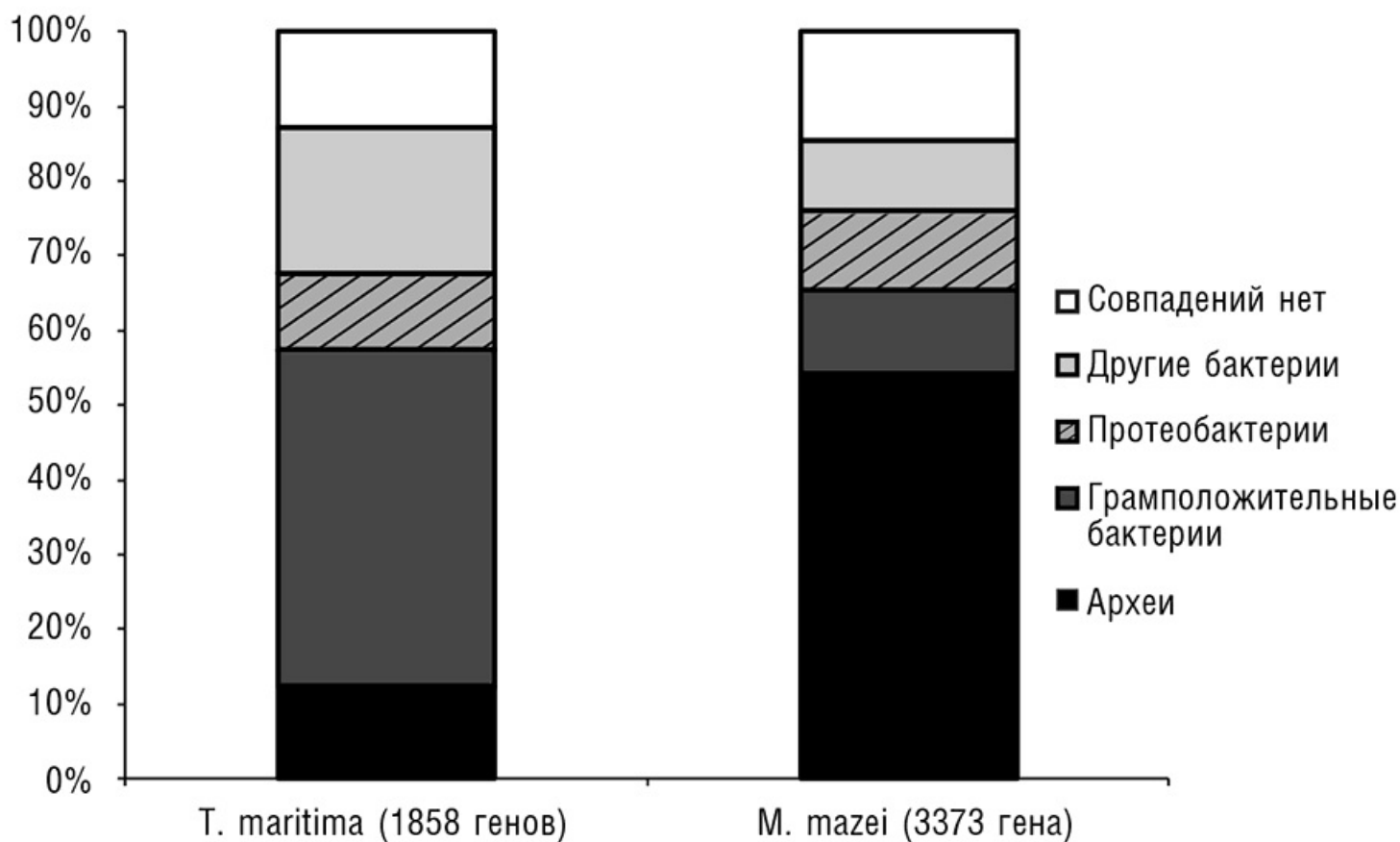


Рис 5–4. Анализ генов в геномах архей и бактерий на предмет таксономической близости наиболее близких гомологов. Показаны данные для гипертермофильной бактерии *Thermotoga maritima* для мезофильной археи *Methanosarcina mazei*. Результаты были получены программой BLASTP путем поиска аминокислотных последовательностей всех белков в каждом геноме с использованием базы данных последовательностей белков NCBI (Altschul et al., 1997).

Ключевая проблема «горизонтальной геномики» – отношение между специфичной для каждого семейства потерей генов и ГПГ. Фундаментальные исследования выявили сложный и нетривиальный характер эволюции прокариот, проявляющийся в «пятнистой» филогенетической структуре, наблюдаемой для многих КОГ (см. рис. 5–5). Такую картину можно объяснить либо ГПГ, либо потерей генов, либо комбинацией этих двух явлений. Простейший (и наиболее, как принято говорить, экономный) эволюционный сценарий можно вычислить, когда известно соотношение скоростей процессов ГПГ и потери генов. Но это отношение (которое, несомненно, различно для разных групп прокариот [\[53\]](#), что будет обсуждаться ниже в этой главе и в гл. 6) является одним из самых больших белых пятен в прокариотической геномике. Известно несколько глобальных реконструкций эволюции прокариот, все из которых основывались на некоторой версии принципа экономии и использовании того или иного сценария с переменным соотношением скоростей приобретения/потери или попытками оценки оптимальной величины этого соотношения (Kunin and Ouzounis, 2003; Mirkin et al., 2003). Результатом этих исследований был вывод о том, что в процессе эволюции прокариот ГПГ может быть почти так же или несколько менее (вероятно, приблизительно вдвое) распространен, как потеря генов. Соответственно, в процессе эволюции большинства КОГ, по всей видимости, произошло по меньшей мере одно событие ГПГ, даже в пределах того ограниченного множества организмов, которые были проанализированы. Конечно, эти исследования проводились на базе сильно упрощенных предположений, таких как

одинаковая для всех организмов частота ГПГ и потери генов внутри групп прокариот; представления, что высокая сложность предковых форм маловероятна (казалось бы, естественная, но, по-видимому, ложная идея; см. гл. 8 об эволюции сложности); и, собственно, концепции дерева видов. Хотя результаты не сильно зависят от топологии дерева видов, базовое представление о дереве с обособленными ветвями, представляющими эволюцию сравниваемых организмов, необходимо при любой реконструкции. В этом заключается фундаментальная проблема, которая достигает буквально философских высот: для того чтобы содержательно обсуждать горизонтальный перенос, *сперва непременно следует определить главное, вертикальное направление эволюции*. Однако, если организмы обмениваются генами с высокой скоростью, в предельном случае совершенно свободно и равномерно, концепция вертикальной эволюции не имеет никакого смысла, так же как и противоположная ей концепция ГПГ. Следовательно, представление эволюции прокариот в виде паутины (сети) кажется логической неизбежностью (см. рис. 5–6). Я должен, однако, оговориться: хотя древовидная компонента эволюции и не обязана существовать, в эволюции прокариот она на самом деле прослеживается. Она схематически изображена на рис. 5–6 и будет главной темой в главе 6.

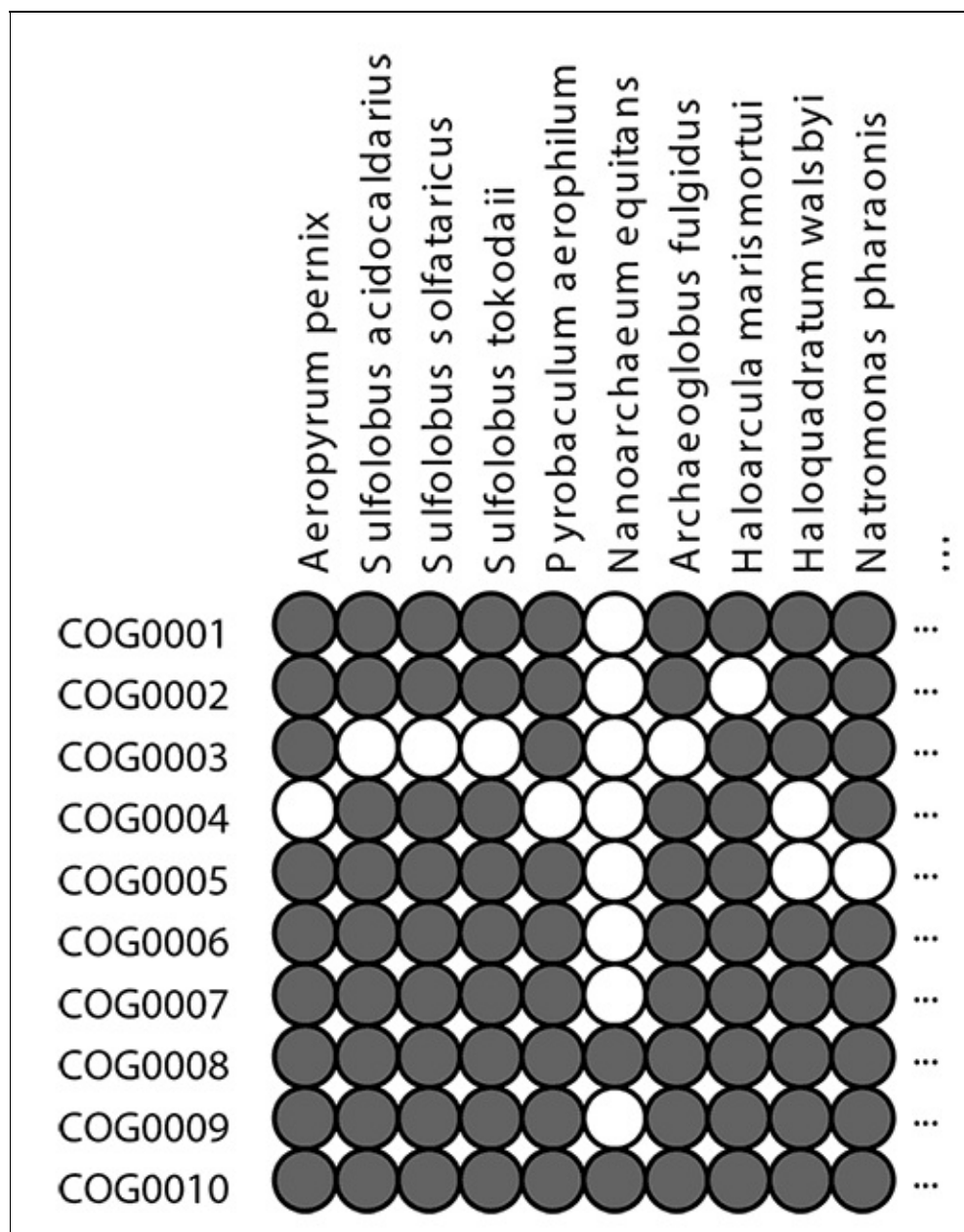


Рис. 5–5. Филетическая карта КОГ. Закрашенные кружки символизируют наличие члена КОГ в геноме, незакрашенные обозначают их отсутствие.

Здесь же мы продолжим обсуждение ГПГ с пониманием того, что древовидная структура действительно наблюдается как важная центральная тенденция эволюции прокариот (см. гл. 6). Широко распространено мнение, что «информационные» гены, кодирующие белки, вовлеченные в процессы трансляции, транскрипции и репликации, намного менее склонны к ГПГ, чем операционные гены, кодирующие ферменты, вовлеченные в метаболизм, транспортные системы и другие «операционные» белки. Обоснование этой точки зрения заключается в так называемой гипотезе сложности (Jain et al., 1999). В соответствии с этой гипотезой, причина низкого темпа ГПГ среди информационных генов заключается в том, что белки, кодируемые этими генами (в отличие от белков, кодируемых большинством операционных генов), обычно являются составными частями сложных белковых машин, которые сильно взаимно адаптированы и, таким образом, не могут быть легко заменены соответствующими ортологами из эволюционно далеких организмов (известными как ксенологи). Однако справедливость гипотезы сложности и ее применимость в общем случае остается неясной, так как и среди информационных генов обнаружено много очевидных случаев ГПГ. Весьма неожиданно, что эти случаи включают не только подавляющее большинство аминоксил-тРНК синтетаз (АРСаз), ферментов, функционирующих в относительной изоляции, но также многие рибосомные белки, являющиеся компонентами молекулярной машины трансляции, рибосомы (Makarova et al., 2001b). Явные свидетельства ГПГ были также обнаружены для таких традиционных маркеров вертикальной филогении, как субъединицы ДНК-зависимой РНК-полимеразы (Iyer et al., 2004a). Разница в режимах эволюции информационных и операционных генов, очевидно, обусловлена как *значительно меньшим распространением замены генов неортологами, так и снижением частоты ГПГ как такового среди информационных генов.*

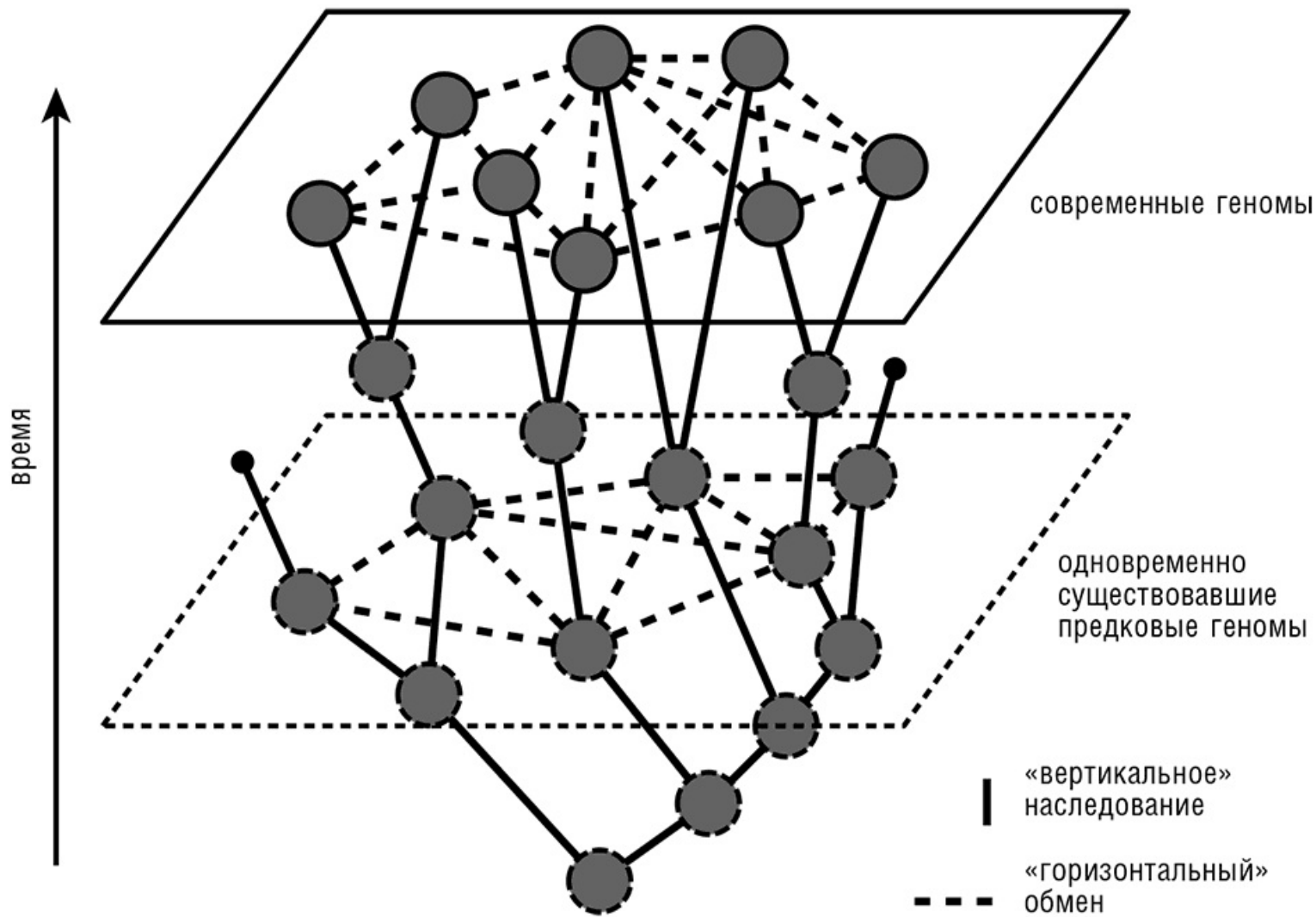


Рис. 5–6. Сетевое представление эволюции прокариот.

В свое время было высказано предположение, что ГПГ между близкородственными организмами (в соответствии с оценкой степени похожести их субъединиц рРНК и других консервативных генов) более широко распространен, чем ГПГ между эволюционно далекими организмами, и такой градиент ГПГ может вносить существенный вклад в наблюдаемую филогенетическую связь между различными группами прокариот (Gogarten et al., 2002). Систематическое изучение способности генов, полученных от других бактерий, комплементировать функции ортологов, показало, что с увеличением степени расхождения последовательностей ортологичных генов комплементация становится менее эффективной (Diaz et al. 2011). Это наблюдение интуитивно привлекательно, так как перенесенный ген должен работать в другой клеточной среде, и, таким образом, можно ожидать, что чем меньше расхождение между набором генов донора и реципиента и структурой генов-ортологов, тем более вероятно, что перенесенный ген успешно приживется и будет зафиксирован отбором. Представляется вероятным, что различия в фиксации перенесенных генов в мире прокариот могут быть фактором, обеспечивающим целостность и стабильность геномов, вопреки интенсивному ГПГ. Мы вернемся к этому вопросу в главе 6 в контексте концепции древа жизни.

В завершение этого краткого обсуждения различных аспектов ГПГ в мире прокариот вернемся к гипотезе эгоистичного оперона, которая постулирует, что «организация бактериальных генов в виде оперонов выгодна для составляющих его генов, так как их близкое расположение облегчает совместный перенос всех генов, необходимых для отбираемого

фенотипа» (Lawrence, 1999). Между функциональностью оперона и эгоистичным характером его эволюции нет противоречия. Оперон является «упакованной» функциональной единицей, часто включающей регулятор собственной активности. В этой роли опероны будут зафиксированы после акта ГПГ с большей вероятностью, чем одиночные гены. Тогда как первоначальная фиксация оперона происходит под влиянием преимуществ совместной регуляции функционально связанных генов, их сохранение и распространение по миру прокариот, осуществляемое при посредничестве ГПГ, происходит в эволюционном режиме, который придает оперонам некоторые (но определенно не все) свойства эгоистичных мобильных элементов. Кроме того, свойство эгоистичности оперонов можно рассматривать как способ преодоления ограничений, налагаемых гипотезой сложности, принимая во внимание, что большинство наиболее распространенных оперонов кодируют субъединицы белковых комплексов (см. предыдущую дискуссию в данной главе). Прекрасной иллюстрацией этой точки зрения является история эволюции мембранных протонных и натриевых АТФ-синтаз, в течение которой оперон, кодирующий множество (до восьми) субъединиц этих сложных молекулярных машин, неоднократно переносился между археями и бактериями (Mulkiđjanian et al., 2008).

Каков, таким образом, главный вывод о распространенности и роли ГПГ в мире прокариот? С моей точки зрения, бесспорно, что *ГПГ является определяющим процессом в эволюции прокариот, влияющим на все аспекты биологии бактерий и архей*. Попытки отрицания ГПГ как маргинального феномена устарели и обречены на неудачу. Метафора паутины эволюции (см. рис. 5–6), по-видимому, становится неотъемлемой частью эволюционной биологии. Однако на количественном уровне вопрос о масштабах ГПГ еще далеко не решен, а также еще далеко не ясно, каковы те главные факторы, которые приводят к горизонтальному переносу индивидуальных генов и оперонов. Эти проблемы являются центральными для нашего понимания эволюции прокариот. В главе 6 мы обратимся к ним прямо.

Мобилом прокариот

Как отмечено в предыдущем разделе, вряд ли любой из членов КОГ принципиально обладает иммунитетом против ГПГ, но некоторые гены в этом отношении «более равны, чем другие». Существенная часть генетического материала прокариот состоит из эгоистичных элементов, для которых горизонтальная мобильность является основным режимом их распространения; совокупность таких элементов в свое время была метко названа мобилом (Frost et al., 2005). Мы также обсуждаем мобилом в контексте мира вирусов (см. гл. 10), но, чтобы описать в общих чертах формирующееся логически последовательное представление об эволюции прокариот, мы должны здесь вкратце рассмотреть наиболее яркие свойства генетических элементов этого класса. Мобилом состоит из бактериофагов, плазмид, транспозонов и генов, которые часто ассоциируются с ними и регулярно становятся «пассажирами», в числе прочих, системы рестрикции-модификации (ОМ) и токсины-антитоксины (ТА). Это представляется вполне естественным, так как вирусы и плазмиды мобильны по определению, так же как и системы защиты. Мобилом неразрывно связан с «главными» хромосомами прокариот. Вирусы (бактериофаги) и многие плазмиды систематически интегрируются в хромосомы, либо обратимо, в этом случае они часто мобилируют хромосомные гены, либо необратимо, когда мобильные элементы «одомашниваются», превращаясь в «обычные» гены, изначально ОРС (см. выше раздел «Фрактальное пространство-время генома, пангеномы и кластеризация прокариот»). Со времен классического эксперимента Жакоба и Волмана в 1950-х было хорошо известно, что конъюгативные плазмиды могут быть посредниками при перемещении больших фрагментов бактериальных хромосом, а вирусы (бактериофаги) давно известны как посредники при переносе генов путем трансдукции (Bushman, 2001). Открытие АПГ, которые, по-видимому, являются специализированными векторами для ГПГ, еще более убедительно подтверждает существование механизмов, регулирующих обмен генетическим материалом между мобилом и хромосомами (Paul, 2008).

Перенос плазмидами устойчивости к антибиотикам и путей вторичного метаболизма является хрестоматийным примером динамики бактериальной мобиломы, но роль плазмид далеко не исчерпывается такими относительно узкими областями биологии. На самом деле граница между хромосомами и плазмидами весьма размыта. Плазмиды являются репликационными (обычно циркулярными, но иногда и линейными), которые, подобно хромосомам прокариот, имеют точку начала репликации и кодируют по крайней мере некоторые из белков, участвующих в собственной репликации и сегрегации (распределении реплицированных плазмид между дочерними клетками в процессе деления). Ключевые белки, вовлеченные в процессы сегрегации плазмид и хромосом, в частности АТФазы семейств FtsK/HerA, гомологичны во всем мире прокариот (а также обнаружены у многих вирусов – см. гл. 10), факт, который подчеркивает общность эволюционного происхождения и стратегий различных прокариотических репликационных элементов (Iyer et al., 2004b; McGeoch and Bell, 2008).

«Канонические» геномы многих бактерий и архей включают, в дополнение к «главной» хромосоме (или хромосомам), один или больше неотъемлемых, относительно больших и стабильных внехромосомных элементов, часто называемых мегаплазмидами. Мегаплазмиды могут удивительно долго сохраняться в процессе эволюции. Например, единственная мегаплазмида бактерии *Thermus thermophilus* гомологична одному или двум мегаплазмидам бактерии *Deinococcus radiodurans*, следовательно, произошла от общего предка этих родственных, но сильно разошедшихся бактерий (Omelchenko et al., 2005). Однако в процессе

эволюции этой древней группы бактерий мегаплазмиды накопили (относительно своих размеров) намного больше отличий в составе генов, чем хромосомы. Кроме того, мегаплазмиды имеют много горизонтально перенесенных генов, включая гены термофильных организмов, которые, очевидно, были приобретены линией *Thermus* важны для их термофильного стиля жизни. Таким образом, хотя мегаплазмиды могут существовать в геномах прокариот в течение продолжительного по эволюционным меркам времени, они демонстрируют большую эволюционную пластичность, чем хромосомы, и служат своеобразными «запасниками» для ГПГ.

Почти все секвенированные геномы прокариот содержат следы интеграции многочисленных плазмид и фагов. В частности, примечательно, что большинство геномов архей обладают несколькими версиями оперона *herA-nurA*, который кодирует ключевые компоненты (АТФазы и нуклеазы) механизма сегрегации плазмид после репликации. Каждый из этих оперонов является остатком отдельного репликона, таким образом, *слияние репликонов*, по-видимому, является у прокариот достаточно частым событием. В ходе эволюции такая интеграция могла быть важным фактором, который сформировал наблюдаемую архитектуру хромосом прокариот (Iyer et al., 2004b; McGeoch and Bell, 2008).

Системы защиты и ответа на стресс, в частности, рестрикции-модификации и системы токсин-антитоксин можно рассматривать в качестве особых частей мобилома. Сравнительный анализ этих систем выявил быструю эволюцию и частый ГПГ; кроме того, они часто обнаруживаются в геномах плазмид и бактериофагов. Несмотря на их громадное молекулярное разнообразие, эти системы работают на одном и том же принципе – каждая содержит токсин, то есть белок, который либо разрушает хромосомную ДНК (ферменты рестрикции), либо блокирует систему трансляции в результате гидролиза мРНК или тРНК (соответствующие токсины обладают эндонуклеазной активностью, специфичной к РНК), либо убивает клетку, вызывая образование отверстий в мембране (Kobayashi, 2001; Van Melderen and Saavedra De Bast, 2009). Вызванная токсинами гибель клетки в случае использования систем рестрикции-модификации предотвращается специфическим метилированием ДНК, а в случае систем токсин-антитоксин путем нейтрализации токсина антитоксином. Эта нейтрализация достигается либо посредством взаимодействия двух белков (токсина и антитоксина), либо посредством подавления трансляции мРНК токсина комплементарной к ней малой РНК, которая в этом случае выполняет роль антитоксина. Эти системы обладают свойствами эгоистичных элементов, которые эволюционировали таким образом, чтобы сделать клетки хозяина зависимыми от них (addicted). Когда клетка теряет соответствующие гены, она обычно погибает либо в результате действия токсина, поскольку токсины более стабильны, чем антитоксины, так что их активность после деградации антитоксинов, запасы которых перестают восполняться, резко возрастает, либо из-за ослабления активности ферментов модификации в результате снижения их концентрации, в то время как рестриктазы сохраняют активность. Из-за этого свойства токсин-антитоксинных систем плазмиды, которые переносят токсин-антитоксинные гены и, таким образом, гарантируют, что их хозяева «подсядут» на них, убивая клетки, которые утратили плазмиды, получают сильное селективное преимущество перед плазмидами, не имеющими токсин-антитоксинных систем. Известные к настоящему времени токсин-антитоксинные системы, вероятно, представляют, как принято говорить, лишь вершину айсберга, так как геномы бактерий и архей несут большое количество различных оперонов, имеющих свойства, сближающие их с токсин-антитоксинными оперонами (а именно пара генов, кодирующих небольшие белки, которая появляется в одинаковой комбинации в различных геномах и геномных окружениях), но которые не были (пока) экспериментально исследованы (Makarova et al., 2009a).

Недавно был открыт еще один очень необычный класс систем мобильной защиты, который присутствует у большинства архей и примерно одной трети бактерий из числа секвенированных геномов (Deveau et al., 2010; Koonin and Makarova, 2009). Эта система состоит из массива *коротких палиндромных повторов, регулярно расположенных группами* (CRISPR), и включает в себя около 50 отдельных семейств генов (ассоциированных с CRISPR и обозначаемых *cas* (*CRISPR-associated*)). Примечательно, что здесь мы сталкиваемся со вторым по размерам (после рибосомного супероперона) массивом связанных между собой генов (генетическим окружением) в геномах прокариот (Rogozin et al., 2002). Система CRISPR-Cas защищает клетки прокариот от фагов и плазмид «ламарковским путем» (мы вернемся к этому вопросу более подробно в гл. 9): фрагмент гена фага или плаزمиды интегрируется в локус CRISPR на бактериальной хромосоме и впоследствии транскрибируется и используется посредством все еще плохо изученных механизмов для подавления репликации эгоистичных агентов. Система CRISPR-Cas демонстрирует выдающуюся пластичность даже между близкородственными штаммами бактерий и архей, а также, по-видимому, часто переносится путем ГПГ.

Избранные примеры, обсуждавшиеся выше, указывают на огромное, все еще не до конца понятое разнообразие мобилома прокариот и подчеркивают значительный вклад, который мобилом вносит в эволюцию геномного пространства-времени прокариот.

Незаменимость ГПГ для эволюции прокариот

Вероятно, еще не все биологи осознают тот факт, что ГПГ является принципиально важным фактором эволюции прокариот и может, по-видимому, рассматриваться как необходимое условие долгосрочного выживания бактерий и архей. Любая популяция, у которой отсутствует рекомбинация генетического материала, в конечном счете имеет тенденцию к вымиранию, так как она не обладает эффективными средствами для устранения неизбежно накапливающихся вредных мутаций. Обычно преимущество популяций, у которых есть механизм полового размножения или его аналог, перед бесполоыми приписывается механизму, известному как храповик Мёллера [Möller's ratchet (Möller, 1964)]. Под действием храповика Мёллера накопление вредных мутаций в условиях отсутствия рекомбинации (одной из форм которой является половое размножение) приводит к постепенной потере приспособленности и гибели бесполой популяции. Эффект храповика Мёллера наиболее сильно проявляется в случае популяций небольшого размера из-за большой роли генетического дрейфа. Майкл Линч с сотрудниками разработали более подробную модель угасания бесполой популяции, известную как мутационная катастрофа (Lynch et al., 1993). С учетом того, что большинство мутаций являются (по меньшей мере) слабовредными, бесполоя популяция входит в «нисходящую спираль» мутационной катастрофы, когда храповик Мёллера действует совместно с генетическим дрейфом. В этом случае размер популяции падает из-за подавления отбора, очищающего от вредных мутаций, что, в свою очередь, в итоге приводит к усилению генетического дрейфа и увеличению вероятности случайной фиксации дополнительных вредных мутаций. Таким образом, мутационная катастрофа, по-видимому, устанавливает пределы для размера генома и продолжительности существования популяций бесполой организмов.

Большинство прокариот не вовлечены в обычные половые отношения, хотя механизм, известный как «бактериальный секс», конъюгация, досконально описан. Однако конъюгация требует наличия специальной плазмиды (описанной в классических ранних экспериментах Вольмана – Жакоба, Ледерберга и Кавалли – Сфорца как F-фактор) или присутствия в хромосоме так называемых элементов интеграции и конъюгации, которые имеются далеко не у всех прокариот (Bushman, 2001). Среди хорошо изученных в настоящее время бактерий конъюгация известна лишь у меньшинства, а у архей она и вовсе не обнаружена (Frost and Koraimann, 2010; Wozniak and Waldor, 2010). Бактерии, у которых регулярно происходит конъюгация, часто формируют большие популяции и, возможно, виды, похожие на классические виды эукариот. В этих случаях половой процесс избавляет бактерии от мутационной катастрофы. Однако, если конъюгации отсутствуют или очень редки, что, по-видимому, характерно для архей и многих бактерий, единственным способом избежать катастрофических последствий является ГПГ, который в этом случае можно рассматривать как одну из форм «незаконной» (негомологичной) рекомбинации. На больших промежутках времени бесполое популяции (прокариот) могут выжить *только в том случае, если они с достаточно высокой частотой получают посредством ГПГ функциональные версии генов, накапливающих вредные мутации* (см. рис. 5–7) ^[54]. Рассмотрение роли ГПГ в эволюции прокариот с популяционно-генетической точки зрения неизбежно ведет к предположению, что *отбор действует в направлении поддержания оптимального уровня ГПГ*. Эта оптимальная частота ГПГ достаточно высока для предотвращения мутационной катастрофы и обеспечения возможностей для адаптивных инноваций, но в то же время достаточно низка, чтобы избежать частого разрушения функционально важных связей между генами (оперонов). Очевидное

предсказание гипотезы оптимизации ГПГ заключается в том, что функционально важные гены, которые быстро эволюционируют и часто утрачиваются в процессе эволюции, должны чаще подвергаться ГПГ. В главе 6 мы увидим, что это предсказание действительно подтверждается сравнительным анализом филогенетических деревьев генов прокариот. Эти наблюдения позволяют нам рационально объяснить эволюцию АПГ как специализированных посредников ГПГ, которые сохраняют темп переноса генов выше того порога, ниже которого наступает катастрофа. Кроме того, ДНК-помпы, участвующие в трансформации (Chen et al., 2005), также можно рассматривать скорее как устройства для обеспечения ГПГ, чем как простые механизмы для приобретения нуклеотидов (в форме ДНК, поглощаемой из окружающей среды), как иногда предполагается.

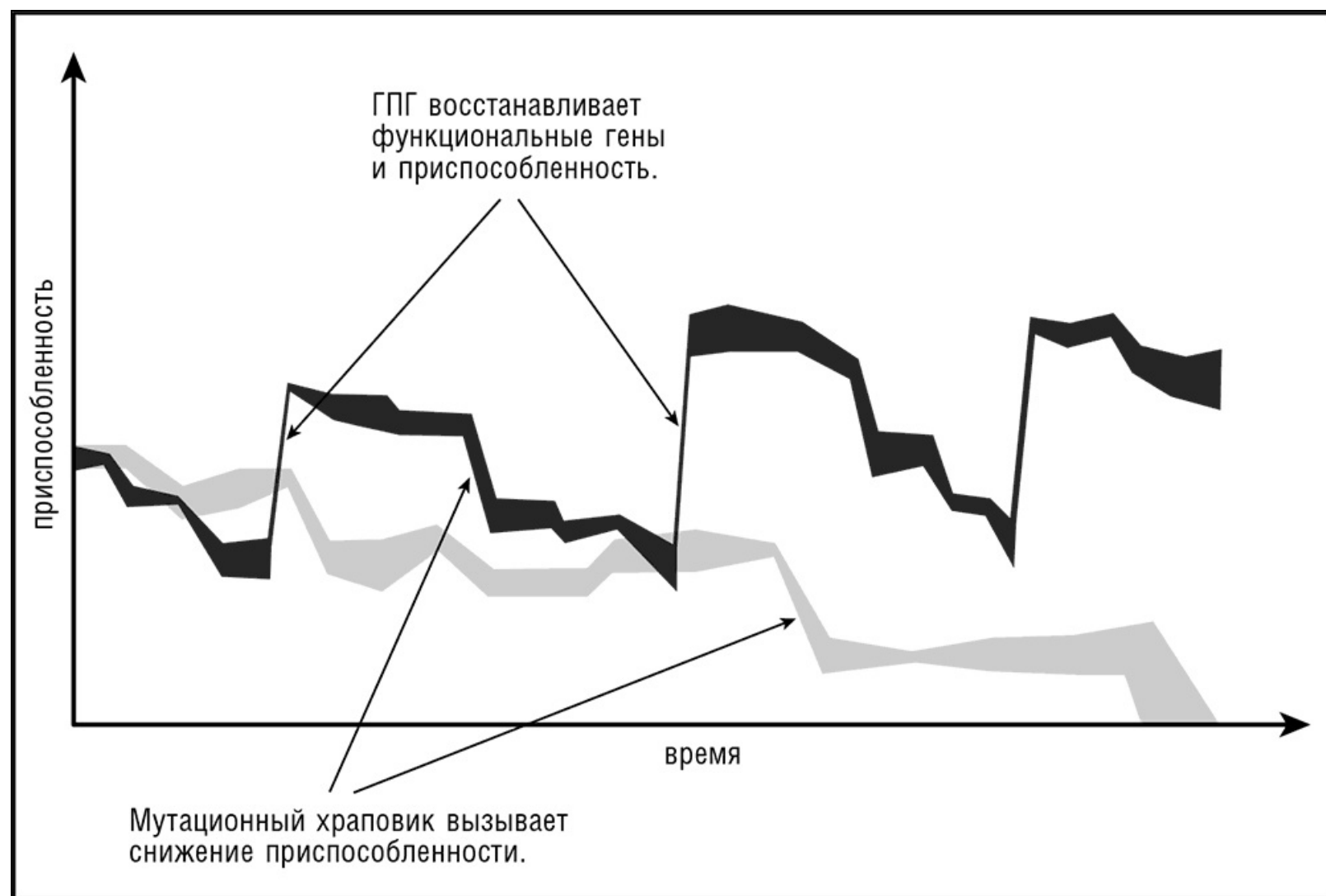


Рис. 5–7. Неизбежность ГПГ: судьба бесполой популяции в случае изоляции и при наличии ГПГ.

Любая бесполой популяция, которая (полностью или почти полностью) изолирована от ГПГ, движется в направлении угасания и последующего вымирания. И это действительно наблюдается у облигатных бактериальных паразитов, особенно тех, что обитают внутри клеток. Внутриклеточные симбионты с самыми маленькими геномами, такие как вышеупомянутая *Hodgkinia cicadicola* или обладающая геномом чуть большего размера *Carsonella ruddii*, приближаются к статусу органеллы эукариотической клетки своего хозяина или даже уже достигли его (McCutcheon et al., 2009) и почти наверняка уже зашли достаточно далеко на пути в вымиранию. Как это часто бывает, при этом существует конкуренция между глобальным

давлением со стороны динамики популяции в целом и локальными адаптациями (шире эта тема будет развернута в гл. 8). Некоторые эндосимбионты насекомых с маленькими, но не самыми крошечными геномами (как правило, около 500 генов), такие как *Wolbachia* или *Wigglesworthia*, сохранили определенные метаболические пути, которые снабжают хозяина необходимыми ему метаболитами, в частности аминокислотами (Wu et al., 2006). Эта адаптация может позволить данным организмам поддерживать относительно большой эффективный размер популяции и, следовательно, по крайней мере временно, избежать гибели. Однако в конечном итоге все же представляется правдоподобным, что такие бактерии имеют относительно короткую (в эволюционном масштабе) продолжительность жизни.

Горизонтальный перенос генов, универсальные законы геномики и хорошо перемешанный резервуар прокариотических генов

В предыдущей главе мы обсудили несколько универсальных зависимостей между геномными переменными (законы геномики), в частности обратно пропорциональную зависимость между численностью функциональных классов генов и размером семейств паралогичных генов данного класса. Теперь мы не можем избежать вопроса, как относятся между собой законы геномики и ГПГ, который является столь заметным процессом в мире прокариот. Действительно, сравнительный геномный анализ показывает, что генные семейства прокариот формируются в большей мере за счет ГПГ, чем за счет дупликации генов (Treangen and Rocha, 2011). Большинство генов, которые выделяются как паралогичные при анализе одиночного генома, в действительности являются псевдопаралогами (Makarova et al., 2005). Вне зависимости от пути происхождения распределение размера семейств с высокой точностью воспроизводится моделями рождения, смерти и инноваций (см. рис. 4–7). Единственное объяснение этого соответствия, по всей видимости, заключается в том, что темп рождения и смерти генов в действительности пропорционален не размеру семейства в данном геноме, а размеру семейства в резервуаре генов-доноров. Поскольку степенные распределения размера семейств очень близки для всех геномов, резервуар доноров в действительности означает всю вселенную прокариотических геномов. Другими словами, этот аспект структуры вселенной геномов может быть описан в рамках универсального степенного закона распределения размеров семейств генов (очевидно, это распределение весьма отличается от структуры ядра – оболочки – облака, так как оно относится к обширным семействам паралогичных генов, а не строго определенным группам ортологов).

Та же логика применима и к масштабированию функциональных классов генов (то есть зависимости численности классов от общего числа генов, см. гл. 4). Учитывая, что ГПГ вносит основной вклад в генетический состав прокариотических геномов, (почти) универсальные законы масштабирования требуют, чтобы вселенная геномов прокариот рассматривалась как единый резервуар генов. Теорию эволюции, отражающую центральную роль ГПГ, еще только предстоит разработать в явном виде. Однако универсальность законов масштабирования и изложенные выше качественные соображения заставляют предполагать, что в среднем *геномная вселенная прокариот является хорошо перемешанным резервуаром генов*. Конечно, существуют существенные локальные неоднородности и «скоростные магистрали» ГПГ (см. гл. 6), но в среднем темп перемешивания генов достаточно высок для того, чтобы обеспечить законы универсального масштабирования.

Характерные геномные профили бактерий и архей с различными стилями жизни и неизоморфное отображение геномного и функционального пространств

Одной из самых больших надежд, связанных со сравнительной геномикой, является возможность, по крайней мере в принципе, определить «геномные профили» организмов с резко отличающимися стилями жизни (фенотипами), то есть наборы генов, которые необходимы и достаточны для поддержания этих стилей жизни. В имеющейся в настоящее время и быстро растущей коллекции геномов прокариот различные стили жизни часто представлены множественными, сильно отличающимися геномами – таким образом похоже, что подошло время всерьез начать изучение геномно-фенотипических связей. До сих пор в этом направлении можно констатировать лишь весьма скромный успех. Когда стиль жизни связан с хорошо изученным биохимическим путем, таким как у метаногенов или у фотосинтезирующих организмов, идентификация геномного профиля может быть относительно простой задачей. Но даже в этом случае анализ генов, кодирующих белки, вовлеченные, например, в фотосинтез, демонстрирует сложное переплетение признаков, специфичных для стиля жизни и для конкретного таксона. Наиболее полный набор генов для фотосинтеза был зафиксирован у цианобактерий, в то время как другие группы фотосинтезирующих бактерий обладают различными подмножествами этого набора генов (Mulkidjanian et al., 2006).

Геномные профили более сложных фенотипов, таких как термофилы или устойчивость к радиации, оказались намного более неуловимыми. Возможно, наиболее настойчивые попытки были посвящены поиску сигналов термофильной адаптации. Примечательно, что был обнаружен один ген, присутствующий во всех секвенированных геномах гипертермофилов, но полностью отсутствующий у мезофилов, и продукт этого гена – белок, абсолютно необходимый для репликации ДНК при экстремально высоких температурах, обратная гираза (Forterre, 2002). Кроме того, геном умеренного термофила *Thermus thermophilus* (штамм HB27) содержит псевдоген обратной гиразы, в то время как близкородственный ему штамм HB8 содержит работоспособный ген обратной гиразы, что демонстрирует процесс избавления от обратной гиразы после вероятного перехода от гипертермофильного к умеренно термофильному стилю жизни (Omelchenko et al., 2005). Однако поиск других генов, специфичных для термофилов, принес ограниченные результаты. Иных генов, кроме обратной гиразы, показывающих явную корреляцию их присутствия-отсутствия с (гипер) термофильностью, обнаружено не было; для гипертермофильных архей и бактерий нашлось лишь несколько генов, которые существенно чаще присутствовали в их геномах по сравнению с геномами мезофилов (Makarova et al., 2003). Кроме того, было предпринято много попыток определить характерные признаки фенотипа термофилов на уровне нуклеотидных и белковых последовательностей и структур. Хотя эти исследования выявили несколько признаков, потенциально характерных для белков термофилов, таких как высокая плотность заряда и большее, чем обычно, представительство дисульфидных мостиков, реальная значимость каждого из этих признаков остается неопределенной (Veeby et al., 2005). На филогенетических деревьях высококонсервативных генов (см. гл. 6) термофилы часто группируются в одном кластере с мезофилами – так, белки бактерии *Thermus* группируются вместе с их гомологами из мезофильной бактерии *Deinococcus* (вспомните, например, известную полимеразу Taq, незаменимое орудие геномной инженерии). Эти факты показывают, что общая эволюционная история оказывает намного более сильное влияние на белковые последовательности, чем термофильные (и другие) адаптации.

Итоговый вывод подобных исследований заключается в том, что сравнительная геномика до сих пор не в состоянии выявить «секреты» (гипер)термофильного стиля жизни. (Интуитивно можно было бы подозревать, что должна быть существенная разница между геномами, кодирующими организмы с оптимальной температурой роста, превышающей 95 °С, и теми, для которых она составляет 37 °С.)

История поиска особенностей геномов, коррелирующих с экстремальной устойчивостью к радиации и высыханию, может быть даже более показательна. Некоторые бактерии и археи, лучше всего из которых описана бактерия *Deinococcus radiodurans*, демонстрируют сопротивляемость экстремальному уровню радиации, что, как считается, является побочным эффектом их адаптации к высыханию. Всесторонний анализ генома *D. radiodurans* напрямую не выявил никаких уникальных признаков генома или систем репарации ДНК, которые могли бы объяснить исключительную способность этих организмов переносить повреждения, вызываемые радиацией, хотя были идентифицированы гомологи белков растений, используемые для повышения сопротивляемости высыханию и в то же время не обнаруженные ни у каких других бактерий (Cox and Battista, 2005; Makarova et al., 2001a). *Deinococcus radiodurans* является популярной экспериментальной моделью, поэтому для описания реакции этой бактерии на высокие дозы облучения впоследствии были предприняты исследования процессов транскрипции и особенностей состава ее белков. Эти исследования возбудили определенный интерес, так как было зафиксировано существенное увеличение экспрессии некоторых экспериментально не исследованных генов, кодирующих белки с предсказанной ролью в процессах, потенциально связанных с радиационной устойчивостью, таких как репарация двойных разрывов в молекулах ДНК (Liu et al., 2003). Однако нокаут этих генов не повлиял на сопротивляемость радиации, в то время как нокаут нескольких других генов, которые не кодируют никаких известных доменов и не экспрессируются на повышенном уровне при облучении, делал организм чувствительным к радиации (Blasius et al., 2008; Cox and Battista, 2005; Makarova et al., 2007a). Сравнительный анализ двух родственных устойчивых к радиации бактерий *D. radiodurans* и *D. geothermalis* не только не смог разрешить проблему геномных особенностей, определяющих устойчивость к радиации, но даже еще более запутал ее (Makarova et al., 2007a). Никаких генов, имеющих явное отношение к сопротивляемости радиации, которые бы были уникальными для этих устойчивых бактерий, открыто не было. Более того, ортологи многих генов *D. radiodurans*, экспрессия которых усиливается в условиях повышенной радиации, попросту отсутствуют у *D. geothermalis*. Тщательное сравнение структуры оперонов и предполагаемых регуляторных областей в двух геномах *Deinococcus* позволило предсказать регулон, ответственный за устойчивость к радиации (Makarova et al., 2007a). Однако роль большинства генов, составляющих этот предсказанный регулон, в обеспечении устойчивости к радиации и высыханию остается невыясненной. Главные свойства генома, определяющие устойчивость к радиации, по-прежнему неуловимы, и растет количество свидетельств, демонстрирующих, что важная роль при этом принадлежит генам, которые повышают устойчивость неожиданными косвенными способами, такими как регулирование внутриклеточной концентрации двухвалентных катионов, оказывающих влияние на степень разрушения белков, вызванного облучением или высыханием (Daly, 2009).

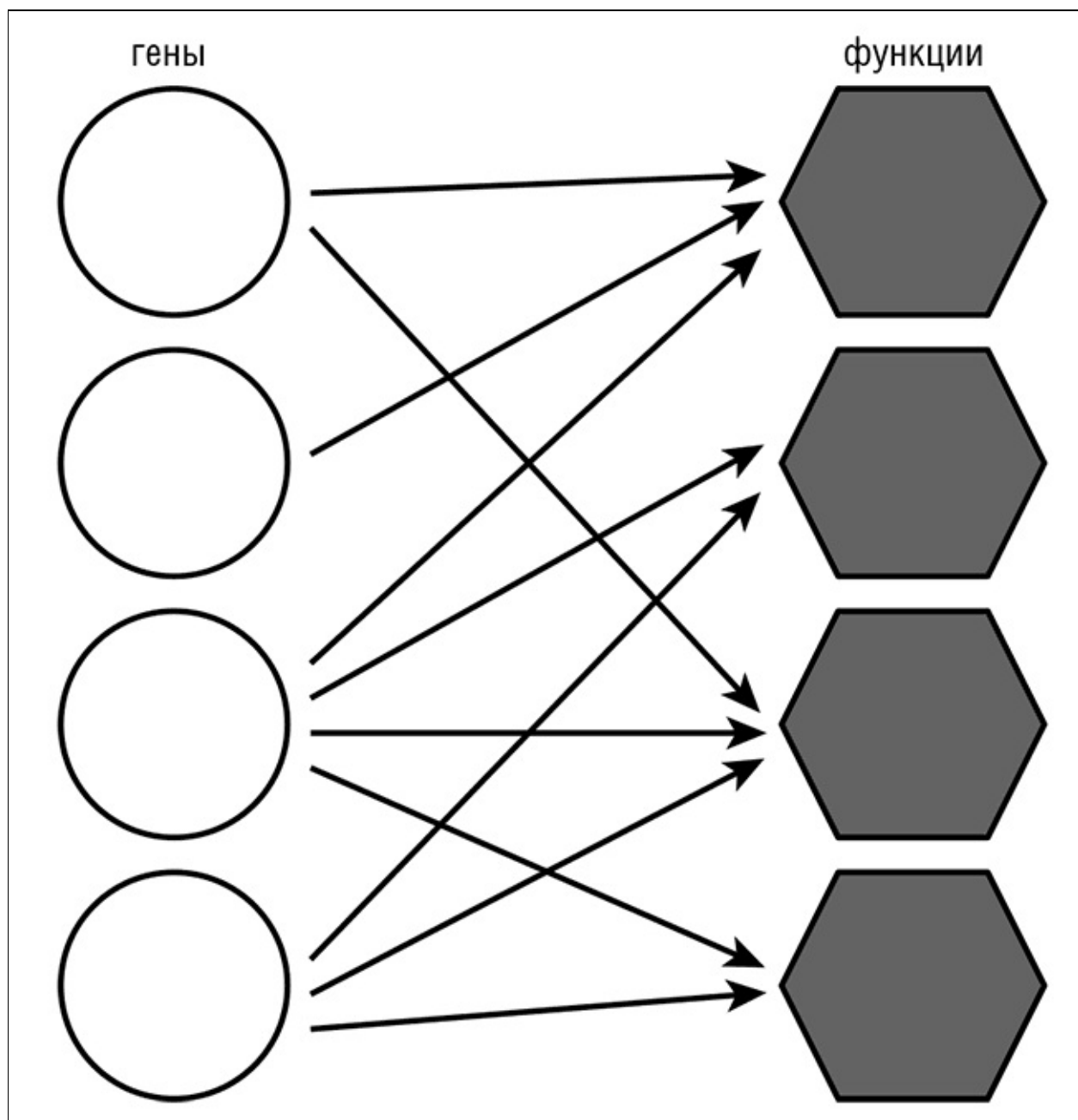


Рис. 5–8. Неизоморфное, многозначное отображение геномного и функционального пространств.

Единственный вывод, который можно сделать в рамках текущего состояния дел по поводу связей между геномом и фенотипом у прокариот, заключается в том, что эти связи являются многогранными, и конкретные наборы генов, ответственных за формирование сложных фенотипов, идентифицировать непросто, несмотря на существование явных генов-сигнатур, связанных с определенными стилями жизни (такими как обратная гираза в случае гипертермофильности). Сложность соотношения геном – фенотип может быть представлена в виде *неизоморфного многозначного отображения между геномным и функциональным пространствами прокариот* (Koonin and Wolf, 2008b). Каждый ген плеiotропен (связан с множеством функций), и каждая функция мультигенна (связана со многими генами; см. рис. 5–8). Мы пришли к этому важному выводу при анализе геномов прокариот, но, без сомнения, он отражает общее правило отсутствия детерминизма при отображении генотип – фенотип (см. гл. 13).

Археи и бактерии в свете сравнительной геномики: как же быть с прокариотами?

Сам термин и концепция прокариот в последнее время ставятся под сомнение как устаревшие и основанные на негативном определении: отсутствии органелл, давших свое имя «высшим» организмам (эукариотам), – ядер (Rase, 2009b, 2006). Вместо представляющегося неадекватным понятия прокариот было решительно предложено классифицировать формы жизни исключительно на базе филогенетического разделения, которое исходно возникло из деревьев рРНК и было подтверждено деревьями нескольких других (почти) универсальных информационных генов (Rase, 2009a). Что касается негативного определения прокариот, то ему было противопоставлено определение, основанное на позитивных признаках, таких как сопряжение транскрипции и трансляции (Martin and Koonin, 2006b) ^[55]. Каковы бы ни были относительные достоинства этих аргументов, сравнительная геномика пролила свой собственный свет на проблему прокариот. Как обсуждалось в этой главе, между археями и бактериями наблюдается очень небольшая степень сохранения репертуара генов (КОГ) и даже еще меньшая степень сохранения в организации специфических оперонов. Действительно, на эволюционных деревьях, построенных на основе сравнения репертуара генов и консервативных пар соседних генов, разделение между бактериями и археями весьма четкое (Wolf et al., 2002).

С этим разделением резко контрастирует единство общей организации геномов бактерий и архей. Несмотря на некоторые исключения, главный принцип организации генома может быть легко выражен следующим сжатым описанием: *бактерии и археи имеют компактные геномы с короткими межгенными промежутками, устроенные таким образом, что множество генов формируют директоны, которые имеют тенденцию функционировать как опероны*. Формирование директонов, многие из которых становятся оперонами, можно рассматривать как непосредственное следствие компрессии генома (дальнейшее обсуждение этой проблемы в гл. 8). Эволюционная устойчивость оперонов затем обеспечивается совместным действием очищающего отбора и частого ГПГ, в соответствии с концепцией эгоистичного оперона. Одинаковый принцип организации геномов бактерий и архей появился как непосредственное следствие давления отбора, которое действует в процессе эволюции этих форм жизни. Эти факторы отбора, в свою очередь, сами калибруются популяционной динамикой (см. гл. 8). Рассматривая все это вместе, я не могу не прийти к концепции *прокариот как формы жизни, характеризующейся особым режимом эволюции, включающим всесторонний и неотъемлемый ГПГ, который создает хорошо перемешанный резервуар генов и приводит к общему типу организации их генома*. Является ли термин «прокариоты» удачным для описания этой части биосферы, остается спорным вопросом (мы внесем некую ясность в эту проблему при обсуждении происхождения эукариот в гл. 7), но, вероятно, все же второстепенным.

Краткий обзор и перспектива

Безусловно, прогресс в наших знаниях о мире прокариот, достигнутый благодаря успехам сравнительной геномики, огромен. Многие из основных закономерностей и моделей, обсуждавшихся здесь, такие как явные отличия архей и бактерий, наряду с фундаментальным сходством в режимах их эволюции и вытекающей из этого организацией генома, оперонная организация генов бактерий и архей и существование ГПГ, были замечены еще в предгеномную эру, но скорее как отдельные случаи, чем как общие тенденции. Сравнительная геномика позволяет исследователю определить, насколько общим является любая наблюдаемая тенденция, и достоверность подобных выводов увеличивается по мере роста размера коллекции секвенированных геномов. В начале геномной эры существовала надежда на открытие нового набора законов геномики. И действительно, при сравнении геномов прокариот были открыты некоторые почти универсальные количественные закономерности. Наилучшими кандидатами на статус законов геномики можно считать специфическую зависимость доли генов разных функциональных классов от размера генома, степенной закон распределения размеров семейств генов и универсальное распределение скоростей эволюции ортологичных генов (см. гл. 4).

Однако в общем и целом по итогам 15 лет работы в области сравнительной геномики, вероятно, правильнее говорить о закономерностях, ограничениях и, возможно, принципах, но не о законах, высеченных на камне. Действительно, в терминах общей организации, основная масса геномов архей и бактерий примечательно похожа и построена в соответствии с одним и тем же простым «генеральным планом» с плотным расположением кодирующих белки и РНК генов, организованных по большей части в виде директонов, как правило, с одной точкой начала репликации. Большинство генов архей и бактерий имеют простую организацию с непрерывной кодирующей последовательностью и короткими регуляторными участками. По-видимому, существует нетривиальная связь между функциями генов и сложностью генома: распределение числа генов, принадлежащих различным функциональным классам, оказывается (почти) одинаковым для широкого спектра доступных геномов, с почти постоянным «замороженным» набором генов, участвующих в процессе трансляции, и резким ростом числа генов-регуляторов и сигнальных белков при увеличении размера генома. Увеличение «бюрократического груза» вместе с энергетическими ограничениями, вероятно, являются важными факторами, устанавливающими верхний предел на размер генома прокариот и, соответственно, сложность их строения. Эти закономерности приближаются к законам геномики настолько близко, насколько это только можно представить, хотя, как всегда это бывает в биологии, существует множество исключений из любого правила.

Важнее, что в пределах этих простых ограничений мы наблюдаем огромное разнообразие и сложность содержания, функционирования и истории прокариотических геномов. Показательные примеры этого встречаются в изобилии. Демонстрация того, что значительное большинство генов в каждом геноме являются не ОРС, а имеют ортологи, является краеугольным камнем сравнительной геномики и эволюционных реконструкций сценариев. Однако обратная сторона медали, а именно фрагментарное распределение кластеров ортологичных генов в геномном пространстве, является не менее фундаментальной закономерностью. Это распределение – результат действия основных факторов, формирующих эволюцию прокариот: ГПГ, потери генов, что часто вызвано упрощением генома, и замены генов неортологами, что отображает неизоморфное соответствие между геномным и функциональным пространствами. Практически неограниченная пластичность архитектуры геномов прокариот, обусловленная частыми геномными перестройками, которые порождают

разнообразные вариации на темы консервативных оперонов, а также открытие ранее неизвестных систем передачи сигналов, регуляции и защиты (лишь несколько примеров которых кратко были рассмотрены в этой главе) еще более увеличивают сложность геномного пространства прокариот, открытого сравнительной геномикой.

Представляется, что главная концептуальная новизна, привнесенная в биологию сравнительной геномикой, – это демонстрация повсеместного распространения ГПГ в мире прокариот, даже с учетом того, что масштабы перемещения генов между организмами, не являющимися близкими родственниками, остаются предметом изучения и дебатов. Независимо от того, какое направление примут эти дебаты в ближайшие годы, широкое распространение ГПГ и очевидное отсутствие непреодолимых барьеров означает, что мир прокариот представляет собой один перемешанный пул генов. Этот пул имеет сложную, разделенную на отдельные районы структуру, в которой удаленные друг от друга части обмениваются генами с меняющейся в широких пределах интенсивностью. Горизонтальный перенос генов влияет на разные классы генов в разной степени, по крайней мере частично в соответствии с гипотезой сложности, но ни один ген, похоже, не избавлен полностью от возможности ГПГ. Критически важно понимать, что существенный уровень ГПГ жизненно необходим для выживания бесполой популяций прокариот на больших интервалах времени, ибо иначе такие популяции вымирают в результате мутационной катастрофы. Таким образом, существенный темп ГПГ является условием *sine qua non* для постоянного выживания и эволюции в мире прокариот. Кроме того, рассмотрение роли ГПГ в формировании эволюционных процессов в геномах прокариот, вместе с универсальными законами масштабирования, приводит нас к заключению, пусть пока только качественному, что пул генов прокариот является в целом хорошо перемешанным, несмотря на все локальные неоднородности.

Важно также отметить, что существенная часть большинства геномов прокариот принадлежит мобилому, обширному множеству генов, которые приходят в геномы и уходят из них с поразительной скоростью и в основном являются эгоистичными генетическими элементами, не имеющими никакого адаптивного значения для генома своего хозяина, хотя иногда эти элементы рекрутируются хозяином для выполнения специфических биологических функций [\[56\]](#).

Взятые вместе, эти открытия формируют новую картину динамичного мира прокариот, который лучше всего может быть представлен сетью генетических элементов, которые обмениваются генами с существенно различающимися скоростями. В этой сети отличие между относительно стабильными хромосомами и мобиломом носит скорее количественный, чем качественный характер. Поразительная универсальность общей организации геномов прокариот, по-видимому, определяется динамическим характером пространства-времени геномной вселенной прокариот, наряду с интенсивным очистительным отбором, поддерживаемым за счет большого эффективного размера популяции большинства прокариот. Сами по себе эти характерные большие размеры популяций во многом определяются высоким уровнем ГПГ, постольку поскольку в противном случае деграция и в конечном счете мутационная катастрофа неизбежны (более подробно об этом см. гл. 8).

Парадокс сегодняшнего состояния дел в эволюционной геномике состоит в том, что, несмотря на громадный прогресс (но также и благодаря ему), проявляющаяся сложность мира прокариот в настоящее время не доступна адекватному анализу. Мы все еще не имеем соответствующего языка (в смысле теории и методологии) для описания динамики и истории геномной сети. Разработка адекватной концептуальной платформы для понимания эволюции прокариот является основным вызовом следующего этапа в развитии этой сферы исследований. В главе 6 описываются некоторые скромные шаги в этом направлении.

Рекомендуемая дополнительная литература

Dauids W., and Z. Zhang.(2008) The Impact of Horizontal Gene Transfer in Shaping Operons and Protein Interaction Networks – Direct Evidence of Preferential Attachment. *BMC // Evolutionary Biology*8: 23.

Количественное изучение скорости ГПГ между генами ядра и оболочки у бактерий. Гены оболочки, в частности, вовлеченные в систему клеточной защиты, имеют более высокую скорость ГПГ.

Doolittle, W. F.(1999) Lateral Genomics // *Trends in Cell Biology*9: M5—8.

Ранняя дискуссия о вкладе ГПГ в концепцию древа жизни.

Doolittle, W. F., and O. Zhaxybayeva.(2009) On the Origin of Prokaryotic Species // *Genome Research*19: 744–756.

Концептуальное обсуждение прокариотических видов с акцентом на многоликость эволюционных процессов у архей и бактерий, приводящих к различным формам объединений организмов, некоторые из которых похожи на формирование видов у эукариот, а другие нет.

Frost, L. S., R. Leplae, A. O. Summers, and A. Toussaint.(2005) Mobile Genetic Elements: The Agents of Open Source Evolution // *Nature Reviews Microbiology*3: 722–732.

Обзор влияния мобильных элементов на эволюцию.

Gogarten, J. P., and J. P. Townsend. (2005) Horizontal Gene Transfer, Genome Innovation and Evolution. *Nature Reviews Microbiology*3: 679–687.

Обзор, демонстрирующий ключевую роль ГПГ в эволюции, по крайней мере в мире прокариот.

Jain, R., M. C. Rivera, and J. A. Lake.(1999) Horizontal Gene Transfer Among Genomes: The Complexity Hypothesis // *Proceedings of the National Academy of Sciences USA* 96: 3,801—3,806.

Ключевое исследование, представляющее гипотезу, что гены, вовлеченные в многочисленные взаимодействия (комплексы), менее склонны к ГПГ, чем гены, имеющие незначительное число партнеров.

Koonin, E. V., and Y. I. Wolf. (2008) Genomics of Bacteria and Archaea: The Emerging Dynamic View of the Prokaryotic World // *Nucleic Acids Research*36: 6,688—6,719.

Обширный концептуальный обзор эволюции прокариот, с акцентом на динамичный характер эволюции, вызванный ГПГ, заменой неортологичными генами, утратой генов, специфичной для линии организмов, и активностью мобильных элементов. Обсуждается состоящая из трех частей вселенная генов прокариот.

Lapierre, P., and J. P. Gogarten. (2009) Estimating the Size of the Bacterial Pan-Genome // *Trends in Genetics*25: 107–110.

Попытка оценить размер содержимого пангенома на основе похожей на описанную в предыдущей статье модели распространения генов среди прокариот.

Lawrence, J.(1999) Selfish Operons: The Evolutionary Impact of Gene Clustering in Prokaryotes and Eukaryotes // *Current Opinion in Genetics & Development*9(6): 642–648.

Обзор данных в поддержку гипотезы «эгоистичного оперона», утверждающей, что опероны поддерживаются больше за счет ГПГ, чем из-за преимуществ совместной регуляции и экспрессии генов.

Sapp, J.(2005) The Prokaryote – Eukaryote Dichotomy: Meanings and Mythology // *Microbiology and Molecular Biology Reviews*69: 292–305.

Исторический анализ революции Вёзе и неадекватность концепции прокариот.

Woese, C. R.(1994) There Must Be a Prokaryote Somewhere: Microbiology's Search for Itself //

Microbiology Reviews 58: 1–9.

Блестящее обсуждение «революции Вёзе» в (микро)биологии самим архитектором этой революции. Акцент сделан на «несвязность» прокариот и взаимоотношения между археями и эукариотами.

Глава 6. Филогенетический лес и поиск неуловимого древа жизни в век геномики

Очень краткая история древа жизни

Концепция древа жизни (ДЖ) в ее современном значении была впервые представлена Дарвином в его записных книжках еще в 1838 году. Двадцатью годами позже Дарвин запечатлел ее в одной-единственной иллюстрации к «Происхождению видов...». Конечно же не ему принадлежит идея изображения генетических взаимоотношений в форме дерева. Деревья столетиями использовались для изображения родословных, таких как настоящие истории семей (королевских, к примеру). Тем не менее именно Дарвину принадлежит плодотворная идея о том, что различные виды связаны между собой деревом, причем листья соответствуют существующим в настоящее время видам, а внутренние вершины ^[57] — вымершим, предковым формам. Более того, Дарвин сформулировал радикальную гипотезу о том, что в конечном счете вся история жизни может быть представлена в виде одного гигантского дерева.

Родство всех существ одного класса иногда изображают в форме большого дерева. Я думаю, что это сравнение очень близко к истине. Зеленые ветви с распускающимися почками представляют существующие виды, а ветви предшествующих лет соответствуют длинному ряду вымерших видов... Разветвления ствола, делящиеся на своих концах сначала на большие ветви, а затем на более и более мелкие веточки, были сами когда-то, когда дерево еще было молодо, побегами, усеянными почками; и эта связь прежних и современных почек, через посредство разветвляющихся ветвей, прекрасно представляет нам классификацию всех современных и вымерших видов, соединяющую их в соподчиненные друг другу группы (Darwin, 1859) (здесь и далее пер. К. А. Тимирязева, С. Л. Соболя, цит. по изд.: *Дарвин Ч. Сочинения*. Т. 3. М: Изд-во АН СССР, 1939).

В шестом издании «Происхождения...» (Darwin, 1872) Дарвин пошел дальше и недвусмысленно ввел понятие ДЖ.

Как почки в процессе роста дают начало новым почкам, а эти, если только сильны, разветвляются и заглушают многие слабые ветви, так, полагаю, было при воспроизведении и с великим деревом жизни, наполнившим своими мертвыми опавшими сучьями кору земли и покрывшим ее поверхность своими вечно расходящимися и прекрасными ветвями.

Для дарвиновских времен это было невероятно смелое предположение, ведь никаких веских свидетельств в пользу общего происхождения всех форм жизни не было, не говоря уже о том, что Дарвин и другие биологи XIX века понятия не имели о том, насколько жизнь на Земле на самом деле многообразна. Тем не менее гипотеза универсального общего предка стала популярной. Через несколько лет после публикации «Происхождения...» Геккель населил абстрактное древо жизни Дарвина реальными формами жизни, почти исключительно животными, с ЧЕЛОВЕКОМ на вершине и амебами и дробянками (название бактерии в XIX веке) у корней (Haesckel, 1997). С тех пор ДЖ стало центральным элементом эволюционной биологии и в каком-то смысле биологии вообще.

На протяжении примерно 140 лет после Дарвина и Геккеля филогенетические деревья (изначально конструируемые на основании фенотипических признаков, но после фундаментальных работ Эмиля Цукеркандля и Лайнуса Полинга в начале 1960-х все чаще полагающиеся на сравнение молекулярных последовательностей) воспринимались как в целом точные отображения эволюции соответствующих организмов. Другими словами, древо, выстроенное для конкретного признака или гена, приравнивалось, по умолчанию, к «древу видов». Принятие 16S-рРНК, молекулы, универсальной для клеточных форм жизни, в качестве золотого стандарта генетической реконструкции привело к трехдоменному древу жизни Вёзе и коллег. Это была достойная кульминация героического периода филогенетики (Rase, 2006;

Woese, 1987; Woese et al., 1990). В древе 16S РНК содержались части с великолепным разрешением ветвей, и, хотя многие другие части остались проработаны довольно слабо, в особенности в глубине древа, ожидалось, что дальнейшее усовершенствование методов филогенетики, вкуче с анализом нескольких дополнительных универсальных генов, позволит получить подробную и исчерпывающую топологию ДЖ в не столь отдаленном будущем (Расе, 1997).

Сложности у концепции ДЖ появились еще до появления геномики, поскольку стало ясно, что среди некоторых распространенных и необходимых генов прокариот наблюдаются множественные горизонтальные переносы генов (ГПГ). Петер Гогартен и коллеги предложили метафору «сеть жизни» в качестве потенциальной замены ДЖ (Hilario and Gogarten, 1993). Однако эти идеи не получили значительной поддержки в догеномную эру, и ГПГ рассматривался в основном как незначительный эволюционный процесс, важный в отдельных областях (таких как распространение устойчивости к антибиотикам), однако в общем эволюционном процессе играющий второстепенную роль – и являющийся несущественным осложнением в процессе построения всеобъемлющего ДЖ. В конце 1990-х сравнение геномов прокариот радикально изменило эту картину, показав, что распределение большинства среди геномов неоднородно (члены большинства КОГ разбросаны среди разнообразных организмов) и топологии филогенетических древ отдельных генов часто не соответствуют друг другу. Эти данные позволили предположить, что ГПГ очень широко распространен среди бактерий и археобактерий и также представлял определенную важность для эволюции эукариот, особенно в контексте эндосимбиоза (см. гл. 7). Таким образом, идеальное ДЖ оказалось химерой, поскольку широко распространенный ГПГ приводит к тому, что древо любого отдельного гена не является точным отображением эволюции целых геномов. Осознание того, что ГПГ среди прокариот является доминирующей формой эволюции, а не редким процессом, привело к идее «выкорчевывания» ДЖ – прежде всего в нескольких влиятельных обзорных статьях Форда Дулиттла (Doolittle, 1999a, b, 2000). Заявленное падение ДЖ привлекло к себе много внимания не только в профессиональных публикациях, но и в научно-популярной литературе (Pennisi, 1999). Это событие часто воспринимается как сдвиг парадигмы эволюционной биологии, если не биологии вообще (O'Malley and Boucher, 2005; см. прил. I).

Взгляды эволюционных биологов на статус ДЖ в свете широкой распространенности ГПГ охватывают весь диапазон – от (i) упорного отрицания значения роли ГПГ в эволюции жизни до (ii) «умеренного» пересмотра концепции ДЖ и вплоть до (iii) полного «выкорчевывания» ДЖ, когда сама эта концепция объявляется лишеной смысла как представление эволюции организмов или геномов (O'Malley and Boucher, 2005). По мере накопления сравнительных геномных данных, установка на отрицание ГПГ быстро становится скорее некоей психологической странностью, нежели научной позицией, которую можно обосновать. Настоящие споры идут, похоже, между «ревизионистами» и «радикальным корчевателями» (ii и iii). Сторонники умеренного подхода придерживаются мнения, что, несмотря на все различия между древами отдельных генов, ДЖ по-прежнему имеет большое значение как главная тенденция, которая, хотя бы в принципе, может быть охарактеризована посредством всестороннего сравнения топологий филогенетических древ (Wolf et al., 2002). Приверженцы радикальных взглядов, напротив, считают, что массовые ГПГ уничтожают самые различия между вертикальными и горизонтальными путями передачи генетической информации, поэтому концепция ДЖ должна быть оставлена в пользу сетевой репрезентации эволюции (в широком ее понимании) (Doolittle and Bapteste, 2007; Gogarten et al., 2002).

Противоречивость концепции ДЖ особенно ярко проявилась в дебатах вокруг автоматически построенного «древа жизни в высоком разрешении», которое Пер Борк с

коллегами получили путем объединения выровненных последовательностей тридцати одного высококонсервативного белка, в основном из участвующих в процессе трансляции (Ciccarelli et al., 2006). Очень скоро это предполагаемое ДЖ было отброшено как «дерево 1 процента» (генов в любом произвольном геноме), которое в целом не отражает историю геномов. По крайней мере для меня эффектно сформулированный аргумент Тал Даган и Билла Мартина (Dagan and Martin, 2006) звучит неотразимо и достоин развернутой цитаты:

«Когда химики или физики обнаруживают, что некая нулевая гипотеза может объяснить только 1 процент полученных ими данных, они немедленно начинают искать другую, лучшую гипотезу. Но похоже, с микробной эволюцией дело обстоит иначе, что не может не беспокоить. Возможно ли, что многие биологи из всех сил хотят отыскать древо жизни, невзирая даже на то, что говорят им факты?»

В настоящей главе я представляю полное разбиение эволюции прокариот на древовидные и сетевидные компоненты, которые, как мне кажется, вполне могут объективно определить роль и место древ в нашем понимании эволюции, а также в определенной степени разрешить полемику вокруг ДЖ. Однако, прежде чем приступить к этому количественному анализу, мы рассмотрим на концептуальном уровне корни «древесного мышления» [\[58\]](#).

Фундаментальные единицы эволюции и присущая им древовидная природа

Как было рассмотрено в главе 2, репликация генетического материала, процесс, по самой своей сути предрасположенный к ошибкам, является одновременно и условием, и прямой причиной эволюции. Критическим в определении статуса древа в биологии является то, что репликация и непременно следующая за ней эволюция суть по природе своей древовидные процессы (Koonin and Wolf, 2009a). И в самом деле, воспроизводящаяся молекула порождает две копии (в случае полуконсервативной репликации дцДНК, которая происходит во всех клеточных организмах и во многих вирусах) или множество копий (в случае консервативной репликации вирусов с геномами, представленными оцДНК или оцРНК) с ошибками, что приводит к древовидному процессу разветвления (см. рис. 6–1). В терминах теории графов – такой процесс может быть изоморфно представлен в виде особой формы направленного ациклического графа, известного как древовидное образование (arborescence) – то есть обобщенное дерево, в котором допустимы множественные разветвления, а все ребра направлены в противоположную от корня сторону (см. рис. 6–1). Хотя случайное вымирание одной или обеих молекул-потомков порождает вершины, не испускающие ребер, такой граф остается древовидным образованием; определение этого класса графов не требует, чтобы листья находились на одном уровне (см. рис. 6–1; здесь и далее вместо термина «древовидное образование» я буду использовать более распространенный термин «дерево»).

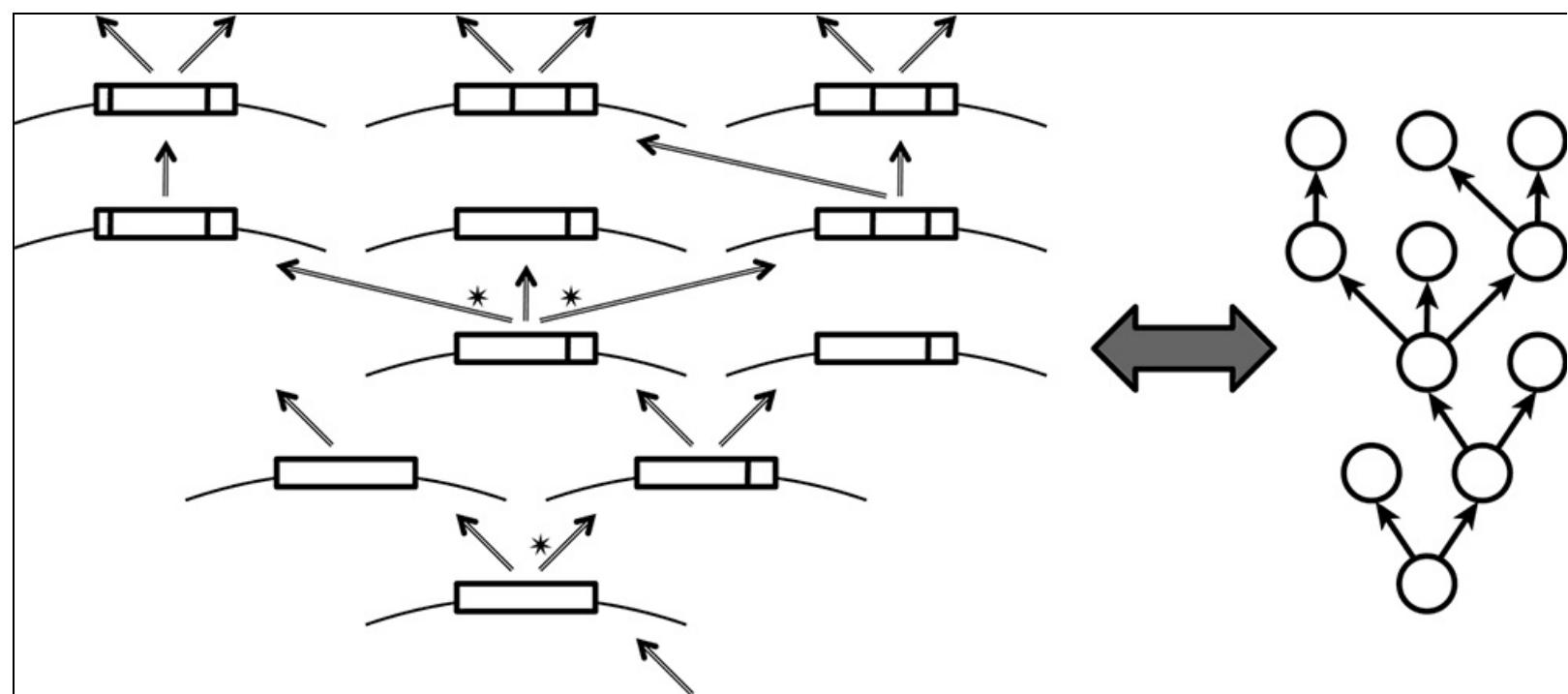


Рис. 6–1. Дерево (древовидное образование) как изоморфное представление предрасположенного к ошибкам процесса репликации генов. Схематическое изображение истории репликации генетического элемента, включающее как раздвоения, так и множественные разветвления (отмечены звездочкой). Зафиксированные мутации показаны штрихами. Адаптировано из Koonin and Wolf, 2009a.

Потенциально серьезным осложнением для древовидного характера эволюции является рекомбинация. Будучи широко распространенной, рекомбинация может превратить представление истории воспроизводящегося ряда поколений из дерева (см. рис. 6–1) в сеть (или,

того хуже, в полную кашу). Возможно ли определить фундаментальный, «атомный» уровень генетической организации, на котором рекомбинацией можно пренебречь? Это представляется сомнительным в случае гомологичной рекомбинации, широко распространенной во время совместной репликации близкородственных последовательностей, в частности у эукариот, вовлеченных в обычные половые отношения, и у «квазиполовых» прокариот. Чаще всего единицей гомологичной рекомбинации является одиночная пара оснований. Однако гомологичная рекомбинация не может происходить между отдаленно родственными последовательностями, поэтому ГПГ между таксономически удаленными прокариотами подразумевает только негомологичную (незаконную) рекомбинацию, которой способствуют специфические механизмы, такие как распространение генов через бактериофаги и плазмиды (см. гл. 5). В отличие от гомологичной рекомбинации, следует ожидать, что эволюционная фиксация событий негомологичной рекомбинации вне генов или между частями генов будет гораздо предпочтительнее; сохранение целостности гена после негомологичной рекомбинации внутри генов крайне маловероятно. Распространенность внутригенной рекомбинации в ходе ГПГ между отдаленно связанными прокариотами не изучалась сколь-нибудь подробно. Тем не менее как минимум одно исследование показывает, что сегменты, кодирующие сравнительно небольшие белковые домены, в значительной степени избегаются рекомбинацией (Chan et al., 2009). Отсюда следует важный и правдоподобный, хотя и не подкрепленный пока что в достаточной мере данными, вывод: *из-за гомологичной рекомбинации эволюционная история гена или домена представляет собой сетевую структуру на малых масштабах, но преимущественно древовидную на больших* (см. рис. 6–2).

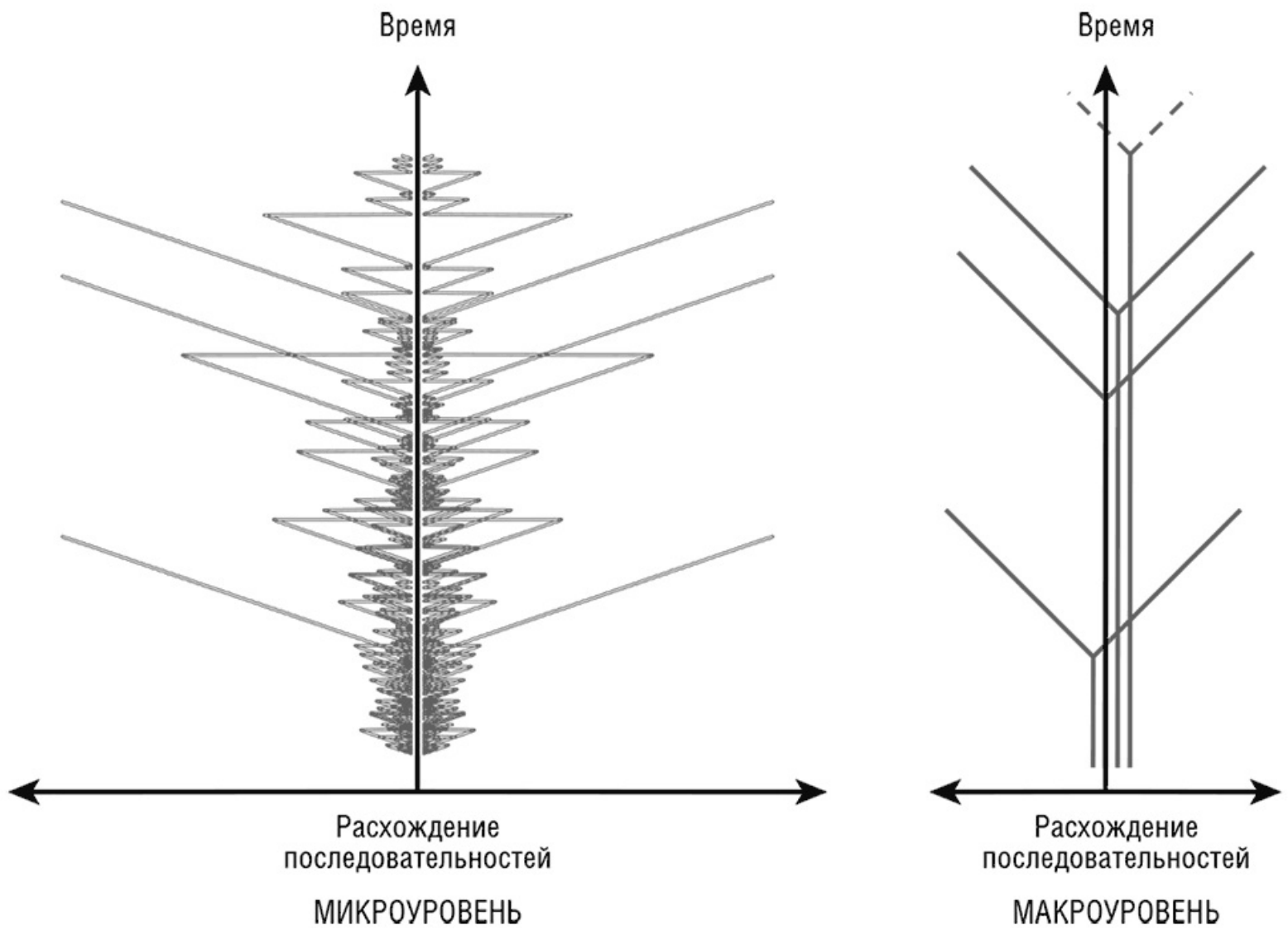


Рис. 6–2. Эволюция гена имеет сетевой вид на малых масштабах, но древовидный на больших. Изображение схематично описывает эволюцию четырех генов. История расхождения каждого гена была вычислена на модели произвольной гомологичной рекомбинации с экспоненциальным снижением частоты рекомбинации с расхождением последовательностей. На каждом шаге моделирования два дочерних гена расходились на постоянную величину (расхождение по принципу молекулярных часов) и либо подвергались гомологичной рекомбинации (что сводило разницу между ними к нулю), либо нет, сохраняя существующее состояние расхождения. После некоторого количества коротких периодов расхождений и рекомбинаций гены стохастически расходились достаточно далеко, чтобы (гомологичная) рекомбинация стала крайне маловероятной. После этой точки они продолжали расходиться без рекомбинации. На большом масштабе это выглядит как простое раздвоение на древе-графе. Адаптировано из Koornin and Wolf, 2009a.

Форд Дулитл и Эрик Батист предположили и продемонстрировали на весьма убедительных примерах, что с помощью дерева можно легко описать отношения между объектами, которые отнюдь не связаны общей родословной, а потому «древесное мышление» не следует считать априорно применимым или, во всяком случае, центральным в биологии (Doolittle and Bapteste, 2007). Несмотря на то что аргумент этот сам по себе обоснован, в нем упускается из вида то принципиально важное обстоятельство, которого мы касались выше, а именно что дерево – это неизбежное формальное следствие истории репликации нуклеиновых кислот и последующей эволюции. Таким образом, деревья нельзя убрать из эволюционной биологии по

фундаментальной причине: *они изначально присущи эволюционному процессу*. И тогда наиболее уместным вопросом становится такой: каковы фундаментальные генетические единицы, эволюция которых лучше всего представима деревом? В практике эволюционной биологии деревья чаще всего выстраиваются для отдельных генов или для наборов генов, которые, как считается, имеют общую историю. Однако обычно подразумевается (или даже недвусмысленно заявляется), что конечной целью является древо видов (организмов). Недостаток ясности в вопросе об основной единице, деревья которой должны строиться и анализироваться, является важным (если не главным) источником всей дискуссии вокруг ДЖ.

На концептуальном уровне ответ на заданный выше вопрос кажется ясным: *фундаментальную единицу эволюции можно в целом удовлетворительно определить как самую малую порцию генетического материала с отчетливой эволюционной траекторией* — то есть такую, которая развивается независимо от других таких же единиц на протяжении достаточно длительного эволюционного периода. На практике, учитывая динамический характер эволюции прокариот, описанный в главе 5, *критерию фундаментальной единицы древовидной эволюции отвечает геномный локус, кодирующий РНК или белок (или индивидуальный эволюционный домен)*. (Очевидно, такая единица соответствует гену, за исключением мультидоменных белков.) В самом деле, как впервые явно отметил Ричард Докинз (Dawkins, 2006), гены в большой степени эгоистичны, то есть подвержены отбору, частично независимому от других генов. В условиях обширного ГПГ ген или оперон потенциально может оказаться в широком спектре организмов. Конечно, обычно это происходит, когда некий ген дает селекционное преимущество организму-носителю, поэтому эволюция генов и эволюция организмов тесно связаны.

Осознание того, что отдельные гены, в противоположность геномам, являются «атомами» эволюции, ставит под сомнение самую идею ДЖ. Тем не менее, как было показано выше, деревья невозможно убрать из какого бы то ни было описания эволюции, по той простой причине, что репликация генетического материала — процесс по сути своей древовидный. Эти два фундаментальных наблюдения вместе приводят к логичному заключению о том, что должно прийти на смену ДЖ: *лес жизни (ЛЖ), то есть совокупность филогенетических деревьев всех генов* (за очевидным исключением ОРС). В таком случае реконструкция истории жизни (ясно, что не всей истории полностью, но ее «скелета») не так проста, как анализ топологии ДЖ. Эта реконструкция требует картирования ЛЖ в поисках «роц» подобных деревьев, которые могут быть отражением долгосрочных тенденций связанной (вертикальной) эволюции наборов генов, и «лиан» ГПГ. Представляется, что всестороннее исследование ЛЖ и есть главная цель филогенетики. В следующем разделе я преимущественно рассматриваю результаты недавних исследований ЛЖ, проведенных совместно с моими коллегами Пере Пуигбо и Юрием Вольфом (Puigbo et al., 2009, 2010). Это ни в коем случае не единственные исследования, сравнивающие филогенетические деревья и пытающиеся провести различия между вертикальной и горизонтальной тенденциями в эволюции. Однако эта работа соответствует современным требованиям, и мне кажется, что мы нашли пригодные способы для представления отношений между деревьями многочисленных генов, поэтому краткое изложение этих результатов дает хорошее представление о структуре ЛЖ. (Изложение в следующих двух разделах носит заметно более специальный характер, чем эта книга в целом; некоторые читатели могут решить сразу перейти к заключительным параграфам каждого раздела, а затем и к краткому обзору главы.)

Лес жизни и почти универсальные филогенетические деревья

В принципе в ЛЖ входят деревья для «всех» генов. Однако на деле работать с тысячью или около того геномных последовательностей прокариот (это число увеличится на несколько сотен к тому времени, когда эта книга будет опубликована) тяжело технически, поскольку максимально правдоподобные (maximum likelihood) методы построения деревьев, обеспечивающие наилучшее разрешение, тяжелы в вычислительном отношении (то есть плохо масштабируются с увеличением числа видов). К счастью, использование всех геномов, видимо, не так уж и важно. Несмотря на динамичную эволюцию прокариот, гены ядра и оболочки в близкородственных организмах (идентифицированные, к примеру, по высокому сходству последовательностей рРНК или других генов ядра) большую часть времени эволюционируют синхронно (а только гены ядра и оболочки распространены достаточно широко для получения информативных филогенетических деревьев). Таким образом, тщательно отобранного представительного набора организмов должно быть достаточно для определения главных тенденций в ЛЖ. Для исследований, которые здесь рассматриваются, мы сконструировали такой набор из 100 геномов прокариот, 41 архейного и 59 бактериальных (в дальнейшем в этой главе мы ссылаемся на эти прокариоты как на виды – с полным осознанием ограничений этой концепции, которые были отмечены в гл. 5). Деревья были построены для всех наборов ортологов с более чем четырьмя членами (минимальное число последовательностей, необходимых для построения бескорневого дерева), таким образом, в общей сложности мы получили 7000 деревьев. Как и ожидалось, с учетом структуры геномного пространства прокариот из ядра, оболочки и облака, описанной в главе 5, большинство из этих деревьев маленькие: только 2040 состояли из более чем 20 видов, и лишь небольшой набор из 102 почти универсальных древ (ПУД) ^[59] включали более 90 процентов проанализированных прокариот.

Обычно филогенетики пытаются определить ГПГ путем сравнения деревьев отдельных генов с заданным заранее «древом видов». Однако, как мы увидели в предыдущем разделе, сама концепция «древа видов» сводится на нет всепроникающим ГПГ и эгоистичностью отдельных генов, которые являются фундаментальными единицами древовидной эволюции. Мы попытались исследовать структуру ЛЖ, не руководствуясь какой-либо предвзятой идеей стандартного древа, с которым следует сравнивать все остальные деревья. С этой целью мы проанализировали полную матрицу топологических расстояний между деревьями; это была довольно большая матрица, включающая почти 24 миллиона попарных сопоставлений деревьев, хотя многие клетки в матрице были пусты, потому что соответствующие деревья состояли из непересекающихся наборов видов.

На рис. 6–3 ЛЖ представлен в виде сети, где каждый узел соответствует дереву. Мы видим, что группа ПУДов занимает особое место в этой сети: около 40 процентов деревьев крайне похожи как минимум на один из ПУДов. (Два дерева считаются топологически подобными, когда различия в соединениях между их ветвями незначительны; из топологических различий высчитываются расстояния между деревьями. Подробности этих вычислений мы опустим.) Напротив, при использовании того же порога сходства 102 случайных дерева того же размера, что и ПУДы, были связаны всего лишь с примерно 0,5 процента деревьев в ЛЖ. Таким образом, существует высокое и неслучайное топологическое подобие между ПУДами и значительной частью ЛЖ.

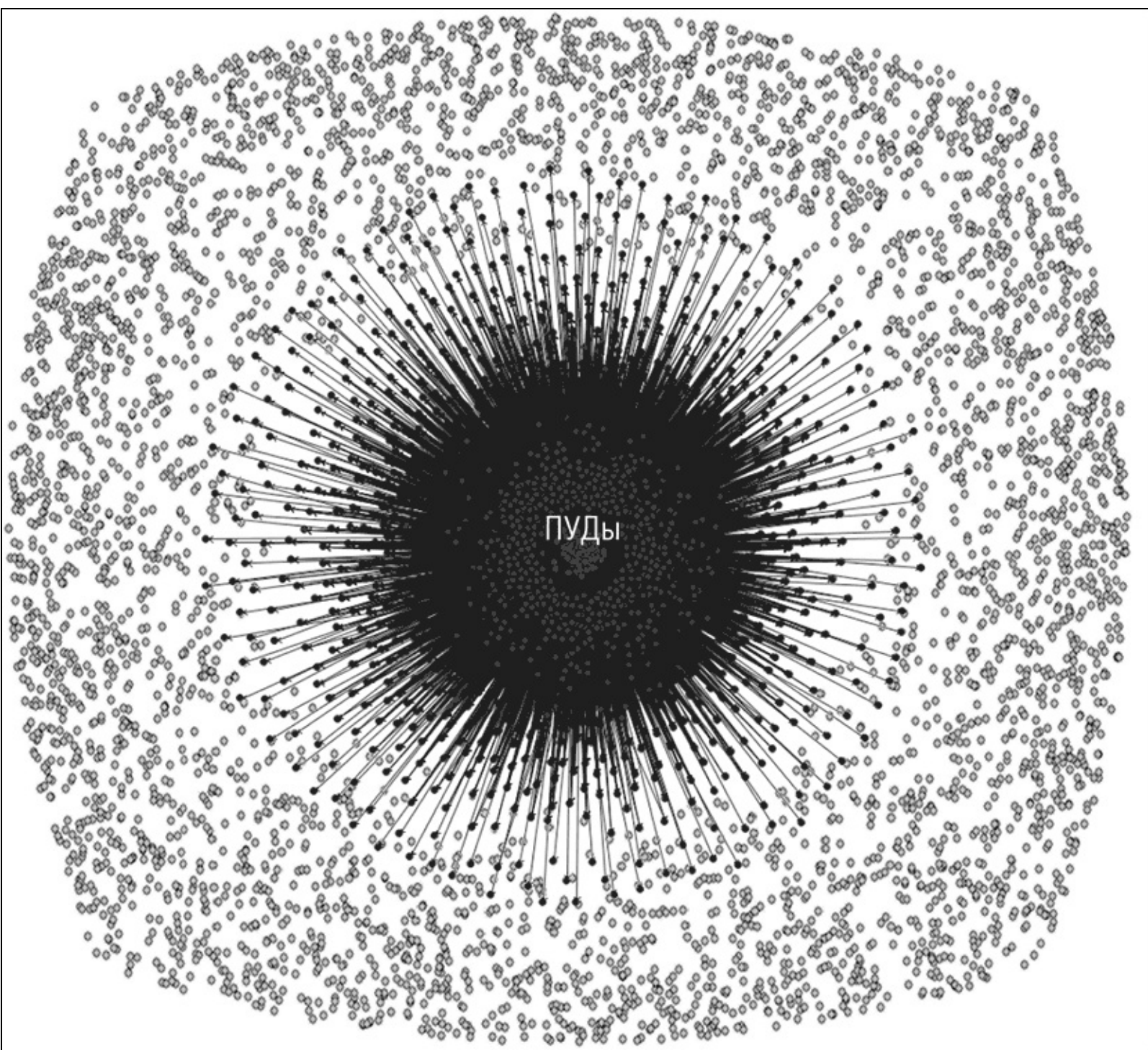


Рис. 6–3. Лес жизни как сеть деревьев. Каждый узел в сети обозначает дерево. 102 почти универсальных дерева (ПУД) показаны в виде темных точек в середине, а остальные деревья – в виде незаштрихованных кружков. ПУДы связаны с деревьями с подобной топологией – то есть обладающими как минимум 50-процентным подобием с как минимум одним ПУДом. Адаптировано из Puigbo et al., 2009.

Зная все расстояния между деревьями в ЛЖ, мы можем применить статистические методы для кластеризации данных – то есть определить, является ли ЛЖ просто облаком случайно расположенных точек (деревьев в топологическом пространстве) или содержит определенные кластеры деревьев с подобными топологиями. Используемый статистический метод разделил ЛЖ на семь кластеров деревьев. Примечательно, что все ПУДы образовали компактную группу в пределах одного из кластеров (см. рис. 6–4). Семь кластеров существенно отличаются по распределению деревьев по количеству видов, распределению археобактерий и бактерий, а также функциональной классификации соответствующих генов. Таким образом, результаты

кластеризации показали, что ЛЖ может быть разделен на крупные, отчетливые группы топологических подобных деревьев; однако на этой стадии по-прежнему неясно, насколько эта кластеризация обусловлена «вертикальными» эволюционными процессами и насколько – горизонтальными. Ключевым наблюдением является то, что все ПУДы занимают компактную, неразрывную область пространства ЛЖ, не разделяются на отчетливые кластеры (в отличие от остальных деревьев в ЛЖ) и отделены примерно одинаковыми расстояниями от всех кластеров (см. рис. 6–4).

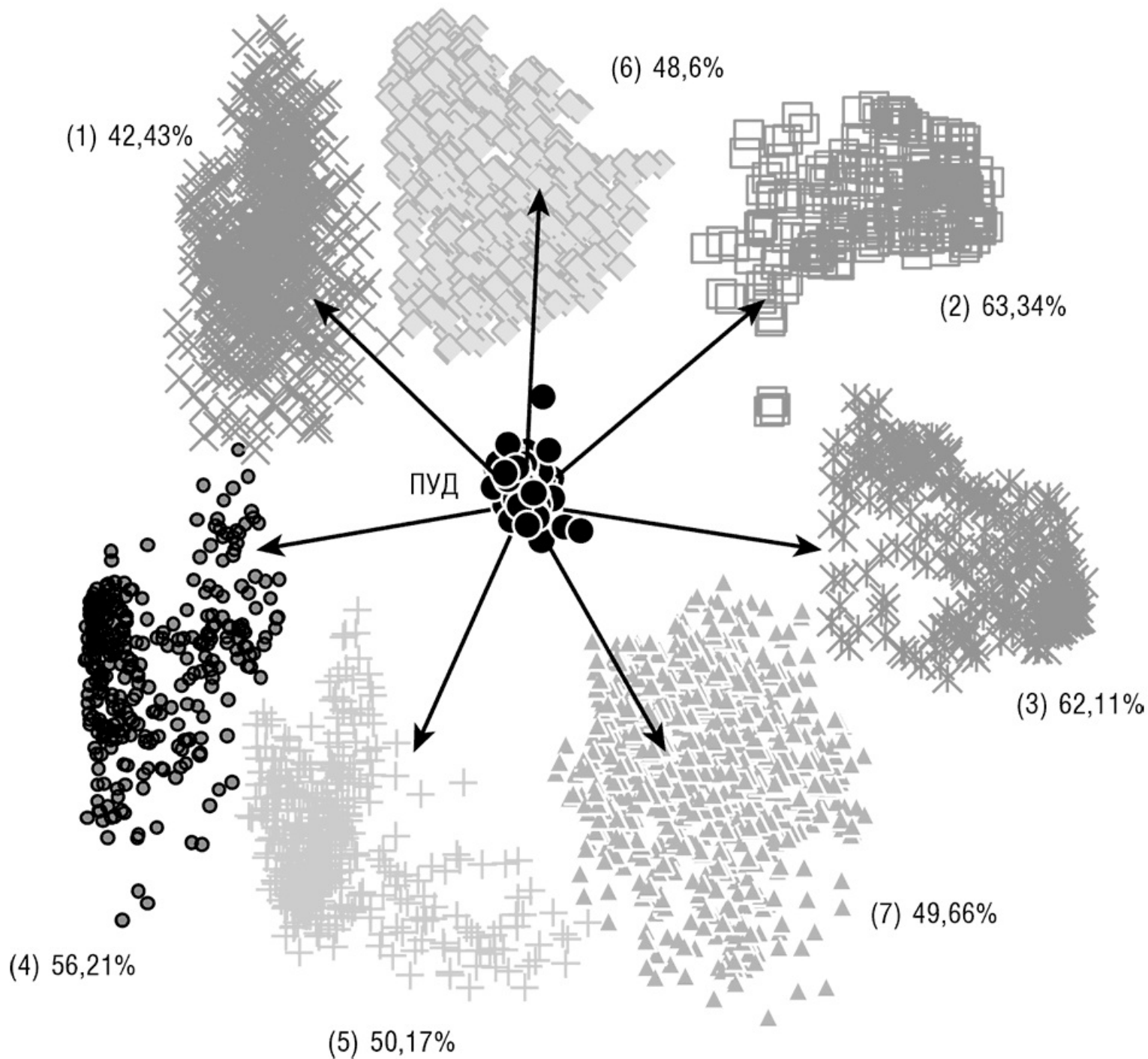


Рис. 6–4. Кластеризация леса жизни в топологическом пространстве. Кластеры были получены с помощью классического многомерного шкалирования (метод кластеризации, являющийся по сути более изощренной версией популярного подхода анализа главных компонент). ПУДы произвольно помещены в центр, показано среднее подобие между ПУДами и каждым из кластеров. Адаптировано из Puigbo et al., 2009.

Результаты первой части нашей экспедиции в чащу леса жизни приводят к важному заключению: топологии ПУДов крайне схожи между собой и, возможно, представляют главную эволюционную тенденцию в ЛЖ. Заявление о главной тенденции может показаться неочевидным, однако оно отражает очень простые и непосредственные наблюдения:

1. Топологии ПУДов сильно схожи с топологиями многих других деревьев в ЛЖ.

2. ПУДы расположены на приблизительно одинаковом расстоянии от кластеров других деревьев. В этом смысле они занимают центральное положение в ЛЖ.

До сих пор мы говорили о ПУДах отвлеченно, не учитывая реальных генов, которые стоят за этим набором больших деревьев. На самом деле природа ПУДов вполне предсказуема: это гены, кодирующие рибосомные белки, а также другие высококонсервативные белки, участвующие в трансляции, наряду с некоторыми ключевыми субъединицами ДНК-зависимой РНК-полимеразы. Это гены, которые, согласно гипотезе сложности, должны быть наименее предрасположены к ГПГ (Jain et al., 1999). Этот набор почти универсальных генов, что несколько парадоксально, также являет собой один из наиболее впечатляющих примеров ГПГ, в частности среди аминоксил-тРНК-синтетаз (АРСаз), некоторые из которых отвечают за устойчивость к антибиотикам, но также и среди некоторого количества рибосомных белков. Как бы то ни было, приведенные здесь наблюдения недвусмысленно показывают, что группа ПУДов внутренне топологически плотная и, более того, связана топологическим подобием со многими другими деревьями в ЛЖ.

В свете практически вездесущего ГПГ ничто не может восстановить ДЖ во всей его былой славе. Однако, если бы нам пришлось искать наиболее осмысленное приближение ДЖ, консенсусная топология ПУДов выглядела бы лучшим из кандидатов. Но прежде чем мы торжественно введем ПУДы в эту должность, следует обсудить более глубокие аспекты эволюции.

В глубь леса жизни: Большой взрыв или сжатый кладогенез?

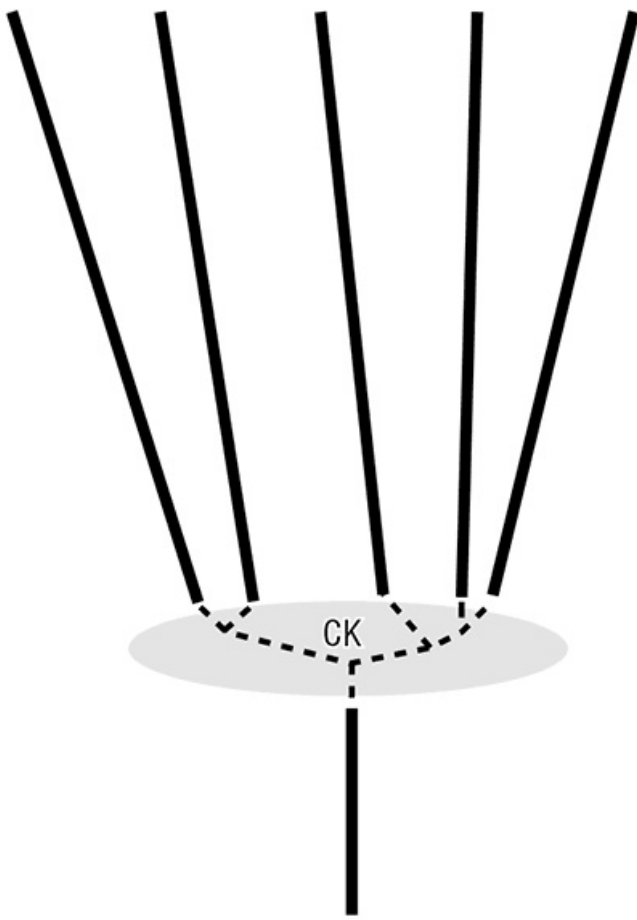
В предыдущем разделе мы увидели, что ПУДы занимают особое положение в ЛЖ. Они топологически подобны друг другу и, следовательно, могут представлять главную тенденцию вертикальной, древовидной эволюции. Однако прежде чем мы признаем, что эти деревья почти универсальных, важнейших генов и в самом деле отражают главную тенденцию в ЛЖ, следует задать следующий ключевой вопрос: пронизывает ли древовидный сигнал всю историю жизни или ограничивается сравнительно недавней эволюцией?

Для такого вопроса у нас есть веская причина. Многие филогенетические исследования, включая изучение суперсети ПУДов (опуская технические детали, суперсеть – консенсусное дерево, сформированное путем «усреднения» топологий ПУДов), явным образом показывают, что глубокорасположенные внутренние узлы филогенетических деревьев, как правило, плохо разрешены по сравнению с узлами, расположенными ближе к листьям (см. рис. 6–5 а). Эта особенность повторяется на многих различных уровнях истории жизни: к примеру, слабое разрешение среди глубочайших ветвей замечено как в филогенетическом древе млекопитающих, которое охватывает промежуток времени примерно в 100 миллионов лет, так и в гипотетическом полном ДЖ, которое охватывает более 3,5 миллиарда лет (Rokas and Carroll, 2006). Во всех этих случаях эволюционные интервалы, включающие первичное разделение крупных групп организмов, оказываются особыми, отличными от «нормальных» эволюционных эпох (аналогия с прерывистым равновесием хотя и поверхностна, но все же соблазнительна – см. гл. 2). Для объяснения этой особенности были предложены две модели:

1. Сжатый кладогенез (см. рис. 6–5 а; Rokas and Carroll, 2006).
2. Более радикальная модель «биологического Большого взрыва» ^[60](ББВ; см. рис. 6–5 б; Koornin, 2007a).

Согласно модели сжатого кладогенеза, эволюция – или, точнее, появление новых групп организмов (клад, то есть отдельных монофилетических ветвей филогенетического дерева) в эпохи трансформаций – происходит быстро, образуя очень короткие внутренние ветви. Соответственно, порядок ветвления в этих частях деревьев крайне тяжело определить с высокой степенью надежности какими бы то ни было филогенетическими методами. Тем не менее в принципе, согласно модели сжатого кладогенеза, существует уникальный порядок ветвей, присущий всему ДЖ (вне зависимости от конкретной интерпретации идеи ДЖ). Модель ББВ утверждает, что переходные эпохи качественно отличаются от «нормальных» древовидных периодов эволюции: модель постулирует, что в результате бурного ГПГ, включая массовый приток генов, вызванный эндосимбиозом и другими процессами, на этих стадиях эволюции полностью отсутствует древовидный сигнал. Отчасти упрощая модель, можно сказать, что в переходные фазы память о предшествующей древовидной эволюции изглаживается и эволюционирующие геномы формируют единый генетический фонд, откуда и возникают новые клады. Какая бы длина ни присваивалась соответствующим внутренним ветвям в процессе построения дерева, согласно модели Большого взрыва это не более чем артефакты; истинная длина всех этих ветвей в точности равна нулю (см. рис. 6–5 б). Вопрос о том, существует ли различимый филогенетический сигнал в самых глубоких узлах деревьев, очевидно важен для обоснования главной древовидной тенденции в ЛЖ, которая предположительно может быть аппроксимирована из топологий ПУДов. К счастью, обе модели могут быть проверены более глубоким анализом тенденций в ЛЖ.

а



б

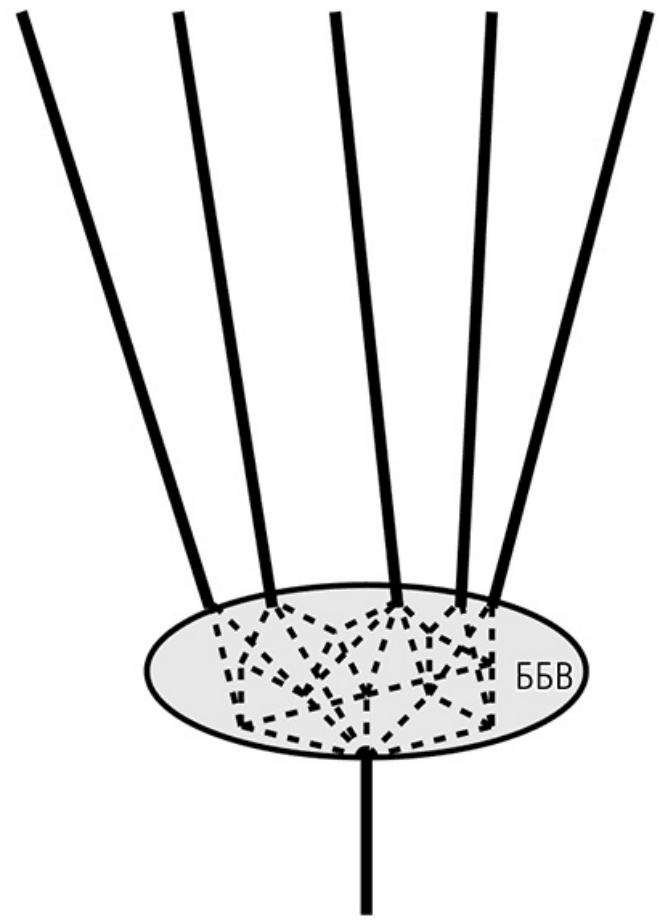


Рис. 6–5. Две модели переходных эпох в эволюции: а– сжатый кладогенез (СК); б– биологический Большой взрыв (ББВ)

Мы ввели новую меру, показатель несовместимости (ПН), которая определяет, насколько репрезентативна топология заданного дерева по сравнению со всем ЛЖ (этот показатель – просто величина, обратная к доле случаев нахождения фрагментов данного дерева во всех деревьях ЛЖ (Puigbo et al., 2009)). Используя ПН, мы объективно изучаем тенденции в ЛЖ, не полагаясь на топологию заранее выбранного «дерева вида». Графики на рис. 6–6 показывают зависимость ПН от филогенетической глубины деревьев во всем ЛЖ и отдельно в ПУДах. Опять же, оставляя в стороне технические тонкости, – чтобы построить эти графики, необходимо разделять деревья на фракции в рамках определенного интервала глубин (особая процедура, детали которой не важны, была разработана для определения глубины по шкале от 0 до 1) и взять среднее значение ПН для этого конкретного интервала. Два графика и разница между ними весьма интересны. График для всего ЛЖ имеет сходство с графиками, описывающими фазовые переходы в различных физических процессах: на определенной глубине значение некоторой переменной (в нашем случае показателя несовместимости) изменяется очень резко (см. рис. 6–6). График для ПУДов сильно отличается: он демонстрирует значительно более низкие значения ПН (то есть топологии ПУДов в среднем более сходны друг с другом, чем топологии других деревьев в ЛЖ) и менее резкие изменения критической глубины, которые вряд ли можно квалифицировать как фазовый переход (см. рис. 6–6).

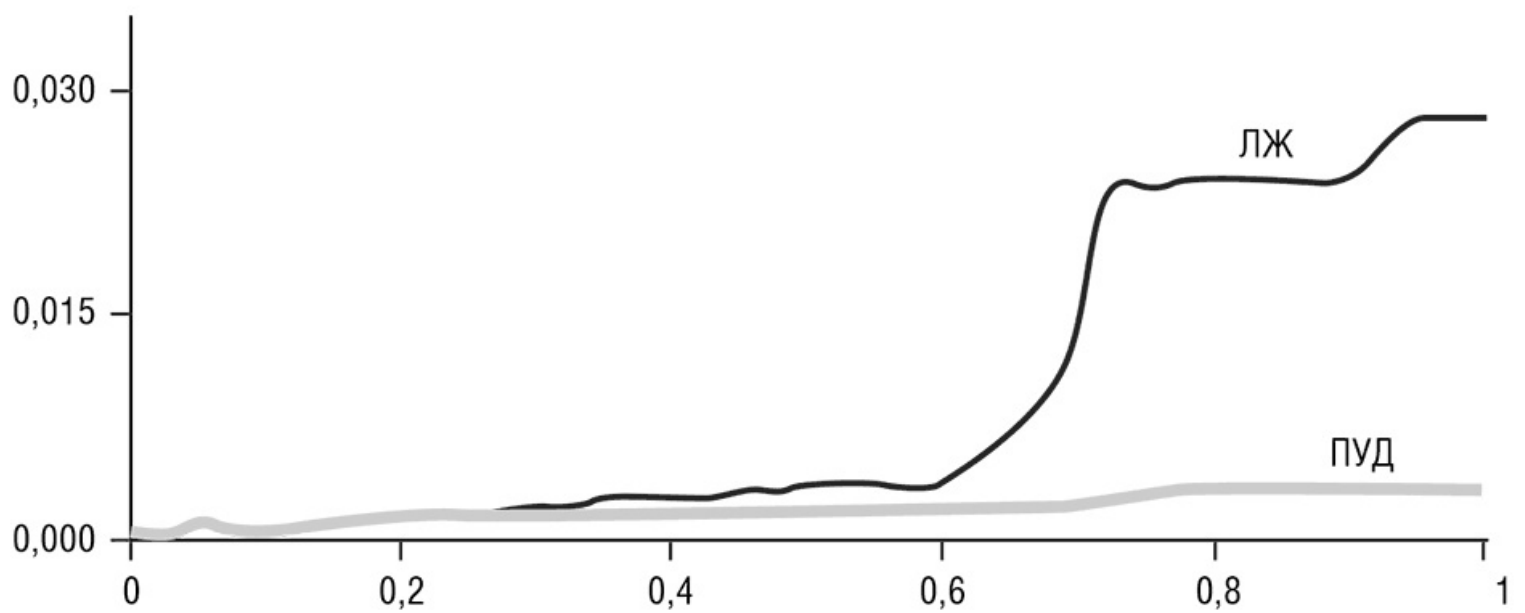


Рис. 6–6. Зависимость показателя несовместимости деревьев от филогенетической глубины леса жизни. Обратите внимание на резкий фазовый переход в графике для всего ЛЖ и значительно более гладкий переход на графике для ПУДов. Адаптировано из Puigbo et al., 2009.

Обнаружение фазового перехода предполагает реальную возможность того, что глубокие части ЛЖ лучше всего описываются моделью ББВ (заметим, что в современной космологии Большой взрыв в буквальном смысле считается фазовым переходом, как разъясняется в прил. II). Чтобы изучить эту возможность, мы разработали компьютерную модель эволюции, которая симулировала Большой взрыв (то есть полное перемешивание порядка расположения ветвей в деревьях) на различных филогенетических глубинах и воспроизводила графики, показанные на рис. 6–6, с различными уровнями дополнительного ГПГ. К большому нашему удивлению, нам не удалось найти комбинацию параметров (глубина Большого взрыва и уровень ГПГ), при которой воспроизвелся бы график, имеющий близкое сходство с рис. 6–6. Кривая, хорошо согласующаяся с эмпирическими наблюдениями, была получена только в симуляции без Большого взрыва в момент или после отделения бактериальных типов – а Большой взрыв (или любое другое событие), который предшествовал бы этому разделению, находится за пределами нашего «горизонта событий» в этом анализе. Таким образом, сопоставление деревьев в ЛЖ, по-видимому, лучше описывается моделью сжатого кладогенеза, хотя, учитывая сложность проблемы, дополнительный анализ определенно необходим.

Если придерживаться модели сжатого кладогенеза, мы должны заключить, что ПУДы и в самом деле представляют главную древообразующую тенденцию, которая сохранялась на протяжении всей эволюции клеточной жизни. Выражаясь более биологическими терминами, около ста белок-кодирующих генов, которые составляют трансляционную и ядро транскрипционной систем (вместе с универсальными рРНК и тРНК), эволюционировали в основном как единый ансамбль со времени последнего универсального общего предка (LUCA) всех форм клеточной жизни (см. гл. 11). Таким образом, эволюция этого набора генов является, вероятно, наилучшим возможным отражением истории организмов, которое можно получить из молекулярных филогений. Что касается переходных эпох в эволюции жизни, их, видимо, лучше всего можно описать как фазы сверхбыстрой, взрывообразной эволюции, которые были запущены затуханием предшествующего многообразия жизненных форм и жесточайшими «бутылочными горлышками» для немногих выживших (см. гл. 9).

Разделение эволюции прокариот на древовидный и сетевидный компоненты

Как мы видели в предыдущем разделе, сигнал древовидной эволюции, который можно определить как консенсусную топологию ПУДов, по-видимому, отражает главную тенденцию в ЛЖ и может быть прослежен во всем диапазоне филогенетических глубин, несмотря на существенный уровень ГПГ. И напротив, общую сумму всех эволюционных схем, которые оказываются несовместимыми с консенсусной топологией ПУДов, будь они вызваны ГПГ или другими процессами (такими как параллельные потери генов, которые тоже часты среди прокариот), можно обозначить как *сетевидный сигнал*. Мы разработали количественную меру, чтобы прямо оценить (по шкале от 0 до 1) вклады древовидных и сетевидных компонент в эволюционные расстояния между видами (Puigbo et al., 2010). Чем ниже показатель (то есть чем ближе он к случайно ожидаемому расстоянию, в предположении, что гены смешиваются свободно), тем более отношения между заданной парой видов определяются сетевидными эволюционными процессами. На карте «дерево – сеть» ПУДов доминировал древовидный сигнал (темная область на рис. 6–7 а): средний показатель для ПУДов составил 0,63, так что эволюция почти универсальных генов прокариот оказывается примерно на две трети древовидной. Исключениями являются радиорезистентная бактерия (*Deinococcus radiodurans*), проявившая главным образом сетевидные отношения с большинством архебактерий, и некоторые из бактериальных таксонов (*Thermotogae*, *Aquificales*, *Cyanobacteria*, *Actinobacteria*, *Chloroflexi*, *Firmicutes* и *Fusobacteriae*), каждый из которых сформировал сильносвязанную сеть с другими бактериями (см. рис. 6–7 а).

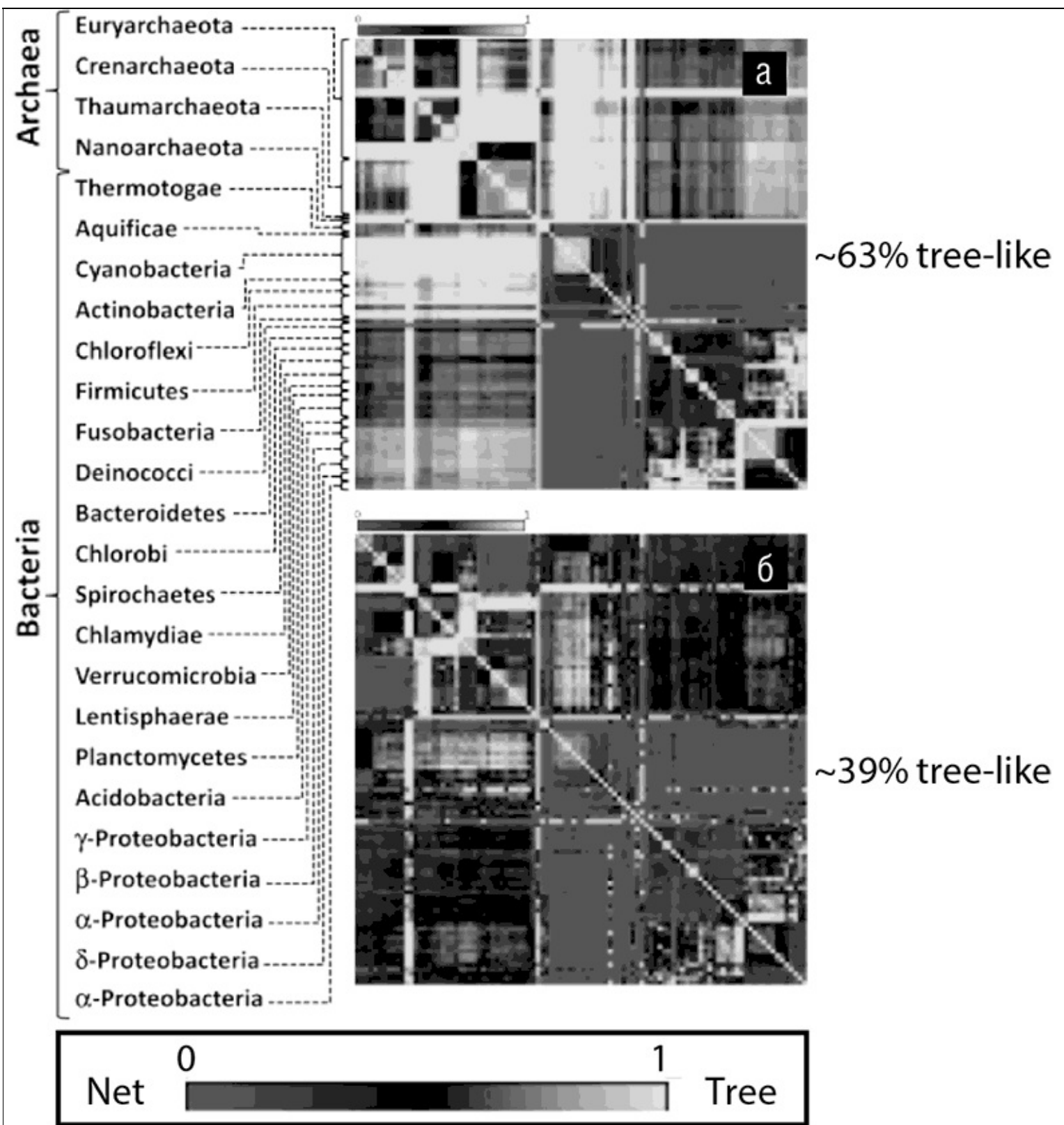


Рис. 6–7. Древоподобный и сетевидный сигналы в эволюции прокариот: а– 102 ПУДа; б– ЛЖ без ПУДов (6799 деревьев). Древоподобный сигнал усиливается от темной (сетевидная эволюция) к светлой (древоподобная эволюция) области. Виды расположены согласно топологии супердрева 102 ПУДов, которое было взято в качестве вертикального (древоподобного) сигнала. На рис. а отмечены крупнейшие группы археобактерий и бактерий. Адаптировано с разрешением из Puigbo et al., 2010.

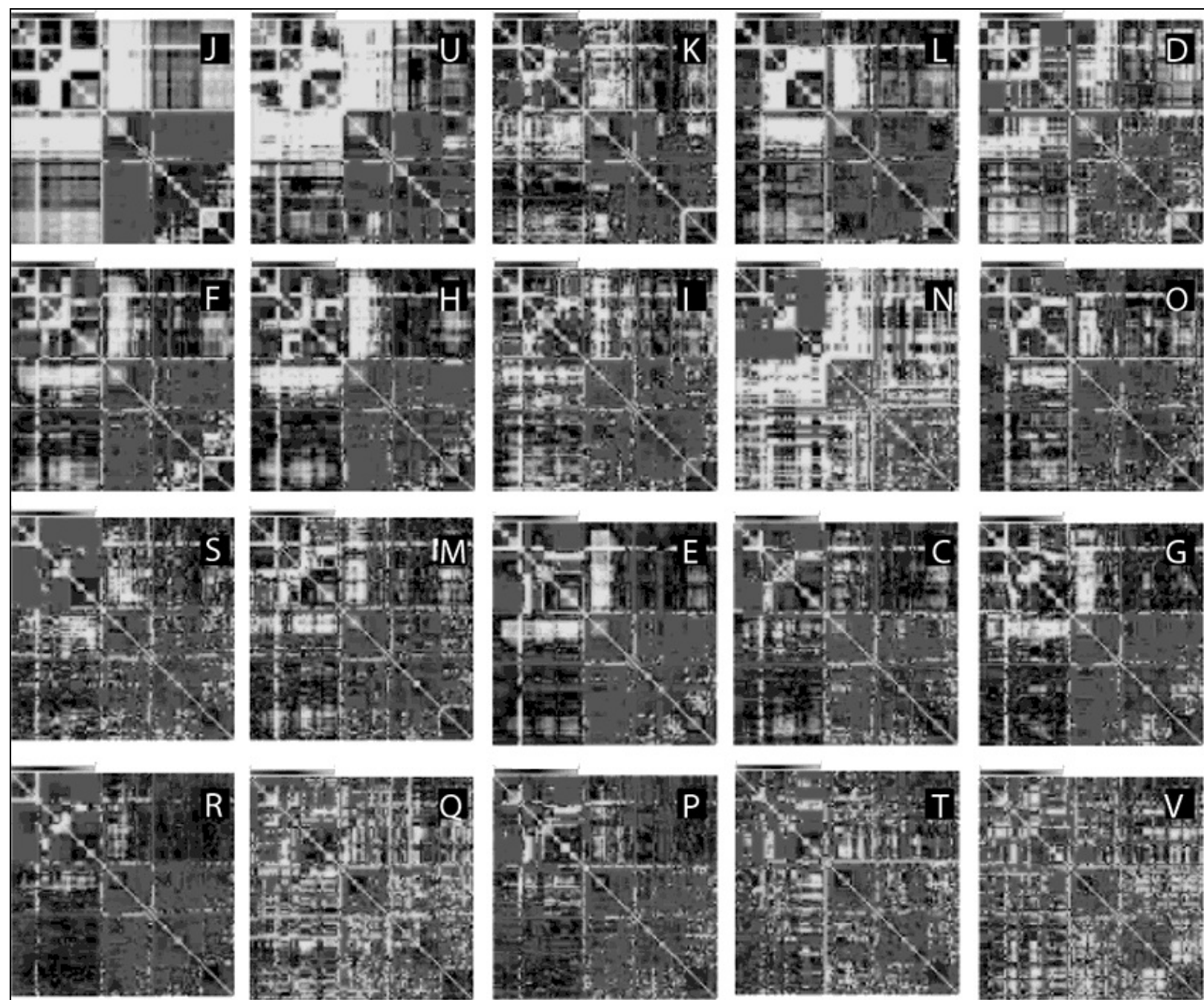


Рис. 6–8. Сигналы древовидной и сетевидной эволюции для различных функциональных классов прокариотических генов. Порядок и нумерация видов как на рис. 6–7. Функциональная классификация генов взята из системы КОГ (Tatusov et al., 2003). Обозначения: *J*– трансляция, рибосомная структура и биогенез; *U*– внутриклеточный обмен, секреция и везикулярный транспорт; *K*– транскрипция; *L*– репликация, рекомбинация и репарация; *D*– контроль клеточного цикла, деление клетки и разделение хромосом; *F*– транспорт и метаболизм нуклеотидов; *H*– транспорт и метаболизм коферментов; *I*– транспорт и метаболизм липидов; *N*– подвижность клетки; *O*– посттрансляционные модификации, белковый обмен и шапероны; *S*– функция неизвестна; *M*– биогенез клеточной стенки, мембраны и оболочки; *E*– транспорт и метаболизм аминокислот; *C*– производство и преобразование энергии; *G*– транспорт и метаболизм углеводов; *R*– только общее предположение о функции; *Q*– биосинтез вторичных метаболитов, транспорт и катаболизм; *P*– транспорт и метаболизм неорганических ионов; *T*– механизмы трансдукции сигнала; *V*– защитные механизмы. Адаптировано с разрешением из Puigbo et al., 2010.

В разительном контрасте с ПУДами, в остальной части ЛЖ доминирует сетевидная эволюция со средним показателем 0,39 (примерно на 60 процентов сетевидный сигнал).

Примечательно, что области древовидной эволюции перемешаны с областями сетевидной эволюции в различных частях ЛЖ (см. рис. 6–7 б). Крупные сетевидные области, которые мы наблюдали среди ПУДов, снова возникают в ЛЖ, однако проявляются и дополнительные подобные области, включающие кренархеоты, которые проявили выраженный сигнал недревовидных отношений с различными бактериями, так же как и некоторые эвриархеоты (см. рис. 6–7 б). Более подробный анализ ЛЖ показывает, что сетевидный сигнал доминирует в эволюции генов, которые присутствуют в небольшом количестве прокариот, в то время как эволюция более распространенных генов более древовидна и сильнее похожа на картину, наблюдаемую среди ПУДов (Puigbo et al., 2010). Эта тенденция очевидно совместима с гипотезой оптимизации ГПГ (см. гл. 5), согласно которой гены, часто теряющиеся в ходе эволюции, должны так же часто передаваться, чтобы избежать исчезновения этих генов и мутационного краха микробной популяции (см. гл. 5).

Различные функциональные классы генов проявили значительные различия в отношении древовидных и сетевидных тенденций в своей эволюции, от доминирования древовидного сигнала среди генов для компонентов механизма трансляции и молекулярных шаперонов до практически полностью сетевидной эволюции генов, кодирующих компоненты систем ионного переноса, передачи сигнала и защитных систем (см. рис. 6–8). Такая схема в целом совместима с гипотезой сложности, но, кроме того, выявляет более тонкую картину, с существенными различиями между, например, ферментами метаболизма нуклеотидов, которые эволюционируют преимущественно древовидно, и белками, участвующими в метаболизме и переносе аминокислот или углеводов, у которых сетевидный сигнал куда более заметен (см. рис. 6–8).

Подводя итог, можно сказать, что количественный анализ древовидного и сетевидного сигналов выявляет несомненный парадокс эволюции прокариот: несмотря на то что древовидная эволюция, безусловно, является сильнейшей тенденцией в ЛЖ, количественно в эволюции прокариот доминирует комбинация сетевидных процессов, таких как ГПГ и специфичная для линии утрата генов. Древовидный процесс отражает значительную часть эволюции среди ПУДов; однако, поскольку ЛЖ состоит преимущественно из небольших деревьев, среди которых древовидный сигнал обнаруживается с трудом, сетевидные процессы, которые управляют эволюцией относительно небольших семейств генов, доминируют количественно.

Древовидная эволюция или неслучайный горизонтальный перенос генов?

Гогартен, Лоуренс и Дулитл предложили еретическую (и весьма своеобразную) гипотезу, чтобы объяснить древовидные сигналы, которые можно наблюдать в филогенетическом анализе отдельных генов или ансамблей генов (Gogarten et al., 2002). Согласно этому предположению, древовидная картина эволюции может на самом деле быть последствием (можно сказать, в несколько провокационной манере, артефактом) неоднородного, неслучайного ГПГ, при котором организмы, «близкие» друг к другу на филогенетическом древе, обмениваются генами часто, а среди организмов, «далеких» друг от друга, ГПГ происходит редко. Как мы уже показывали в главе 5, эта возможность определенно имеет биологический смысл: учитывая, что ГПГ переносит ген в чужеродную внутриклеточную среду, можно ожидать, что чем меньше эта среда отличается от исходной (источника перенесенного гена), тем выше шансы для этого перенесенного гена прижиться. Для этой догадки уже даже существуют некоторые экспериментальные подтверждения, хотя и не систематические (Diaz et al., 2011).

Мы использовали структуру ЛЖ, чтобы симулировать эволюцию с переменными градиентами уровня ГПГ, нисходящими от близких к отдаленным организмам, а также чтобы оценить возможность того, что наблюдаемая нами древовидная картина была простым последствием неслучайного ГПГ. В каждой серии симуляций мы проверяли, могут ли характеристики ПУДов, которые мы наблюдали (такие как среднее расстояние между деревьями и степень разделения между археями и бактериями), быть воспроизведены в различных моделях эволюции. Первая серия симуляций началась с топологии супердрева ПУДов, которую мы взяли в качестве репрезентативного сигнала древовидной эволюции, и заключалась в измерении характеристик итоговых деревьев в зависимости от величины градиента ГПГ. Мы и в самом деле обнаружили, что умеренный градиент ГПГ от листьев древа к его центру воспроизводил эмпирически наблюдаемые свойства ПУДов. Вторую серию симуляций мы начали с деревьев-«звезд», в допущении, что древовидная эволюция является статистической аномалией, и затем постепенно развивали градиент ГПГ, присваивая увеличенные уровни ГПГ случайно соединившимся ветвям. В этой симуляции нам не удалось воспроизвести наблюдаемые характеристики ПУДов, даже при экстремально высоких уровнях ГПГ. И хотя эти симуляции, несомненно, являются сверхупрощенными моделями эволюции, они наводят на мысль, что древовидная тенденция и неслучайный ГПГ сосуществуют в ходе эволюции прокариот. И в самом деле, высокий уровень ГПГ между организмами, чьи гены ядра близко связаны древовидной эволюцией, выливается в самоподкрепляющийся процесс, который поддерживает связанные кластеры прокариот на различных уровнях филогенетической глубины.

Краткий обзор и перспектива

Когда Дарвин ввел метафору ДЖ, свидетельства в пользу этой схемы основывались на наблюдениях за эволюцией животных. Однако он с немалой долей уверенности распространил древовидную модель эволюции на жизнь в целом. В узком смысле Дарвин был прав: никто не отрицает, что эволюция животных имеет древовидную структуру. Однако это не древо жизни, это всего лишь описание эволюции отдельной, сравнительно малочисленной, компактной группы эукариот. Распространение этой концепции на всю полноту клеточной жизни на Земле терпит крах из-за сложной сети обширного ГПГ, который чаще всего встречается среди прокариот, но также внес значительный вклад в эволюцию эукариот, в частности при эндосимбиозе (см. гл. 7).

Тем не менее, несмотря на недавно открытый сетевидный характер эволюции, метафора Дарвина отражает глубинную истину: деревья по-прежнему остаются естественным представлением историй отдельных генов, учитывая фундаментально бинарный характер репликации генов и существенно более низкую частоту внутригенной рекомбинации по сравнению с межгенной рекомбинацией на длинных эволюционных дистанциях. Таким образом, хотя ни одно дерево не может в полной мере представлять эволюцию полных геномов и соответствующих форм жизни, реалистичная картина эволюции обязательно сочетает деревья и сети. Эти компоненты можно обнаружить с помощью анализа леса жизни (ЛЖ), полного множества филогенетических деревьев отдельных генов.

Количественный анализ ЛЖ обнаруживает сложный ландшафт древовидной и сетевидной эволюции. Сигналы этих двух типов эволюции распределены чрезвычайно неслучайным образом среди различных групп прокариот и среди функциональных классов генов. В целом сетевидный сигнал количественно доминирует, и это открытие (почти буквально) поддерживает идеи «латеральной геномики» или «сети жизни». Эти результаты, бесспорно, несовместимы с представлением об эволюции прокариот как о ДЖ, украшенном тонкими, случайными «паутинками» ГПГ (Ge et al., 2005; Kunin et al., 2005). Однако древовидный сигнал, совместимый с консенсусной топологией ПУДов, также несомненно обнаруживается и является сильным; согласно нашим измерениям, до 40 процентов эволюции в мире прокариот следует этому образцу. *Решающая, хоть и в некотором смысле парадоксальная черта эволюции прокариот состоит в том, что, хотя сетевидные процессы доминируют количественно, самой сильной специфической тенденцией является древовидная эволюция, отраженная в консенсусной топологии ПУДов, которая также в значительной степени повторяет древо рРНК.* В принципе эту тенденцию можно назвать «статистическим» или «слабым» ДЖ, хотя я склонен считать, что такая терминология непродуктивна: истинным объектом филогеномики являются ЛЖ и эволюционные структуры, которые можно в нем разглядеть, а не иллюзорное древо жизни.

Древообразующая тенденция эволюции, по-видимому, связана с градиентом ГПГ от близкородственных (в смысле древовидной эволюции) к отдаленным формам жизни. Взаимодействие между древовидной эволюцией и неслучайным ГПГ может создать самоподкрепляющийся эволюционный процесс, который отвечает за внутреннюю устойчивость групп прокариот на различных уровнях филогенетической глубины.

В заключение этой главы мне придется повторить вывод предшествующей: хотя намеченные здесь подходы к количественному анализу ЛЖ информативны и иллюстративны, они конечно же не являются последним словом в методологии филогенетики. По-настоящему пригодный концептуальный аппарат и технический инструментарий для синхронного,

всестороннего анализа древовидных и сетевидных эволюционных процессов еще только предстоит разработать. Как только такие методы появятся, мы начнем распознавать истинную картину эволюции.

Рекомендуемая дополнительная литература

Baptiste, E., M. A. O'Malley, R. G. Beiko, M. Ereshefsky, J. P. Gogarten, L. Franklin-Hall, F. J. Lapointe, J. Duprű, T. Dagan, Y. Boucher, and W. Martin. (2009) Prokaryotic Evolution and the Tree of Life Are Two Different Things. *Biology Direct*: 34.

Еще одна работа на стыке философии и биологии с главным акцентом на том, что «уверенность, что прокариоты связаны посредством... дерева, стала теперь сильнее, чем подтверждающие ее факты».

Ciccarelli, F. D., T. Doerks, C. von Mering, C. J. Creevey, B. Snel, and P. Bork. (2006) Toward Automatic Reconstruction of a Highly Resolved Tree of Life. *Science*: 1,283—1,287.

Кульминация классической традиции в изучении ДЖ. Описана вычислительная процедура, которая автоматически генерирует древо жизни путем объединения выровненных последовательностей 31 универсально консервативного белка (естественно, все эти белки участвуют в процессе трансляции).

Dagan, T., and W. Martin. (2006) The Tree of One Percent. *Genome Biology*7: 118.

Энергичное опровержение «дивного нового ДЖ» Чикарелли и коллег: согласно Дагану и Мартину, дерево, представляющее примерно 1 процент генов в сравненных геномах, даже если и свободно от внутренних противоречий, в принципе не может представлять геномную эволюцию, а следовательно, не может считаться древом жизни.

Doolittle, W. F. (2009) The Practice of Classification and the Theory of Evolution, and What the Demise of Charles Darwin's Tree of Life Hypothesis Means for Both of Them. *Philosophical Transactions of the Royal Society of London Biological Sciences*364: 2,221—2,228.

Резюме этого краткого обзора полемики вокруг ДЖ настолько афористично, что достойно быть процитированным практически полностью: «Споры вокруг статуса дерева жизни (ДЖ) часто продолжаются без договоренности о том, чем же оно должно являться: иерархической схемой классификации, прослеживанием истории геномов и организмов или гипотезой об эволюционных процессах и тех паттернах, которые они могут порождать. Я утверждаю, что для Дарвина эта была гипотеза, ошибочность которой демонстрируется латеральным переносом генов в прокариотах».

Doolittle, W. F. (1999) Phylogenetic Classification and the Universal Tree. *Science*: 2,124—2,129.

Первый принципиальный пересмотр концепции ДЖ, основывающийся на открытии обширного ГПГ среди прокариот. Классический образ трехдомного ДЖ заменяется метафорой сети, также содержащей три основных ствола, но демонстрирующей также множественные горизонтальные связи – в таком большом количестве, что стволы едва различимы.

Doolittle, W. F., and E. Baptiste. (2007) Pattern Pluralism and the Tree of Life Hypothesis. *Proceedings of the National Academy of Sciences USA*104: 2,043—2,049.

Замечательный сплав философии и биологии. Работа иллюстрирует, как одна и та же (древовидная) структура возникает в различных процессах, некоторые из которых не древовидны вообще. Дулитл и Батист заключают: «Плюрализм структуры (признание того, что различные эволюционные модели и представления взаимоотношений будут присущи и истинны для разных таксонов или при разных масштабах или для разных целей) является привлекательной альтернативой донкихотской погоне за единственно истинным ДЖ».

Koonin, E. V., and Y. I. Wolf. (2009) The Fundamental Units, Processes, and Patterns of Evolution, and the Tree of Life Conundrum. *Biology Direct*:33.

Расширенный аргумент в пользу того, что древовидная структура является центральной в эволюции, выведенный из фундаментального процесса бинарной репликации. Ввиду того, что гомологичная рекомбинация внутри генов встречается значительно реже, чем между генами, эволюция отдельных генов оказывается в основе своей древовидной, а описание эволюции генома можно получить, разыскивая тенденции в «лесу» деревьев генов.

O'Malley, M. A., and Y. Boucher. (2005) Paradigm Change in Evolutionary Microbiology. *Studies in History and Philosophy of Biological and Biomedical Sciences* 36: 183–208.

Профессиональный философ и эволюционный биолог представляют классификацию постсовременных взглядов на ДЖ, от отрицания сколько-нибудь существенного вклада ГПГ до категорического отказа от «древесного мышления».

Puigbo, P., Y. I. Wolf, and E. V. Koonin. (2010) The Tree and Net Components of Prokaryote Evolution. *Genome Biology and Evolution* 2: 745–756.

Количественная классификация эволюции прокариот на древовидные и сетевидные компоненты, поддерживающая концепцию латеральной геномики, но также раскрывающая центральную древообразующую тенденцию в эволюции прокариот.

Puigbo, P., Y. I. Wolf, and E. V. Koonin. (2009) Search for a «Tree of Life» in the Thicket of the Phylogenetic Forest. *Journal of Biology* 8: 59.

Всестороннее сравнение топологий примерно 7000 филогенетических деревьев, которые составляют прокариотический ЛЖ. Анализ раскрывает объективную центральную тенденцию в ЛЖ и показывает, что модель сжатого кладогенеза для основных эволюционных переходов лучше совместима с фактами, чем модель биологического Большого взрыва.

Глава 7. Происхождение эукариот: эндосимбиоз, удивительная история интронов и исключительная важность единичных событий в эволюции

Организмы с большими, сложными клетками известны как эукариоты, что означает «обладающие настоящим ядром». Эти организмы включают три царства многоклеточных форм жизни: растения, бурые водоросли и животные, а также обширный набор одноклеточных форм, известных как протисты. Обычно эукариотические клетки крупнее прокариотических и обладают сложной внутриклеточной организацией с разнообразными ограниченными мембраной органеллами, включая ядро, давшее им имя, и митохондрии, произошедшие от эндосимбиотических бактерий. Таким образом, по любому разумному критерию, эукариотические клетки значительно более сложны, чем бактерии и археи. Возникновение такого уровня сложности – загадка эволюционной биологии, представляющая еще более интригующей вследствие особого, «личного» интереса. Ведь изучая происхождение эукариот, мы всматриваемся в исток *нашего* собственно происхождения.

Фундаментальные различия в клеточной организации между эукариотами и прокариотами являют собой, можно сказать, парадоксальное противопоставление особенностям организации генома, которых мы уже касались в главе 3. В то время как эукариотические клетки обладают гораздо более замысловатой, упорядоченной и сложной организацией, чем прокариотические клетки, геномы эукариот намного менее оптимизированы и содержат гораздо больше «мусора», чем геномы прокариот. Объяснение эволюционных основ этого явного парадокса представляет собой принципиальную проблему, и ее решение, вероятно, прольет свет также и на происхождение эукариот – или, точнее говоря, на происхождение клеточной организации эукариот. Эта проблема далека от решения и остается предметом жарких (иногда даже слишком жарких) споров. В данной главе мы рассмотрим эту загадку и, говоря шире, эволюционные связи между археями, бактериями и эукариотами настолько объективно и логично, насколько это возможно, чтобы понять, как могло возникнуть «эукариотическое состояние живого». В заключение этой дискуссии я надеюсь показать, что, хотя некоторые ключевые детали все еще требуют объяснения, контуры приемлемого специфического сценария возникновения эукариот становятся ясными, и этот сценарий согласуется по крайней мере с некоторыми необычными чертами геномов эукариот и удивительной сложностью эукариотической клеточной организации. Более того, в данной главе мы придем к выводу о том, что трехдоменная схема Вёзе не дает адекватной картины истории жизни, и обсудим некоторые общие следствия эволюционной истории эукариот, которые имеют непосредственное отношение к центральной теме этой книги – взаимодействию случайности и необходимости в эволюции.

Эукариотическая клетка, ее внутренняя архитектура и пропасть между прокариотической и эукариотической клеточной организацией

В этой книге нам нет никакой необходимости вдаваться в бесчисленные тонкие детали биологических структур. Однако для того, чтобы обсуждать происхождение эукариот (далее *эукариогенез*) предметно, мы должны полностью оценивать природу и глубину пропасти, разделяющей эукариотические и прокариотические клетки. Действительно, существует резкое различие по сложности организации клетки между эукариотами и прокариотами: типичная эукариотическая клетка примерно в тысячу раз больше по объему, чем обычная бактерия или архея, и обладает исключительно сложной внутриклеточной компартментализацией, которой нет даже у самых «продвинутых» прокариот. Ниже в этой главе мы кратко обсудим некоторые интересные исключения, такие как гигантские прокариотические клетки и прокариотические клетки, содержащие внутриклеточные компартменты. Тем не менее тщательное рассмотрение этих случаев подтверждает фундаментальную дихотомию эукариот и прокариот по клеточной организации.

Компартментализация эукариотических клеток основана на сложной, разносторонней системе внутренних мембран и актинтубулиновом цитоскелете. Удивительное следствие внутриклеточной компартментализации – это физическое отличие эукариотических клеток от прокариотических. У прокариот содержимое клетки является раствором, пусть и вязким, так что макромолекулы (белки и нуклеиновые кислоты) диффундируют более или менее свободно и достигают своих пунктов назначения в клетке в результате сочетания стохастических перемещений и переноса в составе специфических комплексов. Напротив, у эукариот макромолекулы не могут свободно диффундировать и вместо этого добиваются до своего «места работы» при помощи сложных транспортных систем. Это различие ясно демонстрируется простым экспериментом, в котором мембраны прокариотических и эукариотических клеток искусственно делаются проницаемыми (пермеабелизуются): белки и нуклеиновые кислоты «вытекают» из пермеабелизованных бактериальных клеток, но, как правило, не из эукариотических (Hudder et al., 2003). Таким образом, цитозоль эукариотических клеток обладает значительно более низкой энтропией, чем у прокариот, – трудно придумать более фундаментальное различие.

Ядро – органелла, давшая название эукариотам, – содержит в себе геномную ДНК, организованную в виде хроматина и распределенную между несколькими хромосомами; это место транскрипции, сплайсинга и сборки рибосом. Само по себе ядро является частью системы внутренних мембран: оболочка ядра переходит в мембраны эндоплазматического ретикулума. Очевидно, что для функционирования эукариотической клетки ядро должно постоянно взаимодействовать с цитозолем. Действительно, оболочка ядра пронизана пораами, чрезвычайно сложными структурами, отвечающими как за пассивный, так и за активный перенос всех видов молекул (и даже макромолекулярных комплексов, таких как субъединицы рибосом) в ядро и из него. Заметим, что ограничение расположения хроматина и транскрипции ядерным компартментом исключает сопряжение транскрипции и трансляции, отличительную черту экспрессии генов у прокариот. Ниже в этой главе мы рассмотрим фундаментальные последствия этого разобщения.

Изнутри ядро заполнено высокоструктурированным матриксом и в этом отношении походит на эукариотический цитозоль. Эукариотический хроматин, содержащийся в ядре,

никоим образом нельзя считать просто молекулой ДНК, защищенной белками и регулярным образом упакованной в трехмерные структуры. Хроматин является чрезвычайно сложной динамической системой молекулярных машин, состоящих из множества специализированных белков, которые регулируют и координируют процессы репликации и экспрессии в основном посредством так называемого ремоделирования хроматина – модификации структуры хроматина, которая изменяет фактуру доступных участков (Clapier and Cairns, 2009). Хотя картина регуляции прокариотической экспрессии становится все сложнее и уже очень далека от простой схемы Жакоба – Моно (см. гл. 5), ничто в прокариотической клетке не может сравниться со сложностью эукариотического хроматина.

Качественные различия между эукариотами и прокариотами многочисленны и охватывают многообразные аспекты клеточной биологии, в частности те, которые имеют отношение к переработке информации, передаче сигналов и внутриклеточному переносу веществ (см. табл. 7–1). Сложность физической организации эукариотической клетки дополняется чрезвычайно изощренной сетью взаимодействующих сигнальных путей. Основные сигнальные системы эукариот – это киназно-фосфатазный механизм, регулирующий работу белков посредством фосфорилирования и дефосфорилирования; система убиквитина, которая управляет оборотом и локализацией белков посредством их обратимого убиквитинирования; регуляция трансляции посредством микроРНК; и регуляция транскрипции на уровне индивидуальных генов и ремоделирования хроматина.

Таблица 7–1. Краткое сравнение основных структурных и функциональных признаков эукариотических и прокариотических клеток.

Черта/система	Эукариоты	Прокариоты
Эндосимбиоз	<p>Митохондрии и родственные органеллы-эндосимбионты у всех эукариот.</p> <p>Пластиды у растений и многих <i>Chromalveolata</i></p>	<p>Чрезвычайно редкие наблюдения бактериальных эндосимбионтов в бактериях-хозяевах</p>
Внутриклеточные мембраны/компартаментализация	<p>Развитая система внутренних мембран: эндоплазматический ретикулум, комплекс Гольджи.</p> <p>Органеллы, ограниченные мембраной: ядро, вакуоли, пероксисомы.</p> <p>Полностью компартиментализованный цитозоль</p>	<p>Обычно нет ограниченных мембраной органелл.</p> <p>Относительно простые системы внутренних мембран и внутриклеточной компартиментализации в некоторых группах (<i>Verrucomicrobia-Planctomycetes</i>) и специализированных клеточных формах (споры, гетероцисты цианобактерий)</p>

Черта/система	Эукариоты	Прокариоты
Организация хроматина	<p>Высокая сложность организации хроматина, наличие нуклеосом.</p> <p>Сотни ассоциированных белковых комплексов и различных модификаций. Множественные линейные хромосомы</p>	<p>Относительно простая организация хроматина.</p> <p>Обычно одна или несколько кольцевидных хромосом</p>
Цитоскелет	<p>Сложный цитоскелет, состоящий из построенных на основе тубулина микротрубочек и актиновых филаментов, которые взаимодействуют с многочисленными белковыми комплексами и вспомогательными белками цитоскелета</p>	<p>Временные структуры, подобные FtsZ-кольцу, которые формируются во время деления клетки.</p> <p>Микротрубочки у <i>Prostheco bacteria</i>, содержащие тубулин, вероятно приобретенный путем ГПГ от эукариот¹.</p> <p>Возможно, актиновые филаменты у <i>Thermoproteales</i></p>
Сопряжение транскрипции и трансляции	Нет	Есть
Внутриклеточный транспорт белков и нуклеиновых кислот	<p>Главным образом высокоорганизованный транспорт, опосредуемый цитоскелетом и системой внутренних мембран</p>	<p>Главным образом свободная диффузия</p>
Клеточная стенка	<p>У большинства эукариот отсутствует.</p> <p>Клеточные стенки из целлюлозы у растений и из хитина у грибов</p>	<p>Клеточные стенки из пептидогликана у большинства бактерий.</p> <p>Протеинизированные клеточные стенки (S-слои) у архей.</p> <p>Множество производных форм с отсутствующей клеточной стенкой</p>

¹ Совсем недавно тубулин, вероятно являющийся эволюционным предшественником тубулинов эукариот, обнаружен также у архей, принадлежащих к типу Thaumarchaeota. (Yutin N, Koonin EV. Archaeal origin of tubulin. *Biol Direct*. 2012 Mar 29;7:10).

В главе 3 мы обсуждали некоторые из основных отличий в устройстве генома прокариот и эукариот. Далее в настоящей главе мы подробнее рассмотрим эволюцию и возможные источники одной из самых удивительных характерных черт эукариот, экзон-интронной структуры генов. Заметим, что различия проявляются на всех уровнях организации генома, от очевидных признаков, например разделения генома на множество линейных хромосом, до таких тонких деталей, как размер и строение нетранслируемых областей в белок-кодирующих генах (см. табл. 7–1).

У архей или бактерий нет ничего похожего на характерные для эукариот органеллы, особенности устройства генома и функциональные системы. Поэтому сама природа эволюционных связей между прокариотами и эукариотами кажется загадочной. В самом деле, сравнение полных геномных последовательностей однозначно показывает, что несколько тысяч эукариотических генов, отвечающих за ключевые функции клетки (трансляцию, транскрипцию и репликацию), происходят от общего предка с гомологичными генами архей и/или бактерий. Это эволюционное единство клеточных форм жизни делает объяснение того, как общие компоненты дают начало клеткам, столь непохожим в таком множестве черт, чрезвычайно сложной и интересной задачей.

В предыдущем разделе мы подчеркнули некоторые фундаментальные различия между прокариотическими и эукариотическими клетками (см. табл. 7–1). Одно из этих отличий представляется наиболее поразительным и может содержать ключ ко всей проблеме происхождения эукариот. Эта главная черта эукариотической клетки – присутствие митохондрий, играющих важнейшую роль в преобразовании энергии, а также выполняющих многие другие функции в клетках эукариот, такие как участие в различных формах передачи сигналов и программируемой клеточной смерти. Митохондрии – это органеллы характерной формы («дамская туфелька»), окруженные двойной мембраной; внутренняя мембрана содержит электронтранспортную цепь, состоящую из выстроенных в строго определенном порядке белковых комплексов. Удивительно, что митохондрии обладают своим собственным геномом, обычно представленным кольцевой молекулой ДНК, которая варьирует в размерах в разных царствах эукариот (очень небольшие, около 10 Кб, у животных, и более крупные, от 100 Кб до 1 Мб у других эукариот), и кодирующим небольшое число белков (только тринадцать у большинства животных; в основном это субъединицы комплексов электронтранспортной цепи), а также 34 рРНК и тРНК. Более крупные митохондриальные геномы растений, грибов и протист могут содержать больше функциональных генов – до ста у *Reclinomonas americana* из группы *Excavata*, – но в основном большие митохондриальные геномы состоят из встроенных мобильных элементов (Barbrook et al., 2010). Более того, митохондрии обладают собственными системами транскрипции и трансляции, которые обеспечивают экспрессию митохондриального генома. Эти системы во всех отношениях больше напоминают прокариотические, чем эукариотические аналоги. Многие эукариотические клетки содержат более одной митохондрии, и под электронным микроскопом это выглядит так, будто эукариотическая клетка нашпигована множеством паразитических или симбиотических бактерий. И на самом деле так оно и есть.

Сегодня у биологов нет сомнений, что митохондрии произошли от бактерий, бывших эндосимбионтами предков эукариот, и претерпели редуцированную эволюцию, которая превратила их в органеллы, полностью зависимые от клетки хозяина, однако сохранившие некоторые характерные прокариотические черты. Идентифицировать бактериальных предков митохондрий

было относительно нетрудно (Yang et al., 1985): филогенетический анализ митохондриальных рРНК и некоторых белок-кодирующих генов точно поместил их среди альфа-протеобактерий, представителей отдельной ветви *Proteobacteria*, которая, что интересно, включает, наряду с большим числом свободноживущих бактерий, некоторое число внутриклеточных паразитов (таких как *Rickettsia*) и эндосимбионтов (таких как *Wolbachia*). Таким образом, по крайней мере в общих чертах, путь от альфа-протеобактерий до митохондрий кажется ясным. Однако на молекулярном уровне это превращение ни в коей мере не тривиально. В самом деле, большинство митохондриальных геномов претерпели сокращение до крайнего минимума, и это сокращение сопровождалось переносом сотен бывших бактериальных генов в геном хозяина (пока что будем использовать нейтральное определение «хозяин эндосимбионта», но позже в этой главе мы обсудим природу этого хозяина подробно). Белковые продукты большинства этих генов – включая среди прочих все белки, которые составляют митохондриальную систему трансляции, – направляются обратно в митохондрии, где они и выполняют свои функции (см. рис. 7–1). Для того чтобы этот механизм работал, гены, перенесенные в хромосомы хозяина, должны быть транскрибированы, что требует соответствующих регуляторных сигналов; транскрипты должны быть транслированы в цитозоле, что требует полной совместимости с эукариотическими сигналами трансляции; наконец, полученные белки должны быть импортированы в митохондрию, что требует специальных сигналов для импорта и специализированного белкового механизма в наружной мембране митохондрий. Проблема приспособления перенесенных генов эндосимбионта к прохождению через этот сложный путь на первый взгляд кажется совершенно неразрешимой. Однако, по-видимому, имеется очевидное решение; пока что я сохраню интригу и расскажу о нем в разделе о происхождении эукариотической клетки, ниже в этой главе.

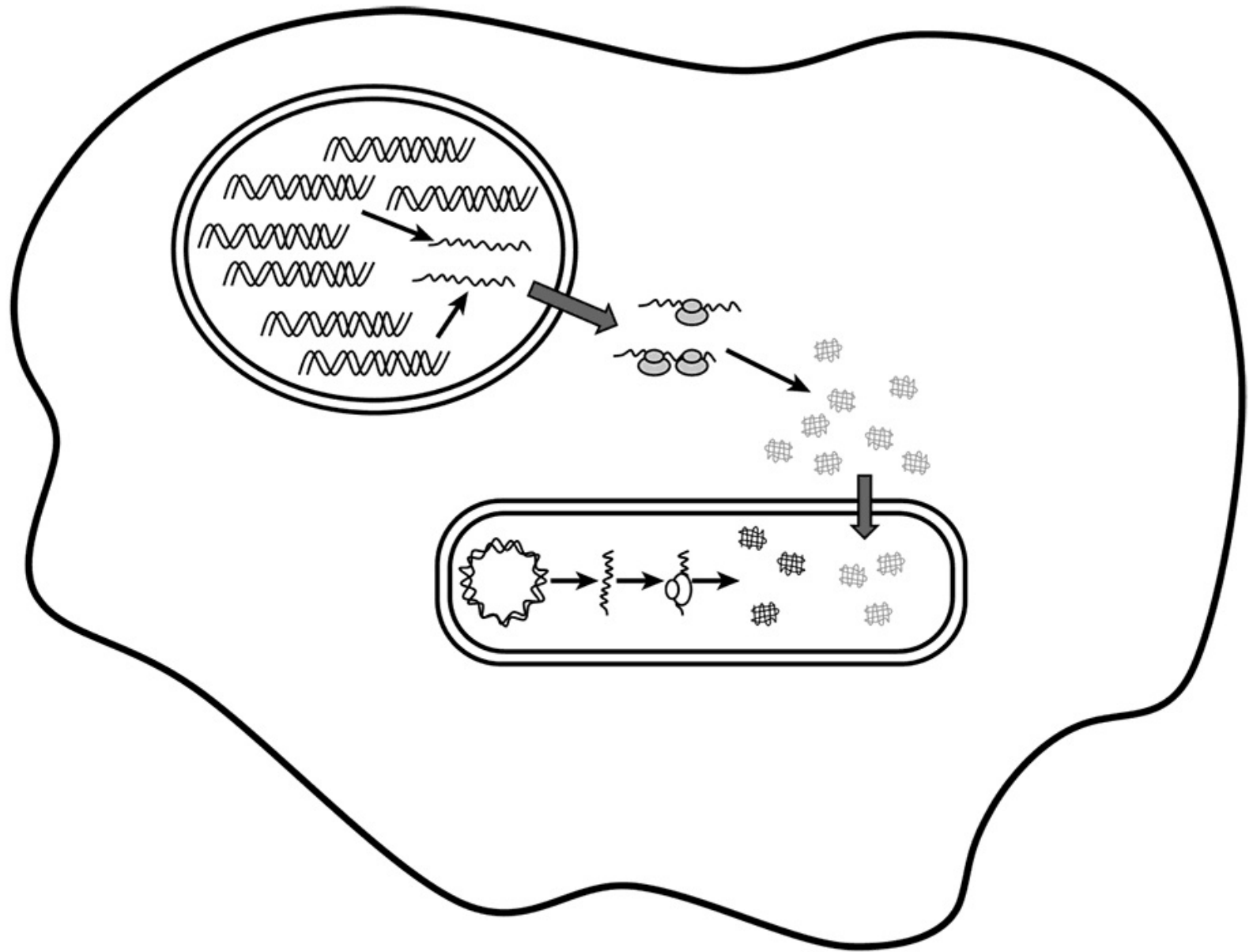


Рис. 7-1. Схематическое изображение митохондрии с ее геномом и системой трансляции, а также транспорта митохондриальных белков в эукариотической клетке.

Эндосимбиоз, митохондрии, гидрогеносомы и пластиды

Важнейшее, хотя и несколько недооцененное открытие первых лет XXI века состоит в том, что все эукариоты, изученные достаточно детально, обладают митохондриями или подобными им органеллами (Shiflett and Johnson, 2010; van der Giezen, 2009). Разнообразные одноклеточные эукариоты (протисты), такие как амёбы, микроспоридии, некоторые анаэробные грибы и различные *Excavata*, лишены типичных митохондрий и долгое время считались примитивными, первично амитохондриальными эукариотическими формами (часто их объединяли в группу *Archaezoa*). Однако недавние ультраструктурные исследования обнаружили ранее неизвестные миниатюрные органеллы, напоминающие митохондрии, во всех этих организмах. Эти протисты – анаэробы, так что эти органеллы, называемые гидрогеносомами, или митосомами, или просто митохондриеподобными органеллами (МПО), не участвуют в аэробном дыхании, подобно митохондриям. Однако все они обладают железо-серными кластерами, которые в митохондриальной электронтранспортной цепи являются главными каталитическими центрами. В МПО эти кластеры вместе с набором остальных ферментов катализируют другие, анаэробные окислительно-восстановительные реакции; в частности, один важный путь производит молекулярный водород, который используется метаболическими системами цитозоля. Несмотря на существенные отличия, МПО содержат некоторое количество белков, общих с типичными митохондриями. Они также используют механизм импорта белков, очень похожий на митохондриальный. Хотя эти миниатюрные МПО лишены генома и системы трансляции, которые всегда присутствуют в настоящих митохондриях, гены для нескольких характерных белков, общих для МПО и митохондрий, обнаружены в ядерных геномах соответствующих организмов. Рассматриваемые в совокупности, все эти факты не оставляют сомнений в том, что МПО представляют собой производные деградировавших митохондрий, которые, вероятно, потеряли свой геном при переходе соответствующих организмов к анаэробному образу жизни. По всей видимости, эта редукция митохондрий произошла в нескольких независимых случаях на протяжении эволюции эукариот. Отсюда следует важнейший вывод: *в настоящее время нам неизвестны амитохондриальные эукариоты*. Конечно, не исключено, что, пока я пишу эти строки, какие-нибудь архезои тихо размножаются, например, в маленьком пруду поблизости. Но с каждым новым эукариотическим организмом, у которого обнаруживаются митохондрии или МПО, это предположение становится все менее вероятным.

История эндосимбиоза у эукариот конечно же не ограничивается одними митохондриями. Вторым ключевым событием в истории эндосимбиоза было приобретение цианобактерий одноклеточным общим предком зеленых водорослей и наземных растений. Эти цианобактериальные эндосимбионты превратились в пластиды, которые в последующем разделились на хлоропласты и хромопласты. После цианобактериального эндосимбиоза ряд протист устроил настоящее буйство по части захвата зеленых водорослей и других обладающих пластидами эукариотических клеток. В результате возникали сложные эндосимбионты, состоящие из пластиды (возможно, это был фактор отбора для эволюционной фиксации эндосимбиоза) и остатка эукариотической клетки (часто называемого нуклеоморфом), первоначального хозяина пластид. Эндосимбиоз, по-видимому, является главным фактором расхождения *Chromalveolata*, одной из супергрупп эукариот (см. ниже) (Bhattacharya et al., 2007; Lane and Archibald, 2008).

В нескольких известных статьях, а также в исключительно смелой книге, написанной в соавторстве с Дорионом Саганом, Линн Маргулис, открывательница концепции эндосимбиоза в

ее современном воплощении, представила эндосимбиоз как единственный доминирующий эволюционный процесс у эукариот (Margulis, 2009; Margulis et al., 2006; Margulis and Sagan, 2003). Маргулис не только предположила, что, помимо митохондрий и пластид, некоторые органеллы, такие как центриоли и жгутики, возникли путем эндосимбиоза, но даже утверждала, что эндосимбиоз лежал в основе всего видообразования у эукариот. Однако, в противоположность случаям митохондрий и пластид, эти идеи имеют очень слабое эмпирическое подтверждение либо не имеют его вовсе. Известные случаи бактериального эндосимбиоза у эукариот достаточно многочисленны, но по большей части носят переходный характер, хотя есть и замечательные исключения, такие как долговременный мутуалистический эндосимбиоз у ряда насекомых (Gibson and Hunter, 2010).

Супергруппы эукариот и корень эукариотического эволюционного дерева

Как отмечалось в главе 6, древообразный эволюционный процесс гораздо лучше отражает эволюцию эукариот, чем прокариот. Главной причиной чего является частичное подавление неспецифического ГПГ, который преобладает в мире прокариот и который эукариоты заменили регулярным половым размножением (см. обсуждение ниже в этой главе). Однако принципиальная уместность метафоры дерева не означает, что корректное древо легко реконструировать. Некоторые царства эукариот, такие как животные, грибы, растения и цилиаты, хорошо определены и, без сомнения, являются монофилетическими; более того, эволюционные связи внутри них в основном соответствуют дереву с вполне определенной топологией. Однако расшифровка эволюционных связей между этими царствами и рядом других групп одноклеточных эукариот (протист) – тяжелая задача, а первичное расхождение от стадии последнего общего предка эукариот (*Last Eukaryote Common Ancestor, LECA*) представляет собой труднейшую проблему среди всех вопросов, связанных с эволюцией эукариот (Koonin, 2010a).

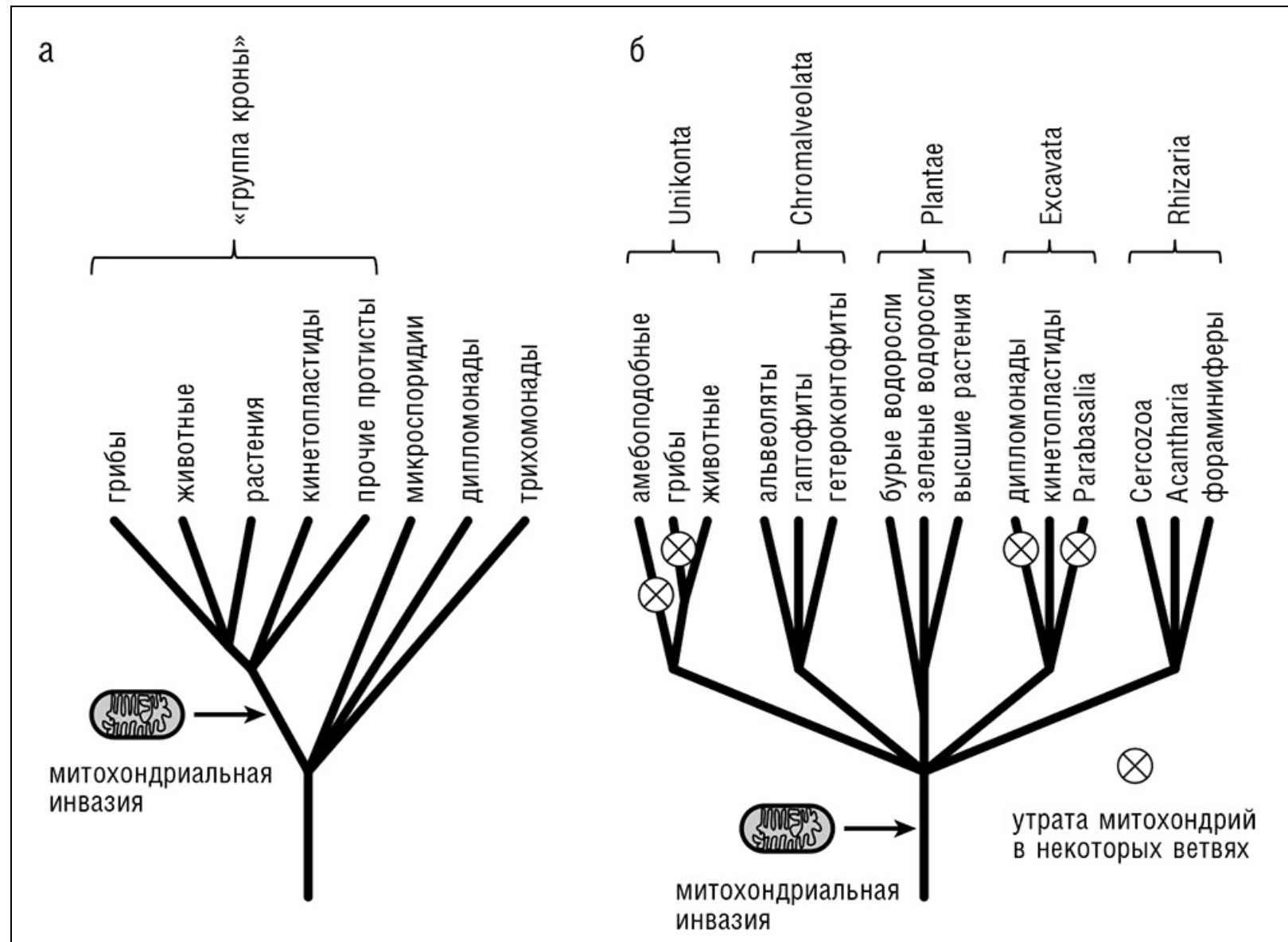


Рис. 7–2. Происхождение эукариот: а — гипотетическое архезойное древо; б — консенсусное «лучевое» древо, включающее пять супергрупп.

Проблема первичного расхождения эукариот связана с повсеместным присутствием митохондрий и МПО у эукариот, как мы уже обсуждали в предыдущем разделе. В течение многих лет большинство эволюционных биологов предпочитало так называемую филогению кроны (или «архезойное» древо), где древо эукариот состояло из «кроны», включающей животных (*Metazoa*), растения (*Viridiplantae*), грибы и некоторые группы протист, в зависимости от метода, использованного для построения древа (Cavalier-Smith, 1998; Patterson, 1999; Roger, 1999). Оставшиеся протисты, не имеющие типичных митохондрий, такие как *Microsporidia*, *Diplomonada* и *Parabasalina*, объединялись в группу *Archezoa* и считались «ранней ветвью эукариот», отделившейся от общего древа до того, как произошел митохондриальный эндосимбиоз (см. рис. 7–2 а). Такая топология древа эукариот была совместима с большинством филогенетических деревьев рРНК и различных консервативных белков. В корневых деревьях, полученных путем включения прокариотической «внешней группы» (outgroup), ветви архезой обычно оказываются в стороне от «кроны», как и следует ожидать, если корень эукариот действительно находится между архезоями и «короной». Однако в течение первого десятилетия XXI века архезойный сценарий расплылся по швам (Embley and Martin, 2006). Главной причиной этого было открытие митохондрий или МПО у всех современных эукариот, подорвавшее представления о раннем отделении «примитивных» протист от общего корня. Одновременно серьезно улучшенная статистика таксонов, появившаяся в результате секвенирования разнообразных геномов, вместе с новыми, более мощными методами филогенетического анализа, показала, что глубокое размещение «рано отделившихся» групп протист, видимое на многих деревьях, представляло собой артефакт притяжения длинных ветвей, обусловленный быстрой эволюцией соответствующих групп (Brinkmann and Philippe, 2007).

Таким образом, нет оснований считать какую-либо группу эукариот примитивными, пресимбиотическими архезоями. Вместо этого, принимая во внимание небольшие геномы и высокую скорость эволюции, характерную для тех групп протист, которые раньше считались рано отделившимися (архезоями), и их паразитический образ жизни, становится все яснее, что большинство, а может быть, и все они произошли от более сложных предковых форм путем редуktивной эволюции (Brinkmann and Philippe, 2007; Koornik, 2010a). Паразиты имеют тенденцию терять гены, органеллы и функции, которые обеспечивает им хозяин (редуктивная эволюция), а также быстро эволюционировать, вследствие непрерывной «гонки вооружений» с защитными системами хозяина (более подробно об этом, в применении к вирусам, будет сказано в гл. 10). Таким образом, архезойная филогения была фактически опровергнута, и изучение ранних стадий эволюции и происхождения эукариот пришлось начать с чистого листа (Embley and Martin, 2006).

В настоящее время филогеномный подход – то есть филогенетический анализ на уровне генома – использует большие наборы консервативных генов. Важнейшим достижением на этой новой стадии было обнаружение «супергрупп», каждая из которых сочетает весьма различные группы эукариотических организмов в монофилетическую кладу (Adl et al., 2005; Keeling, 2007; Keeling et al., 2005). Большинство опубликованных до сих пор результатов филогенетического анализа дают пять супергрупп (или шесть, если не объединять *Amoebozoa* и *Opisthokonta* в супергруппу *Unikonta*; см. рис. 7–2 б). Хотя демонстрация монофилии – нетривиальная задача для каждой из супергрупп, за возможным исключением растений, общая структура древа с небольшим числом супергрупп, образующих лучевую филогению, воспроизводится устойчиво, и последние исследования, по-видимому, подтверждают монофилетическое происхождение каждой из пяти супергрупп. Рассмотрение состава супергрупп может оказаться весьма поучительным и оказать самое радикальное воздействие на наше восприятие эукариот. Из пяти

супергрупп на «лучевом» древе (см. рис. 7–2 б) только три – *Unikonta*, *Plantae* и *Chromalveolata* (в последнем случае имеются в виду бурые водоросли) – включают сложные многоклеточные организмы, и даже в этих трех супергруппах многоклеточные организмы образуют только «кроны» (или только одну ветвь у *Chromalveolata*), тогда как несколько крупных ветвей в этих трех супергруппах и остальные две супергруппы состоят из протист. Хотя животные и растения конечно же наиболее заметные формы жизни, *эукариотическая жизнь в основном определяется огромным разнообразием одноклеточных форм*, тогда как видимые невооруженным глазом, большие многоклеточные существа суть лишь побочные ответвления в трех ветвях эукариот, образованных протистами.

Выяснение родственных связей между супергруппами представляет собой труднейшую проблему. Внутренние ветви чрезвычайно коротки, что означает быстрое (по меркам эволюции) расхождение супергрупп, возможно напомилавшее эволюционный Большой взрыв (см. гл. 6). В двух тщательных филогенетических исследованиях, в каждом из которых было проанализировано более 130 консервативных белков из нескольких десятков видов эукариот, после изучения эффекта исключения быстро эволюционирующих таксонов было получено филогенетическое древо эукариот, состоящее из трех мегагрупп (Burki et al., 2008; Hampl et al., 2009). Эти мегагруппы представлены *Unikonta*, *Excavata* и объединенной группой, составленной из *Plantae*, *Chromalveolata* и *Rhizaria* (см. рис. 7–2 б).

Предпринималось несколько попыток вывести корень филогенетического древа эукариот (см. рис. 7–2 б). Филогенетические подходы сами по себе не дают информации о корне, а использование прокариотических внешних групп приводит к потере разрешения метода, так что требуется независимая информация. Популярная идея заключается в том, чтобы попытаться идентифицировать так называемые *производные состояния признака* (синапоморфии), которые могут расщепить древо на два поддрева и таким образом установить положение корня. Проблема состоит в том, чтобы найти такие признаки, для которых вероятность независимого появления в двух и более линиях очень мала. Первое корневое дерево, являющееся альтернативой филогении группы кроны, было предложено Томом Кавалье-Смитом с соавторами (Richards and Cavalier-Smith, 2005; Stechmann and Cavalier-Smith, 2003). Эти исследователи использовали ясно различимые редкие геномные изменения (РГИ), такие как слияние генов, кодирующих два вездесущих фермента (дигидрофолатредуктазу и тимидилатсинтазу), а позже доменную архитектуру миозинов, и поместили корень древа между *Unikonta* и остальными эукариотами, *Heteroconta* (см. рис. 7–2 б). Это разделение выглядит приемлемым с биологической точки зрения, потому что клетки *Unikonta* имеют единственный жгутик, тогда как у всех остальных эукариотических клеток их два. Тем не менее в этом выводе можно усомниться, так как, используя небольшое число РГИ, трудно исключить параллельное возникновение тех же самых РГИ, таких как слияние или разрыв генов, в других линиях эволюции (феномен, известный под названием гомоплазии).

Игорь Рогозин и коллеги использовали другой РГИ-подход, основанный на редких замещениях высококонсервативных аминокислотных остатков, которые требуют двух нуклеотидных замен, и пришли к заключению, что наиболее вероятное положение корня древа – между растениями и остальными эукариотами (см. рис. 7–2 б; Rogozin et al., 2009). И снова представляется, что такая схема имеет биологический смысл, поскольку цианобактериальный эндосимбиоз, давший начало пластидам, произошел в линии растений и, согласно данному сценарию, мог оказаться тем событием, которое инициировало первичное расхождение эукариот. Несколько крупных ветвей *Chromalveolata* возникли в результате поглощения одноклеточных водорослей предковыми, беспластидным одноклеточным эукариотами (см. рис. 7–2 б) [\[61\]](#).

Другая возможная позиция корня эукариот следует непосредственно из результатов анализа митохондриальных геномов. Как отмечалось ранее в этой главе, представитель *Excavata Reclinomonas americana* имеет, несомненно, самый сложный из известных митохондриальных геномов, содержащий около ста функциональных генов, тогда как у остальных эукариот их менее двадцати. Можно было бы предположить, что *Reclinomonas* представляет собой самую раннюю ветвь эукариот, отделившуюся от ствола дерева эукариот до окончательной деградации генома эндосимбионтов. Данный сценарий поместил бы корень дерева в супергруппу *Excavata*. Однако имеется жизнеспособная и, возможно, более вероятная альтернатива: последний этап деградации митохондриального генома наступил уже после расхождения главных ветвей эукариот и продолжался конвергентными путями в разных линиях независимо. Последний сценарий подразумевает мощный эволюционный процесс, приводящий к потере (либо к переносу в геном хозяина) всех генов эндосимбионта, за исключением малого числа тех, которые обязательно должны остаться в митохондриальном геноме для сохранения жизнеспособности митохондрий; ниже в данной главе мы обсудим возможную природу этого процесса и требования к митохондриальному геному.

Отсутствие согласия относительно положения корня дерева и монофилии по крайней мере в некоторых супергруппах, не говоря о мегагруппах, указывает на то, что, несмотря на некоторые появляющиеся свидетельства, порядок первичных ветвлений в филогении эукариот в настоящее время остается неизвестным. В некотором смысле, если в раннем расхождении эукариот имел место Большой взрыв, то порядок ветвления супергрупп как таковой можно считать не имеющим особого значения. Однако биологические события, повлекшие это раннее расхождение, представляют огромный интерес, поэтому серьезные попытки разобраться в топологии самых глубоких ветвей эукариот, без сомнения, будут продолжаться с использованием более масштабных наборов данных и более совершенных методов.

Реконструкция LECA

Закрепление главных черт клеточной организации и, что более важно, существование большого набора генов, остающихся консервативными у всех или у большинства ранообразных эукариот, не оставляет сомнений, что все ныне существующие эукариоты произошли от одного общего предка (*Last Eukaryote Common Ancestor*, LECA). Как обсуждалось в начале этой главы, все эукариоты, которые были изучены достаточно детально, имеют либо митохондрии, либо МПО. Самое простое (самое экономное) заключение из этого положения вещей состоит в том, что LECA уже обладал митохондриями. И повторяю, правдоподобность этого вывода возрастает с каждой новой описанной группой эукариот, в которой обнаруживаются органеллы, подобные митохондриям.

Реконструкция эволюции геномного репертуара эукариот основана на тех же принципах и методах, что и реконструкция эволюции прокариот, обрисованная в главе 5, – в общем и целом, это принципы наибольшей экономии и наибольшего правдоподобия. Описывая суть этих подходов очень упрощенно, можно сказать, что происхождение генов, представленных у различных ныне живущих представителей главных линий эукариот, и даже предположительно потерянных в некоторых линиях, можно отследить вплоть до LECA. Результаты всех этих реконструкций согласованно указывают на сложность LECA как в отношении числа предковых генов, так и, что может быть даже более важно, в отношении наличия у предка типичных функциональных систем эукариотической клетки. Максимально экономные реконструкции, основанные на филетических паттернах в кластерах ортологичных генов эукариот, относят к LECA приблизительно 4100 генов (Koonin et al., 2004). Такие оценки очень консервативны, поскольку они не принимают во внимание главного аспекта эволюции эукариот – специфической для индивидуальных ветвей утраты предковых генов. Действительно, даже животные и растения, по-видимому наименее склонные к утрате генов эукариотические царства, явно утратили около 20 процентов предполагаемых предковых генов, идентифицированных у свободноживущего представителя *Excavata Naegleria gruberi* (Fritz-Laylin et al., 2010; Koonin, 2010b). Таким образом, эти реконструкции означают, что геном LECA был не менее сложным, чем геном типичных современных свободноживущих одноклеточных эукариот (Koonin, 2010a).

Этот вывод подтверждается сравнительно-геномными реконструкциями предковой композиции ключевых функциональных систем LECA, таких как поры ядра (Mans et al., 2004), сплайсосомы (Collins and Penny, 2005), аппарата РНК-интерференции (Shabalina and Koonin, 2008), системы передачи сигналов посредством убиквитина и протеасом (Hochstrasser, 2009), аппарата внутренних мембран (Field and Dacks, 2009) и аппарата деления клетки (Makarova et al., 2010). Итоги этих исследований ясны и согласуются между собой, даже когда в качестве шаблона для реконструкции используются различные топологии филогенетического древа эукариот: LECA уже обладал всеми этими структурами со всей их функциональностью, возможно столь же сложной, как и у современных эукариот.

Реконструкция других аспектов геномной композиции и архитектуры LECA также указывает на высокую сложность предкового генома. Сравнительный геномный анализ положения интронов в ортологических генах внутри супергрупп и между ними свидетельствует о высокой плотности интронов у предков супергрупп и у LECA, не меньшей, чем у современных свободноживущих одноклеточных эукариот, а вероятнее всего, близкой к богатым интронами

генам животных и растений (ниже в данной главе мы вернемся к примечательной истории эукариотических интронов более подробно).

Систематический анализ широко распространенных паралогичных генов эукариот указывает на то, что *LECA* предшествовали сотни дупликаций, особенно генов, вовлеченных в кругооборот белков, таких как молекулярные шапероны (Makarova et al., 2005). В итоге эти результаты ясно показывают, что *LECA* был типичной, полностью развитой эукариотической клеткой. В последующей эволюции эукариот не проявляется постоянной тенденции к повышению сложности клеток, за исключением специфических для ряда линий «украшательств», обнаруженных в группах многоклеточных (животных, растений и бурых водорослей), а также у некоторых протист, таких как зеленые водоросли или жгутиконосцы.

Стволовая фаза: темные века эволюции эукариот

Демонстрация того, что LECA обладал высокой сложностью клеточной организации, подразумевает наличие ключевой, стволовой (то есть предшествующей расхождению ныне существующих ветвей) фазы в эволюции эукариот (см. рис. 7–3), после появления LECA. Среди остальных эволюционных событий стволовая стадия включала широкомасштабную дупликацию ряда генов, так что набор предковых генов примерно удвоился (Makarova et al., 2005). Как долго продолжалась стволовая фаза в эволюции эукариот? С учетом того, что мы не осведомлены о каком-либо генетическом разнообразии до LECA, интуиция подсказывает, что эта фаза была очень краткой, откуда следует, что события между появлением первой эукариотической клетки и LECA развертывались в быстрой последовательности, возможно взрывообразно (см. рис. 7–3 а). Однако существует совершенно законная и логически корректная альтернатива: стволовая фаза была долгой и в ней происходили постоянные расхождения, но LECA (напомню, что это последний общий предок всех современных эукариот) – это тот, кто выжил в одной из линий, тогда как остальные вымерли (см. рис. 7–3 б). Некоторые из попыток датировать первичное расхождение эукариот – или, другими словами, оценить возраст LECA – дают результаты, согласующиеся с версией о долгой стволовой фазе.

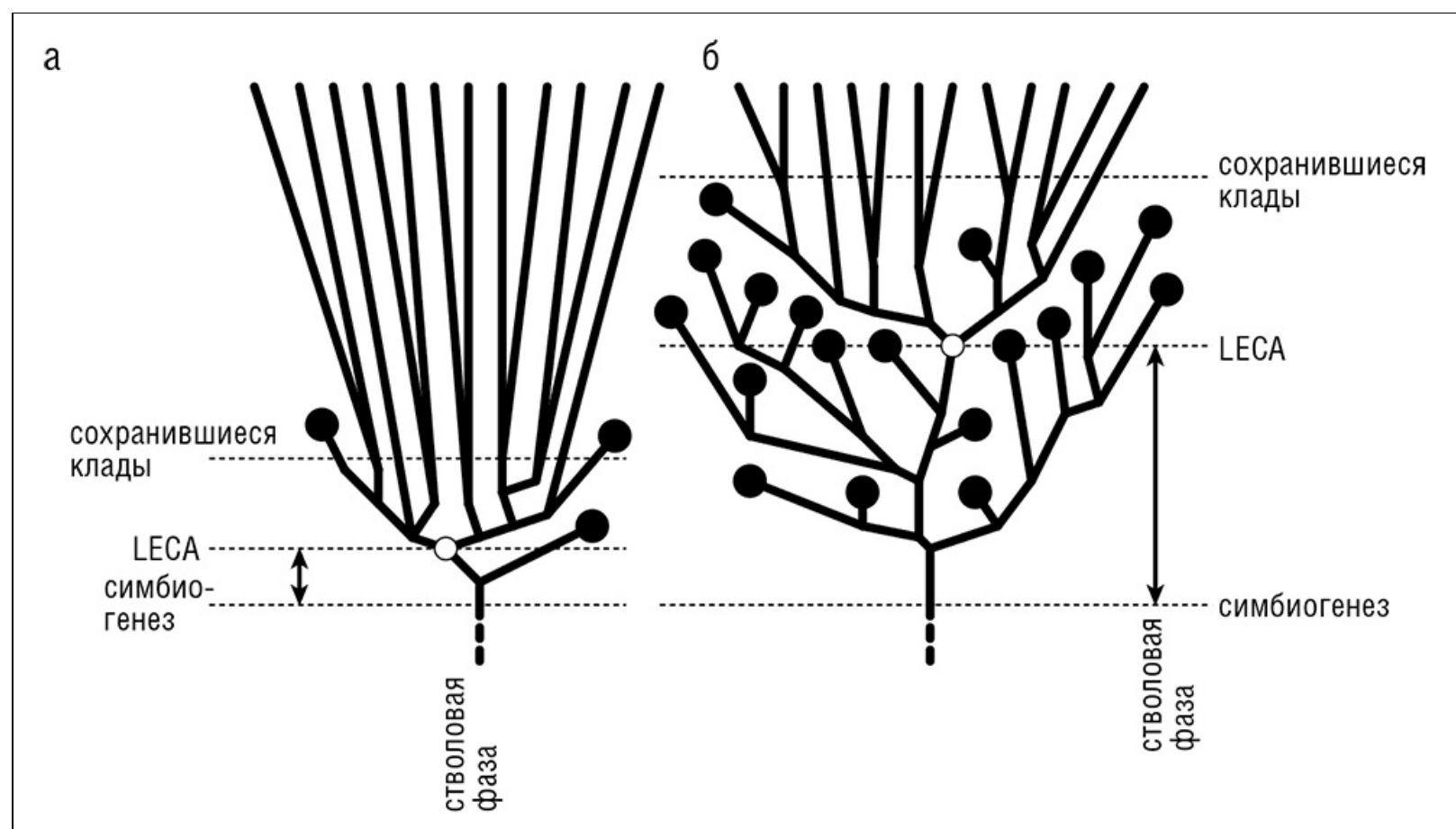


Рис. 7–3. Эволюция эукариот до и после LECA: а– сценарий «взрыва» с короткой стволовой фазой в эволюции; б– сценарий с продолжительной стволовой фазой ствола и значительным вымершим разнообразием, предшествовавшим LECA.

Молекулярная датировка событий эволюционного расхождения – весьма специализированная область исследований с множеством сложных технических проблем (Bromham and Penny, 2003; Graur and Martin, 2004). Нам придется пропустить основную часть

технических подробностей и перейти прямо к результатам. Заметим что принцип состоит в том, чтобы сопоставить молекулярные данные о дивергенции (то есть результаты сравнения последовательностей, привязанные к филогенетическому дереву) с палеонтологическими свидетельствами, используя нескольких точно датированных ископаемых в качестве калибровочных точек (например, самые ранние несомненные ископаемые млекопитающие датируются примерно 120 миллионами лет назад, так что это – самое позднее время начала распространения млекопитающих). Принимая строгие или нестрогие молекулярные часы, можно получить временную оценку для любого события дивергенции относительно взятых калибровочных точек и для данной топологии древа. Такие оценки могут быть достаточно достоверны, когда они представляют собой интерполяцию (то есть вывод о времени дивергенции внутри временного интервала, ограниченного калибровочными точками), но гораздо менее надежны, если получены путем экстраполяции (даты за пределами калибровочного интервала). К сожалению, для древних датировок, таких как возраст ЛЕСА, экстраполяция неизбежна. Временные оценки, полученные разными исследователями, охватывают чрезвычайно широкий диапазон дат между 1000 миллионами лет назад и 2300 миллионами лет назад. Несколько недавних, независимых и технически продвинутых исследований, использовавших нестрогие модели молекулярных часов или РГИ с похожим поведением в ходе эволюции, независимо привели к концепции «молодого ЛЕСА», которая помещает первичное расхождение эукариот в промежуток примерно от 1100 до 1300 миллионов лет назад (Chernikova et al., 2011; Douzery et al., 2004). Конечно, проблема не решена, но это, по-видимому, наилучшая из имеющихся на сегодня оценок возраста ЛЕСА. Из такой оценки следует долгая стволовая фаза в несколько сотен миллионов лет (см. рис. 7–3 б), поскольку не вызывающие сомнений окаменелости эукариот относятся ко времени более чем 1500 миллионов лет назад (Knoll et al., 2006).

Этот вывод заставляет серьезно переоценить наши современные знания о ранней эволюции эукариот. С одной стороны, результаты реконструкции, описывающие ЛЕСА как одноклеточного эукариота современного типа с полностью развитыми характерными функциональными системами эукариотической клетки, становятся менее неожиданными: действительно, похоже, что имелось достаточно времени для эволюции этих черт с момента появления (примитивной) эукариотической клетки. То же самое верно и для ряда генных дубликаций, отнесенных к ЛЕСА: согласно сценарию длинного ствола, они не произошли взрывообразно; было достаточно времени, чтобы гены дублировались постепенно. С другой стороны, стволовая фаза – это настоящие темные века в эволюции эукариот: мы не знаем о них почти ничего и можем надеяться узнать предельно мало. В самом деле, ЛЕСА по сути является «горизонтом событий» для сравнительной геномики: используя только сравнение геномов, мы не можем заглянуть в фазу ствола. Можно получить некоторое представление о том, что тогда происходило, тщательно изучая дубликации предковых эукариотических генов, но это практически единственный источник информации о стволовой фазе. Мы не имеем представления о реальном разнообразии эукариот темных веков, и надежда на то, что мы сможем оценить его в будущем, очень мала. Данные ископаемых выявляют некоторое разнообразие, но эта летопись никогда не будет полной, и трудно даже сказать, насколько она неполна. За немногими исключениям, ископаемые останки эукариот раннего и среднего протерозоя вряд ли представляют какой-нибудь из существующих ныне таксонов – наблюдение, которое следует интерпретировать осторожно, но которое в принципе согласуется со сценарием «молодого ЛЕСА» и идеей о существовании некоторого вымершего разнообразия, которое в настоящее время недоступно для наблюдения (см. рис. 7–3 б). Очевидна возможность, что ЛЕСА был тем организмом-«первооткрывателем», который захватил митохондриального

эндосимбионта, и что эндосимбиоз инициировал распространение дошедших до нас эукариот. Из этого сценария следует, что предки LECA, разнообразные эукариотические *Protozoa*, представляют вымершую, первично амитохондриальную биоту *Archezoa*. Однако есть другая, возможно более убедительная версия – что митохондриальный эндосимбиоз на самом деле дал начало появлению эукариот как таковому, так что эукариоты темных веков уже содержали митохондрии либо МПО. Мы обсудим эту дилемму и доводы в поддержку гипотезы об инициированном эндосимбиозом возникновении эукариот ниже в этой главе, после того как рассмотрим связи между эукариотами и прокариотами, обнаруженные путем сравнения геномов.

Корни эукариот среди архей и бактерий

Поиск архейного и бактериального «родителей» эукариот

Все эукариоты являются гибридными (химерными) организмами как в смысле клеточной организации, так и в отношении набора генов. Действительно, как отмечалось ранее в этой главе, все ныне существующие эукариоты, по-видимому, обладают митохондриями либо МПО, произошедшими от альфа-протеобактерий, а растения и многие группы *Chromalveolata* тому же содержат пластиды, произошедшие от цианобактерий. Набор генов у эукариот представляет собой разнородную смесь генов, вероятно происходящих от архей, генов с наиболее вероятным бактериальным происхождением и генов неизвестного происхождения, считающихся в настоящее время специфическими для эукариот. Может показаться парадоксальным, что, хотя филогенетические деревья, основанные на генах рРНК и соединенных последовательностях белков, участвующих в передаче информации, таких как полимеразы, рибосомальные белки и субъединицы сплайсосом, уверенно объединяют архей и эукариот, независимые сравнения полных геномов сходятся на том, что у эукариот в три и более раза больше генов с ближайшими бактериальными гомологами, чем с ближайшими архейными гомологами (см. рис. 7–4; Esser et al., 2004; Koonin et al., 2004; Makarova et al., 2005). Архейный набор значительно обогащен генами, связанными с функциями обработки информации (трансляция, транскрипция, репликация, сплайсинг), тогда как бактериальный набор кодирует в основном ферменты метаболизма, мембранные белки и компоненты биогенеза мембран, различные сигнальные молекулы и другие «операциональные» белки (подробнее см. ниже в этой главе).

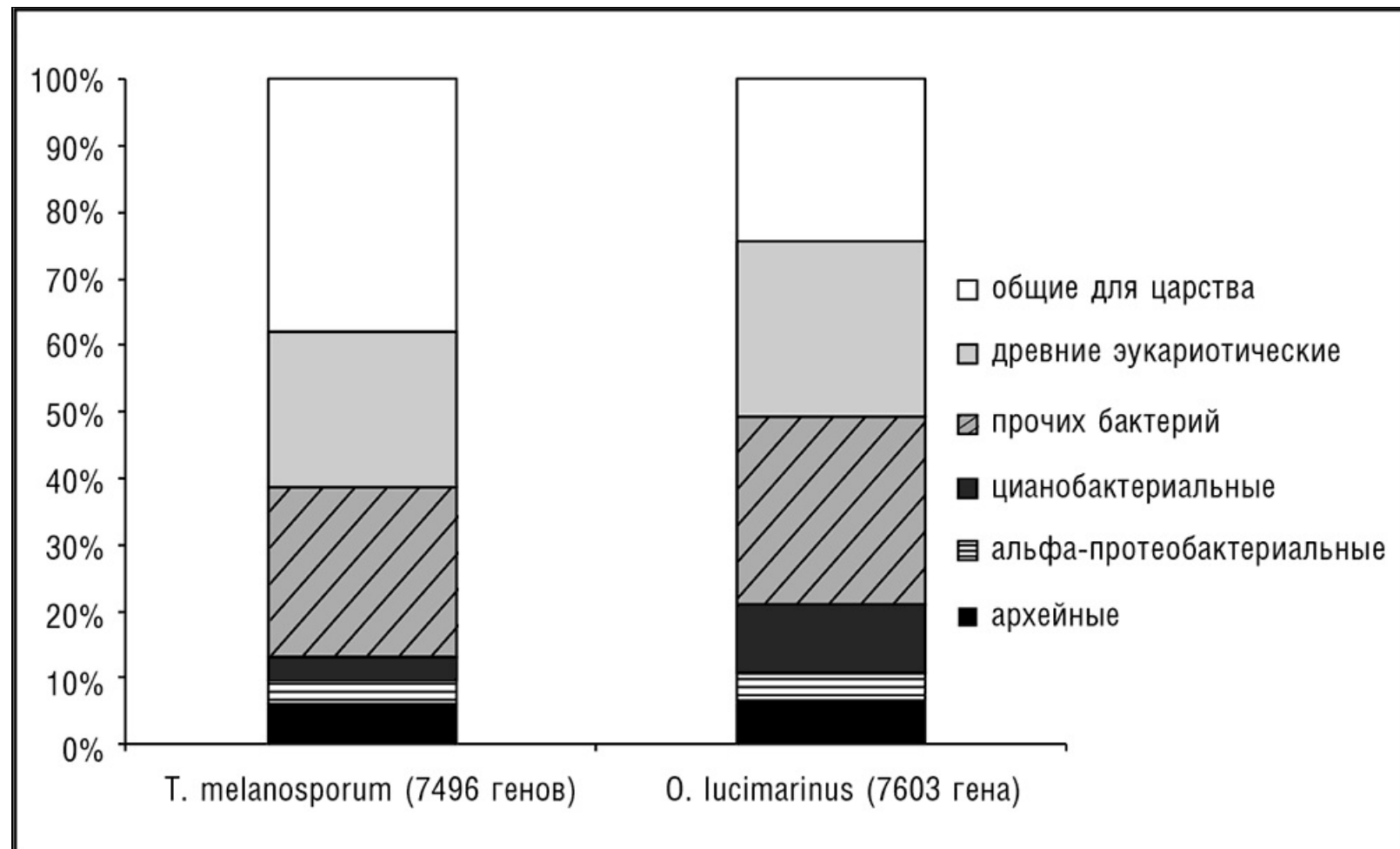


Рис. 7–4. Распределение генов двух дальнородственных эукариот, согласно их

предполагаемому происхождению: архейные, бактериальные либо специфичные для эукариот. Последовательности всех кодируемых белков гриба *Tuber melanosporum* (черный трюфель) и зеленой водоросли *Ostreococcus lucimarinus* сравнивались с базой данных NCBI RefSeq программой BLASTP (Altschul et al., 1997), а предполагаемое филогенетическое родство для каждого белок-кодирующего гена определялось с использованием специально написанной программы. Стоит отметить похожие, относительно небольшие фракции генов очевидно альфа-протеобактериального происхождения и более крупную фракцию цианобактериальных генов у водоросли.

В первом приближении эти наблюдения прекрасно согласуются со сценарием слияния геномов, согласно которому эукариотический геном появился путем слияния двух предковых геномов, архейного (или родственного археям) и бактериального (вероятнее всего, альфа-протеобактериального), учитывая хорошо установленное происхождение митохондриального эндосимбионта (Embley and Martin, 2006; Rivera and Lake, 2004). Слияние геномов проще всего интерпретировать как отражение симбиогенеза. Однако попытки точно указать конкретных архейных и бактериальных «родителей» не дают однозначных результатов и, по-видимому, свидетельствуют о сложных эволюционных связях. Хотя многие из «бактериальных» генов эукариот имеют альфа-протеобактериальные гомологи, они отнюдь не преобладают среди «бактериальных» генов, показывающих явное эволюционное родство с множеством групп бактерий (см. рис. 7–4). Важная причина этого сложного распределения бактериального компонента в наборе эукариотических генов – большой размер пангенома альфа-протеобактерий (см. гл. 5). Таким образом, не идентифицировав реальную альфа-протеобактерию, давшую начало эукариотическим митохондриям, трудно определить ее генетический вклад (Martin, 1999; Esser et al., 2007). Наряду с этой неопределенностью относительно набора генов эндосимбионта невозможно исключить многочисленные источники бактериальных генов, помимо альфа-протеобактериального эндосимбионта, давшего начало митохондриям. В частности, какова бы ни была истинная природа архейного предка эукариот, он, вероятно, жил при умеренных температурах и вообще не в экстремальных условиях и постоянно пребывал в контакте с разнообразным бактериальным сообществом. Современные археи с подобным образом жизни (например, *Methanosarcina*) имеют ряд генов, происходящих из различных бактериальных источников, что указывает на широкомасштабный горизонтальный перенос генов от бактерий (см. гл. 5). Таким образом, архейный хозяин эндосимбионта мог уже иметь много бактериальных генов, что отчасти объясняет наблюдаемое преобладание и разнообразие «бактериальных» генов у эукариот.

Идентификация архейного родителя эукариот даже более трудна, чем идентификация бактериального предка (предков), поскольку не имеется однозначных данных, определяющих предковую ветвь архей, в отличие от ветви альфа-протеобактерий, однозначно идентифицируемого источника митохондрий. Филогеномные исследования, использующие различные методы, указывают на ветви архей (*Crenarchaeota*, *Euryarchaeota*, *Thaumarchaeota* или неизвестная глубокая ветвь) как лучших кандидатов на роль предка эукариот (Cox et al., 2008; Kelly et al., 2010; Pisani et al., 2007; Yutin et al., 2008). Однозначная идентификация таких глубоких эволюционных отношений представляет огромные трудности. Более того, некоторые из этих исследований свидетельствуют о возможности того, что архейное наследие эукариот на самом деле является смешанным, с наибольшей долей генов, унаследованных от глубоких ветвей, за которой следует вклад *Crenarchaeota* (*Thaumarchaeota*) и *Euryarchaeota* (см. рис. 7–5; Yutin et al., 2008). Эти наблюдения наводят на мысль о том, что архейный родитель эукариот принадлежал к (возможно вымершей) глубокой ветви архей с

геномом высокой сложности (Makarova et al., 2010). Такое предположение хорошо согласуется с результатами сравнительно-геномных реконструкций, которые указывают на сложных архейных предков (Csuros and Miklos, 2009; Makarova et al., 2007b; см. обсуждение далее в этой главе).

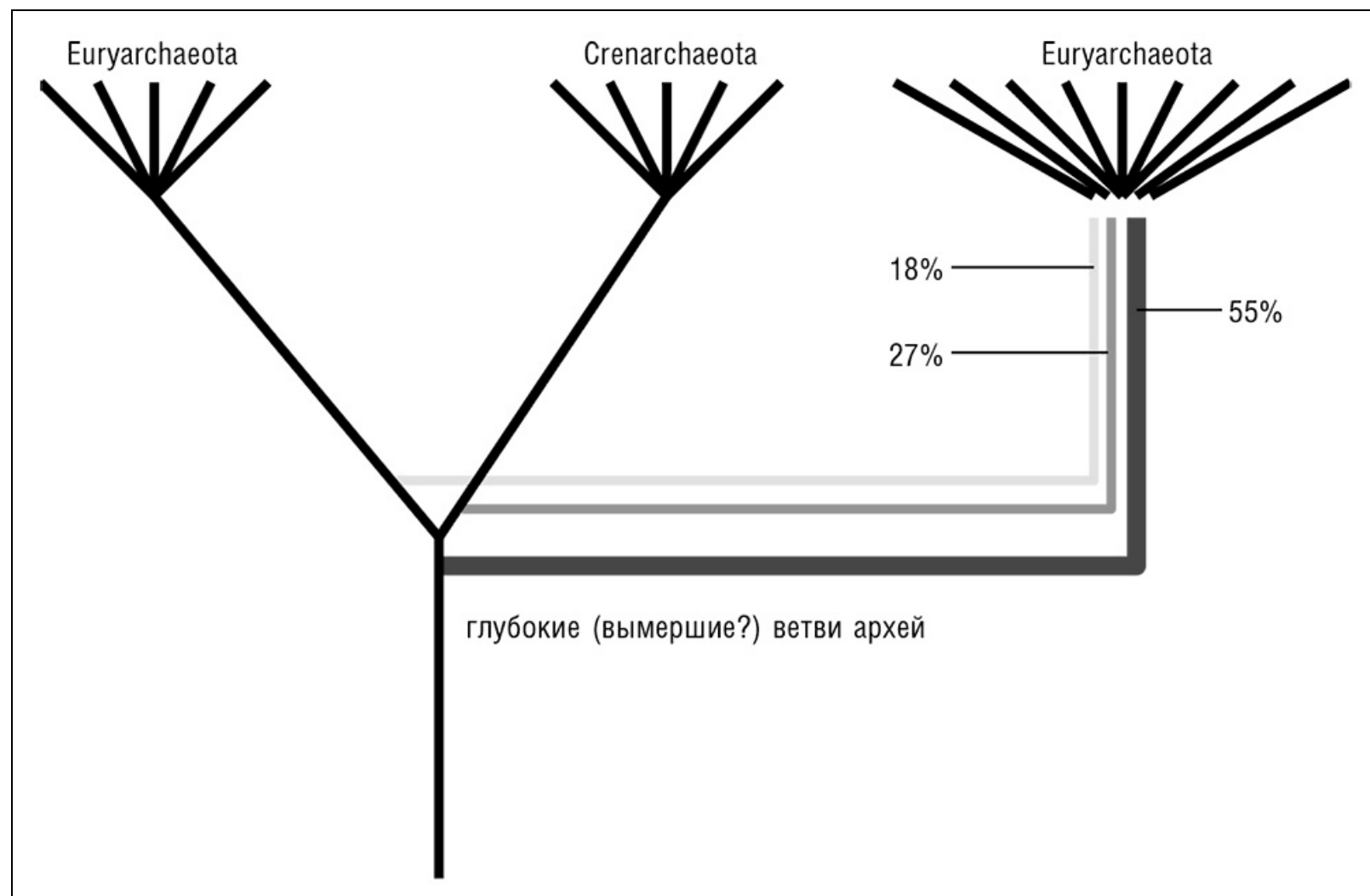


Рис. 7–5. Вклад различных групп архей в происхождение эукариот. Показан процент генов, унаследованных от архей и наиболее вероятно произошедших от *Euryarchaeota*, *Crenarchaeota* и глубоких ветвей. Данные по Yutin et al., 2008.

Происхождение ключевых функциональных систем эукариотической клетки

Другая важная тема, возникающая из сравнительного анализа геномов, – это вопрос о взаимоотношениях между вкладами архей и бактерий в происхождение специфичных для эукариот функциональных систем, и особенно о смешанном архейно-бактериальном происхождении некоторых из этих систем. В принципе существует два типа родственных связей между функциональными системами эукариот и прокариот (см. табл. 7–2):

1. Эукариотические системы, возникшие из гомологичных и функционально аналогичных систем прокариот.
2. Эукариотические системы, возникшие путем сборки из компонентов, которые у прокариот вовлечены в функционально иные, часто разные процессы, иногда вместе с дополнительными белками, которые, по-видимому, специфичны для эукариот.

Таблица 7–2. Предполагаемое происхождение некоторых ключевых функциональных

систем и молекулярных аппаратов эукариот.

Система/комплекс/ функция	Предполагаемое происхождение	Дубликации и другие сложные черты у эукариот
Тип 1: происхождение из функционально аналогичных прокариотических систем		
Аппарат репликации и репарации ДНК	Архейное; ДНК-полимеразы и другие центральные белки репликации происходят либо от <i>Crenarchaeota</i> , либо от <i>Euryarchaeota</i> . Смешанное архейное и бактериальное происхождение в случае ферментов репарации	Несколько ранних дубликаций ДНК полимеразы и других белков репликации
Аппарат транскрипции	Архейное; по крайней мере две субъединицы РНК-полимеразы происходят от <i>Crenarchaeota</i> / <i>Korarchaeota</i>	Дубликации, произошедшие до ЛЕСА, дали начало 3 РНК-полимеразам
Аппарат трансляции, включая рибосомы	Большей частью архейное; некоторые АРСазы замещены бактериальными гомологами	Малочисленные дубликации

Система/комплекс/ функция	Предполагаемое происхождение	Дубликации и другие сложные черты у эукариот
Протеасомы: регулируемый протеолиз	Архейное	Несколько дубликаций субъединиц и значительные структурные «украшения»
Убиквитинная сигнальная система: регулируемый протеолиз и топогенез белка	Архейное	Многочисленные дубликации на всем протяжении эволюции эукариот
Экзосома: регулируемая деградация РНК	Архейное	Многочисленные дубликации до LECA, множественные более поздние дубликации в некоторых линиях эволюции

Тип 2: сборка из различных прокариотических предковых источников

Комплекс ядерной поры: ядерно- цитоплазматический транспорт	Бактериальное; некоторые ключевые белки комплекса состоят из повторов и имеют неизвестное происхождение	Многочисленные дубликации до LECA и более поздние дубликации в некоторых линиях
Хроматин, нуклеосомы	Сложная смесь архейного и бактериального происхождения	Многочисленные дубликации, на всем протяжении эволюции эукариот (включая дубликацию гистонов до LECA), с добавлением специфических для эукариот белков
РНК-интерференция	Гибридное происхождение из архейных и бактериальных компонент	Многочисленные независимые дубликации в разных линиях

Система/комплекс/ функция	Предполагаемое происхождение	Дубликации и другие сложные черты у эукариот
Эндоплазматическая система, эндоплазматический ретикулум	Сложная смесь архейного и бактериального происхождения	Многочисленные независимые дубликации в разных линиях, на всем протяжении эволюции эукариот
Аппарат программированной клеточной смерти	Бактериальное	Многочисленные независимые дубликации в разных линиях

Многие из систем типа 1 долгое время рассматривали как эукариотические инновации. Однако имеет место примечательная тенденция: чем внимательнее мы изучаем быстро прибывающие данные сравнительной геномики архей и бактерий, тем больше обнаруживается эволюционных предшественников для характерных эукариотических систем. Например, у давно известного эукариотического аппарата деградации белков, протеасомы, имеется более простой гомолог с аналогичной функцией у архей, и даже более примитивная версия у бактерий (Groll et al., 2005). Бесспорная эволюционная связь и функциональная аналогия между протеасомами архей и эукариот была точно установлена даже раньше, чем появились последовательности геномов. Сравнительный анализ геномов параллельно привел к предсказанию экзосом у архей: были идентифицированы высококонсервативные наборы предполагаемых архейных оперонов, кодирующих белки, гомологичные субъединицам эукариотических экзосом, молекулярных машин, осуществляющих деградацию РНК у эукариот (Koonin et al., 2001b). Как и следовало ожидать, эта предсказанная архейная экзосома была экспериментально обнаружена через несколько лет (Hartung and Hopfner, 2009).

В течение долгого времени убиквитиновая сигнальная сеть, управляющая деградацией и топогенезом белков в эукариотической клетке посредством присоединения к белкам небольшого, чрезвычайно консервативного белка, называемого убиквитином (Ub), и значительно менее распространенных паралофов убиквитина, считалась типичной специфической для эукариот функциональной системой, уникальной отличительной чертой эукариот (Hochstrasser, 2009). Позже, благодаря растущему разнообразию секвенированных геномов архей и бактерий и прогрессирующим методам обнаружения сходства белков по аминокислотной последовательности и структуре, были обнаружены прокариотические гомологи убиквитина. Эти небольшие белки чрезвычайно распространены у архей, но, по-видимому, работают в реакциях, обеспечивающих встраивание серы, что необходимо для биосинтеза ряда коферментов. Однако тщательное сравнение геномов привело к открытию, у различных бактерий, оперонов, которые комбинируют гены гомологов убиквитина с генами гомологов двух субъединиц убиквитин-лигазы и деубиквитирующего фермента. Хотя эти

белки лишь отдаленно связаны со своими эукариотическими гомологами, колокализация всех этих генов наводила на мысль о вероятности того, что обнаружено бактериальное происхождение Ub-системы (Iyer et al., 2006). Далее в 2010 году появилось сообщение об экспериментах, демонстрирующих, что по крайней мере у некоторых *Archaea* определенная группа гомологов убиквитина функционирует подобно классическому эукариотическому убиквитину, то есть эти небольшие белки присоединяются к различным другим белкам и делают их мишенью для деградации (Humbard et al., 2010).

Убиквитиновая история этим не оканчивается. В декабре 2010 года, когда эта книга уже была практически завершена и редактировалась, было опубликовано удивительное открытие. Его помог сделать недавно секвенированный геном *Candidatus Caldiarchaeum subterraneum*, археи, которая была изолирована из золотоносного рудника и может оказаться представителем нового типа архей либо новой группы в составе *Crenarchaeota* (Nunoura et al., 2010). Геном этого организма содержит оперон, кодирующий четыре белка, которые, по результатам поиска по базам данных, представляются типично эукариотическими – для них обнаруживаются многочисленные высококонсервативные эукариотические гомологи, но не обнаруживается сходства с какими-либо белками из других архей или бактерий. Эти белки – убиквитин и три субъединицы убиквитин-лигазы (E1, E2 и E3). Более того, вслед за убиквитиновым опероном комплементарной цепью ДНК кодируется деубиквитирующий фермент эукариотического типа. Таким образом, этот новый архейный геном кодирует полный набор белков, требующихся для обратимого убиквитинования белков у эукариот. Интересно, что, когда я проводил дополнительный поиск по базе данных белковых последовательностей для того же геномного окружения, мне удалось идентифицировать еще одну субъединицу E3, так что даже пролиферация E3, которая достигает поразительных масштабов у эукариот, началась, по-видимому, уже у *Archaea*. Степень сходства между данными белками и их эукариотическими гомологами неожиданно высока (намного выше, чем для бактериальных белков, кодируемых в подобных оперонах), что наводит на мысль о необычной возможности горизонтального переноса генов от эукариот к археям. Однако это вряд ли самый экономный сценарий, если учитывать локализацию этих генов в одном опероне у *Caldiarchaeum subterraneum*. Остается заключить, что эта архея кодирует предковую убиквитиновую систему. Если это действительно так, мы будем вынуждены заключить, что эта система возникла и полностью сформировалась у *Archaea*, так что эукариоты получили ее в готовом виде, и все, что произошло с убиквитиновой системой в течение эволюции эукариот, сводится к возникновению разнообразия и «украшений». Удивительно, что потребовалось секвенировать более ста геномов архей для того, чтобы обнаружить эту предполагаемую предковую убиквитиновую систему; это показывает, что предковые версии некоторых ключевых эукариотических функциональных средств достаточно экзотичны среди *Archaea*. Я описываю это открытие настолько подробно не только по причине его несомненной важности для понимания происхождения убиквитиновой регуляторной сети, но даже в большей степени из-за его широких последствий для эволюции эукариот, которые я подчеркиваю ниже в этой главе.

Молекулярные машины и системы типа 1 в общем следуют главному направлению эволюции эукариот – а именно возникновению серийных дубликаций генов, а в дальнейшем множественных версий: там, где архейный комплекс состоит из множества копий одного или двух белков, более сложная эукариотическая версия вместо этого содержит множественные паралогичные субъединицы (Makarova et al., 2005; см. табл. 7–2).

В качестве примера систем типа 2 можно привести типичную эукариотическую молекулярную машину, комплекс ядерной поры, для которой у прокариот нет функциональных аналогов. Примечательно, что ничто не указывает на происхождение комплекса ядерной поры

от архей. Он построен из различных белков явно бактериального происхождения в комбинации с белками, состоящими из простых повторов, источник которых трудно установить (Mans et al., 2004). Напротив, аппарат РНК-интерференции, система антивирусной защиты (врожденный иммунитет) и система регуляции экспрессии у эукариот, привлекающая такое внимание в последнее десятилетие, отчасти благодаря исключительному значению в качестве экспериментального инструмента, явно демонстрируют химерное, архейно-бактериальное происхождение (Shabalina and Koonin, 2008). Например, один из ключевых белков РНК-интерференции, эндонуклеаза *Dicer*, состоит из двух доменов бактериальной РНКазы III и геликазного домена явно эуриархейного происхождения; другой важный белок РНК-интерференции, *Argonaute*, также демонстрирует родство с эуриархеями (Shabalina and Koonin, 2008). Еще одна характерная для эукариот молекулярная машина – сплайсосома – в какой-то степени является промежуточной между эукариотическими системами первого и второго типа (Collins and Penny, 2005). *Sm*-белки, составляющие сердцевину сплайсосомы, имеют легко идентифицируемые архейные ортологи, однако они участвуют не в сплайсинге, а в других видах реакций процессинга РНК; более того, сплайсосомы в собственном смысле этого слова обнаруживаются только у эукариот.

Взятые в совокупности, филогенетические наблюдения подсказывают, что архейный предок эукариот сочетал разнообразие черт, обнаруживаемых по отдельности в различных ныне существующих археях. Эволюционные реконструкции, использующие принцип наибольшей экономии и, в особенности, развитые методы наибольшего правдоподобия, указывают на генетическую сложность общего предка всех ныне существующих архей – по меньшей мере, по сравнению с типичными ныне существующими формами, но вполне вероятно обладавшего даже большим разнообразием генов (Csuros and Miklos, 2009; Makarova et al., 2007b). Линии ныне существующих архей, вероятно, возникли путем дифференциальной оптимизации и редуکتивной эволюции сложных предковых форм (больше об этой линии эволюции в гл. 8), тогда как эукариоты в большой степени сохранили предковую сложность (Makarova et al., 2010). Разнообразие источников происхождения различных функциональных систем эукариотических клеток имеет важнейшие последствия для моделей эукариогенеза, которые мы обсудим ниже в этой главе.

Эукариогенез: источники уникальной эукариотической клеточной организации

Сценарий симбиогенеза в сравнении с архезойным сценарием

Многочисленные филогенетические наблюдения, кратко изложенные в предыдущем разделе, не объясняют, откуда возникла эукариотическая клетка, однако создают необходимый фундамент для построения сценариев эукариогенеза. В табл. 7–3 перечисляются самые важные наблюдения, которые должны быть включены в любое эволюционное объяснение происхождения эукариот (эукариогенез) и ранних стадий их эволюции. С учетом этих наблюдений главный вопрос сейчас касается роли эндосимбиоза: был ли он причиной всей цепи событий, приведших к появлению LECA, то есть стволовой фазы эволюции эукариот, как в сценарии симбиогенеза, или же это был шаг в эволюции уже сформировавшейся эукариотической клетки, как в архезойном сценарии? Другими словами, был ли хозяин альфа-протеобактериального эндосимбионта (будущей митохондрии) прокариотом или амитохондриальным эукариотом, то есть архезоем?

Таблица 7–3. Ключевые пункты, которые необходимо принимать во внимание при моделировании эукариогенеза.

- Все ныне существующие эукариоты имеют митохондрии или родственные органеллы, поэтому эндосимбиоз должен предшествовать LECA.
- LECA представлял собой организм высокой сложности, уже имел все характерные функциональные системы эукариот и, вероятно, был типичной эукариотической клеткой, поэтому все ключевые инновации эукариогенеза должны были произойти в стволовой фазе эволюции, до LECA. Среди прочего, эти инновации включают появление интронов и сплайсосом.
- Продолжительность стволовой фазы неизвестна, но есть определенная вероятность, что она была долгой и что существовало значительное разнообразие эукариот до LECA.
- Высококонсервативные гены эукариот представлены в виде химерного набора: меньшая часть – гены, кодирующие системы передачи информации и некоторые другие молекулярные машины, такие как аппарат деления клетки, – восходят к археям, тогда как большая часть – гены метаболических ферментов – происходят от бактерий.
- Некоторые ключевые функциональные системы эукариотической клетки, такие как РНК-интерференция или пути репарации, являются археобактериальными химерами. Другие важнейшие молекулярные машины эукариотической клетки, такие как комплекс ядерной поры, по-видимому, имеют преимущественно бактериальное происхождение.
- Предки эукариотических генов разбросаны среди архейных и бактериальных групп.

С учетом того, что эукариогенез, вероятнее всего, был единичным событием и что промежуточные эволюционные стадии между возникновением первых эукариотических клеток и появлением LECA почти недоступны для сравнительной геномики или любых других

возможных методов, можно усомниться, что на эти вопросы когда-либо будут получены исчерпывающие ответы. Но это не значит, что их следует считать принципиально неразрешимыми. Присутствие митохондрий или МПО у всех ныне существующих эукариот часто привлекается в качестве аргумента в поддержку симбиогенетического сценария, однако этот аргумент сильно, если не безнадежно падает в цене, если принять версию долгой стволовой фазы в эволюции эукариот. Действительно, темные века могли принадлежать вымершим архезоа, из которых только одна группа, «приручившая» альфа-протеобактерии, выжила и дала начало всему современному разнообразию эукариот. Тем не менее симбиогенетический сценарий все же кажется более приемлемым, чем архезойный сценарий, по трем принципиальным причинам.

1. Архезойный сценарий не предлагает конкретных факторов отбора, обуславливающих эволюцию ядра и в особенности развитие комплекса ядерной поры. Ядро разрушает сопряжение транскрипции и трансляции, типичное для бактерий и архей, и влечет за собой эволюцию время– и энергозатратного механизма ядерно-цитозольного транспорта мРНК. Симбиогенетическая гипотеза предлагает приемлемый фактор отбора: защиту от инвазии генома хозяина самосплайсирующимися интронами группы II (это настоящие эгоистичные генетические элементы – о них пойдет речь в гл. 8), эволюционными предшественниками сплайсосомных интронов, которые распространены у альфа-протеобактерий, но не у *Archaea*. Непрестанное воздействие на геном архейного хозяина со стороны ДНК бактериального эндосимбионта из разрушенных клеток последнего могло приводить к активации интронов группы II и их массивному встраиванию в геном хозяина. Такие встроенные интроны должны были фатально нарушать экспрессию генов, если только транскрипция и трансляция не были разобщены и компарментализованы, становясь таким образом движущей силой эволюции ядра (см. подробности в следующем разделе).

2. Как было обрисовано в предыдущем разделе, совокупность геномных, ультраструктурных и функциональных исследований прокариот, особенно архей, показывает, что не только молекулярные компоненты многих характерных эукариотических систем, но также и их современная структура и функции возникли уже у архей и таким образом предшествовали эукариогенезу (системы типа 1 обсуждались ранее; см. табл. 7–2). Однако система внутренних мембран и ядро, так же как сами митохондрии, прерывающие белоккодирующие гены интроны и сплайсосомы, которые отвечают за сплайсинг экзонов (вырезание интронов), являются системами (признаками) типа 2. Эти системы не имеют функциональных аналогов у прокариот, хотя они, по-видимому, и собраны из прокариотических компонентов. Таким образом, приемлемой выглядит единая причинно-следственная цепочка событий (см. рис. 7–6): *эукариогенез был инициирован эндосимбиозом, а система внутренних мембран, включающая ядро, развилась как защита против инвазий интронов группы II и, может быть, вообще бактериальной ДНК* (Martin and Koonin, 2006a; Lopez-Garcia and Moreira, 2006). Вряд ли случайно, что многие ключевые компоненты этих эндомембранных систем явно имеют бактериальное происхождение, тогда как другие являются белками, состоящими из повторов, которые могли возникнуть *de novo*.



Рис. 7–6. Возникновение эукариотической клеточной организации как многоуровневой системы защиты против инвазии интронов: гипотетическая единая причинно-следственная цепь. По Коопин, 2006.

3. Простые оценки, произведенные Ником Лейном и Биллом Мартином, свидетельствуют о том, что возникновение больших, сложных клеток, подобных эукариотическим, вряд ли энергетически возможно без приобретения множественных энергопродуцирующих органелл, способных к автономной репродукции и регуляции (Lane and Martin, 2010). У прокариот белковые комплексы, включая электронтранспортную цепь и мембранную АТФ-синтазу, превращающие энергию протонного либо натриевого градиента в АТФ, расположены на плазматической мембране. Биогенез этих комплексов неразрывно сопряжен с синтезом их субъединиц, высокогидрофобных белков, которые встраиваются в мембрану котрансляционным путем. Поскольку поверхность клетки пропорциональна квадрату ее диаметра, тогда как объем пропорционален кубу диаметра, увеличение размера клетки в какой-то момент делает эту биоэнергетическую модель неэффективной и, таким образом, неприемлемой. В настоящее время мы знаем только два пути к эффективной биоэнергетике, которые потенциально могут инициировать эволюцию больших клеток. Первый путь использовали эукариоты, обладающие множеством энергопроизводящих органелл эндосимбиотического происхождения в каждой клетке. Второй путь реализован в некоторых недавно открытых гигантских бактериях, которые содержат множество копий геномной ДНК в каждой клетке (более 10 тысяч у симбионта рыб *Eupuliscium sp.*; Mendell et al., 2008). Каждая копия хромосомы, по-видимому, прикреплена к мембране, так что синтез мембранных белков, в том числе ответственных за трансформацию

энергии, наверняка тесно сопряжен с встраиванием этих белков в мембрану. В отличие от эукариотического решения, это второе изобретение больших клеток не предполагает резкого уменьшения генома, ключевого признака митохондрий, который вносит вклад в эффективность энергетики эукариотической клетки, и не породило разнообразия сложных жизненных форм.

Против сценария симбиогенеза было выдвинуто несколько возражений (Kurland et al., 2006; Poole and Penny, 2007). Во-первых, прокариотические эндосимбионты в прокариотических хозяевах не распространены, и это внушает мысль, что фагоцитоз, который явно уникален для эукариотических клеток, совершенно необходим для захвата митохондрий. Этот аргумент неубедителен по четырем причинам:

1. Эукариогенез является исключительно редким, возможно единичным событием в истории развития жизни. Как уникальное (или почти уникальное) событие, он совсем не обязательно должен был требовать механизма, который бы рутинно функционировал у хозяина первичного эндосимбионта.

2. Эндосимбиотические бактерии внутри других бактерий не часто, но все-таки встречаются (von Dohlen et al., 2001). Внутриклеточное хищничество бактерий тоже могло стать путем, ведущим к эндосимбиозу (Davidov and Jurkevitch, 2009).

3. Обнаружение эукариотических систем перестройки мембран и гомологов актина белками у архей (Makarova et al., 2010; Yutin et al., 2009) указывает на возможность существования еще не обнаруженных механизмов для захвата других прокариот, сходных с примитивным фагоцитозом.

4. Компьютерное моделирование свидетельствует о том, что дифференциация клеточных форм жизни в хищников и жертв исходно присуща эволюции клеток и поэтому могла возникнуть вскоре после появления первых клеток (de Nooijer et al., 2009).

Во-вторых, потенциально сильным аргументом против сценария симбиогенеза может быть существование значительного числа характерных белков эукариот (ХБЭ), то есть белков, обнаруживаемых только у эукариот. Происхождение ХБЭ загадочно. Однако, как уже упоминалось в предыдущем разделе в отношении характерных эукариотических функциональных систем, тщательные исследования последовательностей и структур приводят к идентификации растущего числа архейных и бактериальных гомологов белков, которые первоначально считались ХБЭ, либо существование таких гомологов становится очевидным с появлением новых секвенированных геномов. Открытие прокариотических гомологов тубулина, актина и убиквитина – хорошо известные примеры; более новые включают так называемые субъединицы GINS эукариотических комплексов репликации ДНК (Marinsek et al., 2006), системы ESCRT-III (Makarova et al., 2010) и субъединицы комплекса TRAPP (Barrowman et al., 2010), играющие ключевые роли в эукариотическом транспорте везикул. Согласно сценарию симбиогенеза, ХБЭ (бывшие и до сих пор сохраняющие этот статус) возникли в результате ускорения эволюции генов, функции которых существенно изменились в ходе эукариогенеза; другим важным путем эволюции ХБЭ могла быть дифференциация простых повторяющихся белковых структур, ведущая к возникновению уникальных глобулярных укладок (Aravind et al., 2006).

Третьим серьезным возражением против сценария симбиогенеза может быть тот факт, что ни архейные, ни бактериальные гены невозможно вывести из какой-либо одной прокариотической группы (хотя происхождение митохондрий от альфа-протеобактерий точно установлено). Однако пангеномы прокариот очень велики, тогда как состав генов у индивидуальных организмов чрезвычайно изменчив, поэтому реконструкция реальных партнеров эндосимбиоза, который привел к эукариогенезу, из ограниченного набора существующих ныне геномов может быть в принципе невозможна (Martin, 1999; Esser et al.,

2007). Более того, многие, если не сказать большинство, современные археи и бактерии, возможно, развились путем оптимизации (см. гл. 5 и 8), в то время как эукариогенез мог быть инициирован симбиозом между двумя прокариотами, обладавшими сложными геномами.

В настоящее время нельзя полностью исключить возможность, что ключевые эукариотические нововведения развились независимо от митохондриального эндосимбиоза и предшествовали ему. Таким образом, в принципе хозяином эндосимбионта мог быть архезой. Однако архезойный сценарий не предлагает убедительной последовательности событий в эволюции сложной внутренней организации эукариотической клетки, не дает разумного обоснования для появления ядра и не объясняет присутствия характерных функциональных систем эукариот в различных группах архей. Напротив, сценарий эндосимбиогенеза может объединить все эти разнородные свидетельства в связную историю, пусть еще далеко не полную и заведомо спекулятивную.

Симбиотический сценарий эукариогенеза: происхождение ключевых эукариотических инноваций, инициированных эндосимбиозом

Мы уже затронули многие аспекты гипотезы об эукариогенезе, инициированном симбиогенезом. Данный раздел объединяет эти разрозненные свидетельства в согласованную схему эукариогенеза. При этом мы не должны забывать, что детали того, что происходило в действительности, восстановить невозможно, и если бы мы занялись разработкой спекуляций об этих деталях, то в конце концов были бы обречены признать, что это просто вымысел. Тем не менее, если мы удерживаем дискуссию на относительно «крупнозернистом» уровне, вполне возможно обнаружение некоей логики даже в единичных эволюционных событиях, каковым является эукариогенез.

Попробуем привязать сценарий эукариогенеза (см. рис. 7–7) к конкретным стадиям и периодам истории жизни на Земле и самой Земли (Kasting and Ono, 2006). Время и место – приблизительно два миллиарда лет назад (палеопротерозой), умеренная температура и соленость, предположительно океанское дно на мелководье. Атмосфера Земли (и, соответственно, океан) в первые 1,5 миллиарда лет истории жизни находились в сильно восстановленном состоянии. Однако примерно в то время, о котором мы говорим, благодаря появлению кислородного фотосинтеза у цианобактерий, началась микрооксигенизация Земли. Концентрация кислорода была, вероятно, на два или три порядка ниже, чем сейчас, но аэробное дыхание уже сделалось возможным. Разнообразие микробной биоты в биосфере было сравнимо с современным, за исключением малочисленности (почти отсутствия) аэробных организмов. Все главные группы архей и бактерий, известные нам сегодня, уже существовали, и даже, вероятно, были другие, ныне вымершие. Экологическая обстановка – важный пункт, который иногда упускается при обсуждении эндосимбиоза: описываемые события, вероятнее всего, происходили в микробных матах, широко распространенных и тесно объединенных сообществах различных бактерий и архей (Allen and Banfield, 2005). В микробных матах уровень горизонтального переноса генов, вероятно, настолько высок, насколько это вообще возможно, и так же должно обстоять дело с частотой заглатывания одной прокариотической клетки другой клеткой, потенциально способного приводить к эндосимбиозу, но вообще редкого среди прокариот.

Главным персонажем этой драмы, хозяином будущего эндосимбионта (назовем его «архейный прародитель эукариот», АПЭ), наиболее вероятно, как отмечалось выше, был представитель мезофильных архей, предположительно с большим геномом, который мог содержать до 5–6 тысяч генов. Имеющиеся данные о дошедших до нас мезофильных археях

удручающе неполны по сравнению с другими группами прокариот. Тем не менее то, что мы все-таки знаем, согласуется с предположением, что генетически сложные организмы с множеством генов, приобретенных путем горизонтального переноса, обычны в этой экологической группе. Действительно, самые большие известные геномы архей, единственные археи, обладающие более чем 5000 генов, обнаружены среди мезофилов (а именно некоторых *Methanosarcina*), и до 20 процентов этих геномов содержат гены сравнительно недавнего бактериального происхождения. Другие известные мезофильные археи, такие как хорошо описанные *Thaumarchaeota*, тип архей, до недавнего времени скрывавшийся за неопределенным наименованием мезофильных *Crenarchaeota*, имеют меньшие геномы, но и они тоже обогащены «бактериальными» генами (Brochier-Armanet et al., 2008). Ни один из известных в настоящее время архей не похож на предполагаемого кандидата на столь важную роль АПЭ. Как обсуждалось ранее в этой главе, архейное наследство современных эукариот – смешанное, с наборами генов, происходящих от различных групп *Crenarchaeota*, *Thaumarchaeota* или *Euryarchaeota*. Существует интересная возможность того, что неуловимый АПЭ сочетал многие (возможно, большинство) из этих генов в пределах одного генома/клетки до того, как произошел эндосимбиоз, хотя и последующее приобретение генов посредством ГПГ также могло быть важно. Отталкиваясь от бурной картины прокариотического мира, описанной в главе 5, можно предположить, что потенциально важными эволюционными посредниками были не вымершие, а неустойчивые группы (линии) архей и бактерий. Такие относительно короткоживущие жизненные формы с очень сложными, мозаичными геномами могли возникать довольно часто (по эволюционным масштабам), но чаще всего они постепенно теряли большие части своих геномов и деградировали до более стабильных, привычных нам форм. Однако некоторые из этих предполагаемых сложных временных состояний могли породить взрывы разнообразия (см. рис. 7–7).

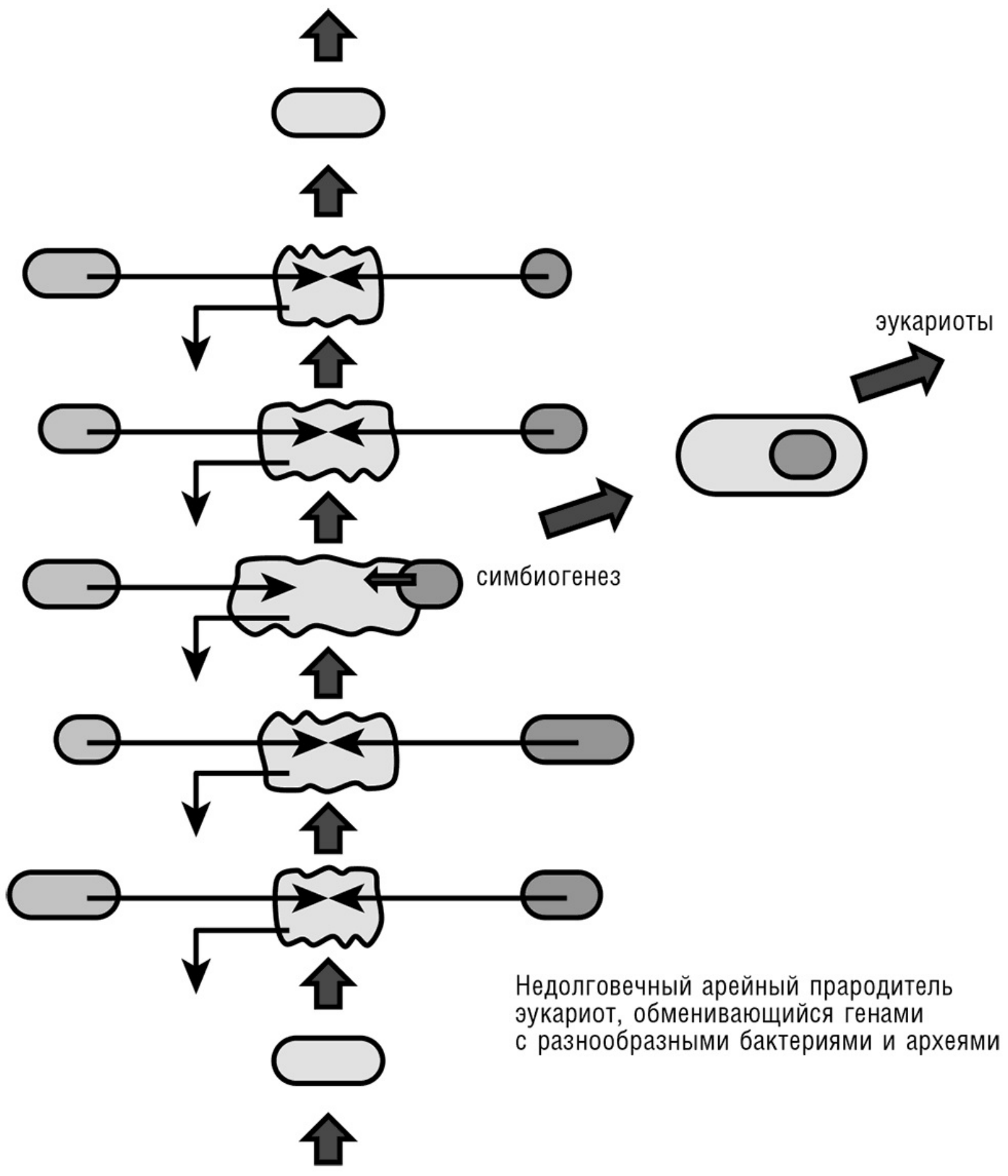


Рис. 7-7. Архейный прародитель эукариот (АПЭ) как переходная сложная архейная форма, предрасположенная к горизонтальному переносу генов, и эндосимбиотический сценарий эукариогенеза.

Когда писалась эта глава, появилось сообщение об удивительном открытии,

иллюстрирующем неизученное разнообразие архей и увеличивающем вероятность того, что близкие родственники неуловимого АПЭ могли дожить до сего дня. Оливье Грос с коллегами сообщили о двух видах *Thaumarchaeota*, населяющих морское мелководье (Müller et al., 2010). Эти археи обладают гигантскими клетками, которые формируют ясно выраженные макроскопические филаменты. Более того, клетки одного из этих видов покрыты симбиотическими альфа-протеобактериями. Бактерии в этом случае являются экто-, а не эндосимбионтами; тем не менее этот тип археобактериальной ассоциации прекрасно может создавать условия, способствующие эндосимбиозу. Эти новооткрытые археи вряд ли являются близкими родственниками АПЭ (хотя, повторюсь, вероятность того, что являются, – ненулевая), но по любым меркам это открытие красноречивее всяких слов говорит о правдоподобности эндосимбиотического сценария эукариогенеза.

История эукариот (см. рис. 7–7) начинается с АПЭ, захватившего альфа-протеобактерию, которую трудно точно идентифицировать. Этот АПЭ мог иметь специфическую склонность к поглощению других прокариотических клеток, хотя, разумеется, это не был настоящий фагоцит, подобный современной амебе. Однако представляется вероятным, что АПЭ был археем, не имеющим клеточной стенки, похожим в этом отношении на ныне существующих термофильных архей из рода *Thermoplasma*. Более того, он, вероятно, обладал некоторой разновидностью цитоскелета, образованной актиноподобными белками, родственными тем, которые обнаружены у другой группы термофильных архей (группа в составе *Crenarchaeota*, известная как *Thermoproteales*); сравнительный анализ последовательностей этих архейных гомологов актинов (которые, к сожалению, еще не исследованы экспериментально) даже свидетельствует о возможности того, что они образуют ветвистые филаменты, ключевые структуры, вовлеченные в эукариотический фагоцитоз (Yutin et al., 2009) [\[621\]](#). Так что не будет безосновательным предположение, что АПЭ «пасся» на бактериальном мате, время от времени захватывая бактериальные клетки. Большая часть поглощенных бактерий встретила свой конец в качестве пищи; другие бактерии могли убивать захватчиков, а некоторые могли становиться временными симбионтами. Фиксация эволюционно стабильного эндосимбионта – чрезвычайно трудная задача, поскольку надо преодолеть много препятствий, чтобы создать такой стабильный эндосимбиоз. Ясно, что, хотя изначальный захват будущего эндосимбионта мог произойти по воле случая, фиксация эндосимбионта могла стать возможной только в той мере, в какой это было связано с явными селективными преимуществами для возникшего химерного организма.

Что могло быть фактором (или факторами) отбора, поддержавшими появление археобактериальной химерной системы? Если предполагать микроаэрофильные условия во время эукариогенеза (или, может быть, даже анаэробные условия в той специфической среде, где имел место эукариогенез), то вряд ли селективным преимуществом была основанная на дыхании биоэнергетика. Вместо этого начальной причиной, способствовавшей стабилизации эндосимбиоза, могла быть метаболическая интеграция хозяина и симбионта, которая постепенно становилась мутуалистической. Специфическая модель такой метаболической связи, так называемая водородная гипотеза, была предложена Биллом Мартином и Миклосом Мюллером (Martin and Müller, 1998). Согласно водородной гипотезе, метаболизм архейного хозяина был основан на утилизации молекулярного водорода, который был побочным продуктом анаэробного гетеротрофного метаболизма симбионта. Анаэробная продукция АТФ, а отчасти и аэробное дыхание могли стать дополнительными преимуществами эндосимбиоза.

Эндосимбиоз мог создать особые условия внутри клеток гибридного организма. Естественно, для того, чтобы быть унаследованными, бактерии, ставшие эндосимбионтами, должны были делиться. Все современные аэробные эукариотические клетки содержат множество митохондрий, и само собой разумеется, что эта связь – множество эндосимбионтов

внутри одной химерной клетки – тянется из очень ранней фазы эволюции, фактически с момента происхождения этой химеры. Число копий эндосимбионта могло расти постепенно, с растущей зависимостью клетки от метаболизма симбионта. В этой ситуации эндосимбионты неизбежно должны были подвергаться лизису, что приводило к выпуску ДНК симбионта в окружающий (хозяйский) цитозоль химерной клетки. Примечательно, что даже в современных растениях и животных, где хромосомы отчасти защищены от чужой ДНК оболочкой ядра и фиксация любой встроившейся ДНК осложнена необходимостью интеграции в зародышевую линию и переживания рекомбинации в мейозе, встройки больших кусков митохондриальной ДНК в ядерный геном довольно распространены (Nazkani-Covo et al., 2010). В химерной протоэукариотической клетке вскоре после эндосимбиоза незащищенная ДНК хозяина должна была подвергаться настоящей бомбардировке ДНК эндосимбионта. Отметим, что эта ситуация принципиально асимметрична, потому что, во-первых, геномы жизнеспособных эндосимбионтов защищены от инвазии ДНК хозяина бактериальной мембраной, тогда как ДНК хозяина открыта; а во-вторых, потому, что количество свободной ДНК эндосимбионта намного больше. Таким образом, *в результате возникает храповик (ratchet) перемещения генов от эндосимбионта к хозяину* (Martin and Koonin, 2006a).

Открытость генома хозяина воздействию ДНК эндосимбионта имеет несколько важнейших последствий. Вставка участка ДНК эндосимбионта в ген хозяина, выполняющий важные для выживания клетки функции, должна была приводить к гибели соответствующих клеток и не фиксироваться в протоэукариотической популяции. Прокариоты обладают геномами с тесным («стена к стене») расположением генов, эти геномы в основном составлены из белок-кодирующих генов (см. гл. 5), и нет причины полагать, что АПЭ был исключением из этого правила. Таким образом, размножение эндосимбионта, сопровождаемое случайным лизисом, должно было оказывать на популяцию химерных клеток исключительное давление, возможно приводящее к втягиванию популяции в «бутылочное горло». Такое «бутылочное горло» может резко увеличивать скорость генетического дрейфа и, следовательно, усиливать роль случайности в эволюции при снижении интенсивности отбора (мы обсудим это важное явление подробнее в гл. 8). Предлагаемый сценарий эукариогенеза выглядит парадоксальным: эндосимбиоз кажется выгодным в плане метаболической кооперации, но в то же время разрушительным по причине высвобождения ДНК эндосимбионта и других эффектов его внутриклеточного размножения. Однако можно полагать, что эта ситуация создает сильное напряжение, но не парадокс: химерная клетка может выдержать атаку чужой ДНК без потери эндосимбионта, если *взаимные связи между хозяином и симбионтом были установлены спустя очень короткое время после вторжения симбионта*. Это напряжение между необходимостью сохранить эндосимбионта и тем гнетом, который он оказывает на химерную клетку, могло оказаться необходимым условием для появления эукариотических нововведений.

Последовательности определенного класса, будучи встроены даже в функционально важные гены, наносят гораздо меньший урон. Это так называемые самосплайсирующиеся интроны группы II, класс обратно-транскрибируемых самореплицирующихся генетических элементов, которые «прыгают» по геномам многих бактерий и некоторых мезофильных архей, а также митохондрий грибов и растений (Lambowitz and Zimmerly, 2004). Они имеют очень интересный, необычный жизненный цикл: используя РНК (рибозимный катализ), они вырезают сами себя из транскриптов соответствующих генов хозяина, а затем, сделав копии собственной ДНК с помощью обратной транскриптазы, которую сами и кодируют, встраиваются в новые сайты на хромосоме хозяина. Сегодня считается доказанным, что интроны группы II, которые в мире эукариот представлены только в некоторых органеллах эндосимбиотического происхождения, являются предшественниками сплайсосомных интронов, которые прерывают

эукариотические белок-кодирующие гены (Keating et al., 2010; Toor et al., 2008). Действительно, терминальные структуры интронов группы II, ответственные за вырезание интрона, имеют близкое сходство с каноническими терминальными структурами сплайсосомных интронов. Еще важнее, что небольшие молекулы РНК в сплайсосоме, катализирующие сплайсинг у всех эукариот, также происходят от интронов группы II. Большинство бактерий контролируют число интронов группы II, удерживая его на уровне нескольких копий на бактериальную хромосому (или удаляя вовсе), вследствие интенсивного очищающего отбора в бактериальных популяциях (см. гл. 8). Интересно, что альфа-протеобактерии относительно богаты этими элементами, они содержат до 30 копий на бактериальный геном. Наиболее высокое содержание интронов группы II наблюдается в митохондриях грибов и растений, где они составляют существенную долю генома. Размножение интронов группы II в геноме эндосимбионта могло начаться вскоре после установления эндосимбиоза и было, вероятно, инициировано неизбежным снижением размера популяции симбионта и последующей невозможностью эффективно избавляться от самореплицирующихся элементов.

Таким образом, интроны группы II могли в значительном количестве присутствовать в ДНК эндосимбионта, бомбардировавшей геном хозяина. Более того, эти элементы обладают способностью активно интегрироваться в другие молекулы ДНК, так что они могли агрессивно атаковать хромосомы хозяина, встраиваясь в гены, а затем перемещаясь куда-то еще (Martin and Koopin, 2006a). Хотя интроны группы II после транскрипции автокаталитически вырезаются из транскрипта, так что окружающие экзоны сшиваются вместе, все же массовое «заражение» генов хозяина может представлять серьезную опасность. В самом деле, сплайсинг – относительно медленный процесс, он намного медленнее трансляции. У прокариот, где транскрипция и трансляция сопряжены, транскрипты со встроенными интронами группы II во многих случаях были бы транслированы раньше, чем дело дошло бы до сплайсинга. При большом количестве интронных вставок последствия могли быть чрезвычайно существенны, возможно, фатальны: накапливались бы неправильно транслированные белки, и это оказывало бы пагубное воздействие на клетку. Еще более серьезными могли бы быть последствия инактивации открытой рамки считывания, кодирующей обратную транскриптазу (ОТ) интронов группы II, действующую в *цис*-положении как кофактор сплайсинга (на этой стадии – не как фермент). Для генов, содержащих интроны с инактивированными генами ОТ, должен был бы идти сплайсинг *in trans* ^[63]. Как известно, он непродуктивен, так что такие интроны с высокой эффективностью прекращали бы синтез соответствующих функциональных белков. Таким образом, заражение генов хозяина интронами группы II должно было создать мощную движущую силу для каскада эволюционных обновлений (Koopin, 2006):

1. Аппарат сплайсинга, способный к эффективному функционированию в *транзакции*.
2. Инструмент защиты, который мог бы разобщить трансляцию и транскрипцию, позволив относительно медленному процессу сплайсинга завершиться до того, как начнется трансляция.
3. Дополнительные «линии защиты» от накопления аберрантных полипептидов.

Действительно, все три типа адаптаций к вторжению интронов развились в эволюции прокариот до появления ЛЕСА: сплайсосома, ядро, дополнительные системы контроля качества, такие как нонсенс-опосредованный распад (НОР) – механизм, удаляющий незрелые транскрипты, и убиквитин-зависимая система деградации белков, которая непосредственно разрушает аберрантные белки (см. рис. 7–6).

Таким образом, в принципе натиск ретроэлементов эндосимбионта на геном хозяина создает давление отбора, необходимое для возникновения ряда определяющих нововведений эукариотической клетки, и самое главное – системы внутренних мембран, главным компонентом которой является ядро. При более близком рассмотрении, однако, проблема

развития этих систем все еще подозрительно напоминает «неупрощаемую сложность». Необходимы особые объяснения, и их непросто находить. Например, сложно устроенный комплекс ядерной поры не может работать и, соответственно, не может быть подхвачен отбором в отсутствие оболочки ядра, но последняя не может сообщаться с цитозолем без комплекса ядерной поры. Скорее всего, эволюция системы внутренних мембран и ядра, хоть и быстрая в масштабах, все же проходила через промежуточные стадии. Пролиферация эндосимбионтов внутри развивающихся химерных клеток могла быть очень постепенной, что позволяло бы протоэукариотам жить достаточно долго, чтобы нововведения с ограниченным положительным эффектом могли зафиксироваться. Можно представить, что серия нововведений началась с образования везикул из мембраны эндосимбионта. Эти везикулы могли сформировать примитивную систему внутренних мембран, включая протоядро, то есть компартмент, заключающий в себе одну или несколько хромосом и имевший не поры современного типа, а только просветы между уплощенными везикулами; каждая из них оставалась связана с системой внутренних мембран. Просветы в протоядерной мембране предупреждали доступ рибосом к сайтам транскрипции, таким образом разъединяя транскрипцию и трансляцию, типично сопряженные у прокариот, и уменьшая тем самым вред от встроившихся ретроэлементов (интронов группы II). В такой ситуации клетка могла пережить дальнейшую пролиферацию (прото)митохондрий и усиление выхода из них ДНК и ретроэлементов. Это, в свою очередь, могло создавать селекционное давление, способствующее дальнейшей эволюции ядра, которая в итоге привела к появлению комплекса поры современного типа, активно контролирующего ядерно-цитоплазматические потоки и соединяющего сплайсинг первичных мРНК с экспортом зрелых мРНК из ядра. Пролиферация внутренних мембран в конечном итоге привела к полной реконструкции мембранной системы прокариотической клетки, причем предковая архейная плазматическая мембрана была замещена бактериальными мембранами, предположительно изнутри, путем распространения эндомембран, происходящих от симбионта.

Можно представить себе подобный сценарий для эволюции сплайсосомы, начиная с системы, состоящей только из РНК, в которой как интроны, так и каталитическая малая РНК, участвующая в сплайсинге, происходят от ретроэлементов. Следующая стадия эволюции могла, в числе прочего, задействовать *Sm*-белок ^[64], стабилизирующий РНК-дуплексы, вовлеченные в сплайсинг (Veretnik et al., 2009), за чем последовала эволюция рибонуклеопротеидной сплайсосомы. Замечательно, что совсем недавние наблюдения указывают на то, что один из ключевых белковых компонентов сплайсосомы, Prp8, является инактивированным производным обратной транскриптазы интрона группы II (Dlakis and Mushegian, 2011). Это неожиданное открытие – еще одно свидетельство разнообразного вклада интронов группы II в происхождение как сплайсосомных интронов, так и самой сплайсосомы. В более общем плане такая пошаговая эволюционная настройка может отчасти объяснять эволюцию сложных систем, столь характерных для эукариотической клетки.

Конечно, многие важные аспекты клеточной организации эукариот не могут быть легко связаны с прямыми последствиями эндосимбиоза. Обратите внимание на эукариотический хроматин, с его множеством линейных хромосом, заменяющих кольцевые хромосомы, обычные у бактерий и архей. Чрезвычайно изощренная организация эукариотического хроматина, с его регулярными нуклеосомными структурами, по крайней мере внешне резко отличается от много более простой структуры прокариотических хромосом (Branco and Pombo, 2007), хотя археи (эуриархеи) обладают простыми нуклеосомами, содержащими гистоны (Bailey et al., 2002). Добавьте к этому фундаментальное изменение устройства генома, в котором главный принцип прокариотических геномов – оперонная организация – отброшен. Эту серию крупных

изменений, связанных с эукариогенезом, трудно приписать специфическому влиянию эндосимбиоза. Тем не менее можно проследить некоторые интересные связи. Линейные хромосомы стоят перед трудной проблемой репликации концов, учитывая, что все известные ДНК полимеразы требуют затравки и не могут стартовать с первого нуклеотида матрицы. Если не работает специальный механизм, обеспечивающий сохранение концевых участков хромосом, то в каждом цикле репликации они укорачиваются, делая репликацию неустойчивой. Все эукариоты используют фермент, называемый теломеразой, который сохраняет набор повторов на концевых участках хромосомы путем обратной транскрипции небольшой молекулы РНК, ассоциированной с ним (Autexier and Lue, 2006). Поразительно, но теломераза – это еще одно (кроме Pgp8) эволюционное производное от обратной транскриптазы интрона группы II, в данном случае сохранившее энзиматическую активность. Теломераза представляется тем связующим звеном, которое обеспечило переход к линейным хромосомам в пределах связанного с эндосимбиозом или, говоря более узко, стимулированного ретроэлементами каскада инноваций (Koonin, 2006; см. рис. 7–6 и 7–7).

Для всей последующей эволюции эукариот критически важным и, очевидно, неизбежным следствием возникновения ядра стало радикальное снижение уровня ГПГ, если не полное его прекращение. Хотя имеются многочисленные сообщения о захватах бактериальных генов одноклеточными эукариотами, уровень ГПГ трудно сравнивать с тем, который наблюдается у непаразитических бактерий и архей (Keeling and Palmer, 2008). Большая часть ДНК, которая попадает в эукариотическую клетку, разрушается, даже не успев войти в ядро и достичь хроматина. Такое резкое замедление ГПГ подсказывает естественный ответ на вопрос, который иначе ставит в тупик: почему эукариоты утратили все опероны своих прокариотических предков? (Архейный хозяин, несомненно, обладал типичной оперонной организацией генов, как и эндосимбионт.) Вспомним концепцию эгоистичного оперона: как только происходит эффективный ГПГ, храповик приводится в движение; значит, коль скоро оперон разрушен, вероятность его воссоздания посредством рекомбинации, а затем сохранения отбором становится чрезвычайно мала. Фактически этот оперон безвозвратно утрачивается в данной линии. Видимо, этот храповой механизм уничтожил все прокариотические опероны на ранних стадиях эволюции эукариот. Опероны, все-таки существующие у некоторых эукариот, таких как нематоды, не имеют ничего общего с прокариотическими оперонами; по всей видимости, они возникли *de novo* не были зафиксированы в далеко разошедшихся эукариотических линиях.

Данный сценарий приводит к простому, но неожиданному предсказанию: те гены, которые могут функционировать только внутри оперонов, но оказывают повреждающее действие, оказавшись вне контекста оперона, будут полностью утрачены у эукариот. Замечательно, что это в точности соответствует случаям систем токсин – антитоксин и рестрикции – модификации, которые повсеместно распространены у бактерий и архей (см. гл. 5), но, по-видимому, полностью отсутствуют у эукариот.

Почти полная элиминация ГПГ также дает эволюционный стимул для широкомасштабной дупликации генов, которая является главным механизмом инновации у эукариот (Lespinet et al., 2002). Популяционное «бутылочное горлышко», обусловленное размножением эндосимбионтов, сделало возможным взрыв дупликаций во время фазы ствола (Makarova et al., 2005; также см. гл. 8), но в более общем смысле дупликации замещают ГПГ как главный источник обновлений в течение всей эволюции эукариот.

Последнее по порядку, но, конечно, не по значению – это то, что низкий уровень ГПГ у эукариот можно считать принципиальным фактором, обусловившим возникновение полового размножения с использованием мейоза, одного из определяющих биологических процессов у эукариот. Действительно, у эукариот вредные мутации обычно не могут быть

скомпенсированы горизонтально перенесенными генами, и это давление должно способствовать эволюции системы регулярной рекомбинации, которая бы предупреждала накопление таких мутаций и последующий мутационный крах. Такая система, противодействующая храповику Мёллера, развилась в форме мейоза и полового размножения. Это мог быть и не единственный фактор, который послужил движущей силой в эволюции полового размножения, но он, несомненно, важен (мы не имеем возможности обсуждать подробно данную проблему, чрезвычайно популярную среди эволюционных биологов [de Visser, Elena, 2007]). Учитывая, что снижение ГПГ в большой степени является следствием эволюции ядра, «изобретение» мейоза и полового размножения – на базе систем репарации и деления клеток у архей – по-видимому, является частью цепи инициированных вторжением интронов адаптаций, направленных на защиту и контроль повреждений (см. рис. 7–6).

Другое важнейшее последствие пролиферации протомитохондриального эндосимбионта внутри химерной протоэукариотической клетки состоит в том, что в известном смысле комплементарным снижению ГПГ из внешних источников является высвобождение случайных участков бактериальной ДНК (в противоположность эгоистичным элементам) посредством лизиса эндосимбионтов. Такие фрагменты ДНК тоже могут встраиваться в хромосому хозяина, хотя и реже, чем эгоистичные элементы. Во многих случаях такие включения окажутся фатальными. Однако, когда полная последовательность гена из эндосимбионта встраивается в межгенный участок на хромосоме хозяина, заметного повреждающего эффекта может и не быть. Более того, встроенный ген может экспрессироваться, если необходимые регуляторные элементы доступны по соседству с местом встройки. Фрагменты митохондриальной ДНК случайно встраиваются в ядерные геномы растений и животных даже теперь (Nazkani-Covo et al., 2010), несмотря на преграды, создаваемые ядром и системами защиты от цитозольной ДНК. Несомненно, частота встроек была намного выше во время эукариогенеза, до полного формирования эукариотической клеточной организации. Храповик переноса генов приводит к удвоению гена, когда функциональные копии одного и того же гена представлены как в геноме эндосимбионта, так и в ядерном геноме. Некоторые из ядерных генов, возникших таким путем, были задействованы в клеточных функциях вне эндосимбионта. Однако в других случаях встроенному гену предшествует последовательность, кодирующая так называемый транзитный пептид, способный опосредовать импорт белка обратно в эндосимбионт. Это еще одно «счастливое совпадение», но оно не так невероятно, как кажется, потому что транзитные пептиды обычно представляют собой простые повторяющиеся последовательности, которые вполне могут возникнуть по чистой случайности (Neupert and Herrmann, 2007). Коль скоро есть некий ядерный ген, кодирующий белок, который функционирует в митохондрии, функционально избыточные митохондриальные гены могут быть утрачены без каких-либо вредных последствий. Эта избыточность создает *другой храповик, который направляет митохондрии прямо на путь редуکتивной эволюции*, притом что ядерный геном постоянно подвергается воздействию ДНК из лизированных эндосимбионтов, приводя к многочисленным «отборочным соревнованиям» для переноса каждого гена эндосимбионта в ядерный геном. Конечный результат: огромное большинство белков, функционирующих в митохондриях, кодируется ядерным геномом, и только те гены, которые непременно должны экспрессироваться внутри митохондрии для ее собственного функционирования (см. обсуждение выше в этой главе), остаются в геноме органеллы. Та же самая схема сохраняется для других эндосимбионтов, в частности для пластид. Конечно, редуکتивная эволюция также приводит к необратимой утрате многих генов эндосимбионта, которые оказываются избыточны без необходимости привлечения белков хозяина, поскольку сама функция этих белков становится несущественной для эндосимбионта или потому, что метаболиты хозяина, такие как

нуклеотиды и аминокислоты, импортируются в эндосимбионт, устраняя необходимость в соответствующих метаболических путях. Эндосимбионты с существенно редуцированным геномом вытеснили эндосимбионтов с большими геномами просто-напросто из-за быстрой репликации генома и соответственно быстрого деления, и таким образом редукция генома зафиксировалась в ходе эволюции.

Приписывание единственной все объясняющей причины любому крупному эволюционному сдвигу неизбежно будет недопустимым упрощением и эпистемологической ошибкой, поскольку «причины» вообще являются конструктами человеческого сознания (см. прил. I). Тем не менее я уверен, что согласованность между многими ключевыми эукариотическими нововведениями, легко интерпретируемая как ответ на эндосимбиоз, и в частности на атаки происходящих из эндосимбионта мобильных элементов, слишком убедительна, чтобы отказаться от нее как от беспочвенной фантазии. Такой сценарий, даже если он не является фальсифицируемым в целом (см. прил. I), тем не менее включает частные фальсифицируемые предсказания и стимулирует новые эксперименты. Действительно, с тех пор, как мы с Биллом Мартином предложили нашу версию эндосимбиогенетического сценария, в которой центральная роль отводится интронам группы II (Martin and Koonin, 2006a), она выдержала две весьма жесткие проверки на фальсифицируемость. Одна из них – открытие сопряжения транскрипции и трансляции у *Archaea* (French et al., 2007). И еще более показательная проверка – определение структуры интрона группы II, которая не оставляет места серьезным сомнениям по поводу происхождения сплайсосомных интронов от прокариотических ретроэлементов, что воспринималось как спекуляция в то время, когда наша гипотеза была сформулирована (Toor et al., 2008).

Еще одно направление экспериментальных исследований могло бы сконцентрироваться на необычных бактериях, таких как *Planctomycetes*, которые обладают внутриклеточными компартментами, заключающими в себе хромосому (Fuerst, 2005). Конечно, эти организмы – прокариоты по всем критериям [65]. Более того, сравнительный анализ геномов показывает, что они не кодируют гомологов белковых субъединиц комплекса ядерной поры (Mans et al., 2004). Наша модель предсказывает, что, хотя *Planctomycetes* и некоторые родственные бактерии обладают «ядроподобным» компартментом, у них сохраняется сопряжение транскрипции и трансляции, типичное для прокариот. Это означает, что функциональные рибосомы входят в «ядерный» компартмент и иницируют трансляцию матричных РНК еще до завершения их транскрипции, или же образующиеся молекулы мРНК одновременно с транскрипцией выталкиваются через отверстия в стенках компартмента.

Если, напротив, эксперименты покажут, что у «ядерных» бактерий транскрипция разобщена с трансляцией, то это будет серьезным возражением против нашей модели. Немедленным ее опровержением было бы открытие ныне существующих архезой, то есть свободноживущих эукариот, не имеющих никаких признаков МПО, однако обладающих всеми остальными типичными признаками эукариотической клеточной организации. (Паразит с сильно редуцированным геномом, однако, не будет здесь убедительным примером.) Вероятность того, что архезои будут однажды открыты, уменьшается с каждым случаем обнаружения еще одной группы простейших, не имеющих нормальных митохондрий, но все-таки содержащих МПО.

Удивительная история эукариотических интронов

«Кусочное» (экзон-интронное) строение белок-кодирующих (и некоторых РНК-кодирующих) генов эукариот – поистине удивительная черта (она не всегда кажется нам таковой только потому, что нам так привычна концепция сплайсинга, да и открыта она, на момент написания этой книги, более 30 лет назад). Почему гены разрываются множеством некодирующих участков, большинство которых не выполняет никаких видимых функций и вырезается из транскрипта сложной молекулярной машиной (предназначенной специально для этой цели) только для того, чтобы быть разрушенными? Кажется, это превосходит все, что может нарисовать себе самое смелое воображение. Когда интроны были открыты в 1977 году, Уолтер Гилберт ^[66] немедленно выступил с привлекательной «гипотезой ранних интронов», послужившей основой так называемой «экзонной теории генов» (Gilbert, 1978). Вкратце, Гилберт предположил, что интроны сопутствовали жизни на самых ранних этапах эволюции и играли ключевую роль в эволюции белок-кодирующих генов, позволяя соединять короткие последовательности, кодирующие первичные пептиды, путем рекомбинации ближайших некодирующих последовательностей. Вслед за формулировкой этой идеи последовали 20 лет попыток подтвердить существование первичных интронов путем анализа разнообразных признаков интронов ныне существующих (de Souza et al., 1998). Мы не станем описывать здесь эту борьбу; достаточно лишь сказать, что убедительных свидетельств найдено не было. Безусловно, гипотеза ранних интронов не поддерживается тем фактом, что у прокариот не обнаружены сплайсосомы и интроны сплайсосомного типа, хотя Гилберт и его коллеги утверждают, что это результат эволюционного упрощения. Возможно, самый серьезный аргумент против «ранних интронов» – это обнаружение предковой связи между бактериальными самосплайсирующимися интронами и сплайсосомными интронами. Из этого открытия следует, что, даже если на самых ранних стадиях эволюции жизни существовали интроны (мы еще вернемся к этому моменту в гл. 10 и 11), эти интроны были совершенно не такими, как современные сплайсосомные интроны, и последние не могут нести какую-либо «память» о первичной эволюции. Сплайсосомные интроны и вся система сплайсинга, таким образом, суть чисто эукариотические черты из тех, что определяют «эукариотическое состояние».

Так почему же эукариотические гены прерываются таким множеством интронов? По-видимому, единственный разумный ответ следующий: потому что их предки агрессивно встраивались в эукариотические гены во время эукариогенеза или вскоре после него, а механизмы, служащие для эффективного удаления их из первичных транскриптов, развились и обеспечили выживание линии организмов со странными «кусочными» генами. После этого давление отбора, направленное на устранение интронов, во многих линиях эукариот оказалось недостаточно сильным, чтобы избавиться от большинства из них, хотя именно это произошло в других линиях, эволюционировавших в условиях сильного очищающего отбора (см. гл. 8). Разумеется, это не отменяет функциональной значимости интронов вообще: известно, что некоторые из них вносят вклад в регуляцию экспрессии (Le Hir et al., 2003), а другие даже содержат встроенные гены (Assis et al., 2008). Более того, интроны обеспечивают возможность альтернативного сплайсинга, ключевого механизма, создающего структурное и функциональное разнообразие белков у многоклеточных эукариот (см. гл. 8). В целом, однако, неоспоримое присутствие интронов, по-видимому, в большой степени зависит от силы очищающего отбора, направленного на их устранение. Популяционно-генетические аспекты утраты и приобретения интронов будут рассмотрены в главе 8; здесь я кратко обсуждаю

результаты реконструкций эволюции интронов и некоторые дополнительные идеи, касающиеся природы геномов самых ранних эукариот в свете рассмотренного выше сценария эукариогенеза.

Эукариоты сильно различаются по характеру интронов: многие протисты и одноклеточные грибы содержат всего несколько интронов на весь геном, тогда как животные, растения и некоторые из простейших богаты интронами, так что кодирующие последовательности большинства их генов прерываются несколькими интронами (Jeffares et al., 2006). Замечательно, что позиции большой доли интронов консервативны у ортологичных генов разных организмов, включая растения и животных (Rogozin et al., 2003). Эволюционные реконструкции, принимающие во внимание консервативные и переменные позиции интронов, приводят к неожиданному выводу о том, что гены LECA были почти так же насыщены интронами, как и у современных млекопитающих, и значительная часть интронов LECA сохранилась по сей день в тех же позициях (см. рис. 7–8; Csuros et al., 2011). Этот вывод может казаться странным, но чем больше геномов становится доступно для анализа все более точными методами реконструкции, тем более убедительным он оказывается. Из этого наблюдения, формально подтвержденного результатами реконструкций, следует, что дальнейшая эволюция вела в первую очередь к утрате интронов, происходившей в большинстве ветвей эукариот, а немногие эпизоды взрывного увеличения их количества, по-видимому, были связаны с появлением новых крупных ветвей, таких как растения и животные (см. рис. 7–8). Резкое увеличение числа интронов у основания супергруппы *Plantae* могло быть обусловлено новой волной интронов группы II, перешедших от цианобактериального симбионта. Источник интронов у основания ветви животных остается загадочным и даже может свидетельствовать о роли скрытого эндосимбиоза в происхождении животных.

Реконструкция (см. рис. 7–8) была проведена с использованием методов Монте-Карло и марковских цепей (Csuros et al., 2011). Показана плотность интронов (число интронов на 1 Кб) для ныне существующих форм и предполагаемая плотность для ключевых предковых форм. Насыщенность черной штриховки приблизительно пропорциональна плотности интронов. Линия человека отмечена кружком. Показаны три супергруппы эукариот (*Chromalveolata*, *Unikontae* *Plantae*) и основные группы внутри каждой из них, для которых известны полные последовательности геномов и соответствующие данные по локализации интронов [\[67\]](#).

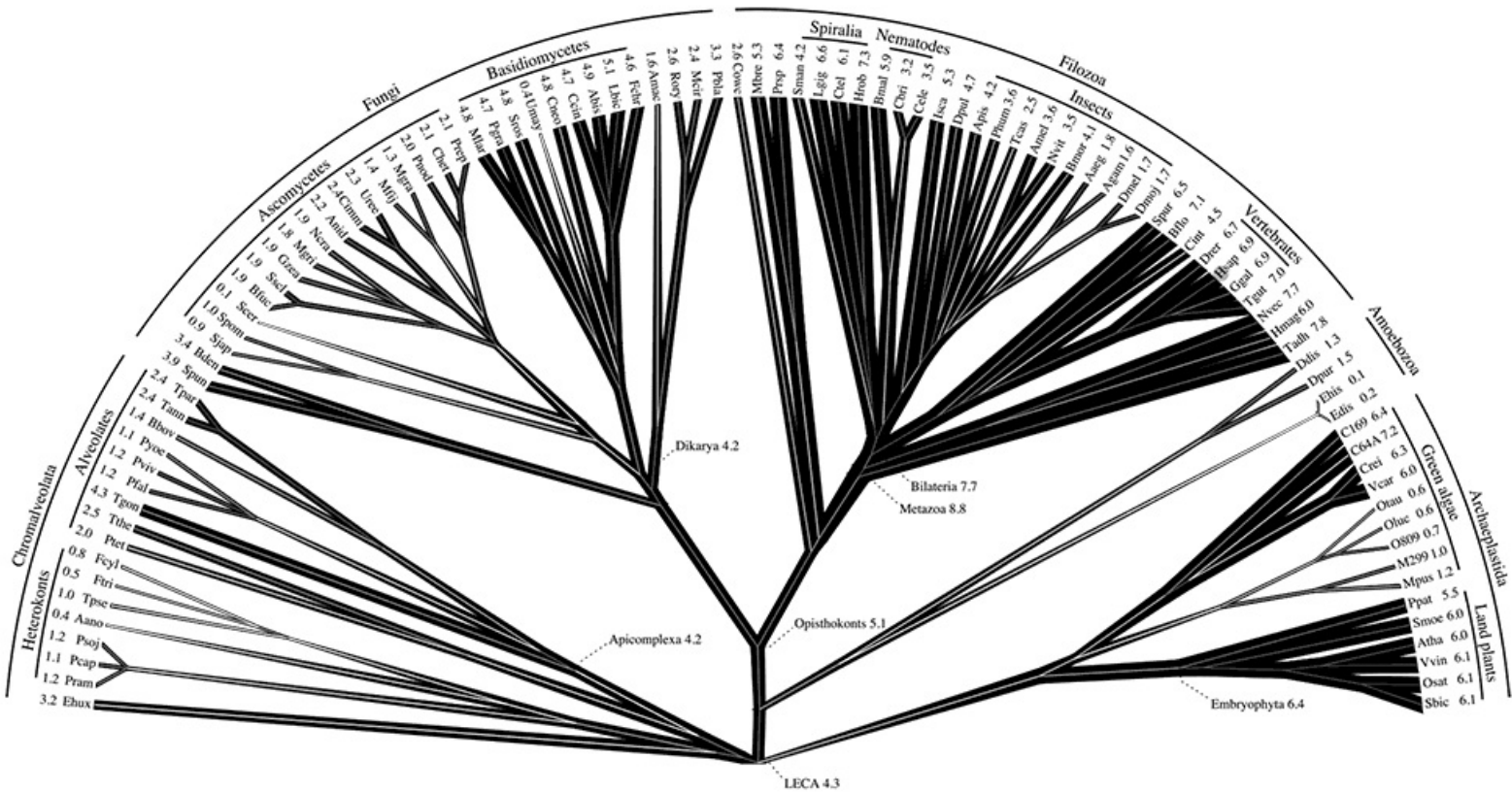


Рис. 7–8. Реконструкция приобретения и утраты интронов в течение эволюции эукариот и плотности интронов у предковых форм.

По: Csuros et al., 2011. Данная статья находится в свободном доступе под Атрибутивной лицензией *Creative Commons*.

Таким образом, LECA, по-видимому, аккумулировал интроны до плотности, близкой к наиболее богатым интронами современным геномам. Что же можно сказать о динамике интронов в течение стволовой фазы, между эукариогенезом и LECA? Простой расчет показывает, что, если инвазия интронов произошла «мгновенно», то протоэукариотический геном должен был большей частью (до 80 процентов) состоять из интронов, учитывая большие и единообразные размеры интронов группы II (около 2,5 Кб; Koopin, 2009b). Скорее всего, это чрезмерное упрощение. Процесс аккумуляции интронов, вероятно, был более постепенным и сопровождался уменьшением встроившихся интронов. Как бы то ни было, интроны, по всей видимости, сыграли ключевую роль в самом начале эволюции эукариот, согласно обсуждаемой модели эукариогенеза.

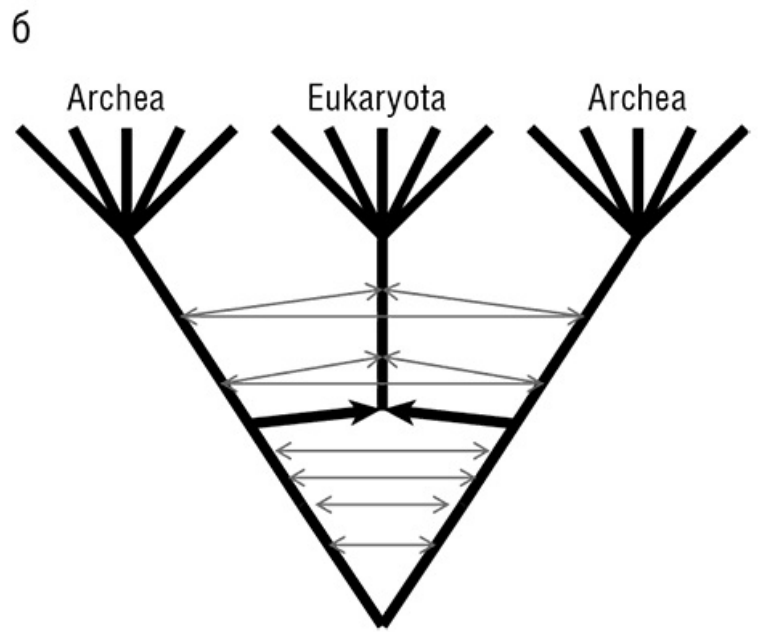
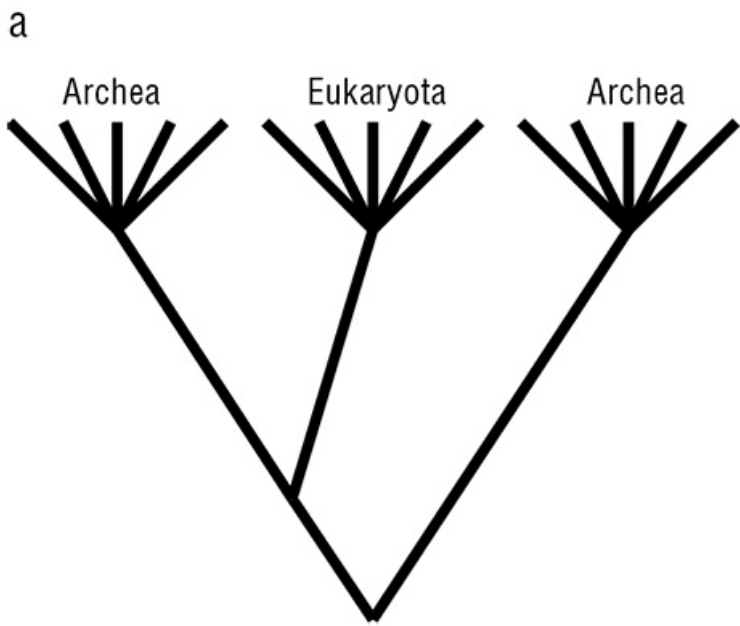


Рис. 7–9. Пересмотр трех доменов жизни: а– традиционное трехдоменное дерево Вёзе; б– циклический граф трех доменов согласно симбиогенетическому сценарию эукариогенеза и ГПГ между доменами.

Три домена жизни: за пределами дерева Вёзе

Симбиогенетический сценарий эукариогенеза ведет к пересмотру трехдоменного дерева жизни, разработанного Вёзе и его коллегами, даже если не учитывать последствия широкомасштабного ГПГ между прокариотами, что обсуждалось в главах 5 и 6. Дерево Вёзе подразумевает архезойный сценарий, а эндосимбиоз рассматривается как относительно позднее событие в истории эукариотического домена, которое не имеет отношения к трехдоменной классификации жизни (см. рис. 7–9 а). Напротив, симбиогенетический сценарий постулирует, что захват первичного эндосимбионта буквально дал начало эукариотическому домену и, в процессе, внес большое число (возможно, большинство) генов в развивающийся эукариотический геном. Согласно последнему сценарию, слияние организмов из двух первичных доменов дало начало третьему домену; тогда конечная диаграмма не будет представлять собой дерево (см. рис 7–9 б). Важное следствие, к которому мы вернемся в главе 11, состоит в том, что, размышляя о происхождении клеток, мы должны думать только о двух прокариотических доменах: археях и бактериях.

Краткий обзор главы

Среди трех доменов жизни эукариоты, без сомнений, обладают самой сложной, поразительно изощренной клеточной организацией, которая может даже навести некоторых на мысли о «неупрощаемой сложности» (Kurland et al., 2006), так как для большинства характерных функциональных систем эукариотических клеток мы не можем найти эволюционные промежуточные формы. Естественно, эукариогенез представляется одной из важнейших проблем эволюционной биологии, задачей, для решения которой «в лоб» мы гораздо лучше оснащены, чем для решения более фундаментальных задач происхождения клетки и, в конце концов, происхождения жизни (см. гл. 11 и 12). Сравнительная геномика на сегодняшний день не смогла ни разгадать загадку эукариогенеза, ни предложить исчерпывающую картину первичного расхождения главных эукариотических линий. Тем не менее филогенетический анализ во многих случаях внес ясность в вопросы, связанные с происхождением и самыми ранними этапами эволюции эукариот. Так, филогеномика прояснила эволюционные связи между царствами эукариот и привела к выявлению пяти или шести супергрупп. Родственные связи между супергруппами и корнем древа эукариот остаются чрезвычайно сложными для расшифровки, возможно по причине сжатого кладогенеза при первичном расхождении главных ветвей эукариот. Продолжающийся сбор данных по геномам различных ветвей жизни – отнюдь не игра в бирюльки; напротив, сравнительный анализ различных геномов продолжает приносить неожиданные биологические открытия, так что ожидать, безусловно, следует еще большего.

Ультраструктурные, функциональные и геномные данные согласованно свидетельствуют о том, что эукариоты являются археобактериальными химерами. Более того, гены явно бактериального происхождения существенно превышают по численности «архейные» гены. Более того, сравнительный анализ растущей коллекции доступных геномов архей все увереннее показывает, что многие ключевые клеточные системы эукариот существуют в примитивных формах уже у архей. Вариабельность этих систем между разными линиями архей, вместе с филогенезом консервативных белков, говорит о том, что архейный предок эукариот принадлежал к глубокой, возможно вымершей ветви архей с геномом высокой сложности и разработанными клеточными функциями. Недавнее открытие возможного прямого предка убиквитиновой системы в новом архейном геноме говорит о том, что мы можем в настоящее время недооценивать, насколько многие типичные функциональные системы эукариот могли быть представлены в течение эволюции *Archaea*. Это и другие подобные открытия внушают доверие к «комбинированному сценарию» происхождения эукариот, согласно которому эти заранее сформировавшиеся системы случайно сочетались в архейном хозяине первичного эндосимбионта. Напротив, системы внутренних мембран эукариот – особенно ядро, с его сложным комплексом ядерной поры – не обнаружены у архей и, по-видимому, были собраны, по крайней мере частично, из предковых бактериальных компонентов. *Представляется важным, что эукариоты унаследовали развитые, сложно организованные системы от архей (естественно, за исключением митохондрий), тогда как многочисленные бактериальные молекулярные компоненты были в основном унаследованы по одному, а новые молекулярные машины возникли путем рекомбинации. Это различие, по-видимому, отражает асимметрию между хозяином и эндосимбионтом: несмотря на все драматические инновации, сопровождавшие эукариогенез, многие клеточные системы архейного хозяина сохранились и изменялись только эволюционным (плавным) путем, посредством дубликаций и приобретения дополнительных деталей.*

Все вместе эти наблюдения, по-видимому, лучше всего совместимы с симбиогенетическим сценарием происхождения эукариот. Согласно этому сценарию, эукариогенез был инициирован эндосимбиозом альфа-протеобактерии с предковыми археями, а система внутренних мембран, в том числе ядро, появилась как защита против вторжения интронов. Более того, остальные ключевые нововведения эукариотической клетки, такие как нонсенс-опосредованный распад ошибочных транскриптов и значительное усложнение убиквитин-зависимой системы деградации аберрантных белков, по-видимому, логически объясняются как дополнительные линии защиты против той же инвазии. Косвенно гипотеза защиты может вносить вклад в понимание эволюции других главных черт эукариот, таких как исчезновение оперонов и переход от кольцевых хромосом к линейным. Подводя итог, можно сказать, что сейчас мы имеем достаточно согласованную, хотя, конечно, слишком схематичную хронику эукариогенеза. Завершая эту главу, я хочу подчеркнуть, что, несмотря на многочисленные детали, остающиеся неясными, история эукариогенеза является идеальной демонстрацией главной темы настоящей книги: взаимодействия между случайностью и необходимостью в эволюции жизни. В самом деле, захват протомитохондриального эндосимбионта, несомненно, был ключевым событием эукариогенеза, и партнеры этого симбиоза были «выбраны» случаем. Тем не менее этот симбиоз, судя по всему, инициировал сложную цепь событий, многие элементы которой были необходимы для того, чтобы обеспечить выживание химерного организма, и мы знаем, что на нашей планете эукариоты действительно выжили и достигли беспрецедентной сложности и разнообразия.

Рекомендуемая дополнительная литература

Doolittle W. F. (1998) You Are What You Eat: A Gene Transfer Ratchet Could Account for Bacterial Genes in Eukaryotic Nuclear Genomes. *Trends in Genetics*14: 307–311.

По-видимому, первое описание хrapовика переноса генов от эндосимбионтов к хозяину, правда в контексте архезойного сценария.

Embley T. M., and W. Martin.(2006) Eukaryotic Evolution, Changes, and Challenges. *Nature*440: 623–630.

Глубокий аналитический обзор различных сценариев эукариогенеза в свете представления о том, что все эукариоты обладают митохондриями или МПО.

Koonin E. V.(2010) The Origin and Early Evolution of Eukaryotes in the Light of Phylogenomics. *Genome Biology*11: 209.

Рассмотрение родственных связей между супергруппами эукариот, природы LECA и эукариогенеза, с упором на высокую сложность LECA.

Koonin E. V.(2006) The Origin of Introns and Their Role in Eukaryogenesis: A Compromise Solution to the Introns-Early Versus Introns-Late Debate? *Biology Direct*1: 22.

Дальнейшее развитие защитного сценария эукариогенеза, запущенного инвазией интронов и включающего единую цепь причинно-следственных связей, обуславливающих происхождение различных специфичных для эукариот функциональных систем.

Kurland C. G., L. J. Collins, and D. Penny.(2006) Genomics and the Irreducible Nature of Eukaryote Cells. *Science*312: 1,011—1,014.

Воодушевленное выступление против эндосимбиотического сценария эукариогенеза и за первичное происхождение сложности эукариот.

Lane N., and W. Martin.(2010) The Energetics of Genome Complexity. *Nature*467: 929–934.

Гипотеза о неизбежности эндосимбиоза, который интерпретируется как единственный путь к эффективной биоэнергетике, необходимой для эволюции больших сложных клеток.

Martin W., and E. V. Koonin.(2006) Introns and the Origin of Nucleus-Cytosol Compartmentalization. *Nature*440: 41–45.

Гипотеза о защите от инвазии интронов как факторе отбора, обусловившем возникновение ядра.

Martin W., and M. Muller.(1998) The Hydrogen Hypothesis for the First Eukaryote. *Nature*392: 37–41.

Ключевая гипотеза о метаболической кооперации как факторе отбора, способствующем мутуалистическим связям между хозяином и эндосимбионтом.

Martin W., T. Dagan, E. V. Koonin, J. L. Dipippo, J. P. Gogarten , and J. A. Lake.(2007) The Evolution of Eukaryotes. *Science*316: 542–543.

Опровержение аргументации Курланда с соавторами (Kurland et al., 2006).

Zimmer C.(2009) Origins. On the Origin of Eukaryotes. *Science*325: 666–668.

Популярное обсуждение различных сценариев эукариогенеза.

Глава 8. Неадаптивная нулевая гипотеза эволюции генома и истоки биологической сложности

Немногие модные слова в последние два десятилетия были настолько популярны и в то же время определялись столь разнообразно, зачастую противоречиво, а иногда и обманчиво, как *сложность* [68]. Несмотря на эту суету, понятие сложности, очевидно, отражает общее, фундаментально важное явление, пронизывающее всю биологию и выходящее за ее рамки. В отличие от многих научных терминов, «сложность» имеет конкретное значение в обыденном языке. Мы узнаем ее, как и порнографию [69], с первого взгляда. Все признают, что млекопитающее или птица сложнее, чем червь, а червь сложнее, чем любой одноклеточный организм. Говоря интуитивно, здесь присутствует дополнительный оттенок, устанавливающий пропорциональность сложности с «развитостью» или «приближением к совершенству».

Поднимаясь уровнем выше чистой интуиции, спросим, что означает большая сложность млекопитающего по сравнению с амебой? Этот вопрос очень важен, если мы стремимся выработать удовлетворительный ответ на известный вопрос: почему вокруг нас существуют слоны и секвойи (даже если их все меньше и меньше), а не одни лишь бактерии и археи с необходимыми и достаточными для функционирования минимальной клетки комплектами генов? Другими словами, какие факторы ведут к появлению сложности в процессе эволюции? В главе 7 мы обсуждали эволюционные сценарии, пытающиеся объяснить, как могла возникнуть поразительно сложная (по сравнению с клетками прокариот) организация эукариотической клетки. В этой главе мы столкнемся с озвученным выше вопросом «почему?» напрямую, и ответы на него будут неожиданными и, возможно, введут некоторых в замешательство.

Точное определение *организационной*– или, в случае биологии, организменной – сложности по самой своей природе дается трудно. Попытки в этом направлении рассматривают различное число составляющих частей в сравниваемых системах [70]. Например, у позвоночных большее количество тканей и типов клеток, чем у червей, и это, естественно, приводит к утверждению, что позвоночные обладают большей организменной сложностью (Bonner, 2004). Для нашего рассуждения, однако, более важен тот факт, что эукариотические клетки имеют гораздо больше внутриклеточных органелл, чем клетки прокариот (те, как правило, вообще не имеют настоящих органелл). Эта разница, безусловно, отражает большую сложность организации эукариотической клетки. Кроме того, можно было бы в принципе измерить число взаимодействий между компонентами или число соединений в сетях передачи сигнала и на этом основании сравнивать сложность организмов или клеток. Однако все эти определения сложности, видимо, упускают «нечто», что мы интуитивно воспринимаем как неотъемлемое свойство сложной организации. В любом случае количественное сравнение организменной сложности, по-видимому, не приносит много пользы в реальных исследованиях. Геномная сложность определяется более естественно и может быть изучена подробнее. Действительно, в конце концов, геномные последовательности представляют из себя длинные строки цифровых символов (букв), а для этого класса объектов хорошо известны формальные, операциональные определения сложности. Вероятно, наиболее известным и наиболее интуитивно осмысленным из них является колмогоровская сложность, которая связана с шенноновской информацией и классическим статистическим определением энтропии по Больцману. Колмогоровская сложность – это просто длина кратчайшей строки символов, в которых может быть закодирована данная последовательность (геном). Очевидно, что наименее сложной последовательностью будет гомополимер (например, polyA), для которого длина сообщения составляет лишь одну букву, а сложность (информационное содержание) – 2 бита (в случае четырех нуклеотидов). Наиболее же сложная последовательность – полностью случайный

полимер с равными частотами для всех четырех нуклеотидов (или 20 аминокислот, если мы примем это определение для аминокислотных последовательностей) в каждой позиции. Классическая формула Шеннона для энтропии (информационного содержания) нуклеотидной последовательности длины L (см. рис. 8–1 а) может быть записана следующим образом:

$$H(L) = \sum_{i=1}^L f_i \log f_i$$

а

TTATGCACATTTACAGCTACATATGCAGAC

$$f_T = 9/30$$

$$f_C = 7/30$$

$$f_A = 10/30$$

$$f_G = 4/30$$

$$H = -(f_T \log_4(f_T) + f_C \log_4(f_C) + f_A \log_4(f_A) + f_G \log_4(f_G)) = 0.96$$

б

...TGC...

...TGC...

...AAC...

...TGC...

...GAC...

...CGG...

...AGC...

...CAC...

...AGC...

...TCC...

$$f_T = 4/10$$

$$f_C = 2/10$$

$$f_A = 3/10$$

$$f_G = 1/10$$

$$H = 0.92$$

$$C = 0.08$$

$$f_T = 0/10$$

$$f_C = 1/10$$

$$f_A = 2/10$$

$$f_G = 7/10$$

$$H = 0.58$$

$$C = 0.42$$

$$f_T = 0/10$$

$$f_C = 9/10$$

$$f_A = 0/10$$

$$f_G = 1/10$$

$$H = 0.23$$

$$C = 0.77$$

Рис. 8–1. Содержание информации и сложность: а– одной последовательности; б– выравнивания гомологичных последовательностей; $f_{об}$ обозначает частоты нуклеотидов в последовательности (а) или столбце выравнивания (б).

Здесь f_i – частота символа i ($i = A, T, G, C$); далее, основание логарифма m считается равным размеру алфавита (4 в случае нуклеотидных последовательностей и 20 для аминокислотных последовательностей) [71]. Определенная таким образом, информация (энтропия) говорит нам очень мало об осмысленном информационном содержании или сложности геномной последовательности. Высокая сложность (энтропия или информационное содержание), очевидно, вовсе не предполагает, что последовательность сложна в каком-либо биологическом

значимом смысле. Совершенно случайная последовательность на самом деле, скорее всего, бессмысленна, в то время как гомополимерная последовательность будет иметь ограниченный биологический смысл. Тем не менее почти случайная высокоэнтропийная последовательность может быть столь же функциональной, как и низкоэнтропийная последовательность, – способа узнать это просто не существует. Требуется биологически содержательное определение сложности, и такая попытка была сделана Крисом Адами (Adami, 2002) и несколько по-другому проинтерпретирована автором этой книги (Koonin, 2004). В соответствии с этим новым определением, энтропия и сложность рассчитываются для выравнивания ортологических последовательностей, а не одной последовательности:

$$H(L) = \sum_{i=1}^L H_i = \sum_{i=1}^L \sum_j f_{ij} \log f_{ij}$$

Здесь $H(L)$ – полная энтропия выравнивания n последовательностей длины L , H_i – энтропия для сайта, а f_{ij} – частоты для нуклеотидов ($j = A, T, G, C$) в сайте i [72]. Очевидно, для полностью консервативного сайта $H(i) = 0$, в то время как для совершенно случайного сайта $H(i) = 1$. Обратите внимание, что это определение энтропии полностью соответствует знаменитому статистическому определению Больцмана:

$$H = k \ln W$$

Здесь W – число микросостояний, соответствующих макросостоянию, для которого энтропия рассчитывается таким образом, что она равна нулю для полностью упорядоченного состояния и максимальна для полностью неупорядоченного состояния. Таким образом, определение *эволюционной энтропии* $H(L)$, введенной предыдущей формулой, представляется физически корректным, следовательно, имеет смысл закрепить термин за обозначением этой величины. Эволюционная энтропия также имеет четкий биологический смысл: сайты с низкой энтропией сохраняются лучше и, как следствие, более важны функционально. Логично, что эти сайты несут больше информации о функционировании и эволюции рассматриваемых организмов – и о взаимодействиях между организмами и окружающей средой, что первоначально имел в виду Адами, – чем сайты с высокой энтропией (слабо сохраняемые, относительно неважные). Величина, которую можно определить как *биологическую (эволюционную) сложность генома*, определяется следующим образом:

$$C(N) = N - \sum_{i=1}^k H(L_i)$$

Тогда биологическая (эволюционная) плотность информации может быть задана как:

$$D(N) = C(N)/N = (N - \sum_{i=1}^k H(L_i)) / N = 1 - \sum_{i=1}^k H(L_i) / N$$

Здесь N —общая длина (число нуклеотидов) генома, L_i —длина геномного сегмента, подверженного измеримому отбору (как правило, ген), k —число таких сегментов в геноме, $H(L)$ — эволюционная энтропия для сегмента L , рассчитанная по предыдущей формуле.

Точные значения H нелегко вычислить для полных геномов, поскольку распределение эволюционных ограничений никогда не известно точно (см. гл. 3). Кроме того, есть степень произвольности в выборе ортологов, включаемых для расчета в выравнивание. Тем не менее эти детали не столь важны, если нам нужна только приблизительная оценка. Действительно, доля сайтов, находящихся под отбором по всему геному, уже оценена с достаточной точностью для некоторых модельных организмов, таких как человек и дрозофила (см. гл. 3). Для других, в частности прокариот и одноклеточных эукариот, в качестве достаточного приближения можно взять долю кодирующих нуклеотидов плюс предполагаемую долю регуляторных сайтов; для участков под отбором за среднее значение энтропии можно принять $H(i) = 0,5$.

Сравнение оценок для $H(N)$, $C(N)$ и $D(N)$ для геномов различных жизненных форм выявляет фундаментальный парадокс. Общая биологическая сложность $C(N)$ монотонно возрастает с размером генома, в частности, для многоклеточных эукариот по сравнению с прокариотами, однако энтропия $H(N)$ возрастает гораздо быстрее, в результате эволюционная плотность информации $D(N)$ резко падает (см. рис. 8–2). Таким образом получается, что организмы, которые обычно воспринимаются как наиболее сложные (к примеру, человек), обладают «энтропийными» геномами с низкой или даже крайне низкой плотностью информации, в то время как организмы, которые мы традиционно считаем примитивными, такие как бактерии, обладают «информационными» геномами, в которых информация плотно упакована и плотность ее высока. Этот парадокс не даст нам много нового по сравнению с уже сказанным в главе 3 об организации различных геномов. Тем не менее поучительно формализовать понятие биологической сложности и выразить его в терминах энтропии, одного из ключевых понятий физики. Формальный разбор понятия сложности указывает на то, что «неладно что-то в Датском королевстве»: геномы организмов, которых мы вполне обоснованно считаем самыми сложными и наиболее «развитыми» (эта идея, возможно, менее оправдана), несут гораздо больше энтропии и, следовательно, имеют гораздо меньшую плотность биологической информации, чем геномы простейших клеточных форм. Перефразируя этот парадокс в более провокационной форме, геномы одноклеточных организмов (особенно прокариот) кажутся несравненно «лучше спроектированными», чем геномы растений и особенно животных.

Парадокс сложности подразумевает, что сложные черты организации геномов «высших» организмов (большие семейства паралогичных генов, сложная регуляция экспрессии генов, альтернативный сплайсинг и многое другое), вероятно, появились не в качестве прямолинейных адаптаций или «улучшений». Объяснение возникновения этих усложнений — большая проблема для эволюционной биологии; возможный ответ пришел в виде новой теории эволюции сложности, предложенной Майклом Линчем в 2003 году (Lynch and Conery, 2003).

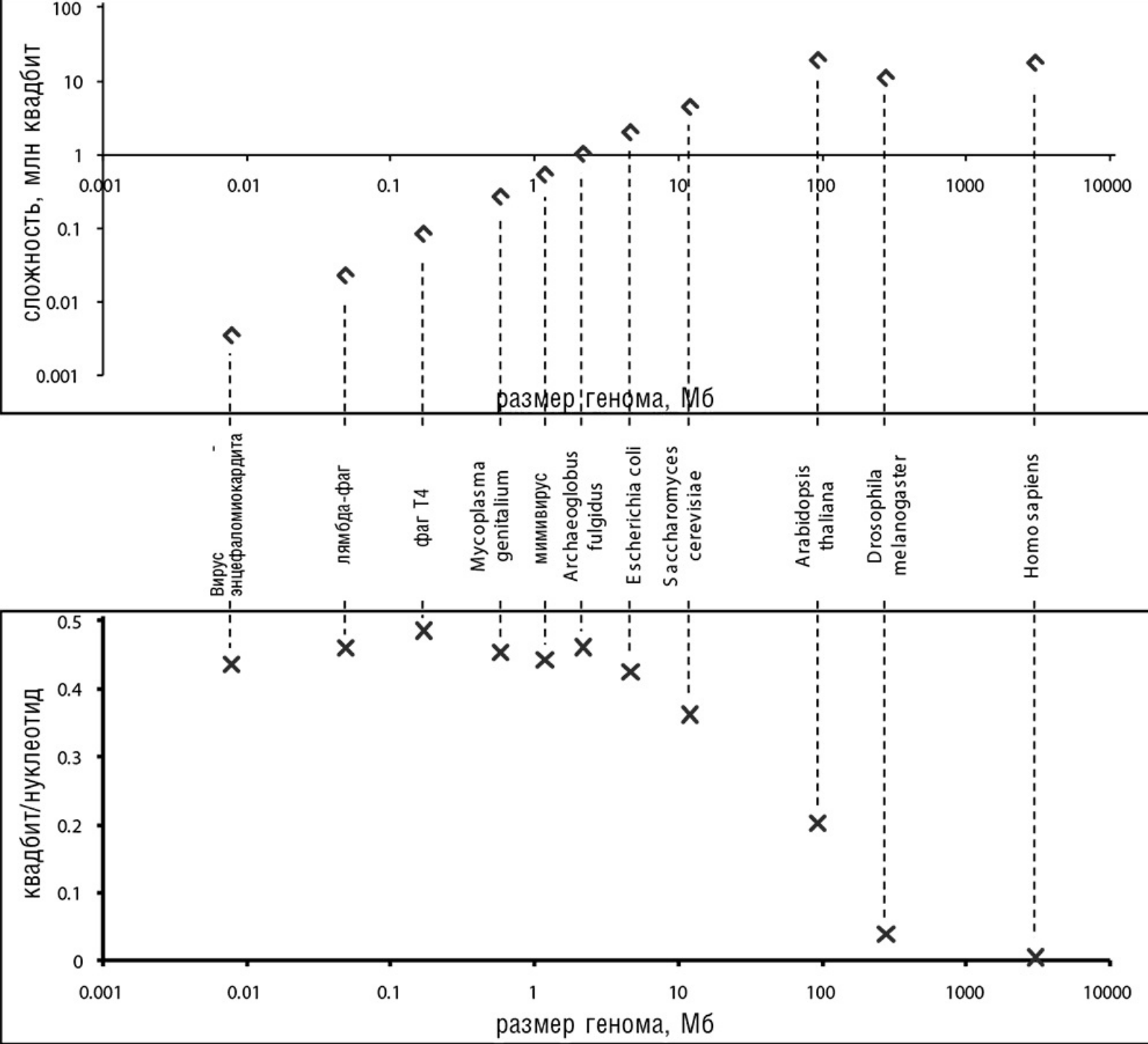


Рис. 8–2. Зависимость эволюционной сложности S_i плотности биологической информации D_{ot} размера генома. Точки – грубые оценки, полученные по приведенным в этой главе формулам, при условии $H(i) = 0,5$ для несинонимичных сайтов в белок-кодирующих областях и $H(i) = 1$ для других сайтов. График выполнен в двойном логарифмическом масштабе.

Эффективный размер популяции как общая мера эволюционных ограничений: неадаптивная теория эволюции генома

Как уже говорилось в предыдущем разделе, наиболее сложные организмы на Земле располагают «высокоэнтропийными» геномами, которые представляются крайне неэффективными и «плохо спроектированными». То, что адаптивная эволюция приводит к таким результатам, на первый взгляд кажется просто невероятным. Неформально мотивацию для новой теории эволюции геномной сложности можно изложить следующим образом. Геномы сложных организмов содержат различные особенности, которые существенны для их организационной сложности, но кажутся бесполезными и, следовательно, на момент своего появления по крайней мере слегка вредны. Наиболее известной такой особенностью в геномах многоклеточных эукариот являются интроны, которые обеспечивают возможность альтернативного сплайсинга, происходящего в большинстве генов млекопитающих и представляющего собой главную основу разнообразия протеома (Blencowe, 2006; Wang et al., 2008), а также дублицированные гены, являющиеся основным источником эволюционных новшеств и разнообразия для эукариот (Lespinet et al., 2002; Lynch and Conery, 2000). Эти геномы также несут в себе многочисленные эгоистичные элементы и прочую ДНК, не подверженную отбору и, в меру нашего понимания, являющуюся «мусором». Сохранение всех этих последовательностей в сложных геномах естественно объясняется слабым (неэффективным) очищающим отбором и, наоборот, большой ролью дрейфа в эволюции данных организмов.

В теории популяционной генетики эффективность очищающего, а также положительного отбора пропорциональна эффективному размеру популяции (N_e) для данного организма, в предположении постоянной скорости мутаций. Только те мутации, для которых $|s| \gg 1/N_e$ (где s – коэффициент отбора, то есть разница в приспособленности между диким типом и соответствующим мутантом), могут быть эффективно зафиксированы (положительный отбор) или отбракованы (очищающий отбор) в ходе эволюции. С другой стороны, мутации с $|s| \ll 1/N_e$ являются «невидимыми» для отбора. Эта простая зависимость, возможно, является основным фактором, определяющим ограничения, которые влияют на различные аспекты эволюции генома и фенотипа, в частности на фиксацию «украшений», характерных для геномов сложных организмов (Lynch, 2007b, 2007c; Lynch and Conery, 2003).

Действительно, различия в N_e , по-видимому, лежат в основе описанного выше качественного различия между архитектурами геномов одноклеточных и многоклеточных организмов. Существенное увеличение размеров генома представляется достижимым только в организмах с небольшими популяциями и сопутствующим слабым отбором. К сожалению, эффективный размер популяции оценить непросто, хотя имеющиеся грубые оценки варьируют в огромном диапазоне: от порядка 10^9 для бактерий до 10^5 и менее для животных (см. табл. 8–1). Более доступные оценки из уровня геномного полиморфизма, которые, как мы увидим в следующем разделе, могут быть даже более актуальными для понимания эволюции геномной сложности, приводятся для произведения $N_e u$, где u – частота мутаций на сайт. Значения $N_e u$ изменяются в масштабах примерно двух порядков величины: от около 0,001 у бактерий до примерно 0,1 у позвоночных (см. табл. 8–1). Предсказывается, что сила отбора варьирует соответственно. Как результат, у прокариот, с их типично большими популяциями, даже очень слабавредные мутации со значениями спорядка 10^{-8} эффективно отбраковываются, и, напротив, в малых популяциях многоклеточных эукариот лишь мутации с относительно большими s

порядка 10^{-4} , влекущие за собой существенные последствия для приспособленности, будут уничтожены очищающим отбором. Как мы увидим в следующих разделах, это различие имеет решающее значение для хода эволюции, поскольку значения s для главных «украшений» в сложных геномах, например интронов, находятся в пределах этого диапазона. Таким образом, они в основном устраняются очищающим отбором у организмов с большим N_e , но не у организмов с малым N_e . *Эволюционное сохранение любого геномного элемента не означает автоматически, что данный элемент удерживается от отбраковки отбором в силу своего функционального значения; как это ни парадоксально, такая эволюционная консервативность может отражать слабый очищающий отбор, недостаточный для устранения неадаптивных предковых особенностей* (Koonin and Wolf, 2010b).

Конечно, N_e не является постоянной на всем протяжении эволюционной истории линии. Напротив, почти неизбежно происходят большие колебания, что приводит к популяционным «бутылочным горлышкам» (участки с низкой N_e), в течение которых эволюция почти полностью зависит от дрейфа, так что многочисленные мало и даже умеренно вредные мутации генома могут быть зафиксированы, обеспечивая сырьем дальнейшую эволюцию. Важно иметь в виду, что даже популяции с большим N_e могут фиксировать слабовредные мутации посредством генетической тяги и хитчхайкинга (см. гл. 2) и, кроме того, нести большой запас нейтральных и слабовредных мутаций, которые не фиксируются, но могут сохраняться в популяции в качестве полиморфизмов в течение длительного времени. Некоторые из этих персистирующих нефиксированных мутаций могут быстро фиксироваться, когда давление отбора меняется и мутация становится выгодной либо когда новая мутация создает полезное сочетание с одним из стойких полиморфизмов.

Эта несложная (и представленная здесь в нарочито упрощенном виде) теория, основанная на популяционной генетике, *задает нулевую гипотезу для эволюции генома* (Koonin, 2004). В следующих разделах мы рассмотрим эту теорию более подробно и, главное, увидим, выдержит ли она проверку данными сравнительной геномики.

Таблица 8–1. Характеристики популяции и особенности организации генома для различных клеточных форм жизни.

Организмы	N_e	$N_e u$	Типичный размер генома, Мб	Плотность генов, ген/Кб	Плотность информации, бит/нуклеотид	Плотность интронов, интрон/Кб
Бактерии, археи	10^8-10^{10}	0,01—1	0,5—10	≈ 1	$\approx 0,4$	—
Одноклеточные эукариоты	$\approx 10^7$	0,01—0,1	5—30	$\approx 0,5$	$\approx 0,3$	0—2
Однолетние растения	$\approx 10^6$	$\approx 0,01$	$(0,1-1)\times 10^3$	0,1—0,2	$\approx 0,2$	5—6
Беспозвоночные	$\approx 10^6$	$\approx 0,01$	$(0,1-1)\times 10^3$	0,1—0,2	$\approx 0,1$	2—7
Деревья	$\approx 10^4$	$\approx 10^{-4}$	$\approx 1\times 10^3$	$\approx 0,01$	$\approx 0,01$	5—6
Позвоночные	$\approx 10^4$	$\approx 10^{-4}$	$(0,5-5)\times 10^3$	$\approx 0,001$	$< 0,001$	5—8

Значения N_e и $N_e u$ по Lynch, 2006; значения плотности интронов по Csuros et al., 2011; для плотности интронов у прокариот прочерк, поскольку сплайсосомы и, соответственно, сплайсосомные интроны у них отсутствуют; значения плотности генов и плотности информации рассчитаны автором.

Генная архитектура эукариот: наглядная демонстрация неадаптивной теории эволюции генома

Эволюция экзон-интронной структуры гена у эукариот (см. также гл. 7) является отличным примером для обсуждения неадаптивной парадигмы популяционной генетики, позволяющим лучше разобраться в теории и ее предсказаниях. Прежде чем мы рассмотрим особенности эволюции генной архитектуры с этой точки зрения, необходимо понять связь между коэффициентом отбора s и грузом вредных мутаций, привнесенным дополнительным элементом генома (Koonin, 2009b; Lynch, 2007b, 2007c). Каждый добавленный в геном элемент увеличивает уязвимость к мутационной инактивации и тем самым «призывает» к отбраковке этого элемента в популяции. Если этот дополнительный элемент требует n нуклеотидов для сохранения функциональности соответствующего гена, это требование, очевидно, открывает возможность для вредных мутаций, так что мутационный груз составляет $s = nu$. Оpoznание и эффективное удаление каждого интрона сплайсосомой требует участия примерно 25–30 нуклеотидов внутри интрона и смежных экзонов, окружающих донорную и акцепторную границы сплайсинга. Тогда условие для фиксации интрона в популяции выглядит как $N_e u \ll 1/n$ или $N_e u \ll 0,04$.

Сравнивая значения $N_e u$ и плотности интронов в табл. 8–1, мы сразу видим отличное соответствие между теорией и наблюдениями. Позвоночные с их низкими значениями $N_e u$, очевидно, находятся значительно ниже порогового значения. Действительно, в генах позвоночных наблюдается самая высокая плотность интронов из всех известных. Кроме того, эволюция позвоночных, по-видимому, включает крайне малый оборот интронов, что совпадает с теоретическим предсказанием о недостаточности силы очищающего отбора для устранения интронов в этих организмах. Беспозвоночные и растения находятся немного ниже порога и имеют промежуточные плотности интронов. В разительном контрасте с ними, большинство одноклеточных эукариот лежат выше порога, даже если и ненамного, и демонстрируют резкое падение плотности интронов (см. табл. 8–1).

Позиции многих интронов сохраняются в ортологичных генах животных и растений (см. гл. 7), таким образом, большинство этих интронов представляют наследие ЛЕСА. Тем не менее представляется, что позиции интронов сохраняются благодаря слабости очищающего отбора, что исключает эффективную отбраковку интронов у организмов с небольшим N_e , а не из-за ограничений на позицию интрона как таковую [73]. Более детальный анализ интронов и интрон-экзонных стыков вскрывает дополнительные факты, кажущиеся необъяснимыми на первый взгляд, но, по всей видимости, отлично согласующиеся с предсказаниями теории (Irimia et al., 2007). Примечательно, что все интроны в бедных интронами геномах одноклеточных эукариот имеют почти одинаковые, по-видимому жестко контролируемые малые размеры и консервативные, оптимизированные сигналы сплайсинга на экзон-интронных стыках. Напротив, в богатых интронами геномах, особенно у позвоночных, интроны часто имеют большую длину и ограничены относительно слабыми, субоптимальными сигналами сплайсинга. Дальнейший анализ эволюции экзон-интронных границ наводит на мысль, что сигналы сплайсинга в богатых интронами геномах все же эволюционировали под действием отбора, направленного на их оптимизацию, но этот отбор был слишком слаб, чтобы компенсировать стохастическое отклонение от консенсусных последовательностей, – что прекрасно согласуется с теорией популяционной генетики (Irimia et al., 2009).

Как говорилось в главе 7, эволюционные реконструкции определенно свидетельствуют о том, что уже ЛЕСА имел высокую плотность интронов, и основная часть дальнейшей эволюции

эукариотных геномов включала в себя потери интронов, которые могли быть либо умеренными, в случае большинства животных и растительных линий, либо чрезвычайно обширными, как у большинства одноклеточных эукариот (Carmel et al., 2007; Csuros et al., 2011). Эпизоды появления новых интронов, по всей видимости, были немногочисленны и разбросаны во времени и были связаны с возникновением новых крупных групп организмов, таких как животные. Последствия этого наблюдения в контексте неадаптивной популяционно-генетической теории эволюции генома весьма интересны. Появляется, по крайней мере в принципе, возможность реконструировать динамику популяций по всей истории всех эукариотических линий исходя из наличных и предполагаемых предковых плотностей интронов. Хотя имеющиеся данные недостаточны для детальной реконструкции, рассмотрение величин на рис. 7–8 уже приводит к интересным выводам. Учитывая, что позвоночные имеют лишь слегка большую плотность интронов, чем у ЛЕСА, что позвоночные и растения совпадают по многочисленным позициям интронов и что повторное встраивание интронов в предковые позиции в сколько-нибудь значительных масштабах крайне маловероятно, по-видимому, бедных интронами промежуточных звеньев вдоль всей эволюционной траектории от ЛЕСА до позвоночных не существовало. Другими словами, *наша эволюционная линия ни разу не проходила через этап высокой эффективной численности популяции и, соответственно, интенсивного отбора за все время эволюции эукариот*. В несколько меньшей степени это относится и к пути от ЛЕСА до растений. Кроме того, эпизоды массового приобретения новых интронов почти наверняка были связаны с популяционными «бутылочными горлышками». Это выглядит весьма логично, если принять во внимание возникновение принципиально новых групп организмов, таких как животные, множества различных инноваций, в том числе обширных дупликаций генов и накопления новых регуляторных элементов, которые возможны только в эволюционном режиме с доминированием дрейфа.

Пожалуй, самый поразительный вывод относится к стволовой фазе эволюции, предшествовавшей ЛЕСА и геномной архитектуре ранних предков эукариот, живших до ЛЕСА. Оценка, основанная на предположении о «мгновенном» вторжении интронов группы II из эндосимбионта в геном хозяина (см. гл. 7), указывает на столь узкое «бутылочное горлышко» ($N_e \approx 1000$, если не меньше), что выживание было бы мало вероятно по чисто стохастическим причинам (Koonin, 2009b). Таким образом, мы вынуждены постулировать до некоторой степени постепенное проникновение интронов в геном хозяина. Тем не менее даже этот сценарий менее разрушительного вторжения предполагает очень длинные и тонкие «бутылочные горлышки» на пути от исходного хозяина эндосимбионта до ЛЕСА (см. рис. 8–3). Такое узкое место, вероятно, будет единственным возможным переходом к появлению эукариотической организации клетки, учитывая многочисленные дупликации и другие новшества, необходимые для эукариогенеза.

Все эти выводы недвусмысленно свидетельствуют в пользу неадаптивной популяционно-генетической теории эволюции генома, что, в сочетании с результатами сравнительной геномики, по-видимому, открывает нам окно в эволюционное прошлое, которое иначе трудно было себе представить.

От мусора к функциональности: важность ослабленного очищающего отбора для эволюции сложности

Что было движущим фактором (или факторами) эволюции геномной (и возможно, связанной с ней организменной) сложности? Неадаптивная популяционно-генетическая теория (Lynch, 2007c; Lynch and Conery, 2003) подсказывает удивительный ответ: *необходимым и, вероятно, достаточным условием для возникновения сложности был неэффективный очищающий отбор в популяциях с небольшим N_e* [74]. Неэффективный отбор способствовал фиксации слегка вредных признаков, которые были бы отбракованы в большой популяции, и накоплению мусора, часть которого затем была задействована в разнообразных функциях.

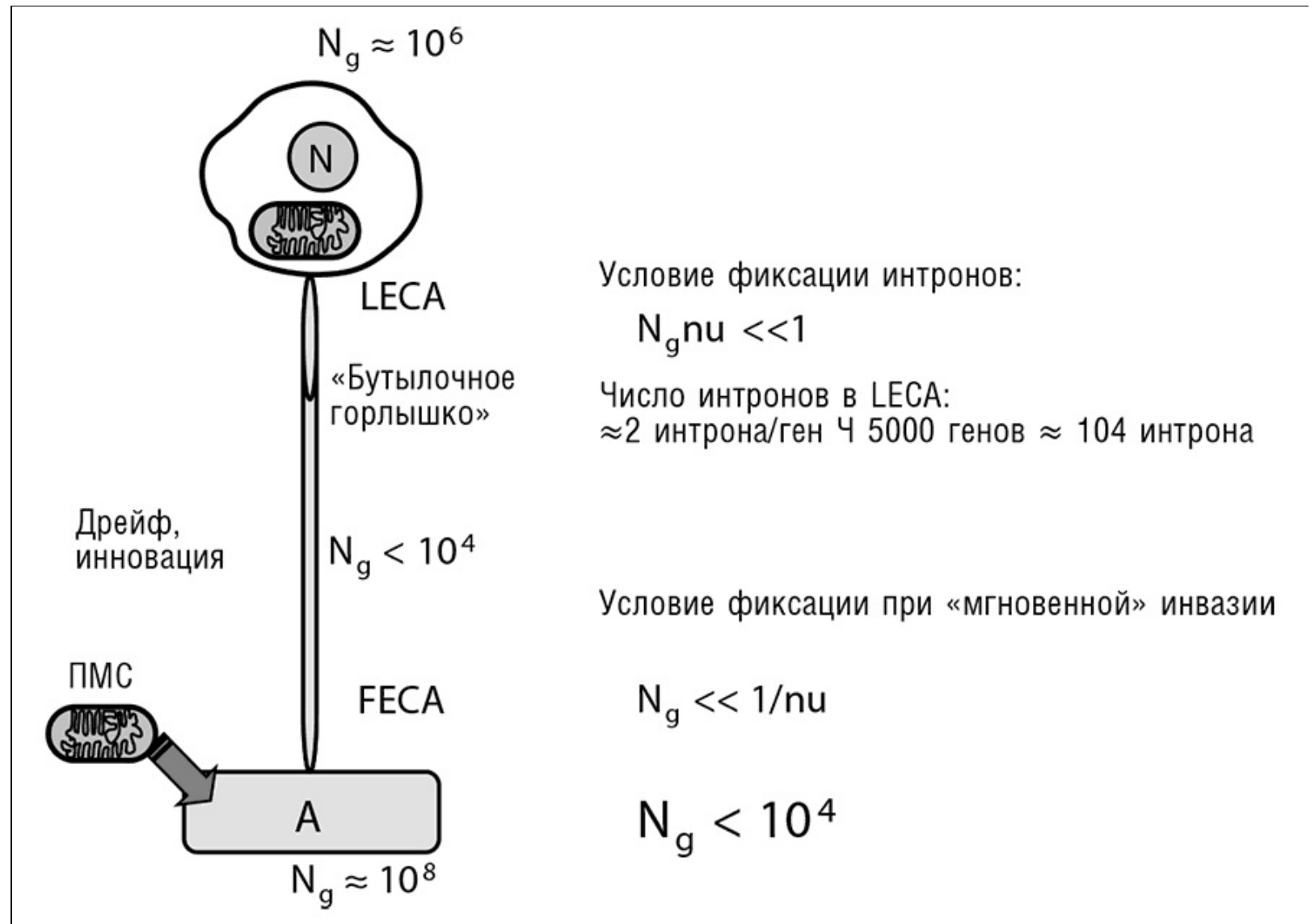


Рис. 8–3. Реконструкция популяционной динамики на протяжении эукариогенеза: эукариогенез делается возможным благодаря крайне узкому «бутылочному горлышку». N_g – эффективное число генов/локусов, n – число нуклеотидов, необходимых для сплайсинга (вначале автокаталитического) интрона (около 25 на интрон), размер мишени для вредных мутаций, u – частота мутаций на нуклеотид на поколение ($\approx 0,5 \times 10^{-9}$); A = архея, предполагаемый хозяин протомитохондриального эндосимбионта (ПМС); N = ядро; FECA = первый общий предок эукариот, химерная клетка, образовавшаяся немедленно после эндосимбиоза.

Перепишем условие фиксации из предыдущего раздела следующим образом:

$$n \ll \frac{1}{N_{eu}}$$

Это простое неравенство задает ограничение на размер мишени вредных мутаций, остающейся невидимой для очищающего отбора, или, другими словами, максимальное число необходимых для функционирования нового геномного элемента нуклеотидов, при котором он имеет шанс зафиксироваться.

Оценки, использующие значения N_e из табл. 8–1, выявляют принципиальные различия между организмами. Так, у позвоночных очищающий отбор «пропускает» до 250 нуклеотидов, в то время как у прокариот фиксация последовательностей длиннее, чем приблизительно 10 нуклеотидов, является маловероятной.

Эти теоретические соображения означают, что существенное увеличение геномной сложности возможно только в режиме ослабленного очищающего отбора. Рассмотрим три основные составляющие геномной сложности у позвоночных, также отвечающие за сложность молекулярного феномена и, насколько мы знаем, дифференциацию тканей и другие аспекты организменной сложности:

1. Альтернативный сплайсинг, который производит большую часть белкового разнообразия в данных организмах [75].

2. Комбинаторная регуляция транскрипции, при которой гены оснащены наборами сайтов связывания транскрипционных факторов. Различные комбинации факторов транскрипции связываются с этими сайтами, обеспечивая сложную регуляцию экспрессии (Venters and Pugh, 2009).

3. Гигантский некодирующий РНом, включающий в себя относительно хорошо изученные микроРНК, ряд других частично охарактеризованных малых РНК, более таинственные длинные некодирующие РНК [76] и огромное количество «темной материи» РНК (Amaral et al., 2008).

Рассмотрев эти замечательные явления более подробно, в каждом из них мы можем безошибочно различить следы неадаптивной эволюции, связанной с ослабленным очищающим отбором.

Как отмечалось в предыдущем разделе, богатые интронами геномы имеют «слабые» сигналы сплайсинга, скорее всего просто потому, что сила очищающего отбора в соответствующих популяциях недостаточна, чтобы жестко контролировать эти нуклеотидные последовательности. Иными словами, аномальные транскрипты, образующиеся с определенной частотой из-за ошибок сплайсинга в богатых интронами организмах, не являются достаточно вредными для того, чтобы быть устраненными очищающим отбором в условиях низкой N_e . Таким образом, неточность в вырезании интронов предоставляет нишу для альтернативного сплайсинга. Иными словами, *неточный сплайсинг – это и есть альтернативный сплайсинг*. Поскольку эволюционирующие небольшие популяции не могли избавиться от него, они «научились» использовать некоторые из альтернативных (первоначально аномальных) транскриптов в различных функциональных ролях. Эти роли часто основаны на том, что альтернативные белки являются модификациями «нормальных» белков и, соответственно, действуют как функциональные варианты исходного белка или же как доминантные отрицательные регуляторы. В соответствии с логикой эволюции, *альтернативный сплайсинг аналогичен горизонтальному переносу генов у прокариот в том, что оба являются выгодными альтернативами дубликации генов, при которых модификация активности достигается за один*

шаг, а не за длительный период эволюции. С учетом реконструкции, приведенной на рис. 7–8, можно предположить, что у LECA ошибки сплайсинга происходили с высокой частотой, давая, соответственно, большое разнообразие транскриптов, но при этом функциональный альтернативный сплайсинг был весьма редок (если вообще происходил). Дальнейшая эволюция различных ветвей эукариот, по-видимому, происходила в соответствии с двумя противоположными сценариями: потеря большинства интронов и усиление сигналов сплайсинга на границах оставшихся интронов, снижающие продукцию аномальных транскриптов до незначительного уровня; сохранение частоты ошибок сплайсинга примерно на том же уровне, что и у LECA (при условии примерно такой же плотности интронов), сопровождаемое эволюцией функционального альтернативного сплайсинга, то есть задействование многих, но, конечно, не всех и, вероятно, даже не большинства аномальных транскриптов для продукции альтернативных функциональных форм белка.

Большинство линий одноклеточных эукариот, эволюционировавших в сторону больших N_e и эффективного очищающего отбора, пошли по первому пути; второй сценарий относится к животным и растениям, которые никогда не достигали больших эффективных размеров популяции и вынуждены были справляться с унаследованным неточным сплайсингом. Третьего пути, по-видимому, не существовало: либо разработать способ устранения аномальных транскриптов, либо использовать их, либо вымереть.

Сайты связывания факторов транскрипции у эукариот состоят из 8—10 нуклеотидов, так что стоимость добавления одного сайта составляет $s \approx 10u$, или примерно 10^{-7} , если взять характерное для позвоночных значение u (Lynch, 2007c). Таким образом, геномы сложных многоклеточных эукариот, по-видимому, могли практически «бесплатно» накапливать сайты связывания транскрипционных факторов, что позволило появиться сложным каскадам сайтов. У одноклеточных эукариот возможности для эволюции в этом направлении были ограничены; для прокариот этот путь к инновациям, судя по всему, был закрыт очищающим отбором.

Некодирующий РНом позвоночных – возможно, главнейшее проявление сложности генома. Белок-кодирующие экзоны составляют около 1,5 процента генома млекопитающих, в то время как экзоны, соответствующие некодирующим РНК, по различным оценкам, занимают более 4 процентов генома – около 80 процентов кодирующего потенциала генома используется для молекул РНК, не транслирующихся в белки (Eddy, 2002). Это коренным образом отличается от кодирующих репертуаров прокариот и даже одноклеточных эукариот, в которых некодирующие РНК составляют лишь небольшую часть. Что еще более поразительно, ряд недавних исследований показывает, что большая часть – вероятно, более 60 процентов – генома млекопитающих транскрибируется на заметном уровне (Lindberg and Lundberg, 2010; Mendes Soares and Valcarcel, 2006). Природа этой «темной материи» далеко не ясна. Иногда считается, что экспрессия подразумевает функциональный смысл транскрибируемой области генома. Однако, учитывая отсутствие какой-либо заметной эволюционной консервации большинства из этих транскрибируемых последовательностей и относительной легкости возникновения ложных (слабых) сайтов инициации транскрипции в случайных последовательностях ДНК, можно сказать, что почти наверняка большая часть темной материи – это транскрипционный шум. Тем не менее эта случайно транскрибируемая часть генома и «мусорная» ДНК в целом представляют собой огромный резервуар для генерации новых микроРНК и других некодирующих, но выполняющих структурные и регуляторные функции РНК, многие из которых плохо сохраняются в процессе эволюции и эволюционируют высокими темпами. Открытие обширного РНОма животных показывает, что сложные геномы многоклеточных организмов и простые геномы одноклеточных форм жизни качественно различаются. Это различие интерпретируется самым естественным образом в рамках неадаптивной

популяционно-генетической теории эволюции генома. Согласно этой теории, эволюция форм жизни с низким N_e и последующим слабым очищающим отбором приводит к накоплению большого количества интронной и межгенной мусорной ДНК, некоторые сегменты которой время от времени задействуются для различных функций. Масштаб преобразования ландшафта экспрессии генома, вызванного, видимо, в первую очередь простыми факторами популяционной генетики, поражает воображение и представляется соразмерным с интуитивно очевидной разницей в сложности (и, очевиднее всего, в размере) между млекопитающим и простейшим. Вспомним обсуждение эволюции последовательностей в главах 3 и 4: широкий набор нефункциональных транскриптов составляет почти нейтральное пространство, открытое для эволюции сложности в многоклеточных организмах. Такое почти нейтральное пространство неизбежно возникает в ходе эволюции организмов с низкой N_e по чисто энтропийным причинам.

Хотя масштаб задействования мусора довольно мал по сравнению с общим количеством некодирующей ДНК, он огромен по отношению к суммарному размеру белок-кодирующих последовательностей. Учитывая популяционное «бутылочное горлышко», через которое, скорее всего, проходил эукариогенез (см. рис. 8–3), вполне вероятно, что значительное количество мусорной ДНК эволюционировало на очень раннем этапе истории эукариот и, возможно, уже присутствовало у LECA – как и интенсивная случайная транскрипция. Можно представить себе, что на следующем этапе эволюции произошло «нарушение симметрии», которое привело к бифуркации, описанной при обсуждении истории интронов: линии с большим N_e установили строгий контроль за геномом, устранив большинство мусорной ДНК. В противоположность им, линии, не достигшие больших N_e , занялись «компенсацией» в виде постепенного приспособливания возрастающего количества частей (бывшего) мусора под функциональную РНК (см. рис. 8–4).

Продолжая в том же духе, неадаптивная теория предлагает простое объяснение для перехода от простого типа регуляции транскрипции по Жакобу – Моно к сложной стратегии регуляции, используемой эукариотами. Вместо того чтобы использовать лишь один сайт связывания для единственного регулятора оперона (или, в редких случаях, несколько сайтов), как у прокариот, транскрипция большинства эукариотических генов регулируется в так называемом комбинаторном режиме, при котором несколько факторов транскрипции взаимодействуют сразу с несколькими, а зачастую и с большим числом сайтов, расположенных перед геном (Ravasi et al., 2010). У прокариот сайты связывания фактора транскрипции содержат достаточно информации для точного распознавания уникального сайта (или нескольких сайтов) в относительно небольшой геномной последовательности. Напротив, у эукариот сайт обычно содержит слишком мало информации для обеспечения точного распознавания (другими словами, геном содержит много сайтов с равным или даже большим сродством к данному транскрипционному фактору; Wunderlich and Mimi, 2009). Эта неадекватность одиночных сайтов связывания у эукариот обусловлена слабостью очищающего отбора, неспособного поддерживать множество точно сохраненных сайтов в геноме (см. обсуждение эволюции интронов ранее в этой главе), а также не может предохранить геном от роста, что увеличивает пространство поиска для транскрипционных факторов. Таким образом, комбинаторная модель может быть единственным решением для проблемы эффективной регуляции. Эволюции такого режима регуляции способствует рост генома, в частности достаточно высокая частота коротких tandemных дупликаций. Эволюция сложной регуляции экспрессии генов, являющейся отличительной чертой эукариот и необходимым условием для эволюции сложных многоклеточных форм, по-видимому, является наиболее ярким примером

превращения мусора в функциональные элементы в ходе эволюции при слабом очищающем отборе. Как и в случае других аспектов эволюции сложности, отбор направлен здесь на предотвращение энтропийного коллапса, а не на непосредственное «улучшение» регуляции.

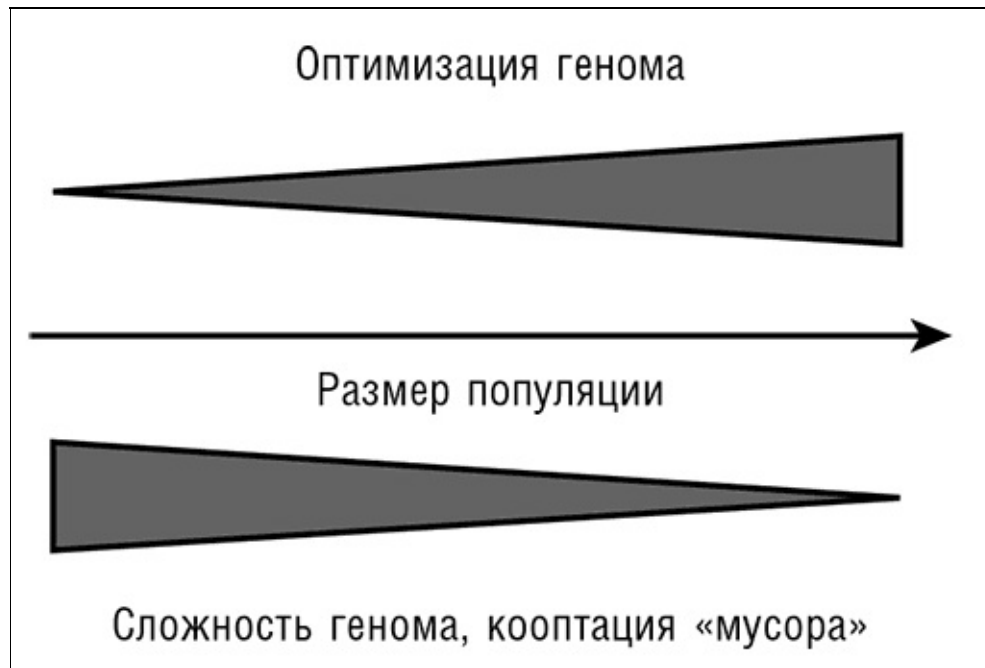


Рис. 8–4. Пути эволюции генома: оптимизация и задействование «мусора».

Эволюция продвинутых адаптаций в малых популяциях со слабым отбором может показаться парадоксальной, и, возможно, не зря: возникновение таких сложных функций, похоже, требует эффективного положительного отбора, что возможно только в популяциях с большим N_e . Это, безусловно, трудный вопрос. Ответ на него, по-видимому, требует противоречащего интуиции мышления в духе «слабого антропного принципа» (см. гл. 12 и прил. II): виды, в которых сложные функции не были зафиксированы, прежде всего через случайный дрейф и конструктивную нейтральную эволюцию (см. обсуждение ниже в этой главе), просто не имели шансов выжить.

Оптимизация генома в качестве основного пути эволюции и сложность как геномный синдром

Мы интуитивно склонны считать, что эволюция происходит от простых форм к сложным. Как писал Дарвин в заключительной 14-й главе «Происхождения...», «...из такого простого начала развилось и продолжает развиваться бесконечное число самых прекрасных и самых изумительных форм» (Darwin, 1859). Конечно, эта интуиция имеет смысл (и создает тяжелую проблему), когда речь заходит о происхождении первых форм жизни (мы обратимся к этой теме в гл. 12). Тем не менее была ли постепенно увеличивающаяся сложность преобладающей тенденцией в истории большинства линий на протяжении всей эволюции жизни? И теория популяционной генетики, и сравнительные геномные реконструкции говорят об обратном [\[77\]](#). В качестве наглядной иллюстрации обратимся еще раз к рис. 7–8. Появление двух ветвей многоклеточных эукариот, по-видимому, сопровождалось умеренным увеличением плотности интронов, что указывает на популяционное бутылочное горлышко, связанное с увеличением общей энтропии генома (величина *N*из первой части этой главы), во многих случаях весьма значительным. Увеличение энтропии создает нейтральное пространство, необходимое для последующего увеличения общей биологической сложности (высокое значение *C*). Напомним, что в этих случаях плотность биологической информации падает (низкое значение *D*): эти линии эволюционируют в «энтропийном режиме». Тем не менее даже среди растений и животных имеются большие группы, к примеру насекомые, *эволюция которых включала оптимизацию генома, или уменьшение эволюционной энтропии*. Этот процесс характеризуется менее стремительным падением в общей сложности и увеличением плотности биологической информации. Обращаясь к большинству ветвей в эукариотном дереве (см. рис. 7–2 и 7–8), включающих одноклеточные формы, мы видим однозначную картину оптимизации генома: энтропия генома резко падает и общая сложность также, хоть и менее резко, уменьшается, в то время как плотность информации быстро возрастает.

Пока еще слишком рано говорить о том, насколько тенденция к оптимизации генома, полученная из реконструкции на рис. 7–8, главенствует в общем контексте эволюции жизни, потому что таксономическая плотность секвенированных геномов из различных ветвей жизни по-прежнему недостаточна. Тем не менее результаты имеющихся ограниченных реконструкций позволяют предположить, что изложенная картина может быть достаточно полной. Например, реконструкция общего предка существующих архей указывает на то, что геном предковой формы был, по крайней мере, столь же сложен (в пересчете на общую сложность *C*, потому как трудно реконструировать энтропию и, следовательно, плотность информации непосредственно), как у типичных современных членов группы (Csuros and Miklos, 2009) [\[78\]](#). Кроме того, проявляется четкая тенденция в самих результатах реконструкций: предполагаемая сложность предковых форм пересматривается в сторону повышения с увеличением числа использованных для реконструкции геномов и с уточнением применяемых моделей наибольшего правдоподобия. Качественно аналогичные результаты были получены в ходе реконструкции геномного набора LESA (см. гл. 7): даже намеренно консервативные подходы, примененные к ограниченному набору геномов, указывают, что LESA был как минимум столь же сложен, как и типичный современный одноклеточный эукариот (Koonin, 2010a).

С учетом результатов этих реконструкций предкового генома и в рамках неадаптивной популяционно-генетической теории эволюции генома возникает искушение предложить общую модель эволюции энтропии и сложности генома. В этой модели эволюция обычно происходит прерывистым образом, через стадии высокой энтропии, связанные с популяционными

«бутылочными горлышками», впоследствии развиваясь в одном из двух различных режимов (см. рис. 8–5):

1. Низкоэнтропийное (высокая плотность биологической информации) состояние, связанное с высоким N_e , по сценарию оптимизации;
2. Высокоэнтропийное (низкая плотность биологической информации) состояние, связанное с низким N_e , в соответствии со сценарием кооптации [79].

Этот паттерн эволюции повторяется на протяжении всей истории жизни [80]. Высокоэнтропийные «бутылочные горлышки» соответствуют появлениям новых крупных групп, в то время как последующие расхождения линий внутри этих групп обычно включают в себя «нарушение симметрии» между этими двумя сценариями. Соответствие между этой моделью и моделью сжатого кладогенеза, изложенной в главе 6, очевидно. Важно отметить, что эпизоды внезапного возрастания энтропии немногочисленны и разнесены во времени друг от друга, тогда как большая часть истории жизни прошла в режиме «нормальной эволюции» между этими эпизодами. В фазе «нормальной эволюции» оптимизация генома, включающая уменьшение генома под действием сильного очищающего отбора в популяциях с большой N_e , по-видимому, встречается чаще, чем ограниченное усложнение, характерное для групп организмов, традиционно рассматриваемых как сложные, куда, безусловно, входит и наша собственная линия млекопитающих.

Режим оптимизации генома легко демонстрируется *in vitro* экспериментах по дарвиновской эволюции. Сол Спигелман и коллеги провели, пожалуй, самую известную серию таких экспериментов в 1960-х годах (Mills et al., 1973; Spiegelman, 1971). Они поместили небольшое количество РНК бактериофага в пробирку с репликазой (фермент фага, ответственный за репликацию генома), нуклеотидами и необходимыми ионами и позволили ему реплицироваться в течение непродолжительного времени. Часть содержимого затем перенесли в другую пробирку, содержащую ту же смесь, и повторили процедуру. В этих условиях давление отбора на РНК фага требует лишь ускорения репликации, и результаты эволюции в этом режиме были весьма радикальны: после примерно 70 повторений размер РНК снизился с ≈ 3500 до ≈ 400 нуклеотидов, то есть до наименьшего размера, при котором молекула способна размножаться при помощи полимеразы.

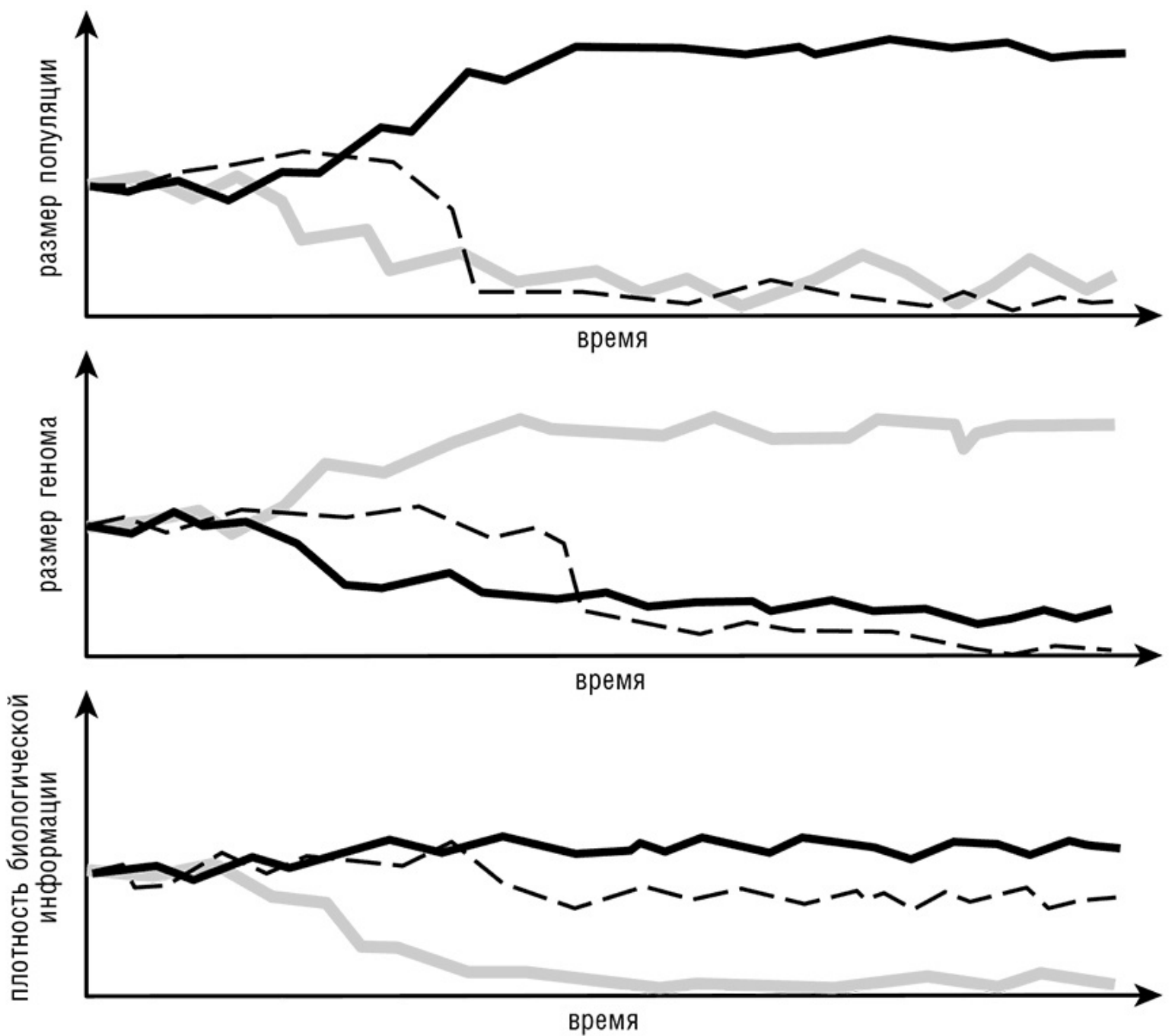


Рис. 8–5. Общая модель динамики эффективного размера популяции, размера генома и плотности биологической информации в соответствии с неадаптивной теорией. Каждый из графиков показывает три пути эволюции генома: сплошная линия – оптимизация генома (свободно живущие автотрофные бактерии и археи, некоторые одноклеточные эукариоты); серая линия – задействование мусорной ДНК и усложнение генома (эукариоты, особенно их многоклеточные формы); пунктирная линия – храповик деградации генома (паразиты и симбионты, особенно внутриклеточные формы).

За пределами нулевой гипотезы: ограничения популяционно-генетического взгляда на эволюцию генома

После прочтения предыдущих разделов этой главы нельзя не усомниться в обоснованности всеобъемлющего объяснения хода эволюции каким-либо одним общим фактором. Эти опасения полностью оправданы. Нужно еще раз подчеркнуть, что наиболее сильным утверждением популяционно-генетической теории эволюции генома является то, что неадаптивная эволюция, управляемая N_e , может быть *подходящей нулевой гипотезой*. Несмотря на свою важность, величина N_e определяет ход эволюции только в грубом приближении и на протяженных временных интервалах. Фактические эволюционные траектории определяются – и ограничиваются – конкретным биологическим контекстом. К примеру, в проведенном моими коллегами и мной широком анализе селективных ограничений в эволюции прокариот нам не удалось обнаружить отрицательной корреляции между силой очищающего отбора и размером генома, предсказываемой с прямолинейной популяционно-генетической точки зрения (Novichkov et al., 2009). Напротив, большие геномы, как правило, развиваются под более сильными ограничениями, чем малые, даже если рассматривать только свободноживущие микробы. Подразумевается, что образ жизни организма может быть критическим фактором эволюции генома, способствующим, в частности, приобретению генов через ГПГ в переменных условиях среды, более или менее независимо от N_e .

Геномика дает множество других указаний на ограниченную применимость популяционно-генетической теории эволюции генома и, в частности, концепции оптимизации генома. Ожидается, что оптимизированные геномы будут найдены в организмах, достигающих высокой численности (то есть высоких значений N_e) в более или менее постоянной среде и, соответственно, подвергающихся, согласно теории, сильному очищающему отбору. Эти геномы, по всей видимости, характеризуются не столько небольшим размером, учитывая непреодолимые ограничения, связанные с образом жизни (например, для автотрофных прокариот нижний порог для количества генов составляет примерно 1300), сколько крайней компактностью и практически полным отсутствием псевдогенов и встроенных в геном эгоистичных элементов. Все подобные элементы, по-видимому, быстро уничтожаются интенсивным очищающим отбором, настолько мощным, что даже короткие межгенные регионы сжимаются до наименьшей длины, необходимой для регуляторных функций. Наиболее распространенная среди известных организмов, морская фотосинтезирующая бактерия *Pelagibacter ubique*, похоже, идеально подходит под этот прогноз – у нее не обнаруживаются псевдогенов или мобильных элементов, очень мало паралогов и чрезвычайно короткие межгенные участки. Однако сравнительная геномика многочисленных штаммов *Prochlorococcus*, группы чрезвычайно распространенных морских фотоавтотрофных цианобактерий, показывает особенности, которые представляются несовместимыми с оптимизацией – а именно геномные острова, содержащие разнообразные гены, характерные для бактериофагов (Novichkov et al., 2009).

В целом взаимодействие между клеточными формами жизни и эгоистичными мобильными элементами существенно меняет структуру генома по сравнению с предсказаниями популяционно-генетической теории. Отношения между хозяевами и эгоистичными элементами (паразитами) часто описываются как «гонка вооружений» (подробнее об этом в гл. 10). Эти взаимодействия могут быть адекватно описаны только с учетом популяционной динамики как для хозяев, так и для паразитов. Конфликт «паразит – хозяин» приводит к равновесию, которое

не может быть выведено из популяционной динамики одного лишь хозяина, так что, по-видимому, даже наиболее оптимизированные геномы содержат существенное число эгоистичных элементов.

Оптимизация генома и уменьшение его размеров – не одно и то же. Бактериальные паразиты и внутриклеточные симбионты, а также единственно известный архейный паразит *Nanoarchaeum equitans* имеют наименьшие геномы среди прокариот, но эти геномы не оптимизированы. Вместо того эти организмы, по-видимому, претерпевают нейтральную деграцию генома. В самом деле, хотя некоторые из этих геномов чрезвычайно малы, так как паразиты и симбионты не нуждаются во многих генах, они часто содержат значительное число псевдогенов. Иногда в них также распространяются эгоистичные элементы. Хорошо изученные примеры этого типа включают *Rickettsia*, *Wolbachia*, патогенные микобактерии и некоторые лактобациллы (Frank et al., 2002; Lawrence et al., 2001). Паразиты и симбионты, как правило, не достигают больших значений N_e . Тем не менее они постепенно теряют гены, ставшие необязательными, в результате действия механизма типа храповика (крайне маловероятно, что некогда утерянный ген будет восстановлен, особенно учитывая образ жизни этих организмов), что подкрепляется предпочтением делеций в процессе мутации (Mira et al., 2001), а также снижением уровня ГПГ (см. гл. 5). Другое ключевое предсказание популяционно-генетической теории для этих организмов выполняется: как правило, они имеют высокие значения Kn/Ks , что свидетельствует о слабом давлении очищающего отбора. Это и ожидается, с учетом их небольших значений N_e . Поэтому представляется, что для некоторых образцов жизни разные предсказания теории могут выполняться или не выполняться независимо друг от друга.

Дарвиновский глаз, нередуцируемая сложность, экзаптация и конструктивная нейтральная эволюция

В предыдущих разделах мы обсудили различные аспекты сложности геномов и факторы, ведущие к ее возникновению. Фенотипическая сложность, с другой стороны, была рассмотрена лишь как следствие геномных процессов. Традиционно же биологов – как в наши дни, так и во времена Дарвина и раньше – занимала именно фенотипическая сложность, то есть сложность структур и функций на уровне организмов. Подробное обсуждение этой проблемы выходит за рамки данной книги, но мы сделаем некоторые замечания по поводу основных концепций, разработанных для объяснения фенотипической сложности.

Дарвин воспринимал эволюцию сложных органов как серьезнейшую проблему, но в то же время считал, что эта проблема разрешима в рамках его теории. Как уже упоминалось в главе 2, суть трудности состоит в очевидной несводимости сложного целого к более простым частям: «Чем может быть полезна половина глаза?» Иными словами, как мог сложный орган, который состоит из нескольких частей, развиться под действием естественного отбора, если отдельные части не имеют никаких известных функций? ^[81] Перед лицом этой трудности Дарвин по-прежнему оставался твердо убежден в силе естественного отбора, что запечатлено в знаменитом отрывке об эволюции глаза:

«Разум мне говорит: если можно показать существование многочисленных градаций от простого и несовершенного глаза к глазу сложному и совершенному, причем каждая ступень полезна для ее обладателя, а это не подлежит сомнению; если, далее, глаз когда-либо варьировал и вариации наследовались, а это также несомненно; если, наконец, подобные вариации могли оказаться полезными животному при переменах в условиях его жизни – в таком случае затруднение, возникающее при мысли об образовании сложного и совершенного глаза путем естественного отбора, хотя и непреодолимое для нашего воображения, не может быть признано опровергающим всю теорию» (Darwin, 1859, пер. К. А. Тимирязева, С. Л. Соболя, цит. по изд.: *Дарвин Ч. Сочинения*. Т. 3. М.: Изд-во АН СССР, 1939).

Повествование Дарвина предлагает одно из возможных концептуальных решений проблемы эволюции организационной сложности. Идею Дарвина можно охарактеризовать как «гипотезу неочевидных промежуточных этапов»: хотя и нельзя сразу представить вероятные промежуточные этапы эволюции по структуре и функциям развитой сложной структуры, такие промежуточные этапы в действительности существовали. Как правило, по крайней мере некоторые из функций этих промежуточных структур могут быть выведены через сравнительное исследование (сравнительная анатомия во времена Дарвина, сравнительная цитология и биохимия в XX веке и, вдобавок, сравнительная геномика в наши дни). Эта идея, безусловно, актуальна и плодотворна и, по-видимому, применима, в частности, для глаз и для других сложных органов животных. Однако дарвиновское объяснение представляется менее плодотворным в случае сложных молекулярных структур, как мы видели в главе 7 на примере супрамолекулярных структур эукариотической клетки.

Вторым основным путем к сложной организации является экзаптация, простая, но мощная концепция, предложенная Стивеном Гулдом и Ричардом Левонтином (см. гл. 2): молекулы или комплексы, которые эволюционировали под действием отбора на определенную функцию, нередко приспособляются (экзаптируются) для других, хоть и часто механистически сходных, функций (Gould, 1997a). Мы столкнулись с многими бесспорными случаями экзаптации при обсуждении фундаментальных инноваций, возникших в ходе эукариогенеза (см. гл. 7), например комплекса ядерных пор. Экзаптация часто дополняется случайной рекомбинацией

уже существующих молекул или устройств, особенно в тех промежутках процесса эволюции, когда рекомбинация стимулируется, как это почти наверняка было при эукариогенезе, потоком генетического материала от симбионта к хозяину. В редких случаях случайные комбинации уже существующих устройств дают новые функции, которые могут решить актуальные проблемы и потому фиксируются отбором.

Третья ключевая идея, которая, возможно, дополняет неадаптивную популяционно-генетическую теорию эволюции генома и может указывать наиболее общий путь к организационной сложности, – это модель конструктивной нейтральной эволюции (КНЭ), предложенная Арлином Стольцфусом в 1999 году (Stoltzfus, 1999) [82]. Суть КНЭ заключена в появлении зависимости между случайно взаимодействующими молекулами, которая делает взаимодействие необходимым и, следовательно, приводит к эволюции организационной сложности. КНЭ является храповиком, как и многие другие эволюционные явления, рассматриваемые в этой книге: *появившаяся однажды зависимость становится фактически необратимой*. Прекрасным примером КНЭ представляется эволюция сплайсосомы у эукариот (см. гл. 7). По модели КНЭ, случайное расщепление некоторых интронов группы II, вторгшихся в геном хозяина на ранней стадии эукариогенеза, привело к возникновению предков snРНК (активный компонент сплайсосомы) и позволило деградировать самосплайсирующимся концевым структурам всех интронов. В одновременном или последующем процессе случайные взаимодействия РНК-связывающих белков, в частности архейного белка Sm, с интронной РНК позволили деградировать интрон-кодируемой обратной транскриптазе. Очевидно, что эти изменения, создающие зависимости между компонентами эволюционирующей сплайсосомы, по сути необратимы, что формирует храповик и фиксирует эволюционирующую сложную организацию. Цитируя недавнее обобщение этого понятия Майклом Греем и коллегами (Gray et al., 2010), можно сказать, что сложность, появляющаяся по КНЭ, видимо, является не столько нередуцируемой, сколько «непоправимой». Прямой параллелью к модели КНЭ является сценарий субфункционализации для эволюции дупликации генов, предложенный Линчем с коллегами (Lynch and Katju, 2004). Согласно этому сценарию, дупликации генов могут быть зафиксированы без прямой адаптации, поскольку после дупликации новые паралоги могут свободно накапливать дифференциальные мутации, которые инактивируют, в каждом из паралогов, некоторые из многих функций, выполняемых предковым геном. Как только это произойдет, оба паралога становятся незаменимыми – еще один храповой механизм конструктивной нейтральной эволюции. Наблюдения, показывающие практически симметричное ослабление очищающего отбора для паралогов сразу же после дупликации, совместимы с моделью субфункционализации (Kondrashov et al., 2002).

Краткий обзор и перспектива: неадаптивная эволюционная парадигма и переоценка концепции эволюционного успеха

Возникновение и эволюция сложности на уровне генотипа и фенотипа и отношение между ними составляют одну из главных проблем биологии, если не сказать главнейшую. Даже если оставить в стороне на время проблему фактического происхождения весьма существенной сложности, связанной с клеточным уровнем организации (см. гл. 11), нельзя не удивиться, почему эволюция жизни не остановилась на стадии простейших автотрофных прокариот, имеющих 1000–1500 генов. Почему же вместо этого эволюция продолжилась, произведя на свет сложных прокариот, обладающих более чем десятком тысяч генов, и, что еще более поразительно, эукариот, с их огромными, тщательно регулируемым геномами, многими типами тканей и даже их способностью к созданию математических теорий эволюции?

Традиционный взгляд на эти проблемы явно или неявно сосредоточивается на сложности как на блистательном проявлении адаптации и силы естественного отбора. Соответственно, более сложные организмы традиционно считаются более развитыми, более успешными и, в некотором смысле, более важными, чем простые существа. Однако Стивен Джей Гулд предложил совершенно иную, стохастическую точку зрения на эволюцию сложности, случайное блуждание, метафорически описываемое им как походка пьяницы, вышедшего из бара на улицу [\[83\]](#): даже если человек под воздействием большого количества алкоголя передвигается совершенно случайно, через какое-то время он в конечном счете окажется довольно далеко от двери бара, например в канаве с другой стороны дороги (Gould, 1997b). То же относится к эволюции сложности: по прошествии достаточного количества времени эволюция, запущенная «со столь простого начала», ожидаемо достигнет высокой сложности в результате чисто стохастических процессов [\[84\]](#). Эта точка зрения на сложность вполне разумна, но является слишком абстрактной для удовлетворительной теории.

Как только стало возможным сравнение геномов простых (прокариот) и сложных (животных и растений) форм жизни, исследователи поняли, что в этих геномах есть что-то странное, вряд ли совместимое с идеей постоянного увеличения геномной сложности параллельно с ростом сложности организмов. Действительно, хотя геномы многоклеточных эукариот могут быть более сложными, чем у прокариот и даже одноклеточных эукариот, но в то же время эти сложные геномы чудовищно неупорядочены и заполнены мобильными элементами и прочим мусором; они представляют собой состояния высокой энтропии, как подчеркивается оценками в этой главе. Концептуальный взгляд на этот парадокс сравнительной геномики привел к неадаптивной теории эволюции генома, разработанной в основном на основе стандартных формул популяционной генетики. Тем не менее, несмотря на этот простой аппарат, теория поставила существующие представления о природе эволюции генома с ног на голову. В соответствии с неадаптивной теорией эволюция сложности генома является не адаптацией как таковой, а скорее следствием первоначального увеличения энтропии, вызванного слабостью очищающего отбора и, напротив, увеличенной силой дрейфа, характерной для популяционных «бутылочных горлышек».

Как это ни парадоксально, увеличение энтропии генома, являющееся необходимым условием для последующего усложнения, может закономерно рассматриваться как «геномный синдром», как неспособность организмов с небольшим эффективным размером популяции справиться с распространением эгоистичных элементов и других процессов, ведущих к росту энтропии. Конечно, эволюция сложности сама представляет собой сложный процесс, и эволюция кооптированных последовательностей включает в себя множество очевидных

адаптаций. Однако первоначальный энтропийный толчок является дезадаптацией, которую популяция изначально не в силах преодолеть. Частично последующая функциональная адаптация изначально нейтральных последовательностей компенсирует бремя возросшей геномной энтропии – другими словами, она позволяет организмам пережить рост своих геномов.

Реконструкции истории геномов и клеток, рассматриваемые в рамках неадаптивной парадигмы эволюции генома, привели к весьма удивительным выводам. Оказалось, что большая – вероятно, практически вся – история жизни шла отнюдь не в ключе «прогрессивной» эволюции в сторону увеличения сложности [85]. Вместо этого многие эволюционирующие линии пошли по пути оптимизации генома, при которой геномная энтропия и общая биологическая сложность генома падали, зачастую значительно, в то время как биологическая плотность информации росла. Некоторые другие линии, такие как наша собственная, пошли по пути задействования мусорных элементов (для регуляторных и структурных ролей, как в случае с РНКомом животных), что привело к резкому увеличению общей сложности, но лишь незначительному уменьшению энтропии. Таким образом, в этих линиях рост плотности биологической информации был весьма скромным, по сравнению с высокоэнтропийными состояниями, связанными с «бутылочными горлышками» в переходные эпохи.

Модели конструктивно нейтральной эволюции и субфункционализации паралогов дополняют неадаптивную теорию геномной эволюции, представляя убедительные сценарии неадаптивной эволюции сложности на уровне молекулярных фенотипов. В более широкой перспективе эти теоретические разработки, совместимые с эмпирическими данными сравнительной геномики, завершают пересмотр эволюционной биологии, начатый нейтральной теорией молекулярной эволюции. Последняя показала, что большинство мутаций, фиксирующихся в процессе эволюции, эффективно нейтральны и поддерживают, таким образом, нейтральность как подходящую нулевую гипотезу для всех молекулярных эволюционных исследований. Новые разработки сделали то же для эволюции генома и молекулярного фенотипа. Очевидно, что от нулевой гипотезы не ожидается полного описания любого процесса, не говоря уже о таком сложном, многогранном процессе, как эволюция жизни. Как мы видели, во многих случаях некоторые из предсказаний неадаптивной теории не исполняются из-за дополнительных сильных ограничений, вытекающих из особенностей образа жизни организмов. И конечно, положительный отбор и вызываемые им адаптации являются важнейшими аспектами эволюции. Однако похоже, что эти факторы проявляются локально на глобальном фоне более фундаментальных процессов, таких как давление очищающего отбора, определяемого эффективным размером популяции и эволюционными храповиками, что может приводить к неадаптивному появлению сложности.

Заканчивая эту главу, скажем несколько слов о понятиях эволюционного успеха и «прогресса». Идея «прогресса» может считаться полностью дискредитированной в узком, антропоморфном смысле слова, но увеличение сложности по-прежнему широко воспринимается как особенность «развитых» форм жизни и основная эволюционная тенденция. Противоположная точка зрения, представленная, вероятно наиболее ярко, Гулдом в нескольких книгах, ассоциирует эволюционный успех группы исключительно с ее распространенностью в биосфере и способностью процветать в различных нишах. В рамках неадаптивной теории естественно соотносить «успех» с большим эффективным размером популяции. С этой точки зрения эволюция сложности не имеет ничего общего с успехом группы и вместо этого запускается неудачами (популяционными «бутылочными горлышками») на каком-то этапе и сохраняющейся неспособностью эволюционировать в сторону большей популяции, эффективно поддерживаемой отбором. По-настоящему успешные и эффективные формы жизни просты и

Оптимизированны.

Рекомендуемая дополнительная литература

Adami C.(2002) What Is Complexity? *Bioessays*24: 1,085—1,094.

Концептуальная статья, которая количественно определяет «физическую сложность» геномов (связанную с эволюционной сложностью, рассматриваемой в этой главе).

Gray, M. W., J. Lukes, J. M. Archibald, P. J. Keeling , and W. F. Doolittle. (2010) Cell Biology. Irremediable Complexity? *Science*330: 920–921.

Обновленное, нетехническое обсуждение гипотезы конструктивной нейтральной эволюции. Достоинно цитаты ради кристально ясного резюме: «Многие из макромолекулярных машин клетки кажутся избыточно сложными, имеющими больше компонент, чем требуют их функции. Как мы можем понять эту сложность в свете эволюции? Одной из возможностей является нейтральный храповой процесс, описанный более десяти лет назад и впоследствии названный конструктивной нейтральной эволюцией. Эта модель дает объяснение эволюции, способное противостоять селекционистским, адапционистским представлениям, понижающим молекулярную биологию».

Koonin E. V.(2004) A Non-adaptationist Perspective on Evolution of Genomic Complexity or the Continued Dethroning of Man. *Cell Cycle*3: 280–285.

Обсуждение смысла эволюционной сложности в свете неадаптивной теории Линча и общих следствий из новых идей об эволюции сложности.

Lynch M.(2007) The Frailty of Adaptive Hypotheses for the Origins of Organismal Complexity. *Proceedings of the National Academy of Sciences USA*. 104 Supplement 1: 8,597—8,604.

Краткое обсуждение новой, плюралистической точки зрения на эволюцию организмов, согласно которой естественный отбор является лишь одним из важных факторов эволюции. Цитата: «Истоки многих аспектов биологического разнообразия, от генно-структурных «украшений» до инноваций на фенотипическом уровне, имеют корни в неадаптивных процессах, при этом популяционно-генетическая среда жестко задает направления, открытые для исследования эволюционным процессом».

Lynch M.(2007) *The Origins of Genome Architecture*. Sunderland, MA: Sinauer Associates.

Основополагающая книга по неадаптивной теории эволюции геномной сложности и ее различным следствиям (подробности см. в гл. 8).

Lynch M.(2006) Streamlining and Simplification of Microbial Genome Architecture. *Annual Review of Microbiology*60: 327–349.

Применение неадаптивной теории эволюции сложности к эволюции геномов прокариот, с акцентом на роль оптимизации генома под действием сильного очищающего отбора в популяциях с большим N_e .

Lynch M., and J. S. Conery.(2003) The Origins of Genome Complexity. *Science*302: 1,401—1,404.

Ключевая статья, представляющая неадаптивную теорию эволюции геномной сложности и описывающая простые популяционно-генетические оценки, лежащие в основе теории.

Novichkov P.S., Y. I. Wolf, I. Dubchak , and E. V. Koonin. (2009) Trends in Prokaryotic Evolution Revealed by Comparison of Closely Related Bacterial and Archaeal Genomes. *Journal of Bacteriology*191: 65–73.

Сравнительное исследование режимов отбора в разных эволюционных линиях прокариот, демонстрирующее ограничения прямолинейного популяционно-генетического подхода.

Stoltzfus A.(1999) On the Possibility of Constructive Neutral Evolution. *Journal of Molecular Evolution*49: 169–181.

Важная концептуальная статья, сделавшая возможным появление понятий неадаптивной эволюции сложности и храповика конструктивной нейтральной эволюции.

Хотелось бы добавить к этому списку две книги, одна из которых без всяких уважительных причин осталась вне поля внимания автора при работе над этой главой, а вторая специфически рекомендуется читателю русского перевода.

John Maynard Smith and Eörs Szathmáry. (1998) The Major Transitions in Evolution. Oxford University Press, Oxford.

Классическая работа, в которой фундаментальные эволюционные инновации, такие как возникновение клетки и возникновение многоклеточных организмов, трактуются в контексте перехода селекции на новый уровень. Так, при возникновении многоклеточных организмов объектом селекции становится уже не индивидуальная клетка, а сообщество клеток с определенным распределением дополняющих друг друга функций (печальные последствия обратного перехода к клеточному уровню селекции очевидны: опухоль по сути представляет собой колонию одноклеточных организмов).

Марков А. В. Рождение сложности. Эволюционная биология сегодня. Неожиданные открытия и новые вопросы. М.: Астрель; Соргус, 2010.

Великолепный, популярный, но серьезный текст о современном состоянии эволюционной биологии, с упором на эволюцию биологической сложности.

Глава 9. Ламарковский, дарвиновский и райтовский режимы эволюции, эволюция эволюционируемости, надежность биологических систем и созидательная роль шума в эволюции

Как уже отмечалось в предисловии к данной книге, одной из ключевых заслуг Дарвина явилось то, что ему удалось выявить основополагающее сочетание случая и неизбежности, сопровождающее эволюцию жизни. Согласно Дарвину, большинство наследственных вариаций случайны, а направленность эволюции всецело задается естественным отбором, обуславливающим закрепление либо отсева случайных мутаций (Darwin, 1859). Как мы неоднократно подчеркивали, на стадии фиксации случайность также вносит свой существенный вклад посредством дрейфа и генетической тяги, которые критически зависят от популяционной динамики (см. гл. 8). Между тем Дарвин отводил значимую, хотя и второстепенную роль фундаментально иному типу вариации, так называемому ламарковскому наследованию.

К ламарковскому наследованию относят не случайно приобретенные фенотипические изменения, в особенности те, что непосредственно возникают в связи с использованием тех или иных органов и, соответственно, предположительно имеют адаптивный характер (полезны для организма). Вызывающий множество споров французский натуралист Жан-Батист Ламарк предполагал, что приобретенные изменения наследуются и представляют тем самым основу для эволюции. Ламарк был автором первой логически последовательной теории эволюции, которую представил в своей «Философии зоологии»: «наследование приобретенных (адаптивных) признаков» играло ключевую роль в этой теории (Lamarck, 1809).

Как неоднократно здесь подчеркивалось, в противоположность Ламарку Дарвин поставил во главу угла случайные, ненаправленные изменения, которые, согласно его теории, обеспечивают основной материал для естественного отбора. Правда, в более поздних изданиях «Происхождения видов...» Дарвин придавал гораздо больший вес ламарковскому механизму эволюции, по-видимому из-за опасений, что случайных вариаций и естественного отбора может оказаться недостаточно для обеспечения эволюционного процесса во всей его полноте.

«Наследование приобретенных (адаптивных) признаков» остается фундаментальной проблемой, значимость которой простирается далеко за рамки драматичной и увлекательной истории биологии XIX и XX веков. Взаимодействие между случайными и направленными изменениями генома (если последние вообще существуют) – главная тема данной книги. Здесь мы подходим к самой сути парадокса случая и необходимости, обращаясь к той части биологически важной неслучайности, которая, вероятно, присутствует уже на начальной стадии эволюционного процесса, в момент возникновения вариации. Говоря более конкретно, ключевой вопрос, касающийся ламарковского механизма наследования и эволюции, звучит следующим образом: могут ли факторы среды вызывать адаптивную эволюцию генома напрямую, без обращения к обходному пути естественного отбора?

По общему мнению, ламарковское наследование было всерьез скомпрометировано знаменитыми опытами Вейсмана, в которых он обрубал хвосты крысам [\[86\]](#). Дискредитация этого типа наследования была усугублена причудливым и трагическим эпизодом предположительно мошеннических экспериментов Пола Каммерера с окраской жаб-повитух, что в итоге привело ученого к самоубийству [\[87\]](#). В XX столетии «ламаркизм» приобрел исключительно дурную репутацию, когда в Советском Союзе стал составной частью лысенковщины [\[88\]](#). Однако в последнее время несколько направлений исследований, похоже, сходятся в своих выводах, указывающих на то, что механизмы, которые в разной степени отвечают критериям наследственности по Ламарку, могут вносить важный вклад в процесс эволюции (Koonin and Wolf, 2009b).

Классическая ламарковская схема подразумевает наследование конкретных адаптивных

характеристик фенотипа, которые индивид приобрел в течение жизни. В этом узком смысле ламарковский сценарий действительно выглядит несостоятельным по причине явного отсутствия механизмов прямого кодирования приобретенных признаков в геномный текст. Вместе с тем эта необратимость потока генетической информации, которую сформулировал Фрэнсис Крик и которая затем стала известна как центральная догма молекулярной биологии (Crick, 1970), строго применима лишь к передаче информации между нуклеиновыми кислотами и белками. Необратимость появляется на этапе узнавания аминокислот специфичными тРНК, что определяет сборку аминокислот в синтезируемые белки на основе последовательности триплетов в соответствующих мРНК. Не существует обратного пути в геном от любых изменений, могущих возникнуть в белковой последовательности [\[89\]](#). Однако ситуация с нуклеиновыми кислотами, в частности РНК, иная: в биогенезе РНК нет необратимой стадии, которая делала бы невозможной передачу информации обратно в геном. Данное различие важно иметь в виду, когда мы обращаемся к разным классам генетических изменений, некоторые из которых могут быть инициированы факторами окружающей среды.

В этой главе я обсуждаю возможные ламарковские и квазиламарковские эволюционные механизмы наряду с другими важными феноменами, такими как эволюция эволюционирования (evolvability), точность передачи биологической информации и роль шума в эволюции. Все это имеет отношение к одному принципиальному вопросу: существует ли определенная эволюционная логика в мутационных процессах, порождающих геномные вариации, или эти процессы определяются одной лишь случайностью?

Ламарковский, дарвиновский и райтовский режимы эволюции и критерии для обнаружения ламарковского наследования

Прежде чем обратиться к широкому кругу явлений, в которых можно усмотреть все или некоторые признаки эволюционного механизма, ассоциирующегося с именем Ламарка, необходимо сформулировать суть ламарковского подхода и критерии, которым эволюционный процесс должен удовлетворять, чтобы считаться ламарковским. Я не рассматриваю подробно различия между оригинальными взглядами Ламарка и многочисленными их последующими интерпретациями; напротив, я хочу сосредоточиться на вычлениении сущности того, что обычно известно как наследование приобретенных признаков и эволюция по Ламарку. Концепция наследственности Ламарка, будучи одним из двух краеугольных камней его эволюционной теории (Gould, 2002), зиждется на двух принципах, которым Ламарк в своей «Философии зоологии» и других работах придавал статус фундаментальных законов:

1. Использование или неиспользование органов.
2. Наследование приобретенных признаков.

Ламарк напрямую связывал «использование или неиспользование» с воздействием среды на «повадки» организма и, через эти повадки, на «форму и свойства» частей тела. При этом он, конечно, полагал, что такие адаптивные, вызванные воздействием среды, изменения наследуются. Ламарк писал: «Природа дает нам несчетные... примеры воздействия среды на поведение, а поведения на формы, организацию и пропорции частей тела животных». Таким образом, его идея наследственности была основана на трехкомпонентной причинно-следственной цепочке: среда – поведение – форма. Ламарк настаивал на важности изменений поведения в качестве промежуточного звена, связывающего окружающую среду с (наследуемым) изменением формы организма:

«Каково бы ни было влияние природных условий, они не вырабатывают непосредственных модификаций формы и организации животных. Однако серьезные перемены в окружающей обстановке влекут за собой изменение стимулов, а смена стимулов с неизбежностью приводит к перестройке поведения. И тогда, если новые потребности становятся постоянными, у животных вырабатываются новые повадки, которые сохраняются столько времени, пока действуют стимулы, породившие их» (Lamarck, 1809).

Ламарк ни в коем случае не был единственным, кто верил в наследование приобретенных признаков: по-видимому, таково было всеобщее убеждение того времени. Однако он был более конкретен, чем другие, в разъяснении цепи причин наследственности и, что более важно, сделал эту схему основой своей незаурядной концепции эволюции. Другим основанием эволюционной теории Ламарка была его убежденность в присущем эволюционирующим организмам стремлению к усложнению организации – или, попросту, прогрессу – которое, в глазах Ламарка, определяет процесс биологической эволюции вместе с наследственностью, как он ее понимал.

Хотя для обозначения этой фундаментальной тенденции Ламарк часто использовал словосочетание «жизненная сила» (*pouvoir de la vie*), его идея была абсолютно материалистической, можно даже сказать механистической, поскольку причину стремления к прогрессу он видел в движении жидкостей в теле животных. По его мысли, эти жидкости проделывали каналы и полости в мягких тканях, что постепенно вело к эволюции в сторону усложнения организации. Для объяснения присутствия просто устроенных форм жизни, несмотря на предположительно прогрессивный характер эволюции, Ламарк постулировал, что постоянным источником примитивных организмов служит самозарождение. Идеи самозарождения и врожденной склонности к усложнению в настоящее время безнадежно

устарели.

Однако по-прежнему широко распространено убеждение, будто эволюция ведет к усложнению организации в результате многочисленных последовательных адаптаций (разумеется, в наши дни ученые, продолжающие выступать за реальность такой тенденции, не назовут ее «врожденной»). Как уже обсуждалось в главе 8, оно явно не соответствует действительности: на самом деле никакой глобальной тенденции к усложнению живых форм в ходе эволюции не прослеживается, даже если максимум наблюдаемой сложности растет в силу стохастических причин. В данной главе мы обратимся к более значимой и интересной проблеме из наследия Ламарка, а именно рассмотрим наследование приобретенных признаков и его вклад в эволюционный процесс.

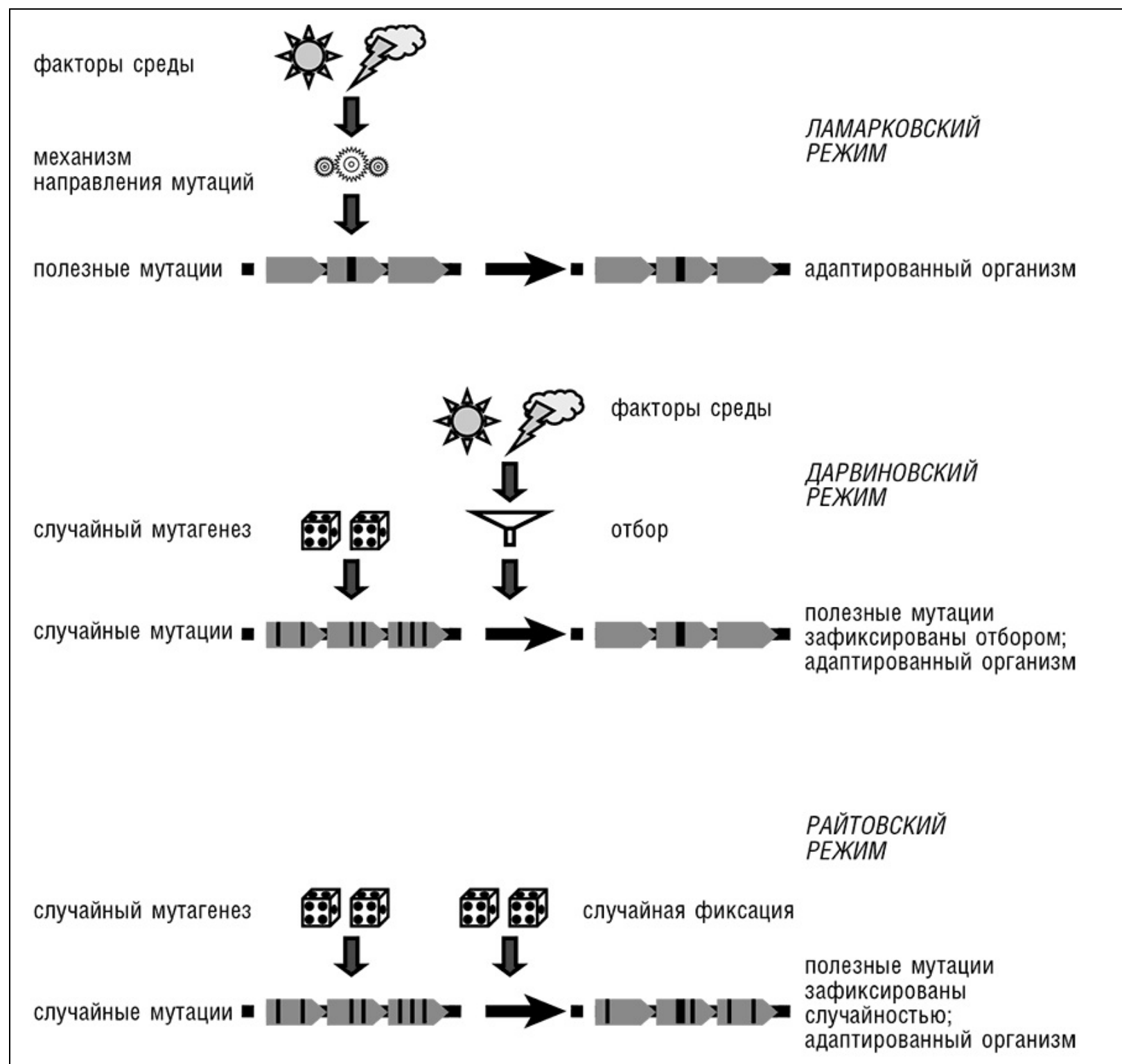


Рис. 9–1. Ламарковский, дарвиновский и райтовский режимы эволюции. Заимствовано из

Схема Ламарка, представленная в терминах, совместимых с современной генетикой (см. рис. 9–1), утверждает следующее:

1. Факторы окружающей среды вызывают (наследуемые) изменения генома.
2. Индуцированные изменения (мутации) затрагивают конкретный ген или группу генов.
3. Индуцированные изменения обеспечивают адаптацию к первоначальному причинному фактору.

Очевидно, что адаптивная реакция на конкретный фактор окружающей среды должна быть опосредована молекулярным механизмом, который бы вызывал изменение в соответствующем гене (или генах). В этом состоит явное отличие от хода дарвиновской эволюции: в ней окружающая среда не является фактором, вызывающим адаптивные изменения, скорее она служит источником внешнего селективного давления, которое способствует фиксации тех случайных изменений, что адаптивны при данных условиях (см. рис. 9–1). Дарвиновская схема проще и менее взыскательна по сравнению со схемой Ламарка, поскольку Дарвину не требовались специальные механизмы, направляющие мутации в соответствующие локусы генома и ограничивающие изменения конкретными мутациями, которые обеспечивают необходимую адаптацию.

С другой стороны, режим эволюции Ламарка был бы более эффективным и быстрым по сравнению с дарвиновским. В самом деле, за предпочтение дарвиновского механизма приходится платить втридорога: многочисленные мутации, что возникают в геномах, по-настоящему вредны, так что их носители погибают, другие почти нейтральны и иногда закрепляются с помощью дрейфа, однако при этом не вносят непосредственного вклада в адаптивную эволюцию. Поэтому ламарковская схема оказалась бы очень полезной для эволюционирующих организмов, будь она практически осуществимой. Трудность обнаружения или даже гипотетического представления механизмов направленного адаптивного изменения в геноме отравила сценарий эволюции Ламарка на долгие десятилетия в мусорную кучу истории науки.

Несмотря на существенные различия в механизмах и невзирая на кажущуюся «расточительность» дарвиновской модальности, которая контрастирует с потенциальной эффективностью эволюции по Ламарку (см., однако, дискуссию ниже в этой главе), схемы Дарвина и Ламарка похожи: у обеих конечные результаты преимущественно адаптивны, и в этом отношении, они радикально отличаются от случайного дрейфа. Последний процесс может быть обозначен как «райтовская модальность эволюции», в честь Сьюэлла Райта, одного из отцов-основателей популяционной генетики и создателя концепции случайного генетического дрейфа (см. рис. 9–1 и гл. 2). В последующих разделах я расскажу о недавних исследованиях некоторых феноменов, которые, на мой взгляд, заставляют нас возвратиться к одной из версий ламарковского сценария в качестве важного вклада в эволюцию на уровне геномов и организмов.

Системы антивирусного иммунитета CRISPR-Cas у прокариот: демонстрация аутентичного механизма по Ламарку

Система антивирусной защиты и адаптивного иммунитета у архей и бактерий, которая в последнее время была изучена благодаря целой серии открытий, часто случайных, по всей видимости, работает непосредственно через предложенный Ламарком механизм. Такая система известна как CRISPR-Cas (или просто CRISPR, для краткости). Аббревиатура CRISPR означает «короткие палиндромные повторы, регулярно расположенные группами», а Cas – «белки, ассоциирующиеся с CRISPR» (Deveau et al., 2010; Karginov and Hannon, 2010; Koonin and Makarova, 2009; van der Oost et al., 2009). CRISPR-повторы содержат короткие уникальные участки-спейсеры, встроенные внутри палиндромных повторяющихся блоков. Геномы архей и бактерий содержат кассеты (группы tandemно организованных, тесно сцепленных, функционально связанных локусов) с многочисленными CRISPR-блоками – во многих случаях более одной кассеты на геном. Хотя CRISPR-повторы были открыты еще в 1980-х, за годы до первых расшифровок полных бактериальных геномов, только гораздо позже стало понятно, что CRISPR-кассеты в геномах практически всегда примыкают к группе *cas*-генов. *Cas*-гены кодируют различные ферменты, участвующие в метаболизме нуклеиновых кислот, включая нуклеазы, геликазы и, возможно, полимеразы [\[90\]](#).

После того как было показано, что некоторые из уникальных спейсеров в CRISPR-кассетах идентичны фрагментам генов бактериофагов и плазмид, была высказана гипотеза о том, что CRISPR-система использует полученные от фагов последовательности в качестве молекул-шаблонов для разрушения мРНК-фагов аналогично эукариотической РНК-интерференции (РНКи) (Makarova et al., 2006). Хотя большую часть деталей механизма еще предстоит выяснить, главные предсказания этой гипотезы к настоящему времени подтверждены: наличие спейсера, последовательность которого в точности комплементарна мишени, то есть соответствующей последовательности в фаговом геноме, необходимо для резистентности [\[91\]](#); РНК-шаблоны, содержащие CRISPR-спейсеры, образуют комплексы с несколькими Cas-белками и используются для борьбы с инфекцией; могут приобретаться новые спейсеры, которые делают бактерию или архею устойчивой к соответствующим фагам. Примечательно, что, судя по всему, некоторые CRISPR-системы нацелены на вирусные мРНК, как и постулирует изначальная гипотеза, тогда как другие уничтожают вирусную ДНК непосредственно (Barrangou et al., 2007; Brouns et al., 2008; Hale et al., 2009; Marraffini and Sontheimer, 2008).

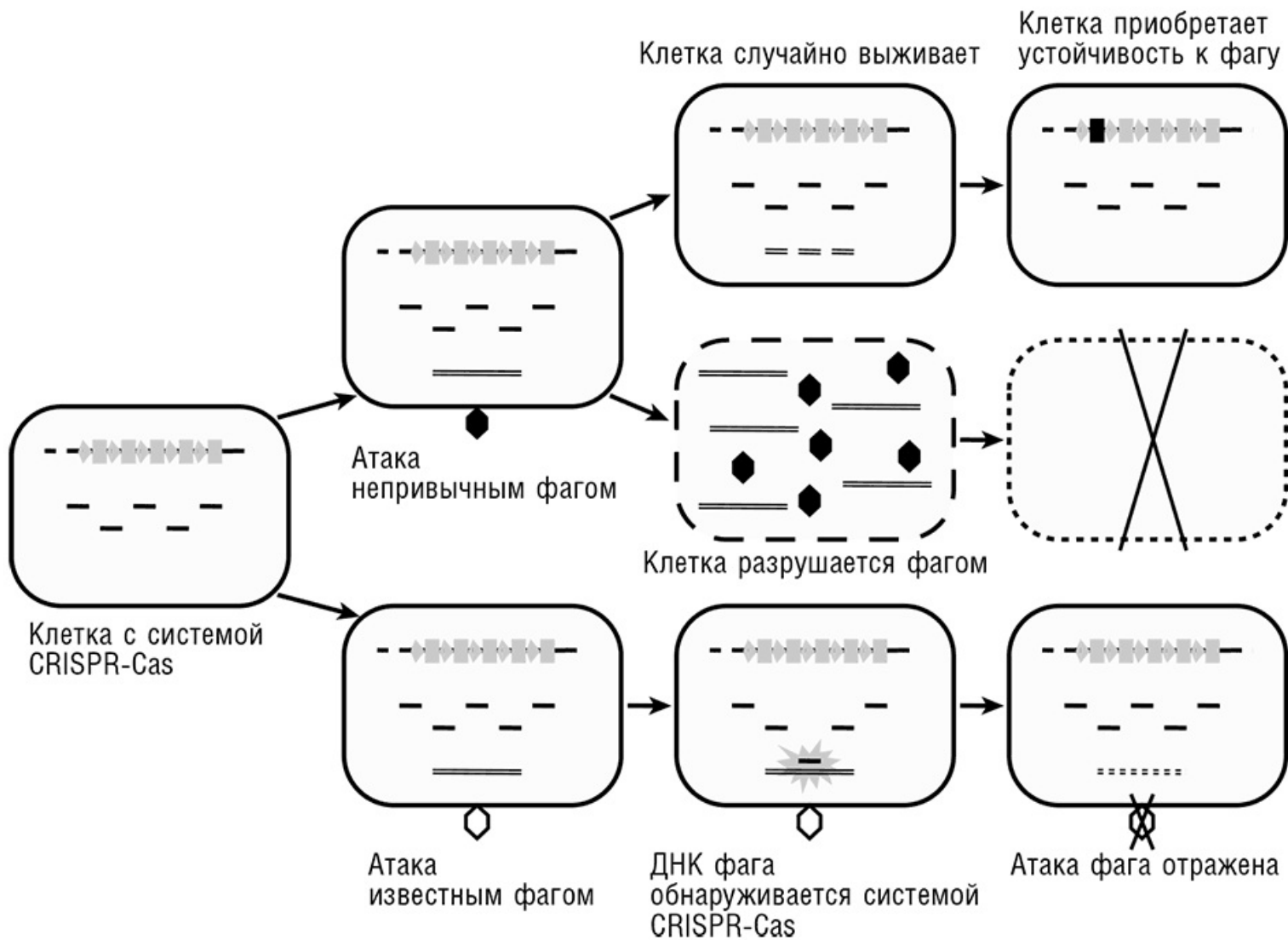


Рис. 9–2. Система CRISPR-Cas и механизм ее действия как пример ламарковской эволюции. Заимствовано из Koonin and Wolf, 2009b.

Механизм наследственности и эволюции генома, реализованный в CRISPR-Cas системе, представляется в полной мере ламарковским (см. рис. 9–2).

- Стимул из внешней среды (эгоистичный генетический элемент, такой как вирус) используется для непосредственного изменения генома.
- Возникающая модификация (уникальный элемент-специфичный спейсер) напрямую влияет на фактор, вызвавший изменение.
- Модификация явно адаптивна и наследуется потомками клетки, столкнувшейся с эгоистичным элементом.

CRISPR-опосредованная наследственность, по-видимому, не очень устойчива: даже близкородственные геномы архей и бактерий не содержат одинаковых вставок. Подразумевается, что, как только бактерии или археи перестают сталкиваться с конкретным агентом (вирусом), соответствующие спейсеры быстро вырождаются. В самом деле, вставки вряд ли могут быть эволюционно стабильными в отсутствие сильного селективного давления, поскольку единственная мутация сразу делает их бесполезными. Кроме того, во многом подобно адаптивной иммунной системе у животных, CRISPR-системы в редких случаях демонстрируют свойство аутоиммунности: спейсеры, идентичные фрагментам обычных генов клетки-хозяина, вставляются в CRISPR-кассеты и, предположительно, нарушают экспрессию соответствующих генов (Stern et al., 2010). Несмотря на некоторую эфемерность

наследственности CRISPR, ее ламарковский характер неоспорим: адаптивная эволюция организмов происходит непосредственно в ответ на внешний фактор среды и результатом является конкретная адаптация (резистентность) к данному специфическому фактору [\[92\]](#).

Другие (квази)ламарковские системы, функционирующие по принципу CRISPR

Интересно и поучительно сравнить особенности наследственности и эволюции в случае CRISPR-системы с соответствующими характеристиками эукариотической РНК-интерференции (RNAi) и, в частности, малых интерферирующих (si) РНК и PIWI-взаимодействующих (pi) РНК, то есть с защитными системами эукариот, в общих чертах функционально аналогичными CRISPR. Для начала вспомним примечательный и довольно загадочный факт: белковый аппарат эукариотической РНК-интерференции не гомологичен Cas-белкам; белковые компоненты этой сложной эукариотической системы были собраны из прокариотических доменов, которые первоначально были вовлечены в исполнение других функций (см. гл. 7; Shabalina and Koopin, 2008). Очевидное отсутствие ортологов для любого из Cas-белков в клетках эукариот позволяет предположить, что эта система каким-то образом исключена из эукариотического мира отбором, хотя лежащее в основе селективное давление представляется смутно. Единственным намеком может служить общая причина утраты оперонов у эукариот, которую мы обсуждали в главе 7: опероны исчезают под действием рекомбинационного хrapовика, и гены, которые требуют особенно тесной координации экспрессии или же вредны вне контекста действия оперона, устраняются путем очищающего отбора [\[93\]](#).

В отличие от CRISPR-Cas, системы РНК-интерференции не используют механизм Ламарка напрямую. Тем не менее они явно демонстрируют характерные «ламарковские» черты. Система siРНК (отдельный вид РНКи) «обучается» внешним агентом (вирусом) путем генерации малых интерферирующих РНК, комплементарных вирусным генам (Kim et al., 2009). Этот процесс, безусловно, имеет сходство с CRISPR-механизмом, но, кроме того, напоминает, по крайней мере метафорически, «изменение повадок» по Ламарку. Более того, система имеет некоторый уровень памяти, поскольку во многих организмах miРНК амплифицируются, и устойчивость к соответствующему вирусу может сохраняться в течение нескольких поколений (Ding, 2010). Подобная стабильность miРНК служит одним из проявлений получающего все более широкое признание РНК-опосредованного наследования, которое иногда называют парамутацией (Hollick, 2010). Ключевое отличие от CRISPR состоит в том, что (насколько известно в настоящее время) miРНК не записываются в геном, так что здесь имеет место лишь *эпигенетическая наследственность ламарковского типа*, но не полноценная генетическая наследственность.

Однако даже это различие размывается в случае piРНК, которые являются производными транспозонов. Это наиболее распространенные малые РНК в животном мире, образующие быстро растущие геномные кластеры, обеспечивающие защиту от мобильных элементов в зародышевой плазме (Bourc'his and Voinnet, 2010). В случае этих малых РНК, как и в ситуации с CRISPR, фрагменты генома мобильного элемента интегрируются в геном хозяина, где они быстро размножаются, видимо, под давлением отбора на эффективную защиту (Assis and Kondrashov, 2009). Такая система, похоже, отвечает всем критериям наследования приобретенных признаков и ламарковского режима эволюции. Здесь особенно примечательно, что изолированная зародышевая плазма, будучи важнейшим изобретением многоклеточных эукариот, которые, по-видимому, блокируют некоторые формы (квази)ламарковского наследования, такие как горизонтальный перенос генов (см. обсуждение далее в этой главе), сама выработала в процессе эволюции особую версию механизма ламарковского типа.

Целый ряд примечательных данных по растениям и животным, полученных совсем недавно, указывает на то, что эукариоты используют обратную транскрипцию для интеграции

ДНК-копий генома РНК вирусов в хромосомы и могут затем использовать эти встроенные последовательности для производства мРНК или белков, обеспечивающих устойчивость к соответствующим вирусам (Feschotte, 2010; Horie et al., 2010; Koonin, 2010c). Эти механизмы еще предстоит исследовать более тщательно, но по идее они должны быть аналогичны CRISPR и, следовательно, являются ламарковскими.

Горизонтальный перенос генов: важная ламарковская составляющая

Одним из главных открытий сравнительной геномики является демонстрация широкого распространения и высокой частоты горизонтального переноса генов среди прокариот, а также значительного уровня горизонтального переноса у одноклеточных эукариот (см. гл. 5 и 7). Прокариоты с легкостью усваивают ДНК из окружающей среды с помощью фагов и плазмид, служащих векторами, или же без векторов, через механизм трансформации, при участии мембранных насосов, специализирующихся на захвате ДНК.

Поглощенная ДНК часто интегрируется в хромосомы прокариот и может быть зафиксирована в популяции, даже если перенесенный генетический материал дает получателю совсем небольшое селективное преимущество или будучи вовсе нейтральным. Явление горизонтального переноса обладает очевидными ламарковскими признаками: ДНК черпаются из окружающей среды, и, естественно, вероятность приобретения генов, которые находятся в избытке в данной среде, гораздо выше, чем вероятность захвата редкого гена. Второй компонент схемы Ламарка, повышение приспособленности за счет приобретенного признака, не реализуется во всех случаях фиксации горизонтального переноса, однако является значимым и достаточно обычным явлением.

Пожалуй, самый простой и привычный пример – эволюция резистентности к антибиотикам (Martinez, 2008; Wright, 2007). Когда чувствительная бактерия попадает в среду, где присутствуют антибиотики, единственный шанс для пришельца выжить заключен в приобретении гена устойчивости путем горизонтального переноса, как правило через плазмиды. Этот распространенный (и исключительно важный в практическом плане) феномен представляет собой ярко выраженный пример наследования по Ламарку. В самом деле, признак – в этом случае активность перенесенного гена, способствующего резистентности к антибиотикам, – приобретает под непосредственным влиянием окружающей среды и очевидным образом оказывается выгодным – часто необходимым в данных конкретных условиях.

Похожая картина наблюдается для генов фотосинтеза в океане: гены бактериородопсина, главного белка светозависимой биоэнергетики (протон-движущей силы) в галофильных археях, а также в многочисленных бактериях, как и гены фотосистем первого и второго типа, участвующие в хлорофиллзависимом фотосинтезе, судя по всему, распространяются горизонтальным переносом с высокой скоростью, часто посредством бактериофагов, выступающих в качестве переносчиков (Alperovitch-Lavy et al., 2011; Falkowski et al., 2008; Sullivan et al., 2006). Эти гены наделяют организм обладателя серьезным селективным преимуществом, так что они фиксируются с высокой частотой.

В целом любой случай горизонтального переноса, при котором приобретенный ген дает реципиенту преимущество с точки зрения воспроизводства в данной среде (которая благоприятствует передаче такого гена), по-видимому, удовлетворяет ламарковским критериям. Исследования по сравнительной геномике показывают, что горизонтальный перенос служит основным способом адаптации бактерий к окружающей среде путем расширения метаболических и сигнальных сетей, куда интегрируются новые горизонтально приобретенные

гены и, таким образом, добавляют новые свойства в уже существующие схемы (Maslov et al., 2009). Количественно горизонтальный перенос, с его ламарковской компонентой, оказывается у прокариот гораздо более важным средством адаптации, нежели дубликация генов (Pal et al., 2005).

Интересным указанием на то, что горизонтальный перенос может быть адаптивным феноменом, служит уже упоминавшееся открытие агентов переноса генов (АПГ). Как отмечалось в главе 5, АПГ являются производными дефектных бактериофагов, которые заключают в себе, по-видимому, случайные фрагменты генома хозяина и переносят их внутри бактериальных и архейных популяций. Интереснейшие наблюдения переноса генов в морских бактериальных сообществах показывают, что АПГ довольно неразборчивы по отношению к бактериям, которых они инфицируют, и обеспечивают очень высокую интенсивность ГПГ (McDaniel et al., 2010). Свойства АПГ еще предстоит детально исследовать, но существует реальная возможность, что эти агенты представляют собой специально предназначенные для горизонтального переноса средства доставки, которые эволюционировали под селективным давлением, направленным на усиление обмена генами. Если это так, напрашивается вывод, что сам ГПГ выступает частично как адаптивный процесс (см. также обсуждение гипотезы оптимизации переноса в гл. 5). Подводя итог сказанному, мы, видимо, не можем избежать вывода, что некоторые из наиболее важных путей эволюции генома – по меньшей мере у прокариот – являются (квази)ламарковскими.

Стресс-индуцированный мутагенез и активизация мобильных элементов: квазиламарковский феномен

Дарвин подчеркивал эволюционную важность случайных, ненаправленных вариаций, в то время как ламарковская эволюция основана на направленной изменчивости, специфически вызываемой экологическими факторами. Реальная эволюция отвергает это противопоставление. Самой яркой иллюстрацией может служить комплекс разнообразных явлений, которые в совокупности известны как стресс-индуцированный мутагенез, одним из важных аспектов которого является активизация мобильных элементов. Явление такого типа впервые было описано Барбарой Макклиток, продемонстрировавшей (в серии классических экспериментов, которые в конечном итоге принесли ей Нобелевскую премию) активизацию «перескакивания генов» в растениях в условиях стресса, а также важность этой стресс-индуцированной мобильности отдельных «управляющих элементов» для возникновения резистентных фенотипов (McClintock, 1984).

Позднее столь же известный и спорный эксперимент Джона Кэрнса (John Cairns) с сотрудниками по восстановлению мутаций в Лас-опероне, индуцированному лактозой, впечатляющим образом вывел ламарковский механизм эволюции на видное место (Brisson, 2003; Cairns et al., 1988; Rosenberg, 2001). Кэрнс и его коллеги обнаружили заметное усиление реверсии мутаций рамки считывания в Лас-опероне в присутствии лактозы и смело предположили, что за наблюдаемым эффектом стоит классический ламарковский механизм эволюции – иначе говоря, лактоза непосредственно и направленно вызвала мутации в Лас-опероне.

Последующие, более тщательные исследования, включая работы Патрисии Фостер и самого Кэрнса, показали, что это не так: стресс, в частности выращивание культуры при недостатке питания, действительно вызывает мутации, но не в специфических локусах (Foster, 2000). Было показано, что все мутации, лежащие в основе реверсии Лас-фенотипа и других подобных фенотипов, индуцированы стрессом (Лас-клетки, высеянные на питательную среду с лактозой в

качестве единственного источника углерода, испытывают голодный стресс), а не берутся из предсуществующего запаса редких, спонтанных мутаций.

Стресс-индуцированный мутагенез – в частности, механизм мутагенного восстановления в кишечной палочке, известный как SOS-репарация, – был открыт задолго до опытов Кэрнса. Более того, Мирослав Радман (Radman, 1975) и Харрисон Эколс (Echols, 1981) независимо друг от друга пришли к плодотворной мысли, что эта мутагенная форма репарации может быть адаптивным механизмом антистрессовой реакции, а не просто сбоем в работе восстановительных систем. Два десятилетия дальнейших исследований подтвердили эту замечательную идею, сомневаться в истинности которой уже нет разумных оснований. Несколько групп убедительных исследований подтверждают адаптивный характер неточной репарации ДНК (Foster, 2007; Galhardo et al., 2007; Rosenberg, 2001).

Активность SOS-каскада и других мутагенных механизмов репарации в бактериях тщательно регулируется, в частности, переключением с точного воспроизведения к подверженному ошибкам восстановлению разрывов двойной спирали под воздействием сигма-фактора РНК-полимеразы, RpoS , с тем чтобы, по-видимому, достичь оптимальной скорости мутаций. Важнее всего, что стресс-индуцированные мутации, возникающие вследствие склонных к ошибкам процессов репарации, хотя и не нацелены на конкретные гены, в то же время не разбросаны по геному беспорядочно. Напротив, эти мутации концентрируются вокруг двухцепочечных разрывов ДНК, которые вызваны различными стресс-факторами и привлекают к себе аппарат мутагенной репарации.

Мутагенная репарация могла возникнуть как специфический адаптивный механизм, который делает возможной координированную эволюцию групп функционально связанных генов (ключевая особенность геномной архитектуры у прокариот) в тех редких клетках, где происходят полезные мутации, одновременно ограничивая ущерб для других частей генома. Стресс-индуцированный мутагенез, в особенности активация ретротранспозонов, был продемонстрирован также у дрожжей и животных, и это дает основание предполагать, что такой путь адаптивной эволюции универсален для клеточных форм жизни.

По крайней мере среди бактерий стресс-индуцированный мутагенез – не редкий или экзотический, а крайне распространенный процесс. Среди сотен изученных природных штаммов *E. Coli* индуцированный мутагенез характерен для более 80 процентов стареющих колоний, и превышение числа мутаций, запускаемых стрессом, над конститутивными мутациями варьирует на несколько порядков (Vjedov et al., 2003).

Примечательно, что стресс-индуцированная и, по всей видимости, адаптивная нестабильность генома лежит также в основе рака. Хорошо известно, что опухоли развиваются (эволюционируют) в условиях постоянного кислородного стресса, который вызывает обширные перестройки генома и мутации. Главным образом благодаря этим стресс-индуцированным изменениям выживают мутанты, способные к неконтролируемому росту в условиях стресса. Несмотря на различия в конкретных механизмах мутагенной репарации и ее регулирования, злокачественные опухоли животных (включая человека) в принципе не так уж отличаются от бактериальной популяции, эволюционирующей в стрессовых условиях.

Адаптивная эволюция, происходящая в результате стресс-индуцированного мутагенеза, не является строго ламарковской, потому что стресс не вызывает мутации непосредственно и исключительно в генах, отвечающих за устойчивость к данному стрессу. Вместо этого в организмах развились механизмы, которые в ответ на стресс вызывают неспецифический мутагенез. Однако этот процесс, как оказывается, тонко настроен таким образом, чтобы минимизировать ущерб от вредных мутаций в тех редких геномах, которые содержат полезную мутацию. Механизмы этого типа лучше всего определить как квазиламарковские.

Действительно, в случае стресс-индуцированного мутагенеза необходимо учитывать следующее:

1. Условия окружающей среды приводят к появлению мутаций.

2. В результате индуцированных мутаций возникает адаптация к факторам стресса, запустившим мутагенез.

3. Мутагенная репарация управляется сложными механизмами регуляции, что не оставляет никаких сомнений относительно адаптивного характера этого процесса.

Существует прямая связь между ламарковским аспектом стрессиндуцированного мутагенеза и горизонтальным переносом, проявляющаяся в явлении переноса детерминант резистентности, индуцируемого антибиотиками. Многие антибиотики вызывают SOS-ответ, что, в свою очередь, приводит к мобилизации интегративных конъюгационных элементов, которые служат переносчиками генов устойчивости к антибиотикам (Barriss et al., 2009). Аналогия с АПГ очевидна и абсолютно уместна. Здесь мы наблюдаем конвергенцию различных механизмов изменения генома в ламарковской модальности эволюции.

Континуум дарвиновских и ламарковских механизмов эволюции

В предыдущих разделах мы обсудили значительное разнообразие явлений. Некоторые из них, по-видимому, строго соответствуют критериям Ламарка, тогда как другие можно квалифицировать как квазиламарковские (см. табл. 9–1). Принципиальное различие между дарвиновским и ламарковским механизмами эволюции состоит в том, что первый опирается на случайную, ненаправленную изменчивость, второй же основан на вариациях, непосредственно обусловленных внешними стимулами и влекущих за собой специфический ответ на этот стимул (см. рис. 9–1). Ни Ламарк, ни Дарвин не знали ничего о механизмах возникновения и закрепления наследственных изменений, так что им было относительно легко допустить мысль, что фенотипические модификации напрямую транслируются в наследственные (или, как сказали бы в наше время, генетические или геномные) изменения. Тем не менее жесткий ламарковский сценарий крайне требователен, поскольку подразумевает, что должен существовать молекулярный механизм для точного перевода фенотипического изменения в соответствующую модификацию генома (мутацию). По всей видимости, фундаментального механизма для такой обратной геномной инженерии не существует, и не лишено смысла предположение, что *подобные механизмы находятся под жестким контролем отбора, направленного против дестабилизации генома*. Более того, передача информации от белков к нуклеиновым кислотам была бы крайне затруднительна физико-химически – эта трудность, по всей видимости, отражает разграничение между матричными и каталитическими биомолекулами, возникшее на ранних этапах эволюции жизни (см. гл. 12). Центральная догма молекулярной биологии (Crick, 1970), согласно которой *поток информации от белков к нуклеиновым кислотам отсутствует*, является частичным воплощением этого разделения. Однако в принципе обратный поток некоторых типов информации от фенотипа – или от окружающей среды, рассматриваемой как расширенный фенотип, – в геном не является невозможным, учитывая широкое распространение обратной транскрипции и ДНК-транспозиции. Для работы истинно ламарковского сценария требуются чрезвычайно изощренные механизмы; в двух примечательных случаях, систем CRISPR-Cas и piRNA (описанных ранее в этой главе), такие механизмы были обнаружены.

Существование других полноценных ламарковских систем вполне представимо и даже вероятно, что предполагает, в частности, открытие вирус-специфических последовательностей, которые потенциально несут устойчивость к генетически родственным вирусам в растительных и животных геномах (см. гл. 10). Тем не менее эти механизмы вряд ли представляют собой доминирующую тенденцию в геномной эволюции, возможно, в силу вышеупомянутой селекции против чрезмерной нестабильности генома. Механизмы, обозначенные в предыдущих разделах как квазиламарковские, напротив, повсеместны. По существу, они оказываются не менее поразительными – и не менее сложными, – чем оригинальный ламарковский сценарий: квазиламарковские процессы переводят случайные мутации в конкретные адаптивные реакции в ответ на стимулы внешней среды.

Тема мощных, часто неблагоприятных воздействий окружающей среды на организмы представляется общим мотивом для различных аспектов (квази)ламарковского режима эволюции, описанных в данной главе: для системы CRISPR-Cas, стресс-индуцированного мутагенеза и других явлений (см. табл. 9–1). Эта связь, скорее всего, не иллюзорна: кажется вполне логичным, что сильные (чрезвычайные) сигналы из окружающей среды инициируют (квази)ламарковские процессы, в то время как относительно слабые («обычное дело») сигналы транслируются в дарвиновскую модальность эволюции (см. рис. 9–3).

Обсуждая эволюционное значение горизонтального переноса генов, Энтони Пул предположил, что ламарковская составляющая горизонтального переноса становится иллюзорной, если посмотреть на эволюцию с «точки зрения гена» (Poole, 2009). Действительно, ламарковская модальность связана в первую очередь с организменным уровнем сложности и не распространяется на самый фундаментальный уровень эволюции, включающий гены, независимо эволюционирующие части гена (те, что кодируют различные белковые домены), а также мобильные элементы (см. гл. 6). Таким образом, ламарковская эволюция выступает в качестве «эмергентного феномена». Это, пожалуй, неудивительно, учитывая сложность механизмов, необходимых для интеграции нового генетического материала в геном при реализации ламарковской схемы.

Таблица 9–1. Ламарковские и квазиламарковские явления.

Явление	Биологическая роль/ функция	Филетическое распространение	Ламарковский критерий		
			Геномные изменения вызываются воздействием среды	Изменения ограничены специфическими геномными локусами	Изменения обеспечивают адаптацию к конкретному воздействию
Ламарковские					
CRISPR-Cas	Защита от вирусов и др. мобильных элементов	Большинство архей и бактерий	Да	Да	Да
piРНК	Защита от мобильных элементов в зародышевой линии	Животные	Да	Да	Да
ГПГ (особые случаи)	Адаптация к новой среде, ответ на стресс, резистентность	Археи, бактерии, одноклеточные эукариоты	Да	Да	Да
Квазиламарковские					
ГПГ (общий случай)	Различные инновации	Археи, бактерии, одноклеточные эукариоты	Да	Нет	Да/нет
Стресс-индуцированный мутагенез	Ответ на стресс и устойчивость к стрессу, адаптация к новым условиям	Повсеместно	Да	Нет или частично	Да, но общая эволюционированность также улучшается

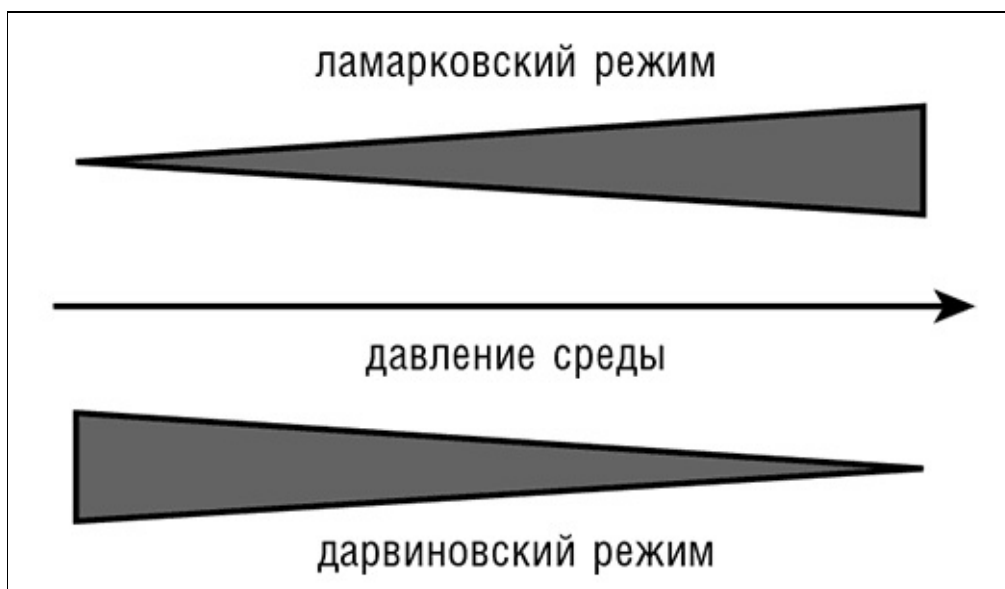


Рис. 9–3. Давление среды и переход от дарвиновского к лamarковскому режиму. Заимствовано из Koonin and Wolf, 2009b.

В целом сравнение между дарвиновским и лamarковским сценариями предполагает, что эволюция представляет собой континуум процессов, от совершенно случайных до истинно адаптивных, которые тонко регулируются для обеспечения определенного ответа на конкретный стимул. Принципиальный вывод, напрашивающийся по итогам многих новейших исследований, упомянутых в этой главе, состоит в том, что изменчивость генома предстает гораздо более сложным явлением, чем это казалось ранее, и регулируется сразу на нескольких уровнях, для того чтобы обеспечить адаптивную реакцию на изменения в окружающей среде. Устранение конфликта между лamarковским и дарвиновским эволюционными сценариями, обладая далеко не только историческим значением, влияет на наши фундаментальные взгляды на роль и место случайности в эволюции. Это представляется подлинной, хоть и недооцениваемой, сменой парадигмы в современной биологии.

Точность передачи информации в биологических системах и ее (не)адаптивная эволюция

Эволюция жизни целиком основана на процессах передачи дискретной информации – между поколениями путем репликации генома и от генома к эффекторным молекулам (РНК и белкам) – как описано в главе 2. Не существует канала передачи информации, свободного от ошибок, как впервые математически сформулировал Клод Шеннон, создавший теорию информации, связав процесс передачи информации с законами термодинамики. Как отмечалось в главе 2, точность репликации генома не должна быть ниже определенного минимума; соответственно, частота мутаций не должна превышать определенный порог, чтобы избежать мутационного вырождения популяции. Очевидно, что частота мутаций не может быть и слишком низкой, дабы оставить возможность хотя бы для минимальной эволюционируемости (потенциала для эволюции – см. обсуждение в следующем разделе). Менее ясно, важен ли этот нижний предел практически в реальных биологических системах. Таким образом, фундаментальный вопрос сводится к следующему: каким образом отбор контролирует частоту мутаций (если он это действительно делает)? Точнее, удерживает ли очищающий отбор частоту мутаций просто ниже порога вырождения, или же происходит по крайней мере, для некоторых организмов и, возможно, в особых ситуациях, отбор на достаточно высокую частоту мутаций, чтобы обеспечить сырье для эволюции?

Отбор на приемлемую точность репликации (и, более глобально, всех процессов передачи информации) является одним из центральных аспектов эволюции. Это непосредственно видно из того огромного разнообразия, сложности и многоуровневой организации репарационных систем, которые обнаруживаются во всех клеточных формах жизни (Aravind et al., 1999; Friedberg et al., 2005). У прокариот до 10 процентов кодирующей емкости генома могут быть заняты компонентами систем репарации, которые действуют на всех этапах репликации ДНК, а также устраняют различные мутационные повреждения, происходящие за пределами процесса репликации. С другой стороны, существует класс репликаторов, у которых (практически) отсутствуют механизмы исправления повреждений; это РНК-содержащие вирусы. Действительно, эти вирусы демонстрируют чрезвычайно высокий уровень ошибок включения нуклеотидов и общий уровень мутаций, явление, хорошо известное в связи с важностью для медицины быстрой эволюции вирусов гриппа и ВИЧ (Holmes, 2009). Эти вирусы, по-видимому, эволюционируют не слишком далеко от порога мутационного вырождения (Drake and Holland, 1999). Все РНК-вирусы обладают небольшими геномами (менее 30 Кб), что отчасти является следствием физической хрупкости длинных молекул РНК, но также связано с отсутствием репарационных механизмов. (Можно было бы утверждать, что сложные системы репарации в таких вирусах не могли развиваться, потому что они не были бы выгодны, учитывая фундаментальную нестабильность генома.) На самом деле РНК-содержащие вирусы с самыми большими геномами (вирусы отряда *Nidovirales*), по всей видимости, обладают различными, хотя и простыми, системами репарации (Eckerle et al., 2007). Кроме того, открытие РНК-деметилаз в геномах различных растительных РНК-содержащих вирусов означает, что даже в этих простейших геномах репарация может эволюционировать, когда вирус распространяется в условиях повышенного давления окружающей среды (Aravind and Koonin, 2001; van den Born et al., 2008).

В рамках концепции мутационного вырождения возникает естественная мысль, что частота мутаций на нуклеотид должна быть обратно пропорциональна размеру генома организма, так чтобы число мутаций на геном на поколение оставалось примерно постоянным. Ян Дрейк

первым выразил эту идею явным образом, и потому ее часто называют гипотезой Дрейка (Drake, 1991). Гипотеза Дрейка довольно хорошо работает для вирусов и прокариот. Однако изучение новейших данных, проведенное Майклом Линчем, неожиданно (по крайней мере, на первый взгляд) выявило противоположную зависимость у эукариот: скорость мутирования в расчете на нуклеотид положительно коррелирует с размером генома (Lynch, 2010). Следуя в русле неадаптивной теории эволюции сложности, Линч показал, что небольшое повышение частоты мутаций не будет «замечено» очищающим отбором в небольших популяциях, которые типичны для многоклеточных эукариот, и, соответственно, не может элиминироваться в ходе эволюции этих организмов. Отсюда появляется «полуадаптивная» гипотеза эволюции мутационного фона: существует селективное давление на понижение частоты мутаций ниже порога вырождения и чуть дальше, поскольку за счет этого популяция становится более устойчивой, однако не на окончательную минимизацию частоты мутаций. В соответствии с этой гипотезой отбора, который бы препятствовал падению частоты мутаций ниже любого минимального значения, не существует; частота мутаций остается относительно высокой по чисто стохастическим причинам (см. также следующие разделы в этой главе).

Однако ситуация не столь проста, как видно из результатов долгосрочных экспериментов с эволюционирующими популяциями кишечной палочки, проведенных Ричардом Ленски и его коллегами. Эти эксперименты показывают, что интенсивный отбор на адаптацию бактериальных популяций к новой среде часто ведет к появлению аллелей-мутаторов (иными словами, бактерий с повышенной частотой мутаций в связи с повреждением одного из ферментов репарации), которые вытесняют предыдущие поколения с низкими показателями мутаций (Sniegowski et al., 1997). Если быть более точным, представляется, что аллели-мутаторы достигают значительного распространения в популяции и даже фиксируются за счет тесного сцепления с адаптивными мутациями, вызываемыми этими мутаторами. Однако, когда внешнее селективное давление убрано, мутаторы становятся невыгодными и вымываются отбором (Denamur and Matic, 2006). Эти результаты приводят к важному обобщению: *в зависимости от целого ряда факторов, таких как экологический стресс и эффективный размер популяции, отбор на низкую или высокую частоту мутаций может происходить и происходит в действительности.*

Частота ошибок транскрипции гораздо выше, нежели частота ошибок репликации, а ошибки трансляции случаются еще более часто (см. рис. 9–4). Хотя экспериментальные измерения уровня ошибок включения аминокислот в ходе трансляции немногочисленны и ограничены несколькими модельными системами, все же ясно, что точность трансляции удивительно низка. В самом деле, частота встраивания ошибочных аминокислот составляет 10^{-4} — 10^{-5} — любопытно, что она близка к частоте ошибок репликации у РНК-вирусов. Получается, что около 20 процентов белковых молекул, синтезируемых в любой из клеток, содержат по меньшей мере одну неверную аминокислоту (Drummond and Wilke, 2009). Последствия ошибок транскрипции и трансляции, иногда метко называемых фенотипическими мутациями, очевидно, менее критичны, чем последствия генетических мутаций, по той причине, что фенотипические мутации, как правило, не наследуются (существуют примечательные исключения, например обратная транскрипция с последующим включением в геном ДНК-копии ошибочно транскрибируемой РНК; Burger et al., 2006). Учитывая сравнительно короткий срок жизни любой РНК или белковой молекулы, никакая фенотипическая мутация сама по себе не может оказать серьезного влияния на выживаемость, поэтому неудивительно, что для фенотипических мутаций допустимы гораздо более высокие частоты ошибок, чем для генетических мутаций. Однако столь же очевидно, что чрезмерно высокие темпы фенотипических мутаций несовместимы с жизнью. А значит, как и в случае с системами

репарации ДНК, многочисленные механизмы, служащие для отслеживания ошибок транскрипции и трансляции, безусловно, существуют. Было показано, что коррекционная активность ДНК-зависимой РНК-полимеразы уменьшает процент ошибок на несколько порядков (Alic et al., 2007; Sydow and Cramer, 2009). Кроме того, также были открыты остающиеся пока слабоизученными процессы посттранскрипционного исправления ошибок метилирования в РНК (Begley and Samson, 2003; Falnes, 2005). Вероятно, из механизмов, контролируемых частоту фенотипических мутаций, лучше всего исследована коррекция аминоксил-тРНК-синтетазой (АРСазой), при которой молекулы аминоксил-тРНК, связанные с ошибочными аминокислотами, гидролизуются и утилизируются (Hussain et al., 2010; Ling et al., 2007). Упомянутая коррекция АРСазой дополняется рибосомной коррекцией на следующей стадии трансляции, при которой рибосома отбраковывает ошибочно связанные тРНК (Blanchard et al., 2004; Daviter et al., 2006). Однако значительное увеличение надежности трансляции, по-видимому, вступает в противоречие с требованием высокой скорости синтеза белка. Существенное повышение точности трансляции может быть легко достигнуто путем мутации конкретных позиций в рРНК или в рибосомных белках, но эти мутации оказываются вредными для клеток, по-видимому, из-за медленной трансляции (Dong and Kurland, 1995; Johansson et al., 2008).

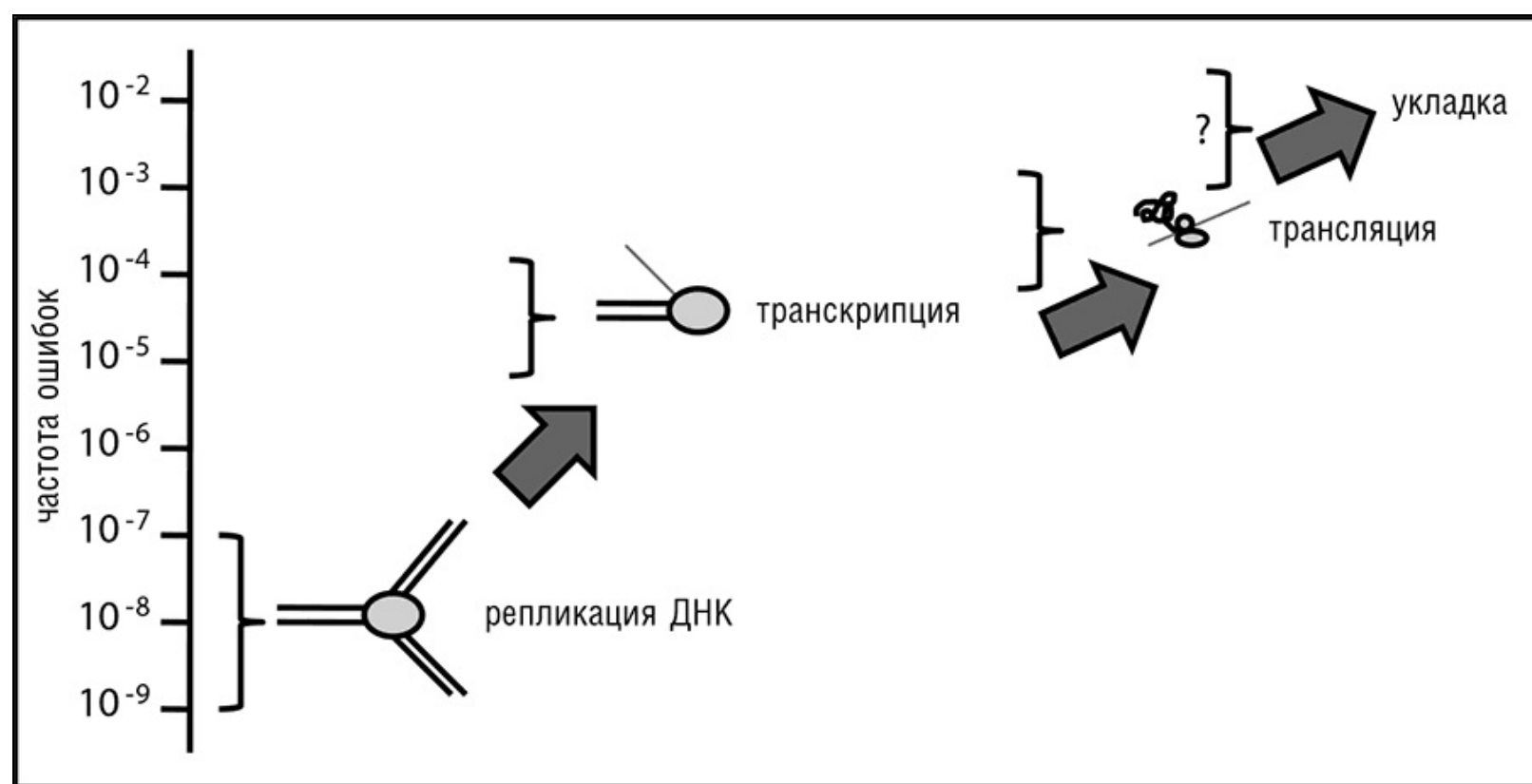


Рис. 9–4. Типичные частоты ошибок на различных стадиях передачи биологической информации.

Эксперименты с точными, но медленными рибосомными мутантами позволяют предположить, что компромисс между скоростью и точностью трансляции связан с механическими ограничениями трансляционной системы и с трудом может быть преодолен мутациями в ее компонентах. В результате для ограничения ошибок трансляции и их пагубных последствий эволюция, похоже, выбрала другие способы адаптации.

Твердо установлено, что «высокостатусные» гены (те, что интенсивно экспрессируются и медленно эволюционируют) обладают более выраженным предпочтением триплетов по

сравнению с «низкостатусными» генами. Оптимальные триплеты, которыми насыщены гены с высоким статусом, обеспечивают более низкую частоту ошибок, равно как и более высокую скорость трансляции, и, таким образом, частично избегают вышеупомянутого компромисса (Drummond and Wilke, 2009, 2008). Асимметрия триплетов между высокостатусными и низкостатусными генами может объясняться ценой селекции; в силу этого заметный отбор оптимальных кодонов может идти только в генах высокого статуса.

Основным вредным эффектом ошибок трансляции считается неправильная укладка белка (Drummond and Wilke, 2009, 2008), хотя ошибки включения аминокислот в каталитических сайтах, безусловно, могут стать дополнительным фактором. Как уже говорилось в главе 4, отбор на устойчивость к неправильной укладке является одним из главных аспектов эволюции белков – возможно, даже ее основной движущей силой. Менее ясно, какой источник неверной укладки наиболее важен – изначально неправильная последовательность или ошибки при ее трансляции. В любом случае, хотя укладка белка обычно не рассматривается в качестве процесса передачи информации, в действительности она им является. В самом деле, *укладка предполагает поток информации, идущий от одномерной аминокислотной последовательности к трехмерной структуре белка.*

Ровно то же самое относится к структурным РНК. Частоту неправильной укладки трудно определить экспериментально, и этого не было сделано для большого числа белков или РНК. Если тенденция, отражающая процент ошибок и показанная на рис. 9–4, – чем дальше от генома, тем менее точен этап передачи информации, – служит каким-либо индикатором, частота ошибок укладки должна быть даже выше частоты ошибок трансляции.

Такое предсказание также следует из здравого смысла, учитывая невероятную сложность процесса укладки и огромное количество ошибочных вариантов, доступных в принципе для укладываемого белка или РНК-молекулы (Bowman et al., 2011; Pande et al., 1998). Имея в виду высокую сложность пространства укладок, эпохальное открытие (сделанное первоначально Кристианом Анфинсеном и впоследствии подтвержденное многочисленными экспериментами), что белки способны самопроизвольно складываться в нативную конформацию, вызвало огромное удивление (Anfinsen, 1973).

Спустя почти 50 лет после открытия Анфинсена все еще остается предметом споров, глобальный или локальный минимум свободной энергии ищут спонтанно укладываемые белки. Но стало ясно, что лишь небольшие белки укладываются спонтанно; большинство белков нуждаются в специальных молекулярных устройствах, других белках, известных как шапероны, чтобы сформировать правильную структуру. Шапероны функционируют удивительным образом: их молекулы образуют «ячейку» (известную также как ячейка Анфинсена), которая изолирует укладываемый белок от цитоплазмы и частично разворачивает его, облегчая тем самым поиск нативной конформации (Ellis, 2003). Большинство шаперонов – синтезируемые в больших количествах, высококонсервативные, высокостатусные белки.

Первоначально некоторые из шаперонов были открыты как «белки теплового шока», то есть белки, которые резко усиливают свою активность при повышенной температуре (и, как было показано позже, при других стрессовых условиях) и противодействуют неправильной укладке других белков, которая усугубляется при стрессе (Vabulas et al., 2010). Хотя это явление менее детально исследовано, белковые шапероны также способствуют укладке молекул РНК (Russell, 2008; Woodson, 2010). В целом контроль над укладкой белков (и, вероятно, РНК) является, без сомнения, одной из основных функций во всех клетках.

Помимо устройств, подобных шаперонам, все клетки задействуют арсенал разнообразных молекулярных машин для контроля управляемого расщепления белков, в частности неправильно уложенных, и РНК. Как и молекулярные шапероны, эти машины – протеасомы, в

случае белков, и экзосомы (деградосомы у бактерий) в случае РНК – повсеместно распространены во всех трех доменах жизни, присутствуют в избытке в большинстве клеток и подвержены регулированию в условиях стресса (Hartung and Hopfner, 2009; Volker and Lupas, 2002; см. также гл. 7). Кроме того, эти машины, наряду с дополнительными вспомогательными системами регулируемого протеолиза, являются основными внутриклеточными потребителями энергии (АТФ). Бактерии дополнительно обладают высококонсервативными системами так называемой транс-трансляции, которые освобождают забуксовавшие рибосомы из аберрантных мРНК, на которых трансляция не в состоянии прекратиться должным образом, и предназначают такие мРНК и их белковые продукты (также аберрантные) к разрушению (Keiler, 2008).

Как мы подробно обсуждали в главе 7, эукариоты обладают важной стадией обработки информации, которая фактически не имеет эквивалента у прокариот: сплайсинг первичных транскриптов. Сопутствующая система контроля качества, по-видимому, эволюционировала одновременно с возникновением эукариот (см. гл. 7): механизм нонсенс-опосредованного распада (НОР), распознающий и уничтожающий аберрантные мРНК, которые содержат стопкодоны внутри экзона помимо последнего, 3'-концевого экзона кодирующей последовательности (Behm-Ansmant et al., 2007; Stalder and Muhlemann, 2008).

Итак, контроль частоты ошибок и их влияния на биологические процессы передачи информации представляется одним из ключевых аспектов эволюции. По причинам, которые мы понимаем лишь частично (в лучшем случае), процент ошибок, по-видимому, не падает сильно ниже максимально допустимого значения: порога мутационного вырождения и соответствующего катастрофического порога фенотипических мутаций, который не изучен подробно, но предположительно существует. В случае частоты мутаций, простая неадаптивная теория популяционной генетики вполне способна объяснить наблюдаемые значения с достаточной степенью надежности (Lynch, 2010). Аналогичная аргументация была применена к фенотипическим мутациям (Burger et al., 2006), но в этом случае решение представляется менее очевидным.

Существует противоречие между относительно высокой частотой ошибок транскрипции, трансляции, сплайсинга, а также, весьма вероятно, укладки и исключительной сложностью устройств предотвращения повреждений, таких как протеасомы, экзосомы, системы НОР и др. Эволюция этих многоуровневых систем контроля предполагает, что *вредное воздействие фенотипических мутаций при той частоте, что наблюдается при воспроизводстве клеток, довольно значительно, однако издержки отбора на повышение точности процессов были бы непосильными, и, соответственно, неоднократно выбирались альтернативные пути эволюции систем предотвращения повреждений.*

В целом, по-видимому, *борьба с энтропией является одним из важнейших аспектов эволюции.* Отбор на уменьшение и контроль энтропии носит универсальный характер и последовательно распространяется по этапам передачи информации, от репликации до укладки и сортировки белков и РНК. Антиэнтропийная эволюция частично ведет к снижению самой частоты мутаций/ошибок, а частично действует на уровне контроля повреждений. Эволюционные эксперименты показывают, что отбор на повышение частоты мутаций действительно происходит. Тем не менее остается неясным, широко ли распространен такой отбор на повышение шума и достаточен ли уровень шума, который антиэнтропийные механизмы не способны устранить, чтобы обеспечить изменчивость, необходимую для эволюции. Мы обсудим эту ключевую проблему в следующих разделах.

Шум в биологических системах и его созидательная роль в эволюции

Никакой информационный канал не может быть свободным от шумов (см. гл. 4 и 8, и предыдущий раздел). Этот фундаментальный, обусловленный законами термодинамики аспект передачи информации делает эволюцию возможной благодаря достаточно высокой частоте ошибок репликации, даже в отсутствие отбора на повышение изменчивости (см. гл. 2). Как отмечалось в гл. 4, практически нейтральные мутации, накапливающиеся по чисто энтропийным причинам, образуют почти нейтральные сети, которые служат резервуарами эволюции, в том числе положительного отбора, особенно в изменяющихся условиях. Подчеркнем еще раз, что не только фиксирующиеся мутации, но и полиморфизмы, которые особенно многочисленны в больших популяциях, вносят вклад в важные для эволюции нейтральные сети. Кроме того, мы неоднократно видели, что увеличение частоты мутаций может быть полезно – и поддерживается отбором, как это имеет место в случае стресс-индуцированного мутагенеза и аллелей-мутаторов.

Здесь я хочу сосредоточиться на фенотипических мутациях и их возможной роли в эволюции. Стандартное, само собой разумеющееся представление сводится к тому, что фенотипический шум не имеет значения для эволюции (до тех пор, пока его вредные эффекты находятся под контролем), по той очевидной причине, что фенотипические мутации не наследуются. Тем не менее такой взгляд может страдать близорукостью благодаря так называемому упреждающему эффекту фенотипических мутаций. Как мы неоднократно наблюдали на страницах этой книги, ключевая проблема в ходе эволюции заключается в том, чтобы лавировать по неровному эволюционному ландшафту, двигаясь от локальных пиков приспособленности или плато к большим высотам на других пиках или плато. Этот путь часто проходит через расщелины, порой глубокие. Чтобы их преодолеть, необходима фиксация двух или более мутаций.

Очевидно, что это нетривиальная задача, учитывая, что первая мутация приводит к снижению приспособленности. В принципе такие трюки могут быть проделаны с помощью дрейфа, но шансы на фиксацию нескольких мутаций чрезвычайно низки, если первая требуемая мутация вредна; они падают до нуля, когда первая мутация оказывается летальной. Именно в этих случаях фенотипические мутации могут прийти на помощь. Когда для приобретения признака требуются две мутации, фенотипические мутации способны обеспечить незначительный уровень второй мутации, как только случилась первая. Если первая мутация вредна или даже летальна, в то время как две мутации, взятые вместе, дают преимущество, фенотипические мутации могут спасти организм и обеспечить для него выгодный фенотип (позволяя достичь труднодоступной высокой точки в ландшафте), пусть и с низкой частотой. Математическое моделирование показывает, что такой упреждающий эффект реализуем в достаточно широком диапазоне эффективных размеров популяции и коэффициентов отбора (Whitehead et al., 2008).

Насколько распространен и важен упреждающий эффект? Мы не знаем, но, если бы он был важен и имел адаптивное значение, можно было бы предсказать, что стресс должен вызывать ошибки трансляции в управляемом режиме, подобно тому как он вызывает мутагенез (см. обсуждение стресс-индуцированного мутагенеза ранее в этой главе). Как это ни удивительно, дело обстоит именно так, как показали недавние исследования, в которых окислительный стресс приводил к резко возросшей частоте ошибочной трансляции, по крайней мере отчасти за счет нарушения корректирующей функции APCаз (Ling and Soll, 2010; Netzer et al., 2009).

Подвержены ли стресс-индуцированные фенотипические мутации регуляции и контролю того же типа, что и стресс-индуцированный генетический мутагенез, только предстоит выяснить. Тот факт, что частота ошибок включения метионина, аминокислоты, которая защищает белки от окислительного повреждения, специфически усиливается в условиях окислительного стресса (Netzer et al., 2009), наталкивает на вывод, что механизмы управления фенотипическими мутациями существуют, так что стресс-индуцированный фенотипический мутагенез можно рассматривать как адаптивную стратегию.

Эволюционная значимость фенотипических мутаций – потенциально широкая и важная тема, и недавние исследования едва успели ее коснуться. Однако практически не вызывает сомнений, что фенотипический шум (энтропия) не является эволюционно несущественным. На самом деле за пределами очевидного селективного давления, держащего его в узде, шум может быть конструктивным фактором в разнообразных эволюционных процессах. Фенотипические мутации являются частью более общей, всеобъемлющей тенденции, которую мы затрагивали в гл. 8. Основная идея состоит в том, что шум, присутствующий на всех уровнях и во всех каскадах передачи биологической информации – типичные генетические и фенотипические мутации, ложная транскрипция, ошибки сплайсинга, неправильная укладка белка, – будучи бременем для эволюционирующих организмов, в то же время повышает их эволюционный потенциал. Простое, но важнейшее следствие заключается в том, что *популяции с низкой интенсивностью очищающего отбора, который не в состоянии устранить или существенно снизить уровень шума, обладают наибольшим эволюционным потенциалом.*

Эволюция эволюционируемости, надежность биологических систем и реализуемость эволюционного предвидения

Эволюционируемость, потенциал биологических систем изменяться в ходе эволюции, является одной из многих важных концепций, введенных Ричардом Докинзом (Dawkins, 2006). В первом десятилетии XXI века «эволюционируемость» стала модным термином и предметом интенсивных дебатов (Brookfield, 2009; Kirschner and Gerhart, 1998; Masel and Trotter, 2010; Radman et al., 1999). Как таковая, способность к эволюционированию не представляет проблемы в качестве неотъемлемого свойства реплицирующихся систем (см. принцип подверженной ошибкам репликации (ПОР), введенный в гл. 2).

Как говорилось ранее в этой главе, эволюция любого организма (с некоторыми оговорками в случае РНК-вирусов) предполагает отбор на уменьшение частоты геномных и фенотипических мутаций. Тем не менее этот отбор относительно неэффективен по причинам, которые для случая геномных мутаций находят объяснение в рамках популяционной генетики, но которые менее ясны в случае фенотипических мутаций, учитывая наличие сложных, специализированных систем, предназначенных для устранения aberrантных продуктов. Таким образом, эволюционируемость, несомненно, является объектом отбора, в очевидном смысле поддержания достаточно низкого уровня шума.

Нулевая гипотеза в отношении эволюционируемости, следовательно, сводится к тому, что остаточный шум, не устраненный очищающим отбором, достаточен (и, более того, необходим), чтобы обеспечить изменчивость, являющуюся первичным материалом для эволюции. Такой «конструктивный» шум представляет собой неадаптивный побочный продукт эволюции. Яблоком раздора служит интригующая возможность, что эволюционируемость эволюционирует нетривиальным образом – иначе говоря, повышенная изменчивость при определенных обстоятельствах и/или в некоторых генных локусах может быть признаком, подверженным отбору. Идея «адаптивной эволюционируемости» звучит как анафема для многих биологов, поскольку отдает «эволюционным предвидением» или временной нелокальностью эволюции, тогда как общепринятая догма состоит в том, что эволюция «близорука», то есть носит строго локальный характер. Однако, невзирая на это общее убеждение, выводы, представленные в данной главе, ясно показывают, что определенные формы эволюционируемости могут эволюционировать и быть адаптивными и что в широком смысле эволюция способна к предвидению.

Попробуем кратко суммировать свидетельства в поддержку адаптивной эволюции эволюционируемости. Возможно, наиболее сильные свидетельства содержатся в обширной серии экспериментальных данных по системам подверженной ошибкам стресс-индуцированной репарации; фактически они представляют собой механизм стресс-индуцированного мутагенеза. Эти системы, в частности аппарат SOS-репарации и мутагенеза у бактерий, по-видимому, не обладают механизмами инициировать мутации именно в тех генах, что задействованы в борьбе с конкретной формой стресса. Тем не менее они определенно способствуют выживанию, увеличивая общую частоту мутаций и повышая тем самым вероятность адаптации. Сложно организованная регуляция этих систем и их широкое распространение у микробов, особенно обитающих в изменчивых условиях, не оставляют сомнений, что стресс-индуцированный мутагенез является адаптивным явлением и представляет собой характерную форму эволюционного предвидения – или, точнее, *эволюционной экстраполяции*.

Эволюционный процесс конечно же не может «знать», что произойдет в каждый

следующий момент, но в среде с часто возникающим стрессом организмы, которые вырабатывают способность временно поднять частоту мутаций, получают дополнительный шанс на выживание. Этот эффект еще более усиливается за счет кластеризации мутаций, чему содействует привлечение стресс-индуцированных репарационных ферментов к поврежденным участкам ДНК (Galhardo et al., 2007; Rosenberg, 2001). В дополнение к этому эволюционные эксперименты показывают, что в условиях стресса повышенная частота мутаций, вызванная мутациями в генах репарации, может обеспечить селективное преимущество организмам, несущим аллели-мутаторы.

Позволив себе антропоморфизм, можно сказать, что эволюция не способна предвидеть, что произойдет в действительности, однако может экстраполировать исходя из трудностей прошлого, что неприятности рано или поздно обязательно случатся, так что единственный шанс выжить состоит в том, чтобы быть подготовленным и включать ускоренное мутирование при появлении проблемы. Это рискованная стратегия, но, по всей видимости, единственная, которая может развиться эволюционно. Реализуется ли аналогичная стратегия на уровне фенотипических мутаций, менее ясно, но такая возможность выглядит реалистичной, учитывая исследования по индуцированию ошибок трансляции, возможно специфических, в условиях стресса.

Ярким доказательством в поддержку эволюции эволюционируемости служат АПГ, специализированные вирусоподобные агенты горизонтального генетического переноса у бактерий и архей (см. гл. 4 и 10). Учитывая, что горизонтальный перенос является доминирующим эволюционным процессом у прокариот, наличие специальных устройств, которые, насколько мы понимаем, эволюционировали исключительно под селективным давлением для усиления переноса, показывает, что механизмы эволюции сами эволюционируют и, в некоторых случаях, могут рассматриваться как адаптации.

Другой важный аспект проблемы эволюционируемости связан с устойчивостью биологических сетей и так называемым явлением капацитации. Некоторые белки, из которых лучше всего исследован молекулярный шаперон HSP90 (HSP является общим сокращением для белка теплового шока, *Heat Shock Protein*, то есть белка, активируемого высокотемпературным стрессом), обладают свойствами эволюционных «конденсаторов», то есть медиаторов эффектов генетической и фенотипической изменчивости (Masel and Siegal, 2009). Инактивация HSP90 приводит к появлению многочисленных мутантных фенотипов, выявляя скрытую изменчивость: белки с аминокислотными заменами, не приводящими к фенотипическому эффекту в присутствии HSP90, неверно укладываются в его отсутствие (Rutherford et al., 2007) [\[94\]](#).

Оказывается, что капацитация довольно общее явление: в дрожжах около 300 генов (более 5 процентов от общего числа) ведут себя как конденсаторы. Общее свойство конденсаторов состоит в том, что они служат центрами сетей взаимодействия, поэтому, как предполагается, нарушение взаимодействий между конденсатором и другими белками высвобождает скрытые вариации (Levy and Siegal, 2008). Эффект капацитации также, вероятно, важен на уровне фенотипических мутаций. Как обратная сторона капацитации, существует возможность, что HSP90 и, вероятно, другие конденсаторы обладают также способностью действовать в качестве усилителей генотипических и фенотипических мутаций, позволяя мутантным белкам, которые иначе свернулись бы неверно, укладываться должным образом и оказывать фенотипический эффект.

Возможно, самый поразительный известный механизм эволюционируемости представлен прионными белками грибов (Halfmann and Lindquist, 2010; Masel and Bergman, 2003). Подробно исследованный дрожжевой прион (PSI+) является фактором терминации (освобождения полипептидной цепи) трансляции и может самопроизвольно переходить в прионную форму,

которая затем запускает автокаталитическое формирование агрегатов амилоида. В составе прионных агрегатов фактор терминации трансляции не активен, что приводит к частому проскоку стоп-кодонов и, соответственно, к образованию разнообразных белков, удлинённых на конце. В результате проявляется скрытая изменчивость в последовательностях 3'-НКР, эволюция которых лишь в слабой степени контролируется очищающим отбором. Образование прионов в сильной степени стимулируется различными видами стресса (Tyedmers et al., 2008), а прионное состояние наследуемо, это удивительная форма наследственности на основе белка [95].

Следует отметить, что механизм действия прионов увеличивает фенотипическую изменчивость через совместный эффект фенотипических мутаций (проскок стоп-кодона во время трансляции) и проявления скрытой генетической вариабельности. Повышенная выживаемость прион-несущих линий в условиях стресса была прямо продемонстрирована в эксперименте (Tyedmers et al., 2008). Пока еще не ясно, насколько типично опосредованное прионами усиление эволюционированности. Что наиболее примечательно, прионная форма фактора терминации трансляции, похоже, не имеет каких-либо иных функций и, таким образом, вероятно, возникла именно под давлением отбора на эволюционированность [96].

Конденсаторы, по всей видимости, работают как настоящие *регуляторы эволюции*. С одной стороны, эти гены обеспечивают устойчивость биологическим системам и смягчают воздействие мутаций. Несколько парадоксальным образом, однако, капацитация также способствует эволюционированности через эффект усиления, позволяя эволюционирующим организмам увеличивать размер почти нейтральных сетей и, следовательно, потенциал для адаптивной эволюции (Wagner, 2008b).

Связь между устойчивостью и эволюционированностью является ключевым аспектом эволюции в целом. Это сопряжение представляется «системным свойством», в котором устойчивость как защищает эволюционирующие системы от вредных эффектов изменчивости, так и увеличивает их эволюционный потенциал (Kaneko, 2007). Количественный анализ популяционно-генетических моделей показывает, что устойчивость может увеличивать или уменьшать эволюционированность в зависимости от динамики численности популяции и структуры адаптивного ландшафта. В частности, было продемонстрировано, что адаптивная эволюция ускоряется с увеличением размера нейтральной сети (устойчивость), если фенотипы с высокой приспособленностью ограниченно доступны для эволюционирующих организмов (иными словами, ландшафт приспособленности имеет сложную структуру с множеством локальных пиков) (Draghi et al., 2010).

В рамках концепции адаптивного ландшафта явление отбора на устойчивость и эволюционированность стало известно как «выживание наиболее плоских» [97]. Моделирование эволюции цифровых организмов показало, что при высокой скорости мутаций генотипы с относительно низкой частотой репликации, но большими почти нейтральными сетями – то есть занимающие сравнительно низкие, но плоские участки поверхности приспособленности – вытеснили генотипы, реплицирующиеся быстрее, но занимающие высокие, крутые пики (Wilke et al., 2001). С другой стороны, при низкой скорости мутаций эволюция происходит за счет *выживания наиболее приспособленных*– генотипов, находящихся на самых высоких пиках; были обнаружены и промежуточные режимы эволюции, в которых выживали как наиболее приспособленные, так и наиболее плоские (Beardmore et al., 2011). Таким образом, устойчивость и эволюционированность приходят к оптимальному соотношению с приспособленностью. Эволюция максимизирует то или другое из этих свойств фенотипов или все сразу (тем самым поощряя разнообразие), в зависимости от условий.

Краткий обзор и перспектива

В этой главе мы обсудили целый спектр разноплановых данных, моделей и гипотез, объединенных общей темой: они уводят эволюционную биологию в сторону от важной, но упрощенной триады «наследственность – изменчивость – отбор», центральной в синтетической теории эволюции. Даже более реалистичная концептуальная структура сегодняшней эволюционной биологии, куда включена важная роль дрейфа, рекомбинации и горизонтального переноса, представляется существенно неполной. Исследования, рассмотренные в данной главе, открывают более сложное, непредвиденное влияние как со стороны случайности и хаотичности, так и со стороны адаптивных, даже направленных процессов (см. рис. 9–5).

Говоря о случайности, энтропия (шум) на всех уровнях биологической передачи информации может быть конструктивным фактором эволюции, в значительной степени из-за устойчивости биологических сетей. В какой степени такая устойчивость является эволюционно возникшим, адаптивным качеством, а не фундаментальным свойством сетей, это глубокий, интересный вопрос, который еще предстоит тщательно расследовать. Важно отметить, что, хотя никто до сих пор не открыл прямого пути от фенотипических мутаций в геном, фенотипический шум также оказывается потенциально важным фактором эволюции благодаря опережающему эффекту, а также специальным механизмам усиления эволюционируемости, которые действуют посредством фенотипических мутаций, как, возможно, происходит в случае прионов грибов [\[98\]](#).

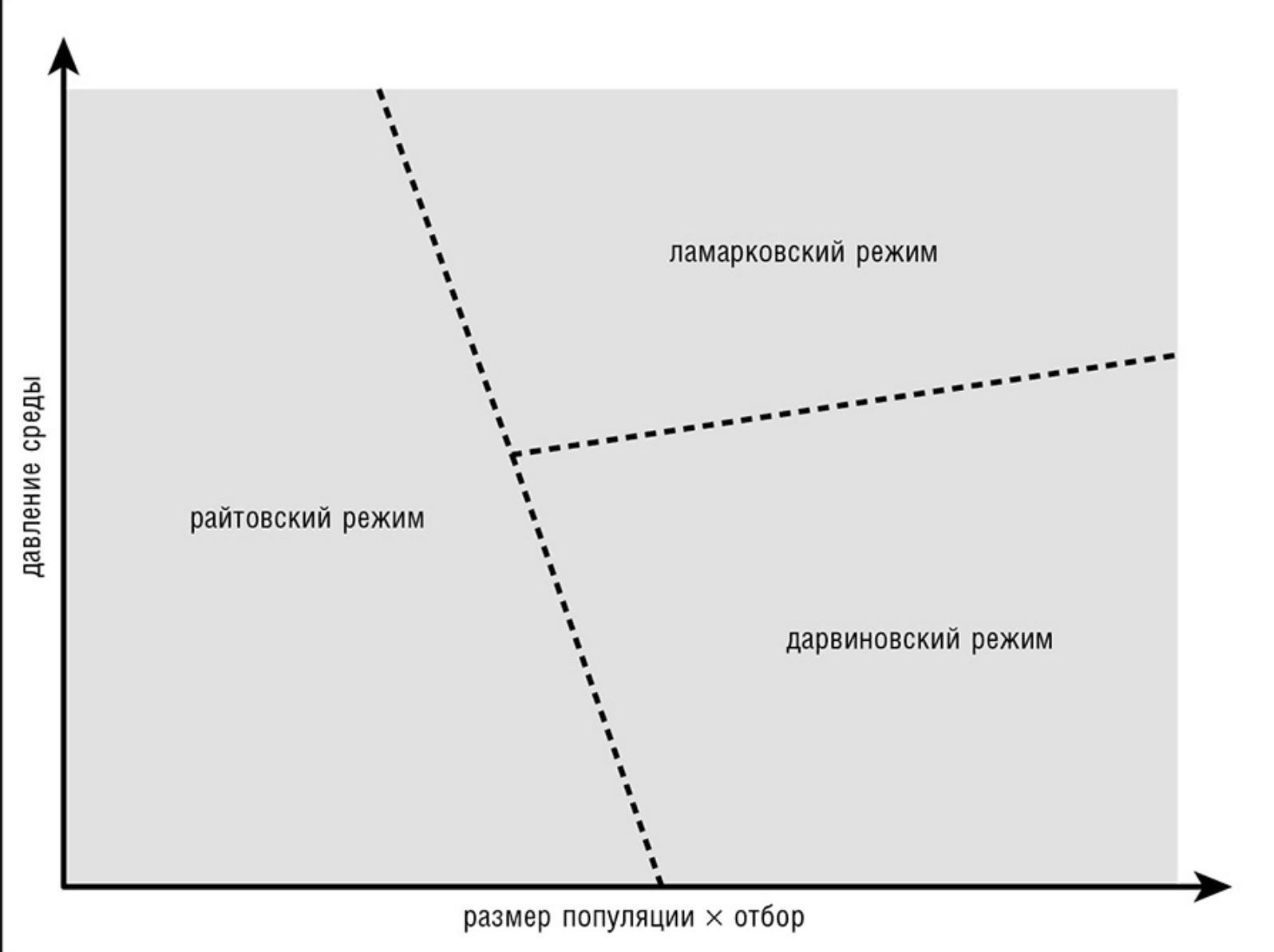


Рис. 9–5. Структура эволюционного процесса: многофакторное представление.

В другом плане многочисленные явления эпигенетической наследственности, как те, что связаны с РНК-интерференцией (см. выше), так и лучше изученные, основанные на наследуемых закономерностях метилирования ДНК, являются важными механизмами эволюции (Johnson and Tricker, 2010; Richards, 2006). Отчасти эпигенетические явления (которые мы не имеем возможности обсуждать здесь подробно) играют ту же роль, что и опережающий эффект фенотипических мутаций: они образуют буфер пластичности, что дает популяции шанс пересечь глубокие долины в неровном адаптивном ландшафте.

Что касается «необходимости», тщательное изучение различных, повсеместно идущих процессов, которые способствуют возникновению геномных вариаций, показывает, что эволюция не полагается всецело на стохастические мутации. Напротив, изменчивость часто управляется сложными молекулярными механизмами, которые инициируют адаптивную реакцию на вызовы окружающей среды разной степени специфичности. Геномная эволюция, как выясняется, охватывает весь спектр сценариев, от чисто дарвиновского, основанного на случайных изменениях, до истинно ламарковского, в котором конкретный механизм ответа на стимул фиксируется в эволюционирующей популяции через специфическое изменение в геноме. В широком смысле все эти пути изменения генома отражают взаимодействие между эволюционирующей популяцией и окружающей средой, где активная роль принадлежит либо

только отбору (чисто дарвиновский сценарий), либо направленной изменчивости, которая сама может стать объектом отбора (ламарковский сценарий).

Механизмы эволюции сами являются объектом отбора и эволюционируют: *способность эволюционировать тоже эволюционирует*. Многие биологи-эволюционисты отнеслись бы с беспокойством к такому заявлению, поскольку оно может трактоваться как принятие идеи «эволюционного предвидения». Тем не менее, невзирая на эти опасения, обширные исследования стресс-индуцированного мутагенеза и появляющееся осознание потенциально ключевой роли специализированных приспособлений, АПГ, в процессе горизонтального переноса генов не оставляют сомнений в том, что эволюционный потенциал организмов сам является предметом отбора и эволюционирует. Эволюция эволюционируемости непосредственно наблюдается в лабораторных экспериментах с эволюционирующими бактериальными популяциями. Повторю еще раз: *эволюция имеет возможность экстраполировать, отталкиваясь от повторяющихся событий прошлого, и эффективно прогнозировать общие черты будущего*.

В завершение этой главы стоит подчеркнуть, что новые пути эволюции, обсуждаемые здесь, не требуют никаких неизвестных фундаментальных механизмов. Таким образом, ни один из этих ранее недооцененных или прямо отрицаемых эволюционных феноменов не идет вразрез с основными принципами молекулярной биологии, в частности центральной догмой Крика, которая провозглашает необратимость передачи информации от нуклеиновой кислоты к белку. Например, CRISPR-система, которая, как представляется на первый взгляд, воплощает лamarковский сценарий эволюции и тем самым нарушает основные табу, действует через комбинацию молекулярных механизмов, которые в принципе универсальны и хорошо известны, даже если детали могут быть уникальными для данной системы. Эти механизмы включают в себя различные дополнительные взаимодействия между нуклеиновыми кислотами, интеграцию фрагментов ДНК в специфические локусы генома, а также узнавание и расщепление различных структур РНК ферментными комплексами – уникальные приспособления, которые эволюционный процесс «насобирал» из обычных компонентов.

Рекомендуемая дополнительная литература

- Draghi J. A., T. L. Parsons, G. P. Wagner, and J. B. Plotkin.*(2010) Mutational Robustness Can Facilitate Adaptation. *Nature*463 (7,279): 353–355.
- Эта работа показывает, «используя общую модель популяционной генетики, что мутационная устойчивость может препятствовать либо способствовать адаптации в зависимости от размера популяции, частоты мутаций и структуры ландшафта приспособленности. В частности, нейтральное разнообразие в устойчивой популяции может ускорять адаптацию, пока число фенотипов, доступных индивиду посредством мутации, меньше, чем общее число фенотипов в адаптивном ландшафте».
- Galhardo R. S., P. J. Hastings, and S. M. Rosenberg.*(2007) Mutation As a Stress Response and the Regulation of Evolvability. *Critical Reviews in Biochemistry and Molecular Biology*42: 399–435.
- Стресс-индуцированный мутагенез интерпретируется как система регуляции способности к эволюционированию, которая облегчает адаптацию и выживание.
- Koonin E. V., and Y. I. Wolf.* (2009) Is Evolution Darwinian or/and Lamarckian? *Biology Direct*4: 42.
- Обсуждение эволюционных явлений, в которых, по-видимому, действуют ламарковские или квазиламарковские механизмы.
- Levy S. F., and M. L. Siegal.*(2008) Network Hubs Buffer Environmental Variation in *Saccharomyces Cerevisiae*. *PLoS Biology*6: e264.
- Новаторское экспериментальное исследование емкости, показывающее, что многочисленные центральные узлы (hubs) молекулярных сетей обладают свойствами эволюционных конденсаторов.
- Lynch M.*(2010) Evolution of the Mutation Rate. *Trends in Genetics*26: 345–352.
- Обзор экспериментально определенных частот мутаций в полном спектре организмов, выявивший парадоксальную зависимость между частотой мутаций и размером генома у эукариот.
- Marraffini L. A., and E. J. Sontheimer.*(2010) CRISPR Interference: RNA-Directed Adaptive Immunity in Bacteria and Archaea. *Nature Reviews Genetics*11: 181–190.
- Обзор молекулярных механизмов системы CRISPR-Cas.
- Masel J., and M. V. Trotter.*(2010) Robustness and Evolvability. *Trends in Genetics*26: 406–414.
- Название «Устойчивость и способность к эволюционированию» говорит само за себя.
- Rajon, E., and J. Masel.*(2011) Evolution of Molecular Error Rates and the Consequences for Evolvability. *Proceedings of the National Academy of Sciences USA*108: 1,082—1,087.
- Важное исследование, в котором проводится различие между локальными адаптациями, направленными на уменьшение эффекта геномных и фенотипических мутаций (эволюция в сторону увеличения устойчивости), и глобальными адаптациями (эволюция в сторону снижения частоты мутаций). Моделирование в рамках популяционной генетики показывает, что локальные адаптации реалистичны только в больших популяциях с интенсивным отбором, в то время как небольшие популяции развивают глобальные адаптации.
- Wagner A.*(2008) Neutralism and Selectionism: A Network-Based Reconciliation. *Nature Reviews Genetics*9: 965–974.
- Важная работа, описывающая эволюцию (почти) нейтральных сетей, которые содержат резервуар потенциально адаптивных модификаций.
- Whitehead D. J., C. O. Wilke D. Vernazobres, and E. BornbergBauer.* (2008) The Look-ahead Effect of Phenotypic Mutations. *Biology Direct*3: 18.

Концептуально важное модельное исследование, демонстрирующее потенциальную эволюционную значимость фенотипических мутаций.

Глава 10. Мир вирусов и его эволюция

Вирусы были открыты как нечто совсем непримечательное, а именно необычная разновидность инфекционных агентов, а возможно, и особый род токсинов, вызывающих болезни растений, например табачную мозаику. Так как эти агенты проходили сквозь тонкие фильтры, задерживающие бактерии, было сделано верное предположение, что они отличаются от (типичных) бактерий. Вскоре после этого были открыты первые вирусы, поражающие животных. В их числе – вирус саркомы Рауса, первый известный вирус с канцерогенными свойствами, были открыты и удивительные патогены, которые, казалось, пожирали бактерии – их назвали бактериофагами, а в итоге они оказались бактериальными вирусами. В дальнейшем, в течение XX столетия, вирусологию ожидало блистательное развитие (Fields et al., 2001) – по двум причинам. Во-первых, вирусы важны для медицины и сельского хозяйства. Во-вторых, вирусы – простейшие генетические системы и потому стали излюбленными моделями, сначала для ранней молекулярной генетики (прежде всего благодаря работам знаменитой «фаговой группы» под руководством Макса Дельбрюка (Caigns, 1966), а затем для геномики. Однако к 1970-м годам генетика, а к концу 1990-х и геномика достаточно окрепли, чтобы продуктивно работать и с клеточными моделями [\[99\]](#). В результате вирусология потеряла ведущую роль в фундаментальной биологии (появляясь, впрочем, в эпизодах).

Первая декада нового тысячелетия была отмечена настоящим возрождением вирусологических исследований, вызванным двумя группами открытий. Первым было обнаружение гигантских вирусов, таких как мимивирус. Их вирусные частицы и геномы достигают размеров клеток, размывая казавшуюся очевидной границу между вирусами и клетками (Raoult et al., 2004; Van Etten et al., 2010). Вторая, еще более замечательная серия открытий, была сделана в рамках метагеномики: к великому удивлению биологов, оказалось, что вирусы – наиболее распространенные биологические объекты на Земле (Edwards and Rohwer, 2005). Эти достижения стимулировали намного более широкий интерес к эволюции вирусов. Я рассматриваю эти результаты как открытие обширного древнего мира вирусов, который является неотъемлемой и основополагающей частью живого с момента его зарождения на Земле. Все это время мир вирусов интенсивно взаимодействовал с клеточными формами жизни, выработавшими колоссальное многообразие систем противовирусной защиты, но сохранил свою обособленность и, во многих отношениях, остается ключом ко всей истории жизни. В этой главе мы обсудим вирусный мир, его развитие, а также гонку вооружений между вирусами и клетками, которая пронизывает всю эволюцию [\[100\]](#). С точки зрения автора, вирусы представляют собой одну из «империй жизни», другая же, разумеется, представлена клеточными организмами. В главе 11 обсуждается вклад вирусов в возникновение и эволюцию клеток.

Необыкновенное разнообразие и повсеместное распространение вирусов

Что такое вирус?

Определения в биологии даются трудно и никогда не являются достаточно полными. Тем не менее, прежде чем рассматривать в этой главе различные аспекты эволюции вирусов, их надо как-то определить. Действительно, нетрудно дать грубое общее определение. За последнее столетие знания о вирусах развились от туманных предположений их первооткрывателей, Дмитрия Ивановского и Мартина Бейеринка, до тончайших молекулярных деталей. Здесь мы определим вирусы следующим, очень общим образом: *«облигатные внутриклеточные паразиты или симбионты, обладающие собственными геномами, кодирующими информацию, необходимую для вирусной репродукции (отсюда некоторая степень автономии от генетических систем хозяина), но не кодирующими всю систему трансляции и мембранного аппарата»*. Это определение подойдет ко всякому «по-настоящему» эгоистичному генетическому элементу. Ключевая фраза тут *«кодирующие информацию, необходимую для вирусной репродукции, а следовательно, обладающие некоторой степенью автономии от генетических систем хозяина»*. Таким образом, обычные гены и опероны под это определение не подойдут, даже если обладают некоторыми эгоистическими чертами, потому что не кодируют никаких специальных приспособлений для самовоспроизведения. Таким образом, это определение охватывает все обширное множество биологических объектов, кодирующих «нечто», необходимое для самовоспроизведения, но не систему трансляции и мембраны.

Я специально не указываю, что это «приспособление для самовоспроизведения» обязано быть белком, так что определенно подойдут и вириоды (растительные патогены с геномами размером всего около трехсот нуклеотидов, использующие для репликации транскрипционный аппарат хозяина). Также я не указываю, что вирусный геном обязательно должен кодировать капсид (то есть белковую оболочку вириона). Это могло бы показаться несколько контринтуитивным, поскольку исторически вирусы были известны в основном как частицы (вирионы), начиная с первой удачной кристаллизации вируса табачной мозаики Уэнделлом Стенли в 1934 году. Капсид настолько привлекает внимание, что Патрик Фортер и Дидье Рауль недавно определили вирусы как «капсид-кодирующие организмы», в противоположность клеточным формам жизни, определенным как «рибосомокодирующие организмы» (Raoult and Forterre, 2008). Это определение указывает правильное направление, если требуется разграничить клетки и вирусы как две основные формы жизни, но оно безосновательно узко и не очерчивает границы мира вирусов объективно.

Конечно, наличие капсида – очень важная и часто встречающаяся черта. Однако в этой главе мы рассмотрим четкие эволюционные взаимосвязи, сопровождаемые сходством геномной архитектуры и циклов репликации, существующие между традиционными, капсид-кодирующими вирусами и «голыми» эгоистичными элементами, такими как плазмиды и разнообразные мобильные элементы. В рамках данного определения (хотя, как и все определения, оно имеет границы применимости) все эти агенты принадлежат обширному миру вирусов.

Время от времени вспыхивают дискуссии на занимательную тему – являются ли вирусы «живыми». Последняя версия этого спора привлекла существенное внимание (Moreira and Lopez-Garcia, 2009). Сам по себе этот вопрос носит исключительно семантический характер и, следовательно, не является значимым. Приведенное здесь определение ясно отводит вирусам

место во владениях биологии; как будет изложено далее в этой главе, сравнительная геномика выявила множественные связи между геномами вирусов и клеточных форм жизни. Неудачным последствием отрицания статуса «живых» за вирусами будет то, что тогда вирусы не будут иметь очевидного отношения к эволюции клеточных форм жизни. В этой и следующей главах мы увидим, что это противоречило бы действительности.

Разнообразие стратегий репликации-экспрессии среди вирусов

Все клеточные формы жизни обладают геномами, представленными дцДНК, которая транскрибируется в мРНК (транслирующиеся во многочисленные белки), а также в разнообразные некодирующие РНК. Единообразие генетического цикла клеточных форм жизни находится в разительном контрасте с вариабельностью циклов репликации-экспрессии у вирусов, некоторые из которых обладают РНК-геномами разной полярности, в то время как другие обладают геномами в виде оцДНК (Baltimore, 1971; см. рис. 10-1 [\[101\]](#)). Некоторые вирусы и вирусоподобные элементы освоили переход от РНК к ДНК, сочетая действие кодируемой вирусом обратной транскриптазы (ОТ) и ДНК-зависимой РНК-полимеразы хозяина. Вирусы с плюс-цепью РНК отличаются использованием простейшего воображимого генетического цикла, в то время как обратно транскрибирующиеся элементы связывают мир РНК и мир ДНК. Такая пластичность циклов вирусной репликации может иметь важные эволюционные следствия, о чем будет рассказано в следующей главе.

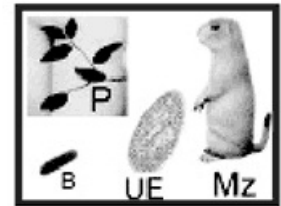
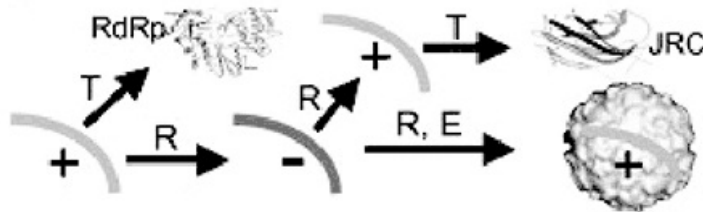
Когда разнообразие вирусных стратегий репликации-экспрессии стало очевидным, возникло искушение «сыграть в Менделеева» и создать исчерпывающую таблицу возможных циклов репликации-экспрессии, заполнить ее наблюдаемыми разновидностями, а затем попытаться предсказать, какие клетки таблицы заполнятся благодаря будущим открытиям, а какие окажутся «запрещенными» по каким-либо фундаментальным причинам. Насколько мне известно, первую такую попытку предпринял мой учитель вирусологии, Вадим Израилевич Агол (Agol, 1974, первоначальная публикация вышла в мало кому теперь известном русском научном журнале, в котором я ее и прочел). Именно эта статья, благодаря ее захватывающей ясности и ценнейшей попытке (по крайней мере, такой воспринимал ее я в то время, между первым и вторым курсом университета) воспользоваться глубокими, пусть даже и простыми, соображениями симметрии в биологии, прежде всего и побудила меня изучать вирусы. Я никогда не сожалел об этом решении; спустя годы я выработал собственную версию классификации стратегий геномов (Koopin, 1991) [\[102\]](#). Помимо незыблемой центральной догмы, кажется, есть один *фундаментальный запрет: одноцепочечная ДНК никогда не транслируется, поэтому РНК всегда вовлечена в цикл самовоспроизводства любого генетического элемента*. В отличие от случая с белками в рамках центральной догмы, здесь, по-видимому, нет прямого химического обоснования для такого запрета (и на самом деле, экспериментально трансляция оцДНК была продемонстрирована Hulen and Legault-Demare, 1975; McCarthy and Holland, 1965). Однако единственная известная нам система трансляции, очевидно, эволюционировала в направлении синтеза белков лишь на матрице РНК (подробнее см. гл. 12). Помимо этого исключения, все возможные на базе молекул РНК и ДНК циклы репликации-экспрессии, по-видимому, реализованы в мире вирусов, хотя некоторые экзотические формы генома, такие как гибрид РНК-ДНК, встречаются редко (см. рис. 10-1).

КЛАСС

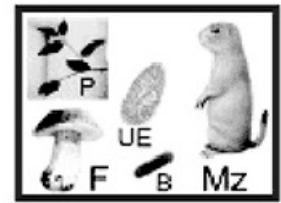
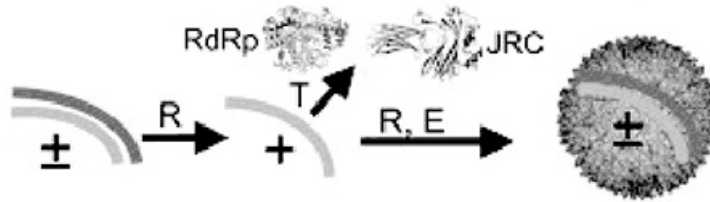
ЦИКЛ РЕПЛИКАЦИИ

ХОЗЯЕВА

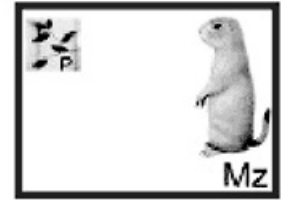
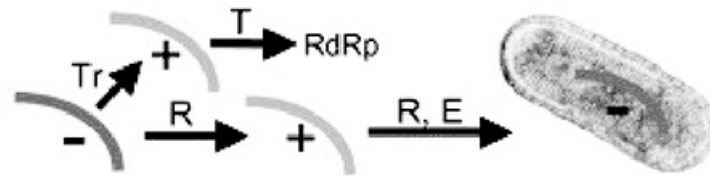
(+)РНК, 3—30 Кб



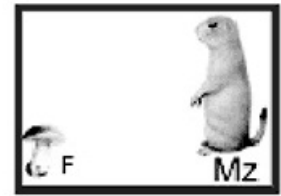
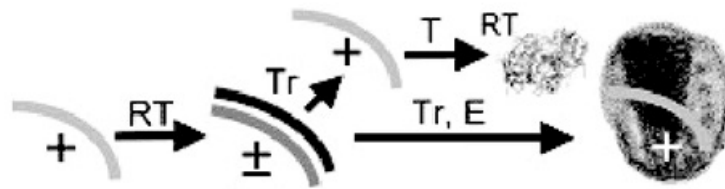
дцРНК, 4—25 Кб



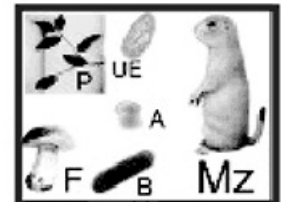
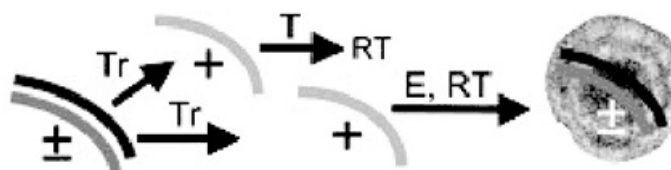
(-)РНК, 11—20 Кб



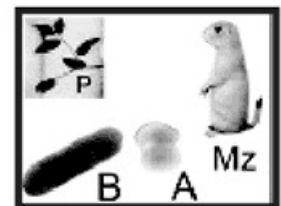
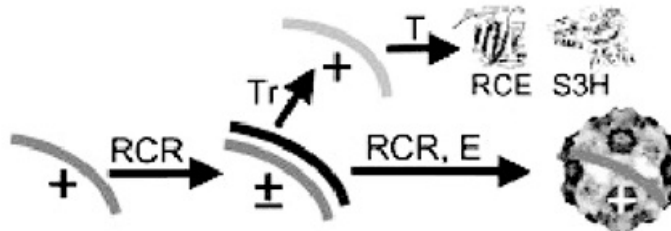
ретроидные РНК, 7—12 Кб



ретроидные ДНК-вирусы
и элементы, 2—10 Кб



оцДНК-вирусы,
плазмиды, 2—11 Кб



дцДНК-вирусы,
плазмиды, 5—1200 Кб

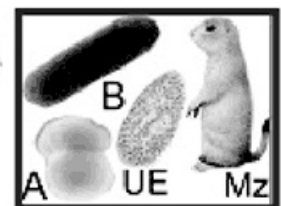
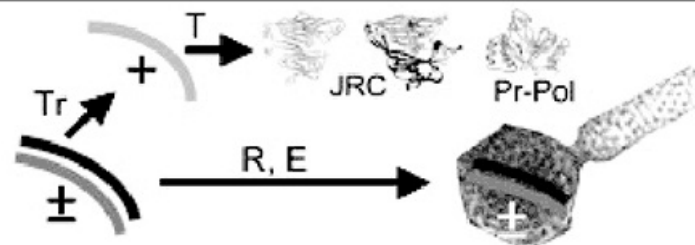


Рис. 10-1. Разнообразие стратегий репликации-экспрессии у вирусов и вирусоподобных элементов. Для каждого класса вирусов и схожих элементов указаны приблизительные размеры генома (Кб, в тысячах оснований). «+» обозначает плюс-цепь (однонаправленную с мРНК), «—» обозначает минус-цепь. Tr – транскрипция, T – трансляция, R – репликация, E – капсидирование вируса, А – археи, В – бактерии, F – грибы, Mz – животные, P – растения, UE – одноклеточные эукариоты. Для каждого класса вирусов (элементов) показаны типичные структуры их белков-маркеров, а также ультрамикроскопические изображения характерных представителей. RdRp –

РНК-зависимая РНК-полимераза; JRC – белок капсида с укладкой типа рулета; RT – обратная транскриптаза; RCR – (инициирующая) эндонуклеаза репликации способом катящегося кольца. Справа показан диапазон возможных хозяев. Размер соответствующего изображения и сокращения приблизительно соответствует представленности данного класса вирусов в этом таксоне (по Koopin et al., 2006, с изменениями).

Диапазон сложности геномов, их функционального содержания и разнообразие геномной архитектуры вирусов

Диапазон размера геномов вирусов с различными геномными стратегиями весьма велик: размеры геномов самого крупного известного вируса – мимивируса – и самых мелких вирусов (таких как цирковирусы) отличаются на три порядка. Если мы включим сюда вироиды, которые не кодируют белки, но являются полноценными эгоистичными генетическими элементами и даже патогенами, размах колебаний расширяется до четырех порядков (см. рис. 10-1). Если иметь в виду, что в геномах вирусов обычно белок-кодирующие гены расположены вплотную, то колебания количества генов укладываются приблизительно в тот же диапазон. Размер генома сильно зависит от природы генома и цикла репликации-экспрессии. Видимо, лишь у вирусов, содержащих дцДНК, геном может достичь больших (по вирусным меркам) размеров, более 35 Кб, и (на данный момент) вплоть до 1,1 Мб (Van Etten et al., 2010). Все классы РНК-вирусов, все элементы, способные к обратной транскрипции, и все вирусы, содержащие оцДНК, обладают небольшими геномами, никогда не превышающими 35 Кб, – и то к этому пределу приближается лишь одна группа довольно редких РНК-вирусов животных (коронавирусы и родственные им формы, составляющие порядок *Nidovirales* (Gorbalenya et al., 2006). Причина этого очевидна: большая химическая стабильность и регулярная структура дцДНК благоприятна для функций хранения информации и репликации, а далее эволюция систем репарации дцДНК еще больше усилила функциональное разделение между дцДНК и другими формами нуклеиновых кислот. В дополнение к разнообразию стратегий репликации-экспрессии и широкому диапазону размеров, вирусные геномы принимают все возможные молекулярные конфигурации, включая как линейные, так и циркулярные молекулы ДНК или РНК и как единичные, так и разделенные на множественные сегменты геномы (хромосомы, см. рис. 10-1).

Функциональный набор вирусных генов разительно различается в зависимости от цикла экспрессии и, что еще важнее, от размера генома и генетической сложности. Малые вирусные геномы кодируют практически исключительно белки, непосредственно участвующие в репликации генома, а также субъединицы вириона. Часто – и для всех известных РНК-вирусов и элементов, способных к обратной транскрипции, это именно так, – вирус кодирует полимеразу, участвующую в репликации его собственного генома. Это легко объяснить: клеточные хозяева вируса обычно не кодируют РНК-зависимой РНК-полимеразы или обратной транскриптазы, которая могла бы осуществить репликацию или обратную транскрипцию длинных молекул РНК. РНК-зависимые РНК-полимеразы и обратные транскриптазы, которые кодируются в геномах клеточных форм жизни и осуществляют «обычные» клеточные функции, такие как теломераза и РНК-зависимая РНК-полимераза, участвующая в РНК-интерференции эукариот, способны синтезировать лишь короткие олигонуклеотиды (см. гл. 7 и обсуждение ниже в этой главе). У ДНК-вирусов, напротив, есть возможность использовать аппарат репликации (и транскрипции) хозяина, и они широко пользуются этой возможностью. Так, многие вирусы этого типа, в частности большинство известных вирусов, поражающих архей, так же как и многочисленные умеренные бактериофаги (такие как фаг лямбда, классический модельный организм), не кодируют даже репликативной полимеразы и никаких других белков, непосредственно участвующих в репликации. В таких случаях последовательности ДНК, ответственные за распознавание и привлечение репликативного аппарата хозяина, оказываются основными факторами, определяющими возможность автономной репродукции вируса (см. определение выше), хотя вирусные белки выполняют другие важные функции в вирусной репродукции, такие как подавление или репрограммирование экспрессии генов и метаболизма

хозяина. Вирусы с самыми большими геномами, напротив, вдобавок к белкам, составляющим аппарат репликации вирусного генома, кодируют целую коллекцию разнообразных белков, участвующих в процессах репарации, мембранного транспорта, ряде метаболических путей, а в некоторых случаях и трансляции. Обычно (и во всех случаях, когда затрагивается трансляция) вирус кодирует не систему или путь целиком, а только ферменты для одного-двух шагов, которые дополняют или видоизменяют соответствующие функциональные системы клетки хозяина.

Вирусы, содержащие различные формы нуклеиновых кислот, распределены по таксонам хозяев не равномерным и не случайным образом. В частности, необычайное разнообразие содержащих дцДНК бактериофагов и вирусов архей находится в резком контрасте с отсутствием дцДНК-вирусов у растений. РНК-вирусы, напротив, чрезвычайно распространены и разнообразны у растений и животных, но среди бактерий на данный момент представлены лишь двумя небольшими семействами и до сих пор не обнаружены у архей [\[103\]](#) (см. рис. 10-1). В некоторых случаях биологическая подоплека характерного спектра хозяев вируса совершенно ясна. Например, в растениях крупные вирусы столкнулись бы с серьезными проблемами с межклеточным распространением, так как плазмодесмы (каналы, соединяющие окруженные клеточной стенкой клетки растений) непроницаемы для больших частиц и даже больших молекул ДНК. Однако в большинстве случаев причины предпочтительного распределения вирусов среди тех или иных групп хозяев остаются неизвестными. Например, сложно сказать, почему РНК-вирусы столь распространены среди растений и животных, но не среди прокариот; далее в этой главе мы вернемся к рассмотрению этого вопроса с другой, эволюционной позиции.

Метагеномика вирусов, экспериментальная вирусология, агенты переноса генов и повсеместное распространение вирусов

Вирусы – вездесущие спутники клеточных форм жизни: при более-менее детальном изучении любой клеточный организм оказывается населенным вирусами. У тех организмов – таких как нематоды, – у которых настоящие вирусы пока не были открыты, в геноме встроены многочисленные мобильные элементы [104].

С недавних пор изучение виромов (всего множества вирусов, обнаруживаемых в данной среде обитания) стало областью науки, испытывающей бурное развитие (Edwards and Rohwer, 2005; Kristensen et al., 2010). Выделение вирома – методически довольно незамысловатая операция, так как вирусные частицы (по крайней мере подавляющее большинство, исключая, возможно, их гигантские разновидности) проходят сквозь фильтры, непроницаемые даже для мельчайших клеток. Так что достаточно просто собрать частицы в фильтрате и анализировать состав. Изучение виромов принесло большие неожиданности. Первой оказалась сама концентрация вирусных частиц. Поразительно, но по крайней мере в морской среде вирусы (прежде всего бактериофаги) оказались наиболее распространенной из биологических форм, причем в совокупности количество вирусных частиц превосходит количество клеток по меньшей мере на порядок. Это не вполне корректный способ сравнения, так как одна инфицированная вирусом клетка способна продуцировать сотни вирусных частиц, и тем не менее такие данные показывают, что вирусы исключительно широко распространены и активны в окружающей среде. Сейчас вирусы рассматривают в качестве крупных геохимических действующих сил, так как уничтожение вирусами микроорганизмов оказывает сильное влияние на формирование органического осадка (Suttle, 2007, 2005). Вторая крупная неожиданность – это огромное генетическое разнообразие виромов, а также их непредвиденная генетическая композиция. Набор генов в ДНК-виrome резко отличается от набора генов известных бактериофагов. Виром в основном состоит из редких и уникальных генов, для которых не удается найти и подобрать гомологичных последовательностей в имеющихся базах данных. Хотя гены, специфичные для бактериофагов, встречаются в виромах чаще, чем в микробиомах, они представляют ничтожное меньшинство; большинство генов, для которых обнаружены гомологи, представляются случайным набором бактериальных генов. Исключив возможность массивной контаминации, которая, учитывая тщательно разработанные протоколы, используемые для выделения виромов, маловероятна, мы приходим к выводу, что в виромах в основном представлены не обычные вирусы, а некие другие биологические объекты (Kristensen et al., 2010).

Какова природа этой «темной материи», которая доминирует в виромах? На самом деле мы этого не знаем, однако легко предложить правдоподобную гипотезу. Вспомним агенты переноса генов (АПГ), которые уже обсуждались в главах 5 и 9. Агенты переноса генов – особая разновидность псевдовирусов (Lang and Beatty, 2007). Они образуют вирусные (фагоподобные) частицы, которые состоят из белков, кодируемых дефектным профагом, находящимся в соответствующей бактериальной или архейной хромосоме. Однако частицы АПГ не содержат ДНК профага (так что они не являются настоящими бактериофагами), но вместо этого служат оболочками для, судя по всему, случайных фрагментов бактериальной хромосомы. Легко представить, что *темная материя виромов в основном состоит из АПГ* (Kristensen et al., 2010), так что по сути дела это всего лишь «псевдовиромеры». Эта простая гипотеза, которая, конечно, нуждается в эмпирической проверке, имеет далеко идущие следствия. В самом деле, если вирусные (или вирусоподобные) частицы – самые распространенные биологические объекты на

Земле и большая часть из них – агенты переноса генов, тогда неизбежен логический вывод, что в биосфере доминируют АПГ. В сочетании с присутствием многочисленных, часто «скрытых» профагов и других мобильных элементов в бактериальных и архейных геномах (Cortez et al., 2009) и еще большим изобилием (чаще неактивных) эгоистичных элементов в геномах многих эукариот (включая человека), эти наблюдения наводят на мысль, что мир вирусов в большой степени «выстраивает» геномы клеточных форм жизни и таким образом определяет эволюцию жизни в целом. Необходимо еще тщательно продумать и изучить в дальнейшем фундаментальные следствия такого вывода, и мы не раз вернемся к ним в этой и следующих главах.

Хотя метагеномика вирусов – еще молодая область науки, она уже замечательным образом преобразила наше понимание мира вирусов, причем это далеко не сводится к удивительным данным по генетическому составу вирионов. Одна группа открытий была сделана в результате анализа последовательностей, полученных в ходе глобального исследования океана, грандиозного метагеномного проекта Крейга Вентера (Yoosaph et al., 2007). Оказывается, что помимо несметного числа бактериальных последовательностей (основной цели проекта) база данных глобального исследования океана содержит многочисленные последовательности, гомологичные консервативным генам больших нуклеоцитоплазматических ДНК-вирусов (NCLDV, подробнее о них позже в этой главе), заражающих эукариот. Независимо от конкретного источника вирусной ДНК (гигантские вирусы, зараженные пикозеукариоты, проходящие бактериальные фильтры, или, вероятнее всего, и те и другие), разнообразие нескольких семейств NCLDV вышло далеко за границы, очерченные традиционной вирусологией (Monier et al., 2008). Вторая группа замечательных открытий сделана в метагеномике морских РНК-вирусов: были открыты многочисленные РНК-вирусы, заражающие одноклеточных морских эукариот, и весьма неожиданно оказалось, что все они принадлежат к одному суперсемейству вирусов, ранее обнаруженных у животных и растений, – пикорнаподобным вирусам (Koonin et al., 2008). В совокупности эти открытия показывают, что наши представления о мире вирусов пока крайне поверхностны – быть может, истинные размеры и разнообразие этого мира превзойдут самые смелые фантазии [\[105\]](#).

Эволюция вирусов: полифилия, монофилия и гены-сигнатуры

В предыдущих разделах этой главы было дано представление о мире вирусов и показано, что он соизмерим по масштабу с миром клеточных форм жизни – и, возможно, количественно доминирует в биосфере. Более того, вирусный мир оказывает определяющее воздействие на эволюцию клеток с помощью агентов переноса генов и огромного арсенала мобильных элементов. Так что изучать эволюцию вирусов необходимо, если мы стремимся к хоть сколько-нибудь глубокому пониманию эволюции жизни в целом.

Сравнительная геномика не свидетельствует в пользу монофилетического происхождения всех вирусов. Здесь мы понимаем под «монофилией» «*происхождение от общего предкового вируса или вирусоподобного эгоистического элемента*» (Koonin et al., 2006). У многих групп вирусов просто нет общих генов, что позволяет решительно отменить любые идеи об общем происхождении. Идея «общих генов», будучи приложена к вирусам, оказывается непростой: в мире вирусов общность не обязательно ограничивается однозначными отношениями ортологии между генами, которые легко выявить благодаря очевидному сходству последовательностей. Вместо этого, как будет изложено в следующих разделах, отдаленные отношения гомологии между вирусными белками, а также между вирусными белками и их гомологами из клеточных форм жизни могут дать более сложные, но важные свидетельства об эволюции вирусов. Несмотря на эту сложность, среди больших групп вирусов изобилуют случаи, когда между ними либо нет никаких гомологичных генов, какое бы определение мы ни взяли, либо имеются лишь отдаленно родственные домены, чьи эволюционные траектории очевидно различны. Например, большинство вирусов гипертермофильных кренархеот не имеют общих генов ни с какими другими вирусами (Prangishvili et al., 2006b), а РНК-вирусы разделяют с ДНК-вирусами и плазмидами, реплицирующимися способом «катящегося кольца», лишь чрезвычайно отдаленно родственные домены в соответствующих белках репликации.

С другой стороны, можно уверенно показать монофилию нескольких крупных групп вирусов, включая обширные совокупности РНК-вирусов и сложных ДНК-вирусов (см. табл. 10-1). Некоторые из этих монофилетических классов переходят даже границы геномных стратегий: так, монофилетический класс элементов, размножающихся через стадию обратной транскрипции, включает в себя как РНК-вирусы, так и вирусы, мобильные элементы и плазмиды с ДНК-геномами. В классе элементов, реплицирующихся по механизму катящегося кольца, сочетаются вирусы, содержащие одноцепочечную или двухцепочечную ДНК, а также плазмиды. Более того, на основании сходства в структурах комплексов, участвующих в репликации РНК, а также присутствия гомологии, пусть и отдаленной, в ферментах репликации была предложена правдоподобная гипотеза об общем происхождении вирусов, содержащих позитивную РНК-нить, вирусов, содержащих дцРНК, и элементов, способных к обратной транскрипции (Ahlgquist, 2006). Однако в целом вывода, что вирусы имеют много различных линий происхождения, по-видимому, невозможно избежать (Koonin et al., 2006; см. табл. 10-1).

Краткая естественная история вирусных генов

Анализ последовательностей выявил несколько категорий вирусных генов, принципиально отличающихся по происхождению (Koonin et al., 2006). Можно обсуждать, какая степень подробности классификации оптимальна, но четко различаются по меньшей мере пять классов, укладываемых в три более крупные категории.

Гены с четко опознаваемыми гомологами у клеточных форм жизни:

1. Гены с близкими гомологами у клеточных организмов (обычно это хозяева данного вируса) присутствуют у узкой группы вирусов.

2. Гены, консервативные среди большой группы вирусов или даже нескольких групп и имеющие относительно отдаленные клеточные гомологи.

Вирус-специфичные гены:

3. Гены-сироты [106] – гены без выявленных гомологов, кроме как у близкородственных вирусов.

4. Вирус-специфические гены, консервативные для (относительно) широкой группы вирусов, но не имеющие гомологов у клеточных форм жизни.

5. Гены – вирусные сигнатуры.

Вирусные сигнатуры – это гены, общие для многих разнообразных групп вирусов, имеющие лишь отдаленные гомологи у клеточных организмов, для которых имеются убедительные свидетельства в пользу монофилии (общего происхождения) всех вирусных членов соответствующих семейств генов. Словосочетание «гены-сигнатуры» было использовано, чтобы подчеркнуть, что эти гены являются признаками «вирусного состояния».

Относительный вклад каждого из этих классов генов в наборы генов различных вирусов зависит от размера вирусного генома и его генетической сложности, которые разнятся более чем на три порядка (см. рис. 10-1). Вирусы с малыми геномами, такие как большинство РНК-вирусов, часто обладают лишь несколькими генами, большинство из которых принадлежат к классу генов-сигнатур. Напротив, у вирусов с большими геномами, например поксвирусов, широко представлены все пять классов. Чтобы проиллюстрировать разнообразие генетического состава, на рис. 10-2 показана разбивка набора генов трех вирусов, обладающих малым, средним и крупным геномами соответственно, на пять классов генов. Интересно, что крупные и среднего размера геномы бактериофагов и вирусов архей заполнены «сиротами», которые зачастую составляют более 80 процентов генов этих вирусов. Быстро эволюционирующие «гены-сироты» фагов, возможно, поставляют много, если не большинство «генов-сирот», которые находят в геномах прокариот (несмотря на отсутствие выявляемой консервативности последовательностей), так что они играют ключевую роль в эволюции прокариот (Daubin and Ochman, 2004).

Происхождение пяти классов вирусных генов, скорее всего, различно. Два класса генов, имеющих легко распознаваемые гомологи в клеточных формах жизни, вероятно, представляют собой соответственно относительно недавние (класс 1) и древние (класс 2) заимствования из геномов клеточных хозяев. Откуда произошли вирус-специфические гены – это значительно более трудный и интересный вопрос. Первая возможность заключается в том, что эти гены произошли от других генов вирусов или хозяев, но резкое ускорение эволюции, связанное с появлением новых, специфических вирусных функций, стерло все следы их происхождения. Это соображение согласуется с тем фактом, что многие (возможно, большинство) генов 4-го класса (вирус-специфических генов, консервативных среди группы вирусов) – компоненты вириона, то есть несут важнейшую для вируса функцию, для которой нет эквивалента у клеточных форм жизни. Мы пока отложим обсуждение других путей возникновения и эволюции вирус-специфических генов и обсудим эволюцию вирусных генов-сигнатур. Гены-сигнатуры, пересекающие границы, пролегающие между самыми разными эволюционными линиями вирусов, – глубоко интересный и значимый объект для понимания эволюции и происхождения вирусов.

Таблица 10-1. Наиболее крупные монофилетические классы вирусов и эгоистичных генетических элементов.

Класс вирусов	Группы вирусов ¹	Хозяева	Доводы в пользу монофилии
РНК-вирусы с позитивным геномом	Суперсемейство I: пикорнаподобные Суперсемейство II: альфаподобные Суперсемейство III: флавиподобные Точные родственные отношения РНК-бактериофагов внутри этого класса остаются неясными (возможно, это четвертая эволюционная линия)	Животные, растения, простейшие, бактерии (одно семейство бактериофагов)	Консервативная RdRp JRC у большинства вирусов суперсемейства I и подгруппах из суперсемейств II и III Реконструирован предок, обладающий RdRp и JRC
Вирусы и мобильные элементы, размножающиеся через стадию обратной транскрипции	Ретровирусы, гепаднавирусы, каулимовирусы, баднавирусы LTR-содержащие и LTR-несодержащие ретроэлементы Ретроны Самосплайсирующиеся интроны II группы, предшественники сплайсосомных интронов эукариот	Животные, грибы, растения, простейшие, бактерии, археи	Консервативная RT
Мелкие ДНК-вирусы, плазмиды и транспозоны, реплицирующиеся по механизму «катящегося кольца»	Гемини-, цирко-, парво-, паповавирусы, фаги (например, ФХ174), плазмиды архей и бактерий, хелитроны (обширная группа эукариотических транспозонов)	Животные, растения, археи, бактерии	Консервативные RCRE, JRC, S3H (в вирусах эукариот)

Хвостатые бактериофаги (Caudovirales)	Семейства: Myoviridae (например, T4), Podoviridae (например, T7), Siphoviridae (например, λ)	Бактерии, эуриархеи	Сложные, пересекающиеся наборы генов, консервативные в подгруппах хвостатых фагов Полагают, что гены всех хвостатых фагов представляют собой единый пул
Нуклеоцитоплазматические крупные ДНК-вирусы (NCLDV)	Поксвирусы, асфарвирусы, иридовирусы, марсельвирусы фикоиднавирусы, мимивирусы	Животные, водоросли, простейшие	Во всех этих вирусах находят основной набор из 11 консервативных генов, включающий JRC, S3H, FtsK-подобную упаковывающую АТФазу Реконструирован предок приблизительно с 50 генами

Сокращения: JRC – белок капсида с укладкой типа рулета; LTR – long terminal repeat (длинный концевой повтор); RdRp – РНК-зависимая РНК-полимераза; RCRE – (инициирующая) эндонуклеаза репликации по механизму «катящегося кольца»; RT – обратная транскриптаза; S3H – геликаза суперсемейства 3.

1 Уже после выхода этой книги автору посчастливилось участвовать в работе, где из горячего источника в Йеллоустонском парке, населенного исключительно гипертермофильными кренархеотами, методами метагеномики был выделен и секвенирован РНК-геном нового вируса, очень отдаленно родственного разным семействам позитивных РНК-вирусов эукариот (Bolduc B, Shaughnessy DP, Wolf YI, Koonin EV, Roberto FF, Young M. Identification of novel positive-strand RNA viruses by metagenomic analysis of archaea-dominated Yellowstone hot springs. *J Virol.* 2012 May;86(10):5562-73). Скорее всего, этот вирус представляет новое суперсемейство позитивных РНК-вирусов, которые, возможно, размножаются в археях.

Прямые данные, подтверждающие (или отвергающие) это предположение, еще предстоит получить. Если гипотеза подтвердится, наши представления об эволюции этого класса вирусов могут потребовать существенного пересмотра (см. ниже).

Гены – вирусные сигнатуры: сигналы из древнего мира вирусов

Помимо больших классов, перечисленных в табл. 10-1, отчетливых вертикальных родственных отношений между крупными группами вирусов не существует. Однако значительное число генов, кодирующих белки, которые играют ключевую роль в репликации генома, экспрессии и морфогенезе вирионов, входят в пересекающиеся наборы генов, принадлежащие группам вирусов, казалось бы не родственных ни в каком другом отношении, хотя ни один из этих генов не присутствует у *всех* вирусов (см. табл. 10-2). Большинство генов – вирусных сигнатур не имеют высококонсервативных гомологов у клеточных форм жизни (исключения – легко распознаваемые провирусы и мобильные генетические элементы), хотя отдаленные гомологи бывают. Два наиболее широко распространенных среди вирусов гена кодируют белок капсида с так называемой укладкой типа рулета и геликазу суперсемейства 3. Оба этих белка пересекают барьер между ДНК– и РНК-вирусами и появляются в поразительно широком ассортименте групп вирусов, от некоторых самых мелких РНК-вирусов с позитивным геномом до крупных нуклеоцитоплазматических ДНК-вирусов, класса вирусов, включающего гигантский мимивирус (см. табл. 10-2). Уточняя, можно сказать, что белок капсида с укладкой типа рулета – главный строительный блок икосаэдрических (сферических) вирусных капсидов, наиболее частой формы капсидов, которые крайне различны по величине, но довольно сходны по симметрии и общей форме среди всего огромного диапазона вирусов, которые используют разнообразные стратегии репликации-экспрессии и заражают хозяев, представляющих все или почти все разнообразие клеточной жизни. Так, геликаза суперсемейства 3 участвует в репликации геномов огромного разнообразия РНК– и ДНК-вирусов.

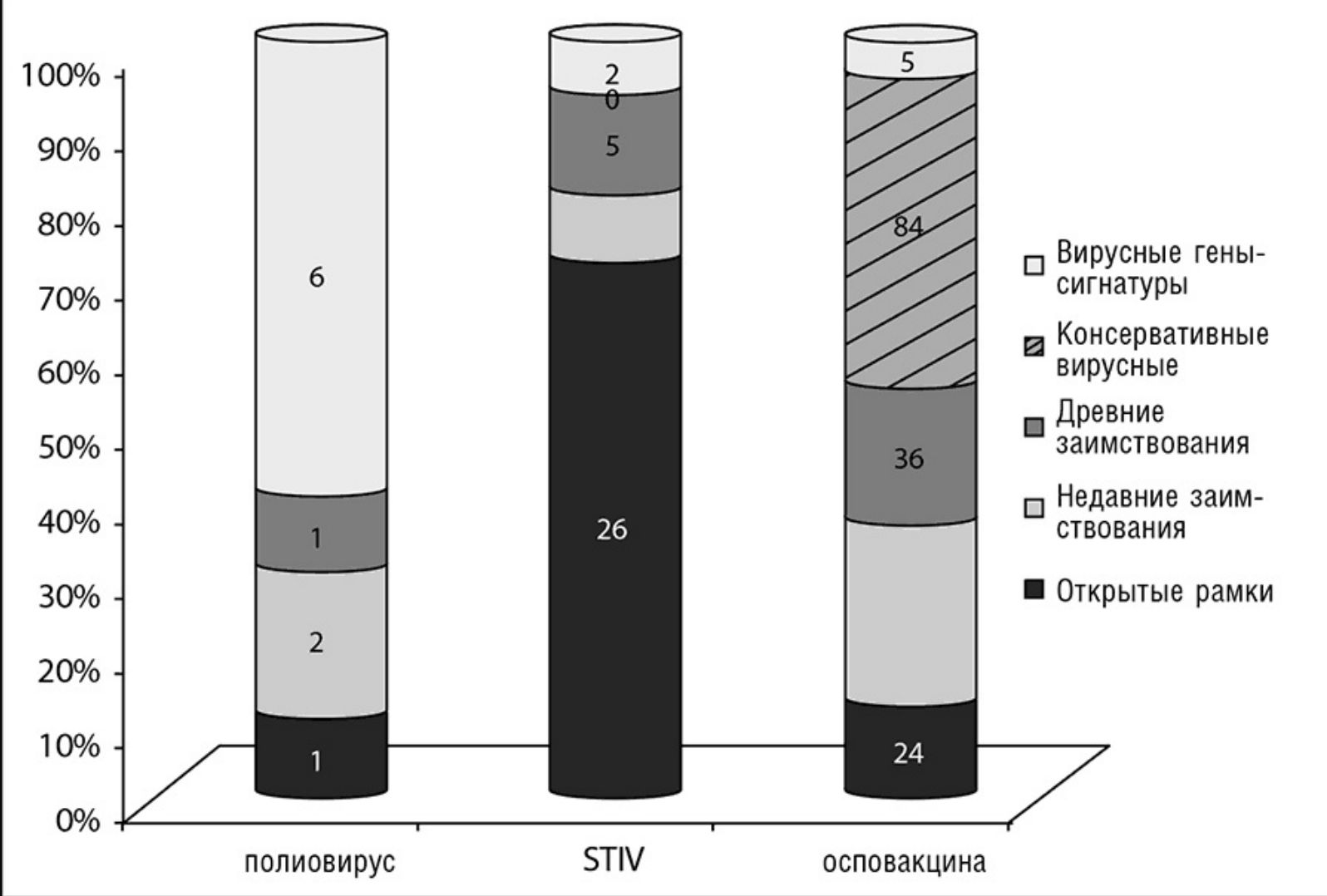


Рис. 10-2. Разбиение вирусных генов на пять эволюционных классов: вирус, обладающий малым геномом: полиовирус (7,4 Кб); вирус, обладающий геномом промежуточного размера: *Sulfolobus turreted icosahedral virus* (STIV); вирус, обладающий крупным геномом: вирус оспавакцины (195 Кб). Данные по Koonin et al., 2006.

Другие белки, перечисленные в табл. 10-2, не столь широко распространены, как белок капсида с укладкой типа рулета и геликазы суперсемейства 3, но все же за их счет формируются множественные непредвиденные связи между группами вирусов, во всех других отношениях казавшихся неродственными. В качестве примера можно рассмотреть эндонуклеазу инициации репликации по механизму «катящегося кольца», которая объединяет огромное разнообразие небольших репликонов одноцепочечных и двухцепочечных ДНК, включая вирусы, плазмиды и транспозоны, которые размножаются в животных, растениях, бактериях и археях. Детальный анализ аминокислотных последовательностей показал, что ДНК-связывающий домен репликативного белка вирусов полиомы и папилломы (например, Т-антигена SV40) представляет собой неактивную производную форму эндонуклеазы инициации репликации способом «катящегося кольца» (Iyer et al., 2005). Таким образом, посредством детального исследования одного из белков-сигнатур хорошо известная связь между различными небольшими репликонами оцДНК (вирусов и плазмид) распространилась также на группу репликонов дцДНК такого же размера. Подобное расширение набора вирусных групп, охватываемых определенным геном-сигнатурой, произошло также в результате детального анализа АТФазы, отвечающей за упаковку вирусной ДНК в капсид, и праймазы архей и эукариот, участвующей в инициации репликации ДНК (Iyer et al., 2005; Iyer et al., 2004b; см.

табл. 10-2).

Репликация генома РНК-вирусов с позитивным геномом, дцРНК-вирусов, вирусов с негативным геномом и обратно транскрибирующихся вирусов (элементов) катализируется другим классом ферментов – вирусных сигнатур: РНК-зависимыми РНК-полимеразами и обратной транскриптазой. Полимеразы РНК-вирусов с позитивным геномом и обратная транскриптаза образуют монофилетическую группу внутри обширного класса так называемых palm-доменов («домены-ладони»), характерных для различных полимераз (Iyer et al., 2005; Koonin et al., 2008). РНК-зависимые РНК-полимеразы дцРНК-вирусов и вирусов с негативным геномом, скорее всего, сильно измененные производные того же полимеразного домена (Delarue et al., 1990; Gorbalenya et al., 2002; Koonin et al., 1989). Этот ген-сигнатура может привести нас к самым ранним этапам эволюции жизни, к миру РНК (см. гл. 11 и 12 – там гораздо подробнее) и началу мира вирусов. Palm-домен, вероятно, является изначальным белком-полимеразой, сменившим рибозимные полимеразы (гипотетического) мира РНК. Это предположение поддерживается не только широким распространением palm-домена среди современных форм жизни, но еще и структурной, а следовательно, и эволюционной связью между palm-доменом и доменом, содержащим РНК-распознающий мотив (RRM, RNA Recognition Motif), древний РНК-связывающий домен, который первоначально, возможно, способствовал репликации рибозимов (Aravind et al., 2002). РНК-зависимые РНК-полимеразы и обратные транскриптазы исключены из основного хода репликативного цикла клеточных форм жизни, хотя большинство эукариотических геномов, особенно растений и животных, включает в себя множество копий RT-кодирующих ретроэлементов; у прокариот тоже есть немного подобных элементов (см. также гл. 5 и 7). Эти элементы, однако, эгоистичны и с эволюционной точки зрения принадлежат миру вирусов. Возможно, наиболее впечатляющее вторжение обратной транскриптазы в мир клеток – это каталитическая субъединица эукариотической теломеразы, важнейшего фермента, участвующего в репликации концов хромосом [\[107\]](#). Конечно, не стоит забывать, что все интроны эукариот произошли от ретроэлементов прокариот (см. гл. 7). Примечательно, что единственная другая известная РНК-зависимая РНК-полимераза, не родственная полимеразам, содержащим palm-домен, и являющаяся компонентом системы РНК-интерференции эукариот (см. гл. 7), также, по-видимому, имеет вирусное происхождение (Iyer et al., 2003).

Таблица 10-2. Белки, кодируемые наиболее распространенными генами – вирусными сигнатурами.

Белок	Функция	Группы вирусов	Гомологи среди клеточных форм жизни	Комментарии
Белок капсида с укладкой типа рулета (JRC)	Основная субъединица капсида икосаэдрических вирусов	Пикорнавирусы, корновирусы, кармовирусы, фаги, содержащие дцРНК, нуклеоцитоплазматические крупные ДНК-вирусы (NCLDV), герпесвирусы, аденовирусы, паповавирусы, парвовирусы, икосаэдрические ДНК-фаги и вирусы архей	Домены типа рулета четко просматриваются в нуклеоплазминах эукариот, а также в доменах межбелкового взаимодействия определенных ферментов	Некоторые икосаэдрические вирусы, такие как фаги, содержащие оцРНК, а также альфа-вирусы, обладают неродственными белками капсида. У поксвирусов JRC не является белком вириона, но формирует промежуточные структуры в процессе морфогенеза вириона
Геликаза суперсемейства 3 (S3H)	Инициация и элонгация в ходе репликации генома	Пикорнавирусы, корновирусы, РНК-вирусы эукариот, нуклеоцитоплазматические крупные ДНК-вирусы (NCLDV), бакуловирусы, некоторые фаги (например, P4) и плазмиды (в частности, плазмиды архей)	S3H — особое, сильно разветвленное семейство АТФаз класса AAA+	Характерно слияние с праймазой в ДНК-вирусах и плазмидах

Белок	Функция	Группы вирусов	Гомологи среди клеточных форм жизни	Комментарии
ДНК-праймаза архей и эукариот	Инициация репликации генома	NCDLV, герпесвирусы, бакуловирусы, некоторые фаги	Все вирусные праймазы, по-видимому, составляют монофилетическую группу в семействе ДНК-праймаз архей и эукариот	Характерно слияние с S3H у большинства NCDLV, некоторых фагов и плазмид архей
UL9-подобная геликаза суперсемейства 2	Инициация и элонгация в ходе репликации генома	Герпесвирусы, некоторые NCDLV, некоторые фаги	Вирусные UL9-подобные геликазы представляют отдельную ветвь обширного суперсемейства ДНК- и РНК-геликаз	Слияние с праймазой у асфавирусов, мимивирусов
Эндонуклеаза инициации репликации по механизму «катящегося кольца» (RCRE)/ белок, связывающий сайт старта репликации (origin)	Инициация репликации генома	Мелкие ДНК-вирусы эукариот (парво-, гемини-, цирко-, паповавирусы), фаги, плазмиды и хелитроны	Никакой эндонуклеазы типа RCRE или origin-связывающего белка, как у паповавирусов, в клетках не существует. Однако эти белки являются производными palm-домена («домена-ладони»), который обнаруживают в большинстве ДНК- полимераз клеток	Паповавирусы обладают неактивной формой RCRE, функционирующей как origin-связывающий белок
Упаковывающая АТФаза семейства FtsK	Упаковка ДНК в вирион	NCDLV, аденовирусы, полиднавирусы, некоторые фаги (такие как P9 и M13), транспозоны нематод	Это отдельная ветвь в FtsK-HerA суперсемействе нуклеотидтрифосфатаз, обладающих P-петлей, которые включают АТФазы — ДНК-насосы бактерий и архей	

Белок	Функция	Группы вирусов	Гомологи среди клеточных форм жизни	Комментарии
АТФазная субъединица терминазы	Упаковка ДНК в вирион	Герпесвирусы, хвостатые фаги	Терминазы составляют семейство, производное от содержащих Р-петлю нуклеотидтрифосфатаз, которое отдаленно родственно суперсемействам геликаз I/II и AAA+ АТФаза	
РНК-зависимая РНК-полимераза (RdRp)/Обратная транскриптаза (RT)	Репликация РНК-генома	РНК-вирусы с позитивным геномом, вирусы, содержащие дцРНК, обратно транскрибирующиеся вирусы/элементы, возможно, РНК-вирусы с негативным геномом	Это другая крупная группа palm-доменов, отличная от тех, которыми обладают ДНК-полимеразы	RdRp вирусов, содержащих дцРНК, гомологичны полимеразам РНК-вирусов с позитивным геномом. Происхождение RdRp РНК-вирусов с негативным геномом остается неясным, хотя анализ мотивов в аминокислотной последовательности и особенно структурный анализ говорит о ее происхождении от RdRp РНК-вирусов с позитивным геномом

Перечень генов – вирусных сигнатур в табл. 10-2 консервативен. Вероятнее всего, другие гены также заслуживают статуса сигнатур, но отыскать явные свидетельства в пользу этого непросто. Секвенирование новых вирусных геномов в совокупности с всесторонним сравнительным анализом могло бы помочь выявить новые гены, которые, несмотря на относительно узкое распространение среди вирусов, могут считаться «сигнатурами». В самом деле, так может обстоять дело со многими, если не с большинством, генов класса 4, вирусных генов, консервативных для больших групп вирусов, но не клеточных форм жизни.

Комбинация свойств белков – вирусных сигнатур во многом необычна и требует эволюционистского объяснения. В самом деле, все гены-сигнатуры без исключения ответственны за важнейшие, центральные аспекты вирусного жизненного цикла, включая репликацию генома, формирование вириона и упаковку геномной ДНК в вирион (см. табл. 10-2). Обладание этими генами связывает между собой совершенно различные классы вирусов, которые часто имеют совершенно разные стратегии самовоспроизведения и различаются по размеру генома на три порядка. Наконец, у всех генов – вирусных сигнатур есть отдаленные гомологи в клеточных формах жизни (см. табл. 10-2), но вирусные версии, по-видимому, имеют общее происхождение.

Две сразу возникающие гипотезы о происхождении вирусных генов-сигнатур предлагают противоречащие друг другу эволюционные сценарии, ответственные за существование и распространение этих генов (Koonin et al., 2006).

1. Гены-сигнатуры – наследие последнего универсального общего предка вирусов (*Last Universal Common Ancestor of Viruses*, LUCAV). Этот сценарий предполагает, что, невзирая на все свидетельства в пользу противоположного (см. выше), все ныне живущие вирусы на самом деле монофилетичны, хотя их дальнейшая эволюция включала в себя масштабную потерю генов в некоторых эволюционных линиях, а также обильное заимствование новых генов от хозяев в других.

2. Напротив, в рамках гипотезы о полифилетическом происхождении вирусов распространение генов-сигнатур по всему диапазону групп вирусов может объяснить

горизонтальный перенос генов.

При более детальном рассмотрении ни одна из этих гипотез не кажется правдоподобным всеобъемлющим объяснением существования и распределения вирусных генов-сигнатур. Действительно, относительно малое число и мозаичное распространение генов-сигнатур (см. табл. 10-2) вряд ли свидетельствует о существовании LUCAV, хотя очевидно, что множество разнообразных вирусов, если не все они, разделяют некую общую историю. С другой стороны, исключительно отдаленное (но все же различимое) сходство между белками-сигнатурами из различных групп вирусов, обладающих абсолютно разными стратегиями репликации, мало совместимо со сценарием ГПГ.

Ниже в этой главе мы рассмотрим сценарий происхождения и эволюции вирусов, который не включает (в традиционном понимании) LUCAV, но сочетает в себе аспекты общего происхождения и гипотезу ГПГ и естественным образом связан с определенными моделями эволюции клеток. Наиболее простое объяснение того факта, что белки-сигнатуры, участвующие в вирусной репликации и формировании вириона, присутствуют в широком спектре вирусов, но, очевидно, отсутствуют в какой-либо из клеточных форм жизни, заключается в том, что у последних таких генов просто-напросто никогда и не было. Вместо этого наиболее правдоподобный сценарий постулирует, что *гены-сигнатуры предшествуют клеткам и происходят непосредственно из первичного, доклеточного пула генов*. Как можно представить, в таком первичном пуле отбор будет действовать прежде всего в отношении функций, непосредственно связанных с репликацией, что согласуется со свойствами большинства генов-сигнатур (см. табл. 10-2). Если учесть распространение генов-сигнатур среди многочисленных групп резко различных вирусов, важнейший вывод будет заключаться в том, что крупные классы вирусов и сами возникли на доклеточных этапах эволюции. Этот вывод – ключевой момент концепции *древнего мира вирусов*. Основная черта мира вирусов состоит в *непрерывном потоке генетической информации через огромное разнообразие эгоистичных элементов от доклеточного этапа эволюции до наших дней*.

Конкурирующие концепции происхождения и эволюции вирусов

Прежде чем мы обсудим возникающую концепцию происхождения вирусов из первичного пула генов во всей ее полноте, нам надо вкратце рассмотреть существующие гипотезы происхождения и эволюции вирусов. Традиционно эти идеи вращаются вокруг трех тем (см. рис. 10-3):

1. Происхождение вирусов из первичных генетических элементов.
2. Дегенерация одноклеточных паразитов до вирусного состояния.
3. Сценарий «сбежавших генов», который возводит вирусы к генам клеточных организмов, которые сбежали из клеточного генома и переключились на эгоистичный режим самовоспроизводства.

«Первичная» гипотеза была довольно модной в самые первые дни вирусологии, и примечательно, что Феликс д'Эррель (D'Herelle, 1922), первооткрыватель бактериофагов и один из основателей вирусологии, еще в 1922 году предположил, что фаги могут быть эволюционными предшественниками клеток. Несколько лет спустя, в 1928 году, Дж. Б. С. Холдейн предложил ту же гипотезу в классическом эссе на тему происхождения жизни (Haldane, 1928, мы вернемся к поистине пророческим идеям Холдейна в гл. 11). Однако, как только стало ясно, что все вирусы – облигатные внутриклеточные паразиты, «первичная» гипотеза была, по сути, отвергнута в силу простого и, при поверхностном рассмотрении, неопровержимого аргумента, что внутриклеточные паразиты не могут предшествовать появлению полноценных клеток. Напротив, присутствие во многих вирусах (особенно с большими геномами) многочисленных генов, произошедших от хозяев (в противоположность вирус-специфическим генам), может быть истолковано в поддержку гипотезы «сбежавших генов» или даже «дегенерации клетки». Во дни расцвета молекулярной биологии, когда фундаментальные различия между вирусами и клетками были четко осознаны – так что происхождение вирусов от клеток (пусть и дегенерировавших) было сочтено маловероятным, – гипотеза сбежавших генов стала, в большей или меньшей степени благодаря исключению прочих, общепринятой концепцией происхождения вирусов (Luria and Darnell, 1967). Однако недавно открытие гигантских вирусов и особенно тот факт, что эти вирусы обладают некоторыми важнейшими «клеточными» генами, например генами для множества компонентов системы трансляции, привели к воскрешению гипотезы клеточной дегенерации (Claverie, 2006). В самом деле, в терминах размера генома и генетической сложности открытие гигантских вирусов уничтожает границу между вирусами и клеточными формами жизни.

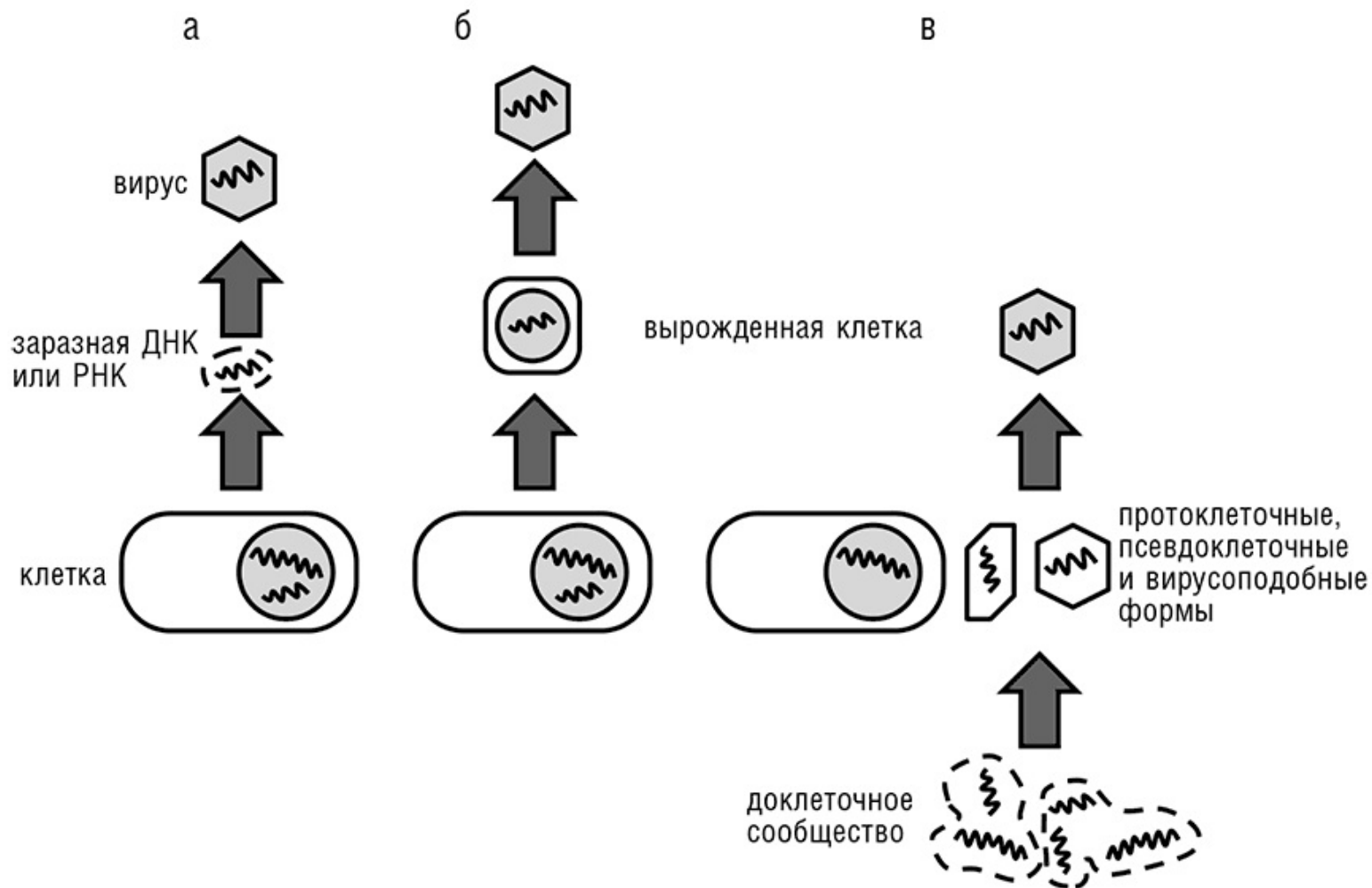


Рис. 10-3. Три конкурирующие гипотезы происхождения вирусов: а– сценарий сбежавших генов; б– сценарий клеточной дегенерации; в– сценарий первичного пула генов.

Невзирая на все эти аргументы, существование вирусных генов-сигнатур, по-видимому, успешно фальсифицирует как гипотезу дегенерации клетки, так и гипотезу сбежавших генов (или, по крайней мере, вызывает серьезные сомнения в них). Что касается гипотезы дегенерации клетки, давайте рассмотрим NCDLV (Koonin and Yutin, 2010), класс крупных вирусов, к которым эта концепция легче всего приложима и на самом деле уже прилагалась вскоре после открытия гигантского мимивируса (см. табл. 10-1). Среди девяти генов, которые объединяют (практически) все NCDLV, три наиболее важных (белок капсида с укладкой типа рулета, геликаза суперсемейства 3 и упаковывающая АТФаза) – вирусные гены-сигнатуры. Даже простейшая предковая форма NCDLV не могла бы без них функционировать. Соответственно, для обоснования клеточного происхождения этого предкового NCDLV потребовалось бы привлечь *ad hoc* определенно неэкономные сценарии, например согласованную потерю всех генов-сигнатур у всех известных клеточных форм жизни или их происхождение от вымершей крупной эволюционной линии клеток. Та же логика в основном отвергает концепцию сбежавших генов, ввиду того что у генов-сигнатур никогда не было клеточного «дома», из которого они бы сбежали. Иными словами, чтобы спасти гипотезу «сбежавших генов», нужно постулировать существование вымерших клеточных доменов, откуда могли бы сбежать гены-сигнатуры.

Таким образом, наиболее экономный сценарий эволюции вирусов, по-видимому, включает доклеточный мир вирусов. Наиболее вероятным кажется, что основные классы вирусов – по крайней мере все стратегии репликации и экспрессии генома – возникли уже в доклеточную

эру. Возможно, называть гипотетические первичные эгоистические элементы вирусами нецелесообразно, учитывая отсутствие клеток на этом этапе их эволюции. Однако, если их называть «вирусоподобными» агентами или как-то вроде этого, это никак не изменит того факта, что нет ни следа клеточного происхождения вирусов, и ни в коей мере не опровергнет гипотезу древнего мира вирусов.

Следует иметь в виду два замечания. Во-первых, рассматривая три сценария происхождения вирусов, мы говорим о вирусах (или вирусных геномах) как о «независимо эволюционирующих генетических элементах». Много, и, в некоторых вирусах, возможно, большинство, вирусных генов может быть клеточного происхождения (см. рис. 10-2), но вирусы как (квази)автономные сущности, по-видимому, не имеют клеточных корней. Во-вторых, хотя для каждой конкретной эволюционной линии вирусов сценарии исключают друг друга, разные группы вирусов в принципе могут иметь разное происхождение. В любом случае пока у нас нет достаточно убедительного свидетельства, что сценарии сбжавших генов или дегенерации клетки – лучшее объяснение происхождения каких-либо известных вирусов.

Непрерывность мира вирусов и связи с миром клеточных организмов

Анализ вирусных генов-сигнатур меток, по-видимому, предполагает, что они произошли из первичного пула генов. Существование этого первичного пула оказывается логически неизбежным, безотносительно того, какой именно выбрать сценарий ранних этапов эволюции. Мы всерьез обратимся к этим вопросам в главах 11 и 12. Здесь я хочу подчеркнуть другой и достойный всяческого внимания аспект эволюции вирусов, а именно очевидную непрерывность мира вирусов от доклеточной эры до наших дней и далее в любом обозримом будущем. В самом деле, если гены-сигнатуры 1) происходят из первичного пула генов, 2) имеют лишь очень отдаленные гомологи среди генов клеточных форм жизни и, очевидно, никогда не входили в состав клеточных геномов и 3) необходимы для репродукции вирусов, то неизбежно заключение, что эти гены переходили от вируса к вирусу (или вирусоподобному элементу) на протяжении всех четырех миллиардов лет эволюции жизни. В этом случае мы должны заключить, что *вирусные геномы, хотя и не монофилетические в обычном смысле, произошли благодаря перемешиванию и подгонке друг к другу генов в гигантской генетической сети, которую представляет собой мир вирусов.* Многочисленные гены клеточных форм жизни также пронизывают эту сеть, прежде всего благодаря геномам крупных вирусов, таких как NCDLV и крупные бактериофаги, которые позаимствовали множество генов от своих хозяев на разных этапах эволюции. Однако большинство заимствованных генов сами по себе не критичны для репликации и экспрессии вирусного генома (исключая некоторые случаи возможного неортологичного замещения генов-сигнатур); обычно эти гены участвуют во взаимодействии между вирусом и хозяином. Таким образом, несмотря на интенсивный взаимообмен генами с хозяевами, представляется, что вирусы всегда происходят от других вирусов, даже если, во многих случаях, пути эволюции – круглые. Мы теперь можем сформулировать принцип, симметричный знаменитому обобщению Рудольфа Вирхова, отражающему тот факт, что клетки размножаются делением, но никогда, насколько нам известно, не появляются *de novo: omnis cellula e cellula*. Будучи приложен к миру вирусов, принцип формулируется таким образом: *omnis virus e virus*.

Конечно, существуют вирусы и различные вирусоподобные агенты, испытывающие принцип непрерывности на прочность. Наиболее ярким примером могут быть многочисленные, разнообразные вирусы гипертермофилов кренархеот (Prangishvili et al., 2006a, 2006b) и группа полиднавирусов, заражающих насекомых. Некоторые из этих вирусов не имеют общих генов ни с какими другими вирусами; вирусы кренархеот обычно имеют немногочисленные гены с узнаваемым происхождением, а полиднавирусы содержат ряд генов, произошедших от генов хозяина (Diru et al., 2006). Однако даже среди этих необычных вирусов некоторые сохраняют один или два гена-сигнатуры, более того, у полиднавирусов заметны четкие признаки инактивации этих уцелевших генов-сигнатур. Очевидно, полиднавирусы произошли от разных групп полноценных вирусов насекомых, в частности асковирусов и нудивирусов, пройдя путь, очень напоминающий эволюцию агентов переноса генов (Bezier et al., 2009; Bigot et al., 2008). Геном нудивируса встроился в геном предковой формы осы и активно экспрессируется, продуцируя белки, необходимые для формирования вирионов и упаковки ДНК. Однако вирусные гены (по крайней мере большинство из них) не включаются в вирионы, в которые вместо этого упаковываются случайные фрагменты генома хозяина. Эта поразительная конвергенция эволюции псевдовирусов прокариот и эукариот предполагает сильное давление отбора, направленное на «изобретение» специальных приспособлений для горизонтального

переноса генов и перестройки генома.

Эти наблюдения на геномах необычных вирусов указывают на значительную пластичность, но также и на столь же примечательную устойчивость вирусного состояния: у вирусов есть способности радикально менять свое генетическое наполнение вплоть до утраты всех генов, непосредственно вовлеченных в репликацию вирусного генома, но все же оставаться явными вирусами, если говорить об их цикле размножения и образе жизни, типичном для внутриклеточных паразитов [\[108\]](#). Агенты переноса генов и полиднавирусы делают важный дальнейший шаг в эволюции этой пластичности, так как в этих случаях содержание вирусного генома полностью замещено. Более того, эти псевдовирусы избавились от независимой репликации; гены, необходимые для формирования вириона и упаковки ДНК, реплицируются лишь внутри генома хозяина, так что отношения между вирусом и клеточным хозяином становятся скорее симбиотическими, чем паразитическими. Такой симбиоз мог бы быть крайне распространенным феноменом и может принимать разнообразные обличья, насколько можно предположить исходя из того примера, что устранение всех профагов из бактериального генома существенно снижает устойчивость бактерии к различным формам стресса (Wang et al., 2010). Таким образом, многие профаги могут быть не паразитами, а скорее настоящими симбионтами.

В этой главе я прежде всего подчеркиваю непрерывность и относительную автономность мира вирусов, но пути передачи генетической информации, прямо связывающие вирусные и клеточные империи, конечно, переплетены и крайне важны для эволюции обеих империй. Важнейшим примером этого может служить вся история взаимодействия между обратнотранскрибирующимися элементами и эукариотами, которую мы уже многократно упоминали. От приобретения в древности интронов, сплайсосомы и теломеразы до совсем недавнего использования провирусных последовательностей в качестве промоторов геномы эукариот формировались благодаря многочисленным, непрерывным вкладам со стороны обратнотранскрибирующихся элементов. В ходе эволюции эти агенты утратили свой эгоистичный характер: по сути дела, они покинули мир вирусов и вступили в мир клеточных форм жизни. Ясно, что аналогичные процессы распространены среди прокариот, у которых гены профагов систематически фиксируются в хромосомах и теряют связь с вирусными геномами. Как упоминалось ранее, вирусы с крупными геномами, наоборот, систематически и непрерывно заимствуют многочисленные гены хозяина, которые часто остаются в вирусных геномах на протяжении сотен миллионов лет. Прекрасный пример этого явления представляют собой NCDLV, в частности поксвирусы, обладающие десятками и даже сотнями генов, заимствованных у хозяина на разных этапах эволюции и прежде всего участвующих во взаимодействии вируса и хозяина. Таким образом, несомненно, вирусные и клеточные миры соединены многими дорогами с двусторонним движением.

Фундаментальная неизбежность паразитов

В настоящей главе мы достаточно подробно обсуждали эмпирические свидетельства существования древнего, частично автономного, непрерывного во времени мира вирусов, который, очевидно, является ключевым компонентом биосферы с ее зарождения в доклеточную эру до наших дней. Эти наблюдения, по сути, подтверждаются сильными теоретическими указаниями на то, что возникновение паразитов – *неизбежное последствие эволюционного процесса, в котором имеется различие между геномом (генотипом) и фенотипом*. В свою очередь, различие генотипа и фенотипа – неотъемлемая часть принципа подверженной ошибкам репликации, который мы обсуждали в главе 2. Паразиты неизбежно появляются в компьютерных симуляциях эволюции простых репликаторов (Szathmary and Maynard Smith, 1997; Takeuchi and Hogeweg, 2008); не углубляясь в детали, этот исход легко понять интуитивно.

Рассмотрим РНК-белковую репликаторную систему, напоминающую репликацию современных РНК-вирусов – то есть такую, в которой каждый репликатор кодирует собственную репликазу (мы сознательно игнорируем необходимость в системе трансляции и мономерах или, скорее, имплицитно включаем все это как неопределенный ресурс – подробнее на эту тему см. гл. 12). Представляется неизбежным, что эволюция такой системы приведет к дифференциации на «хозяев» и «паразитов». Хозяевами будут полноценные геномы, кодирующие репликазу, а паразитами, как правило, будут редуцированные геномы, потерявшие ген репликазы, но сохранившие все элементы последовательности, необходимые для эффективного распознавания репликазой хозяина (см. рис. 10-4). Это обязательно произойдет из-за давления отбора, направленного на ускорение репликации, поскольку очевидная стратегия для того, чтобы увеличить скорость репликации, – избавиться от большого гена репликазы, сохраняя лишь те части генома, которые распознают репликазу, использовать репликазу, продуцируемую геномом хозяина, и, наконец, развивать все более эффективные участки распознавания для этой репликазы. В самом деле, компьютерные модели показывают, что в отсутствие компартиментализации паразиты вытеснят хозяев, так что вся система в целом рухнет и, в конце концов, вымирает; однако, если включить в модель компартиментализацию, имитирующую протоклетки (см. подробнее в гл. 11 и 12), системы хозяин – паразит приобретают устойчивую эволюционную стратегию (Takeuchi and Hogeweg, 2008).

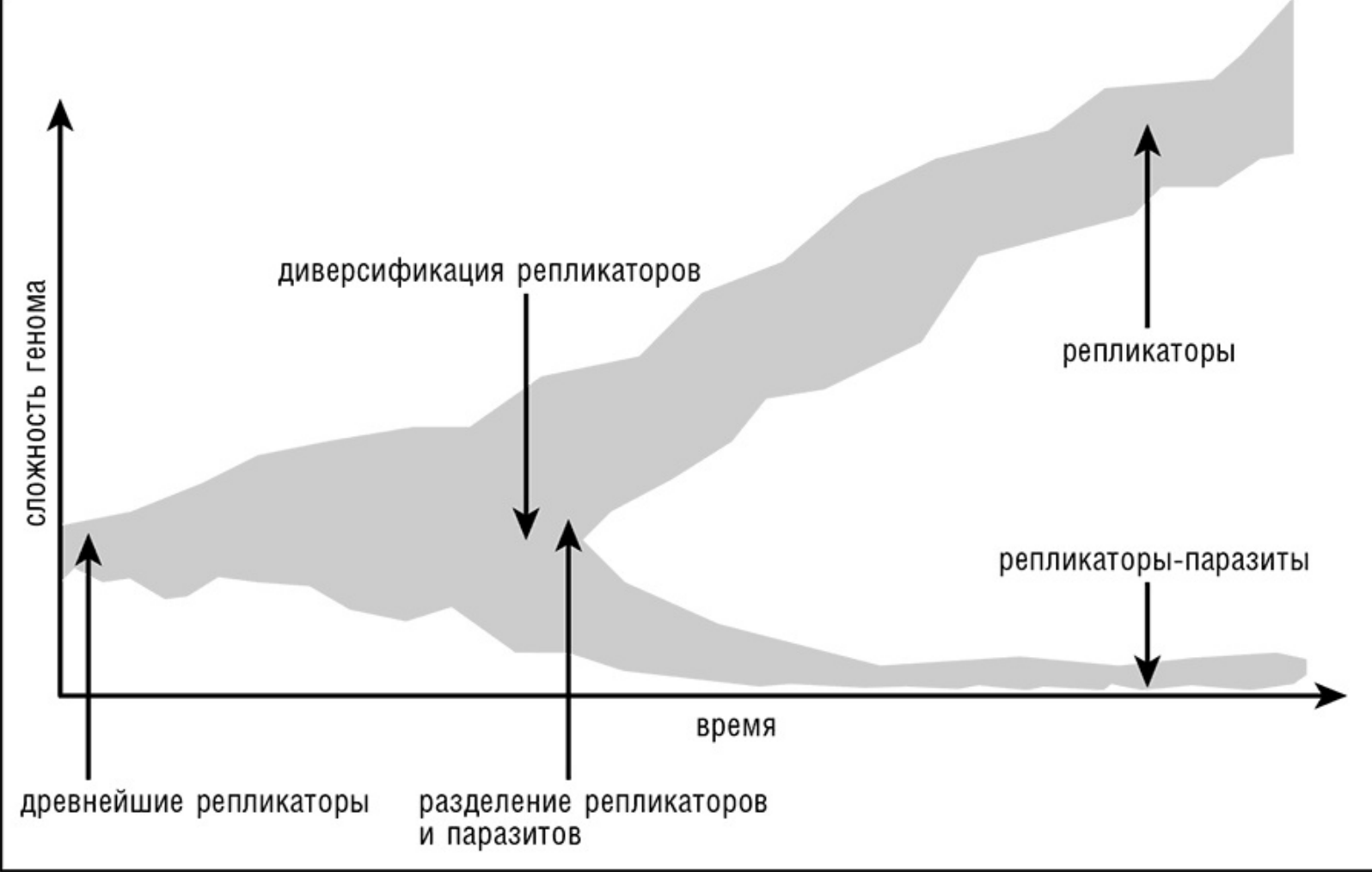


Рис. 10-4. Дифференциация популяции эволюционирующих репликаторов на хозяев и паразитов.

Возникновение паразитов как имманентное свойство систем репликаторов – не только предмет теории и моделирования. В знаменитых экспериментах Спигельмана и его коллег, описанных в главе 8, продемонстрирована та же тенденция в простейшей мыслимой ситуации, в которой единственный фактор отбора, действующий на эволюционирующую популяцию репликаторов (молекул РНК), – это необходимость распознавать репликазу и размножаться. В этих экспериментах геном РНК-бактериофага быстро теряет все кодирующие последовательности (включая самый большой ген, репликазу) и становится в полном смысле слова эгоистическим репликатором, малой РНК, состоящей почти исключительно из элементов, необходимых для репликации (Mills et al., 1973). Этот исход находится в полном согласии с интуицией и симуляционными результатами, которые показывают неизбежное возникновение паразитов. В современном мире вирусов такие эгоистичные репликаторы непрерывно возникают *de novo* и распространяются в форме так называемых дефектных интерферирующих вариантов вирусных геномов (продуцирующих дефектные интерферирующие частицы), которые особенно распространены среди РНК-вирусов (Bangham and Kirkwood, 1993). Дефектные интерферирующие производные вирусов успешно паразитируют на родительском вирусе.

Вечная гонка вооружений между хозяином и паразитом и эволюция систем защиты и взлома защиты

Согласно гипотезе Красной Королевы [\[109\]](#), коэволюционирующие системы паразит – хозяин могут поддерживать стабильную эволюционную траекторию, лишь постоянно изменяясь в непрерывной гонке вооружений. Хозяева развивают новые защитные механизмы, и паразиты отвечают, развивая механизмы взлома защиты, а также новые механизмы для атаки, уклоняющиеся от защиты, и так до бесконечности, если рассматривать эволюцию жизни в целом, или до вымирания хозяина либо паразита в каждом конкретном случае. Математическое моделирование происхождения и эволюции репликаторных систем не только неизбежно ведет к возникновению паразитов, но и показывает, что паразиты движут эволюцию механизмов репликации (Szathmary and Demeter, 1987). «Эта гонка вооружений – один из главнейших движущих факторов всей эволюции» (Forterre and Prangishvili, 2009). Справедливость этого утверждения кажется очевидной, если рассмотреть известные механизмы защиты и взлома защиты. Множественные, многослойные системы защиты составляют существенную часть геномов всех клеточных организмов, единственное исключение – некоторые внутриклеточные паразиты; с другой стороны, взлом защиты – одна из основных функций генов у вирусов с большими геномами.

У прокариот ассортимент противовирусной защиты включает в себя систему специфического иммунитета CRISPR-Cas, которую мы затрагивали в главе 10; чрезвычайно разнообразные системы рестрикции-модификации (незаменимый инструментарий генетической инженерии); и, очевидно, большое разнообразие дополнительных, менее тщательно описанных антивирусных систем. Геномы прокариот содержат множественные «островки защиты от вирусов», которые обогащены за счет известных систем защиты, но также содержат многочисленные неизученные гены (Merkl, 2006). Во многих случаях при тщательном анализе белковых последовательностей этих незнакомцев обнаруживаются домены, типичные для защитных функций, например сильно изменившиеся рестрикционные ферменты и другие нуклеазы (Makarova et al., 2009b). Таким образом, остается мало сомнений в том, что осталось открыть еще немало новых систем защиты. Нет хороших оценок того, какая доля бактериального или архейного генома обычно предназначена для противовирусной защиты, и эти значения, вероятно, сильно варьируют, тем не менее у большинства непаразитических организмов, как можно ожидать, они довольно велики – порядка 10 процентов всего набора генов [\[110\]](#).

Эволюционная динамика систем защиты может быть исключительно сложной, особенно у прокариот, потому что защитные механизмы, такие как CRISPR-Cas или системы рестрикции-модификации, не только препятствуют репликации эгоистичных элементов, но, более того, в целом предотвращают горизонтальный перенос генов (Marraffini and Sontheimer, 2008). Поскольку ГПГ – основной путь для введения новшеств среди архей и бактерий, а также, по-видимому, необходим для выживания бесполой микробных популяций (см. гл. 5), против систем защиты существует давление отбора. Вместе с движением по принципам Красной Королевы это давление – возможная причина эволюционной нестабильности систем защиты, особенно их чрезвычайно частой потери.

Защитные системы эукариот, разумеется, еще более разнообразны и сложны. Они включают в имеющийся почти у всех эукариот аппарат РНК-интерференции и различные другие механизмы врожденного и приобретенного иммунитета (в частности, система интерферона, стимулируемая дцРНК), которые мы не можем обсудить здесь детально. Следует

упомянуть, что у эукариот ГПГ не играет роли, сравнимой с его ролью у прокариот, отсюда отсутствие давления отбора, направленного на устранение систем защиты, которые, следовательно, оказываются более устойчивыми в ходе эволюции. Взятые в совокупности, системы защиты занимают значительную часть любого эукариотического генома – достаточно вспомнить главный комплекс гистосовместимости и кластеры генов иммуноглобулинов у позвоночных или огромные кластеры генов стрессового ответа у растений.

Особый и поистине радикальный тип борьбы с паразитами – запрограммированная клеточная смерть (ПКС), которая, по-видимому, происходит в различных формах у большинства клеточных организмов за исключением некоторых бактериальных паразитов. Наиболее хорошо описанные проявления этого феномена – изолированные системы ПКС животных и растений (также известные под названием «апоптоз» – в основном в применении к животным), в которых участвуют каскады самоубийственных протеолитических и нуклеолитических реакций и которые обычно запускаются вирусной инфекцией (а также другими паразитами и другими формами стресса). Существование ПКС у одноклеточных организмов, особенно прокариот, – более спорный вопрос (Bidle and Falkowski, 2004; Koonin and Aravind, 2002), но накапливающиеся свидетельства показывают, что токсин-антитоксиновая система у бактерий и архей все-таки запускает ПКС в ответ на вирусную инфекцию или другие формы стресса (Van Melderen, 2010) [\[111\]](#).

В соответствии с законом Красной Королевы, вирусы никогда далеко не отставали от своих клеточных хозяев (кроме тех, что вымерли). Лучший известный тому пример – эукариотические вирусы с большими геномами, такие как поксвирусы или бакуловирусы: до половины генов у этих вирусов функционируют как устройства взлома защиты, действующие против всех уровней защиты хозяина. Основная стратегия взлома защиты, применяемая этими вирусами, проста и эффективна: вирус «крадет» ген, кодирующий компонент защиты хозяина. После мутирования в вирусном геноме белковый продукт этого гена превращается из эффектора в доминантно-негативный ингибитор соответствующей системы защиты. Более маленькие вирусы не могут позволить себе сравнимый ассортимент генов взлома защиты, но тем не менее несут гены белков-охранников, которые по большей части участвуют в агрессии, например протеазы, расщепляющие белковые факторы, необходимые для трансляции РНК хозяина, но не вирусных РНК (Agol and Gmyl, 2010) [\[112\]](#). На другом, более фундаментальном уровне знаменательное проявление эффекта Красной Королевы – быстрая антигенная вариация у некоторых вирусов, например вируса гриппа и ВИЧ, которая позволяет этим вирусам обгонять в развитии иммунный ответ хозяина.

Важно подчеркнуть, что паразиты и защитные системы связаны не только «геномными войнами», но также и более прямым путем: эгоистичные элементы систематически привлекаются для несения защитных функций, и в то же время защитные системы могут развить эгоистичные черты. Системы рестрикции-модификации традиционно расценивались как средства защиты; и, конечно, они делают бактерии устойчивыми к чужеродной ДНК. Однако эти системы также являются особым рода эгоистичными элементами (Kobayashi, 2001, 1998). Хотя они не кодируют никаких приспособлений для собственной репликации и потому не могут считаться настоящими эмигрантами в мир вирусов, они часто обитают на плаزمиде, успешно входя, таким образом, в мир вирусов в качестве симбионтов плазмид (а эти последние точно принадлежат миру вирусов). Системы рестрикции-модификации делают клетки зависимыми от них самих и плазмид-носителей, способствуя, таким образом, собственной репликации. Токсин-антитоксиновые системы (см. гл. 5) обладают такими же свойствами. В более общем виде – хорошо известно, что и у бактерий, и у эукариот вирусы имеют тенденцию защищать хозяев от суперинфекции другими вирусами. Таким образом, взаимодействие между

вирусами («геномными паразитами») и системами защиты хозяина – весьма сложная сеть с множеством петель обратной связи, занимающая центральное положение в эволюции как паразитов, так и хозяев.

Наконец, хотя эта глава посвящена миру вирусов, необходимо сделать замечание касательно паразитизма среди клеточных форм жизни. Эта форма паразитизма относительно редка среди прокариот (хотя тут надо отметить примечательную крошечную паразитическую архею *Nanoarchaeum equitans*), но становится чрезвычайно распространенной у эукариот, чьи крупные, сложные клетки и особенно многоклеточные организмы – прекрасные мишени для микробных паразитов. В ходе эволюции эти паразиты, особенно внутриклеточные, прогрессивно теряли собственные системы противовирусной защиты и развивали механизмы, преодолевающие защиту хозяина. В этом отношении они начинают напоминать вирусы, но, безусловно, так вирусами и не становятся.

Краткий обзор и перспектива

В этой главе мы обсудили древний обширный мир вирусов. Понятие о месте вирусов в биосфере и ее эволюции радикально изменилось за последние несколько лет, во многом благодаря успехам метагеномики разнообразных виромов. Сейчас мы понимаем, что представление о вирусах как непримечательных вутрикеточных паразитах чрезвычайно далеко от действительности. Напротив, вирусы представляют собой наиболее распространенные биологические объекты на Земле, как физически, так и генетически. Сравнительная геномика вирусов и вирусоподобных элементов, таких как плазмиды и транспозоны, *раскрывает сложную сеть эволюционных взаимоотношений, в которой узлы соответствуют генам-сигнатурам, присутствующим у множества разнообразных групп вирусов* имеющих лишь отдаленные гомологи у клеточных форм жизни. Эти открытия сравнительной геномики не поддерживают гипотезу единого предка для всех вирусов, напротив, они предполагают «олигофилетический» сценарий, по которому основные классы вирусов со всеми типами циклов репликации-экспрессии произошли непосредственно из первичного, доклеточного пула генетических элементов. Оглядываясь назад, этот вывод может показаться не таким уж неожиданным, если учесть, что теория и эксперимент указывают на то, что возникновение паразитов – имманентная черта эволюционирующих систем репликаторов и что соревнование хозяев и паразитов движет эволюцию тех и других.

Таким образом, недавно нам стало известно о существовании обширного мира вирусов, который эволюционировал непрерывно и квазиавтономно от клеточных организмов на протяжении всей истории живого на Земле. В общей классификации форм жизни *вирусы и другие эгоистичные элементы, с одной стороны, и клеточные формы жизни, с другой, представляют две основные «империи»*. Несмотря на важность генов-сигнатур и относительную автономность мира вирусов, эти две империи соединены множеством дорог с двухсторонним движением генетического обмена, так что вирусы во многом формируют эволюцию клеточных геномов и наоборот. Более того, взаимодействие между вирусной и клеточной империей происходит в соответствии с принципом Красной Королевы в бесконечной гонке вооружений, которая является одним из основных формообразующих факторов в эволюции всех геномов и одним из самых отчетливых проявлений дарвиновской «борьбы за существование».

Несколько парадоксально, недавние успехи в изучении вирусов не только расширили просторы вирусного мира, но также показали, как мало мы знаем о его истинной структуре и составе. Анализ виромов предполагает, что мир вирусов по большей части состоит из неописанной «темной материи», которая может быть весьма отличной от известных вирусов. Понимание природы этой темной материи вполне может привести к существенным изменениям в общей картине эволюции жизни.

Рекомендуемая дополнительная литература

Forterre P., and D. Prangishvili.(2009) The Great Billion-Year War Between Ribosome – and Capsid-Encoding Organisms (Cells and Viruses) As the Major Source of Evolutionary Novelties. *Annals of the New York Academy of Sciences*1,178: 65–77.

Концептуально важная статья, которая подчеркивает ключевую роль гонки вооружений между вирусами и их хозяевами в эволюции сложности.

Holmes E. C.(2009) *The Evolution and Emergence of RNA Viruses*. Oxford: Oxford University Press.

Обзор эволюции вирусов с упором на макроэволюционные процессы, быструю смену и возникновение новых вирусов.

Koonin, E. V., T. G. Senkevich, and V. V. Dolja. (2006) The Ancient Virus World and Evolution of Cells. *Biology Direct*1: 29.

Анализ взаимоотношений различных групп вирусов, демонстрирующий широкое распространение вирусных генов-сигнатур. Выработана концепция мира вирусов, то есть непрерывного потока генетической информации через различные эгоистичные генетические элементы. Мир вирусов развивался бок о бок с клеточными формами жизни на протяжении всей истории жизни на Земле, но сохранил автономию от эволюционирующих клеточных организмов, несмотря на множественные обмены генами.

Koonin E. V., Y. I. Wolf, K. Nagasaki, and V. V. Dolja. (2008) The Big Bang of Picorna-like Virus Evolution Antedates the Radiation of Eukaryotic Supergroups. *Nature Reviews Microbiology*6: 925–939.

Эволюционный сценарий для самой крупной группы РНК-вирусов, заражающих эукариот, надсемейства пикорнаподобных вирусов, которое было значительно расширено благодаря метагеномике. Результаты сравнительной геномики указывают на сборку вирусного генома предковой формы пикорнаподобных вирусов из различных прокариотических элементов, за которой последовала фаза взрывной эволюции.

Koonin E. V., and N. Yutin. (2010) Origin and Evolution of Eukaryotic Large Nucleo-cytoplasmic DNA Viruses. *Intervirology*53: 284–292.

Обзор эволюции NCDLV, включающий анализ различных источников происхождения консервативных предковых генов NCDLV.

Kristensen D. M., A. R. Mushegian, V. V. Dolja, and E. V. Koonin. (2010) New Dimensions of the Virus World Discovered Through Metagenomics. *Trends in Microbiology*18: 11–19.

Обзор и статистический анализ данных по метагеномике вирусов, приводящий к заключению, что известные виромы в самом деле состоят преимущественно из вирусных последовательностей, а не случайной бактериальной контаминации. Предложена гипотеза, что «темная материя» виромов может быть представлена агентами переноса генов.

Krupovic M., and D. H. Bamford.(2008) Virus Evolution: How Far Does the Double Beta-barrel Viral Lineage Extend? *Nature Reviews Microbiology*6: 941–948.

Статья, предлагающая концепцию первичного статуса белков капсида в качестве маркеров эволюционных линий вирусов. На основании этой идеи делается вывод, что множество вирусов бактерий, архей и эукариот, обладающих белками капсида, содержащими домен типа двойного бета-бочонка (рулета), принадлежат к древней, широко разошедшейся эволюционной линии. Опубликован и контраргумент – что эволюция вирусов не может быть сведена к эволюции белков капсида (Koonin EV, Wolf YI, Nagasaki K, Dolja VV. The complexity of the virus world. *Nat Rev Microbiol*.2009 Feb 9).

Moreira, D., and P. Lopez-Garcia. (2009) Ten Reasons to Exclude Viruses from the Tree of Life. *Nature Reviews Microbiology*7: 306–311.

Морейра и Лопес-Гарсия утверждают, что, так как вирусы – облигатные внутриклеточные паразиты и не кодируют всю информацию, необходимую для собственной репликации, у них, по сути, отсутствует самостоятельная эволюционная история, они несущественны для понимания эволюции клеток и не должны считаться «живыми». Ряд последовавших писем в *Nature Review Microbiology*, включая наше с коллегами, оспаривает эти положения (исключая метафизическую проблему «бытия живым»), демонстрируя эволюционные связи между большими группами вирусов и широкое распространение генов-сигнатур по миру вирусов.

Raoult D., and P. Forterre. (2008) Redefining Viruses: Lessons from Mimivirus. *Nature Reviews Microbiology*6: 315–319.

Отталкиваясь от открытия гигантских вирусов, Рауль и Фортер описывают два фундаментальных типа организмов: капсид-кодирующие организмы (вирусы) и организмы, кодирующие рибосомы (клеточные формы жизни). Однако эта классификация игнорирует многочисленные эволюционные связи между полноценными вирусами и различными мобильными элементами, у которых отсутствует капсид.

Van Etten J. L., L. C. Lane, and D. D. Dunigan. (2010) DNA Viruses: The Really Big Ones (Giruses). *Annual Review of Microbiology*64: 83–99.

Открытие гигантских вирусов (иногда называемых «гирусами») в классе NCCLV размыло границы между вирусами и клетками в отношении размера генома и подчеркнуло роль ГПГ в эволюции крупных вирусов.

Глава 11. Последний универсальный общий предок, происхождение клеток и первичный резервуар генов

В предыдущих главах мы рассмотрели фундаментальные аспекты эволюции прокариот и эукариот, выделив древний мир вирусов и таким образом заложив фундамент для обсуждения ключевого момента в эволюции жизни, клеточной организации и разных типов клеток. Как точно предсказал Дарвин (Darwin, 1859) и вполне подтвердила сравнительная геномика, все дошедшие до нас клетки произошли от одного общего предка, который стал известен как LUCA (*Last Universal Common Ancestor* – последний общий предок всего живого) [\[113\]](#). Тем не менее не существует единого мнения о природе этого общего предка и о том, насколько походят на него современные клетки. Приводились как аргументы в пользу того, что LUCA был неотличим от современных прокариот, так и сценарии, описывающие LUCA как намного более примитивный организм (Glansdorff et al., 2008).

Трудность этой проблемы невозможно переоценить. Действительно, все известные нам клетки отличаются сложностью и исключительным совершенством организации. Самые простые клеточные формы жизни, бактериальные (и единственный известный архейный) паразиты и симбионты (см. гл. 5), очевидно, произошли в результате дегенерации более сложных организмов, однако даже они обладают несколькими сотнями генов, кодирующими компоненты высокоразвитой мембраны, системы репликации, транскрипции и трансляции, сложный аппарат деления клетки и по крайней мере несколько центральных метаболических путей. Как мы уже обсуждали, самые простые из свободноживущих клеточных организмов намного сложнее, с геномом размером не менее 1300 генов. Единственные известные автономно воспроизводящиеся агенты, которые существенно проще, чем клетки, – это вирусы, но они являются облигатными внутриклеточными паразитами и среди них не находится ничего похожего на промежуточную стадию между клеткой и вирусом (в каком бы направлении ни шла эволюция от одних к другим). Таким образом, принимая во внимание принципы *omnis cellula e cellula* и *omnis virus e virus*, требуются какие-то радикально новые решения: принцип униформизма неприменим к происхождению клеток, которое было событием, фундаментально отличающимся от знакомых нам эволюционных процессов. Итак, здесь мы обсудим прежде всего реконструкцию генного состава LUCA и выводы о происхождении клеток, следующие из этой реконструкции.

Сравнительно-геномная реконструкция генового репертуара LUCA

Почему мы уверены, что LUCA существовал? Есть множество аргументов, поддерживающих предположение о существовании LUCA, и самым сильным из них является сохранившаяся в универсальном (одинаковом для всех организмов) виде система экспрессии генов. Действительно, все известные нам клеточные формы жизни используют один и тот же генетический код (20 универсальных аминокислот и стоп-сигналов, закодированных в 64 кодонах), и небольшие отклонения содержатся лишь в сильно деградировавших геномах бактериальных паразитов и органелл. Все клетки используют одинаковые рибосомы, которые состоят из трех универсально консервативных молекул РНК и примерно 50 белков, из которых 20 универсальны. В дополнение к этому, универсально сохраняющиеся компоненты системы трансляции включают около 30 транспортных РНК, несколько факторов трансляции, 18 аминоацил-тРНК-синтетаз и несколько тРНК-модифицирующих ферментов (Anantharaman et al., 2002). За пределами системы трансляции универсально консервативны только гены трех субъединиц РНК-полимеразы. Таким образом, существует всего около 100 универсально сохранившихся генов, и все они участвуют в экспрессии генов (случается, что некоторые из этих генов, особенно те, что кодируют малые рибосомные белки, отсутствуют в аннотациях вновь расшифрованных геномов [Charlebois and Doolittle, 2004]; однако, скорее всего, эти исчезновения являются ошибками расшифровки или аннотации, а не реальными потерями в геноме). Мы уже знакомы с этими (почти) универсальными генами из главы 6, где мы наблюдали весьма сходные (хотя и не тождественные) топологии их филогенетических деревьев. Универсальное сохранение генетического кода и механизма экспрессии и единство эволюционных судеб его компонентов не оставляют места для сомнений в том, что эта система является наследством некоей формы LUCA. Основной вопрос состоит, таким образом, не в том, существовал ли LUCA, а в том, каким он был, о каких свойствах этой системы мы можем сделать выводы с достаточной уверенностью, а какие (пока что) остаются неясными [\[114\]](#).

Ясно, что ни один организм, даже самый примитивный, не может состоять из одних механизмов экспрессии, так что необходима реконструкция остальной части генового репертуара LUCA. И методы реконструкции, и их подводные камни здесь те же, что были описаны в главах 5 и 7. В контексте эволюции прокариот, которая важна для реконструкции LUCA, следует рассматривать три типа элементарных событий: (1) «рождение» гена – появление нового гена, обычно в процессе генной дупликации, за которой следует радикальная дивергенция; (2) обогащение генами за счет ГПГ; и (3) утрата гена. Напомним, что надежной реконструкции пути эволюции и наборов предковых генов препятствует неточность, связанная с относительными вероятностями событий и скоростями разных процессов, в особенности разница скоростей утраты генов и ГПГ. В принципе даже ген, найденный во всех современных клеточных формах жизни, может не быть унаследован от LUCA: его универсальность может быть последствием многочисленных горизонтальных переносов. Более того, предположения о геномном содержании предковых форм, основанные на принципе наибольшей экономии и даже наибольшего правдоподобия, консервативны, и неясно, насколько они недооценивают результат. Вопреки всем сложностям и неточностям эволюционных реконструкций, анализ наибольшей экономии совместно с менее формальными попытками реконструкции далекого прошлого конкретных функциональных систем не оставляет серьезных сомнений в том, что LUCA уже содержал несколько сотен генов (Mirkin et al., 2003; Ouzounis et al., 2006; Snel et al., 2002). В дополнение к вышеупомянутой «золотой сотне» генов, участвующих в экспрессии, этот

разнообразный генный комплект состоит из множества метаболических ферментов, включающих как пути основного энергетического обмена и биосинтеза аминокислот, нуклеотидов и коферментов, так и некоторые важные белки мембран, такие как субъединицы частицы узнавания сигналов (SRP) и H^+ -АТФазы.

Однако в реконструкции генного состава LUCA по-прежнему зияют дыры. Более всего удивляет (1) отсутствие ключевых компонентов механизма репликации ДНК, а именно полимераз, ответственных за инициацию (праймазы) и элонгацию репликации ДНК и заполнение лакун после удаления праймера, а также основных геликаз (Leipe et al., 1999); и (2) отсутствие основных ферментов липидного биосинтеза (Pereto et al., 2004). Эти незаменимые белки не попадают в реконструированный генный репертуар, поскольку данные процессы у бактерий и архей катализируются разными, неродственными ферментами, а в случае мембранных фосфолипидов еще и образуют химически разные мембраны [115].

Таким образом, реконструированный генный набор LUCA исключительно неоднороден: в то время как отдельные системы по своей сложности почти не отличаются от аналогичных систем существующих организмов, другие системы выглядят рудиментарными или вовсе отсутствуют. Столь странная картина напоминает концепцию асинхронной «кристаллизации» различных клеточных систем на ранних стадиях эволюции, которую предложил Карл Вёзе (Woese, 1998), и подсказывает, что нужно взглянуть на проблему LUCA с более общих позиций.

По-видимому, имеет смысл думать о LUCA в двух разных плоскостях (Koonin, 2009c):

- генетическая сложность, которая может быть выражена в числе генов;
- степень структурного и биологического сходства с современными клетками.

Для краткости и удобства назовем эти свойства «клеточностью».

Можно ожидать, что эти характеристики коррелируют между собой, но они не связаны жестко. В принципе вполне можно представить, что LUCA был клеточной формой жизни значительно более простой, нежели любая из современных клеток (по крайней мере свободноживущих) в смысле генетического содержания; или же, напротив, что относительная генетическая сложность была свойственна клеткам еще до появления современного типа клеточной организации – для этого сценария мы будем использовать обозначение LUCAS (*Last Ancestral Universal Common State*– последнее универсальное предковое состояние [всего живого]); см. рис. 11-1).

«Униформистское предположение» о том, что LUCA был более-менее обычной, похожей на современные прокариотической клеткой, негласно принимается в большинстве описаний ранней клеточной эволюции, но редко проговаривается явно. Тем не менее любая реконструкция LUCA должна учитывать и эволюцию тех свойств, которые трудно проследить вглубь до единого предка архей и бактерий, особенно два главных из них: репликацию ДНК и биогенез (и химию) мембран. Униформистские гипотезы, которые скрыто или явно основаны на предположении о «клеточности» LUCA, объясняют несохранение этих ключевых систем (на самом глубоком уровне реконструкции) одним из двух способов:

- каким-то образом LUCA комбинировал разные варианты этих систем, а впоследствии в архейной и бактериальной линиях эволюции был утерян один из вариантов, свой для каждой линии;
- LUCA обладал определенным вариантом каждой из этих систем, впоследствии замененным на неортологичный у архей или бактерий.

Точнее, говоря о биогенезе мембран, предполагают, что LUCA имел смешанную гетерохиральную мембрану, а две хиральные разновидности появились в результате последующей специализации архей и бактерий соответственно (Pereto et al., 2004). Касательно же репликации ДНК была высказана гипотеза о том, что одна из современных систем репликации является предковой, в то время как другая развивалась у вирусов и впоследствии заменила собою исходную либо у архей, либо у бактерий (Forterre, 1999, 2006).

В противоположность этому, более радикальные предположения относительно природы LUCA исходят из положения «что наблюдаемо, то и есть»: LUCA вообще не обладал этими ключевыми чертами, общими для бактерий и архей, по крайней мере не обладал ими в их современной форме (Koonin, 2009c; Koonin and Martin, 2005). Возможность того, что LUCA существенно отличался от всех известных ныне клеток, была впервые высказана в концепции «прогеноты» – гипотетической примитивной формы жизни, в которой связь между генотипом и фенотипом еще не полностью сформировалась (Doolittle and Brown, 1994; Woese and Fox, 1977). В своей первоначальной форме идея прогеноты включает примитивную неточную трансляцию – не слишком жизнеспособное предположение, с учетом обширной диверсификации белков до LUCA, которую однозначно демонстрирует анализ различных белковых семейств (см. гл. 12).

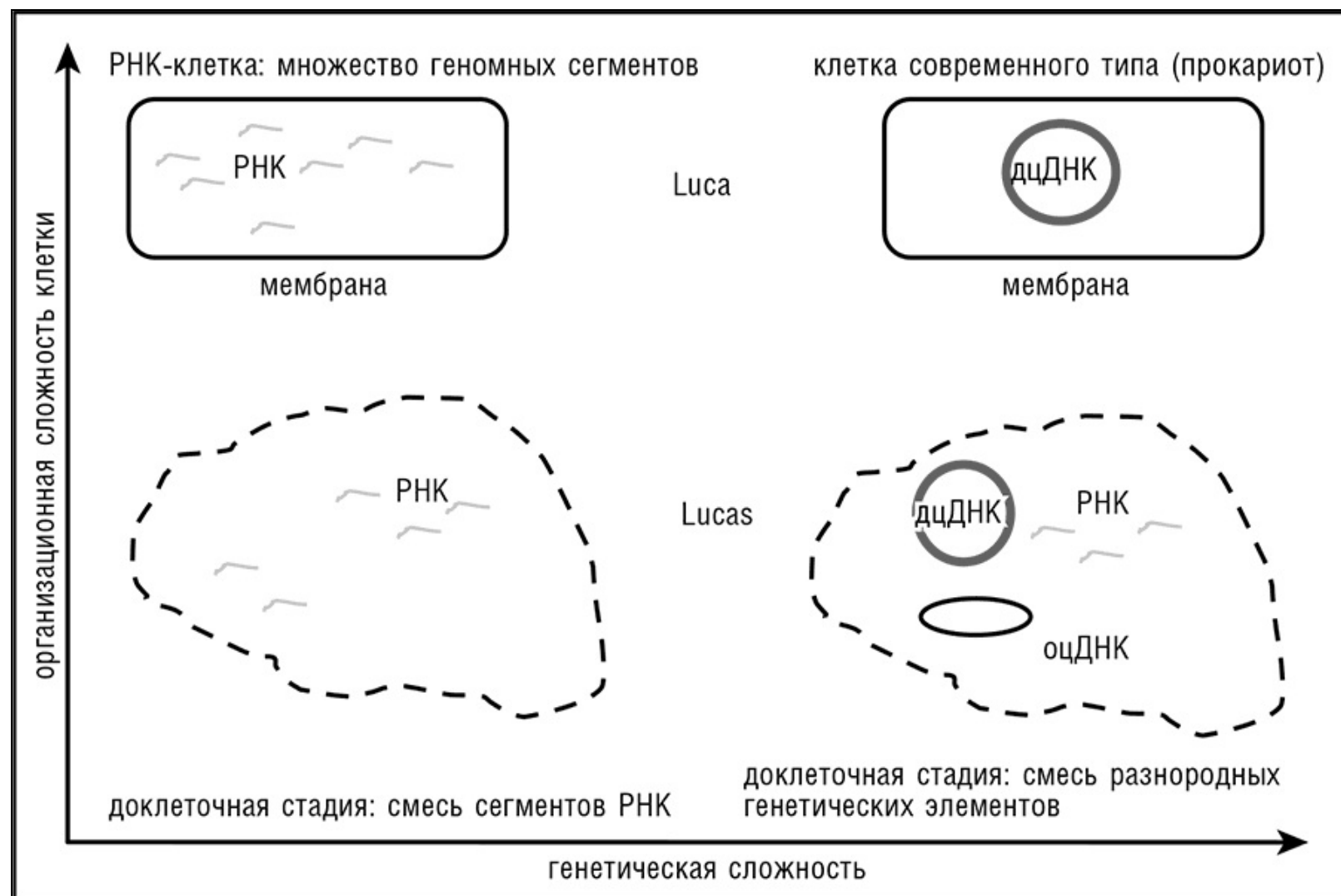


Рис. 11-1. Пространство логических возможностей для LUCA(S): генетически сложный или генетически простой, клеточный или неклеточный. Предполагается, что доклеточный пул молекул РНК генетически проще, чем смешанный пул различных генетических элементов (LUCAS), и что предполагаемая РНК-клетка проще, чем клетки современного типа (LUCA).

Более реалистично предполагать, что развитие основных черт клетки («кристаллизация» в

смысле Вёзе) была асинхронной, так что LUCA сильно напоминал современные клетки в одних аспектах, но был исключительно «примитивен» в других. Результаты сравнительной геномики дают подсказки для различения «развитых» и примитивных свойств LUCA. Так, основываясь на основных областях несходства между археями и бактериями, была высказана гипотеза о том, что LUCA

- не имел типичного большого ДНК-генома;
- не был типичной, окруженной мембраной клеткой (см. рис. 11-2; Koonin and Martin, 2005).

Загадкой ДНК-генома и репликации, которую требовалось разгадать, была комбинация негомологичных и консервативных компонентов в аппаратах репликации ДНК архей и бактерий, притом что ядро механизма транскрипции универсально консервативно. Чтобы объяснить эту смешанную картину консерватизма и разнообразия, было сделано предположение о том, что у LUCA был «ретровирусоподобный» цикл репликации с консервативным механизмом транскрипции, вовлеченным в транскрипцию дцДНК провирусного типа и консервативными компонентами системы репликации ДНК, которые в данном случае играли вспомогательную роль (Leipe et al., 1999). Эта умозрительная схема сочетала в некоем гипотетическом цикле репликации консервативные белки, вовлеченные в транскрипцию и репликацию, с белками, которые, по крайней мере в современной биосфере, однозначно принадлежат к миру вирусов, как, например, обратная транскриптаза (см. гл. 10). Этот предполагаемый изначальный цикл репликации-экспрессии, хотя формально и подходит под требования универсального сохранения упомянутых белков при неуниверсальности других ключевых компонент механизма репликации ДНК, не имеет, однако, никаких аналогий в существующих генетических системах.

РНК-репликаторы в неорганических полупроницаемых ячейках

РНК-репликаторы и механизмы трансляции в неорганических полупроницаемых ячейках

Отделение репликаторов от репликационных паразитов в неорганических полупроницаемых ячейках

Появление репликации ДНК и паразитов репликации ДНК в неорганических полупроницаемых ячейках

Бактерии, археи, ДНК- и РНК-вирусы

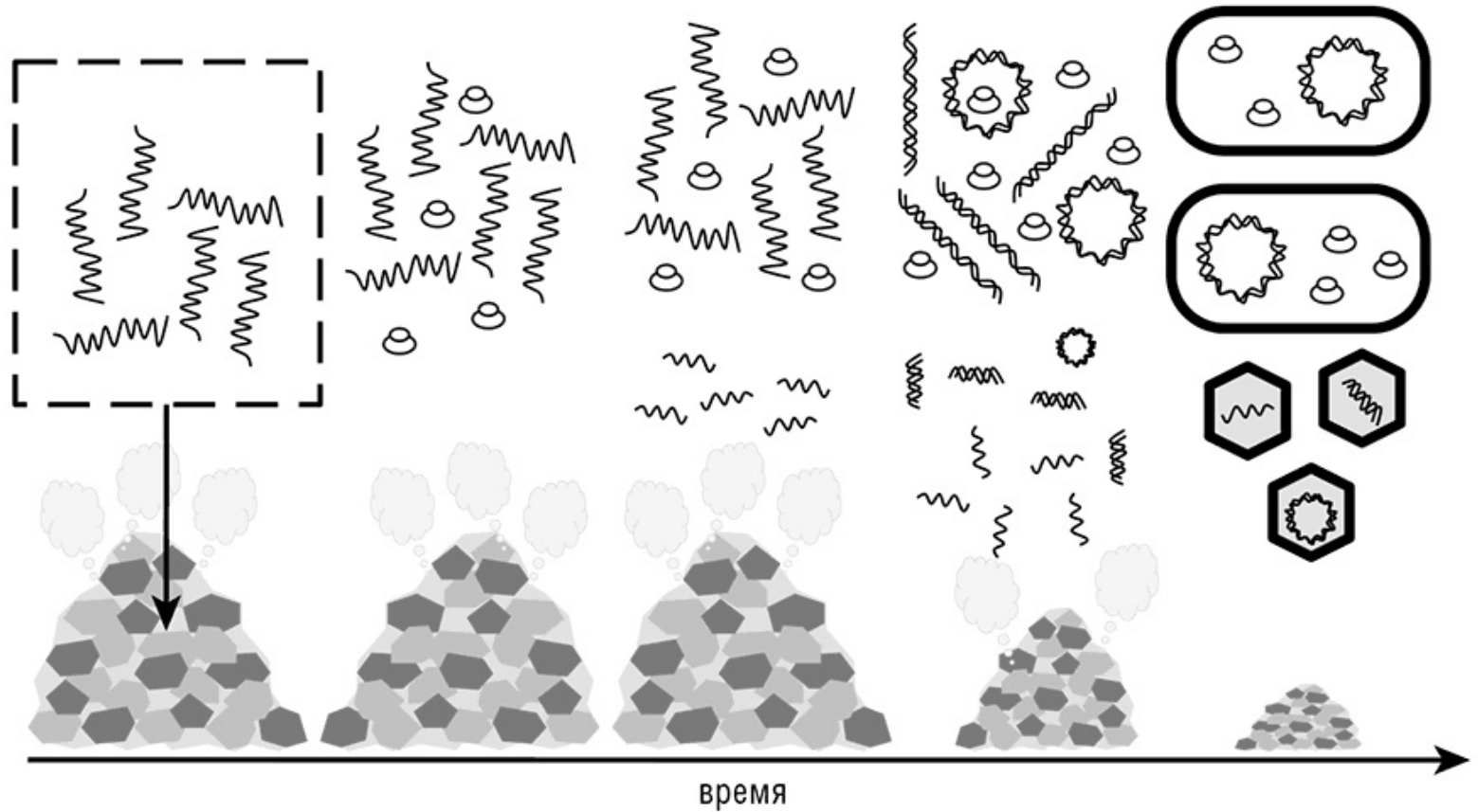


Рис. 11-2. Сценарий вирусного мира для доклеточной эволюции.

Еще одна область несходства архей и бактерий, липидный биосинтез (вместе с химией липидов), требует еще более радикальной гипотезы неклеточного, но компартментализированного LUCA (Martin and Russell, 2003; Koonin and Martin, 2005). Конкретное предположение состоит в том, что LUCA(S) был разнородной популяцией генетических элементов, существовавших в сети неорганических ячеек (в гл. 12 эти возможные инкубаторы жизни обсуждаются в подробностях). Основным затруднением для моделей безмембранного LUCA является то, что некоторые мембранные белки и даже молекулярные комплексы, такие как протонная АТФаза и частица узнавания сигнала, практически универсальны во всех современных формах клеточной жизни и, по всей вероятности, присутствовали и в LUCA.

Более тщательное рассмотрение основных проблем концепции клеточного LUCA, а именно негомологичности основных элементов системы репликации ДНК и радикального различия между фосфолипидами и ферментами липидного биосинтеза у архей и бактерий, приводит к идее о том, что две эти проблемы тесно связаны между собой (Koonin, 2009c). LUCA (реконструированный методами сравнительной геномики) как сложный организм без обширного ДНК-генома, сравнимого с геномами современных архей и бактерий, должен был бы

иметь геном, состоящий из нескольких сотен РНК-сегментов (или ДНК провирусного типа), длиной в несколько тысяч оснований каждый. Это ограничение диктуется исключительно низкой стабильностью молекул РНК в сравнении с ДНК, а также тем фактом, что самые длинные РНК-геномы (принадлежащие коронавирусам) не превышают 30 Кб.

Предполагалось, что LUCA мог быть РНК-клеткой в собственном смысле слова, из которой далее произошли три основные линии РНК-клеток (предки бактерий, архей и эукариот), геномы которых независимо заменились ДНК в результате приобретения механизмов ДНК-репликации от различных вирусов (Forterre, 2006). Этот сценарий, однако, выглядит надуманным, поскольку необходимость в сотнях «хромосом» выглядит непреодолимым препятствием для РНК-клетки. Требуемая точность распределения копий такого фрагментированного генома между дочерними клетками при клеточном делении требует исключительной точности механизма сегрегации генома, превышающей даже точность соответствующего аппарата современных прокариот. В противном случае изменение в наборе генов при каждом делении клетки, безусловно, сделает размножение невозможным. Механизмы деления, которые функционируют в современных бактериях (и, вероятно, в археях), закачивают дцДНК в дочерние клетки при помощи специальных АТФаз, которые, по всей вероятности, эволюционировали совместно с дцДНК-геномами. Таким образом, если LUCA действительно имел, вместо большого генома из дцДНК, «коллективный» геном из множества РНК-сегментов, то он должен был быть формой жизни, весьма отличной от современных клеток, возможно даже неклеточной.

Другим широко обсуждаемым аспектом ранних форм жизни, включающих и LUCA, является безудержный ГПГ, который часто рассматривается как предпосылка к эволюции сложной жизни (еще одно важное предположение Карла Вёзе [Woese, 2002]). Несомненно, ГПГ – путь к быстрым новоприобретениям, а на ранних стадиях эволюции жизни новоприобретения должны были быть исключительно быстрыми. Вёзе и его коллеги предположили и проиллюстрировали на математических моделях, что сама универсальность генетического кода могла быть связана с критической ролью ГПГ на ранних стадиях эволюции (Vetsigian et al., 2006). При обширном ГПГ единый вариант кода обязательно распространится во всей популяции предковых форм, поскольку любой организм с отличающимся генетическим кодом будет неспособен получать выгоду от ГПГ и, будучи изолированным от других организмов, непременно будет отбракован отбором [116]. Постоянный бурный ГПГ – неотъемлемая черта моделей неклеточного компартиментализированного LUCA(S), но, безусловно, не может быть принят как данность в рамках моделей клеточного LUCA. В следующем разделе мы рассмотрим модель неклеточного LUCA(S) подробнее.

Неклеточный компартиментализированный LUCA(S): сообщество разнородных репликаторов и лаборатория ранней эволюции

Майкл Рассел и коллеги предположили, что сети микроячеек в современных и древних гидротермальных источниках, состоящие в основном из сульфида железа, могли быть идеальной средой обитания для ранней жизни (Martin and Russell, 2003; Russell, 2007). Эти сети неорганических ячеек обеспечивают градиенты температуры и pH, способствующие первичным реакциям, и предоставляют универсальные каталитические поверхности для примитивной биохимии [\[117\]](#). Подробности устройства этой среды остаются предметом исследований и споров (некоторые частности разбираются в гл. 12), но мало кто сомневается, что такие сети неорганических ячеек подходят на роль инкубаторов добиологической и доклеточной биологической эволюции, от смеси органических молекул к предполагаемому изначальному миру РНК и далее до происхождения клеток, выходящих из ячеек (см. рис. 11-2). Эти отсеки могли быть населены самой разнородной популяцией генетических элементов, вначале сегментами РНК, затем более крупными и сложными молекулами РНК, включающими один или несколько белок-кодирующих генов, а еще позже – также и сегментами ДНК постепенно увеличивающегося размера (см. рис. 11-2).

Таким образом, ранние формы жизни, возможно включая LUCA(S), рассматриваются как сложные ансамбли генетических элементов, населявшие сети неорганических ячеек. Главная особенность этой модели состоит в том, что все стратегии репликации и экспрессии, использующиеся современными вирусами (см. рис. 10-1), и, соответственно, генетические элементы, кодирующие все разнообразные механизмы репликации и экспрессии, эволюционировали одновременно и в конечном итоге сосуществовали в сети, в некоторых случаях в одной и той же ячейке. Таким образом, ранняя, несколько искусственная схема, где универсально консервативные компоненты механизма репликации ДНК были вовлечены в первичный цикл репликации ретровирусного типа (Leire et al., 1999), становится ненужной.

Рассматриваемая модель объясняет отсутствие гомологии между мембранами, системами мембранного биогенеза и механизмами репликации ДНК архей и бактерий тем, что LUCA(S) не имел единого большого ДНК-генома и не был типичной клеткой, ограниченной мембраной. Согласно этой модели, первичные, доклеточные, «коммунальные» формы жизни рассматриваются как «лаборатории», в которых были «изобретены» и опробованы различные стратегии репликации и экспрессии генома, а также рудиментарные формы биогенной компартиментализации (см. рис. 11-2 и обсуждение далее в этой главе).

Главным пунктом этого сценария ранней эволюции жизни является вирусоподобная природа предполагаемой доклеточной жизни. В соответствии с этой моделью, *жизнь началась как первичный вирусный мир*. Как уже упоминалось в главе 10, мысль о том, что вирусы могли быть соотнесены с первыми формами жизни, почти так же стара, как сама вирусология. Следуя умозрительным предположениям д'Эрреля, Холдейн выдвинул эту точку зрения в более определенной форме в своем классическом эссе 1928 года о происхождении жизни (Haldane, 1928). Со своей неизменной и поразительной способностью к предвидению Холдейн предположил, что первые самовоспроизводящиеся агенты были вирусами или чем-то похожим на вирусы и что вирусный этап в эволюции жизни предшествовал появлению клеток: «Жизнь, возможно, пребывала в вирусной стадии многие миллионы лет, прежде чем подходящие совокупности элементарных частей объединились в первую клетку». Впоследствии, однако,

идея изначального происхождения вирусов и более смелая идея первичного вирусного этапа в эволюции жизни была фактически отброшена, когда стало ясно, что вирусы – облигатные внутриклеточные паразиты, большинство функций которых зависимы от клетки-хозяина. Ее место в концепциях о происхождении и эволюции вирусов, как обсуждалось в главе 10, заняли сценарии клеточной дегенерации и беглых клеточных генов.

Вирусологический ренессанс в первом десятилетии третьего тысячелетия привел к распространению гипотез и моделей, основанных на идее о важном вкладе вирусов в происхождение и эволюцию клеток. По сценарию «трех РНК-клеток и трех ДНК-вирусов» П. Фортера, современные, основанные на ДНК клетки возникли, когда три различных ДНК-вируса заменили первоначальный РНК-геном в трех клеточных линиях (предках бактерий, архей и эукариот соответственно; Forterre, 2006). Сами эти ДНК-вирусы, по предположению Фортера, эволюционировали как паразиты первичных РНК-клеток. Однако, как уже говорилось в этой главе, гипотеза РНК-клетки вряд ли жизнеспособна. Более правдоподобный сценарий, который, возможно, примиряет выводы сравнительной геномики и общую логику доклеточной эволюции, возрождает идею Холдейна на новом уровне и включает эволюцию различных вирусоподобных элементов и даже вирусоподобных частиц, предшествующую появлению клеток современного типа.

Проблема возникновения клеток наиболее ярко отражает те трудности, с которыми сталкиваются все объяснения эволюции сложных биологических структур (см. гл. 8). Действительно, среди современных биологических объектов мы не видим никаких промежуточных форм между макромолекулами и клетками, и представление о том, какими были и как работали такие промежуточные структуры, является огромной проблемой. Как мы уже неоднократно отмечали в настоящей книге, наименьшая клетка, не являющаяся паразитом или воспроизводящимся внутри других клеток симбионтом, должна нести по меньшей мере 400 генов, в то время как автотрофные клетки вряд ли могут существовать с геномом менее 1000 генов. Эти гены находятся на одной большой хромосоме (в большинстве прокариот) или на нескольких меньших хромосомах и (или) крупных плаزمиде (в меньшем числе бактерий и архей), но никогда на сегментах ДНК размером порядка оперона или гена. Фактор отбора, направляющий эволюцию к большим, протяженным геномам прост, если мы вспомним постулат Вирхова *omnis cellula e cellula*: эволюционирующей клетке требуется точно раздваивать геном, что практически невыполнимо с геномом, состоящим из сотен сегментов. Эта эволюционная логика подсказывает, что первые клетки на самом деле должны были иметь одну хромосому, и не только потому, что это так для большинства современных архей и бактерий, но, что более важно, из-за вероятной простоты и сравнительной неточности древнего механизма клеточного деления. Эволюционный рост сложных геномов, кодирующих наименьший комплект генов, необходимых для деятельности клетки, требует некоей разновидности первичной, абиогенной компартментализации, не требующей сложного мембранного аппарата современной клетки. Эту сложность не следует недооценивать: напомним, что все клеточные мембраны являются не только сложно организованными транспортными устройствами, но и машинами по преобразованию энергии, превращающими разность электрохимических потенциалов (протонный или натриевый градиент) в химическую энергию АТФ [\[118\]](#).

В настоящее время серьезно рассматриваются две формы первичной, абиогенной компартментализации: липидные везикулы и сети неорганических ячеек. Сценарий липидных везикул привлекателен тем, что в нем абиогенные мембраны являются прямыми предками современных биологических мембран. Эта возможность в настоящее время широко изучается экспериментально, в первую очередь в лаборатории Джека Шостака, и были получены интересные результаты в области транспорта полярных соединений, в том числе нуклеотидов,

через липидные мембраны (Mansy et al., 2008). Однако трудности, с которыми сталкивается эта модель, остаются очень большими. Эти проблемы достаточно очевидны и включают в себя не только транспорт мономеров со скоростями, достаточными для поддержания репликации и трансляции генетических элементов до появления белковых транспортеров, но и создание и поддержание мембранного потенциала для производства энергии. Кроме того, везикулярная модель, по-видимому, не способствует интенсивному ГПГ, важному аспекту всей микробной эволюции, но особенно необходимому на доклеточной стадии.

Не отказывая модели липидных везикул в потенциальной важности, рассмотрим модель происхождения клеток из вирусного мира как начального состояния, эволюционирующего в сетях неорганических ячеек. Эта модель, возможно, встречает меньше проблем, чем модель липидных везикул, и содержит ряд привлекательных черт, в том числе возможный ключ к разгадке происхождения биологических мембран и биоэнергетики. Как и биологическая эволюция в целом, доклеточная эволюция, несомненно, обуславливалась сочетанием случайного дрейфа и естественного отбора. Возможности для дрейфа изобилуют в рамках этой модели; пожалуй, самым простым и очевидным примером будет «засев» пустой ячейки случайным генетическим элементом. Отбор сразу же выходит на сцену с появлением репликаторов (см. гл. 2) – первоначально, как сейчас предполагается, молекул РНК, реплицируемых рибозимами, а впоследствии, после появления трансляции, молекул РНК, реплицирующихся с помощью белков (см. гл. 12). Одним из центральных аспектов модели вирусоподобной компартментализированной доклеточной стадии эволюции является *постепенный переход от отбора на уровне отдельных генетических элементов к отбору ансамблей таких элементов* [\[119\]](#), кодирующих ферменты, непосредственно участвующие в репликации, а также белки, отвечающие за вспомогательные функции, такие как трансляция и синтез предшественников нуклеиновых кислот. Отбор на уровне ансамблей генетических элементов, очевидно, является одной из форм *группового отбора*, который сам по себе является предметом давних споров среди биологов-эволюционистов и иногда отбрасывается как выдумка. Не вдаваясь глубоко в теоретические построения, отметим, что начальная эволюция, от малых генетических элементов к большим геномам, сравнимым с геномами современных клеточных форм жизни, не представляется возможной без группового отбора в той или иной форме (Koonin and Martin, 2005). Исследуя математические модели, Э. Саттари с коллегами продемонстрировали возможность группового отбора в ансамблях репликаторов, самовоспроизводящихся в разделенной на ячейки среде (Fontanari et al., 2006; Szathmary and Demeter, 1987). Некоторые из решений, возможно доступных ансамблям «эгоистичных кооператоров», известны из теории группового отбора. Наиболее очевидным и важным представляется взаимный альтруизм, когда члены группы реализуют взаимодополнительные функции, стимулирующие репродукцию друг друга [\[120\]](#). Так, в первичном ансамбле генетический элемент, кодирующий репликазу, катализировал бы репликацию элементов, кодирующих, например, компоненты систем трансляции и синтеза предшественников, – то есть функции, способствующие его собственной репликации.

Ансамбли эгоистичных кооператоров могли возникать двумя (отнюдь не взаимоисключающими) путями: (1) физическим соединением генетических элементов и (2) компартментализацией. Первый путь представляет собой начало эволюции оперонов, включая супероперон, включающий гены рибосомных белков и субъединиц РНК-полимеразы, единственный массив генов, в значительной степени сохраненный между археями и бактериями (см. гл. 5). Путь компартментализации включает эволюцию вирусоподобных частиц, которые могли содержать достаточно стабильные наборы сегментов генома, напоминающие ныне существующие сегментированные РНК-вирусы. В отличие от клеток, вирусоподобные частицы

с малыми геномами, в особенности широко распространенные икосаэдрические (сферические) капсиды, являются простыми симметричными структурами, которые во многих случаях формируются путем самосборки единственного белка капсида. Таким образом, весьма привлекательна идея о том, что простые вирусоподобные частицы были первой формой настоящей, биологической компартментализации и играли большую роль на доклеточной стадии эволюции. В дополнение к преимуществам компартментализации, вирусоподобные частицы могут защитить генетические элементы (особенно РНК) от деградации и могут служить транспортом для перемещения генов между ячейками и сетями.

Большая часть сферических вирусов с относительно сложными геномами обладает молекулярными механизмами для упаковки ДНК и РНК внутри капсида; по крайней мере в некоторых случаях эти молекулярные механизмы работают и в обратном направлении, обеспечивая экспорт вирусных транскриптов из капсидов (Rao and Feiss, 2008). Вирусные механизмы упаковки и экспорта используют моторные АТФазы как минимум трех известных семейств, которые, по-видимому, имеют общую архитектуру, формируя гексамерные каналы, по которым активно (то есть за счет энергии гидролиза АТФ) перемещаются молекулы ДНК или РНК. Примечательно, что одна из групп вирусных упаковочных АТФаз является ветвью суперсемейства FtsK-HerA, которое также включает в себя прокариотические АТФазы, ответственные за транспорт ДНК в дочерние клетки во время клеточного деления, в то время как другое семейство сходно с двигательными АТФазами бактериальных нитей (pili) (Iyer et al., 2004b). В имеющих мембраны вирионах многих вирусов аппарат упаковки перемещает ДНК и РНК как сквозь капсид, так и сквозь липидную мембрану вириона. Заманчиво предположить, что вирусные механизмы упаковки были эволюционными предшественниками клеточных насосов и двигательных АТФаз. $H^+(Na^+)$ – АТФазы/АТФ-синтазы, важнейшие, универсальные мембранные ферменты, краеугольный камень современной клеточной энергетики, также образуют подобные гексамерные каналы и, возможно, появились как часть механизма упаковки и экспорта в некоем, до сих пор не описанном (возможно, вымершем) классе вирусоподобных агентов.

Мембраны ионных градиентно-зависимых АТФ-синтаз представляют собой удивительный молекулярный «электродвигатель», роторный мотор, который переводит ионный градиент в механическую энергию вращения, а затем в химическую энергию β - γ -фосфодиэфирной связи АТФ. Анализ методами сравнительной геномики заставляет предположить, что предки двух основных ветвей мембранных АТФаз/синтаз, так называемые F-АТФазы, как правило обнаруживаемые в бактериях (и в эндосимбиотических органеллах эукариот), и V-АТФазы, характерные для архей и эндомембранных систем эукариот, произошли от общего предка, который функционировал как транслоказа белков или РНК (Mulkidjanian et al., 2007).

В предложенном эволюционном сценарии стадия транслоказы предвещает расхождение ветвей бактерий и архей, но мембранные АТФ-синтазы как таковые существенно различаются у архей и бактерий и, возможно, возникли дважды, независимо в разных ветвях клеточной жизни. Эти аналогии позволяют устремиться еще дальше в доклеточное эволюционное прошлое и предположить происхождение этой древней транслоказы от РНК-геликазы и мембранной поры или канала (см. рис. 11-3). Реконструкция эволюции мембранной АТФ-синтазы является основой для определения последовательности доклеточной эволюции: широкая диверсификация ферментов, содержащих петлю связывания фосфата (P-loop; см. гл. 12), которая породила, кроме множества АТФаз, еще и отдельное семейство РНК-геликаз (включающих бактериальный фактор терминации транскрипции Rho), произошла *еще до развития мембранной энергетики*, по крайней мере в той ее форме, которая наблюдается в современных клетках.

Весьма привлекательной кажется мысль о том, что, возможно, древнейшие вирусные

мембраны могли быть промежуточными шагами в эволюции биологической компартментализации, что относит вглубь по исторической шкале время появления полноценных клеточных мембран. Действительно, в эволюции современных сложных мембран есть парадокс. Мембраны всех современных клеток являются исключительно изошренными устройствами: двойной липидный слой непроницаем даже для небольших молекул, а перенос молекул между внутриклеточным пространством клетки и внешним миром происходит при помощи мембранных белковых комплексов, таких как каналы, поры, транслоказы и вышеупомянутые градиент-зависимые АТФазы, ответственные за клеточную энергетику. Мембрана – это еще одна сложнейшая система, понимание происхождения которой сталкивается с классической дарвиновской проблемой: жизнеспособные промежуточные стадии сложно себе представить. «Дырявая» мембрана не может обеспечить целостность клеточного содержимого, в то время как непроницаемая мембрана будет бесполезна, поскольку не позволит импортировать строительные блоки для репликации. Вирусоподобные частицы могут разрешить этот парадокс, поскольку они получают выгоду от непроницаемой мембраны, если вирион снабжен транслоказой нуклеиновой кислоты [\[121\]](#). Подобно тому как репликацию генома вирусоподобных объектов можно рассматривать как изначальный полигон для отработки репликационных стратегий, две из которых впоследствии были вовлечены в две основных линии клеточной жизни, вирусные частицы могли быть «лабораторией» для отработки молекулярных устройств, которые в дальнейшем встроились в мембраны развивающихся клеток.

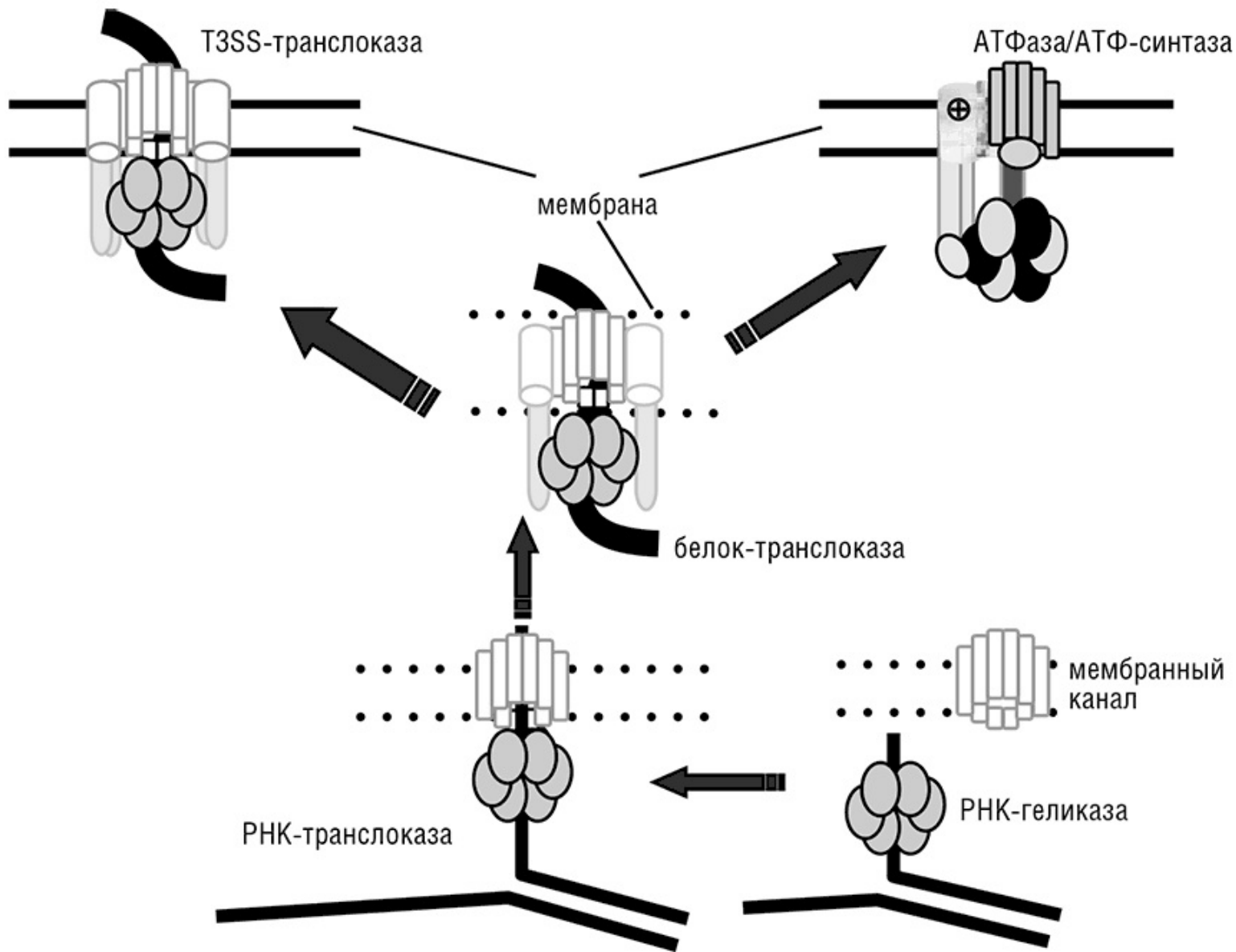


Рис. 11-3. Модель эволюции молекулярных моторов и мембранной биоэнергетики: от РНК-геликазы и мембранного канала к РНК-белковой транслоказе, и далее к ионно-градиентной АТФ-синтазе. Сплошные линии показывают современные непроницаемые для ионов мембраны, прерывистые линии показывают гипотетические «протекающие» древние мембраны. Изогнутая стрелка показывает утечку катионов из клетки. ТЗSS – система секреции типа 3, белковая транслоказа, распространенная у современных бактерий. В белковых транслоказах центральную позицию временно занимает перемещаемый белок, в то время как в мембранных АТФазах эту позицию занимают соответствующие белковые субъединицы. Сценарий эволюции взят из Mulkidjanian et al., 2007.

От отбора генных ансамблей прямая дорога к отбору содержимого ячеек, когда ячейки, поддерживающие быструю репликацию, «заражают» смежные ячейки и фактически распространяют свои коллективные «геномы»; первичные вирусоподобные частицы могли содействовать этому процессу (Koonin and Martin, 2005). Доклеточный эквивалент ГПГ – перенос генетического содержимого между ячейками – тоже часть этой модели, в согласии с идеей о том, что массовый ГПГ был неотъемлемой частью ранней стадии эволюции жизни. После того как в ходе эволюции эгоистичных кооператоров в сети неорганических ячеек был достигнут определенный уровень сложности, стал возможен «побег» протоклеточных организмов, имеющих сравнительно большие ДНК-геномы и мембраны, содержащие

механизмы транспорта и транслокации (изначально развившиеся, согласно модели, в вирусоподобных агентах). Нельзя сказать, как много подобных попыток провалилось моментально и сколько продержалось дольше, но только археи и бактерии (согласно симбиотическому сценарию более позднего возникновения эукариот, как описано в гл. 7) дожили до сегодняшнего дня. Первый успешный «побег» клеточной жизни из гипотетического доклеточного «супа» соответствует *дарвиновскому порогуклеточной эволюции*, описанному Вёзе, – порогу, за которым ГПГ должен был существенно сократиться и начаться эволюция отдельных линий (видов) клеточных организмов (Woese, 2002).

Как и в других моделях ранних стадий эволюции биологической сложности и, возможно, даже более явно, сценарий «первобытного вирусного мира», описанный здесь, сталкивается с проблемой победы эгоистичных элементов. В главе 10 мы говорили о том, что появление паразитов – черта, присущая любой эволюционирующей системе репликаторов. Если бы первобытные паразиты стали слишком агрессивными, они могли бы уничтожать своих хозяев внутри ячейки и далее выживать, только инфицируя следующую ячейку (где они снова представляли бы опасность). Можно вообразить разрушительную «пандемию», прокатившуюся по всей сети и уничтожившую все ее содержимое, и, скорее всего, именно такой была судьба многих, если не большинства, первобытных «организмов». Примечательно, что математическое моделирование репликаторов заставляет предполагать, что важной движущей силой, определившей появление ДНК, которая привела к разделению роли матрицы и катализаторов на доклеточной стадии эволюции, могла быть повышенная сопротивляемость паразитов в системах со специализированными, выделенными матрицами (Takeuchi et al., 2011). Условием для выживания доклеточных форм жизни было, во-первых, появление умеренных паразитов, которые не убивали хозяина, и, во-вторых, эволюция защитных механизмов, вероятнее всего основанных на РНК-интерференции. Повсеместное распространение умеренно эгоистичных элементов и защитных систем, основанных на РНК-интерференции, во всех ветвях клеточной жизни наводит на мысль, что эти явления появились на очень ранней, даже, возможно, доклеточной стадии эволюции.

Согласно этому сценарию, в первобытном генетическом резервуаре не существовало четко очерченных границ между эгоистичными генными элементами, которые позже стали вирусами, и большими генными ансамблями, которые в дальнейшем дали начало геномам клеточных форм жизни, хотя расхождение этих двух форм началось, когда паразиты начали «кормиться» на ансамблях «эгоистичных кооператоров». Появление клеток стало и настоящим началом мира вирусов, каким мы представляем его сегодня.

Модель доклеточной эволюции в первобытном вирусном мире, обрисованная здесь, предлагает, по-видимому, правдоподобные, хотя и весьма умозрительные решения многих загадок, связанных с происхождением клеток. Сравнительная геномика вирусов и других эгоистичных элементов дает, как мне кажется, серьезную эмпирическую поддержку этой модели. Учитывая, что, согласно такому сценарию, первые клетки произошли из неклеточного предкового состояния в ходе множественных независимых случаев возникновения протоклеток, кажется осмысленным говорить не о едином предке всех живых форм (LUCA), а о предковом состоянии (LUCAS), описывающем первобытный резервуар вирусоподобных генетических элементов.

Краткий обзор и перспектива

Все существующие формы жизни размножаются как клетки или внутри клеток. Хотя в главе 10 мы рассмотрели сильные аргументы сравнительной геномики в пользу того, что мир вирусов развивался постепенно и квазиавтономно от клеточных форм жизни на всем протяжении эволюции жизни на Земле, факт остается фактом: вирусы не могут размножаться вне клеток. Мы не знаем всех промежуточных стадий эволюции; даже самые простые клетки обладают сложной трансформирующей энергией мембраной, включающей разнообразные транспортные системы, а также обширными ДНК-геномами и сложной системой генной репликации и клеточного деления. Не существует униформистского объяснения эволюции клеток – доклеточная биота, безусловно, разительно отличалась от всей известной нам жизни. В настоящей главе мы обсуждали в основном мир вирусов в качестве сценария эволюции как вирусов, так и клеток. Согласно этой гипотезе, доклеточная стадия эволюции жизни происходила в сети неорганических ячеек, содержащих разнообразную смесь вирусоподобных генетических элементов, которые постепенно превратились в ансамбли «эгоистичных кооператоров» и истинных паразитов. Предполагается, что эти ансамбли генетических элементов были предковой стадией, из которой появились клетки; возможно, речь идет о множестве независимых «попыток», но только две из них (предки бактерий и архей соответственно) дали стабильные клеточные линии, успешные в долговременной эволюционной перспективе.

Исходя из этого гипотетического статуса первобытных форм жизни, давших начало клеткам, предлагается заменить понятие всеобщего предка (LUCA) на всеобщее предковое состояние (LUCAS). LUCA(S) мог довольно сильно отличаться от современных клеток, как нам подсказывает отсутствие гомологии ключевых компонентов репликации ДНК и биогенеза мембран (а также различий в химических структурах липидов) у архей и бактерий. Эти фундаментальные различия между двумя основными доменами клеточной жизни подразумевают неклеточную природу LUCAS. Однако не следует принимать эту модель безоговорочно: несмотря на всю правдоподобность, сценарий неклеточного LUCAS тоже сталкивается с существенными трудностями. Например, в рамках этого сценария сложно объяснить универсальное сохранение частицы узнавания сигнала, рибонуклеопротеиновой машины, которая еще до окончания трансляции встраивает образующиеся белки в мембраны [\[122\]](#).

Какой бы интригующей ни была возможность существования неклеточного LUCAS и как бы ни было важным реконструировать детали этого ключевого предкового состояния, это все же второстепенно для модели мира вирусов. Даже если модель неклеточного LUCAS будет убедительно опровергнута и появятся веские доводы в пользу клеточного LUCA, это не отменит модели доклеточной эволюции, которую мы обсуждаем, – только отбросит ее назад и будет свидетельствовать в пользу единственности успешного возникновения клетки. То же самое справедливо для модели сети неорганических ячеек (подробно рассматриваемых в гл. 12). Даже если эта модель окажется неправдоподобной, в то время как, скажем, модель клеточной эволюции из липидных везикул получит достоверное экспериментальное подтверждение, – и это вряд ли отменит необходимость существования первобытного резервуара генетических элементов. Кратко говоря, *вирусоподобный характер генетического резервуара на доклеточной стадии эволюции жизни является логической необходимостью.*

Рекомендуемая дополнительная литература

Doolittle W. F., and *J. R. Brown.*(1994) Tempo, Mode, the Progenote, and the Universal Root. Proceedings of the National Academy of Sciences USA 91: 6,721—6,728.

Обсуждение природы LUCA, в частности, был ли он прогенотой, на заре эры геномики.

Glansdorff N., *Y. Xu*, and *B. Labedan.*(2008) The Last Universal Common Ancestor: Emergence, Constitution, and Genetic Legacy of an Elusive Forerunner. Biology Direct 3: 29.

Подробный обзор гипотез и идей касательно LUCA. Согласно модели, предпочитаемой Глансдорфом с коллегами, LUCA был сообществом разнообразных РНК-клеток.

Koonin E. V.(2009) On the Origin of Cells and Viruses: Primordial Virus World Scenario. Annals of the New York Academy of Sciences 1,178: 47–64.

Концептуальный анализ, совмещающий модели доклеточной эволюции в сетях неорганических ячеек и вирусного мира, приводящий к предположению, что LUCAS был сообществом вирусоподобных агентов.

Koonin E. V.(2003) Comparative Genomics, Minimal Gene-Sets, and the Last Universal Common Ancestor. Nature Reviews Microbiology 1: 127–136.

Обзор и критический анализ реконструкций минимального и предкового генных наборов.

Koonin E. V., and *W. Martin.*(2005) On the Origin of Genomes and Cells Within Inorganic Compartments. Trends in Genetics 21: 647–654.

Модель ранней эволюции жизни, от образования первых полимеров до возникновения клеток. LUCA(S) рассматривается как неклеточное сообщество разнообразных репликаторов. Постулируется множество событий возникновения клеток, из которых, однако, только два привели к формам, выжившим в течение длительного времени и давшим начало археям и бактериям.

Morange M.(2010) Some Considerations on the Nature of LUCA, and the Nature of Life. Research in Microbiology 162: 5–9.

Обсуждение эпистемологических аспектов исследования ранних стадий эволюции жизни, включая LUCA.

Mulkijanian A. Y., *K. S. Makarova*, *M. Y. Galperin*, and *E. V. Koonin.*(2007) Inventing the Dynamo Machine: The Evolution of the F-type and V-type ATPases. Nature Reviews Microbiology 5: 892–899.

Сценарий происхождения мембранных АТФаз (АТФ-синтаз), использующих трансмембранный ионный градиент для синтеза АТФ, из геликазы и белкового комплекса мембранной поры. Такой сценарий предполагает, что значительное разнообразие белков и, в частности, возникновение геликаз, содержащих фосфат-связывающую петлю, предшествует по времени возникновению мембранной биоэнергетики современного типа, так что на ранних стадиях жизни должен был действовать иной энергетический механизм.

Mushegian A.(2008) Gene Content of LUCA, the Last Universal Common Ancestor. Frontiers in Bioscience 13: 4,657—4,666.

Современные методы и результаты реконструкции генного репертуара LUCA.

Woese C. R.(2000) Interpreting the Universal Phylogenetic Tree. Proceedings of the National Academy of Sciences USA 97: 8,392—8,396.

Влиятельнейшая статья, в которой Вёзе помещает LUCA в корень универсального древа жизни, начальную стадию эволюции, когда интеграция генетических элементов стала достаточно тесной, чтобы поддерживать эволюцию клеточной линии.

Woese C. R.(2002) On the Evolution of Cells. Proceedings of the National Academy of Sciences

USA 99: 8,742—8,747.

Дальнейшая разработка концепций предыдущей статьи. Вёзе утверждает, что первые клеточные формы были «коммунальными», на стадии эволюции, когда безудержный ГПГ был необходим для появления эволюционных новшеств, а клеточных линий не существовало. Вводится понятие «дарвиновского порога» как стадии эволюции, когда стабильность генома стала достаточной, чтобы началось вертикальное наследование.

Глава 12. Происхождение жизни. Возникновение трансляции, репликации, метаболизма и мембран: биологический, геохимический и космологический подходы

В предыдущей главе мы обсудили возможные сценарии возникновения клеток и (будем надеяться) достигли определенной степени убедительности, рассмотрев сценарий клеточной эволюции из первичного вирусного мира. Однако эта модель имеет дело с относительно поздними стадиями эволюции, на которых репликация генетического материала и трансляция уже сформировали белковое разнообразие. Ценность этих моделей останется сомнительной, пока мы не разработаем возможного объяснения происхождения фундаментальных процессов передачи информации.

Происхождение жизни – наиболее сложная проблема, стоящая перед эволюционной биологией и, можно утверждать, перед биологией в целом. Несомненно, проблема эта столь сложна, а текущее положение вещей столь трудно, что некоторые исследователи предпочитают отказываться этой проблеме в научности на том основании, что единичные события не подлежат научному исследованию. Такая позиция, однако, является глубоко неудовлетворительной, в особенности из-за того, что, хоть мы и знаем с уверенностью, что жизнь на этой планете возникла лишь однажды (см. гл. 11), у нас нет ни малейшего представления, уникальна ли (или, напротив, обычна) жизнь во Вселенной в целом. Если принимать вопрос происхождения жизни как научный, то нельзя отрицать, что это вопрос огромной значимости, в сравнении с которым прочие биологические проблемы, пожалуй, малосущественны.

Естественно потребовать, чтобы, коли мы начинаем рассуждать о происхождении некоего явления, само явление было определено. В научной и философской истории давалось множество определений жизни [\[123\]](#), и сам этот вопрос отдает эссенциализмом (см. доп. А) [\[124\]](#). Однако в контексте обсуждения в предыдущих главах прийти к определению того, что следует считать живым, удивительно просто: *любой стабильный во времени репликатор является формой жизни*. Любая репликаторная система может – и непременно будет – эволюционировать благодаря комбинации дрейфа и естественного отбора (принцип подверженной ошибкам репликации, гл. 2). Не отмеченный явно, но важный аспект такого определения – наличие обратной связи генотип – фенотип, когда некоторые мутации, то есть ошибки репликации, влияют на ее эффективность (см. гл. 2). Такого рода обратная связь вполне вообразима в гипотетическом мире РНК. Во всех известных формах жизни, однако, разделение генотипа и фенотипа более явное: в то время как генотип выражается в молекулах нуклеиновых кислот, фенотип заключается в белковых молекулах, обладающих исключительно исполнительной функцией, но не несущих информационной (матричной) [\[125\]](#) функции.

Следовательно, хотя происхождение трансляции в принципе и не является неотъемлемой частью вопроса о происхождении жизни (поскольку обитатели мира РНК должны считаться полноценными формами жизни), на практике два этих вопроса связаны прочной и, вероятно, неразрывной связью. В этой главе мы обсудим загадку происхождения репликации и трансляции. Поскольку механизмы трансляции универсально сохраняются, этот вопрос следует считать ключевым в проблеме происхождения жизни.

Вопрос о происхождении жизни по природе своей не может быть только биологическим, поскольку до того, как возникла жизнь (даже в простейшем ее воплощении), существовала

«предбиологическая» химия, которая должна рассматриваться с точек зрения химии, геохимии и геофизики. Данные этих областей обширны и сложны и в основном лежат за пределами моей профессиональной компетенции. Поэтому здесь мы приведем лишь краткий обзор, подчеркивающий наиболее важные результаты.

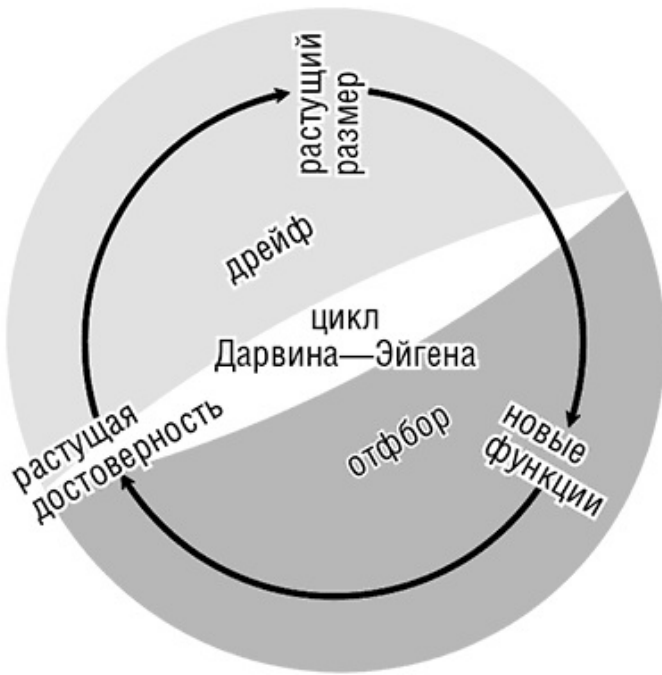
Наконец, не без понятной робости, мы коснемся чрезвычайно общих аспектов вероятности «уникальных событий» в рамках современных космологических теорий. Это рассуждение позволит нам, по крайней мере, выработать определенные идеи касательно распространенности жизни в космосе.

Происхождение репликации и трансляции и мир РНК

Цикл Дарвина – Эйгена

Главной целью разработанной Манфредом Эйгеном теории самореплицирующихся систем было построение простой модели, объясняющей происхождение биологической информации и, следовательно, самой жизни. Теория Эйгена вскрыла существование фундаментального предела, ограничивающего достоверность репликации (*порог Эйгена*): если произведение частоты ошибок (мутаций) и информационной емкости системы (размер генома) ниже порога Эйгена, происходит стабильное наследование и, следовательно, эволюция; если же эта величина выше порога, то мутационная катастрофа и вымирание неизбежны (Eigen, 1971). Порог Эйгена лежит где-то в интервале от 1 до 10 мутаций на цикл репликации (Tejero et. al., 2011), но, каков бы он ни был, для стабильного поддержания репликации, что является необходимым условием начала биологической эволюции, система должна непрерывно оставаться над этим порогом (см. рис. 12-1 а). Само происхождение первых организмов создает по меньшей мере видимость парадокса, поскольку для репликации необходима некоторая минимальная сложность, а репликация с высокой точностью требует кодирования еще большего объема информации (Penny, 2005). В то же время уровень точности репликации в данной точке эволюционной траектории ограничивает объем информации, которая может быть закодирована в геноме. Таким образом, на первый взгляд рост сложности генома представляется невозможным. Однако комбинация естественного отбора и генетического дрейфа превращает этот, казалось бы, порочный круг в (казалось бы) бесконечную спираль возрастающей сложности (цикл Дарвина – Эйгена в терминологии Д. Пенни; Penny, 2005). Даже малые приобретения в точности репликации идут системе на пользу, хотя бы потому, что они увеличивают число жизнеспособных копий генома и тем понижают репродукционные издержки. Сам по себе большой геном является скорее обузой для системы из-за высоких репродукционных издержек. Однако умеренное увеличение генома, такое как дупликация частей генома или рост вследствие рекомбинации, способно закрепляться в малых популяциях. Репликаторы, обладающие достаточно высокой точностью, могут использовать этот случайно закрепленный и изначально бесполезный генетический материал для эволюции новых функций, не срываясь при этом с «обрыва Эйгена» (рис. 12-1 б). Среди этих новых, улучшающих приспособленность функций будут и повышающие точность репликации, что позволяет затем увеличить объем кодируемой информации. Таким образом, цикл Дарвина – Эйгена повторяет себя в спиральной прогрессии, приводя к непрерывному увеличению размера генома (рис. 12-1 а).

а



б

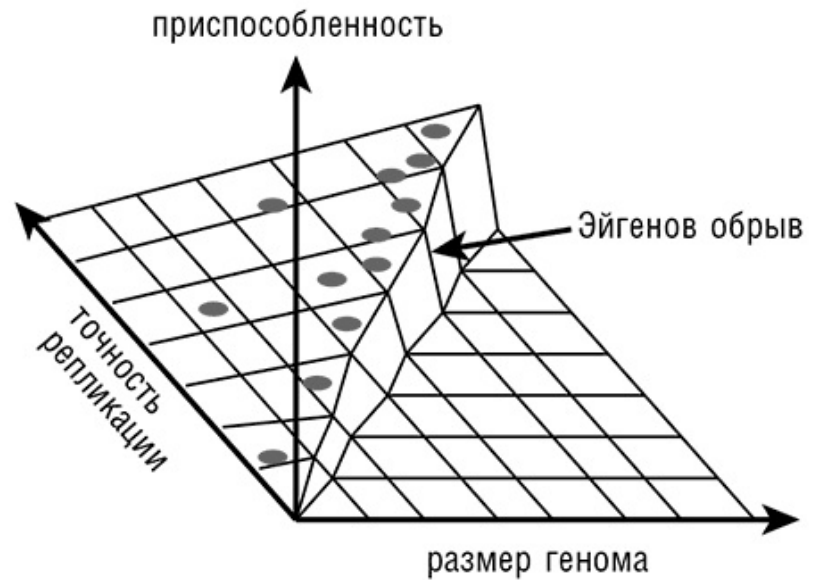


Рис. 12-1. Точность репликации и эволюция: а — цикл Дарвина – Эйгена; б — эволюция и край обрыва Эйгена.

Решающим в изучении происхождения жизни является вопрос о том, как начался цикл Дарвина – Эйгена, то есть как именно была достигнута наименьшая сложность, необходимая для приемлемой репликации. Даже в простейших современных системах, таких как РНК-вирусы, точность репликации в которых составляет всего 10^{-3} (то есть в среднем одна ошибка на 1000 нуклеотидов), и вириодах, реплицирующихся с наименьшей известной на сего дня среди всех репликонов точностью (около 10^{-2} ; Gago et al., 2009), репликация катализируется сложными белковыми полимеразы. Сами эти полимеразы (репликазы) синтезируются в результате трансляции соответствующих мРНК при посредстве чрезвычайно сложного рибосомного аппарата. Отсюда следует драматический парадокс происхождения жизни: для достижения минимальной начальной сложности, необходимой для того, чтобы биологическая система начала движение по Дарвину – Эйгену, эта система должна обладать гораздо большей начальной сложностью. В рамках обычного эволюционного мышления невозможно даже представить решения этой головоломки, поскольку это мышление относится исключительно к системам, уже находящимся на спирали. Таким образом, решение непременно окажется неординарным. В следующих разделах этой главы мы сначала рассмотрим потенциал подхода сверху вниз, то есть основанного на анализе современных генов, для получения указаний на возможное происхождение систем репликаторов. Затем мы обсудим подход снизу вверх.

Изучение эволюции белковых доменов дает аргументы в пользу сложного мира РНК: взгляд сверху вниз

Как уже упоминалось выше, система трансляции – единственный сложный ансамбль генов, сохранившийся во всех современных клеточных формах жизни. Современная трансляционная система, содержащая около 60 универсальных белок-кодирующих генов и 40 генов структурной

РНК, представляет собой наилучшим образом сохранившийся остаток LUCA(S) и лучшее свидетельство того, что в какой-то форме LUCA(S) действительно существовал (см. гл. 11). В силу исключительной консервативности системы трансляции анализ ортологических последовательностей дает крайне мало для понимания ее происхождения, которое находится за горизонтом сравнения современных форм. Реконструкция генного состава LUCA(S) указывает на сложную трансляционную систему, включающую по меньшей мере 18 из 20 аминокислот-тРНК-синтетаз (АРСаз), несколько факторов трансляции, не менее 40 рибосомных белков и некоторое число ферментов, модифицирующих рРНК и тРНК. По-видимому, у LUCA(S) ядро системы трансляции было уже полностью сформировано (Anantharaman et al., 2002).

Однако сравнение последовательностей и структур белковых компонентов и РНК внутри самой системы трансляции оказывается информативным благодаря большому числу паралогов среди соответствующих генов. Когда в реконструированном генном наборе LUCA(S) обнаруживается пара паралогичных генов, давшая ей начало дупликация, естественно, относится к более ранней стадии эволюции, и, таким образом, реконструкция последовательности древнейших дупликаций позволяет взглянуть на очень ранние этапы эволюции. История паралогичных АРСаз особенно показательна. АРСазы образуют два класса по 10 специфичностей каждый (то есть каждый класс отвечает за распознавание и активацию 10 аминокислот), с неродственными каталитическими доменами и различными наборами вспомогательных доменов. Каталитические домены АРСаз классов I и II принадлежат, соответственно, к укладкам Россмана и биотин-синтазы [\[126\]](#). Анализ эволюционной истории этих белковых укладок приводит к существенным выводам о ранней эволюции трансляционных систем, и не только об этом (Aravind et al., 2002). Каталитические домены АРСаз класса I образуют лишь небольшую ветвь эволюционного древа доменов укладки Россмана. Таким образом, появлению общего предка АРСаз предшествовало множество ветвлений эволюционного пути от примитивного древнего домена к состоянию с высоким разнообразием, соответствующему LUCA(S). Существенное разнообразие доменов укладки Россмана возникло еще до серии дупликаций, приведших к возникновению, еще до эпохи LUCA(S), АРСаз различных специфичностей (см. рис. 12-2 а). Подобная эволюционная картина обнаруживается и при анализе домена биотин-синтазы, давшего начало АРСазам класса II. Даже при рассмотрении только этих двух укладок становится очевидно, что *огромное структурное и функциональное разнообразие белковых доменов возникло еще до того, как появилась полноценная, состоящая из РНК и белков система трансляции современного типа.*

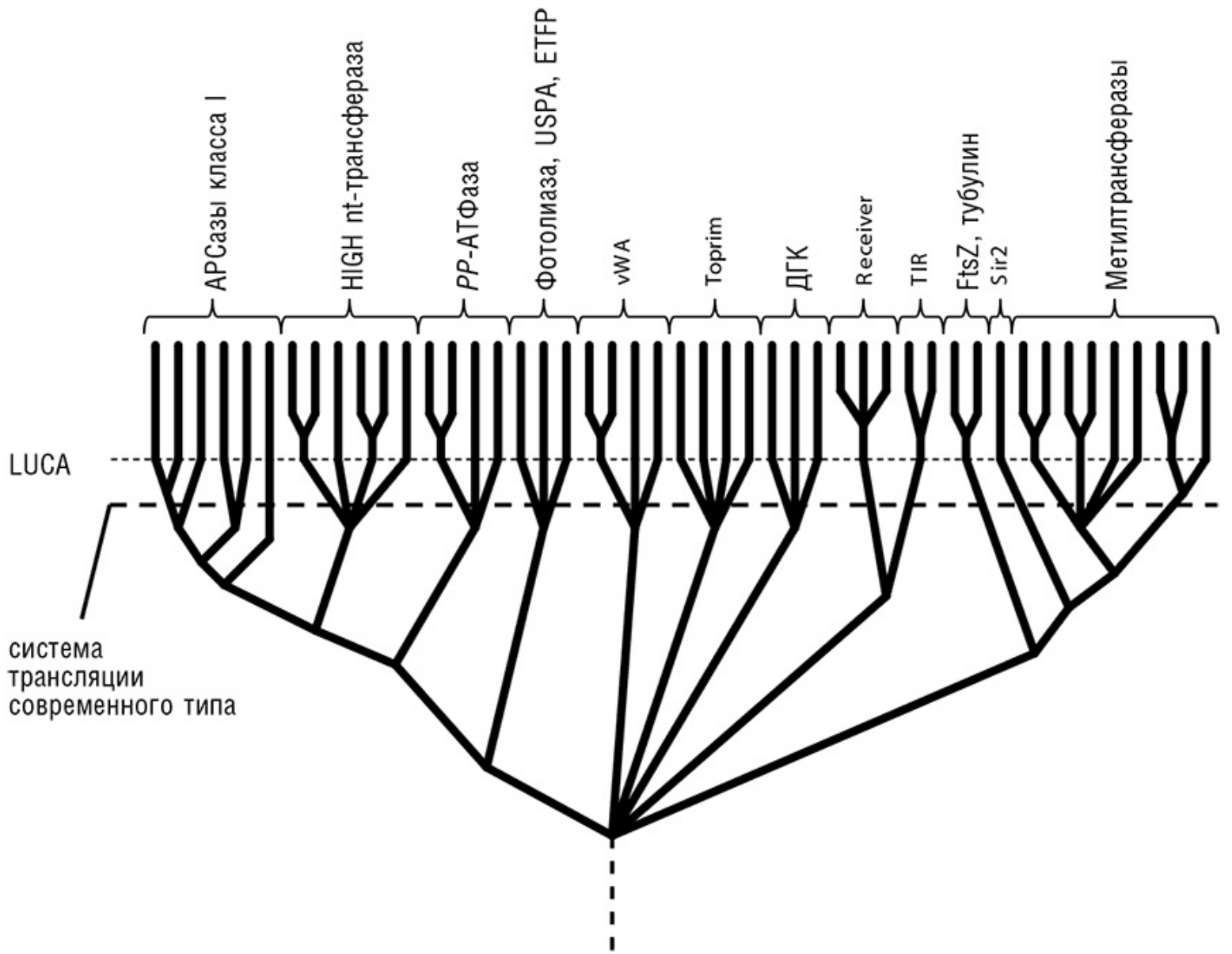


Рис. 12-2а. Разнообразие белковых доменов, кристаллизация системы трансляции и LUCA(S): эволюция нуклеотид-связывающих доменов укладки Россмана. Указаны только достаточно хорошо изученные белки. *USPA*– универсальный фактор стресса А; *ETFP*– электронтранспортный флавопротеин; *vWA*– фактор фон Виллебранда А; *Toprim*– каталитический домен топоизомераз, праймаз и некоторых нуклеаз; *ДГК*– дегалогеназа галоидных кислот; *Receiver*– компонент двухкомпонентной сигнальной системы прокариот; *TIR*– широко распространенный домен белок-белкового взаимодействия в сигнальных системах прокариот и эукариот; *Sir2*—деацетилаза белков (в частности, гистонов). Подробное описание см. Aravind et al., 2002 и ссылки в этой статье.

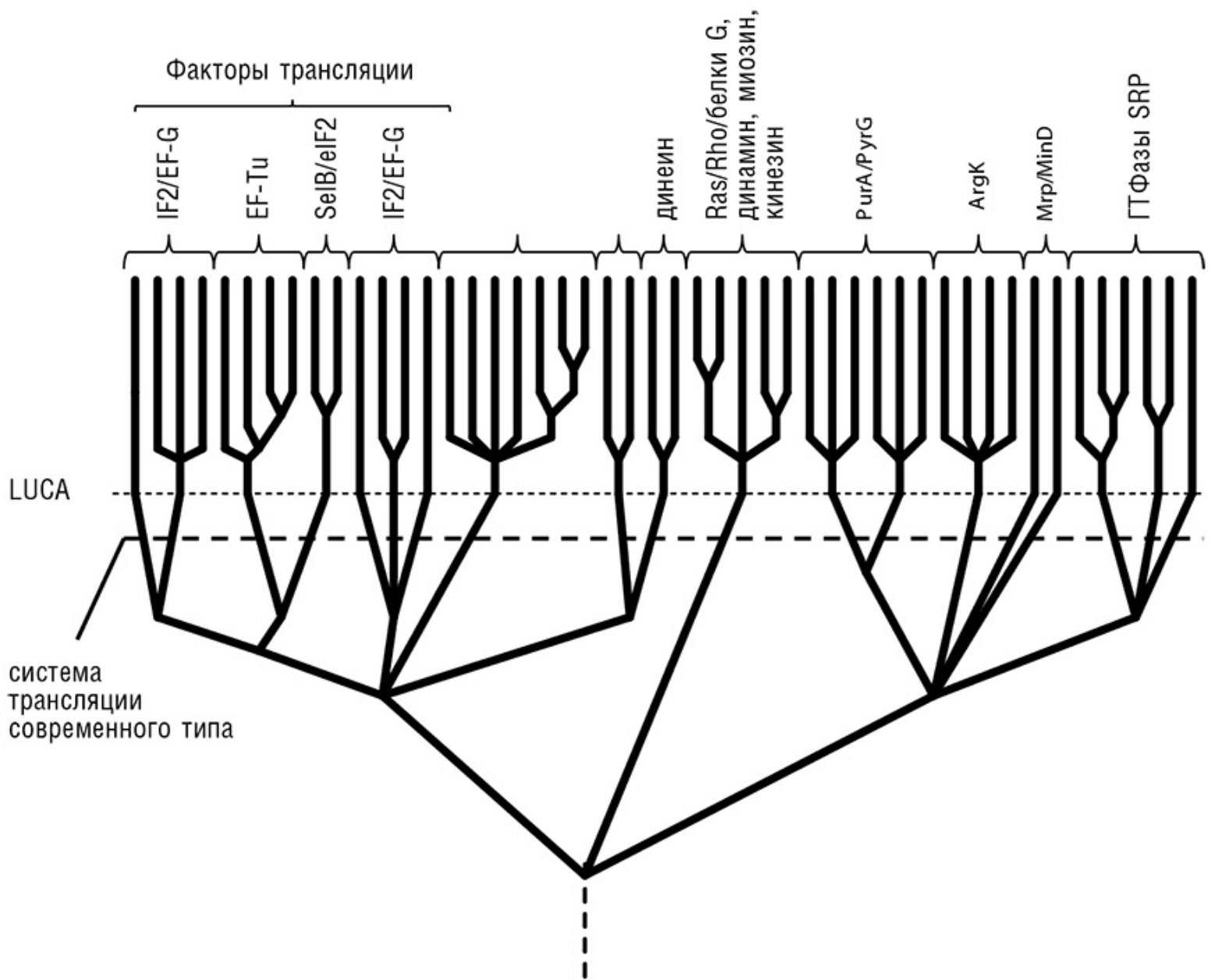


Рис. 12-26. Эволюция нуклеотид-связывающих доменов ГТФаз и родственных им АТФаз, по данным из Leire et al., 2002. Указаны только хорошо изученные белки. Динеин, динамин, кинезин и миозин – моторные ГТФазы и АТФазы, ассоциированные с цитоскелетом; Ras/Rho-сигнальные ГТФа-зы, ассоциированные, в частности, с эндомембранной системой эукариот; G-белки – ассоциированные с мембраной ГТФазы, функционирующие совместно с G-белок-сопряженными рецепторами; PurA и PyrG – ферменты метаболизма нуклеотидов; ArgK, аргининкиназа, – фермент метаболизма аминокислот; Mrp и MinD – АТФазы, участвующие в клеточном делении прокариот; SRP – частица узнавания сигналов.

Эволюционный анализ обширнейшего класса фосфат-связывающих доменов (известных также как Р-петли) ГТФаз, в котором множество трансляционных факторов образует компактные семейства, приводит, в сущности, к тому же результату: в последовательности эволюционных бифуркаций (ветвлений эволюционного древа), представляющей историю ГТФазного домена, трансляционные факторы также возникли сравнительно поздно (см. рис. 12-2 а; Leire et al., 2002). ГТФазы образуют лишь одну из многих больших ветвей эволюционного древа укладки, содержащей Р-петлю, которая включает огромное разнообразие белковых доменов, связывающих НТФ (нуклеозидтрифосфаты – чаще всего субстратом является АТФ, гораздо реже ГТФ и изредка – другие НТФ) и расщепляющих β - γ -фосфодиэфирную связь (см.

рис. 12-2 б). Эта укладка является самым распространенным доменом во всех прокариотах (Wolf et al., 1999b), и во всех реконструкциях генетического разнообразия LUCA(S) выявляется несколько десятков содержащих Р-петлю белков. Таким образом, интенсивная эволюция домена, содержащего Р-петлю, предшествовала не только LUCA(S), но и, что еще более удивительно, современной системе трансляции. Сама Р-петля (известная также под названием мотива Уолкера А [\[1271\]](#), рис. 12-3), богатая глицином петля, оборачивающая фосфатный конец НТФ-субстрата, является наиболее консервативным мотивом среди всех белковых последовательностей, несомненно зафиксированным на самых ранних стадиях белковой эволюции (Gorbalenya and Koonin, 1989; Trifonov et al., 2006).

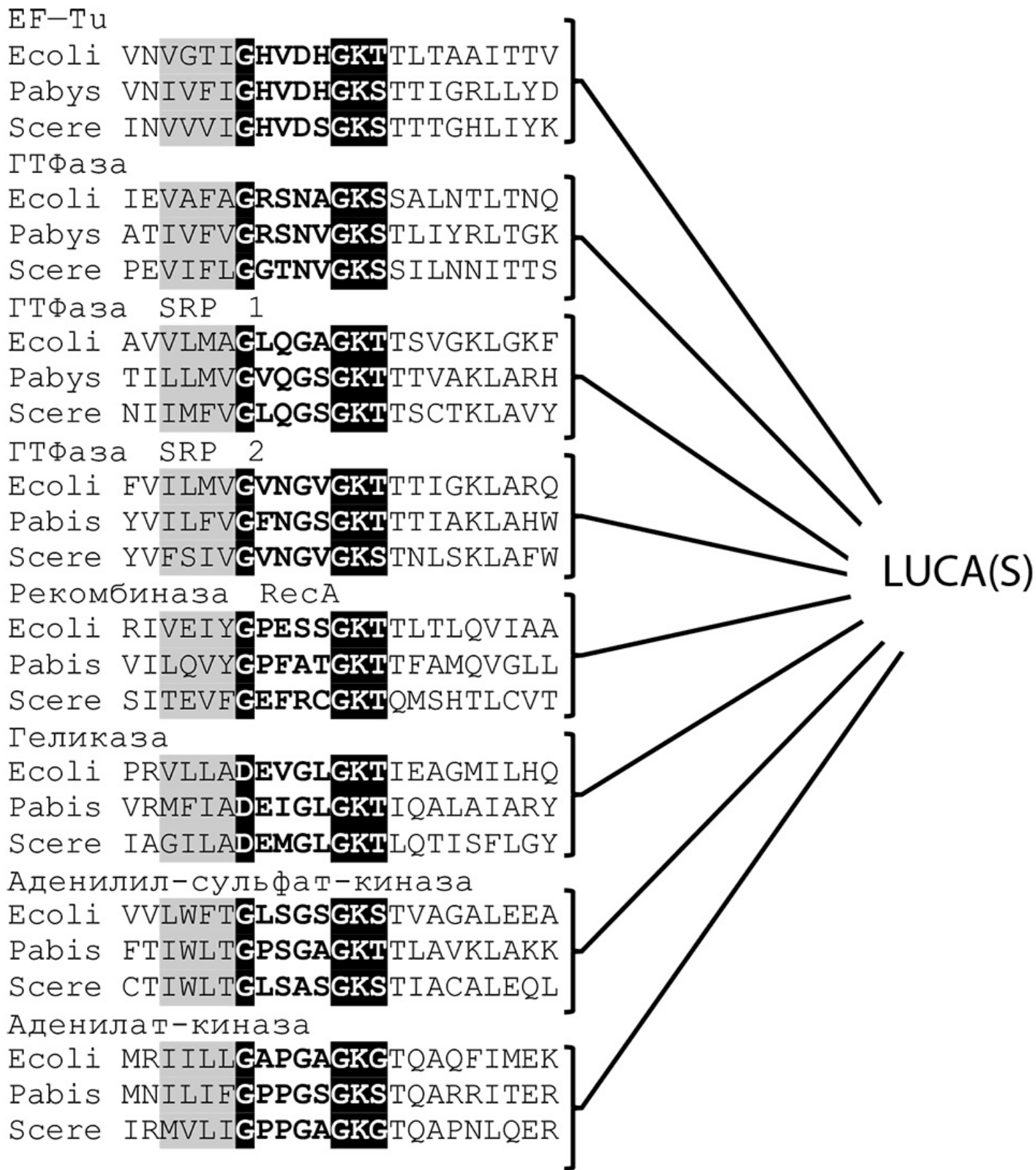


Рис. 12-3. P-петля, древнейший и самый распространенный в белковых последовательностях мотив. На рисунке изображено выравнивание P-петель восьми древних НТФаз, каждая из которых, по данным эволюционных реконструкций, была представлена в LUCA(S) (Mirkin et al., 2003). Для каждой линии представлены три последовательности: бактерии (*Escherichia coli*, Ecoli), археи (*Pyrococcus abyssi*, Pabys) и эукариота (дрожжи *Saccharomyces cerevisiae*, Scere). Белыми буквами на черном фоне обозначены аминокислотные остатки, напрямую

взаимодействующие с фосфатным концом НТФ, а серым фоном – характерная гидрофобная β -последовательность, предшествующая ФСП. SRP – частица узнавания сигналов.

Таким образом, из сравнительного анализа древних паралогических отношений между белковыми компонентами системы трансляции следует бесспорный, хоть и парадоксальный, вывод: за интересным исключением основных рибосомных белков, *все белки, играющие главную роль в современной трансляции, возникли в результате долгой и сложной эволюции различных белковых доменов*. Возникает замкнутый круг: для возникновения этих белков требуется точная и эффективная система трансляции. Древняя система трансляции могла быть и не столь совершенной, как современная, но несомненно, что она отличалась от нее не более чем на порядок по точности и скорости, иначе белковая эволюция не стала бы возможной. Однако из всего, что нам известно о современной системе трансляции, такой уровень точности невообразим без сложного и специализированного белкового аппарата.

Итак, система трансляции ярко высвечивает парадокс Дарвина – Эйгена, присущий любому анализу возникновения сложных биологических систем: для работы эффективной и точной системы трансляции современного типа требуется множество различных белков, но для того, чтобы эти белки могли возникнуть, нужна система трансляции почти столь же совершенная, как современная. По-видимому, существует только одно возможное разрешение этого парадокса, а именно через отказ от первой части противопоставления: следует заключить, что *трансляционная система, сравнимая с современной по точности и скорости, работала в отсутствие значительного разнообразия белков и, возможно, вообще без белков*. Таким образом, основываясь на сравнительном анализе составных частей системы трансляции, мы должны сделать предположение о существовании сложного и разнообразного мира РНК, в котором уже действовала некая разновидность цикла Дарвина – Эйгена.

Это не все, что может нам дать сравнительный анализ: сравнение самих РНК также вскрывает важные явления и загадывает поразительные загадки. Так, анализ большой рибосомной субъединицы РНК 23S выявил иерархический сценарий последовательности дубликаций, способной привести от простой древней шпильки РНК к современной, сложной, универсально консервативной структуре рРНК (Vokov and Steinberg, 2009).

Консерватизм структуры определенных элементов последовательностей (таких как псевдоуридиновая петля) и даже сайтов модификации тРНК всех специфичностей (и, разумеется, всех видов) не оставляет сомнений в том, что все они эволюционировали от общего предка (Eigen et al., 1989). Отсюда, при сравнении современных последовательностей и структур, возникает второй парадокс эволюции трансляции. Если на некоторой стадии эволюции существовал единственный предок тРНК всех специфичностей, то как могла подобная трансляционная система действовать, точнее, как она могла обеспечивать специфичность кодирования аминокислотных последовательностей нуклеотидными? Если же на этом этапе трансляционной системы еще не существовало, то что привело к эволюции специфичной к аминокислотам тРНК?

Мы обратимся к этим и смежным вопросам ниже, но сначала следует хотя бы кратко рассмотреть центральное понятие в области исследования происхождения жизни: мир РНК.

Рибозимы и мир РНК

Центральная догма молекулярной биологии (Crick, 1970) постулирует, что информация передается от ДНК к белку посредством РНК (Фрэнсис Крик дополнил ее возможностью обратной передачи информации от РНК к ДНК после открытия обратной транскрипции):

ДНК↔РНК ⇒ белок

Очевидно, размышляя о происхождении первых живых форм, мы оказываемся перед вопросом типа «курица или яйцо»: что появилось первым – ДНК или белок? ген или его продукт? В такой постановке вопрос, естественно, неразрешим из-за парадокса Дарвина – Эйгена: чтобы реплицировать и транскрибировать ДНК, нужны функционально активные белки, но производство этих белков, в свою очередь, требует точной репликации, транскрипции и трансляции нуклеиновых кислот. Если строго следовать центральной догме, невозможно вообразить, каким мог быть начальный материал для цикла Дарвина – Эйгена. Даже вынесение ДНК из триады и постулат о том, что изначальный генетический материал состоял только из РНК (и сведение, таким образом, триады к диаде), хоть и является ценной идеей (см. следующую главу), но не помогает разрешить парадокс. Чтобы эволюция в сторону усложнения началась, система должна каким-то образом вступить в цикл Дарвина – Эйгена до того, как установится связь между РНК-матрицами (информационной частью системы) и белком (исполнительной частью).

Блестяще остроумное и, по-видимому, единственное решение было предложено независимо К. Вёзе, Ф. Криком и Л. Оргелом в 1967–1968 годах (Crick, 1968; Orgel, 1968; Woese, 1967): ни курица, ни яйцо, но то, что между ними, – одна РНК. Уникальным свойством РНК, делающим ее вероятным и, скорее всего, наилучшим кандидатом на главную роль в древнейшей репликационной системе, является ее способность сочетать в себе информационные и каталитические функции. Было очень заманчиво предположить, что первые репликаторные системы – первые формы жизни – состояли только из молекул РНК, действующих и как носители информации (геномы и гены), и как катализаторы различных реакций, включая в том числе синтез их самих и их предшественников. Это смелое предположение получило блестящее подтверждение с открытием и последующим изучением рибозимов (ферментов РНК): Томас Чек и коллеги в 1982 году открыли автокаталитическое расщепление интрона рРНК инфузории *Tetrahymena*, а в 1983 году Сидней Альтман и коллеги показали, что РНКазы являются рибозимом. Следом за этими эпохальными открытиями изучение рибозимов выросло в огромную самостоятельную и растущую область исследований (Cech, 2002; Doudna and Cech, 2002; Fedor and Williamson, 2005).

Открытие рибозимов сделало чрезвычайно привлекательной идею о том, что первые репликаторы целиком состояли из молекул РНК, катализировавших свою собственную репликацию. В 1986 году Уолтер Гилберт ввел термин «мир РНК» для обозначения этой гипотетической стадии эволюции жизни, и гипотеза мира РНК завоевала широкую популярность, став ведущей и самой популярной гипотезой о ранних стадиях эволюции. (Различные аспекты гипотезы мира РНК и подтверждающие ее данные основательно рассмотрены в одноименной книге, вышедшей в 2010 г. в 4-м издании: Atkins et al., 2010.)

Популярность гипотезы мира РНК еще более стимулировала исследования рибозимов, нацеленные на поиск разнообразных каталитических активностей РНК – в первую очередь, пожалуй, активности РНК-репликазы. Заслуживает внимания тот факт, что главным экспериментальным подходом к получению рибозимов с желаемой активностью является отбор *in vitro*, который, во всяком случае концептуально, воспроизводит дарвиновскую эволюцию, происходившую, как полагают, в первичном мире РНК (Ellington et al., 2009). Эксперименты по направленному отбору строятся таким образом, чтобы в случайной популяции последовательностей РНК амплифицировались только те из них, которые катализируют

заданную реакцию. В многостадийных экспериментах по отбору были получены рибозимы, катализирующие весьма обширное разнообразие реакций.

В табл. 12-1 перечислены некоторые из наиболее биологически значимых реакций, катализируемых рибозимами. Примечательно, что все три элементарные реакции, необходимые для трансляции, – (1) активация аминокислот через образование аминоктил-АМФ, (2) аминоктирование (т)РНК и (3) транспептидация (реакция пептидилтрансферазы) – успешно моделируются с помощью рибозимов. Реакция авто-аминоктирования, принципиально важная для возникновения первичных РНК-адаптеров (аналог АРСазы в мире РНК), была отобрана *in vitros* относительной легкостью. Поразительно, что лучшие из полученных рибозимов катализируют эту реакцию с большей скоростью и специфичностью, чем соответствующие АРСазы, и что были отобраны очень короткие олигонуклеотиды, обладающие этой активностью (Turk et al., 2010).

Таблица 12-1. Некоторые из функций рибозимов, потенциально важные для биологической эволюции.

Реакция	Свойства рибозима
Синтез аминоктил-аденилатов	Низкоэффективное образование лейцил- и фенилаланил-аденилатов наблюдалось со 114-нуклеотидным рибозимом
Авто-аминоктирование	Авто-аминоктирование 43-нуклеотидного рибозима фенилаланином с Фен-АТФ в качестве субстрата. 77-нуклеотидная РНК катализирует ту же реакцию с более высокими избирательностью и скоростью авто-аминоктирования, чем Фен-РСаза
3'-аминоктирование РНК <i>in-trans</i>	Минимальный рибозим, способный к неизбирательному аминоктированию тРНК, состоит из 29 нуклеотидов. Был получен 45-нуклеотидный рибозим с широким спектром действия на разнообразные тРНК и аминокислоты. Сообщается также о высокоизбирательном и эффективном аминоктировании более длинных рибозимов

Реакция	Свойства рибозима
Рибозимы-пептидилтрансферазы, отобранные <i>in vitro</i>	Отобрано несколько рибозимов, образующих дипептиды из аминокислоты, связанной эфирной связью с АМФ, или олигонуклеотида и свободной аминокислоты. Обнаружено структурное сходство между рибозимами пептидилтрансферазы и соответствующим участком 23S рРНК. Описано образование Фен-Фен-тРНК 29-нуклеотидным аминокацилирующим рибозимом
Рибосомная пептидилтрансфераза	В больших рибосомных субъединицах центр пептидилтрансферазы картируется в участке, содержащем только РНК, подводя к заключению о том, что реакция катализируется рибозимом; идентификация активных остатков, однако, пока не удается
РНК-лигаза	Многие рибозимы способны лигировать молекулы РНК, прикрепленные рядом к молекуле рибозима при помощи спаривания оснований
РНК-полимераза	Из числа РНК-лигирующих рибозимов были отобраны способные к удлинению предварительно отожденного праймера на 10—14 нуклеотидов

Заимствовано с дополнениями из Wolf and Koonin, 2007.

По понятным причинам, основные усилия сосредоточились на подтверждении полимеризации нуклеотидов и, в конечном счете, репликации РНК, катализируемой рибозимами, то есть ключевых процессов в гипотетическом первичном мире РНК. Результаты экспериментов, нацеленных на создание рибозимов-репликаз, до сих пор неоднозначны (Cheng and Unrau, 2010). Получены рибозимы, способные достраивать праймер после его гибридизации с матрицей (Johnston et al., 2001); на начальном этапе такие рибозимы могли действовать только путем спаривания части молекулы рибозима с матрицей, но затем были получены, путем дополнительного отбора, и общие рибозимы-полимеразы данного класса (Lincoln and Joyce, 2009). Результаты последнего прорыва в области рибозимов-полимераз были опубликованы в то время, когда окончательно редактировалась настоящая глава: рибозим-эндонуклеаза получена с

использованием рибозима-полимераза, в свою очередь построенной рекомбинацией двух уже существующих рибозимов, что в принципе представляет собой правдоподобный путь добиологической эволюции (Wochner et al., 2011). Невзирая на серьезные успехи, полученные до сих пор рибозимы-полимеразы все еще весьма далеки от достаточно точных (в смысле порога Эйгена) процессивных репликаз, способных катализировать репликацию экзогенных матриц и самих себя. Ферменты с такими свойствами, по всей видимости, являются *conditio sine qua non* для развития гипотетического мира РНК. Кроме того, даже имеющиеся рибозимы с ограниченными способностями РНК-полимераз довольно сложны: их молекулы состоят из приблизительно 200 нуклеотидов, и эволюция таких молекул в добиологических условиях была бы нетривиальной.

Концепция мира РНК опирается не только на каталитические способности рибозимов. Процессы мира РНК и по сей день просматриваются в живых формах, пусть и многократно затмеваемые разнообразием белков с их каталитическими и структурными функциями (Doudna and Cech, 2002). Реакции, катализируемые рибозимами, пусть малочисленные и гораздо менее разнообразные, чем катализируемые белковыми ферментами, имеют важнейшее значение в современных клетках. Первейшим примером естественных рибозимов является сама рибосома, где ключевая пептидилтрансферазная реакция катализируется рРНК большой субъединицы без непосредственного участия белков (Berlinger and Rodnina, 2007) [\[128\]](#). В вездесущем ферменте процессинга тРНК, РНКазе *P*, собственно катализатором служит молекула РНК, в то время как белковые субъединицы играют роль кофакторов, стабилизирующих каталитическую РНК и способствующих реакции (McClain et al., 2010). Далее, самосплайсирующиеся интроны групп I и II, широко распространенные у бактерий и в органеллах растений, грибов и простейших, являются рибозимами, катализирующими свое собственное вырезание из РНК-транскриптов, часто при помощи особых белков, именуемых матуразами (см. также гл. 7). Практически не подлежит сомнению, что бесчисленное множество эукариотических сплайсосомных интронов, а также малые ядерные (мя)РНК, составляющие активные элементы эукариотических сплайсосом, произошли от интронов группы II (см. гл. 7). Таким образом, сплайсинг, вездесущий процесс в эукариотических клетках, основан на рибозимной каталитической реакции.

В мельчайших из известных инфекционных агентов, вириодах, катализируемые рибозимами реакции непосредственно участвуют в репликации: хотя полимеризация нуклеотидов и катализируется белковой полимеразой, процессинг промежуточных продуктов репликации с образованием зрелых геномов катализируется рибозимом, содержащимся в самой молекуле РНК вириода (Flores et al., 2004). Существование и центральное значение в современных клетках этих (и, вероятно, других, все еще неоткрытых) РНК-катализируемых реакций предполагает большую роль РНК-катализаторов в начале эволюции жизни. Все эти свидетельства, конечно, далеки от доказательства реальности древнего мира РНК, как он был определен ранее – *сообщество разнообразных РНК, обладающих разнообразными каталитическими свойствами и реплицируемых рибозимами-полимеразами*. Тем не менее свойства современных РНК, в первую очередь рибозимная активность, полностью совместимы с таким эволюционным этапом и значительно увеличивают его привлекательность. В частности, тот основополагающий факт, что пептидилтрансферазная реакция в рибосоме катализируется рибозимом, наводит на мысль о том, что система трансляции возникла как рибозимная машина.

Таким образом, три независимые группы свидетельств сходятся в поддержку важнейшей роли РНК – а именно РНК-катализа – на самых ранних этапах истории жизни и совместимы с реальностью сложного древнего мира РНК, который Вёзе, Крик и Оргел постулировали вначале на чисто логических основаниях.

1. Сравнительный анализ белковых компонентов системы трансляции и их гомологов, выполняющих другие функции, наводит на мысль о том, что широкое разнообразие белкового мира сформировалось, когда система трансляции состояла в основном из РНК.

2. В современных клетках действует несколько классов рибозимов, и их свойства совместимы с гипотезой о том, что они являются реликтами первичного мира РНК.

3. Хотя рибозимы и менее универсальны, чем белковые ферменты, и обычно значительно уступают им в каталитической активности, они, как было показано в экспериментах по отбору, способны катализировать значительное разнообразие реакций, включая и те, что играют центральную роль в эволюции трансляции (см. табл. 12-1).

Невзирая на все эти аргументы в ее поддержку, гипотеза мира РНК сталкивается с серьезными трудностями. Во-первых, несмотря на все приложенные усилия, отобранные *in vitro* рибозимы оказываются (относительно) слабыми катализаторами большинства реакций; отсутствие эффективных процессивных рибозимов-полимераз представляется особенно тяжелой проблемой, но имеется также и серьезная нехватка других видов активности, в частности необходимых для синтеза нуклеотидов. Нужно признать, что было бы нереалистичным ожидать от экспериментов по эволюции рибозимов *in vitro* легкого воспроизведения фактической сложности изначального мира РНК. Хотя эти эксперименты и ставят на службу мощь отбора, они, очевидно, выполняются в совершенно другом временном масштабе и в условиях, неспособных точно воспроизвести (неизвестные) условия в начале жизни (мы обсудим потенциальные экологические ниши для возникновения жизни ниже в этой главе).

Исследование Э. Сатмари и сотрудников дает количественную оценку сложности, которая может быть достигнута в мире РНК, и точности репликации, необходимой для достижения этого уровня сложности (Kun et al., 2005). Оценка, основанная на функциональной устойчивости к мутациям хорошо известных рибозимов, показывает, что при частоте ошибок 10^{-3} на нуклеотид за цикл репликазы (это примерно соответствует точности РНК-зависимой РНК-полимеразы современных вирусов) РНК-«организм», состоящий из примерно сотни «генов» размером с тРНК (80 нуклеотидов), будет устойчивым. Такой уровень точности всего лишь на порядок выше, чем у самых точных рибозимов-полимераз, полученных отбором *in vitro*. Данную величину можно положить приближенной верхней границей сложности ансамблей совместно развивающихся «эгоистичных кооператоров», которые могли представлять собой «организмы» мира РНК.

Даже в лучшем случае мир РНК вряд ли обладал потенциалом эволюции дальше чрезвычайно простых «организмов». Для достижения большей сложности потребовались изобретение трансляции и «белковый прорыв» (перенос основной каталитической активности на белки). Однако силы отбора, лежащие в основе возникновения системы трансляции в мире РНК, остаются неясными, и реконструкция пути к трансляции крайне сложна. Это отсутствие ясности в отношении непрерывности эволюции от мира РНК к РНК-белковому миру является второй по значимости проблемой гипотезы мира РНК, возможно даже более существенной, чем ограниченный каталитический арсенал и (как правило) низкая эффективность рибозимов. Далее мы обсудим возможные пути выхода из этой ситуации.

Природа и происхождение генетического кода

Природа и происхождение значения кодонов универсального генетического кода имеют решающее значение для понимания того, как могла возникнуть трансляция. Эволюция кода остро интересовала исследователей еще до того, как код был полностью расшифрован, и уже самые ранние работы по этой теме отчетливо различали три (возможно, и не взаимоисключающие) эволюционных модели: (1) стерическая комплементарность, обеспечивающая специфическое взаимодействие между аминокислотами и соответствующими им кодонными или антикодонными триплетами нуклеотидов, (2) «застывшая случайность», то есть закрепление случайного кода, который стало практически невозможно существенно изменить позже, и (3) адаптивная эволюция кода, начавшаяся с первоначально случайного соответствия кодонов аминокислотам (Crick, 1968).

Структура кода явственно неслучайна: кодоны родственных аминокислот в основном смежны в кодовой таблице, что обуславливает высокую (хотя и не наивысшую возможную) устойчивость кода к мутациям и ошибкам трансляции, как впервые отметил Вёзе (Woese, 1967), а С. Фриленд и Л. Херст впоследствии показали количественно (Freeland and Hurst, 1998). Высокая надежность кода опровергает сценарий «застывшей случайности» в его крайней форме (случайная привязка кодонов к аминокислотам без последующей эволюции), однако и стереохимическая модель, и модель отбора, и их сочетание, и «застывшая случайность» с последующей адаптацией способны в принципе объяснить наблюдаемые свойства кода (Koonin and Novozhilov, 2009).

Положение осложняется тем, что неизвестно, существует ли стереохимическое соответствие между аминокислотами и соответствующими триплетами нуклеотидов (кодонами). Ответ на этот кажущийся простым вопрос оказывается на удивление трудноуловимым. Ранние попытки обнаружить избирательность взаимодействия (поли)аминокислот и полинуклеотидов были безрезультатны, что говорит о том, что если специфическое сродство и существует, то отнюдь не строгое, и взаимодействие, вероятно, слабо и зависимо от внешних факторов. Хотя и имеются некоторые сообщения о неслучайных взаимодействиях аминокислот с олигонуклеотидами, но в целом попытки продемонстрировать такое взаимодействие напрямую оказались неудачными (Saxinger and Ponnampereuma, 1974).

Интерес к стереохимической гипотезе возродился с появлением метода усиления отбора (SELEX) для выделения олигонуклеотидов (аптамеров), избирательно связывающихся с аминокислотами (Yarus et al., 2005, 2009). Для восьми аминокислот с большими боковыми цепями были выделены аптамеры, значительно обогащенные триплетами нуклеотидов, идентичными либо кодонам, либо антикодонам соответствующей аминокислоты. Результаты этих экспериментов, правда, несколько неубедительны, так как для одних аминокислот аптамеры содержат в основном кодоны, тогда как для других – в основном антикодоны. Рассматриваемые в совокупности, данные по сродству аптамеров к аминокислотам считаются серьезным аргументом в поддержку стереохимической гипотезы происхождения кода, однако реальная надежность и значимость этих результатов остаются под большим вопросом.

Наличие как кодонов, так и антикодонов в аптамерах для нескольких аминокислот затруднительно интерпретировать в терминах стереохимической комплементарности. Кроме того, аминокислоты, для аптамеров которых получены подробные данные, имеют сложные боковые цепи (которые, предположительно, необходимы для избирательного взаимодействия с аптамерами) и, вероятно, являются поздними добавлениями к генетическому коду (Trifonov, 2004). До тех пор, пока не будут получены аналогичные результаты для простых,

предположительно древних аминокислот, эксперименты по отбору аптамеров затруднительно рассматривать как убедительный аргумент в пользу стереохимической гипотезы происхождения кода.

Итак, ключевой вопрос о том, играло ли роль в происхождении кода прямое взаимодействие между аминокислотами и специфическими нуклеотидными триплетами, все еще ждет ответа. В следующем разделе, при обсуждении происхождения трансляции, мы будем максимально объективны, рассматривая происхождение кода либо через избирательное взаимодействие между аминокислотами и соответствующими нуклеотидными триплетами, либо через случайное соответствие между аминокислотами и кодонами, то есть из «застывшей случайности».

Происхождение трансляции: ключевые идеи и модели [\[129\]](#)

За 40 лет, прошедших со времени открытия механизма трансляции и расшифровки генетического кода, были предложены многочисленные – неизбежно спекулятивные, иногда надуманные и часто весьма остроумные теоретические модели происхождения и эволюции различных компонентов трансляционной системы и разных аспектов процесса трансляции. Представить здесь тщательный критический анализ этих моделей нереально. Вместо этого мы рассмотрим несколько центральных идей, имеющих прямое отношение к происхождению трансляции, а затем обсудим несколько более подробно единственные два известных мне убедительных сценария.

Главное общее положение эволюции трансляции состоит в том, что отбор на синтез белков не мог быть основной движущей силой эволюции трансляционной системы. Чтобы эта сложная система эволюционировала дарвиновским путем, необходимы многочисленные шаги, а белки появляются лишь на последних из них; до этого момента организм «не знает», какие эволюционные преимущества может нести с собой использование белков. Как отмечалось в главе 9, существует много ситуаций, в которых эволюция как будто обладает возможностью некоего предвидения; эти случаи, однако, основаны фактически на экстраполяции, тогда как в случае трансляции нет еще ничего, из чего можно были бы экстраполировать.

Возникновение сложного механизма трансляции путем случайного дрейфа также не реалистично – во всяком случае, в рамках стандартных представлений эволюционной биологии (см. обсуждение в конце этой главы). Таким образом, единственный возможный путь для появления трансляции, по-видимому, *экзаптация: промежуточные этапы эволюции трансляционной системы должны быть отобраны для иных функций, нежели синтез белка.* Различные сценарии возникновения трансляции исходят из различных предположений о природе экзаптируемых функций.

Простая и потенциально плодотворная идея состоит в том, что в мире РНК аминокислоты и пептиды функционировали в качестве кофакторов рибозимов. Сатмари первым разработал гипотезу, основанную на этом предположении, и предложил, что «кодирующие коферментные олигонуклеотиды» (ККО; олигонуклеотиды-рибозимы с различными каталитическими активностями, использующие аминокислоты в качестве кофакторов) могли быть эволюционными предшественниками тРНК (Szathmary, 1993, 1999). Предполагается, что ККО собирались с участием содержащихся в них протоантикодонов на эволюционных предшественниках мРНК, однако детали этой стадии остаются неясными. Гипотеза ККО связывается с представлением о том, что тРНК возникла в ходе эволюции в результате двух последовательных дупликаций связывающих аминокислоты шпилек. Вариант гипотезы ККО,

предложенный Р. Найтом и Л. Ландвебер, включает эволюцию аминокацилирующих рибозимов (возможность, хорошо подтверждаемая экспериментальными данными, см. табл. 12-1) и возникновение безматричного, опосредованного рибозимами синтеза пептидов в качестве промежуточного этапа эволюции трансляции (Knight and Landweber, 2000).

Майкл Ярус (Yarus, 1998) предложил модель прямого связывания РНК в качестве альтернативы ККО как сценария происхождения трансляции. Согласно этой модели, изначальной формой взаимодействия аминокислот и прото-тРНК было непосредственное связывание, предположительно через кодонные триплеты. Затем непосредственное связывание вытеснилось адапторным механизмом, вероятно, при участии аминокацилирующих рибозимов, как в варианте ККО Найта – Ландвебер.

Взяв за основу гипотезу ККО, мы с Юрием Вольфом разработали обобщенную, но детализированную модель возникновения трансляционной системы в мире РНК (Wolf and Koopin, 2007). Эта модель включает в себя как дарвиновский отбор, так и аспекты конструктивной нейтральной эволюции (см. гл. 8), наряду с экзапацией и субфункционализацией.

Отправной точкой всех сценариев происхождения трансляции является реплицирующийся ансамбль эгоистичных кооператоров, состоящий из молекул РНК с различными рибозимными активностями и существующий в сети неорганических ячеек (см. дальнейшее обсуждение в следующем разделе). Эти рибозимы исполняют, в числе прочих функций, и функцию репликасы; вероятно, представлены и другие функции, такие, например, как синтез предшественников РНК. Наш эволюционный сценарий включает в себя следующие этапы (см. рис. 12-4).

1. Рибозим является частью ансамбля эгоистичных кооператоров в ячейке. Этот рибозим достаточно сложен для катализа реакции ($X \rightarrow Y$), скорость которой влияет на приспособленность ансамбля, и имеет определенное число позиций, способных к эволюции, так что возможна эволюция новых функций. Две или более абиогенных аминокислоты, присутствующие в ячейке, связываются с R . Избирательное связывание аминокислот обеспечивается активным центром, случайно присутствующим в R . Участие стереохимического протокода (кодон/антикодон) на данном этапе возможно, но не повлияет на ситуацию существенным образом. Присоединенные аминокислоты стимулируют реакцию $X \rightarrow Y$, катализируемую R . *In vitro* были получены рибозимы, сильно стимулируемые пептидами что дает экспериментальное обоснование этому принципиальному шагу (Robertson et al., 2004). В контексте эгоистично-кооперативной эволюции (см. гл. 11) естественный отбор будет отбирать аминокислоты, стимулируемые R , приводя к постепенному совершенствованию пространственного выравнивания аминокислот на R и отбору последовательности и структуры оптимальных для связывания аминокислот.

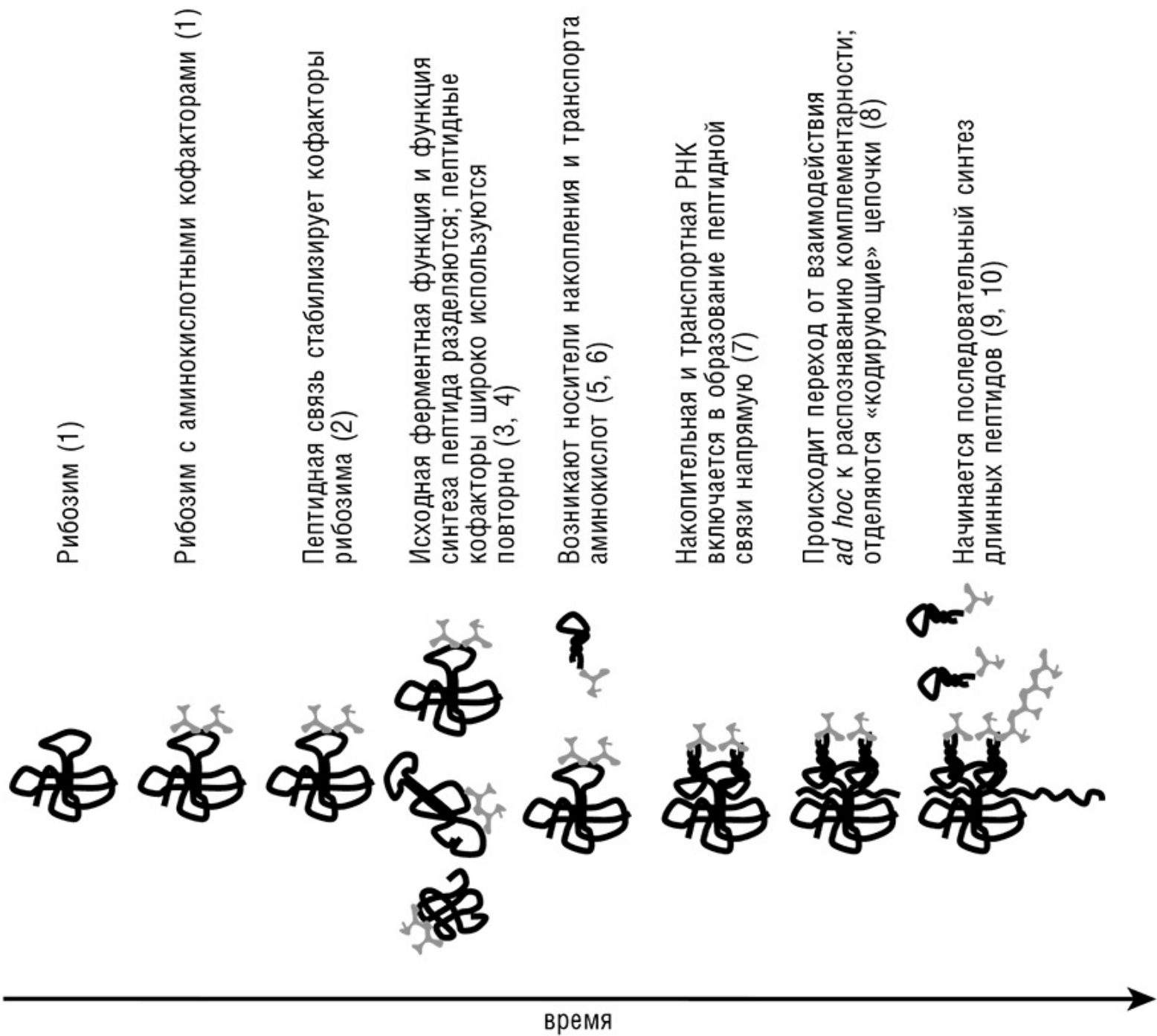


Рис. 12-4. Концептуальный сценарий происхождения трансляционной системы посредством экзэптации и субфункционализации. Шаги модели, описанные в тексте, обозначены цифрами в скобках.

2 . Rприобретает дополнительную активность лигазы пептидной связи, формируя олигопептид Риз соседних аминокислот, связанных с R. Отбором *in vitro* были получены рибозимы с высокой активностью пептидной лигазы, хотя и с низкой избирательностью. По-видимому, рибозимы этого класса способны синтезировать только короткие пептиды, состоящие из, самое большее, четырех или пяти аминокислот. Селекционным преимуществом этого новоприобретения будет повышение стабильности реактивного комплекса, приводящее к дальнейшему усилению реакции $X \rightarrow Y$. Естественно задаться вопросом, откуда на этом шаге берется энергия, необходимая для формирования пептидной связи. В экспериментально описанных рибозимных пептидных лигазах один из субстратов является активированным

производным (аминоацил-аденилат), так что используется энергия эфирной связи. Это напоминает современную трансляцию, в которой АРСазы используют аминоацил-аденилаты для аминоацилирования специфической тРНК, а высокоэнергетичная эфирная связь аминоацил-тРНК используется для транспептидации. Гипотетические древнейшие пептид-лигазы, возможно, действовали таким же образом, используя аминоацил-аденилаты или другие активированные производные аминокислот, произведенные другими рибозимами. И действительно, были получены рибозимы, катализирующие каждую из этих реакций, от аденилирования аминокислот до синтеза пептидов (см. табл. 12-1). Эти рибозимы, несомненно, зависят от энергии фосфодиэфирной связи в АТФ или иной формы энергии.

3. Спонтанная диссоциация или распад R высвобождает пептид P обратно в ячейку. Если R обладает неспецифической способностью стимулировать и (или) стабилизировать рибозимы, он может быть захвачен другим рибозимом E , катализирующим другую реакцию ($U \rightarrow V$). Интересным примером мог бы быть пептид, содержащий пару отрицательно заряженных аминокислот и образующий комплекс с двухвалентным катионом, аналогично разнообразным, неродственным современным ферментам метаболизма нуклеиновых кислот (полимеразы, нуклеазы, лигазы, топоизомеразы, и др.). Если R повышает каталитическую активность E , он снова увеличивает приспособленность всего ансамбля.

4. В то время как активность E по-прежнему зависит от наличия P , копия R (R_L) может потерять исходную функцию катализа $X \rightarrow Y$ при сопутствующем усилении функции аминокислотной лигазы, в то время как другая копия (R_0) сохраняет исходную функцию, все еще усиливаемую пептидом P . Заметим, что это типичная субфункционализация, основной путь эволюции дублированных генов в современных геномах (см. гл. 8). Субфункционализация, возможно, была важна уже в мире РНК, когда выгода усиленного катализа R_0 и E перевешивала увеличение затрат на репликацию.

5. Повсеместный катализ при помощи пептидов в разделенной на ячейки добиологической системе делает аминокислоты ценным ресурсом для эволюционирующих эгоистичных кооперативов. Учитывая, что аминокислоты являются небольшими полярными молекулами, способными диффундировать сквозь стенки ячеек, накопление аминокислот в ячейке должно было быть полезным. Таким образом, связывающие аминокислоты малые РНК (T) развиваются под эволюционным давлением в сторону накопления аминокислот; эти молекулы могут рассматриваться как аналоги связывающих аминокислоты аптамеров (см. предыдущий раздел). Первоначально РНК T связывают аминокислоты неспецифически. Затем постепенно эволюционирует автокаталитическое аминоацилирование 3'-конца РНК T , что приводит к увеличению сродства к аминокислотам и избирательности в их связывании. Как и в случае пептид-лигазы на шаге 2, этой реакции необходим источник энергии; в этом качестве выступают активированные производные аминокислот, такие как аминоацил-аденилаты.

6. Различные виды РНК T , избирательно связывающие разные аминокислоты, эволюционируют путем дубликации и диверсификации, с сохранением вариантов под давлением отбора в сторону эффективного накопления широкого арсенала аминокислот. Детали связывания аминокислот РНК T будут различаться в зависимости от того, принимается ли гипотеза избирательного распознавания аминокислот специфическими (анти)кодонами. Если такого избирательного распознавания нет, то рассматривается сценарий «застывшей случайности», при котором сайт связывания в РНК T не имеет сродства к кодону или антикодону, а последовательность экспонированной петли (предтечи антикодонной петли) случайна. Независимо от конкретной модели (даже если принимается застывшая случайность), данный шаг, устанавливающий соответствие между аминокислотами и тринуклеотидами, является

критически важным для становления генетического кода.

7. Рибозим R_L развивает способность связывать комплексы аминоксил – РНК T , а не отдельные аминокислоты, что приводит к большей стабильности и пространственной точности связи. Главная биохимическая активность R_L смещается от лигирования аминокислот к транспептидации (передача растущего пептида от одного вида РНК T_k другому), что приводит, благодаря высокой энергии связи аминоксил-РНК, к увеличению выхода пептидов. Примечательно, что 50S субъединица бактериальной рибосомы, в качестве предка которой предполагается рибозим R_L , может катализировать реакцию транспептидации со скоростью, сравнимой со скоростью полной рибосомы (Wohlgemuth et al., 2006).

8. Эволюционирует вспомогательная субъединица РНК R_S под давлением отбора в сторону повышения эффективности связи и точности расположения комплекса аминоксил- T_n а R_L . Механизм распознавания РНК T переходит от слабоизбирательного взаимодействия между РНК T_i и R_L к избирательному спариванию оснований между протоантикодонной петлей T_i и РНК R_S . Этот шаг является решающим в возникновении полноценной трансляции, механизма, основанного на адаптерах (прото-тРНК, РНК T_v этой модели), сопрягающих аминокислоты с соответствующими им кодонами.

9. Поскольку происхождение тРНК всех специфичностей от единого предка очевидно, эволюционный путь от набора примитивных РНК T_k современным тРНК требует специального объяснения. На описанных выше ранних этапах эволюции системы трансляции различные виды РНК T могли эволюционировать почти параллельными конвергентными путями. Тем не менее общее происхождение тРНК подразумевает последующее «бутылочное горлышко», через которое прошел только один победитель, молекула в форме «L» с акцепторным триплетом С–С–А на 3'-конце. Давление отбора при этом эволюционном «захвате» могло происходить в сторону пространственной комплементарности и усиленного взаимодействия между аминокислированной РНК T_i пептидил-трансферазой R_L . Такой отбор изначально действовал на единственную РНК T , возможно имевшую сродство к наиболее распространенной аминокислоте. Впоследствии остальные тРНК должны были эволюционировать путем дубликации и специализации.

10. Следующим шагом в эволюции системы трансляции могло быть физическое отделение матричной цепи M от R_S , в результате чего произошло дальнейшее разделение функций кодирования и катализа. В этот момент нить M освобождается от эволюционных ограничений, связанных с функциями катализа и связывания в первичной трансляции, поскольку эти функции перешли на физически различные молекулы РНК R_L и R_S и прото-тРНК. Единственным требованием к M остается ее способность принимать растянутую конформацию для размещения спаренных оснований кодона и антикодона при связывании аминоксил- тРНК. Эволюционные преимущества такого разделения очевидны: промежуточный ассоциат $R_S R_L$ (который, на данный момент, можно обоснованно назвать *проторибосомой*) в присутствии в ячейке различных олиго- и полинуклеотидов обеспечит синтез все большего разнообразия пептидов, расширяя, таким образом, каталитические возможности ансамбля. Кроме того, этот шаг позволяет отбору действовать в сторону увеличения потенциала репликации (в частности, появления высокоспецифичных сайтов узнавания репликазы) тех видов M , которые кодируют полезные пептиды, приводя к повышению концентрации этих видов РНК в ячейке. По сути дела, в эгоистичном кооперативе запускается разновидность цикла Дарвина – Эйдена.

11. Освобождение (прото)тРНК из $R_S R_L$ при транспептидации вызывает трехнуклеотидный сдвиг, характерное движение современной рибосомы, позволяющее

синтезировать длинные пептиды, или, фактически, первые белки. *Это и есть белковый прорыв.*

В соответствии с таким эволюционным сценарием путь от белкового прорыва к трансляционной системе современного типа есть по большей части история принятия на себя белками функции первичных рибозимов. Каталитические возможности и способности к связыванию у белков несравненно выше, нежели у РНК и пептидов, и в процессе эволюции они постепенно, но необратимо вытеснили в этом качестве первичные рибозимы.

Теперь нам предстоит рассмотреть существенно иную эволюционную модель, первоначально обрисованную Анатолием Альтштейном (Altstein, 1987), а затем независимо и более полно разработанную Энтони Пулом и соавторами (Poole et al., 1998). В этой модели рибосомы и механизм трансляции производятся от древнего рибозима-репликазы. В модели постулируется, что проторибосома изначально функционировала как «трипликаза», сложный рибозим, объединяющий функции РНК-полимеразы и РНК-лигазы и синтезировавший молекулы РНК, комплементарные матрице, тринуклеотидными шагами. Эта «трипликазная» проторибосома способствует сборке тРНК-подобных молекул (аналогично РНК Тв предыдущей модели) на матрице РНК путем спаривания оснований (прото)антикодонов с комплементарными триплетами (кодонами) матрицы, отщепления остальной пре-тРНК и соединения (лигирования) соседних триплетов. Пул с соавторами считают правдоподобным механизм репликации на основе комплементарного взаимодействия тринуклеотидов (в противоположность мононуклеотидам) с матрицей, поскольку каталитическая эффективность рибозимов мала. Комплекс РНК-матрицы с комплементарным тринуклеотидом будет сохраняться намного дольше, чем комплекс с мононуклеотидом, оставляя трипликазе больше возможностей лигировать соседние триплеты. Механизм трипликазы кажется особенно правдоподобным в свете опытов Фредерика и Ноллера, показавших, что мРНК пропускается через рибосому тринуклеотидными шагами, с согласованным движением тРНК, но без участия трансляционных факторов (Frederick and Noller, 2002).

Переход от трипликазы к современному типу трансляционно-репликационной системы требует нескольких сложных этапов, а именно появления генетического кода (в данном случае на уровне аминокислотной избирательности прото-тРНК) и обратной связи между трансляцией и репликацией РНК (происхождение белковых РНК-полимераз или белковых кофакторов рибозима-полимеразы). Кроме того, необходима субфункционализационная стадия, на которой трипликаза порождает отдельные проторибосомы и репликазы, а последняя переходит от лигирования триплетов к обычному механизму репликации с присоединением нуклеотидов по одному.

Скептический обзор моделей эволюции репликации и трансляции

В предыдущих разделах мы обрисовали современное состояние гипотезы мира РНК и подробнее обсудили происхождение репликации и трансляции. Зададимся теперь простым прямым вопросом: убедительны ли имеющиеся свидетельства в поддержку любой из этих моделей. Само собой, вопрос уже подразумевает отрицательный ответ. У нас есть определенные косвенные указания, пусть даже и далекие от полного согласованного сценария самых ранних этапов эволюции передачи биологической информации. Во-первых, не будем забывать о логической неизбежности мира РНК: какой еще могла быть отправная точка эволюции системы трансляции? Во-вторых, сравнительный анализ составных частей системы трансляции, несомненно, указывает на гораздо большую роль РНК в древней трансляции по сравнению с современной системой, в частности что РНК играла решающую роль в установлении соответствия аминокислот кодонам. В-третьих, рибозимы впечатляют своей каталитической

универсальностью и эффективностью (хотя и уступают в этом белкам). Тридцать лет назад не было известно ни одной каталитической функции молекул РНК, а теперь описаны десятки видов активности рибозимов, включая некоторые из ключевых реакций трансляции, такие как высокоэффективное аминокислотирование.

На этом, к сожалению, хорошие новости заканчиваются, а остальное напоминает скорее отрезвляющий ледяной душ. Невзирая на все достижения «рибозимологии», ни одна рибозим-полимераза даже не приближается к уровню эффективности, который необходим, чтобы всерьез рассматривать репликационные системы, состоящие из одних РНК, в качестве ключевого промежуточного звена в эволюции жизни. Ни один рибозим не способен катализировать синтез нуклеотидов или даже сахаров, входящих в их состав. Даже если мы закроем глаза на эти проблемы, путь от предполагаемого мира РНК к системе трансляции невероятно труден. Общие представления касательно функций абиогенных аминокислот и, возможно, пептидов в мире РНК, такие как роль кофакторов рибозимов (см. обсуждение в предыдущих параграфах), оказываются плодотворными и совместимы с экспериментальными данными. Тем не менее разбиение эволюции системы трансляции на последовательные шаги, каждый из которых давал бы биологически объяснимое селективное преимущество, чрезвычайно затруднительно даже в качестве спекулятивной схемы, не говоря уже об экспериментальной проверке. Гипотезы трипликазы и проторибосомы привлекательны как попытка одновременно объяснить происхождение репликации и трансляции в одном цикле, но реалистичен ли подобный сценарий? Сама трипликаза должна быть чрезвычайно сложной, замысловатой молекулярной машиной, и это вызывает подозрение, что, невзирая на всю ее привлекательность, трипликаза может быть не самым правдоподобным решением вопроса о происхождении трансляции.

Итак, принимая во внимание все соображения, моя оценка текущего состояния дел в области изучения происхождения репликации и трансляции довольно мрачна. Несмотря на содержательные теоретические модели и наводящие на размышления экспериментальные результаты, мы в настоящее время не имеем правдоподобного решения этих вопросов и даже не видим пути к такому решению. Конечно, область исследования рибозимов молода, и есть основания ожидать, что достаточно скоро в ней будет достигнут значительный прогресс. Тем не менее ближе к концу этой главы мы обсудим радикальную альтернативу. Сначала, однако, мы должны рассмотреть область самой науки о происхождении жизни, и прежде всего ее химические, геологические и геохимические аспекты. Самоочевидно, что в столь кратком разделе мы будем весьма поверхностны и очертим лишь некоторые ключевые идеи и наработки.

Происхождение жизни с точки зрения химии и геохимии

Происхождение жизни сделалось научной проблемой с тех пор, как Луи Пастер продемонстрировал невозможность самопроизвольного зарождения живых форм. По удивительному совпадению, результаты его экспериментов были опубликованы в 1859 году, в тот же год, когда Дарвин издал «Происхождение видов...», где, в числе других основополагающих идей, предложил и гипотезу об общем предке всех живых организмов. Повидимому, первое последовательное представление о происхождении жизни было изложено в 1924 году русским биохимиком Александром Ивановичем Опариным в виде популярной брошюры (Oparin, 1924), впоследствии многократно переизданной в постоянно расширяемых вариантах (Oparin and Fesenkov, 1956). Сценарий Опарина был наивным и мало обоснованным (при менее доброжелательном отношении его даже можно назвать химически несостоятельным), но он включал некоторые ключевые положения, которые принимаются в области происхождения жизни и по сей день ^[130]. Центральное предположение Опарина состояло в том, что среда, в которой возникла жизнь (предположительно примитивный океан в целом, но, возможно, и какая-то разновидность дарвиновского «теплого пруда»), представляла собой сложный раствор абиогенных органических молекул, включая аминокислоты и сахара, – все мономеры, необходимые для синтеза биополимеров. Опарин назвал эту гипотетическую среду, породившую жизнь, *первичным бульоном* (и л и супом). Подобный сценарий возникновения жизни был предложен позже в статье Дж. Б. С. Холдейна, уже цитированной в главе 11 по поводу гипотезы о «вирусной» стадии, предшествовавшей появлению первых клеток (Haldane, 1928). Опарин и Холдейн утверждали, что атмосфера первичной Земли была восстановительной и что первыми организмами были анаэробные гетеротрофы, питавшиеся смесью мономеров, изобиловавших в гипотетическом первичном бульоне.

Опарин и другие ранние исследователи происхождения жизни осознавали важность появления клеточной (или подобной) организации на ранней стадии жизни или даже до появления первых *bona fide* форм жизни. Согласно популярной тогда идее, доклеточная эволюция началась с так называемых коацерватных капель, образовавшихся в результате взаимодействия между определенными противоположно заряженными полимерами. Опарин и его соратники и последователи были биохимиками по профессии и, таким образом, отдавали в ходе своих размышлений приоритет обмену веществ. Ранние сценарии происхождения жизни придерживались того, что в коацерватах возникла простая сеть (до)метаболических реакций, обеспечивавшая обратную связь процессам роста и, возможно, деления этих пузырьков. С этого момента развивался все более сложный метаболизм, приведший в конечном итоге к появлению автотрофии. В этих качественных моделях доклеточной эволюции не уделялось никакого серьезного внимания происхождению генетической информации и процессов информационного обмена. Эти процессы (не вполне еще понятые к тому времени), как предполагалось, *каким-то образом* развились как следствие или побочный продукт эволюции обмена веществ. Некоторые ограниченные опыты с коацерватными каплями и другими подобными пузырьками, такими как так называемые микросферы, состоящие из беспорядочных полимеров аминокислот, известных под названием протеиноидов, продемонстрировали способность этих пузырьков поддерживать сеть простых реакций, расти и делиться, но не более того (Fox, 1976).

Не следует относиться к ранним гипотезам о происхождении жизни слишком пренебрежительно. Создатели этих качественных моделей были полностью рациональны в своем мышлении и осознавали важность обмена веществ, источников энергии и деления среды

на ячейки, напоминающие клетки. Тем не менее они крайне недооценили или даже просто проигнорировали полную практическую невозможность существования «первичного бульона» в объеме всего древнего океана. Любой сколь-нибудь реалистичный сценарий происхождения жизни должен включать четко определенную *доклеточную, абиогенную ячеистую структуру*, неорганические катализаторы для ускорения «предбиохимических» реакций, предшествовавших появлению настоящих ферментов, градиенты температуры и (или) электрохимического потенциала, необходимые для выработки энергии в доступной форме, и решение чрезвычайно сложной проблемы о происхождении генетической информации (см. обсуждение выше в этой главе). В общем, ранние концепции недооценивали масштаб проблемы происхождения жизни и не исследовали специфические абиогенные условия, необходимые в качестве предпосылки начала биологической эволюции. Впоследствии несколько групп исследователей попытались уйти от понятия однородного первичного супа, заменив его той или иной формой неорганических ячеек, и стремились обратиться к проблеме происхождения жизни в целом, сочетая моделирование, эксперимент и наблюдения в природных условиях. Общей идеей этих гипотез является наличие единой структуры, которая может одновременно обеспечивать компартментализацию, градиенты энергии и катализаторы. Мы не можем здесь детально обсудить все эти исследования, не говоря уже обо всей области пребиотической химии. Вместо этого мы сосредоточимся лишь на нескольких моделях, представляющихся особенно продуктивными.

Г. Вахтершаузер предположил, что жизнь возникла в зонах вулканической активности, на поверхности, богатой железом, никелем, кобальтом и другими переходными металлами, каталитически активными и способствовавшими фиксации CO_2 , что приводило к росту органических сверхструктур. Были получены некоторые многообещающие результаты по синтезу пептидов в древних условиях, постулируемых в этой модели (Huber et al., 2003; Wächtershäuser, 1997). Таким образом, еще до того, как возникли организмы, могла существовать некая форма неорганической хемоаутотрофии. Синтезированные органические молекулы могли способствовать реакциям, катализируемым неорганическими катализаторами, приводя к химическому отбору, предвещавшему отбор биологический. Клеточная организация и генетические механизмы, согласно этому сценарию, появились в результате первичной химической эволюции.

Как уже говорилось в главе 11, Майкл Рассел и коллеги разработали, по всей вероятности, наиболее реалистичные, последовательные геохимические основы для возникновения жизни и доклеточной эволюции (Martin et al., 2008; Russell, 2007; Russell and Hall, 1997), опираясь на предшествующие исследования Дж. Бароса (Baross and Hoffman, 1985). Основная идея заключается в том, что оптимальные условия для возникновения жизни обеспечивались теплыми (но не горячими) щелочными источниками на дне первичного океана. Согласно Расселу и Холлу, гидротермальные источники образуют непрерывно-проточные реакторы, производящие насыпи осажденных карбонатов, кремния, глины и железо-никелевых сульфидов (см. рис. 12-5). Первоначально Рассел и Холл предсказали существование и размеры таких насыпей из геохимических соображений (Russell and Hall, 1997). Это предсказание было блестяще подтверждено открытием в области Срединно-Атлантического хребта образования, получившего название Потерянного города (*Lost City*) из-за украшающих его гигантских карбонатных шпилей (Kelley et al., 2005). Структура этих шпилей поистине замечательна: по сути, они являются сетью неорганических ячеек, образованных прежде всего сульфидами железа-никеля, способных катализировать различные органические реакции и поддерживающих постоянный поток протонов (протонодвижущая сила), который может обеспечивать энергией идущие в ячейках реакции. Как уже говорилось в главе 11, сети неорганических ячеек могли

быть инкубатором жизни, от простейших реакций органического синтеза до появления и выхода из них первых клеток.

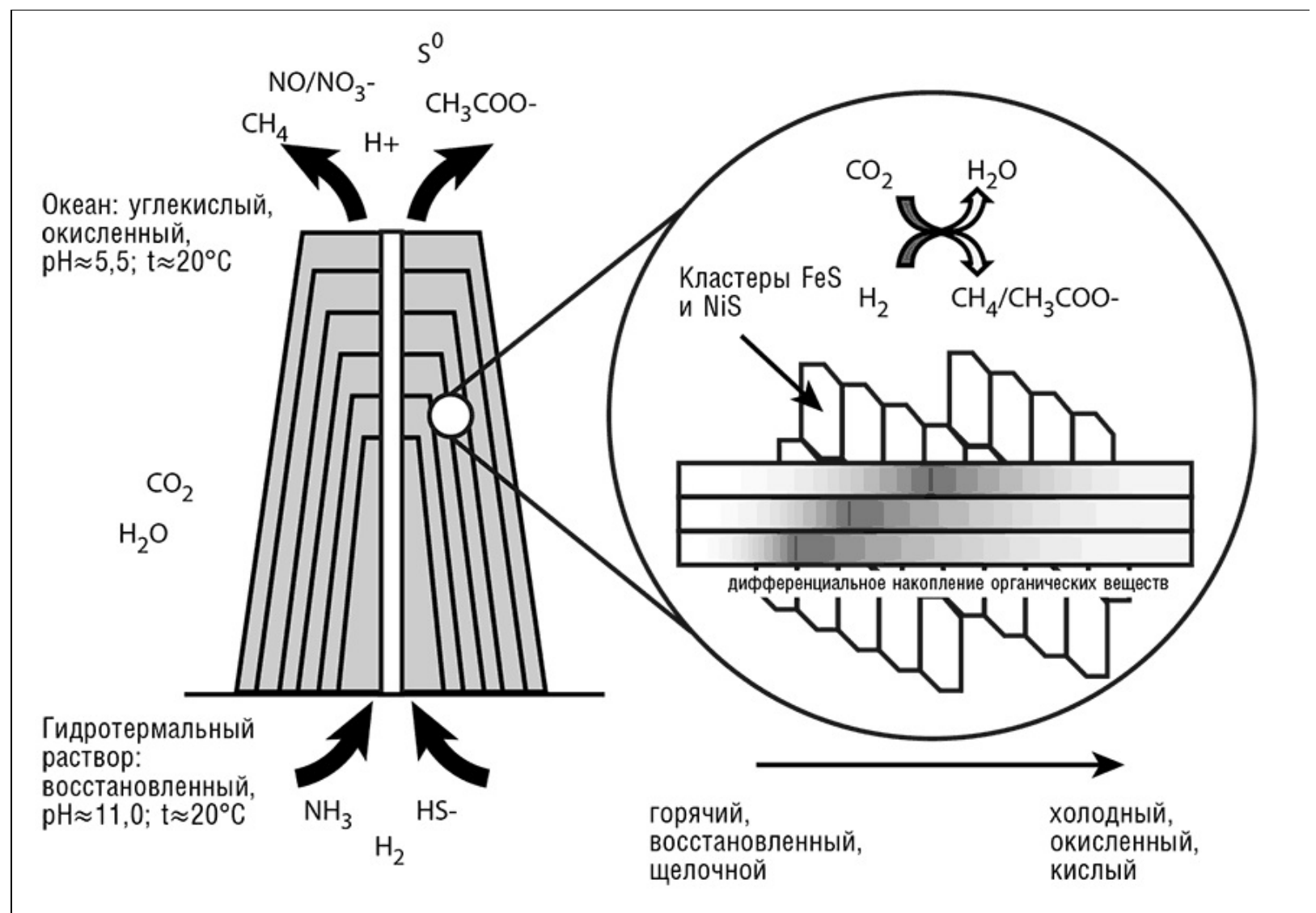


Рис. 12-5. Сети неорганических ячеек: проточные реакторы первичной химии и биохимии. В основном по данным Martin and Russell, 2007.

С целью анализа свойств этих сетей и составляющих их ячеек было проведено значительное число экспериментов и построено множество математических моделей. Было показано, что мембраны сетей полупроницаемы, так что сложные органические молекулы, включая азотистые основания, аминокислоты, сахара и жирные кислоты, оказываются изолированными в ячейках, в то время как небольшие органические молекулы и радикалы, такие как ацетат и метан, свободно сквозь них диффундируют (Mielke et al., 2010). Таким образом, ячейки представляются идеальной средой для различных реакций органического синтеза, от самых простых и фундаментальных, таких как восстановление CO_2 молекулярным водородом с получением ацетата, до сложных реакций, например синтеза полинуклеотидов и пептидов, – иными словами, «происхождения биохимии» (Martin and Russell, 2007). Опытов по химии неорганических сетей пока проведено недостаточно, но компьютерное моделирование показывает, что при наличии температурного градиента, неизбежно представленного в гидротермальных источниках, может быть достигнута чрезвычайно высокая концентрация малых молекул и полимеров (Baaske et al., 2007). Такая концентрация органических молекул способна существенно облегчить различные реакции, маловероятные в ином случае, включая

полимеризацию нуклеотидов и лигирование РНК (Koonin, 2007c). Хотя мембраны неорганических ячеек (Рассел заходит столь далеко, что называет их протоклетками), по всей видимости, непроницаемы для биологически важных мономеров, сами эти мембраны относительно неустойчивы (словами Рассела, «хлипкие мембраны»), так что ячейки должны разрушаться и соединяться вновь, распространяя свое содержимое. Как мы уже говорили в главе 11, в результате, вероятно, происходит отбор на уровне ячеек.

При изучении событий, случившихся лишь однажды, притом около 4 миллиардов лет назад, важно не держаться одной-единственной модели, а, так сказать, распределять свои ставки. Армен Мулкиджанян предложил существенно отличную модель неорганических ячеек, приписав роль инкубаторов самых ранних форм жизни другому виду отложений, пребывающих на меньших глубинах и при более низких температурах и состоящих в основном из сульфида цинка (Mulkidjanian, 2009). В этой модели источником энергии органических реакций является ультрафиолетовое излучение, достигающее тех небольших глубин, где находятся отложения сульфида цинка. Кроме того, сульфид цинка значительно превосходит сульфид железа в качестве катализатора и в то же время является менее реактивным химическим веществом, не оказывающим разрушительного влияния на такие лабильные молекулы, как РНК [\[131\]](#).

Несмотря на то что модели Вахтершаузера, Бароса, Рассела и Мулкиджаняна отличаются существенными аспектами, они объединены общими чертами, более важными, чем их различия. В каждой из этих моделей жизнь образовалась в особых местах обитания и в особых условиях, в которых неорганические ячейки способствуют накоплению органических молекул в высокой концентрации, обладают каталитическими поверхностями и обеспечивают градиент усвояемой энергии, способной поддерживать первичный органический синтез. Эти общие признаки, вероятно, должны быть частью любого стремящегося к реалистичности сценария происхождения жизни.

Радикальная альтернатива: космология вечной инфляции, переход от случайности к биологической эволюции в истории жизни и переоценка роли крайне редких событий в эволюции [132]

Как уже отмечалось выше, ситуация в области происхождения жизни в целом видится довольно мрачной. Даже при (весьма нетривиальном) предположении, что мономеры, такие как НТФ, изобильны и доступны, проблема синтеза достаточно стабильных полимеров с регулярной структурой (РНК) чудовищно сложна, а проблема происхождения репликации и трансляции таких первоначальных молекул РНК еще сложнее. В этой книге неоднократно подчеркивалось, что эволюция путем естественного отбора и дрейфа может начаться только после возникновения достаточно точной репликации, причем даже на этой стадии эволюция трансляции остается весьма проблематичной.

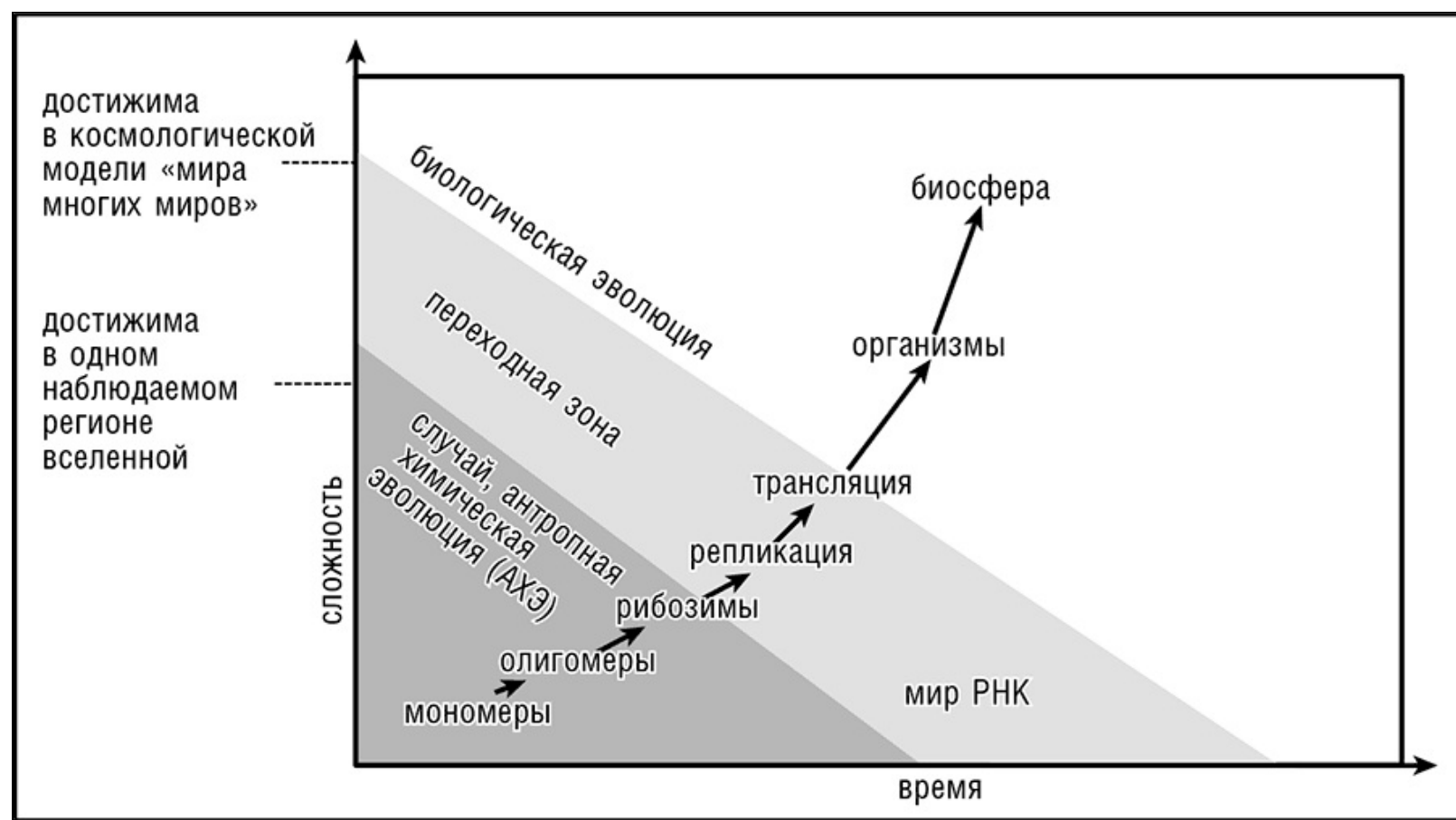


Рис. 12-6. Добиологическая и биологическая стадии происхождения жизни: переход от антропной причинности к биологической эволюции.

Появлению первой репликаторной системы, ознаменовавшему собой «дарвиновский прорыв» [133], неизбежно предшествовала череда сложных, трудных шагов, на которых биологические эволюционные механизмы еще не действовали (см. рис. 12-6). Даже с учетом среды, способной облегчить эти процессы, такой как сети неорганических ячеек в гидротермальных источниках, умножение вероятностей этих шагов делает появление первых репликаторов ошеломляюще невероятным (см. прил. II).

Глубокие трудности проблемы происхождения жизни могут показаться непреодолимыми,

что заставляет задаться весьма общими вопросами, выходящими за рамки биологии. Не существует ли скрытых от нас факторов, имеющих решающее значение для зарождения жизни, которые, будь они нам сейчас известны, существенно изменили бы картину и сделали бы возникновение жизни гораздо более вероятным? Возможно ли, что процессы, формирующие основу для возникновения жизни, действительно так сложны, как мы себе представляем, но число попыток настолько велико, что появление жизни в результате одной или нескольких из них возможно или даже неизбежно? Другими словами, возможно ли, что сами наши понятия о вероятности неполны?

Первая возможность имеет отношение к нахождению условий, существовавших на первичной Земле и каким-то образом сделавших происхождение жизни «легким». Ячейки Рассела продвигают нас на этом пути, но, видимо, недостаточно далеко: даже в подобных проточных реакторах, богатых энергией и катализаторами, сочетание всех необходимых процессов было бы крайне редким.

Вторая возможность может быть рассмотрена в контексте всей Вселенной. Зададимся вопросом, сколько существует планет с благоприятными условиями для возникновения жизни, иными словами, сколько попыток возникновения жизни всего было сделано. В данном разделе мы рассмотрим эту линию рассуждений с точки зрения современной физической космологии.

В XX веке космология прошла драматический путь от странного (и не особенно авторитетного) философского направления к полноценному, быстро развивающемуся разделу физики, прочно опирающемуся на многочисленные научные наблюдения. Ведущее направление в сегодняшней космологии сосредоточивается на так называемой инфляции, периоде экспоненциально быстрого первоначального расширения Вселенной (Carroll, 2010; Guth, 1998a; Guth and Kaiser, 2005; Vilenkin, 2007). В наиболее правдоподобных, самосогласованных моделях инфляция вечна, с бесконечным числом островных вселенных (или просто вселенных), возникающих через распад небольших регионов изначального океана ложного (высокоэнергетического) вакуума и составляющих бесконечную мультивселенную (см. прил. II). Модель «мира многих миров» (МММ) делает поразительное предсказание, что все макроскопические, «крупнозернистые» истории событий, не запрещаемых физическими законами сохранения, были реализованы (или, возможно, будут реализованы) где-то в бесконечной мультивселенной, и не только однажды, но бесконечное число раз (Garriga and Vilenkin, 2001; Vilenkin, 2007). Например, существует бесконечное число (макроскопически) точных копий Земли, со всем, что существует на ней, хотя вероятность того, что данная наблюдаемая область Вселенной содержит одну из этих копий, исчезающе мала. Эта картина выглядит крайне нелогичной, и даже сумасбродной, но таково прямое следствие вечной инфляции, доминирующей модели эволюции Вселенной в современной космологии.

Модель МММ тесно связана с антропным принципом (иногда называемым антропным отбором), спорной, но мощной и популярной среди космологов концепцией (Barrow and Tipler, 1988; Carter, 1974; Livio and Rees, 2005). Согласно антропному принципу, единственная причина, по которой наша Вселенная имеет свои характерные параметры, состоит в том, что в противном случае в этой Вселенной не было бы никаких наблюдателей [\[134\]](#). В такой форме так называемый «слабый» антропный принцип может быть реалистично определен только в контексте огромной (или, еще лучше, бесконечной) мультивселенной. В модели МММ антропный отбор имеет простой смысл: среди огромного числа наборов параметров, существующих в мультивселенной (в бесконечном числе экземпляров каждый), наша Вселенная должна иметь именно такие параметры, которые способствуют возникновению и поддержанию сложных форм жизни. Иногда говорят, что наша Вселенная принадлежит «биофильной

области» мультивселенной (Livio and Rees, 2005). Сам термин «антропный принцип» можно считать неудачным, так как он может быть истолкован как предполагающий некую особую важность людей или, в более общем случае, сознательных наблюдателей или, что еще хуже, привлечь телеологические интерпретации. Такой взгляд на антропный принцип совершенно неверен; на самом деле он представляет собой просто «отбор наблюдателя» (Bostrom, 2002). Тот факт, что в этой Вселенной имеется жизнь, строго ограничивает ее характеристики: по меньшей мере, наша часть Вселенной должна содержать галактики и планетные системы, а не только массивные черные дыры или частицы разреженного газа, являющиеся в принципе более вероятными состояниями с высокой энтропией.

В сравнении со старыми космологическими концепциями, рассматривавшими конечную Вселенную, модель МММ меняет самые определения *возможного, вероятного и случайного* по отношению к любому историческому сценарию. Проще говоря, вероятность реализации любого сценария, не запрещенного законами сохранения, в бесконечной мультивселенной в точности равна единице. С другой стороны, вероятность того, что данный сценарий реализуется в данной (островной) вселенной, равна частоте этого сценария в мультивселенной и может быть исчезающе малой. С несколько иной точки зрения известная идея второго начала термодинамики, будучи истинной лишь в статистическом смысле, обретает буквальный смысл в бесконечной мультивселенной: любое нарушение второго начала, дозволенное законами сохранения, действительно случилось, причем бесконечное число раз. Таким образом, *спонтанное возникновение сложных систем, которое могло бы считаться практически невозможным в конечной вселенной, становится не только возможным, но и неизбежным в МММ*, хотя априорная вероятность подавляющего большинства историй, происходящих в данной вселенной, исчезающе мала. Эта новая сила случайности, подкрепленная антропным рассуждением, имеет глубокие последствия для нашего понимания любого явления во Вселенной, не исключая и жизнь на Земле (Koonin, 2007b).

История жизни не может не включать в себя важнейший переход от *случая биологической эволюции* (см. рис. 12-6). Синтез нуклеотидов и полинуклеотидов по крайней мере среднего размера не мог возникнуть биологически и, следовательно, произошел абиогенным путем, то есть фактически случайно, при помощи химического отбора, как, например, выживание более стабильных видов РНК. На другом конце спектра не может быть никаких разумных сомнений в том, что первые клетки произошли в результате биологической эволюции на ее доклеточной стадии (см. гл. 11). Где-то между этими крайностями имеется переход, *порог биологической эволюции*. Чаще всего, с появлением концепции мира РНК, этот порог неявно связывается с появлением репликации молекул РНК. Трансляция, как полагают, возникла позже, в результате специфического процесса, движимого отбором. Как уже говорилось в предыдущем разделе, как катализируемая рибозимами репликация, так и, особенно, эволюция трансляции в мире РНК сталкиваются с огромными трудностями. Модель МММ значительно расширяет интервал на оси организационной сложности, где порог может быть преодолен через появление сложности, возникшей случайно (см. рис. 12-6). В этой модели возможность того, что прорывная стадия наступления биологической эволюции была состоянием высокой сложности, не может быть сброшена со счетов, какой бы невероятной (то есть редкой в смысле мультивселенной) и нелогичной она ни казалась. Например, в рамках этой модели прорыв может быть обеспечен случайным появлением центральных элементов сопряженной системы репликации и трансляции, напоминающей, по крайней мере в принципе, современные РНК-вирусы (Koonin, 2007b).

Модель МММ не только допускает, но и гарантирует, что где-то в бесконечной мультивселенной (более того, в каждой отдельной бесконечной вселенной) такая сложная

система появится; кроме того, существует бесконечное число этих систем. Таким образом, уместный вопрос состоит не в том, возникла ли система произвольной сложности спонтанно, по воле случая (поскольку МММ гарантирует это), но в том, *каков был наиболее вероятный прорывный этап на Земле, который должен быть отнесен к случайности в соответствии с антропным подходом.* Я полагаю, что, учитывая серьезные проблемы эволюционных сценариев, разработанных, чтобы объяснить происхождение репликации и трансляции путем биологической эволюции, *следует принимать всерьез возможность того, что порог биологической эволюции соответствует очень сложной стадии – возможно, системе репликации вкупе с трансляцией и с белковыми полимерами, отвечающими за репликацию РНК.* Эта гипотеза (которую я называю антропной химической эволюцией, АХЭ), конечно, не исключает особой важности рибозимов в ранней биологии, в частности, в первичной системе трансляции, как предполагается исходя из сравнительного анализа белковых компонентов аппарата трансляции (см. выше в этой главе). Тем не менее следствием из модели АХЭ является то, что мир РНК, каким он описывается в настоящее время, – огромное сообщество реплицирующихся молекул РНК, наделенных различными каталитическими возможностями, но не содержащий ни системы трансляции, ни генетически кодируемых белков, – мог никогда и не существовать.

Согласно гипотезе АХЭ, основные элементы системы трансляции, а именно состоящая только из РНК рибосома и избирательные адаптеры, по крайней мере для некоего подмножества из двадцати современных белковых аминокислот, возникли случайно, в соответствии с антропным принципом. Согласно этой модели, прорывная система, запустившая биологическую эволюцию, была примитивной, но достаточно эффективной РНК-машинной трансляции, способной транслировать экзогенную РНК, генерируя функциональные белки, в том числе репликазы. Достаточное разнообразие случайно синтезированных РНК, включая и те, что кодировали белки, обладающие активностью репликазы (хоть первоначально и низкой), является еще одной антропно определенной особенностью тех мест на древней Земле, где зародилась жизнь. Как уже говорилось в предыдущем разделе, сети неорганических ячеек в гидротермальных источниках могли играть роль добиотических химических реакторов. Существование такой сети само по себе также является частью антропного сценария.

В таких условиях появление основанного на РНК механизма трансляции привело бы к производству репликазы и, с последующей репликацией РНК, к критическому переходу от антропной причинности к биологическому отбору (см. рис. 12-6). В принципе начало биологической эволюции можно себе представить и с репликазой как единственным изначально активным белком. Однако, учитывая вероятное существование реакторов, производящих РНК, которые обсуждались ранее в этой главе, вполне возможно, что с появлением трансляции другие случайные последовательности РНК послужили основой предковых форм наиболее распространенных белковых укладок, порождая несколько белковых функций (например, РНК-связывающие белки и примитивные ферменты, катализирующие синтез нуклеотидов) и тем самым придавая минимально необходимую устойчивость формирующейся биологической системе. Появление этих укладок можно назвать Большим взрывом белковой эволюции.

Как уже отмечалось, современный универсальный генетический код гораздо более надежен, чем был бы случайный, по отношению к мутационным и, вероятно также, к трансляционным ошибкам. Эта устойчивость проявляется и в очевидной неслучайности структуры кода, выражающейся в первую очередь в том, что серия кодонов, которые отличаются только третьей позицией, кодирует либо одну и ту же, либо две подобные аминокислоты, и в других особенностях соответствия кодонов аминокислотам (Koonin and Novozhilov, 2009). Примечательно, что предполагаемый предковый «дублетный» код, в котором третья позиция не

несла никакой информации, мог быть даже более надежным, чем современный (Novozhilov and Koonin, 2009). Надежность, как обычно предполагается, эволюционировала в ходе оптимизации кода. Однако модель АХЭ предлагает альтернативную точку зрения, при которой базовая структура кода возникла по чистой случайности, поскольку только коды с определенным минимальным уровнем надежности позволили бы функциональной репликазе появиться в прорывной системе. Конечно, этот сценарий не исключает и последующей корректировки кода биологической эволюцией, что, по всей вероятности, в действительности произошло.

Таким образом, гипотеза АХЭ снимает парадоксы происхождения репликации и трансляции предположением о том, что оба этих процесса, в их примитивных формах, не произошли в результате биологической эволюции, а возникли случайно, как связанная система, в соответствии с антропным принципом.

Гипотеза АХЭ, несомненно, должна показаться большинству эволюционных биологов нелепой и возмутительной, поскольку она уклоняется от поисков «механизмов» доклеточной эволюции. Тем не менее существуют смягчающие обстоятельства. Во-первых, постулируемая возможность возникновения сопряженной репликационно-трансляционной системы не требует никаких неизвестных процессов. Напротив, для нее необходимы только хорошо известные, обычные реакции, такие как полимеризация нуклеотидов и аминокислот и фосфорилирование/дефосфорилирование нуклеотидов; при этом нужны только распространенные в химии и биохимии взаимодействия. Как уже отмечалось в этой главе, элементарные реакции, необходимые для трансляции (активация аминокислот, аминоацилирование РНК и транспептидация), легко моделируются с помощью рибозимов, в противоположность репликации РНК, которой, как известно, трудно достичь в отсутствие белков. Во-вторых, исключая разве что полную неадекватность нынешнего понимания условий на первичной Земле, любые мыслимые сценарии эволюции жизни обязательно требуют сочетания маловероятных условий и событий до начала биологической эволюции. Список таких событий включает в себя абиогенный синтез довольно сложных и не очень устойчивых органических молекул, таких как нуклеотиды, накопление этих молекул в соответствующих ячейках до высоких концентраций и их полимеризацию с получением полинуклеотидов достаточного размера и разнообразия. Таким образом, независимо от космологических соображений, некоторые формы антропной причинности представляются неизбежным аспектом эволюции жизни (см. рис. 12-6).

Я привел сценарий АХЭ, чтобы показать, что диапазон сложности, открытой для антропной причинности, может быть гораздо шире, чем предполагалось ранее, настолько, что первичная репликационно-трансляционная система могла возникнуть без биологического отбора. Случайное происхождение изоцирированной системы, способной исполнять сложные биологические функции, может показаться бессмыслицей. Я полагаю, однако, что это всего лишь семантическая ловушка. До наступления биологической эволюции не может быть «функции», только сложность, а модель МММ гарантирует возникновение любого уровня сложности (это гарантированно произойдет «где-то» в бесконечной Вселенной, но антропный принцип прямо помещает эти события на Землю).

Все эти соображения провоцируют довольно-таки кошмарный вопрос: имеет ли какое-либо значение в бесконечно избыточном мире МММ биологическая эволюция вообще и дарвиновский отбор в частности? Разве не возникнет система любой, даже наивысшей, сложности просто случайно? Ответ: да, но этот вопрос упускает важное обстоятельство. В модели МММ случайное появление бесконечного числа сложных биот неизбежно, но несравнимо реже, чем развитие по сценарию АХЭ, которое включает в себя переход от случайности к биологической эволюции после стадии прорывной системы (см. рис. 12-6).

Начало биологической эволюции канализирует исторический процесс, сокращая многочисленные траектории, которые возможны в принципе, до относительно малого числа стабильных и более вероятных – тех, что совместимы с дарвиновским режимом эволюции сложных систем (см. рис. 12-7). Этот переход приводит к гораздо большей скорости эволюционных изменений, чем та, что достигается случаем, и, как только возникает возможность для биологической эволюции, антропная причинно-следственная связь отходит на второй план в истории жизни. Конечно, «второй план» не означает потери важности: непредвиденные случайности имеют решающее значение, особенно на переходных этапах эволюции (см. обсуждение ранее в этой книге, особенно в гл. 7). Таким образом, в любой реконструкции происхождения жизни и ранней эволюции порог должен быть связан с самой нижней точкой, то есть минимально сложной системой, способной к биологической эволюции.

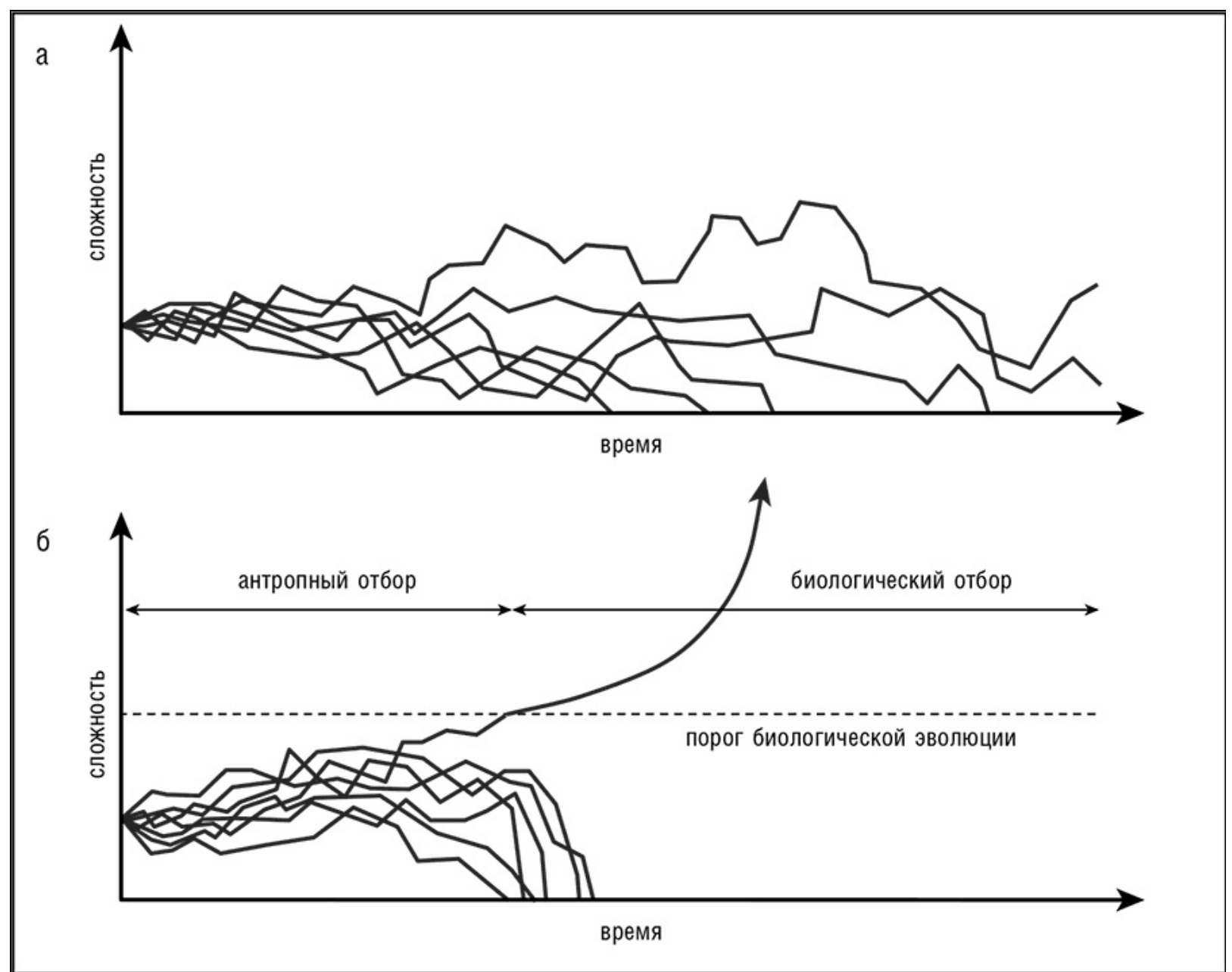


Рис. 12-7. Отсечение эволюционных траекторий на пороге биологической эволюции: а – только химическая эволюция; б – возникновение биологической эволюции.

Сильная форма гипотезы АХЭ, согласно которой прорывным этапом в истории жизни была примитивная совместная репликационно-трансляционная система (см. рис. 12-6), в принципе легко опровержима. Такая система должна рассматриваться как верхняя граница сложности для

прорывного этапа. Как только возможность биологической эволюции при более низком уровне сложности, таком как мир РНК, будет убедительно продемонстрирована, а путь от мира РНК к системе трансляции очерчен либо опытным путем, либо в убедительной модели, сильная форма гипотезы АХЭ будет опровергнута. Демонстрация самостоятельного возникновения жизни на нескольких планетах в нашей Вселенной будет иметь тот же результат. В приложении приведен грубый, но, хочется надеяться, поучительный расчет верхней границы вероятности возникновения совместной репликационно-трансляционной системы в наблюдаемой части Вселенной; эта вероятность, несомненно, исчезающе мала. Противоположное предсказание состоит в том, что любые формы жизни, которые могут быть обнаружены на Марсе, или, возможно, Европе (спутнике Юпитера, на котором была обнаружена жидкая вода), или даже на любой экзопланете в ходе будущих планетарных исследований, будут иметь общее с земной жизнью происхождение. Любая из этих находок опровергнет сильную гипотезу АХЭ, но не сделает модель МММ незначимой для нашего понимания происхождения жизни. Действительно, любое такое открытие (важное само по себе) просто понизит порог биологической эволюции в масштабе рис. 12-7.

Самым простым и сильным опровержением гипотезы АХЭ будет опровержение самой МММ [\[135\]](#). Тем не менее здесь требуется важное разъяснение. Для обоснованности представленных здесь концептуальных рамок не требуется верная во всех деталях МММ. Только два общих положения существенны: (1) пространственно бесконечная вселенная, такая, как любая островная вселенная в МММ; мультивселенная, хоть и является неотъемлемой частью теории вечной инфляции, не требуется в качестве аргумента; (2) конечность числа различных макроскопических историй. Даже сильная форма гипотезы АХЭ, представленная здесь, не будет опровергнута, если некоторые конкретные детали МММ окажутся неправильными, но только если одно из этих общих предположений окажется неверным.

Краткий обзор и перспектива

Проблема происхождения жизни является не только одной из самых трудных задач во всей науке, но и одной из важнейших. Исследование происхождения жизни превратилось в активно развивающуюся область междисциплинарных исследований, хотя некоторые ученые часто смотрят на нее со скептицизмом и даже с насмешкой. Такое отношение вполне понятно и в каком-то смысле может быть оправдано, учитывая «постыдный», редко упоминаемый секрет: несмотря на многие интересные результаты, если судить только по простым критериям достижения конечной цели или даже приближения к ней, исследование происхождения жизни, можно сказать, терпит фиаско – мы все еще не имеем даже правдоподобной и непротиворечивой модели, не говоря уже об обоснованном сценарии возникновения жизни на Земле. Конечно, это связано не с отсутствием экспериментальных и теоретических усилий [\[136\]](#), а с чрезвычайной имманентной трудностью и сложностью проблемы. Для возникновения жизни необходима последовательность чрезвычайно маловероятных событий, от синтеза и накопления нуклеотидов до появления трансляции; при перемножении их вероятностей окончательный результат кажется почти что чудом.

Не все так мрачно: важные указания на возможные пути возникновения жизни все же были обнаружены. Некоторые особенности природной среды, которые существуют до сих пор, такие как сети неорганических ячеек в гидротермальных источниках, вероятно, существовали и 4 миллиарда лет назад и могли быть подходящим инкубатором для всех ранних шагов эволюции жизни, начиная с синтеза и концентрирования мономеров до происхождения трансляции. Гипотеза мира РНК, серьезно поддерживаемая, пусть и не напрямую, внушительным объемом данных по каталитической активности рибозимов, является привлекательным, и, видимо, единственным мыслимым выходом из парадоксальных ситуаций, связанных с происхождением трансляции.

Тем не менее трудности остаются огромными. Несмотря на все усилия, мы в настоящее время не имеем связной и убедительной модели, описывающей путь от простых органических молекул к первой форме жизни. Хуже всего, что мощные механизмы биологической эволюции не были доступны ни на одной стадии, предшествующей появлению репликаторных систем. Учитывая все эти принципиальные трудности, кажется разумным серьезно рассматривать радикально альтернативные гипотезы возникновения жизни. Космологическая модель вечной инфляции и связанная с ней гипотеза МММ может предложить решение загадки происхождения жизни, потому что в бесконечной мультивселенной с конечным числом различных макроскопических историй (так что каждая повторяется бесконечное число раз) случайное появление даже очень сложных систем не только возможно, но и неизбежно. Таким образом, интервал на шкале организационной сложности, в котором может лежать переход от антропоидной селекции к биологической эволюции, резко расширяется. В частности, можно представить, что прорывным шагом к наступлению биологической эволюции было случайное появление примитивной сопряженной системы репликации и трансляции. То, что это крайне редкое событие произошло на Земле и привело к возникновению жизни такой, как она нам известна, может быть обосновано одной лишь антропоидной причинностью. Согласно этой модели, мир РНК как таковой, с его разнообразным населением реплицирующихся молекул РНК, но без трансляции, никогда не был этапом ранней эволюции жизни на Земле. Однако этот сценарий ни в коем случае не отменяет центральной роли РНК в становлении биологической эволюции и ранней эволюции жизни. В самом деле, модель АХЭ включает в себя сложный ансамбль нереплицирующихся молекул РНК, возникший по воле случая, положившего начало

биологической эволюции.

Учитывая огромную сложность и трудность проблемы происхождения жизни и отсутствие биологических механизмов эволюции (отбора и дрейфа) на любом этапе, предшествовавшем довольно сложным репликаторным системам, я полагаю, что возможность возникновения жизни на основе сочетания чрезвычайно маловероятных событий, которое теория МММ делает неизбежным, не должна сбрасываться со счетов. Эта возможность кажется в высшей степени нелогичной, но мы слишком хорошо знаем, что интуиция – плохой путеводитель по временным и пространственным масштабам, лежащим далеко за пределами человеческого опыта. Кроме того, модель АХЭ – не досужие домыслы. Напротив, это вполне опровержимая гипотеза, и опровержение, случись оно в форме демонстрации возможности развития трансляции в мире РНК или открытия независимо возникшей жизни в нашей Вселенной, будет по-настоящему переломным достижением.

Рекомендуемая дополнительная литература

Aravind L., R. Mazumder, S. Vasudevan, and E. V. Koonin. (2002) Trends in Protein Evolution Inferred from Sequence and Structure Analysis. *Current Opinion in Structural Biology*12: 392–399.

В статье приводятся аргументы в пользу того, что основные белковые укладки достигли существенного разнообразия еще до появления системы трансляции современного типа. Реконструкция эволюции укладки Россмана, присутствующей в АРСазах класса I, используется, чтобы показать, что основанные большей частью на РНК системы трансляции должны быть достаточно эффективны для белковой эволюции.

Crick F. H. (1968) The Origin of the Genetic Code. *Journal of Molecular Biology*38: 367–379.

Блестящая пророческая статья, остающаяся актуальной более чем через 40 лет после опубликования. В ней Крик очерчивает идею мира РНК (без использования этого термина) и возможные сценарии эволюции генетического кода, до сих пор определяющие данную область исследований.

Koonin E. V. (2007) The Cosmological Model of Eternal Inflation and the Transition from Chance to Biological Evolution in the History of Life. *Biology Direct*2: 15.

В статье вопрос о происхождении жизни помещается в контекст инфляционной космологии, согласно которой каждое физически возможное макроскопическое состояние существует в бесконечном числе экземпляров в бесконечной вселенной. Согласно этой концепции, крайне маловероятные события, такие как случайное появление полной совместной системы репликации и трансляции, не могут быть исключены как возможные шаги добиологической эволюции.

Koonin E. V., and W. Martin. (2005). On the Origin of Genomes and Cells Within Inorganic Compartments. *Trends in Genetics*21: 647–654.

В статье развивается единый сценарий доклеточной эволюции генетических элементов, ансамблей «эгоистичных кооператоров» и все увеличивающихся геномов в неорганических ячейках.

Martin W., J. Baross, D. Kelley, and M. J. Russell. (2008) Hydrothermal Vents and the Origin of Life. *Nature Reviews Microbiology*6: 805–814.

Обзор условий в гидротермальных источниках и свойств сетей неорганических ячеек, делающих их пригодными в качестве инкубаторов жизни.

Martin W., and M. J. Russell. (2003) On the Origins of Cells: A Hypothesis for the Evolutionary Transitions from Abiotic Geochemistry to Chemoautotrophic Prokaryotes, and from Prokaryotes to Nucleated Cells. *Philosophical Transactions of the Royal Society London B Biological Sciences*358: 59–83.

Основополагающая статья, где идея о том, что ранние стадии эволюции жизни могли быть ограничены сетями неорганических ячеек в гидротермальных источниках, впервые помещена в биологический контекст.

Robertson M. P., and G. F. Joyce. (2012) The Origins of the RNA World. *Cold Spring Harbor Perspectives in Biology*4:a003608.

Обзорная статья, описывающая концепцию мира РНК и активность рибозимов, важных для происхождения репликаторных систем. Авторы признают трудности, с которыми сталкивается гипотеза мира РНК, и делают предположение, что класс репликаторных систем иной, пока неизвестной природы мог предшествовать РНК.

Russell M. J. (2007) The Alkaline Solution to the Emergence of Life: Energy, Entropy, and Early Evolution. *Acta Biotheoretica*55: 133–179.

Подробная статья, где обосновывается представление о том, что «проточные реакторы» гидротермальных источников обладают низкой энтропией и способствуют образованию сложных формаций, таких как репликаторные системы, и, в конечном счете, клеток.

Vetsigian K., C. Woese, and N. Goldenfeld.(2006) Collective Evolution and the Genetic Code. *Proceedings of the National Academy of Sciences USA*. 103: 10,696—10,701.

Убедительная аргументация, подтвержденная математическим моделированием, в пользу того, что универсальность генетического кода обусловлена решающей ролью ГПГ на ранних стадиях эволюции жизни.

Wolf Y. I., and E. V. Koonin.(2007) On the Origin of the Translation System and the Genetic Code in the RNA World by Means of Natural Selection, Exaptation, and Subfunctionalization. *Biology Direct*2: 14.

Развернутый гипотетический сценарий возникновения системы трансляции путем экзаптации древних рибозимов, стимулируемых аминокислотами и пептидами.

Mulkiđjanian AY, Bychkov AY, Dibrova DV, Galperin MY, Koonin EV. Origin of first cells at terrestrial, anoxic geothermal fields. Proc Natl Acad Sci USA. 2012 Apr 3;109(14):E821-30.

В статье рассматриваются разнообразные данные, поддерживающие сценарий происхождения клеток в окрестностях наземных геотермальных источников.

Глава 13. Постсовременное состояние эволюционной биологии

В предыдущих двенадцати главах мы рассмотрели многообразие аспектов эволюции жизни. Подобное рассмотрение, конечно, никоим образом не может быть всеобъемлющим, но это и не было моей целью. Так или иначе, информация, представленная в предыдущих главах, достаточна (и необходима) для выражения сути этой книги: за 50 лет, прошедших со времени утверждения СТЭ в науке, эволюционная биология кардинальным образом изменилась и вошла в новую эпоху – эпоху «постмодерна».

Согласно синтетической теории эволюции, эволюция жизни – это процесс активной адаптации популяций к изменяющимся условиям среды. Теперь стало очевидным, что, хотя подобная адаптация и является, несомненно, существеннейшим компонентом эволюционного процесса, в количественном отношении она не доминирует. И хотя я полностью сознаю, что любая попытка предложить широкое обобщающее определение приводит к чрезмерному упрощению, вместе с тем предлагаю следующее.

Эволюция жизни – это преимущественно стохастический процесс, основанный на исторической случайности, ограниченный прежде всего разнообразными условиями поддержания основ биологической организации и модулируемый механизмом адаптации.

Ограничения, формирующие ход эволюции, имеет смысл рассматривать максимально широко. Они включают все виды предупреждения и устранения повреждений и локальной оптимизации, такие как снижение уровня ошибок во всех информационных процессах и расход энергии, а также непрерывную «гонку вооружений» между паразитами и хозяевами, которая стимулирует эволюцию различных адаптаций, благодаря эффекту Красной Королевы. В этой заключительной главе кратко излагаются разнообразные аспекты постсовременного состояния эволюционной биологии и рассматриваются осуществимость и возможные очертания постсовременной теории эволюции. В табл. 13-1 представлена постсовременная переоценка нескольких фундаментальных положений концепции Дарвина и СТЭ.

Таблица 13-1. Постсовременная переоценка некоторых центральных положений дарвиновской и синтетической теорий эволюции.

Положение	Постсовременное состояние
<p>Материал для эволюции поставляется в первую очередь случайными наследуемыми вариациями</p>	<p>Верно лишь отчасти. Репертуар случайных изменений значительно расширен и включает дупликацию генов, областей генома и целых геномов; потерю генов и вообще генетического материала; ГПГ, в том числе массовый перенос генов при эндосимбиозе; вторжение мобильных эгоистичных элементов и пополнение генома их последовательностями и т. п. Более того, (квази)направленное (ламарковское) изменение генома признается существенным фактором эволюции</p>
<p>Фиксация (редких) полезных изменений естественным отбором — главная движущая сила эволюции</p>	<p>Верно лишь отчасти. Естественный (положительный) отбор является важным, но лишь одним из нескольких фундаментальных факторов эволюции и не доминирует количественно. Доминирующими в эволюции являются нейтральные процессы в сочетании с очищающим отбором. Также важно прямое воздействие сигналов окружающей среды на геном, то есть (квази)ламарковские явления</p>
<p>Изменения, фиксируемые естественным отбором, крайне незначительны. Эволюцию описывает градуализм</p>	<p>Неверно. Дупликация или горизонтальный перенос даже одного гена — отнюдь не «бесконечно малые» изменения, равно как и делеция или приобретение крупных сегментов ДНК, перестановка частей генома, полногеномная дупликация и, что наиболее существенно, эндосимбиоз. Градуализм не может быть основным режимом эволюции</p>

Положение	Постсовременное состояние
<p>Униформизм: эволюционные процессы в целом оставались одними и теми же в течение эволюции жизни</p>	<p>Верно лишь отчасти. Современные эволюционные процессы важны с момента происхождения репликации. Тем не менее крупные скачки эволюции, такие как происхождение эукариот, могут быть вызваны фактически уникальными событиями, такими как эндосимбиоз; самые ранние этапы эволюции (до LUCA) частично опирались на различные процессы, не участвующие в последующей «нормальной» эволюции</p>
<p>Эволюция путем естественного отбора стремится производить все более сложные адаптивные свойства организма, а значит, прогресс есть общая тенденция в эволюции</p>	<p>Неверно. Сложность генома, вероятно, эволюционировала как «геномный синдром», вызванный слабостью очищающего отбора в малой популяции, а не как адаптация. Никакой тенденции к увеличению сложности в эволюции нет, и понятие эволюционного прогресса является необоснованным</p>
<p>Вся эволюция жизни может быть изображена единым большим деревом</p>	<p>Неверно. Открытие фундаментального вклада ГПГ и мобильных генетических элементов в эволюцию генома отвергает понятие ДЖ в его первоначальном смысле. Тем не менее деревья остаются важной формой представления эволюции отдельных генов и многих фаз эволюции в группах относительно близкородственных организмов. Возможность восстановления ДЖ в правах как центральной тенденции эволюции не исключена</p>
<p>Все сегодняшние формы клеточной жизни восходят к малому числу древних форм (возможно, одной-единственной, LUCA)</p>	<p>Верно. Сравнительная геномика не оставляет никаких сомнений в общем происхождении клеточной жизни. Однако она также указывает на возможность, что LUCA(S) сильно отличался от современных клеток</p>



Плюрализм паттернов и процессов эволюции: смена концепций отбора, вариации и древа жизни

Роль и статус отбора

Двоякий смысл слова «постмодерн» в предисловии к этой книге, возможно, не ускользнул от внимательного читателя. Что бы мы ни думали о постмодернистской философии (см. прил. I), ее мировоззрение, безусловно, подчеркивает богатство и чрезвычайное разнообразие процессов и закономерностей, составляющих реальность. Сложность этих разнообразных тенденций такова, что, по мнению некоторых философов постмодернистского толка, ни одно серьезное обобщение не имеет права на существование. В сегодняшней эволюционной биологии разнообразие процессов и закономерностей является, пожалуй, главной темой, и, коль скоро мы принялись говорить парадоксами, можно сказать, что *главной темой является отсутствие всеобъемлющей главной темы.*

Центр внимания, лежавший в СТЭ на естественном отборе, действующем на случайные генетические изменения, сместился ко множеству взаимодополняющих фундаментальных эволюционных процессов и моделей (см. рис. 13-1). В новой эволюционной биологии естественный отбор является лишь одним из процессов, формирующих эволюционирующий геном, и, видимо, не доминирует количественно. Эволюцию в значительной степени определяют нейтральные процессы, такие как генетические дрейф и тяга.

Развивая тему плюрализма, заметим, что относительный вклад адаптивных и нейтральных процессов отнюдь не постоянен по всей широте спектра форм жизни. Как очень точно сформулировал Майкл Линч (естественно, перефразируя Добржанского), «ничто в эволюции не имеет смысла, кроме как в свете популяционной генетики» (Lynch, 2007b). Действительно, динамика популяций – или, проще говоря, эффективный размер популяции, в краткосрочной и долгосрочной перспективе – является ключевым фактором, определяющим давление отбора. Эффективный размер популяции может отличаться на порядки даже у близкородственных организмов, поэтому различия в интенсивности отбора весьма велики (например, у насекомых и у млекопитающих). Эти различия определяют различные эволюционные режимы: при высоком *Неэволюция* определяется в первую очередь (или даже исключительно) отбором; при низких же *Нена* первое место выходит дрейф. В реальном процессе эволюции (почти) все линии спуска проходят через несколько «бутылочных горлышек», где доминирующей силой эволюции становится случайный дрейф; отсюда следует *неизбежный и большой вклад случайности в эволюцию всех живых организмов.*

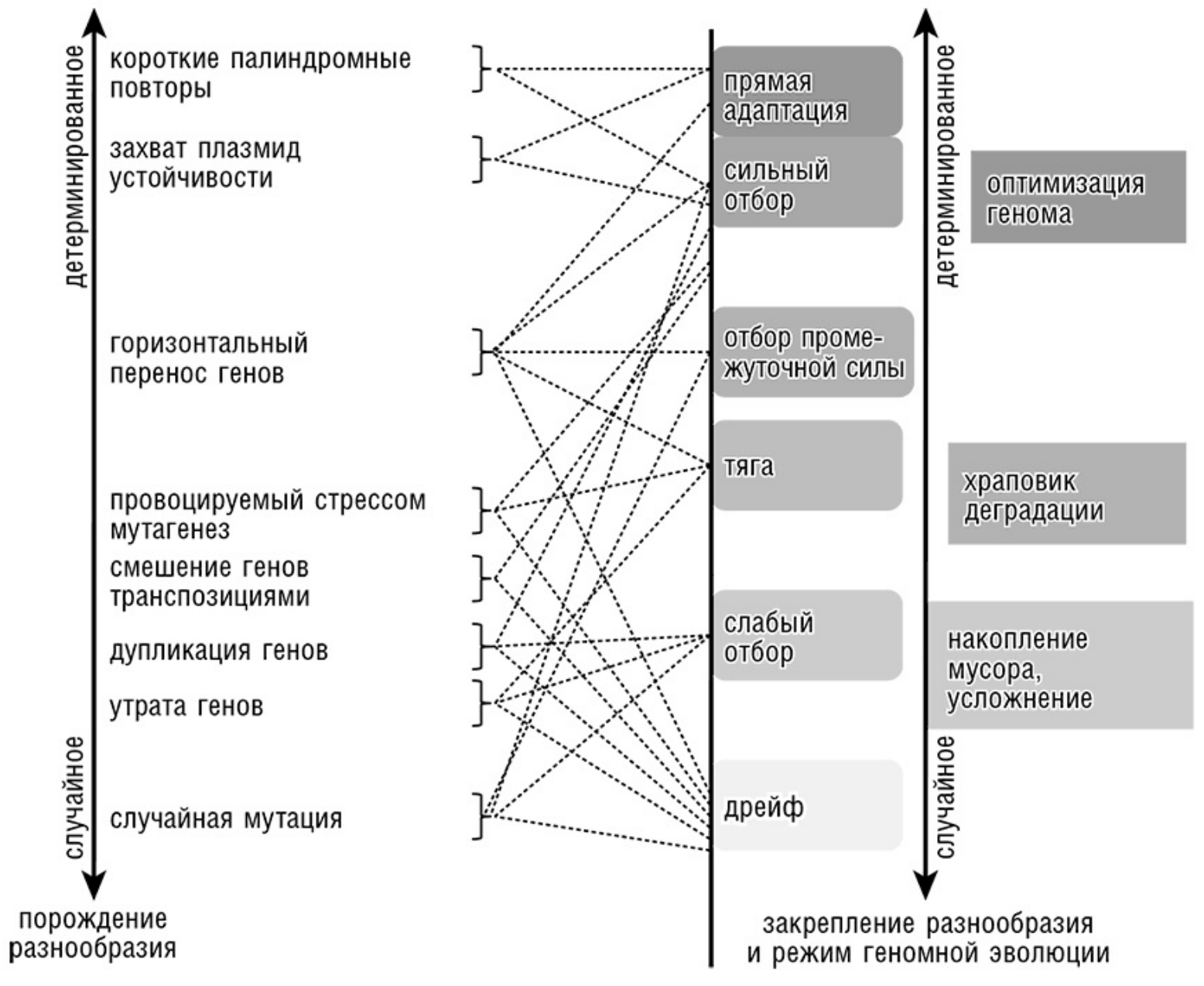


Рис. 13-1. Множественные процессы, порождающие геномное разнообразие, влияющие на его закрепление в эволюции и формирующие эволюцию генома в целом.

В эволюции фенотипов роль отбора, возможно, еще больше, чем в эволюции геномов. Однако достижения системной биологии существенно расширили концепцию фенотипа. Наряду с традиционными признаками организмов, мы теперь изучаем и эволюцию молекулярных фенотипических признаков, таких как экспрессия гена и изобилие белка; молекулярная эволюция фенотипических признаков, как оказывается, включает в себя заметную нейтральную составляющую. Более того, *храповик конструктивной нейтральной эволюции*, по-видимому, приводит в действие неадаптивную эволюцию сложных фенотипических признаков, выходящих, с традиционной (нео) – дарвинистской точки зрения, типичными адаптациями.

Даже когда отбор и адаптация несомненно участвуют в процессе, проявление этих факторов эволюции часто (возможно даже, в большинстве случаев) довольно сильно отличается от (нео)дарвиновской идеи «улучшения». Весьма часто *адаптации имеют отношение к сохранению целостности клеточной организации, предотвращению сбоев и контролю за повреждениями*. В некотором смысле сказанное выше есть тривиальная констатация факта, если мы учтем разнообразие и сложность молекулярных машин, предназначенных для контроля качества каждого из основных процессов передачи информации, примерами которых являются

системы репарации ДНК и деградации белков и молекулярные шапероны. Кроме того, важным, чтобы не сказать главным движущим началом эволюции белок-кодирующих генов представляется отбор на устойчивость к ошибкам укладки. В многоклеточных организмах значение отбора для предотвращения сбоя очевидно проявляется на уровне клеток и тканей, примером чего служит сложнейшая система апоптоза, запрограммированной клеточной гибели.

В ретроспективе все эти выводы могут показаться совершенно очевидными, если учесть, как развиты, сложны и в том или ином смысле оптимизированы клетки или даже отдельные белки и РНК. Как только эти сложные системы возникли – а эволюционные реконструкции ясно показывают, что они присутствовали в течение подавляющей части истории жизни, более 3,5 миллиарда лет, – контроль качества и предотвращение повреждений действительно стали основными «задачами» эволюции, невзирая на важность появления время от времени новых адаптаций. Признание этого факта возлагает огромную нагрузку на ранние, доклеточные стадии эволюции, когда изменения были быстрыми, а роль положительного отбора, наряду с конструктивной нейтральной эволюцией, должна была быть куда значительнее, чем в течение последующих 3,5 миллиарда лет. В некотором смысле почти все «действительно интересное» в эволюции жизни произошло на сравнительно короткой ранней стадии, предшествующей установлению основ клеточной организации (см. гл. 11 и 12, а также более подробно ниже в данной главе). Конечно, существуют и заметные исключения, такие как появление эукариотических клеток или многоклеточных эукариотических организмов, но нет никаких сомнений, что большинство фундаментальных эволюционных инноваций «втиснуто» в короткий отрезок первых 5 процентов истории жизни.

Меняющиеся концепции вариации и конец градуализма

В дополнение к меняющимся концепциям отбора и его роли в эволюции были радикально переработаны и идеи о том, что является эволюционно важными геномными и фенотипическими вариациями (см. рис. 13-1). Случайные изменения, ведущие к бесконечно малым полезным фенотипическим изменениям, которые Дарвин рассматривал как ключ ко всей эволюции, остаются важными, но становятся лишь одним из классов изменений, существенных для эволюции, притом не доминирующим – во всяком случае, количественно.

Начнем с того, что почти нейтральные мутации, «невидимые» отбору и не закрепленные дрейфом или же сохраняющиеся в популяции, не будучи закрепленными, встречаются, по-видимому, чаще, чем слегка полезные «дарвиновские» мутации. Важный момент, который не был четко осознан до недавнего времени, состоит в том, что практически нейтральные мутации далеко не безразличны для эволюции. В действительности эти мутации образуют нейтральные сети, составляющие важнейший резервуар эволюционной пластичности.

Кроме того, формы генетических вариаций, которые ни в коем случае нельзя рассматривать как ведущие к бесконечно малым полезным эффектам, имеют решающее значение для эволюции. Это понимание ставит крест на градуализме, который Дарвин и архитекторы СТЭ полагали центральным для всей эволюции. Неградуалистские типы эволюционно важных генетических изменений включают дубликации генов и целых геномов, потерю генов и ГПГ, в частности храповики обширного, направленного переноса генов, провоцируемого эндосимбиозом. Горизонтальный перенос генов является наиболее распространенным эволюционным процессом у прокариот, и, по всей видимости, имеют место специфические адаптации, поддерживающие ГПГ на оптимальном уровне. Скорость ГПГ значительно уменьшается у эукариот, где обширная дубликация с последующей субфункционализацией компенсирует сокращение ГПГ. Кроме того, эндосимбиоз и последующий перенос генов от

симбионтов к хозяину сыграли решающую роль в эволюции эукариот.

Не только бесконечно малый эффект изменений, но и его полная случайность также уходит в прошлое. Такие механизмы, как стресс-индуцированный мутагенез, высокоразвитые, сложные системы которого существуют во всех клеточных формах жизни, являются адаптивными и неслучайными, и адаптивная система иммунитета у прокариот, по-видимому, действует через самое настоящее ламарковское наследование. Как правило, эволюционные процессы охватывают континуум ламарковского, дарвиновского и райтовского режимов эволюции, и относительный вклад каждого из них в любом конкретном эпизоде из истории жизни зависит от динамики популяции и давления окружающей среды. Наконец, становится все более ясным, что фенотипические мутации существенны для эволюции и могут способствовать адаптации в сочетании с генетическими мутациями, в частности посредством эффекта предвидения. Все это вместе – важность неслучайных мутаций и (квази)ламарковские механизмы эволюции, наряду с вкладом фенотипических мутаций, – показывает, что *способность к эволюции тоже эволюционирует, что, в свою очередь, по-видимому, опровергает, хотя бы частично, одно из заветных убеждений эволюционных биологов – веру в то, что эволюция не имеет предвидения.*

От древа жизни к паутине жизни

Приверженность эволюционной биологии древу жизни как единому, окончательному представлению об истории форм жизни на Земле уступает место плюралистической картине, в которой разнообразные сетевые процессы дополняют древовидные процессы генной эволюции. Эти процессы включают ГПГ, особенно распространенный у прокариот, но также вносящий огромный вклад в эволюцию эукариот, особенно через эндосимбиоз, а также через различные формы слияния генома и обмена генетическим материалом между хозяевами и паразитами (см. гл. 5 и 7). В первом приближении можно сказать, что в наших представлениях об истории жизни древо жизни уступило место паутине жизни (или ризоме жизни; Raoult, 2010) [\[137\]](#).

Когда мы глубже всматриваемся в эволюционные процессы, становится ясно, что эволюция жизни может быть осмысленно отражена лишь сложной, динамической картиной взаимодействующих процессов, среди которых древовидная эволюция будет, вероятно, самым фундаментальным, поскольку эта форма эволюции является прямым следствием двоичной репликации генетического материала как базового универсального процесса всей жизни (гл. 6). Более того, имеется неоспоримая динамическая согласованность между эволюционной историей больших ансамблей генов (самые большие из таких коэволюционирующих ансамблей известны как геномы), возрождающая концепцию ДЖ, но уже в виде центральной статистической тенденции в «лесу» филогенетических деревьев отдельных генов. Несмотря на статистическую согласованность генетических историй, «ядро» универсально консервативных, повсеместно распространенных генов клеточных форм жизни чрезвычайно мало из-за потери генов в отдельных линиях эволюции и неортологичного замещения генов – фундаментальных эволюционных явлений, важность которых не могла быть оценена в догеномную эпоху. Суммируя, можно сказать, что *паутина эволюции представляет собой высокодинамичное геномное пространство-время*, в котором геном каждого вида есть лишь преходящее, метастабильное скопление генов.

Неожиданная значимость простых физических и математических моделей для понимания эволюции: биологическая эволюция как предмет статистической физики

В предыдущем разделе мы рассмотрели множественность структур и процессов, которые являются определяющим аспектом новой концепции эволюции. Упрощенно говоря, эта множественность значительно увеличивает энтропию эволюционной биологии. Тем не менее анализ данных, получаемых в областях геномики и системной биологии, выявляет в равной степени заметную противоположную, «антиэнтропийную» тенденцию к структурированию эволюционной теории. Было обнаружено несколько универсальных распределений и зависимостей, таких как распределение скорости генной эволюции, связь между эволюцией и экспрессией генов и распределение степени связности узлов разнообразных сетей. Более того, некоторые из этих универсалий могут быть легко получены из простых математических моделей эволюции, очень похожих на модели, используемые в статистической физике. Эти модели становятся все более общими по мере того, как они объединяются и совместно объясняют универсальные зависимости, первоначально казавшиеся несвязанными, например распределение скорости генной эволюции совместно с антикорреляцией между эволюционной скоростью и экспрессией, или законы масштабирования для семейств генов совместно с масштабированием по функциональным классам.

Имеется удивительно простое общее объяснение этой неожиданной применимости простых моделей для объяснения эволюции геномов. Эволюционная геномика имеет дело с большими ансамблями объектов (генов, белков), которые можно для многих целей рассматривать как слабо взаимодействующие и перемещающиеся (то есть эволюционирующие) по независимым траекториям. Соответственно, *принципы статистической физики столь же применимы к генетическим ансамблям, сколь и к ансамблям молекул*. Естественно, статистический подход к эволюционным явлениям подвержен тем же ограничениям, что и аналогичные подходы в физике, а именно эти модели недостаточны для объяснения конкретных биологических явлений, часто связанных с небольшим набором генов, а не с большим ансамблем. Кроме того, взаимодействие между генами, эпистаз, часто вносит существенный вклад в эволюцию [\[138\]](#). Тем не менее примечательно, что современная геномика и системная биология, раскрывая чрезвычайно сложную, многогранную картину эволюции, в то же время позволяют выработать и простые обобщенные модели. Очень заманчиво предложить еще один новый вариант знаменитой фразы: *ничто в эволюции – и в популяционной генетике – не имеет смысла, кроме как в свете статистической физики*.

Воспроизводимость эволюции: детерминизм и стохастика эволюционного процесса

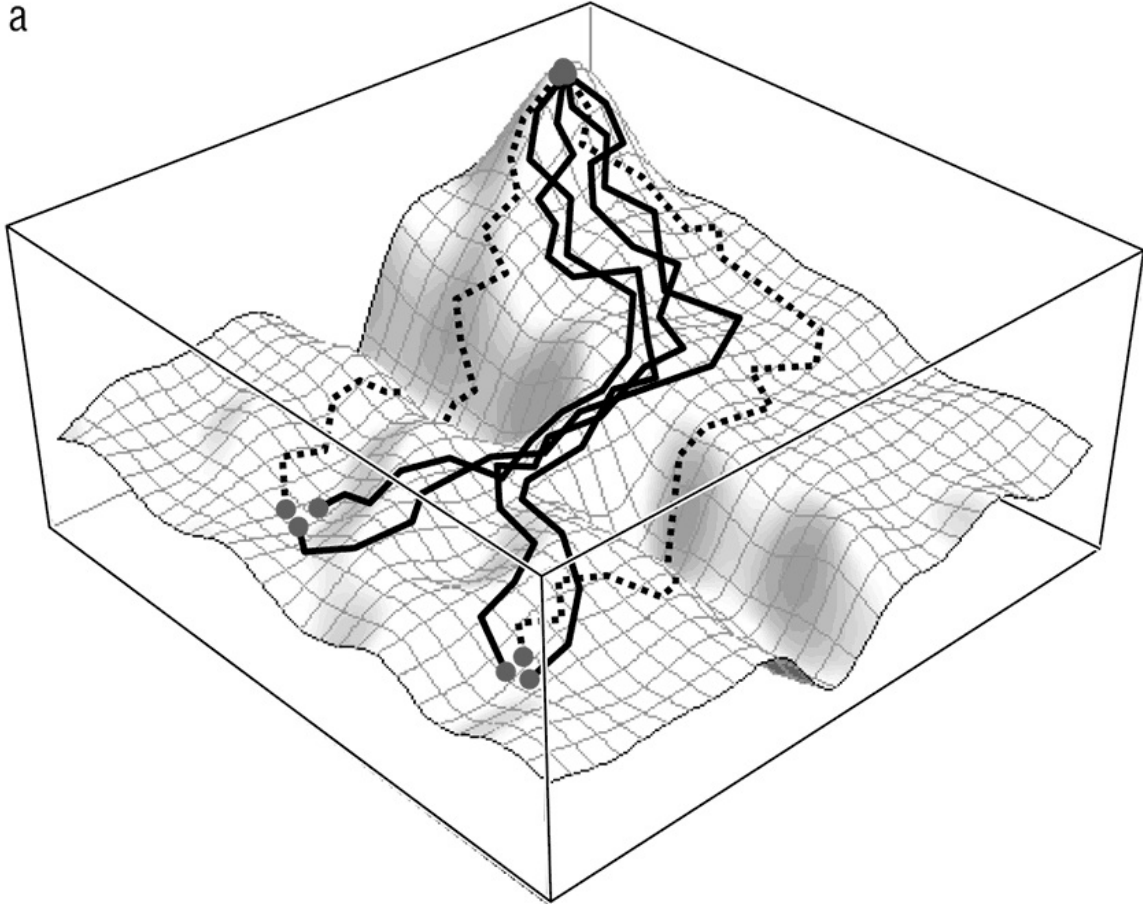
Пространство генотипов, даже если рассматривать только относительно простые, небольшие геномы, невообразимо огромно (скажем, для прокариот с геномом в 1 миллион пар оснований имеется $4^{1\ 000\ 000}$ возможных последовательностей, число, значительно превосходящее по величине все, что существует на самом деле в наблюдаемой части Вселенной, например общее число протонов или электронов). Какая часть этих генотипов на самом деле жизнеспособна и, следовательно, могла бы сыграть роль в эволюции? Или, чтобы задать вопрос осмысленным в контексте эволюции образом, какова доля всех возможных траекторий в пространстве генотипов, которые открыты эволюционному процессу? Вышесказанное – формализованная постановка любимого вопроса Стивена Джей Гулда (Gould, 1997b): что бы мы увидели, если бы имели возможность заново проиграть пластинку эволюции? Ответ, данный не только Гулдом, но и Франсуа Жакобом в знаменитой статье об эволюции-«ремесленнике» (Jacob, 1977), Дэном Деннетом в «Опасной идее Дарвина» (Dennett, 1996) и многими другими, был таков: мы бы не увидели ничего подобного реально существующей биосфере, потому что вся эволюция – сплошная цепь непредвиденных стечений обстоятельств. При описании общей картины эволюции Деннет вполне обоснованно обращается к физическому явлению детерминированного хаоса: каждое событие, которое происходит в процессе эволюции, безусловно, имеет конкретные физические причины, но малые возмущения способны вызвать большие изменения в ходе эволюции, так что далекие результаты становятся совершенно непредсказуемыми.

По-прежнему трудно дать уверенный общий ответ на этот ключевой вопрос эволюции, однако имеющиеся ограниченные прямые исследования эволюционных траекторий как индивидуальных белков, так и бактериальных популяций принесли неожиданные результаты (O'Maille et al., 2008; Ostrowski et al., 2008; Weinreich et al., 2006). Похоже, что в большинстве случаев лишь небольшая часть из теоретически возможных путей на самом деле доступна для эволюции, так что эволюция представляется менее стохастической, более детерминированной и более предсказуемой, чем предполагалось ранее (см. рис. 13-2). Эти наблюдения позволяют предположить, что адаптивные ландшафты по меньшей мере некоторых из развивающихся генов и геномов являются существенно неровными, так что большинство путей прерывается глубокими оврагами низкой приспособленности и, таким образом, запрещены (O'Maille et al., 2008). Основной подоплекой этого, вероятно, является эпистаз, то есть взаимодействие между различными частями одного и того же гена или между различными генами: на пересеченном ландшафте одна мутация часто приводит к фатальному падению приспособленности, но вторая, путем эпистаза, способна привести в область высокой приспособленности. Эпистаз представляется одним из важнейших факторов, поддерживающих целостность эволюционирующих биологических систем, которая проявляется в многих аспектах эволюции (Kogenaru et al., 2009). Как отмечалось в предыдущем разделе, эпистаз неизбежно ограничивает применимость представления эволюционирующих геномов ансамблями слабо взаимодействующих «частиц». Эпистатическое взаимодействие сильно ограничивает диапазон доступных эволюционных траекторий – но насколько сильно, еще предстоит выяснить с помощью дальнейшего моделирования и экспериментальных исследований эволюции. Вполне может оказаться, что модель детерминированного хаоса верна и что обнаруженные ограничения на практике мало влияют на предсказуемость эволюции, то есть на результат метафорического повторного проигрывания пластинки. Доступные траектории, даже если они и составляют лишь

малую долю теоретически возможных, все же могут оказаться столь многочисленными и разнообразными, что эволюция окажется на практике непредсказуемой. Важнейшей и пока нерешенной проблемой оказывается взаимосвязь между доступными траекториями. Если эти траектории кластеризуются на небольшом участке геномного пространства-времени, эволюция может быть квазидетерминированной [\[139\]](#); если же доступные траектории беспорядочно разбросаны, (не)предсказуемость эволюции не будет сильно зависеть от подобных ограничений (см. рис. 13-2).

Скорее всего, результаты детального анализа эволюционных ландшафтов и траекторий на них будут различаться для эволюции на различных уровнях и в различных ситуациях, в согласии с концепцией плюрализма эволюционных процессов, обсуждавшейся выше. Кроме того, следует еще раз подчеркнуть, что соотношение между детерминизмом и стохастичностью определяется давлением отбора, то есть эффективным размером популяции. В эффективно бесконечной популяции эволюция фактически детерминирована, в то время как в небольших популяциях эволюция происходит стохастически в рамках фундаментальных ограничений. Чтобы исключить всякую возможность недоразумений, отметим, что, даже если эволюция и может быть описана как квазидетерминированный процесс, такое описание не имеет ничего общего с телеологическими представлениями. Тем не менее канализация в смысле Уоддингтона (см. гл. 2) представляется интересной аналогией.

a



6

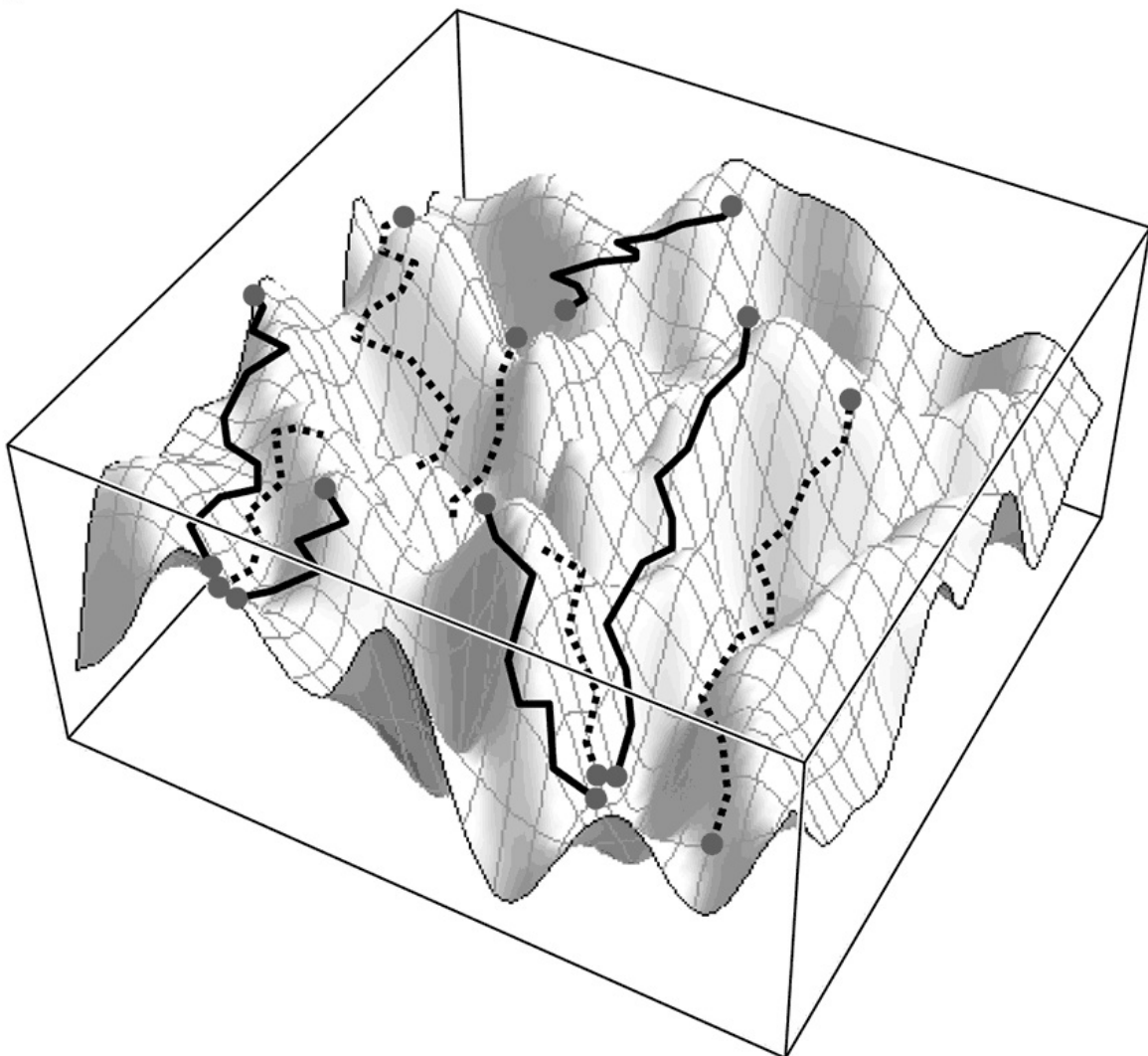


Рис. 13-2. Пересеченный адаптивный ландшафт и доступные эволюционные траектории: а — квазидетерминированная эволюция: канализация доступных траекторий; б — стохастическая эволюция: случайное распределение доступных траекторий. Сплошные линии — монотонно восходящие траектории, доступные для эволюции, движимой исключительно отбором. Пунктирные линии — немонотонные траектории, доступные только с участием генетического дрейфа.

Принято считать, что геном (генотип) определяет фенотип организма (с некоторым участием эпигенеза), фенотип жестко контролируется отбором, а фенотипические изменения не имеют эволюционных последствий. Сравнительная геномика и системная биология показывают, что все эти утверждения не являются истиной в последней инстанции, и такие упрощенные обобщения оставляют в стороне ключевые биологические явления. Отсутствие простой детерминированной связи между фенотипом и генотипом выражается по крайней мере в двух взаимодополнительных аспектах их взаимоотношений:

1. Фенотипические мутации и другие формы шума, такие как случайная транскрипция практически всех геномных последовательностей у животных, неотъемлемо присущи биологическим системам и вносят вклад в их эволюцию (см. гл. 9) [\[140\]](#). Эти эволюционно важные фенотипические изменения частично контролируются геномом, но связь между геномом и шумом стохастическая по своей природе.

2. Отображение генома на фенотип является неизоморфным и сложным (говоря проще, это многозначное отображение, то есть не отображение одного элемента в один, но многих в многие); все гены плейотропны, и все фенотипические свойства («функции» или «антравольты») зависят от активности многих генов. В целом геномфенотипическое отображение является исключительно сложным графом (см. гл. 5, в основном рис. 5–9). Ребра этого графа имеют различные веса, которые отображают разный вклад разных генов в один и тот же признак.

Повсеместное распространение и эволюционное значение фенотипической изменчивости делают связь генома с фенотипом принципиально недетерминированной. Многозначное отображение ограничивает эволюцию, возможно являясь даже ее главным ограничителем (см. предыдущий раздел), но при этом делает связь генома с фенотипом невероятно сложной, а реконструкцию фенотипа по геному крайне затруднительной. Некоторые простые фенотипические признаки, конечно, предсказуемы. Например, если бактерия не имеет лактозного оперона, она не сможет употреблять лактозу. Тем не менее даже такие простые признаки часто могут реализовываться несколькими разными путями. Любой сложный фенотип крайне трудно поддается предсказанию, например термофилия и устойчивость к радиации у прокариот, как мы видели в главе 5. Сложность связи генома с фенотипом и, как следствие, трудность построения надежных представлений о фенотипе организма только на основании анализа генома резко усугубляется для эукариотических и особенно многоклеточных организмов. Удивительное и труднообъяснимое, но на данный момент несомненное отсутствие сильной связи между очевидной биологической важностью генов и скоростью их эволюции (см. гл. 4) особенно подчеркивает возникающее понимание того, что фенотипические последствия эволюции генома нетривиальны и, в общем случае, трудно предсказуемы. Недооценка этой сложности может привести к нереалистичным надеждам на быстрый успех в проектах, направленных на исследование и манипуляции со сложными фенотипами, таких как изучение полногеномных ассоциаций, «война с раком» или индивидуализированные лекарства [\[141\]](#).

Эта книга прежде всего о концепциях, идеях и моделях, а не о методах. Тем не менее, прежде чем закончить эту последнюю главу, необходимо сказать несколько слов о новом поколении подходов, которые не только позволили поразительно глубоко проникнуть в ключевые эволюционные процессы, но и должны начать менять сам облик эволюционной биологии в следующем десятилетии (или близко к тому). Эти исследовательские стратегии подпадают под категорию «экспериментальной эволюционистики». В сегодняшних эволюционных экспериментах ход эволюции популяции организмов или молекул можно проследить непосредственно, обрабатывая, за счет применения нового поколения методов секвенирования, тысячи и даже миллионы молекул ДНК или РНК. Эксперименты Ричарда Ленски и его коллег по долгосрочной лабораторной эволюции популяций *E. coli*, о которых мы не раз говорили в этой книге, являются замечательным примером такого рода опытов (Ostrowski et al., 2008; Barrick et al., 2009; Woods et al., 2011). Эти эксперименты уже дали бесценную информацию о различных режимах отбора и дрейфа, распространенности параллельных мутаций, эволюции способности к эволюции и многом другом. Однако, благодаря появившейся в настоящее время возможности секвенирования тысяч полных бактериальных геномов, самые многообещающие результаты ожидаются в не столь отдаленном будущем, когда будут всесторонне изучены эволюционные траектории популяций в различных условиях окружающей среды и при разных давлениях отбора. Концептуально эти эксперименты продолжают направление исследований, начатое в пророческой работе Шпигельмана и коллег с РНК-бактериофагами в 1960-х годах (см. гл. 8). Эксперименты Шпигельмана почти на полвека опередили свое время и оказали сравнительно небольшое влияние на развитие науки, но в первые десятилетия XXI века статус экспериментального исследования эволюции быстро меняется.

Другое направление эволюционного экспериментирования включает в себя изучение ландшафтов отбора эволюционирующих белков или молекул РНК, которые мы кратко обсудили ранее в этой главе. Имеющиеся в настоящее время экспериментальные данные описывают только крошечные доли ландшафтов, но возможность исследовать большие области уже представляется реалистичной (Kogenaru et al., 2009; Loewe, 2009) [\[142\]](#). В конечном итоге подробная реконструкция полных ландшафтов отбора изменит наши представления о том, что значит «понимать» эволюционный процесс.

Дивный новый мир вирусов и прокариот

СТЭ ориентируется в основном на изучение эволюции животных и растений – многоклеточных эукариот, размножающихся чаще всего половым путем. Одноклеточные эукариоты и прокариоты, не говоря уж о вирусах, не считались важными для эволюционной биологии. Пожалуй, включение огромного мира микробов в систему эволюционных представлений следует считать главным толчком для перехода от СТЭ к «постсовременному» состоянию эволюционной биологии. Первые попытки разобраться в эволюционных связях между бактериями были в высшей степени обескураживающими, но последующее сравнение последовательностей консервативных генов таких как рибосомные РНК, а затем и полный анализ геномов привели к замечательным успехам. Сравнительная геномика бактерий и архей изменила центральные концепции эволюционной биологии, включая древо жизни (см. гл. 6), и показала крайне динамичный характер геномов и пангеномов (см. гл. 5).

Изучение мира вирусов привело к не менее, а может быть, и более драматическому изменению наших взглядов на эволюцию жизни на Земле. Крошечные паразиты, вирусы являются наиболее многочисленными и генетически разнообразными биологическими объектами на всей планете. Мир вирусов, по всей вероятности, существует с самых ранних, доклеточных стадий эволюции и постоянно взаимодействует с миром клеточных организмов, внося существенный вклад в его эволюцию, но сохраняя при этом собственную автономию.

Империи и домены жизни

В свое время Карл Вёзе, используя древо рРНК, классифицировал живое на три домена, что являлось огромным концептуальным скачком как для эволюционной биологии, так и биологии в целом. Положение вещей, однако, радикально изменилось за последние 30 лет, и в рамках этой классификации живых организмов не могут быть объяснены новые сложные реалии эволюции, открытые сравнительной геномикой. Первым важным открытием, противоречащим трехдоменной схеме, была демонстрация химерной природы генома эукариот. Трехдоменное древо отображает эволюцию только подмножества генов, вовлеченных в информационные процессы и количественно составляющих незначительное меньшинство эукариотических генов, даже внутри группы высококонсервативных генов, прослеженных до последнего общего предка всех эукариот. Конечно, доменная классификация живых организмов является лишь соглашением среди исследователей, поэтому классификация эукариот как отдельного домена не является правильной или неправильной сама по себе. Тем не менее такая классификация способна ввести в заблуждение, особенно в сопровождении тройственной схемы древа жизни, поскольку объединение двух организмов и их изначально отдельных геномов, которое, по всей видимости, дало начало эукариотам, в этой схеме не отражено. Схема же, включающая такое объединение в явном виде, с математической точки зрения деревом не является (см. рис. 13-3).

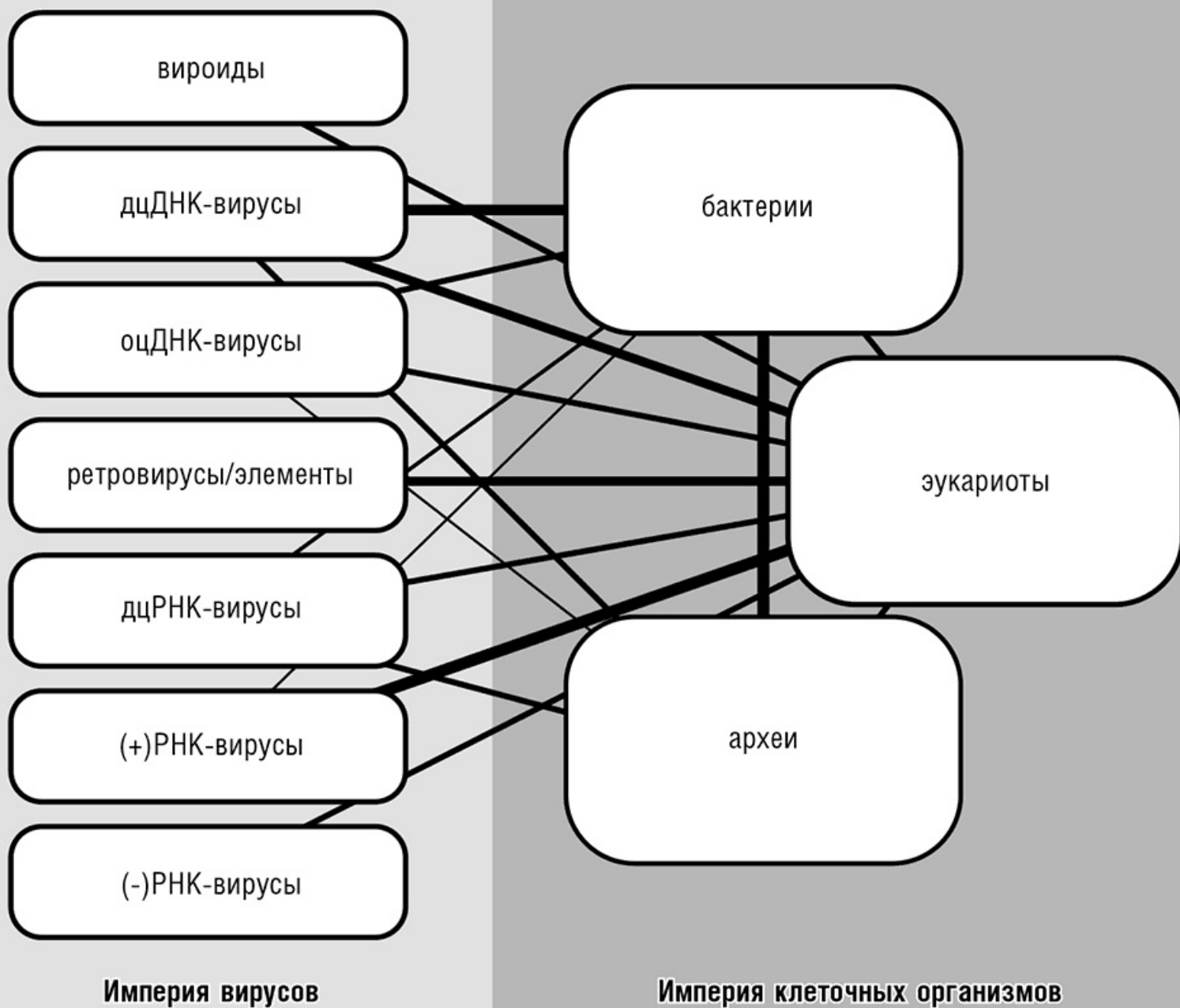


Рис. 13-3. Империи и домены жизни. Соединительные линии показывают потоки генетической информации между доменами, включая как вирусные инфекции, так и все пути горизонтального переноса генов.

Возможно, даже более фундаментальным отходом от трехдоменной схемы стало открытие мира вирусов с его неожиданным и ошеломляющим простором и удивительной эволюционной связностью [\[143\]](#). Как обсуждалось в главе 10, вирусоподобные паразиты неизбежно возникают в любой репликаторной системе, таким образом, не будет преувеличением утверждение, что *жизнь невозможна без присутствия вирусов*. Более того, кажется почти неизбежным, что доклеточная эволюция жизни прошла через вирусоподобную стадию развития. Также важно заметить, что главные группы вирусов различаются между собой не меньше (а возможно, и больше), чем три (или два) домена клеточных форм жизни. Эти различия обусловлены тем, что разные группы вирусов используют различные циклы репликации/экспрессии, в отличие от клеточных организмов, которые в этом отношении устроены идентично (подробности см. в гл. 10).

Любая классификация по сути всего лишь соглашение, в то время как эволюционные

сценарии стремятся реконструировать, по возможности, историю, которая на самом деле произошла. Очевидно, наилучшей классификацией живых организмов будет та, что наиболее точно отражает лучший из существующих эволюционных сценариев. С этой точки зрения наибольший смысл имеет разделение всех известных форм жизни на две «империи» [\[144\]](#): вирусов и клеточных форм (Koonin, 2010d). Эти две империи сильно отличаются друг от друга, но постоянно обмениваются генами (см. рис. 13-3). Можно говорить о трех или только о двух доменах клеточной жизни, но архео-бактериальное объединение должно быть необходимой частью классификационной схемы. Несмотря на то что нашей целью не является высказывать какие-либо формальные предложения относительно новых «вирусных доменов», отдельные большие группы вирусов, характеризующиеся общими наборами консервативных генов, несомненно, сравнимы по своему статусу с доменами клеточной жизни.

Парадокс биологической сложности, обманчивость прогресса и значение неадаптивных храповиков

Клетки и организмы как устройства для самовоспроизведения генов

По наблюдениям автора этой книги, многие исследователи и просто образованные читатели с пренебрежением относятся к концепции эгоистичного гена Ричарда Докинза (Dawkins, 2006) – во всяком случае, в ее крайних формах – возможно, потому, что она противоречит здравому смыслу и «унижает наше достоинство». Тем не менее, если мы принимаем тезис о том, что репликация генетического материала – основное свойство живых систем (см. гл. 2), у нас нет логического выхода, кроме как принять и концепцию эгоистичного гена. В частности, Докинз весьма решительно называет организмы транспортными средствами для репликации и эволюции генов, и я думаю, что эта простая идея описывает ключевой аспект биологической эволюции. Конечно, это не подразумевает никакого метафизического или телеологического контекста, например будто клетки и организмы существуют «с целью» репликации генов. «Целеориентированный» подход вообще неконструктивен (см. прил. I). Таким образом, предположение, будто фенотип существует с целью репликации, столь же бессмысленно, как и противоположное ему. Тем не менее между геномом и фенотипом существует фундаментальная, логически неизбежная асимметрия: все фенотипические признаки организмов, как и сами клетки и организмы, рассматриваемые как сложные физические системы, возникают и эволюционируют лишь ввиду того, что они способствуют процессу самовоспроизведения генома – либо увеличивая эффективность процесса, либо, по крайней мере, не уменьшая ее.

Как обсуждалось в главе 10, «сердцевину» фенотипа всех организмов в значительной мере составляют *антиэнтропийные устройства*, уменьшающие уровень ошибок при передаче информации во время самого процесса репликации и подчиненных ему процессов транскрипции, трансляции, укладки РНК и белков, а также снижающие вредные эффекты таких ошибок, когда они все-таки случаются. В остальном большая часть «сердцевины» фенотипа – это *вспомогательные устройства*, которые производят и транспортируют строительные блоки для репликации и сохранения фенотипа. Таким образом, невозможно отрицать, что эволюция фенотипа направлена на обеспечение репликации генов (генома).

С точки зрения одной лишь репликации, возникновение сложности являет собою загадку: почему существует так много живых форм, неизмеримо более сложных, чем это необходимо для минимального, простейшего приспособления для репликации? Мы не знаем точно, каким должно быть самое простое устройство для самовоспроизведения, но у нас есть замечательные кандидаты на это звание – простейшие автотрофные бактерии и археи, такие как *Pelagibacter ubique* и *Prochlorococcus sp.* Эти организмы обходятся приблизительно 1300 генами, не потребляют органических молекул и полностью независимы от других живых форм [\[145\]](#). Кстати, эти организмы также являются наиболее «успешными» на Земле. Они имеют самые большие популяции, которые эволюционировали под сильнейшим давлением отбора и в результате этого обладают наиболее «оптимизированными» геномами. Биосфера, состоящая из одних таких высокоэффективных одноклеточных организмов, легко представима, более того, земная биота до появления эукариот (то есть первые 2 миллиарда лет эволюции или около того) должна была быть очень похожа на такой мир, гораздо более, чем сегодняшняя биосфера (хотя более сложные прокариоты, несомненно, существовали и в то время).

Так почему же организмы столь сложно устроены?

Один из ответов, казалось бы, самоочевидный и принимавшийся биологами и всеми интересующимися эволюцией, состоит в том, что более сложные организмы являются также более приспособленными. Эта точка зрения, несомненно, ошибочна. Более того, и в этом состоит парадокс сложности, общая тенденция прямо противоположна: чем сложнее организм, тем меньший эффективный размер популяции он имеет и, по единственно разумному определению эволюционного успеха, тем менее успешным он является. Понимание этого может подсказать нам, что решение загадки появления сложности может быть поразительно простым: взглянув на ту же тенденцию с противоположной стороны, мы заметим, что чем меньше эффективный размер популяции, тем ниже интенсивность отбора и, следовательно, больше вероятность неадаптивного развития сложности. В этом и состоит суть популяционно-генетической неадаптивной концепции эволюции генома, предложенной Линчем.

Теперь мы можем сформулировать более конкретный ответ на вопрос в заголовке этого раздела – или, скорее, несколько взаимодополняющих ответов.

1. В самом общем смысле сложность появляется просто «потому, что может»: принимая во внимание универсальный процесс «беспорядочного блуждания», при наличии достаточного времени, стабильно возрастает вероятность случайного усложнения биологической организации.

2. Другое, более конкретное описание роли случайности проявления как определяющего фактора эволюции дается популяционной генетикой: сложность может возникнуть в результате случайной фиксации фактически нейтральных (слегка вредных) мутаций через генетический дрейф в популяциях с малым эффективным размером. Таким образом, из комбинации (1) и (2) получается, что сложность возникает, «потому что может», при условии слабого давления очищающего отбора, который не может отбраковать незначительно вредные изменения, такие как дупликация генов, включение мобильных элементов в нескольких сайтах генома и т. д.

3. На фоне этих случайных факторов эволюция сложной организации стала возможной благодаря *храповикам конструктивной нейтральной эволюции*: как только два или более гена делаются зависимыми друг от друга, в результате дифференциального накопления слегка вредных мутаций в обоих, оба они фиксируются в эволюции и оказываются неразрывно связанными, что приводит к повышенной сложности. Таков механизм эволюции генетических дупликаций, а также многих горизонтально перенесенных генов при субфункционализации; по-видимому, это один из основных путей эволюции.

4. Сложные формы обычно не более приспособлены, чем простые, однако сложность может облегчить адаптацию к новым нишам, как, например, в случае наземных растений. Появление сложности, таким образом, может нести в себе и адаптивную компоненту, в дополнение к основным неадаптивным факторам, упомянутым ранее.

Необходимо подчеркнуть существенную роль нейтральных храповиков в придании эволюции очевидной направленности без фактического повышения приспособленности. Храповик конструктивной нейтральной эволюции может быть ключевым элементом в появлении разнообразных сложных биологических свойств: храповики переноса гена от эндосимбионта к хозяину внесли существенный вклад в эукариогенез, а храповики необратимой потери генов явились ведущим фактором в редуktивной эволюции паразитов и симбионтов. Храповики можно рассматривать как узкие, крутые горные хребты на адаптивном ландшафте: как только эволюционирующая популяция оказывается на таком гребне, она начинает следовать квазидетерминированному пути, поскольку падение с хребта приводит к резкому снижению

приспособленности и неизбежному вымиранию. В этом процессе сложность может возникать без вклада адаптации.

Величайшая загадка происхождения жизни

Благодаря достижениям в областях геномики и системной биологии, за первое десятилетие XXI века мы накопили много больше знаний о ключевых аспектах эволюции, чем за предыдущие полтора века. Хотя крупные скачки эволюции, такие как происхождение эукариот или происхождение животных, остаются чрезвычайно сложными проблемами, появляется все больше и больше зацепок для их решения. Вне всякого сомнения, был достигнут значительный прогресс даже в этих трудных областях эволюционной биологии [\[146\]](#). Мы даже начали разрабатывать сценарии возникновения клетки, выходящие за рамки чистых спекуляций.

Тем не менее происхождение жизни, или, если быть более точным, происхождение первых репликаторных систем и происхождение трансляции, остается великой загадкой, и прогресс в решении этих проблем очень скромнен – и даже, если говорить о трансляции, незначителен. Существует несколько потенциально плодотворных наблюдений и идей, таких как открытие вероятных инкубаторов жизни – сетей неорганических ячеек в гидротермальных источниках – и химическая многосторонность рибозимов, поддерживающая гипотезу мира РНК. Однако эти достижения остаются только предварительными, пусть и важными, поскольку мы даже близко не подошли к убедительному сценарию добиологической эволюции от первых органических молекул до первой репликаторной системы и от них до собственно биологических объектов, в которых хранение информации и функция разделены между различными классами молекул (нуклеиновых кислот и белков соответственно).

На мой взгляд, несмотря на все достижения, эволюционная биология является и будет оставаться крайне неполной, пока нет хотя бы правдоподобного, пусть не полностью убедительного, сценария происхождения жизни. Поиск решения этой великой загадки может вести нас в неожиданных (и глубоко противоречащих, с точки зрения биологов, здравому смыслу) направлениях, в частности к полной переоценке важнейших понятий случайности, вероятности и возможного вклада чрезвычайно редких событий, что может быть проиллюстрировано на примере космологического подхода, рассматриваемого в главе 12.

Нужна ли и полезна ли новая теория биологической эволюции? Появится ли постсовременная синтетическая теория?

Принцип дополнительности Бора, как представляется автору этой книги, занимает центральное место в нашем понимании эволюции, прежде всего в смысле взаимной дополнительности между случайностью и детерминированными факторами (необходимостью) – той самой дополнительности, которая является лейтмотивом этой книги. Дополнительность проявляется на всех уровнях и во всех аспектах эволюции и может стать основным руководящим принципом на пути к новой теоретической биологии. Очевидные случаи, которые мы обсудили в этой книге, включают дополнительность между:

- случайными и (квази)направленными мутациями;
- отбором и дрейфом;
- эгоистичным и альтруистичным поведением разнообразных генетических элементов (персистентных вирусов, ретроэлементов, систем токсинов – антитоксинов и рестрикции – модификации и др.);
- устойчивостью и способностью к эволюции.

Дополнительность также имеет важное значение для эпистемологии новой эволюционной биологии. Учитывая ключевую роль исторической случайности (в том числе «концепции ремесленника» Жакоба) и огромную сложность биологических явлений, невозможно представить себе некий набор уравнений, представляющий собою общую теорию биологической эволюции, даже в том ограниченном смысле, в каком общая теория относительности Эйнштейна является теорией гравитации, а стандартная модель физики элементарных частиц – теорией материи и энергии. Кроме того, никакое сочетание простых физических и математических моделей не может описать эволюцию, так как ее недетерминированные, обусловленные историческим стечением обстоятельств компоненты напрямую не формализуемы. Лучшее, на что мы можем надеяться и над чем работать, – это новая, геномная форма популяционно-генетической теории, которая образует необходимую основу для всех последующих исследований эволюции генома и фенома. Такая теория – дело непростое, но, как мы попытались показать в этой книге, перспектива ее появления становится все более реалистичной. Более того, такая теория будет подкреплена и полным описанием в явном виде адаптивных ландшафтов для различных эволюционных режимов, изученных в прямых экспериментах. Тем не менее, при всей своей важности, такая теория никогда не сможет объяснить «всей эволюции», не более чем, скажем, статистическая физика может «объяснить» геологию. По-видимому, постсовременная синтетическая эволюционная теория может появиться только в виде *сложной системы взаимодополняющих взглядов, опирающихся на модели, заимствованные из статистической физики и популяционной генетики, а также на реконструкции фактического эволюционного прошлого.*

Словосочетание «постсовременная синтетическая теория эволюции», неоднократно повторенное в этой книге, – не просто фанатерия, но и очевидный оксюморон, ибо философия постмодернизма есть в первую голову отрицание самой возможности любого синтеза (см. прил. I). Тем не менее этот выбор слов вполне преднамеренный, так как сложность эволюции жизни вызывает призрак постмодернистского мировоззрения, сколь бы возмутительно это ни звучало. И тем не менее мы можем ожидать все более глубокого понимания предмета через новую систему взаимодополняющих, взаимодействующих моделей, теорий и обобщений. Любопытно

отметить, что некоторые из ведущих физиков-теоретиков сегодня размышляют о будущем физики в подобном свете [\[147\]](#).

Рекомендуемая дополнительная литература

Goldenfeld N., and C. Woese.(2007) Biology's Next Revolution. *Nature*445: 369.

Влияние «латеральной геномики» архей и бактерий на общие представления эволюционной биологии, в кратком изложении.

Gould, Stephen Jay.(1997) Full House: The Spread of Excellence from Plato to Darwin. New York: Three Rivers Press.

Блистательная книга, главная тема которой – разоблачение мифа прогресса в биологии. Гулд также обсуждает происхождение сложности чисто стохастическим путем (случайными блужданиями) и фундаментальную непредсказуемость эволюции.

Koonin E. V.(2009) Darwinian Evolution in the Light of Genomics. *Nucleic Acids Research*37: 1011–1034.

В статье, опубликованной к 200-летию Дарвина, обрисовываются новые открытия в различных областях, в частности переоценка древа жизни и демонстрации несостоятельности панадапционизма и градуализма, и отсюда потребность в новой синтетической теории. «Предварительный показ» настоящей книги.

Koonin E. V.(2010) The Two Empires and the Three Domains of Life in the Postgenomic Age. *Nature Education*3: 27.

Популярная статья, где вводится деление форм жизни на вирусную и клеточную империи.

Lewontin R. C.(2002) Directions in Evolutionary Biology. *Annual Review of Genetics*36: 1—18.

В статье подчеркивается множественность эволюционных факторов и недостаточность естественного отбора в качестве объяснения эволюции.

Goldenfeld N, Woese CR. Life is Physics: Evolution as a Collective Phenomenon Far From Equilibrium. Annu Rev Cond Mat Physics 2011; 2: 375–399.

По-видимому, эту статью можно считать научным завещанием Карла Вёзе. Эволюция здесь рассматривается как физический феномен, который можно и должно изучать с использованием методов физики твердого тела.

Koonin EV, Wolf YI. Evolution of microbes and viruses: a paradigm shift in evolutionary biology? Front Cell Infect Microbiol. 2012;2:119 Анализ вклада микробной геномики в развитие новой эволюционной биологии.

Приложение I. Философия постмодерна, метанарративы; природа и цели научных исследований

Я не просто дилетант, а скорее невежда в темах, затронутых в этом приложении. И все же краткий непрофессиональный обзор необходим. В самом деле, слово «постмодерн», неоднократно использующееся в этой книге, по-видимому, требует некоторого (псевдо) – философского пояснения. Эволюционная биология в целом – такой предмет, при обсуждении которого неизбежны экскурсы в эпистемологию.

Философия постмодерна, (не)доверие к метанарративам и (не)осуществимость синтеза

«Постсовременная синтетическая теория» в предисловии к данной книге – намеренный оксюморон. Действительно, основной пафос постмодернистской философии заключается в ее недоверии к любому обобщению, «большой картине», всеобъемлющей истории или теории всего, которые ученые и особенно философы склонны измышлять. Прочитав Жана-Франсуа Лиотара, одну из выдающихся фигур в постмодернистском поколении философов: «Упрощая до крайности, мы считаем «постмодерном» недоверие в отношении метанарративов. Оно является, конечно, результатом прогресса науки; но и прогресс, в свою очередь, предполагает это недоверие» (Lyotard, 1979). На постмодернистскую философию часто смотрят с презрением и еще чаще с насмешкой: в самом деле, попытки прочитать постмодернистский опус могут оставить несколько головокружительное впечатление. Тем не менее акцент постмодерна на множественности и разнообразии моделей отлично согласуется с последними результатами эволюционной биологии, показывающими чрезвычайно сложную гамму различных эволюционных процессов и сопротивляющимися любым попыткам втиснуть эволюцию в какую-нибудь «незатейливую» схему вроде естественного отбора случайных вариантов. Видные философы постмодерна Жиль Делёз и Феликс Гваттари говорят о ризоме как о ключевой метафоре пронизывающей весь мир множественности моделей (Deleuze and Guattari, 1987). Микробиолог Дидье Рауль заимствовал этот термин, чтобы назвать недавно открытую сложность эволюционных процессов *ризомой жизни* (Raoult, 2010), и трудно отрицать, что метафора подходит, особенно учитывая ревизию концепции древа жизни перед лицом вездесущего горизонтального переноса генов.

Проблема с постмодернизмом та же, что и с большинством философских систем. Он убедителен в критике, но не способен предложить конструктивную альтернативу. На самом деле, возможно к их чести, философы-постмодернисты отрицают саму необходимость таких альтернатив и, кажется, вполне удовлетворены, просто размышляя о ризоме. Однако наука – и эволюционная биология в частности – так работать не может. Чтобы был какой-то прогресс, мы вынуждены создавать нарративы и объединять их в метанарративы, которые философы науки (особенно физики) часто называют парадигмами, следуя классической «Структуре научных революций» Томаса Куна (Kuhn, 1962). Парадигмы и метанарративы необходимы хотя бы для того, чтобы противопоставлять им новые наблюдения и оценивать годность существующих парадигм и нужду в новых.

Все сказанное мной до сих пор о парадигмах достаточно общо и тривиально, но в случае эволюционной биологии есть и кое-что особенное. Учитывая, что многое в эволюционной биологии относится к уникальной истории единственной известной нам реализации жизни и что очень многое в этой истории есть дело обстоятельств и случая, сжатый метанарратив кажется принципиально невозможным. Лучшее, на что можно надеяться, – переплетение множества нарративов разного уровня абстракции. Говоря образно, любое описание течения эволюции имеет чрезвычайно высокую алгоритмическую (колмогоровскую) сложность (см. гл. 6) и потому сопротивляется любому обобщению. У эволюционной биологии есть, однако, и другой аспект. Давно известно, что простая популяционная генетическая теория способна довольно хорошо описать микроэволюционные процессы. Сравнительная геномика и системная биология добавили новые классы количественных переменных, которые могут использоваться для проверки моделей эволюции. Взятые вместе, эти модели могут составить другой тип метанарратива, который может претендовать на роль теории в том смысле, в котором этот

термин используется в физике (см. гл. 4). Если не сейчас, то, возможно, в будущем это направление в эволюционной теории и в самом деле будет законной частью физики.

Однако «полная физическая теория эволюции» остается принципиально иллюзорной целью. Теории физического типа – и снова разные для разных аспектов эволюционного процесса, а не единая всеобъемлющая теория – могут включать в себя только типичные аспекты эволюции, «необходимую» часть противопоставления Моно (даже если необходимость принимает вероятностную форму и таким образом сама зависит от случайности). Лучшая возможность «понять» эволюцию – примириться с тем, что взаимодополнительность физического (теоретического) и исторического метанарративов эволюции – устойчивое внутреннее свойство эволюционной биологии, а не временная ситуация, вызванная несовершенством теории. Широко распространено мнение, что исторические нарративы с их неизбежной склонностью к описательности в лучшем случае маргинальны в науке, своего рода «коллекционирование марок» [\[148\]](#); в худшем – ненаучные домыслы, учитывая, что редкие и уникальные события критически важны в эволюции, как подчеркнуто в этой книге. Я считаю эту позицию глубоко неудовлетворительной и несостоятельной в контексте эволюционной биологии. Вопреки всем трудностям и неизбежной неточности, связанной с уникальными событиями, они критически важны для эволюции жизни, так что эволюционные биологи должны сконцентрировать усилия на том, чтобы узнать как можно больше о каждом важном эволюционном переходе, включающем события такого рода. Я полагаю, пришло время понять, что физическая и историческая точки зрения на эволюционные процессы фундаментально отличаются, и последняя являются не «низшей», а дополнительной по отношению к первой.

Возвращаясь к постмодерну, следует ли нам доверять эволюционным метанарративам, хотя бы они и были необходимы для прогресса исследования? В определенном смысле ответ тривиально отрицательный: любая парадигма включает в себя упрощения, и старые парадигмы последовательно замещаются новыми под давлением накапливающихся новых данных. Именно это мы видим сегодня в эволюционной биологии: парадигма синтетической теории эволюции разрушается и уступает место новому взгляду, о котором и рассказывается в этой книге. Нет причин полагать, что какая бы то ни было парадигма может быть истиной, а тем более окончательным отражением «реальности». Это, конечно, относится к любым научным исследованиям, но эволюционные нарративы, кажется, обладают специфическими свойствами, которые требуют особой осторожности в интерпретации. Мы кратко остановимся на этих проблемах в следующих разделах.

Вопросы «почему?» и семантические ловушки: что на самом деле мы говорим об эволюции?

Многие исследования в эволюционной биологии сосредоточены на попытках ответить на разного рода вопросы «почему?». Эти вопросы изобилуют на всех уровнях исследований: от классической организменной биологии (почему самцы больше и сильнее самок у некоторых видов животных, но меньше у других?) до геномики (почему так много интронов в генах одних эукариотов, но мало в генах других?), системной биологии (почему некоторых белков синтезируется намного больше, чем других?) и проблемы происхождения жизни (почему в белках всех организмов ровно 20 аминокислот?). Эволюционные биологи часто (хотя далеко не всегда) избегают задавать вопрос «почему?» прямо, но этот вопрос, похоже, всегда присутствует за кадром и влияет на самую логику исследования. До недавних пор, а иногда и в наши дни любое «почему?» вызывало почти автоматическую реакцию в виде сочинения истории о приспособлении. «Сан-марковская» критика Гулда и Левонтина, нейтральная теория и, позднее, неадаптивная теория эволюции сложности Линча изменили ситуацию, так что теперь мы склонны к более сбалансированным, сложным повествованиям, которые в дополнение к отбору включают в себя и неадаптивные факторы, такие как дрейф генов, эффект «движения в повозке» и различные нейтральные храповики.

Но лучше ли предыдущих эти новые истории? С одной стороны, кажется, что да, потому что они принимают во внимание вклад многих процессов и, по крайней мере в самых тщательных исследованиях, этот вклад устанавливается измерением некоторых величин, а не исходя из качественных рассуждений. Тем не менее все эволюционные сценарии, сконструированные с целью ответить на вопрос «почему?», следует считать тем, чем они и являются: нарративами, сочиненными учеными. По самой своей природе нарративы непременно упрощают и редуцируют исследуемый феномен – в нашем случае эволюционный процесс – до небольшого числа дискретных «факторов». Эти факторы, как, например, естественный отбор, являются абстракциями, выведенными из наблюдений. Эволюционные биологи измеряют не отбор, а величины некоторых специфических переменных, таких как K_{ai} и K_s . Из соотношений между этими измеряемыми величинами делаются выводы относительно очищающего отбора, положительного отбора и нейтральности, и конструируется метанарратив.

«Диалектика» этой ситуации состоит в том, что эволюционные нарративы, безусловно, являются упрощенными «мифами», которые имеют огорчительный (и в современных исследованиях непреднамеренный) телеологический привкус (например, «отобранный для» или, что еще хуже, «отобранный для такой-то цели»), и все же язык этих нарративов, похоже, лучше всего подходит для описания эволюции и формулировки опровержимых гипотез, которые стимулируют дальнейшие исследования. На данный момент мы вряд ли можем отказаться от этих историй (в самом деле, большая часть этой книги написана как раз в такой манере) именно потому, что они необходимы для прогресса в исследованиях, хотя бы они и оставляли ученых с чувством неловкости и неудовлетворенности. Важно не забывать, что эволюционные нарративы – это семантические приспособления, сконструированные для структурирования и упрощения нашего мышления об эволюции и для стимулирования новых гипотез. Этим нарративам следует осмотрительно не доверять и ни в коем случае не считать их «точным представлением реальности» (что бы это ни значило – см. следующий раздел).

Интересен вопрос, будет ли способна эволюционная биология в не очень отдаленном будущем разработать новый язык, в котором будет меньше мифа и больше измеряемых величин.

Такая надежда не выглядит невероятной. В конце концов, язык современных эволюционных нарративов, который прочно включает в себя, скажем, различие между очищающим и позитивным отбором, формализм популяционной генетики, структуру ландшафтов приспособленности или скорость горизонтального переноса генов, гораздо ближе к конкретным измерениям, чем язык синтетической теории эволюции, не говоря уже о языке Дарвина.

Таким образом, мы не можем ожидать, чтобы эволюционная биология (да и любая область науки) избавилась от (мета)нарративов. Однако новые нарративы, похоже, заметно «лучше», чем старые – то есть они меньше упрощают и лучше соотносятся с наблюдениями.

Природа и цели науки: зачем вообще изучать эволюцию?

Вопрос в заголовке раздела может выглядеть абсурдным, но, хотя в конце этой книги он может быть задан только в шутку, в принципе это законный и важный вопрос, который заслуживает честного, серьезного размышления. Дадим для начала тривиальный, но необходимый ответ: изучение эволюции жизненно важно для биологии хотя бы только потому, что такие эволюционные концепции, как переменные ограничения и очищающий отбор, лежат в основании огромной части экспериментов в современной биологии. В самом деле, каждый эксперимент по направленному мутагенезу основан на эволюционном мышлении: только эволюционный анализ может подсказать экспериментатору, какие позиции в гене нужно мутировать, чтобы повлиять на его активность каким-то конкретным образом, даже если исследователь не думает при этом в терминах эволюции. Еще более тонкий эволюционный анализ необходим, скажем, в изучении эволюции вирусов или развития рака, так что знание определенных аспектов эволюции буквально спасает тысячи жизней и экономит миллионы долларов (например, благодаря предсказанию эпидемий гриппа и улучшению вакцин).

Однако, несмотря на всю его биологическую разумность, это поверхностный ответ. Во всех этих исследованиях модели эволюции могут использоваться, а зачастую и используются, как любой другой инструмент или метод, безотносительно «эволюционной реальности». Итак, интересно ли нам, как шла эволюция «на самом деле» и что в действительности происходило в глубоком прошлом с жизнью на Земле? Задавая эти вопросы, мы касаемся глубочайших проблем природы и целей всей науки. По-видимому, самый распространенный взгляд состоит в том, что наука стремится понять, как «работает» мир, в котором мы живем. Однако сам смысл «понимания реальности» не слишком ясен. Все, что может сделать наука, – это разработать модели, часто (но не обязательно) в форме уравнений, и посмотреть, опровергают ли наблюдения эти модели, – или, другими словами, оценить их предсказательную силу. Научный процесс не говорит нам о мире ничего прямо; он говорит лишь о совместимости конкретных наблюдений с принятыми моделями. Все аспекты любого мировоззрения («картины реальности») могут рассматриваться как *метафизические следствия* моделей и в этом качестве могут считаться несущественными.

В случае эволюционной биологии можно использовать математическую теорию, описывающую связь между данными (в первую очередь это результат секвенирования, но также, например, сравнительная экспрессия или данные протеомики), для предсказания фенотипического результата мутаций или появления новых вирусных штаммов с конкретными свойствами, не обращаясь к какой-либо «реалистической» картине процесса эволюции. Это не шутка – я пытаюсь честно изобразить, как видят природу научного процесса и научную картину мира многие, если не большинство физиков и философов науки, включая неоспоримо выдающиеся фигуры. Стивен Хокинг и Леонард Млодинов в своей последней популярной книге о физике и космологии (Hawking and Mlodinow, 2010) очень удачно назвали эту позицию «модельно-зависимый реализм». Согласно этой точке зрения, ученые конструируют модели и конкурирующие модели сравниваются по своей способности объяснять данные и предсказывать результаты экспериментов. Модель, которая точнее всех объясняет наибольший массив наблюдений и делает это с максимально возможной простотой (но не проще), становится победителем (обычно до тех пор, пока не проигрывает новой, еще более точной и элегантной модели). Хокинг и Млодинов отчеканили броский термин «модельно-зависимый реализм» в 2010 году, но эта точка зрения, конечно, гораздо старше и общепринята среди физиков. Нильса Бора, например, цитируют так: «Нет квантового мира. Есть лишь абстрактное квантово-

физическое описание. Неправильно думать, что задачей физики является ответить на вопрос, откуда взялась природа... физику беспокоит вопрос, что мы можем сказать о природе» (Pais, 1994).

Можно многое сказать в пользу модельно-зависимого реализма. Чтобы защитить этот серьезный взгляд на науку, физики (включая не только Хокинга и Млодинова, но и Ричарда Фейнмана) обращаются к сравнению мифологических идей стабильности Солнечной системы и соображений, выдвигаемых птолемеевой и ньютоновой физиками. Древнегреческие астрономы заменили мифологию рациональной, но произвольной концепцией эпициклов, множества сфер, обращающихся вокруг неподвижной Земли с прикрепленными к ним небесными телами. Коперник и Кеплер заменили схему эпициклов моделью планет, обращающихся по эллиптическим орбитам вокруг неподвижного Солнца. Ньютон представил теоретическое обоснование для этой модели в виде своего закона гравитации, согласно которому сила притяжения держит тела на стабильных траекториях. Можно утверждать, однако, как предложил Фейнман, что Ньютонова картина мира столь же мифологична, как и Птолемеева, а то, что мы считаем иначе, – дело привычки. В самом деле, что это за «силы», которые будто бы действуют в пустом пространстве на расстоянии, и чем они лучше, скажем, богов, предпринимающих некоторые периодические действия, позволяющие миру продолжать свое существование? Если обдумать это трезво, силы настолько же непонятны, как и боги. Действительно, сам Ньютон превосходно выразился, что он не «измышляет гипотез», утверждение, которое надо понимать в том смысле, что сэр Исаак обдуманно отказывается сказать что-то о том, «как устроен мир», – и очень правильно отказывается, считает Фейнман. В модельно-зависимом реализме наука разрабатывает модели, которые затем сравниваются с наблюдениями; победителем среди конкурирующих моделей становится та, что лучше остальных совпадает с наблюдениями, точнее других предсказывает новые экспериментальные результаты и проста настолько, насколько это возможно (но не проще). «Истинность» модели (ее способность описывать «реальность») не является частью концепции науки – важны предсказательная сила, элегантность и простота [\[149\]](#).

Я возражаю против мировоззрения модельно-зависимой реальности. Хотя все «картины реальности» являются мифами, Ньютонов миф все-таки лучше Птолемеева, потому что он включает меньше произвольных допущений. В конце концов, Ньютон постулирует небольшое количество сущностей, таких как гравитация и масса, которые, как он открыто признает, невозможно объяснить, в противоположность бесконечной последовательности случайных сущностей – эпициклов, постулируемых Птолемеевой космогонией. Ньютоново мировоззрение, хотя и включает сущности, недоступные для «понимания», представляется более экономным и менее искусственным, чем предыдущие взгляды. Далее можно указать, что эйнштейновская интерпретация гравитации в общей теории относительности представляет собой следующий шаг вперед: Эйнштейн предложил физически правдоподобное описание до того таинственной «силы» в терминах деформации пространства-времени – не только потому, что общая теория относительности лучше объясняет конкретные тонкие эффекты гравитации. Я полагаю, что этот аспект эволюции физики важен для нашего понимания функционирования науки в целом. Никакая модель не может претендовать на точное отображение «реальности», которая принципиально неизвестна; однако последовательные модели мира не только увеличивают точность предсказаний, но и описывают нашу ускользающую реальность все менее нелепо и все более правдоподобно. Другими словами, если сформулировать просто и резко, фраза «новые модели лучше описывают реальность, чем предыдущие» имеет некоторый смысл.

Интересно, что физики, которые во взглядах на науку и природу полностью придерживаются модельно-зависимого реализма, склонны выражать, по другим причинам и,

возможно, в более сильных формах, ту же позицию, что и (антинаучные) философы-постмодернисты: «глобальные картины» (метанарративы) считаются полностью неуместными (или, по меньшей мере, не рассматриваются как часть науки). Я полагаю, что такая позиция распространяет скептицизм так далеко, что становится нереалистичной и непродуктивной для развития науки. Более сильная форма реализма, чем та, что принята в модельно-зависимой версии, особенно важна в областях науки, которые являются частично историческими, – тех, где речь идет о событиях, невозпроизводимых в прямых экспериментах, из которых некоторые вполне могут быть уникальными (по крайней мере, в наблюдаемой части Вселенной). Мы изучаем эволюцию не только ради конкретных предсказаний, сколь бы они ни были важны, но, скорее, чтобы получить некое «понимание» истории жизни и ее фундаментальных тенденций, тех, что могут быть неотъемлемо присущи жизни в целом и снова проявятся в других реализациях жизни, если таковые будут когда-либо обнаружены.

Философ сэра Карл Раймунд Поппер, основатель парадигмы фальсифицируемости (опровержимости) в эпистемологии, был изначально настроен чрезвычайно скептически по отношению к теории Дарвина из-за ее очевидной нефальсифицируемости, настолько, что объявлял дарвинизм «ненаучным». Позднее, однако, Поппер изменил свою позицию и предположил, что, хотя дарвинизм не является фальсифицируемой теорией сам по себе, он является метафизической программой, способной породить великое множество фальсифицируемых гипотез. В этом контексте Поппер не использовал слово «метафизический» в качестве уничтожающей характеристики; напротив, он считал эту программу плодотворной и продуктивной, и даже необходимой. Он только имел в виду, что дарвиновская концепция эволюции в целом не является фальсифицируемой (и, вероятно, не верифицируемой). Поппер был довольно красноречив на этот счет, хотя его понимание эволюционной биологии было скорее поверхностным:

«Но все же теория бесценна. Я не вижу, как без нее наше знание могло бы вырасти настолько, как это случилось после Дарвина. При попытках объяснить эксперименты с бактериями, которые адаптируются, например, к пенициллину, становится ясно, что теория естественного отбора может нам сильно помочь. Хотя она является метафизической, теория проливает свет на очень конкретные и практические исследования. Она позволяет нам рационально изучать адаптации к новым условиям окружающей среды (таким как среды, загрязненные пенициллином): она предполагает существование механизма адаптации и даже позволяет изучать его в деталях. И это единственная теория на настоящий момент, которая все это делает» (Popper, 1982).

Я полагаю, в позиции Поппера есть значительные достоинства. И относительно «современной синтетической теории» эволюционной биологии, и относительно дарвиновской теории одинаково верно, что фальсификация всех их концептуальных основ или всех их утверждений едва ли возможна. Однако, если не обращать внимание на ее метафизический характер как целого, эта теория порождает множество конкретных фальсифицируемых утверждений, особенно учитывая быстро увеличивающийся объем данных в геномике и системной биологии. Более того, я думаю, что рассмотрение процессов эволюции с дополнительных ракурсов (см. гл. 13) может привести нас настолько близко к пониманию эволюции, как она «на самом деле» происходила, и жизни, как она «в действительности» развивалась, насколько это вообще возможно.

Приложение II. Эволюция космоса и жизни: вечная инфляция, теория «мира многих миров», антропный отбор и грубая оценка вероятности возникновения жизни [\[150\]](#)

Краткое введение в инфляционную космологию для неспециалистов

Теория «мира многих миров» (МММ), принципиально важная для космологического взгляда на происхождение жизни, представленного в главе 12, является следствием инфляционной космологии. В конце XX века она заменила собою классическую модель Большого взрыва. Инфляция – период начального, экспоненциально быстрого расширения Вселенной (Guth, 2001, 1998b). Алан Гут разработал инфляционную модель, чтобы объяснить некоторые ключевые астрономические факты, необъяснимые в рамках космологии Большого взрыва:

- плоскую геометрию наблюдаемой части Вселенной (метagalактики);
- крупномасштабную однородность космического реликтового излучения;
- локальные неоднородности реликтового излучения;
- отсутствие наблюдаемых магнитных монополей.

В самых правдоподобных, самосогласованных инфляционных моделях инфляция вечна, с бесконечным количеством островных (дочерних) вселенных (далее просто вселенных), возникающих при распаде небольших областей первичного «океана» ложного (высокоэнергичного) вакуума и составляющих вместе бесконечную мультивселенную (Guth, 2001; Vilenkin, 2007). Предсказания, полученные на основе инфляционных моделей, находятся в превосходном количественном согласии с указанными выше фактами (Guth and Kaiser, 2005). Более того, версия теории струн с населенным ландшафтом независимо приводит к похожей модели мультивселенной (Bousso, 2006; Bousso and Polchinski, 2004; Susskind, 2003, 2006a) [\[151\]](#). Таким образом, хотя модель вечной инфляции не может считаться доказанной, на данный момент она является наиболее предпочтительным сценарием космической эволюции. Наблюдателям, находящимся внутри, их вселенная кажется замкнутой и бесконечной, и она содержит в себе бесконечное число не связанных между собой областей (метagalактик). Для таких наблюдателей (вроде нас) их вселенная расширяется из сингулярности (Большой взрыв), которая соотносится с концом инфляции в данной части мультивселенной.

Модель МММ утверждает, что все макроскопические, «крупнозернистые» последовательности событий, которые не запрещены физическими законами сохранения, уже произошли (или произойдут) где-то в бесконечной мультивселенной (и даже в одной из островных вселенных) – и при этом не единожды, а бесконечное число раз (Garriga and Vilenkin, 2001; Vilenkin, 2007). Например, существует бесконечное число (макроскопически) точных копий Земли, включая все, что на ней находится, хотя вероятность, что конкретная наблюдаемая область вселенной (далее Н-область) содержит такую копию, будет бесконечно малой. Такая картина абсолютно противоречит интуиции, но является прямым следствием вечной инфляции (Guth, 2001; Linde, 1986; Vilenkin, 1983).

Гаррига и Виленкин показали, что за конечное время содержимое каждой Н-области может принять конечное количество состояний; соответственно, у любой Н-области может быть конечное, хотя и невероятно большое (порядка 10 в степени 10^{150}) число уникальных, макроскопических, «крупнозернистых» историй (Garriga and Vilenkin, 2001). В сущности, конечность числа крупнозернистых историй является следствием квантовой неопределенности (принципа Гейзенберга; Carroll, 2010; Vilenkin, 2007). Такой же вывод независимо получается при совершенно другом подходе: через так называемый голографический предел,

устанавливающий максимальное количество энтропии, которое может содержаться в любой конечной области вселенной (’t Hooft, 1993; Bouso, 2006; Carroll, 2010; Garriga and Vilenkin, 2001). Все это – вечная инфляция, конечность числа уникальных крупнозернистых историй и неизбежные квантовые флуктуации в момент Большого взрыва (точка отсчета для каждой вселенной) – ведет к простому, но поразительному заключению, что каждая история, допустимая физическими законами сохранения, повторяется бесконечное число раз в мультивселенной и, более того, в каждой из бесконечного числа бесконечных (островных) вселенных (Bouso, 2006; Vilenkin, 2007).

Теория МММ тесно связана с антропным принципом (антропным отбором), спорной, но все более популярной среди космологов концепцией. Согласно антропному принципу, единственной «причиной», по которой нашей метagalактике присущи ее индивидуальные свойства, является то, что иначе не существовало бы наблюдателя, который бы мог видеть Вселенную (Barrow and Tipler, 1988; Livio and Rees, 2005; Rees, 2001). Следует особо подчеркнуть, что я обсуждаю здесь только так называемый «слабый» антропный принцип, единственную форму этой концепции, сколь-нибудь приемлемую для научного рассмотрения. «Сильный» антропный принцип – это телеологическая идея, согласно которой наше (человеческое) существование есть, в некотором таинственном смысле, цель эволюции Вселенной (Barrow and Tipler, 1988); в таком виде эту концепцию невозможно отнести к науке. Похоже, антропный принцип может быть реалистически определен только в контексте огромной (или бесконечной) мультивселенной (Susskind, 2006b). В частности, в теории МММ у антропного отбора есть прямая интерпретация: свойства нашей метagalактики отобраны среди огромного числа наборов параметров, существующих в мультивселенной (каждый в виде бесконечного числа копий), поскольку они благоприятны для возникновения и развития сложных форм жизни.

По сравнению с более старыми космологическими концепциями, которые рассматривали конечную Вселенную, теория МММ меняет сами понятия «возможно», «вероятно» и «случайно», относящиеся к любому историческому сценарию (см. табл. II-1). Проще говоря, вероятность осуществления любого сценария, не нарушающего законы сохранения, в бесконечной Вселенной строго равна 1. Верно и обратное: вероятность того, что данный конкретный сценарий реализован в данной метagalактике, равна частоте таких сценариев во Вселенной. Слегка изменив точку зрения, мы придаем формулировке второго начала термодинамики, вместо привычного статистического, буквальный смысл: любое нарушение этого закона, допустимое другими законами сохранения, будет иметь место – и при этом бесконечное число раз. Таким образом, спонтанное зарождение сложных систем, которое должно бы в конечной Вселенной рассматриваться как практически невозможное, становится не просто возможным, но неизбежным в теории МММ, пусть даже вероятности огромного большинства историй и были бы в данной метagalактике исчезающе малыми. Новое могущество случая, подкрепленное антропным отбором, должно иметь глубокие последствия для нашего понимания любого феномена во Вселенной, и жизнь на Земле не может быть исключением.

Вероятность случайного возникновения различных революционных систем в Н-области: грубая прикидка верхних пределов

Общие предположения: в Н-области содержится 10^{22} звезд, у каждой десятой есть пригодная для жизни планета; то есть имеется 10^{21} таких планет (несомненно, это сильное преувеличение; в действительности большинство звезд не имеет планет вовсе, не говоря о пригодных для жизни). Каждая планета размером с Землю, у каждой имеется пригодный для обитания слой толщиной 10 км (10^6 см); отсюда объем этого слоя $4/3\pi [R^3 - (R-l)^3] \approx 5 \times 10^{24} \text{ см}^3$, где R – радиус планеты, l – толщина обитаемого слоя. Синтез РНК происходит в 1 % объема обитаемого слоя – то есть в объеме $5 \times 10^{22} \text{ см}^3$ (опять сильное преувеличение – в действительности «фабрика РНК» будет очень мало). Положим концентрацию нуклеотидов в объеме V и скорость синтеза молекул РНК размера n (свободный параметр, зависящий от специфики модели революционной стадии, далее n -мер) за 1 молекулу/ $\text{см}^3/\text{сек}$ (и снова сильное преувеличение для любой молекулы сколько-нибудь значительного размера; более того, не учтена обратная зависимость от n , которая должна быть достаточно сильной). Время после Большого взрыва в данной Н-области (как верхний предел) для всех планет $10^{10} \text{ лет} \approx 3 \times 10^{17} \text{ секунд}$. Тогда количество уникальных n -меров, опробованных за время после Большого взрыва, будет:

$$S \approx 5 \times 10^{22} \times 10^{21} \times 3 \times 10^{17} \approx 1,5 \times 10^{61}$$

Предположим, что для начала биологической эволюции требуется уникальный n -мер. Количество возможных последовательностей, состоящих из n нуклеотидов, $N = 4^n \approx 10^{0,6n}$.

Можно ожидать, что уникальный n -мер возникнет в Н-области Ераз:

$$E = S/N = 1,5 \times 10^{61} / 10^{0,6n}$$

и

$$n = \log(E \times 1,5 \times 10^{61}) / 0,6$$

Подставив $E = 1$, получаем $n \approx 102$ (нуклеотида). Заметим, что, так как величина n прямо пропорциональна логарифму S , оценка будет мало зависеть от начальных предположений о величине переменных; например, изменение S на порядок величины приведет к увеличению или уменьшению n менее чем на 2 нуклеотида.

Можно представить себе рибозим-репликазу, состоящую из приблизительно ста нуклеотидов; таким образом, в принципе спонтанное появление таковой в конечной вселенной, состоящей из единственной Н-области, нельзя исключать в нашей «игрушечной» модели (и снова, скорость синтеза РНК, принятая здесь, намеренно сильно переоценена).

Для появления примитивной системы сопряженной репликации-трансляции, что в данном контексте рассматривается как революционная стадия, требования гораздо жестче. Как

минимум, необходимо спонтанное появление следующего:

- Две рРНК, с общим размером не менее 1000 нуклеотидов.
- Примерно 10 примитивных адаптеров по 30 нуклеотидов каждый, в целом около 300 нуклеотидов.
- По меньшей мере одна РНК, кодирующая репликазу, размером примерно 500 нуклеотидов (оценка снизу). В принятой модели, $n = 1800$, и в результате $E < 10^{-1018}$.

Другими словами, даже в нашей игрушечной модели, которая предполагает сильно преувеличенную скорость синтеза РНК, вероятность случайного зарождения системы трансляция – репликация в единственной Н-области будет $P < 10^{-1018}$. Очевидно, эта версия революционной стадии может рассматриваться только в контексте вселенной с бесконечным (или, по меньшей мере, очень большим) количеством Н-областей.

Модель, рассмотренная здесь, ни в коем случае не предполагалась реалистичной. Она только иллюстрирует разницу в требованиях, накладываемых на вероятность возникновения разных версий революционных систем, и следовательно, связь этой версии с разными космологическими моделями вселенной.

Таблица II-1. Новые определения и новые интерпретации известных определений в модели МММ.

Термин	Определение
Инфляция	Экспоненциальное расширение мультивселенной под давлением отталкивающей гравитации ложного (высокоэнергетического) вакуума. Возможно, инфляция вечна — то есть, раз начавшись, она никогда не закончится
Мультивселенная (мультимир, метавселенная, мультиверс)	Вся ткань реальности, состоящая из бесконечно расширяющегося ложного вакуума с бесконечным числом небольших распадающихся областей, дающих начало вселенным
Вселенная (островная вселенная, дочерняя вселенная)	Часть мультивселенной, расширяющаяся с момента Большого взрыва, состоящего в локальном переходе ложного вакуума в низкоэнергетический (истинный) вакуум. Вселенная бесконечна с точки зрения наблюдателя, находящегося внутри ее, но конечна для воображаемого внешнего наблюдателя
Наблюдаемая область (Н-область)	Конечная область в пределах вселенной, доступная для наблюдений из выбранной точки (пространство, сигнал из которого может достичь выбранной точки за время существования вселенной). Наша Н-область содержит примерно 10^{20} звезд
Большой взрыв	В традиционной космологии XX века расширение Вселенной из сингулярности. Природа «взрыва» никогда не была разъяснена. В космологии вечной инфляции Большой взрыв соответствует концу инфляции данной области мультивселенной, в результате распада ложного вакуума и формирования Вселенной в форме расширяющегося пузыря низкоэнергетического (истинного) вакуума
Макроскопическая (крупнозернистая) история	Любая последовательность физических событий, допустимых законами физики, определенная с точностью до квантовой неопределенности и произошедшая в какой-либо метагалактике за конечное время. Было показано, что число всех возможных макроскопических историй конечно, хотя и огромно. Отсюда следует, что даже в пределах единственной вселенной каждая история повторяется бесконечное число раз



Термин	Определение
Вероятность, случайность, возможность	<p>Учебники определяют вероятность как предел, к которому стремится частота конкретного события, когда количество испытаний стремится к бесконечности.</p> <p>В бесконечной вселенной (и, очевидно, в мультивселенной) с конечным числом историй реализуется бесконечное число испытаний, следовательно, вероятность равна частоте.</p> <p>Вероятность любой допустимой истории, включая зарождение жизни, поэтому $P = 1$. Однако вероятность p наблюдения какой-либо конкретной истории в конкретной N-области лежит в интервале от 0 до 1, как и указано в учебниках, и может быть чрезвычайно мала для огромного числа историй, включая происхождение жизни.</p> <p>Таким образом, понятия возможности и случайности в их традиционной форме применимы только к конечным областям вселенной, тогда как во всей бесконечной вселенной необходимо реализуются все допустимые истории</p>
Антропный принцип, антропный отбор, антропное мышление	<p>Идея того, что история нашего мира (нашей N-области, нашей Галактики, нашей Солнечной системы и так далее) не зависела от какого-то специального «механизма», но была просто «выбрана» из бесконечного набора всех историй, гарантированно реализовавшихся в бесконечной вселенной, благодаря тому, что оказалась благоприятной для возникновения сложной жизни.</p> <p>Антропный отбор — эпистемологический, а не онтологический принцип, и его не надо воспринимать как активный процесс какого-либо рода. Такова формулировка слабого антропного принципа, принятого в данной работе. Сильный антропный принцип предполагает возникновение сознания как цель космической истории; это концепция телеологическая и ненаучная</p>

- (1948) *O Polozhenii V Biologicheskoi Nauke. Stenograficheskii Otchet Sessii Vsesoyuznoi Akademii Selskohozyastvennyh Nauk Imeni V. I. Lenina. (On the Situation in Biological Science. A Transcript of the Session of the V. I. Lenin All-Union Academy of Agricultural Sciences, July 31-August 7, 1948.* Moscow, USSR: The State Agricultural Literature Publishers.
- 't Hooft, G.(1993) Dimensional Reduction in Quantum Gravity. gr-qc/9310026.
- Adami C.(2002) What Is Complexity? *Bioessays*24: 1,085—1,094.
- Adl, S. M., A. G. Simpson, M. A. Farmer, R. A. Andersen, O. R. Anderson, J. R. Barta, S. S. Bowser, G. Brugerolle, R. A. Fensome, S. Fredericq, T. Y. James, S. Karpov, P. Kugrens, J. Krug, C. E. Lane, L. A. Lewis, J. Lodge, D. H. Lynn, D. G. Mann, R. M. McCourt, L. Mendoza, O. Moestrup, S. E. Mozley-Standridge, T. A. Nerad, C. A. Shearer, A. V. Smirnov, F. W. Spiegel, and M. F. Taylor. (2005) The New Higher Level Classification of Eukaryotes with Emphasis on the Taxonomy of Protists. *J Eukaryot Microbiol*52: 399–451.
- Agol V. I.(1974) Towards the System of Viruses. *Biosystems*6: 113–132.
- Agol V. I., and A. P. Gmyl. (2010) Viral Security Proteins: Counteracting Host Defences. *Nat Rev Microbiol*8: 867–878.
- Ahlquist P.(2006) Parallels among Positive-Strand RNA Viruses, Reverse-Transcribing Viruses and Double-Stranded RNA Viruses. *Nat Rev Microbiol*4: 371–382.
- Alic N., N. Ayoub, E. Landrieux, E. Favry, P. Baudouin-Cornu, M. Riva , and C. Carles.(2007) Selectivity and Proofreading Both Contribute Significantly to the Fidelity of RNA Polymerase III Transcription. *Proc Natl Acad Sci USA*104: 10,400—10,405.
- Allen, E. E., and J. F. Banfield.(2005) Community Genomics in Microbial Ecology and Evolution. *Nat Rev Microbiol*3: 489–498.
- Alperovitch-Lavy, A., I. Sharon, F. Rohwer, E. M. Aro, F. Glaser, R. Milo, N. Nelson, and O. Beja.(2011) Reconstructing a Puzzle: Existence of Cyanophages Containing Both Photosystem-I and Photosystem-II Gene Suites Inferred from Oceanic Metagenomic Datasets. *Environ Microbiol*13: 24–32.
- Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller , and D. J. Lipman.(1997) Gapped Blast and Psi-Blast: A New Generation of Protein Database Search Programs. *Nucleic Acids Res*25: 3,389—3,402.
- Altstein, A. D.(1987) Origin of the Genetic System. *Mol Biol*21: 309–322.
- Amaral, P. P., M. E. Dinger, T. R. Mercer, and J. S. Mattick.(2008) The Eukaryotic Genome as an RNA Machine. *Science*319: 1,787—1,789.
- Anantharaman, V., E. V. Koonin, and L. Aravind.(2002) Comparative Genomics and Evolution of Proteins Involved in RNA Metabolism. *Nucleic Acids Res*30: 1,427—1,464.
- Anfinsen, C. B.(1973) Principles That Govern the Folding of Protein Chains. *Science*181: 223–230.
- Aquadro, C. F.(1997) Insights into the Evolutionary Process from Patterns of DNA Sequence Variability. *Curr Opin Genet Dev*7: 835–840.
- Aravind, L., V. Anantharaman, S. Balaji, M. M. Babu, and L. M. Iyer.(2005) The Many Faces of the Helix-Turn-Helix Domain: Transcription Regulation and Beyond. *FEMS Microbiol Rev*29: 231–262.
- Aravind, L., L. M. Iyer and E. V. Koonin. (2006) Comparative Genomics and Structural Biology of the Molecular Innovations of Eukaryotes. *Curr Opin Struct Biol*16: 409–419.
- Aravind, L., and E. V. Koonin. (1999) DNA-Binding Proteins and Evolution of Transcription

Regulation in the Archaea. *Nucleic Acids Res*27: 4,658—4,670.

Aravind, L., and E. V. Koonin. (2001) The DNA-Repair Protein Alkb, Egl-9, and Lepreca Define New Families of 2-Oxoglutarate – and Iron-Dependent Dioxygenases. *Genome Biol*2: RESEARCH0007.

Aravind, L., R. Mazumder, S. Vasudevan, and E. V. Koonin. (2002) Trends in Protein Evolution Inferred from Sequence and Structure Analysis. *Curr Opin Struct Biol*12: 392–399.

Aravind, L., R. L. Tatusov, Y. I. Wolf, D. R. Walker, and E. V. Koonin. (1998) Evidence for Massive Gene Exchange Between Archaeal and Bacterial Hyperthermophiles. *Trends Genet*14: 442–444.

Aravind, L., D. R. Walker, and E. V. Koonin. (1999) Conserved Domains in DNA Repair Proteins and Evolution of Repair Systems. *Nucleic Acids Res*27: 1,223—1,242.

Assis, R., and A. S. Kondrashov. (2009) Rapid Repetitive Element-Mediated Expansion of piRNA Clusters in Mammalian Evolution. *Proc Natl Acad Sci USA*106: 7,079—7,082.

Assis, R., A. S. Kondrashov, E. V. Koonin, and F. A. Kondrashov. (2008) Nested Genes and Increasing Organizational Complexity of Metazoan Genomes. *Trends Genet*24: 475–478.

Atkins, J. F., R. F. Gesteland, and T. R. Cech, eds. (2010) *RNA Worlds: From Life's Origins to Diversity in Gene Regulation*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Auster, P. (1991) *The Music of Chance*. New York: Penguin.

Autexier, C., and N. F. Lue. (2006) The Structure and Function of Telomerase Reverse Transcriptase. *Annu Rev Biochem*75: 493–517.

Baaske, P., F. M. Weinert, S. Duhr, K. H. Lemke, M. J. Russell, and D. Braun. (2007) Extreme Accumulation of Nucleotides in Simulated Hydrothermal Pore Systems. *Proc Natl Acad Sci USA*104: 9,346—9,351.

Babu, M., N. Beloglazova, R. Flick, C. Graham, T. Skarina, B. Nocek, A. Gagarinova, O. Pogoutse, G. Brown, A. Binkowski, S. Phanse, A. Joachimiak, E. V. Koonin, A. Savchenko, A. Emili, J. Greenblatt, A. M. Edwards, and A. F. Yakunin. (2011) A Dual Function of the CRISPR-Cas System in Bacterial Antivirus Immunity and DNA Repair. *Mol Microbiol*79: 484–502.

Ball, P. (2011) A Metaphor Too Far. *Nature* doi:10.1038/news.2011.115.

Bailey, K. A., F. Marc, K. Sandman, and J. N. Reeve. (2002) Both DNA and Histone Fold Sequences Contribute to Archaeal Nucleosome Stability. *J Biol Chem*277: 9293–9301.

Baltimore, D. (1971) Expression of Animal Virus Genomes. *Bacteriol Rev*35: 235–241.

Bangham, C. R., and T. B. Kirkwood. (1993) Defective Interfering Particles and Virus Evolution. *Trends Microbiol*1: 260–264.

Barabasi, A. L. (2002) *Linked: The New Science of Networks*. New York: Perseus Press.

Barabasi, A. L., and Z. N. Oltvai. (2004) Network Biology: Understanding the Cell's Functional Organization. *Nat Rev Genet*5: 101–113.

Barbrook, A. C., C. J. Howe, D. P. Kurniawan, and S. J. Tarr. (2010) Organization and Expression of Organellar Genomes. *Philos Trans R Soc Lond B Biol Sci*365: 785–797.

Baross, J. A., and S. A. Hoffman. (1985) Submarine Hydrothermal Vents and Associated Gradient Environments As Sites for the Origin and Evolution of Life. *Origins of Life*15: 327–345.

Barrangou, R., C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero, and P. Horvath. (2007) CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*315: 1,709—1,712.

Barrick, J. E., D. S. Yu, S. H. Yoon, H. Jeong, T. K. Oh, D. Schneider, R. E. Lenski, and J. F. Kim. (2009) Genome Evolution and Adaptation in a Long-Term Experiment with *Escherichia coli*. *Nature*461: 1243–1247.

Barrow, J. D., and F. J. Tipler. (1988) *The Anthropic Cosmological Principle*. Oxford: Oxford University Press.

- Barrowman, J., D. Bhandari, K. Reinisch, and S. Ferro-Novick. (2010) Trapp Complexes in Membrane Traffic: Convergence Through a Common Rab. *Nat Rev Mol Cell Biol*11: 759–763.
- Barton, N. H. (2000) Genetic Hitchhiking. *Philos Trans R Soc Lond B Biol Sci*355: 1,553—1,562.
- Barton, N. H., and J. B. Coe. (2009) On the Application of Statistical Physics to Evolutionary Biology. *J Theor Biol*259: 317–324.
- Basu, M. K., E. Poliakov, and I. B. Rogozin. (2009) Domain Mobility in Proteins: Functional and Evolutionary Implications. *Brief Bioinform*10: 205–216.
- Beardmore, R. E., I. Gudelj, D. A. Lipson, and L. D. Hurst. (2011) Metabolic Trade-Offs and the Maintenance of the Fittest and the Flattest. *Nature*[Epub ahead of print].
- Beeby, M., B. D. O'Connor, C. Ryttersgaard, D. R. Boutz, L. J. Perry, and T. O. Yeates. (2005) The Genomics of Disulfide Bonding and Protein Stabilization in Thermophiles. *PLoS Biol*3: e309.
- Begley, T. J., and L. D. Samson. (2003) Molecular Biology: A Fix for RNA. *Nature*421: 795–796.
- Behe, M. J. (2006). *Darwin's Black Box: The Biochemical Challenge to Evolution*. New York: Free Press.
- Behm-Ansmant, I., I. Kashima, J. Rehwinkel, J. Sauliere, N. Wittkopp, and E. Izaurralde. (2007) mRNA Quality Control: An Ancient Machinery Recognizes and Degrades mRNAs with Nonsense Codons. *FEBS Lett*581: 2,845—2,853.
- Beiko, R. G., and N. Hamilton. (2006) Phylogenetic Identification of Lateral Genetic Transfer Events. *BMC Evol Biol*6: 15.
- Beiko, R. G., T. J. Harlow, and M. A. Ragan. (2005) Highways of Gene Sharing in Prokaryotes. *Proc Natl Acad Sci USA*102: 14,332—14,337.
- Beringer, M., and M. V. Rodnina. (2007) The Ribosomal Peptidyl Transferase. *Mol Cell*26: 311–321.
- Bezier, A., M. Annaheim, J. Herbinere, C. Wetterwald, G. Gyapay, S. Bernard-Samain, P. Wincker, I. Roditi, M. Heller, M. Belghazi, R. Pfister-Wilhem, G. Periquet, C. Dupuy, E. Huguet, A. N. Volkoff, B. Lanzrein, and J. M. Drezen. (2009) Polydnviruses of Braconid Wasps Derive from an Ancestral Nudivirus. *Science*323: 926–930.
- Bhattacharya, D., J. M. Archibald, A. P. Weber, and A. Reyes-Prieto. (2007) How Do Endosymbionts Become Organelles? Understanding Early Events in Plastid Evolution. *Bioessays*29: 1,239—1,246.
- Bidle, K. D., and P. G. Falkowski. (2004) Cell Death in Planktonic, Photosynthetic Microorganisms. *Nat Rev Microbiol*2: 643–655.
- Biebricher, C. K., and M. Eigen. (2005) The Error Threshold. *Virus Res*107: 117–127.
- Bigot, Y., S. Samain, C. Auge-Gouillou, and B. A. Federici. (2008) Molecular Evidence for the Evolution of Ichnoviruses from Ascoviruses by Symbiogenesis. *BMC Evol Biol*8: 253.
- Bjedov, I., O. Tenaillon, B. Gerard, V. Souza, E. Denamur, M. Radman, F. Taddei, and I. Matic. (2003) Stress-Induced Mutagenesis in Bacteria. *Science*300: 1,404—1,409.
- Blanchard, S. C., R. L. Gonzalez, H. D. Kim, S. Chu, and J. D. Puglisi. (2004) tRNA Selection and Kinetic Proofreading in Translation. *Nat Struct Mol Biol*11: 1,008—1,014.
- Blasius, M., S. Sommer, and U. Hubscher. (2008) *Deinococcus Radiodurans*: What Belongs to the Survival Kit? *Crit Rev Biochem Mol Biol*43: 221–238.
- Blencowe, B. J. (2006) Alternative Splicing: New Insights from Global Analyses. *Cell*126: 37–47.
- Bokov, K., and S. V. Steinberg. (2009) A Hierarchical Model for Evolution of 23s Ribosomal RNA. *Nature*457: 977–980.
- Bolotin, A., B. Quinquis, A. Sorokin, and S. D. Ehrlich. (2005) Clustered Regularly Interspaced Short Palindrome Repeats (Crisprs) Have Spacers of Extrachromosomal Origin. *Microbiology*151: 2,551—2,561.

- Bonner, J. T.(2004) Perspective: The Size-Complexity Rule. *Evolution*58: 1,883—1,890.
- Bostrom, Nick.(2002) *Anthropic Bias: Observation Selection Effects in Science and Philosophy*. New York and London: Routledge.
- Bourc'his, D., and O. Voinnet.(2010) A Small-RNA Perspective on Gametogenesis, Fertilization, and Early Zygotic Development. *Science*330: 617–622.
- Bousso, R.(2006) Holographic Probabilities in Eternal Inflation. *Phys Rev Lett*97: 19,1302.
- Bousso, R., and J. Polchinski.(2004) The String Theory Landscape. *Sci Am*291: 78–87.
- Bowman, G. R., V. A. Voelz, and V. S. Pande.(2011) Taming the Complexity of Protein Folding. *Curr Opin Struct Biol*21: 4—11.
- Branco, M. R., and A. Pombo.(2007) Chromosome Organization: New Facts, New Models. *Trends Cell Biol*17: 127–134.
- Brinkmann, H., and H. Philippe.(2007) The Diversity of Eukaryotes and the Root of the Eukaryotic Tree. *Adv Exp Med Biol*607: 20–37.
- Brisson, D.(2003) The Directed Mutation Controversy in an Evolutionary Context. *Crit Rev Microbiol*29: 25–35.
- Brochier-Armanet, C., B. Boussau, S. Gribaldo, and P. Forterre.(2008) Mesophilic Crenarchaeota: Proposal for a Third Archaeal Phylum, the Thaumarchaeota. *Nat Rev Microbiol*6: 245–252.
- Bromham, L., and D. Penny.(2003) The Modern Molecular Clock. *Nat Rev Genet*4: 216–224.
- Brookfield, J. F.(2009) Evolution and Evolvability: Celebrating Darwin 200. *Biol Lett*5: 44–46.
- Brouns, S. J., M. M. Jore, M. Lundgren, E. R. Westra, R. J. Slijkuis, A. P. Snijders, M. J. Dickman, K. S. Makarova, E. V. Koonin, and J. van der Oost.(2008) Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science*321: 960–964.
- Bruggeman, F. J., and H. V. Westerhoff.(2007) The Nature of Systems Biology. *Trends Microbiol*15: 45–50.
- Burger, R., M. Willensdorfer, and M. A. Nowak.(2006) Why Are Phenotypic Mutation Rates Much Higher Than Genotypic Mutation Rates? *Genetics*172: 197–206.
- Burki, F., K. Shalchian-Tabrizi, and J. Pawlowski.(2008) Phylogenomics Reveals a New «Megagroup» Including Most Photosynthetic Eukaryotes. *Biol Lett*4: 366–369.
- Bushman, F.(2001) *Lateral DNA Transfer: Mechanisms and Consequences*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Cairns, J., J. Overbaugh, and S. Miller.(1988) The Origin of Mutants. *Nature*335: 142–145.
- Cairns, J., G. S. Stent, and J. D. Watson, eds.(1966) *Phage and the Origins of Molecular Biology*, Cold Spring Harbor, NY: CSHL Press.
- Carmel, L., Y. I. Wolf, I. B. Rogozin, and E. V. Koonin.(2007) Three Distinct Modes of Intron Dynamics in the Evolution of Eukaryotes. *Genome Res*17: 1,034—1,044.
- Carroll, L.(1872). *Through the Looking Glass and What Alice Found There*. London: Macmillan.
- Carroll, S.(2010) *From Eternity to Here: The Quest for the Ultimate Theory of Time*. New York: Penguin.
- Carter, B.(1974) Large Number Coincidences and the Anthropic Principle in Cosmology in *Confrontation of Cosmological Theories with Observational Data*. M. S. Longair, editor. Dordrecht: D. Reidel, pp. 291–298.
- Casino, P., V. Rubio, and A. Marina.(2010) The Mechanism of Signal Transduction by Two-Component Systems. *Curr Opin Struct Biol*20: 763–771.
- Cavalier-Smith, T.(1998) A Revised Six-Kingdom System of Life. *Biol Rev Camb Philos Soc*73: 203–266.
- Cech, T. R.(2002) Ribozymes, the First 20 Years. *Biochem Soc Trans*30: 1,162—1,166.
- Chan, C. X., A. E. Darling, R. G. Beiko, and M. A. Ragan.(2009) Are Protein Domains Modules of

Lateral Genetic Transfer? *PLoS One*4: e4524.

Chargaff, E.(1978) *Heraclitean Fire: Sketches from a Life Before Nature*. New York: Rockefeller University Press.

Charlebois, R. L.,and W. F. Doolittle.(2004) Computing Prokaryotic Gene Ubiquity: Rescuing the Core from Extinction. *Genome Res*14: 2,469—2,477.

Chen, I., P. J. Christie, and D. Dubnau.(2005) The Ins and Outs of DNA Transfer in Bacteria. *Science*310: 1,456—1,460.

Cheng, L. K.,and P. J. Unrau.(2010) Closing the Circle: Replicating RNA with RNA. *Cold Spring Harb Perspect Biol*2: a002204.

Chernikova, D., S. Motamedi, M. Csuros, E. V. Koonin, and I. B. Rogozin.(2011) A Late Origin of the Extant Eukaryotic Diversity: Divergence Time Estimates Using Rare Genomic Changes. *Biol Direct*,in press.

Ciccarelli, F. D., T. Doerks, C. von Mering, C. J. Creevey, B. Snel, and P. Bork. (2006) Toward Automatic Reconstruction of a Highly Resolved Tree of Life. *Science*311: 1,283—1,287.

Clapier, C. R.,and B. R. Cairns.(2009) The Biology of Chromatin Remodeling Complexes. *Annu Rev Biochem*78: 273—304.

Claverie, J. M.(2006) Viruses Take Center Stage in Cellular Evolution. *Genome Biol*7: 110.

Collins, L.,and D. Penny.(2005) Complex Spliceosomal Organization Ancestral to Extant Eukaryotes. *Mol Biol Evol*22: 1,053—1,066.

Cortez, D., P. Forterre, and S. Gribaldo.(2009) A Hidden Reservoir of Integrative Elements Is the Major Source of Recently Acquired Foreign Genes and Orfans in Archaeal and Bacterial Genomes. *Genome Biol*10: R65.

Costanzo, G., S. Pino, F. Ciciriello, and E. Di Mauro.(2009) Generation of Long RNA Chains in Water. *J Biol Chem*284: 33,206—33,216.

Costanzo, G., R. Saladino, C. Crestini, F. Ciciriello, and E. Di Mauro.(2007) Nucleoside Phosphorylation by Phosphate Minerals. *J Biol Chem*282: 16,729—16,735.

Cox, C. J., P. G. Foster, R. P. Hirt, S. R. Harris, and T. M. Embley.(2008) The Archaeobacterial Origin of Eukaryotes. *Proc Natl Acad Sci USA*105: 20,356—20,361.

Cox, M. M.,and J. R. Battista.(2005) *Deinococcus Radiodurans – the Consummate Survivor*. *Nat Rev Microbiol*3: 882–892.

Crick, F.(1970) Central Dogma of Molecular Biology. *Nature*227: 561–563.

Crick, F. H.(1958) On Protein Synthesis. *Symp Soc Exp Biol*12: 138–163.

Crick, F. H.(1968) The Origin of the Genetic Code. *J Mol Biol*38: 367–379.

Csuros, M.,and I. Miklos.(2009) Streamlining and Large Ancestral Genomes in Archaea Inferred with a Phylogenetic Birth-and-Death Model. *Mol Biol Evol*26: 2,087—2,095.

Csuros, M., I. B. Rogozin, and E. V. Koonin. (2011) Reconstructed HumanLike Intron Density in the Last Common Ancestor of Eukaryotes. *PLoS Comput Biol*,in press.

D’Herelle, F.(1922). *The Bacteriophage; Its Role in Immunity*. Baltimore: Williams and Wilkins.

Dagan, T., and W. Martin. (2006) The Tree of One Percent. *Genome Biol*7: 118.

Daly, M. J.(2009) A New Perspective on Radiation Resistance Based on *Deinococcus Radiodurans*. *Nat Rev Microbiol*7: 237–245.

Danchin, A.(2003). *The Delphic Boat: What Genomes Tell Us*. Cambridge, MA: Harvard Univ Press.

Darwin, C.(1859) *On the Origin of Species*. London: Murray. Darwin, C. (1872) *Origin of Species*. New York: The Modern Library.

Daubin, V., and H. Ochman.(2004) Bacterial Genomes As New Gene Homes: The Genealogy of Orfans in *E. Coli*. *Genome Res*14: 1,036—1,042.

- Davidov, Y., and E. Jurkevitch.(2009) Predation Between Prokaryotes and the Origin of Eukaryotes. *Bioessays*31: 748–757.
- Daviter, T., K. B. Gromadski, and M. V. Rodnina. (2006) The Ribosome's Response to Codon-Anticodon Mismatches. *Biochimie*88: 1,001—1,011.
- Dawkins, R.(1996) *The Blind Watchmaker: Why the Evidence of Evolution Reveals a Universe Without Design*. London: W.W. Norton & Co.
- Dawkins, R.(2006) *The Selfish Gene: 30th Anniversary Edition – with a New Introduction by the Author*. Oxford: Oxford University Press.
- Dayhoff, M. O., W. C. Barker, and L. T. Hunt.(1983) Establishing Homologies in Protein Sequences. *Methods Enzymol*91: 524–545.
- de Nooijer, S., B. R. Holland, and D. Penny.(2009) The Emergence of Predators in Early Life: There Was No Garden of Eden. *PLoS ONE*4: e5507.
- de Souza, S. J., M. Long, R. J. Klein, S. Roy, S. Lin, and W. Gilbert.(1998) Toward a Resolution of the Introns Early/Late Debate: Only Phase Zero Introns Are Correlated with the Structure of Ancient Proteins. *Proc Natl Acad Sci USA*95: 5,094—5,099.
- de Visser, J. A., and S. F. Elena. (2007) The Evolution of Sex: Empirical Insights into the Roles of Epistasis and Drift. *Nat Rev Genet*8: 139–149.
- Deckert, G., P. V. Warren, T. Gaasterland, W. G. Young, A. L. Lenox, D. E. Graham, R. Overbeek, M. A. Snead, M. Keller, M. Aujay, R. Huber, R. A. Feldman, J. M. Short, G. J. Olsen, and R. V. Swanson.(1998) The Complete Genome of the Hyperthermophilic Bacterium *Aquifex Aeolicus*. *Nature*392: 353–358.
- Delarue, M., O. Poch, N. Tordo, D. Moras, and P. Argos.(1990) An Attempt to Unify the Structure of Polymerases. *Protein Eng*3: 461–467.
- Delaye, L., and A. Moya.(2010) Evolution of Reduced Prokaryotic Genomes and the Minimal Cell Concept: Variations on a Theme. *Bioessays*32: 281–287.
- Deleuze, G., and F. Guattari. (1987) *Thousand Plateaus: Capitalism and Schizophrenia*. Minneapolis, MN: University of Minnesota Press.
- Denamur E. and I. Matic.(2006) Evolution of Mutation Rates in Bacteria. *Mol Microbiol*60: 820–827.
- Dennett, D. C.(1996) *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon & Schuster.
- Deppenmeier, U., A. Johann, T. Hartsch, R. Merkl, R. A. Schmitz, R. Martinez-Arias, A. Henne, A. Wiezer, S. Baumer, C. Jacobi, H. Bruggemann, T. Lienard, A. Christmann, M. Bomeke, S. Steckel, A. Bhattacharyya, A. Lykidis, R. Overbeek, H. P. Klenk, R. P. Gunsalus, H. J. Fritz, and G. Gottschalk.(2002) The Genome of *Methanosarcina Mazei*: Evidence for Lateral Gene Transfer Between Bacteria and Archaea. *J Mol Microbiol Biotechnol*4: 453–461.
- Deveau, H., J. E. Garneau, and S. Moineau.(2010) CRISPR/Cas System and Its Role in Phage-Bacteria Interactions. *Annu Rev Microbiol*64: 475–493.
- Diaz, R., C. Vargas-Lagunas, M. A. Villalobos, H. Peralta, Y. Mora, S. Encarnacion, L. Girard, and J. Mora.(2011) argC Orthologs from Rhizobiales Show Diverse Profiles of Transcriptional Efficiency and Functionality in *Sinorhizobium Meliloti*. *J Bacteriol*193: 460–472.
- Ding, S. W.(2010) RNA-Based Antiviral Immunity. *Nat Rev Immunol*10: 632–644.
- Dlakic, M. and A. Mushegian.(2011) Prp8, the Pivotal Protein of the Spliceosomal Catalytic Center, Evolved from a Retroelement-Encoded Reverse Transcriptase. *RNA*[Epub ahead of print].
- Dobzhansky, T.(1951). *Genetics and the Origin of Species*. New York: Columbia University Press.
- Dobzhansky, T.(1973) *Nothing in Biology Makes Sense Except in the Light of Evolution*. *The*

Dong, H., and C. G. Kurland.(1995) Ribosome Mutants with Altered Accuracy Translate with Reduced Processivity. *J Mol Biol*248: 551–561.

Doolittle, R. F.(1995) The Multiplicity of Domains in Proteins. *Annu Rev Biochem*64: 287–314.

Doolittle, W. F.(1999a) Lateral Genomics. *Trends Cell Biol*9: M5—8.

Doolittle, W. F. (1999b) Phylogenetic Classification and the Universal Tree. *Science*284: 2,124—2,129.

Doolittle, W. F.(2000) Uprooting the Tree of Life. *Sci Am*282: 90–95.

Doolittle, W. F., and E. Bapteste.(2007) Pattern Pluralism and the Tree of Life Hypothesis. *Proc Natl Acad Sci USA*104: 2,043—2,049.

Doolittle, W. F., and J. R. Brown.(1994) Tempo, Mode, the Progenote, and the Universal Root. *Proc Natl Acad Sci USA*91: 6,721—6,728.

Doolittle, W. F., and C. Sapienza.(1980) Selfish Genes, the Phenotype Paradigm and Genome Evolution. *Nature*284: 601–603.

Doolittle, W. F., and O. Zhaxybayeva.(2009) On the Origin of Prokaryotic Species. *Genome Res*19: 744–756.

Doudna, J. A., and T. R. Cech.(2002) The Chemical Repertoire of Natural Ribozymes. *Nature*418: 222–228.

Douzery, E. J., E. A. Snell, E. Bapteste, F. Delsuc, and H. Philippe.(2004) The Timing of Eukaryotic Evolution: Does a Relaxed Molecular Clock Reconcile Proteins and Fossils? *Proc Natl Acad Sci USA*101: 15,386—15,391.

Draghi, J. A., T. L. Parsons, G. P. Wagner, and J. B. Plotkin.(2010) Mutational Robustness Can Facilitate Adaptation. *Nature*463: 353–355.

Drake, J. W.(1991) A Constant Rate of Spontaneous Mutation in DNA-Based Microbes. *Proc Natl Acad Sci USA*88: 7,160—7,164.

Drake, J. W., and J. J. Holland.(1999) Mutation Rates Among RNA Viruses. *Proc Natl Acad Sci USA*96: 13,910—13,913.

Dronamraju, K. R., ed.(1968) *Haldane and Modern Biology*. Baltimore, Johns Hopkins University Press.

Drummond, D. A., J. D. Bloom, C. Adami, C. O. Wilke, and F. H. Arnold. (2005) Why Highly Expressed Proteins Evolve Slowly. *Proc Natl Acad Sci USA*102: 14,338—14,343.

Drummond, D. A., A. Raval, and C. O. Wilke.(2006) A Single Determinant Dominates the Rate of Yeast Protein Evolution. *Mol Biol Evol*23: 327–337.

Drummond, D. A., and C. O. Wilke.(2008) Mistranslation-Induced Protein Misfolding As a Dominant Constraint on Coding-Sequence Evolution. *Cell*134: 341–352.

Drummond, D. A., and C. O. Wilke.(2009) The Evolutionary Consequences of Erroneous Protein Synthesis. *Nat Rev Genet*10: 715–724.

Dupuy, C., E. Huguet, and J. M. Drezen.(2006) Unfolding the Evolutionary Story of Polydnviruses. *Virus Res*117: 81–89.

Echols, H.(1981) SOS Functions, Cancer, and Inducible Evolution. *Cell*25: 1–2.

Eckerle, L. D., X. Lu, S. M. Sperry, L. Choi, and M. R. Denison.(2007) High Fidelity of Murine Hepatitis Virus Replication Is Decreased in nsp14 Exoribonuclease Mutants. *J Virol*81: 12,135—12,144.

Eddy, S. R.(2002) Computational Genomics of Noncoding RNA Genes. *Cell*109: 137–140.

Edwards, R. A., and F. Rohwer.(2005) Viral Metagenomics. *Nat Rev Microbiol*3: 504–510.

Eigen, M.(1971) Self-organization of Matter and the Evolution of Biological Macromolecules. *Naturwissenschaften*58: 465–523.

- Eigen, M., B. F. Lindemann, M. Tietze, R. Winkler-Oswatitsch, A. Dress, and A. von Haeseler.(1989) How Old Is the Genetic Code? Statistical Geometry of tRNA Provides an Answer. *Science*244: 673–679.
- Eisen, J. A., J. F. Heidelberg, O. White, and S. L. Salzberg.(2000) Evidence for Symmetric Chromosomal Inversions Around the Replication Origin in Bacteria. *Genome Biol*1: RESEARCH0011.
- Eldredge, N., and S. J. Gould.(1997) On Punctuated Equilibria. *Science*276: 338–341.
- Ellington, A. D., X. Chen, M. Robertson, and A. Syrett.(2009) Evolutionary Origins and Directed Evolution of RNA. *Int J Biochem Cell Biol*41: 254–265.
- Ellis, R. J.(2003) Protein Folding: Importance of the Anfinsen Cage. *Curr Biol*13: R881—883.
- Embley, T. M., and W. Martin.(2006) Eukaryotic Evolution, Changes and Challenges. *Nature*440: 623–630.
- Esser, C., N. Ahmadinejad, C. Wiegand, C. Rotte, F. Sebastiani, G. Gelius-Dietrich, K. Henze, E. Kretschmann, E. Richly, D. Leister, D. Bryant, M. A. Steel, P. J. Lockhart, D. Penny, and W. Martin.(2004) A Genome Phylogeny for Mitochondria Among Alpha-Proteobacteria and a Predominantly Eubacterial Ancestry of Yeast Nuclear Genes. *Mol Biol Evol*21: 1,643—1,660.
- Esser, C., W. Martin, and T. Dagan. (2007) The Origin of Mitochondria in Light of a Fluid Prokaryotic Chromosome Model. *Biol Lett*3: 180–184.
- Falkowski, P. G., T. Fenchel, and E. F. Delong. (2008) The Microbial Engines That Drive Earth's Biogeochemical Cycles. *Science*320: 1,034—1,039.
- Falnes, P. O.(2005) RNA Repair – the Latest Addition to the Toolbox for Macro-molecular Maintenance. *RNA Biol*2: 14–16.
- Fedor, M. J., and J. R. Williamson.(2005) The Catalytic Diversity of RNAs. *Nat Rev Mol Cell Biol*6: 399–412.
- Felsenstein, J.(1996) Inferring Phylogenies from Protein Sequences by Parsimony, Distance, and Likelihood Methods. *Methods Enzymol*266: 418–427.
- Felsenstein, J.(2004). *Inferring Phylogenies*. Sunderland, MA: Sinauer Associates.
- Feschotte, C.(2010) Virology: Bornavirus Enters the Genome. *Nature*463: 39–40.
- Field, M. C., and J. B. Dacks.(2009) First and Last Ancestors: Reconstructing Evolution of the Endomembrane System with Escrts, Vesicle Coat Proteins, and Nuclear Pore Complexes. *Curr Opin Cell Biol*21: 4—13.
- Fields, B. N., P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, S. E. Straus, and D. M. Knipe.(2001). *Fields Virology*. New York: Lippincott Williams & Wilkins.
- Fisher Box, J.(1978) R. A. Fisher: The Life of a Scientist. New York: Wiley.
- Fisher, R. A.(1930) *The Genetical Theory of Natural Selection*. New York: Oxford University Press.
- Fitch, W. M.(1970) Distinguishing Homologous from Analogous Proteins. *Syst Zool*19: 99—113.
- Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J. F. Tomb, B. A. Dougherty, J. M. Merrick et al. (1995) Whole-Genome Random Sequencing and Assembly of *Haemophilus Influenzae* Rd [See Comments]. *Science*269: 496–512.
- Flores, R., S. Delgado, M. E. Gas, A. Carbonell, D. Molina, S. Gago, and M. De la Pena.(2004) Viroids: The Minimal Non-Coding RNAs with Autonomous Replication. *FEBS Lett*567: 42–48.
- Fontanari, J. F., M. Santos, and E. Szathmary.(2006) Coexistence and Error Propagation in Pre-Biotic Vesicle Models: A Group Selection Approach. *J Theor Biol*239: 247–256.
- Forterre, P.(1999) Displacement of Cellular Proteins by Functional Analogues from Plasmids or Viruses Could Explain Puzzling Phylogenies of Many DNA Informational Proteins. *Mol Microbiol*33: 457–465.
- Forterre, P.(2002) A Hot Story from Comparative Genomics: Reverse Gyrase Is the Only

Hyperthermophile-Specific Protein. *Trends Genet*18: 236–237.

Forterre, P. (2006) Three RNA Cells for Ribosomal Lineages and Three DNA Viruses to Replicate Their Genomes: A Hypothesis for the Origin of Cellular Domain. *Proc Natl Acad Sci USA*103: 3,669—3,674.

Forterre, P., and D. Prangishvili.(2009) The Great Billion-Year War Between Ribosome – and Capsid-Encoding Organisms (Cells and Viruses) As the Major Source of Evolutionary Novelty. *Ann N Y Acad Sci*1,178: 65–77.

Foster, P. L.(2000) Adaptive Mutation: Implications for Evolution. *Bioessays*22: 1,067—1,074.

Foster, P. L. (2007) Stress-Induced Mutagenesis in Bacteria. *Crit Rev Biochem Mol Biol*42: 373–397.

Fowles, J.(1969). *The French Lieutenant's Woman*. Boston: Little & Brown.

Fox, S. W.(1976) The Evolutionary Significance of Phase-Separated Microsystems. *Orig Life*7: 49–68.

Frank, A. C., H. Amiri, and S. G. Andersson.(2002) Genome Deterioration: Loss of Repeated Sequences and Accumulation of Junk DNA. *Genetica*115: 1—12.

Fraser, C. M., J. D. Gocayne, O. White, M. D. Adams, R. A. Clayton, R. D. Fleischmann, C. J. Bult, A. R. Kerlavage, G. Sutton, J. M. Kelley, et al.(1995) The Minimal Gene Complement of *Mycoplasma Genitalium*. *Science*270: 397–403.

Fredrick, K., and H. F. Noller.(2002) Accurate Translocation of mRNA by the Ribosome Requires a Peptidyl Group or Its Analog on the tRNA Moving into the 30s P-Site. *Mol Cell*9: 1,125—1,131.

Freeland, S. J., and L. D. Hurst.(1998) The Genetic Code Is One in a Million. *J Mol Evol*47: 238–248.

French, S. L., T. J. Santangelo, A. L. Beyer, and J. N. Reeve.(2007) Transcription and translation are coupled in Archaea. *Mol Biol Evol*24: 893–895.

Friedberg, E. C., G. C. Walker, W. Siede, R. D. Wood, R. A. Schulz, and T. Ellenberger. (2005) *DNA Repair and Mutagenesis*. Washington, D.C.: ASM Press.

Fritz-Laylin, L. K., S. E. Prochnik, M. L. Ginger, J. B. Dacks, M. L. Carpenter, M. C. Field, A. Kuo, A. Paredez, J. Chapman, J. Pham, S. Shu, R. Neupane, M. Cipriano, J. Mancuso, H. Tu, A. Salamov, E. Lindquist, H. Shapiro, S. Lucas, I. V. Grigoriev, W. Z. Cande, C. Fulton, D. S. Rokhsar, and S. C. Dawson.(2010) The Genome of *Naegleria Gruberi* Illuminates Early Eukaryotic Versatility. *Cell*140: 631–642.

Frost, L. S., and G. Koraimann.(2010) Regulation of Bacterial Conjugation: Balancing Opportunity with Adversity. *Future Microbiol*5: 1,057—1,071.

Frost, L. S., R. Leplae, A. O. Summers, and A. Toussaint.(2005) Mobile Genetic Elements: The Agents of Open Source Evolution. *Nat Rev Microbiol*3: 722–732.

Fuerst, J. A.(2005) Intracellular Compartmentation in Planctomycetes. *Annu Rev Microbiol*59: 299–328.

Futuyma, D.(2005). *Evolution*. Sunderland, MA: Sinauer Associates.

Gago, S., S. F. Elena, R. Flores, and R. Sanjuan.(2009) Extremely High Mutation Rate of a Hammerhead Viroid. *Science*323: 1,308.

Galhardo, R. S., P. J. Hastings, and S. M. Rosenberg.(2007) Mutation As a Stress Response and the Regulation of Evolvability. *Crit Rev Biochem Mol Biol*42: 399–435.

Galperin, M. Y.(2005) A Census of Membrane-Bound and Intracellular Signal Transduction Proteins in Bacteria: Bacterial Iq, Extroverts and Introverts. *BMC Microbiol*5: 35.

Garriga, J., and A. Vilenkin.(2001) Many Worlds in One. *Phys. Rev. D*64: 043,511.

Garriss, G., M. K. Waldor, and V. Burrus.(2009) Mobile Antibiotic Resistance Encoding Elements Promote Their Own Diversity. *PLoS Genet*5: e1000775.

- Gavrilets, S.(2004) *Fitness Landscapes and the Origin of Species*.Princeton, NJ: Princeton University Press.
- Ge, F., L. S. Wang, and J. Kim.(2005) The Cobweb of Life Revealed by Genome-Scale Estimates of Horizontal Gene Transfer. *PLoS Biol*3: e316.
- Gell-Mann, M.(1995) *The Quark and the Jaguar: Adventures in the Simple and the Complex*New York: St. Martin's Griffin.
- Gibson, C. M.,and M. S. Hunter.(2010) Extraordinarily Widespread and Fantastically Complex: Comparative Biology of Endosymbiotic Bacterial and Fungal Mutualists of Insects. *Ecol Lett*13: 223–234.
- Gilbert, W.(1978) Why Genes in Pieces? *Nature*271: 501.
- Gillespie, J. H.(2000) The Neutral Theory in an Infinite Population. *Gene*261: 11–18.
- Giovannoni, S. J., H. J. Tripp, S. Givan, M. Podar, K. L. Vergin, D. Baptista, L. Bibbs, J. Eads, T. H. Richardson, M. Noordewier, M. S. Rappe, J. M. Short, J. C. Carrington, and E. J. Mathur.(2005) Genome Streamlining in a Cosmopolitan Oceanic Bacterium. *Science*309: 1,242—1,245.
- Glansdorff, N., Y. Xu, and B. Labedan.(2008) The Last Universal Common Ancestor: Emergence, Constitution and Genetic Legacy of an Elusive Forerunner. *Biol Direct*3: 29.
- Gliboff, S.(2005) Protoplasm... Is Soft Wax in Our Hands: Paul Kammerer and the Art of Biological Transformation. *Endeavour*29: 162–167.
- Gogarten, J. P., W. F. Doolittle, and J. G. Lawrence.(2002) Prokaryotic Evolution in Light of Gene Transfer. *Mol Biol Evol*19: 2,226—2,238.
- Gogarten, J. P., H. Kibak, P. Dittrich, L. Taiz, E. J. Bowman, B. J. Bowman, M. F. Manolson, R. J. Poole, T. Date, T. Oshima, et al.(1989) Evolution of the Vacuolar H⁺-Atpase: Implications for the Origin of Eukaryotes. *Proc Natl Acad Sci USA*86: 6,661—6,665.
- Gogarten, J. P., and J. P. Townsend. (2005) Horizontal Gene Transfer, Genome Innovation and Evolution. *Nat Rev Microbiol*3: 679–687.
- Gorbalenya, A. E., V. M. Blinov, A. P. Donchenko, and E. V. Koonin.(1985) An NTP-Binding Motif Is the Most Conserved Sequence in a Highly Diverged Monophyletic Group of Proteins Involved in Positive Strand RNA Viral Replication. *Molek Genetika, Microbiol Virusol*11: 30–36.
- Gorbalenya, A. E., L. Enjuanes, J. Ziebuhr, and E. J. Snijder.(2006) Nidovirales: Evolving the Largest RNA Virus Genome. *Virus Res*117: 17–37.
- Gorbalenya, A. E., and E. V. Koonin. (1989) Viral Proteins Containing the Purine NTP-Binding Sequence Pattern. *Nucleic Acids Res*17: 8,413—8,440.
- Gorbalenya, A. E., F. M. Pringle, J. L. Zeddam, B. T. Luke, C. E. Cameron, J. Kalkmakoff, T. N. Hanzlik, K. H. Gordon, and V. K. Ward. (2002) The Palm Subdomain-Based Active Site Is Internally Permuted in Viral RNA-Dependent Rna Polymerases of an Ancient Lineage. *J Mol Biol*324: 47–62.
- Gould, S. J.(1997) The Exaptive Excellence of Spandrels as a Term and Prototype. *Proc Natl Acad Sci USA*94: 10,750—10,755.
- Gould, S. J.(1997). *Full House: The Spread of Excellence from Plato to Darwin*. San Francisco: Three Rivers Press.
- Gould, S. J.(2002). *The Structure of Evolutionary Theory*. Cambridge, MA: Harvard University Press.
- Gould, S. J., and R. C. Lewontin.(1979) The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proc R Soc Lond B Biol Sci*205: 581–598.
- Graur, D., and W. Martin.(2004) Reading the Entrails of Chickens: Molecular Timescales of Evolution and the Illusion of Precision. *Trends Genet*20: 80–86.
- Gray, M. W., J. Lukes, J. M. Archibald, P. J. Keeling, and W. F. Doolittle. (2010) Cell Biology. Irremediable Complexity? *Science*330: 920–921.

- Grilli, J., B. Bassetti, S. Maslov, and M. C. Lagomarsino. (2011) Joint Scaling Laws in Functional and Evolutionary Categories in Prokaryotic Genomes. *arXiv:1101.5814v1 [q-bio.GN]*.
- Grishin, N. V., Y. I. Wolf, and E. V. Koonin. (2000) From Complete Genomes to Measures of Substitution Rate Variability Within and Between Proteins. *Genome Res*10: 991—1,000.
- Groll, M., M. Bochtler, H. Brandstetter, T. Clausen, and R. Huber. (2005) Molecular Machines for Protein Degradation. *ChemBiochem*6: 222–256.
- Guth, A. (1998) *The Inflationary Universe*. Boston: Basic Books.
- Guth, A. H. (1998) *The Inflationary Universe: The Quest for a New Theory of Cosmic Origins*. New York: Perseus Book Group.
- Guth, A. H. (2001) Eternal Inflation. *Ann N Y Acad Sci*950: 66–82.
- Guth, A. H., and D. I. Kaiser. (2005) Inflationary Cosmology: Exploring the Universe from the Smallest to the Largest Scales. *Science*307: 884–890.
- Haeckel, E. (1997). *The Wonders of Life: A Popular Study of Biological Philosophy*. London: General Books, LLC.
- Haldane, J. B. S. (1928) The Origin of Life. *Rationalist Annual*148: 3—10.
- Hale, C. R., P. Zhao, S. Olson, M. O. Duff, B. R. Graveley, L. Wells, R. M. Terns, and M. P. Terns. (2009) RNA-Guided RNA Cleavage by a CRISPR RNA-Cas Protein Complex. *Cell*139: 945–956.
- Halfmann, R., and S. Lindquist. (2010) Epigenetics in the Extreme: Prions and the Inheritance of Environmentally Acquired Traits. *Science*330: 629–632.
- Hampl, V., L. Hug, J. W. Leigh, J. B. Dacks, B. F. Lang, A. G. Simpson, and A. J. Roger. (2009) Phylogenomic Analyses Support the Monophyly of Excavata and Resolve Relationships Among Eukaryotic «Supergroups.» *Proc Natl Acad Sci USA*106: 3,859—3,864.
- Harrison, P. M., and M. Gerstein. (2002) Studying Genomes Through the Aeons: Protein Families, Pseudogenes, and Proteome Evolution. *J Mol Biol*318: 1,155—1,174.
- Hartl, D. L., and A. G. Clark. (2006) *Principles of Population Genetics*. Sunderland, MA: Sinauer Associates.
- Hartung, S., and K. P. Hopfner. (2009) Lessons from Structural and Biochemical Studies on the Archaeal Exosome. *Biochem Soc Trans*37: 83–87.
- Hawking, S. W. (1988) *A Brief History of Time: From the Big Bang to Black Holes*. London: Bantam.
- Hawking, S. W., and L. Mlodinow. (2010) *The Grand Design*. London: Bantam.
- Hazkani-Covo, E., R. M. Zeller, and W. Martin. (2010) Molecular Poltergeists: Mitochondrial DNA Copies (Numts) in Sequenced Nuclear Genomes. *PLoS Genet*6: e1000834.
- Hilario, E., and J. P. Gogarten. (1993) Horizontal Transfer of Atpase Genes – the Tree of Life Becomes a Net of Life. *Biosystems*31: 111–119.
- Hochstrasser, M. (2009) Origin and Function of Ubiquitin-Like Proteins. *Nature*458: 422–429.
- Hollick, J. B. (2010) Paramutation and Development. *Annu Rev Cell Dev Biol*26: 557–579.
- Holmes, E. C. (2009) *The Evolution and Emergence of RNA Viruses*. Oxford: Oxford University Press.
- Horie, M., T. Honda, Y. Suzuki, Y. Kobayashi, T. Daito, T. Oshida, K. Ikuta, P. Jern, T. Gojobori, J. M. Coffin, and K. Tomonaga. (2010) Endogenous Non-Retroviral RNA Virus Elements in Mammalian Genomes. *Nature*463: 84–87.
- Huber, C., W. Eisenreich, S. Hecht, and G. Wachtershauser. (2003) A Possible Primordial Peptide Cycle. *Science*301: 938–940.
- Hudder, A., L. Nathanson, and M. P. Deutscher. (2003) Organization of Mammalian Cytoplasm. *Mol Cell Biol*23: 9,318—93,26.
- Hulen, C., and J. Legault-Demare. (1975) In Vitro Synthesis of Large Peptide Molecules Using

- Glucosylated Single-Stranded Bacteriophage T4d DNA Template. *Nucleic Acids Res*2: 2,037—2,048.
- Humbard, M. A., H. V. Miranda, J. M. Lim, D. J. Krause, J. R. Pritz, G. Zhou, S. Chen, L. Wells, and J. A. Maupin-Furlow. (2010) Ubiquitin-Like Small Archaeal Modifier Proteins (Samps) in *Haloferax Volcanii*. *Nature*463: 54–60.
- Hurst, L. D. (2002) The Ka/Ks Ratio: Diagnosing the Form of Sequence Evolution. *Trends Genet*18: 486.
- Hurst, L. D., and N. G. Smith. (1999) Do Essential Genes Evolve Slowly? *Curr Biol*9: 747–750.
- Hussain, T., V. Kamarthapu, S. P. Kruparani, M. V. Deshmukh, and R. Sankaranarayanan. (2010) Mechanistic Insights into Cognate Substrate Discrimination During Proofreading in Translation. *Proc Natl Acad Sci USA*107: 22,117—22,121.
- Huxley, J. (2010) *Evolution: The Modern Synthesis: The Definitive Edition*. Cambridge: MIT Press.
- Huxley, T. H. (1860) Darwin on the Origin of Species. *Westminster Review*: 541–570.
- Irimia, M., D. Penny, and S. W. Roy. (2007) Coevolution of Genomic Intron Number and Splice Sites. *Trends Genet*23: 321–325.
- Irimia, M., S. W. Roy, D. E. Neafsey, J. F. Abril, J. Garcia-Fernandez, and E. V. Koonin. (2009) Complex Selection on 5' Splice Sites in Intron-Rich Organisms. *Genome Res*19: 2,021—2,027.
- Iwabe, N., K. Kuma, M. Hasegawa, S. Osawa, and T. Miyata. (1989) Evolutionary Relationship of Archaeobacteria, Eubacteria, and Eukaryotes Inferred from Phylogenetic Trees of Duplicated Genes. *Proc Natl Acad Sci USA*86: 9,355—9,359.
- Iyer, L. M., A. M. Burroughs, and L. Aravind. (2006) The Prokaryotic Antecedents of the Ubiquitin-Signaling System and the Early Evolution of Ubiquitin-Like Beta-Grasp Domains. *Genome Biol*7: R60.
- Iyer, L. M., E. V. Koonin, and L. Aravind. (2003) Evolutionary Connection Between the Catalytic Subunits of DNA-Dependent RNA Polymerases and Eukaryotic RNA-Dependent RNA Polymerases and the Origin of RNA Polymerases. *BMC Struct Biol*3: 1.
- Iyer, L. M., E. V. Koonin, and L. Aravind. (2004a) Evolution of Bacterial RNA Polymerase: Implications for Large-Scale Bacterial Phylogeny, Domain Accretion, and Horizontal Gene Transfer. *Gene*335: 73–88.
- Iyer, L. M., E. V. Koonin, D. D. Leipe, and L. Aravind. (2005) Origin and Evolution of the Archaeo-Eukaryotic Primase Superfamily and Related Palm-Domain Proteins: Structural Insights and New Members. *Nucleic Acids Res*33: 3,875—3,896.
- Iyer, L. M., K. S. Makarova, E. V. Koonin, and L. Aravind. (2004b) Comparative Genomics of the Ftsk-Hera Superfamily of Pumping Atpases: Implications for the Origins of Chromosome Segregation, Cell Division and Viral Capsid Packaging. *Nucleic Acids Res*32: 5,260—5,279.
- Jacob, F. (1977) Evolution and Tinkering. *Science*196: 1,161—1,166.
- Jacob, F. (1993). *The Logic of Life*. Princeton: Princeton University Press.
- Jacob, F., and J. Monod. (1961) Genetic Regulatory Mechanisms in the Synthesis of Proteins. *J Mol Biol*3: 318–356.
- Jain, R., M. C. Rivera, and J. A. Lake. (1999) Horizontal Gene Transfer Among Genomes: The Complexity Hypothesis. *Proc Natl Acad Sci USA*96: 3,801—3,806.
- Jeffares, D. C., T. Mourier, and D. Penny. (2006) The Biology of Intron Gain and Loss. *Trends Genet*22: 16–22.
- Jensen, L. J., P. Julien, M. Kuhn, C. von Mering, J. Muller, T. Doerks, and P. Bork. (2008) EggNog: Automated Construction and Annotation of Orthologous Groups of Genes. *Nucleic Acids Res*36: D250—254.
- Johansson, M., M. Lovmar, and M. Ehrenberg. (2008) Rate and Accuracy of Bacterial Protein

Synthesis Revisited. *Curr Opin Microbiol*11: 141–147.

Johnson, L. J., and P. J. Tricker. (2010) Epigenomic Plasticity Within Populations: Its Evolutionary Significance and Potential. *Heredity*105: 113–121.

Johnson, N.(2009) Simply Complexity: A Clear Guide to Complexity Theory. New York: Oneworld Publications.

Johnston, W. K., P. J. Unrau, M. S. Lawrence, M. E. Glasner, and D. P. Bartel. (2001) RNA-Catalyzed Rna Polymerization: Accurate and General RNA-Templated Primer Extension. *Science*292: 1,319—1,325.

Jordan, I. K., K. S. Makarova, J. L. Spouge, Y. I. Wolf, and E. V. Koonin. (2001) Lineage-Specific Gene Expansions in Bacterial and Archaeal Genomes. *Genome Res*11: 555–565.

Jordan, I. K., L. Marino-Ramirez, Y. I. Wolf, and E. V. Koonin. (2004) Conservation and Coevolution in the Scale-Free Human Gene Coexpression Network. *Mol Biol Evol*21: 2,058—2,070.

Jordan, I. K., I. B. Rogozin, Y. I. Wolf, and E. V. Koonin. (2002) Essential Genes Are More Evolutionarily Conserved Than Are Nonessential Genes in Bacteria. *Genome Res*12: 962–968.

Juhas, M., J. R. van der Meer, M. Gaillard, R. M. Harding, D. W. Hood, and D. W. Crook. (2009) Genomic Islands: Tools of Bacterial Horizontal Gene Transfer and Evolution. *FEMS Microbiol Rev*33: 376–393.

Kaneko, K.(2007) Evolution of Robustness to Noise and Mutation in Gene Expression Dynamics. *PLoS One*2: e434.

Karev, G. P., Y. I. Wolf, A. Y. Rzhetsky, F. S. Berezovskaya, and E. V. Koonin. (2002) Birth and Death of Protein Domains: A Simple Model of Evolution Explains Power Law Behavior. *BMC Evol Biol*2: 18.

Karginov, F. V., and G. J. Hannon.(2010) The CRISPR System: Small RNA-Guided Defense in Bacteria and Archaea. *Mol Cell*37: 7—19.

Kasting, J. F., and S. Ono.(2006) Palaeoclimates: The First Two Billion Years. *Philos Trans R Soc Lond B Biol Sci*361: 917–929.

Kauffman, S.(1996) At Home in the Universe: The Search for the Laws of Self-Organization and Complexity. Oxford: Oxford University Press.

Keating, K. S., N. Toor, P. S. Perlman, and A. M. Pyle.(2010) A Structural Analysis of the Group II Intron Active Site and Implications for the Spliceosome. *RNA*16: 1–9.

Keeling, P. J.(2007) Genomics. Deep Questions in the Tre of Life. *Science*317: 1,875—1,876.

Keeling, P. J., G. Burger, D. G. Durnford, B. F. Lang, R. W. Lee, R. E. Pearlman, A. J. Roger, and M. W. Gray.(2005) The Tree of Eukaryotes. *Trends Ecol Evol*20: 670–676.

Keeling, P. J., and J. D. Palmer.(2008) Horizontal Gene Transfer in Eukaryotic Evolution. *Nat Rev Genet*9: 605–618.

Keiler, K. C.(2008) Biology of Trans-Translation. *Annu Rev Microbiol*62: 133–151.

Kelley, D. S., J. A. Karson, G. L. Fruh-Green, D. R. Yoerger, T. M. Shank, D. A. Butterfield, J. M. Hayes, M. O. Schrenk, E. J. Olson, G. Proskurowski, M. Jakuba, A. Bradley, B. Larson, K. Ludwig, D. Glickson, K. Buckman, A. S. Bradley, W. J. Brazelton, K. Roe, M. J. Elend, A. Delacour, S. M. Bernasconi, M. D. Lilley, J. A. Baross, R. E. Summons, and S. P. Sylva. (2005) A Serpentinite-Hosted Ecosystem: The Lost City Hydrothermal Field. *Science*307: 1,428—1,434.

Kelly, S., B. Wickstead, and K. Gull.(2010) Archaeal Phylogenomics Provides Evidence in Support of a Methanogenic Origin of the Archaea and a Thaumarchaeal Origin for the Eukaryotes. *Proc Biol Sci*278: 1009–1018.

Khesin, R. B.(1984) Inconstancy of the Genome. Moscow: Nauka.

Kim, V. N., J. Han, and M. C. Siomi.(2009) Biogenesis of Small RNAs in Animals. *Nat Rev Mol Cell Biol*10: 126–139.

- Kimura, M.(1983) *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press.
- Kimura, M.(1991) Recent Development of the Neutral Theory Viewed from the Wrightian Tradition of Theoretical Population Genetics. *Proc Natl Acad Sci USA*88: 5,969—5,973.
- Kipling, R.(2009) *Just So Stories: For Little Children*. Oxford: Oxford University Press.
- Kirschner, M.,and J. Gerhart.(1998) Evolvability. *Proc Natl Acad Sci USA*95: 8,420—8,427.
- Knight, R. D.,and L. F. Landweber. (2000) The Early Evolution of the Genetic Code. *Cell*101: 569—572.
- Knoll, A. H., E. J. Javaux, D. Hewitt, and P. Cohen. (2006) Eukaryotic Organisms in Proterozoic Oceans. *Philos Trans R Soc Lond B Biol Sci*361: 1,023—1,038.
- Kobayashi, I.(1998) Selfishness and Death: Raison D’etre of Restriction, Recombination, and Mitochondria. *Trends Genet*14: 368—374.
- Kobayashi, I.(2001) Behavior of Restriction-Modification Systems As Selfish Mobile Elements and Their Impact on Genome Evolution. *Nucleic Acids Res*29: 3,742—3,756.
- Kogenaru, M., M. G. de Vos, and S. J. Tans. (2009) Revealing Evolutionary Pathways by Fitness Landscape Reconstruction. *Crit Rev Biochem Mol Biol*44: 169—174.
- Kondrashov, F. A., I. B. Rogozin, Y. I. Wolf, and E. V. Koonin. (2002) Selection in the Evolution of Gene Duplications. *Genome Biol*3: RESEARCH0008.
- Koonin, E. V.(1991) Genome Replication/Expression Strategies of Positive-Strand RNA Viruses: A Simple Version of a Combinatorial Classification and Prediction of New Strategies. *Virus Genes*5: 273—281.
- Koonin, E. V.(2003) Comparative Genomics, Minimal Gene-Sets and the Last Universal Common Ancestor. *Nature Rev Microbiol*1: 127—136.
- Koonin, E. V.(2004) A Non-Adaptationist Perspective on Evolution of Genomic Complexity or the Continued Dethroning of Man. *Cell Cycle*3: 280—285.
- Koonin, E. V.(2005) Orthologs, Paralogs and Evolutionary Genomics. *Annu Rev Genet*39: 309—338.
- Koonin, E. V.(2006) The Origin of Introns and Their Role in Eukaryogenesis: A Compromise Solution to the Introns-Early Versus Introns-Late Debate? *Biol Direct*1: 22.
- Koonin, E. V.(2007a) The Biological Big Bang Model for the Major Transitions in Evolution. *Biol Direct*2: 21.
- Koonin, E. V.(2007b) The Cosmological Model of Eternal Inflation and the Transition from Chance to Biological Evolution in the History of Life. *Biol Direct*2: 15.
- Koonin, E. V.(2007c) An RNA-Making Reactor for the Origin of Life. *Proc Natl Acad Sci USA*104: 9,105—9,106.
- Koonin, E. V.(2009a) Evolution of Genome Architecture. *Int J Biochem Cell Biol*41: 298—306.
- Koonin, E. V.(2009b) Intron-Dominated Genomes of Early Ancestors of Eukaryotes. *J Hered*100: 618—623.
- Koonin, E. V.(2009c) On the Origin of Cells and Viruses: Primordial Virus World Scenario. *Ann N Y Acad Sci*1,178: 47—64.
- Koonin, E. V.(2010a) The Origin and Early Evolution of Eukaryotes in the Light of Phylogenomics. *Genome Biol*11: 209.
- Koonin, E. V.(2010b) Preview. The Incredible Expanding Ancestor of Eukaryotes. *Cell*140: 606—608.
- Koonin, E. V.(2010c) Taming of the Shrewd: Novel Eukaryotic Genes from RNA Viruses. *BMC Biol*8: 2.
- Koonin, E. V.(2010d) The Two Empires and the Three Domains of Life in the Postgenomic Age.

*Nature Education*3: 27.

Koonin, E. V., and L. Aravind.(2002) Origin and Evolution of Eukaryotic Apoptosis: The Bacterial Connection. *Cell Death Differ*9: 394–404.

Koonin, E. V., L. Aravind, and A. S. Kondrashov.(2000a) The Impact of Comparative Genomics on Our Understanding of Evolution. *Cell*101: 573–576.

Koonin, E. V., and V. V. Dolja. (1993) Evolution and Taxonomy of Positive-Strand RNA Viruses: Implications of Comparative Analysis of Amino Acid Sequences. *Crit Rev Biochem Mol Biol*28: 375–430.

Koonin, E. V., N. D. Fedorova, J. D. Jackson, A. R. Jacobs, D. M. Krylov, K. S. Makarova, R. Mazumder, S. L. Mekhedov, A. N. Nikolskaya, B. S. Rao, I. B. Rogozin, S. Smirnov, A. V. Sorokin, A. V. Sverdlov, S. Vasudevan, Y. I. Wolf, J. J. Yin, and D. A. Natale.(2004) A Comprehensive Evolutionary Classification of Proteins Encoded in Complete Eukaryotic Genomes. *Genome Biol*5: R7.

Koonin, E. V., A. E. Gorbalenya, and K. M. Chumakov.(1989) Tentative Identification of RNA-Dependent RNA Polymerases of dsRNA Viruses and Their Relationship to Positive Strand RNA Viral Polymerases. *FEBS Lett*252: 42–46.

Koonin, E. V., and K. S. Makarova.(2009) CRISPR-Cas: An Adaptive Immunity System in Prokaryotes. *F1000 Biol Rep*1: 95.

Koonin, E. V., K. S. Makarova, and L. Aravind.(2001a) Horizontal Gene Transfer in Prokaryotes: Quantification and Classification. *Annu Rev Microbiol*55: 709–742.

Koonin, E. V., and W. Martin.(2005) On the Origin of Genomes and Cells within Inorganic Compartments. *Trends Genet*21: 647–654.

Koonin, E. V., and A. R. Mushegian.(1996) Complete Genome Sequences of Cellular Life Forms: Glimpses of Theoretical Evolutionary Genomics. *Curr Opin Genet Dev*6: 757–762.

Koonin, E. V., A. R. Mushegian, and K. E. Rudd.(1996) Sequencing and Analysis of Bacterial Genomes. *Curr Biol*6: 404–416.

Koonin, E. V., and A. S. Novozhilov.(2009) Origin and Evolution of the Genetic Code: The Universal Enigma. *IUBMB Life*61: 99–111.

Koonin, E. V., T. G. Senkevich, and V. V. Dolja. (2006) The Ancient Virus World and Evolution of Cells. *Biol Direct*1: 29.

Koonin, E. V., and Y. I. Wolf. (2006) Evolutionary Systems Biology: Links Between Gene Evolution and Function. *Curr Opin Biotechnol*17: 481–487.

Koonin, E. V., and Y. I. Wolf. (2008a). Evolutionary Systems Biology. Evolutionary Genomics and Proteomics. Ed. by M. Pagel and A. Pomiankowski. Sunderland, MA: Sinauer Associates, Inc.: 11–25.

Koonin, E. V., and Y. I. Wolf. (2008b) Genomics of Bacteria and Archaea: The Emerging Dynamic View of the Prokaryotic World. *Nucleic Acids Res*36: 6,688–6,719.

Koonin, E. V., and Y. I. Wolf.(2009a) The Fundamental Units, Processes and Patterns of Evolution, and the Tree of Life Conundrum. *Biol Direct*4: 33.

Koonin, E. V., and Y. I. Wolf. (2009b) Is Evolution Darwinian or/and Lamarckian? *Biol Direct*4: 42.

Koonin, E. V., and Y. I. Wolf.(2010a) The Common Ancestry of Life. *Biol Direct*5: 64.

Koonin, E. V., and Y. I. Wolf. (2010b) Constraints and Plasticity in Genome and Molecular-Phenome Evolution. *Nat Rev Genet*11: 487–498.

Koonin, E. V., Y. I. Wolf, and L. Aravind.(2000b) Protein Fold Recognition Using Sequence Profiles and Its Application in Structural Genomics. *Adv Protein Chem*54: 245–275.

Koonin, E. V., Y. I. Wolf, and L. Aravind.(2001b) Prediction of the Archaeal Exosome and Its Connections with the Proteasome and the Translation and Transcription Machineries by a

Comparative-Genomic Approach. *Genome Res*11: 240–252.

Koonin, E. V., Y. I. Wolf, and G. P. Karev. (2002) The Structure of the Protein Universe and Genome Evolution. *Nature*420: 218–223.

Koonin, E. V., Y. I. Wolf, K. Nagasaki, and V. V. Dolja. (2008) The Big Bang of Picorna-Like Virus Evolution Antedates the Radiation of Eukaryotic Supergroups. *Nat Rev Microbiol*6: 925–939.

Koonin, E. V., and N. Yutin. (2010) Origin and Evolution of Eukaryotic Large Nucleo-Cytoplasmic DNA Viruses. *Intervirology*53: 284–292.

Kristensen, D. M., A. R. Mushegian, V. V. Dolja, and E. V. Koonin. (2010) New Dimensions of the Virus World Discovered Through Metagenomics. *Trends Microbiol*18: 11–19.

Krylov, D. M., Y. I. Wolf, I. B. Rogozin, and E. V. Koonin. (2003) Gene Loss, Protein Sequence Divergence, Gene Dispensability, Expression Level, and Interactivity Are Correlated in Eukaryotic Evolution. *Genome Res*13: 2,229–2,235.

Kuhn, T. S. (1962) *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Kun, A., M. Santos, and E. Szathmary. (2005) Real Ribozymes Suggest a Relaxed Error Threshold. *Nat Genet*37: 1,008–1,011.

Kunin, V., L. Goldovsky, N. Darzentas, and C. A. Ouzounis. (2005) The Net of Life: Reconstructing the Microbial Phylogenetic Network. *Genome Res*15: 954–959.

Kunin, V., and C. A. Ouzounis. (2003) The Balance of Driving Forces During Genome Evolution in Prokaryotes. *Genome Res*13: 1,589–1,594.

Kurland, C. G. (2005) What Tangled Web: Barriers to Rampant Horizontal Gene Transfer. *Bioessays*27: 741–747.

Kurland, C. G., B. Canback, and O. G. Berg. (2003) Horizontal Gene Transfer: A Critical View. *Proc Natl Acad Sci USA*100: 9,658–9,662.

Kurland, C. G., L. J. Collins, and D. Penny. (2006) Genomics and the Irreducible Nature of Eukaryote Cells. *Science*312: 1,011–1,014.

Lamarck, J.-B. (1809) *Philosophie Zoologique, Ou Exposition Des Considerations Relatives A L'histoire Naturelle Des Animaux*. Paris: Dentu.

Lambowitz, A. M., and S. Zimmerly. (2004) Mobile Group Ii Introns. *Annu Rev Genet*38: 1–35.

Lane, C. E., and J. M. Archibald. (2008) The Eukaryotic Tree of Life: Endosymbiosis Takes Its Toll. *Trends Ecol Evol*23: 268–275.

Lane, N., and W. Martin. (2010) The Energetics of Genome Complexity. *Nature*467: 929–934.

Lang, A. S., and J. T. Beatty. (2007) Importance of Widespread Gene Transfer Agent Genes in Alpha-Proteobacteria. *Trends Microbiol*15: 54–62.

Lang, B. F., M. W. Gray, and G. Burger. (1999) Mitochondrial Genome Evolution and the Origin of Eukaryotes. *Annu Rev Genet*33: 351–397.

Lapierre, P., and J. P. Gogarten. (2009) Estimating the Size of the Bacterial Pan-Genome. *Trends Genet*25: 107–110.

Lathe, W. C. III, B. Snel, and P. Bork. (2000) Gene Context Conservation of a Higher Order Than Operons. *Trends Biochem Sci*25: 474–479.

Lawrence, J. (1999) Selfish Operons: The Evolutionary Impact of Gene Clustering in Prokaryotes and Eukaryotes. *Curr Opin Genet Dev*9: 642–648.

Lawrence, J. G., R. W. Hendrix, and S. Casjens. (2001) Where Are the Pseudo-genes in Bacterial Genomes? *Trends Microbiol*9: 535–540.

Le Hir, H., A. Nott, and M. J. Moore. (2003) How Introns Influence and Enhance Eukaryotic Gene Expression. *Trends Biochem Sci*28: 215–220.

Leipe, D. D., L. Aravind, and E. V. Koonin. (1999) Did DNA Replication Evolve Twice Independently? *Nucleic Acids Res*27: 3,389–3,401.

- Leipe, D. D., E. V. Koonin, and L. Aravind.(2004) Stand, a Class of P-Loop NTPases Including Animal and Plant Regulators of Programmed Cell Death: Multiple, Complex Domain Architectures, Unusual Phyletic Patterns, and Evolution by Horizontal Gene Transfer. *J Mol Biol*343: 1—28.
- Leipe, D. D., Y. I. Wolf, E. V. Koonin, and L. Aravind.(2002) Classification and Evolution of P-Loop GTPases and Related ATPases. *J Mol Biol*317: 41–72.
- Lenski, R. E., C. Ofria, R. T. Pennock, and C. Adami.(2003) The Evolutionary Origin of Complex Features. *Nature*423: 139–144.
- Lespinet, O., Y. I. Wolf, E. V. Koonin, and L. Aravind.(2002) The Role of Lineage-Specific Gene Family Expansion in the Evolution of Eukaryotes. *Genome Res*12: 1,048—1,059.
- Levy, S. F., and M. L. Siegal. (2008) Network Hubs Buffer Environmental Variation in *Saccharomyces Cerevisiae*. *PLoS Biol*6: e264.
- Li, W. H.(1997). *Molecular Evolution*. Sunderland, MA: Sinauer.
- Lincoln, T. A., and G. F. Joyce . (2009) Self-Sustained Replication of an Rna Enzyme. *Science*323: 1,229—1,232.
- Lindberg, J., and J. Lundeberg. (2010) The Plasticity of the Mammalian Transcriptome. *Genomics*95: 1–6.
- Linde, A. D.(1986) Eternally Existing Self-Reproducing Chaotic Inflationary Universe. *Phys Lett B*175: 395.
- Lindell, D., J. D. Jaffe, Z. I. Johnson, G. M. Church, and S. W. Chisholm. (2005) Photosynthesis Genes in Marine Viruses Yield Proteins During Host Infection. *Nature*438: 86–89.
- Ling, J., H. Roy, and M. Ibba. (2007) Mechanism of Trna-Dependent Editing in Translational Quality Control. *Proc Natl Acad Sci USA*104: 72–77.
- Ling, J., and D. Soll. (2010) Severe Oxidative Stress Induces Protein Mistranslation Through Impairment of an Aminoacyl-Trna Synthetase Editing Site. *Proc Natl Acad Sci USA*107: 4,028—4,033.
- Liu, Y., J. Zhou, M. V. Omelchenko, A. S. Beliaev, A. Venkateswaran, J. Stair, L. Wu, D. K. Thompson, D. Xu, I. B. Rogozin, E. K. Gaidamakova, M. Zhai, K. S. Makarova, E. V. Koonin, and M. J. Daly. (2003) Transcriptome Dynamics of *Deinococcus Radiodurans* Recovering from Ionizing Radiation. *Proc Natl Acad Sci USA*100: 4,191—4,196.
- Livio, M., and M. J. Rees. (2005) Cosmology. Anthropic Reasoning. *Science*309: 1,022—1,023.
- Lobkovsky, A. E., Y. I. Wolf, and E. V. Koonin. (2010) Universal Distribution of Protein Evolution Rates As a Consequence of Protein Folding Physics. *Proc Natl Acad Sci USA*107: 2,983—2,988.
- Loewe, L.(2009) A framework for evolutionary systems biology. *BMC Syst Biol*3: 27.
- Lopez-Garcia, P., and D. Moreira. (2006) Selective Forces for the Origin of the Eukaryotic Nucleus. *Bioessays*28: 525–533.
- Lozada-Chavez, I., S. C. Janga, and J. Collado-Vides. (2006) Bacterial Regulatory Networks Are Extremely Flexible in Evolution. *Nucleic Acids Res*34: 3,434—3,445.
- Luria, S. E., and J. Darnell. (1967). *General Virology*. New York: John Wiley.
- Luscombe, N. M., J. Qian, Z. Zhang, T. Johnson, and M. Gerstein. (2002) The Dominance of the Population by a Selected Few: Power-Law Behaviour Applies to a Wide Variety of Genomic Properties. *Genome Biol*3: RESEARCH0040.
- Lynch, M.(2006) The Origins of Eukaryotic Gene Structure. *Mol Biol Evol*23: 450–468.
- Lynch, M.(2007a) The Evolution of Genetic Networks by Non-Adaptive Processes. *Nat Rev Genet*8: 803–813.
- Lynch, M.(2007b) The Frailty of Adaptive Hypotheses for the Origins of Organismal Complexity. *Proc Natl Acad Sci USA*104 Suppl 1: 8,597—8,604.
- Lynch, M.(2007c) *The Origins of Genome Architecture*. Sunderland, MA: Sinauer Associates.

Lynch, M.(2010) Evolution of the Mutation Rate. *Trends Genet*26: 345–352.

Lynch, M., R. Burger, D. Butcher, and W. Gabriel. (1993) The Mutational Meltdown in Asexual Populations. *J Hered*84: 339–344.

Lynch, M.,and J. S. Conery. (2000) The Evolutionary Fate and Consequences of Duplicate Genes. *Science*290: 1,151—1,155.

Lynch, M.,and J. S. Conery. (2003) The Origins of Genome Complexity. *Science*302: 1,401—1,404.

Lynch, M.,and V. Katju. (2004) The Altered Evolutionary Trajectories of Gene Duplicates. *Trends Genet*20: 544–549.

Lyotard, J. F.(1979) *La Condition Postmoderne: Rapport Sur Le Savoir*. Paris: Minuit.

Makarova, K. S., L. Aravind, N. V. Grishin, I. B. Rogozin, and E. V. Koonin . (2002) A DNA Repair System Specific for Thermophilic Archaea and Bacteria Predicted by Genomic Context Analysis. *Nucleic Acids Res*30: 482–496.

Makarova, K. S., L. Aravind, Y. I. Wolf, R. L. Tatusov, K. W. Minton, E. V. Koonin, and M. J. Daly.(2001a) Genome of the Extremely Radiation-Resistant Bacterium *Deinococcus Radiodurans* Viewed from the Perspective of Comparative Genomics. *Microbiol Mol Biol Rev*65: 44–79.

Makarova, K. S., N. V. Grishin, S. A. Shabalina, Y. I. Wolf, and E. V. Koonin . (2006) A Putative RNA-Interference-Based Immune System in Prokaryotes: Computational Analysis of the Predicted Enzymatic Machinery, Functional Analogies with Eukaryotic Rnai, and Hypothetical Mechanisms of Action. *Biol Direct*1: 7.

Makarova, K. S., M. V. Omelchenko, E. K. Gaidamakova, V. Y. Matrosova, A. Vasilenko, M. Zhai, A. Lapidus, A. Copeland, E. Kim, M. Land, K. Mavrommatis, S. Pitluck, P. M. Richardson, C. Detter, T. Brettin, E. Saunders, B. Lai, B. Ravel, K. M. Kemner, Y. I. Wolf, A. Sorokin, A. V. Gerasimova, M. S. Gelfand, J. K. Fredrickson, E. V. Koonin, and M. J. Daly. (2007a) *Deinococcus Geothermalis*: The Pool of Extreme Radiation Resistance Genes Shrinks. *PLoS ONE*2: e955.

Makarova, K. S., V. A. Ponomarev, and E. V. Koonin . (2001b) Two C or Not Two C: Recurrent Disruption of Zn-Ribbons, Gene Duplication, Lineage-Specific Gene Loss, and Horizontal Gene Transfer in Evolution of Bacterial Ribosomal Proteins. *Genome Biol*2: RESEARCH 0033.

Makarova, K. S., A. V. Sorokin, P. S. Novichkov, Y. I. Wolf, and E. V. Koonin . (2007b) Clusters of Orthologous Genes for 41 Archaeal Genomes and Implications for Evolutionary Genomics of Archaea. *Biol Direct*2: 33.

Makarova, K. S., Y. I. Wolf, and E. V. Koonin. (2003) Potential Genomic Determinants of Hyperthermophily. *Trends Genet*19: 172–176.

Makarova, K. S., Y. I. Wolf, and E. V. Koonin . (2009a) Comprehensive Comparative-Genomic Analysis of Type 2 Toxin-Antitoxin Systems and Related Mobile Stress Response Systems in Prokaryotes. *Biol Direct*4: 19.

Makarova, K. S., Y. I. Wolf, S. L. Mekhedov, B. G. Mirkin, and E. V. Koonin . (2005) Ancestral Paralogs and Pseudoparalogs and Their Role in the Emergence of the Eukaryotic Cell. *Nucleic Acids Res*33: 4,626—4,638.

Makarova, K. S., Y. I. Wolf, J. van der Oost, and E. V. Koonin . (2009b) Prokaryotic Homologs of Argonaute Proteins Are Predicted to Function As Key Components of a Novel System of Defense Against Mobile Genetic Elements. *Biol Direct*4: 29.

Makarova, K. S., N. Yutin, S. D. Bell, and E. V. Koonin . (2010) Evolution of Diverse Cell Division and Vesicle Formation Systems in Archaea. *Nat Rev Microbiol*8: 731–741.

Mans, B. J., V. Anantharaman, L. Aravind, and E. V. Koonin . (2004) Comparative Genomics, Evolution, and Origins of the Nuclear Envelope and Nuclear Pore Complex. *Cell Cycle*3: 1,612—1,637.

- Mansy, S. S., J. P. Schrum, M. Krishnamurthy, S. Tobe, D. A. Treco, and J. W. Szostak. (2008) Template-Directed Synthesis of a Genetic Polymer in a Model Protocell. *Nature*454: 122–125.
- Margulis, L.(2009) Genome Acquisition in Horizontal Gene Transfer: Symbiogenesis and Macromolecular Sequence Analysis. *Methods Mol Biol*532: 181–191.
- Margulis, L., M. Chapman, R. Guerrero, and J. Hall. (2006) The Last Eukaryotic Common Ancestor (LECA): Acquisition of Cytoskeletal Motility from Aerotolerant Spirochetes in the Proterozoic Eon. *Proc Natl Acad Sci USA*103: 13,080—13,085.
- Margulis, L., and D. Sagan. (2003). *Acquiring Genomes: The Theory of the Origin of Species*. New York: Basic Books.
- Marinsek, N., E. R. Barry, K. S. Makarova, I. Dionne, E. V. Koonin, and S. D. Bell. (2006) GINS, a Central Nexus in the Archaeal DNA Replication Fork. *EMBO Rep*7: 539–545.
- Marraffini, L. A., and E. J. Sontheimer. (2008) CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA. *Science*322: 1,843—1,845.
- Martin, W.(1999) Mosaic Bacterial Chromosomes: A Challenge En Route to a Tree of Genomes. *Bioessays*21: 99—104.
- Martin, W., J. Baross, D. Kelley, and M. J. Russell. (2008) Hydrothermal Vents and the Origin of Life. *Nat Rev Microbiol*6: 805–814.
- Martin, W., and E. V. Koonin. (2006a) Introns and the Origin of Nucleus-Cytosol Compartmentalization. *Nature*440: 41–45.
- Martin, W., and E. V. Koonin. (2006b) A Positive Definition of Prokaryotes. *Nature*442: 868.
- Martin, W., and K. V. Kowallik. (1999) Annotated English Translation of Mereschkowsky's 1905 Paper «über Natur Und Ursprung Der Chromatophoren Im Pflanzenreiche». *Eur J Phycol*34: 287–296.
- Martin, W., and M. Muller. (1998) The Hydrogen Hypothesis for the First Eukaryote. *Nature*392: 37–41.
- Martin, W., and M. J. Russell. (2003) On the Origins of Cells: A Hypothesis for the Evolutionary Transitions from Abiotic Geochemistry to Chemoautotrophic Prokaryotes, and from Prokaryotes to Nucleated Cells. *Philos Trans R Soc Lond B Biol Sci*358: 59–83; discussion 83–55.
- Martin, W., and M. J. Russell. (2007) On the Origin of Biochemistry at an Alkaline Hydrothermal Vent. *Philos Trans R Soc Lond B Biol Sci*362: 1,887—1,925.
- Martinez, J. L.(2008) Antibiotics and Antibiotic Resistance Genes in Natural Environments. *Science*321: 365–367.
- Masel, J., and A. Bergman. (2003) The Evolution of the Evolvability Properties of the Yeast Prion [Psi+]. *Evolution*57: 1,498—1,512.
- Masel, J., and M. L. Siegal. (2009) Robustness: Mechanisms and Consequences. *Trends Genet*25: 395–403.
- Masel, J., and M. V. Trotter. (2010) Robustness and Evolvability. *Trends Genet*26: 406–414.
- Maslov, S., S. Krishna, T. Y. Pang, and K. Sneppen. (2009) Toolbox Model of Evolution of Prokaryotic Metabolic Networks and Their Regulation. *Proc Natl Acad Sci USA*106: 9,743—9,748.
- Mayr, E.(1963). *Animal Species and Evolution*. Cambridge: Harvard University Press.
- McCarthy, B. J., and J. J. Holland. (1965) Denatured DNA as a Direct Template for in Vitro Protein Synthesis. *Proc Natl Acad Sci USA*54: 880–886.
- McClain, W. H., L. B. Lai, and V. Gopalan. (2010) Trials, Travails, and Triumphs: An Account of RNA Catalysis in RNase P. *J Mol Biol*397: 627–646.
- McClintock, B.(1984) The Significance of Responses of the Genome to Challenge. *Science*226: 792–801.
- McCutcheon, J. P., B. R. McDonald, and N. A. Moran. (2009) Convergent Evolution of Metabolic Roles in Bacterial Co-Symbionts of Insects. *Proc Natl Acad Sci USA*106: 15,394—15,399.

- McDaniel, L. D., E. Young, J. Delaney, F. Ruhnau, K. B. Ritchie, and J. H. Paul. (2010) High Frequency of Horizontal Gene Transfer in the Oceans. *Science*330: 50.
- McDonald, J. H., and M. Kreitman. (1991) Adaptive Protein Evolution at the Adh Locus in *Drosophila*. *Nature*351: 652–654.
- McGeoch, A. T., and S. D. Bell. (2008) Extra-Chromosomal Elements and the Evolution of Cellular DNA Replication Machineries. *Nat Rev Mol Cell Biol*9: 569–574.
- Medvedev, Z. A. (1969) The Rise and Fall of T. D. Lysenko. New York: Columbia University Press.
- Mendell, J. E., K. D. Clements, J. H. Choat, and E. R. Angert. (2008) Extreme Polyploidy in a Large Bacterium. *Proc Natl Acad Sci USA*105: 6,730—6,734.
- Mendes Soares, L. M., and J. Valcarcel. (2006) The Expanding Transcriptome: The Genome as the «Book of Sand». *Embo J*25: 923–931.
- Merkl, R. (2006) A Comparative Categorization of Protein Function Encoded in Bacterial or Archeal Genomic Islands. *J Mol Evol*62: 1—14.
- Mielke, R. E., M. J. Russell, P. R. Wilson, S. E. McGlynn, M. Coleman, R. Kidd, and I. Kanik. (2010) Design, Fabrication, and Test of a Hydrothermal Reactor for Origin-of-Life Experiments. *Astrobiology*10: 799–810.
- Mills, D. R., F. R. Kramer, and S. Spiegelman. (1973) Complete Nucleotide Sequence of a Replicating RNA Molecule. *Science*180: 916–927.
- Mira, A., A. B. Martin-Cuadrado, G. D’Auria, and F. Rodriguez-Valera. (2010) The Bacterial Pan-Genome: A New Paradigm in Microbiology. *Int Microbiol*13: 45–57.
- Mira, A., H. Ochman, and N. A. Moran. (2001) Deletional Bias and the Evolution of Bacterial Genomes. *Trends Genet*17: 589–596.
- Mirkin, B. G., T. I. Fenner, M. Y. Galperin, and E. V. Koonin. (2003) Algorithms for Computing Parsimonious Evolutionary Scenarios for Genome Evolution, the Last Universal Common Ancestor, and Dominance of Horizontal Gene Transfer in the Evolution of Prokaryotes. *BMC Evol Biol*3: 2.
- Mojica, F. J., C. Diez-Villasenor, J. Garcia-Martinez, and E. Soria. (2005) Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *J Mol Evol*60: 174–182.
- Molina, N., and E. van Nimwegen. (2009) Scaling Laws in Functional Genome Content Across Prokaryotic Clades and Lifestyles. *Trends Genet*25: 243–247.
- Monier, A., J. M. Claverie, and H. Ogata. (2008) Taxonomic Distribution of Large DNA Viruses in the Sea. *Genome Biol*9: R106.
- Monod, J. (1972). Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology. New York: Vintage.
- Moreira, D., and P. Lopez-Garcia. (2009) Ten Reasons to Exclude Viruses from the Tree of Life. *Nat Rev Microbiol*7: 306–311.
- Mulkijanian, A. Y. (2009) On the Origin of Life in the Zinc World: 1. Photosynthesizing, Porous Edifices Built of Hydrothermally Precipitated Zinc Sulfide as Cradles of Life on Earth. *Biol Direct*4: 26.
- Mulkijanian, A. Y., M. Y. Galperin, K. S. Makarova, Y. I. Wolf, and E. V. Koonin. (2008) Evolutionary Primacy of Sodium Bioenergetics. *Biol Direct*3: 13.
- Mulkijanian, A. Y., E. V. Koonin, K. S. Makarova, S. L. Mekhedov, A. Sorokin, Y. I. Wolf, A. Dufresne, F. Partensky, H. Burd, D. Kaznadzey, R. Haselkorn, and M. Y. Galperin. (2006) The Cyanobacterial Genome Core and the Origin of Photosynthesis. *Proc Natl Acad Sci USA*103: 13,126—13,131.
- Mulkijanian, A. Y., K. S. Makarova, M. Y. Galperin, and E. V. Koonin. (2007) Inventing the

- Dynamo Machine: The Evolution of the F-Type and V-Type Atpases. *Nat Rev Microbiol*5: 892–899.
- Müller, F., T. Brissac, N. Le Bris, H. Felbeck, and O. Gros. (2010) First Description of Giant Archaea (Thaumarchaeota) Associated with Putative Bacterial Ectosymbionts in a Sulfidic Marine Habitat. *Env Microbiol*12: 2,371—2,383.
- Müller, H. J.(1964) The Relation of Recombination to Mutational Advance. *Mutat Res*106: 2–9.
- Mushegian, A. R.,and E. V. Koonin . (1996a) Gene Order Is Not Conserved in Bacterial Evolution. *Trends Genet*12: 289–290.
- Mushegian, A. R.,and E. V. Koonin.(1996b) A Minimal Gene Set for Cellular Life Derived by Comparison of Complete Bacterial Genomes. *Proc Natl Acad Sci USA*93: 10,268—10,273.
- Nelson, K. E., R. A. Clayton, S. R. Gill, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, J. D. Peterson, W. C. Nelson, K. A. Ketchum, L. McDonald, T. R. Utterback, J. A. Malek, K. D. Linher, M. M. Garrett, A. M. Stewart, M. D. Cotton, M. S. Pratt, C. A. Phillips, D. Richardson, J. Heidelberg, G. G. Sutton, R. D. Fleischmann, J. A. Eisen, O. White, S. L. Salzberg, H. O. Smith, J. C. Venter, and C. M. Fraser. (1999) Evidence for Lateral Gene Transfer Between Archaea and Bacteria from Genome Sequence of *Thermotoga Maritima*. *Nature*399: 323–329.
- Netzer, N., J. M. Goodenbour, A. David, K. A. Dittmar, R. B. Jones, J. R. Schneider, D. Boone, E. M. Eves, M. R. Rosner, J. S. Gibbs, A. Embry, B. Dolan, S. Das, H. D. Hickman, P. Berglund, J. R. Bennink, J. W. Yewdell, and T. Pan. (2009) Innate Immune and Chemically Triggered Oxidative Stress Modifies Translational Fidelity. *Nature*462: 522–526.
- Neupert, W.,and J. M. Herrmann. (2007) Translocation of Proteins into Mitochondria. *Annu Rev Biochem*76: 723–749.
- Novichkov, P. S., M. V. Omelchenko, M. S. Gelfand, A. A. Mironov, Y. I. Wolf, and E. V. Koonin . (2004) Genome-Wide Molecular Clock and Horizontal Gene Transfer in Bacterial Evolution. *J Bacteriol*,in press.
- Novichkov, P. S., Y. I. Wolf, I. Dubchak, and E. V. Koonin. (2009) Trends in Prokaryotic Evolution Revealed by Comparison of Closely Related Bacterial and Archaeal Genomes. *J Bacteriol*191: 65–73.
- Novozhilov, A. S.,and E. V. Koonin. (2009) Exceptional Error Minimization in Putative Primordial Genetic Codes. *Biol Direct*4: 44.
- Nunoura, T., Y. Takaki, J. Kakuta, S. Nishi, J. Sugahara, H. Kazama, G. J. Chee, M. Hattori, A. Kanai, H. Atomi, K. Takai, and H. Takami. (2010) Insights into the Evolution of Archaea and Eukaryotic Protein Modifier Systems Revealed by the Genome of a Novel Archaeal Group. *Nucleic Acids Res*Dec. 15 [Epub ahead of print].
- O'Maille, P. E., A. Malone, N. Dellas, B. Andes Hess Jr., L. Smentek, I. Sheehan, B. T. Greenhagen, J. Chappell, G. Manning, and J. P. Noel . (2008) Quantitative Exploration of the Catalytic Landscape Separating Divergent Plant Sesquiterpene Synthases. *Nat Chem Biol*4: 617–623.
- O'Malley, M. A.(2009) What Did Darwin Say About Microbes, and How Did Microbiology Respond? *Trends Microbiol*17: 341–347.
- O'Malley, M. A.,ed. (2010) Special Issue: The Tree of Life. Biology and Philosophy.
- O'Malley, M. A.,and Y. Boucher. (2005) Paradigm Change in Evolutionary Microbiology. *Stud Hist Philos Biol Biomed Sci*36: 183–208.
- Ochman, H.(2002) Distinguishing the Orfs from the Elfs: Short Bacterial Genes and the Annotation of Genomes. *Trends Genet*18: 335–337.
- Ohno, S.(1970). Evolution by Gene Duplication. New York: Springer-Verlag.
- Ohta, T.(2002) Near-Neutrality in Evolution of Genes and Gene Regulation. *Proc Natl Acad Sci USA*99: 16,134—16,137.
- Omelchenko, M. V., M. Y. Galperin, Y. I. Wolf, and E. V. Koonin. (2010) Non-Homologous Isofunctional Enzymes: A Systematic Analysis of Alternative Solutions in Enzyme Evolution. *Biol*

Omelchenko, M. V., Y. I. Wolf, E. K. Gaidamakova, V. Y. Matrosova, A. Vasilenko, M. Zhai, M. J. Daly, E. V. Koonin, and K. S. Makarova. (2005) Comparative Genomics of *Thermus Thermophilus* and *Deinococcus Radiodurans*: Divergent Routes of Adaptation to Thermophily and Radiation Resistance. *BMC Evol Biol*5: 57.

Oparin, A. I.(1924) *The Origin of Life*. Moscow: Moscow Worker. Oparin, A. I., and V. V. Fesenkov. (1956). *Life in the Universe*. Moscow: USSR Academy of Sciences Publisher.

Orgel, L. E.(1968) Evolution of the Genetic Apparatus. *J Mol Biol*38: 381–393.

Orgel, L. E.,and F. H. Crick. (1980) Selfish DNA: The Ultimate Parasite. *Nature*284: 604–607.

Ostrowski, E. A., R. J. Woods, and R. E. Lenski. (2008) The Genetic Basis of Parallel and Divergent Phenotypic Responses in Evolving Populations of *Escherichia Coli*. *Proc Biol Sci*275: 277–284.

Ouzounis, C. A., V. Kunin, N. Darzentas, and L. Goldovsky. (2006) A Minimal Estimate for the Gene Content of the Last Universal Common Ancestor – Exobiology from a Terrestrial Perspective. *Res Microbiol*157: 57–68.

Pace, N. R.(1997) A Molecular View of Microbial Diversity and the Biosphere. *Science*276: 734–740.

Pace, N. R. (2006) Time for a Change. *Nature*441: 289.

Pace, N. R. (2009a) Mapping the Tree of Life: Progress and Prospects. *Microbiol Mol Biol Rev*73: 565–576.

Pace, N. R.(2009b) Problems with «Procaryote». *J Bacteriol*191: 2,008—2,010; discussion 2,011.

Pais, A.(1994) *Niels Bohr's Times: In Physics, Philosophy and Polity*. Oxford: Oxford University Press.

Pal, C., B. Papp, and L. D. Hurst.(2001) Highly Expressed Genes in Yeast Evolve Slowly. *Genetics*158: 927–931.

Pal, C., B. Papp, and M. J. Lercher. (2005) Adaptive Evolution of Bacterial Metabolic Networks by Horizontal Gene Transfer. *Nat Genet*37: 1,372—1,375.

Pande, V. S., A. Grosberg, T. Tanaka, and D. S. Rokhsar.(1998) Pathways for Protein Folding: Is a New View Needed? *Curr Opin Struct Biol*8: 68–79.

Patterson, D. J.(1999) The Diversity of Eukaryotes. *Am Nat*154: S96—124.

Paul, J. H.(2008) Prophages in Marine Bacteria: Dangerous Molecular Time Bombs or the Key to Survival in the Seas? *ISME J*2: 579–589.

Pennisi, E.(1999) Is It Time to Uproot the Tree of Life? *Science*284: 1,305—1,307.

Pennisi, E.(2009) History of Science. The Case of the Midwife Toad: Fraud or Epigenetics? *Science*325: 1,194—1,195.

Penny, D.(2005) An Interpretative Review of the Origin of Life Research. *Biol Philos*,in press.

Pereto, J., P. Lopez-Garcia, and D. Moreira.(2004) Ancestral Lipid Biosynthesis and Early Membrane Evolution. *Trends Biochem Sci*29: 469–477.

Perna, N. T., G. Plunkett III, V. Burland, B. Mau, J. D. Glasner, D. J. Rose, G. F. Mayhew, P. S. Evans, J. Gregor, H. A. Kirkpatrick, G. Posfai, J. Hackett, S. Klink, A. Boutin, Y. Shao, L. Miller, E. J. Grotbeck, N. W. Davis, A. Lim, E. T. Dimalanta, K. D. Potamousis, J. Apodaca, T. S. Anantharaman, J. Lin, G. Yen, D. C. Schwartz, R. A. Welch, and F. R. Blattner. (2001) Genome Sequence of Enterohaemorrhagic *Escherichia Coli* O157:H7. *Nature*409: 529–533.

Pisani, D., J. A. Cotton, and J. O. McInerney.(2007) Supertrees Disentangle the Chimerical Origin of Eukaryotic Genomes. *Mol Biol Evol*24: 1,752—1,760.

Poole, A. M.(2009) Horizontal Gene Transfer and the Earliest Stages of the Evolution of Life. *Res Microbiol*160: 473–480.

- Poole, A. M., D. C. Jeffares, and D. Penny. (1998) The Path from the Rna World. *J Mol Evol*46: 1—17.
- Poole, A. M., and D. Penny. (2007) Evaluating Hypotheses for the Origin of Eukaryotes. *Bioessays*29: 74–84.
- Popper, K. R. (1982) *Unended Quest: An Intellectual Autobiography*. London: Open Court Publishing Company.
- Prangishvili, D., P. Forterre, and R. A. Garrett. (2006a) Viruses of the Archaea: A Unifying View. *Nat Rev Microbiol*4: 837–848.
- Prangishvili, D., R. A. Garrett, and E. V. Koonin. (2006b) Evolutionary Genomics of Archaeal Viruses: Unique Viral Genomes in the Third Domain of Life. *Virus Res*117: 52–67.
- Price, M.N., P. S. Dehal, and A. P. Arkin. (2010) FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One*5: e9490.
- Prusiner, S. B. (1998) Prions. *Proc Natl Acad Sci USA*95: 13,363—13,383.
- Puigbo, P., Y. I. Wolf, and E. V. Koonin. (2009) Search for a «Tree of Life» in the Thicket of the Phylogenetic Forest. *J Biol*8: 59.
- Puigbo, P., Y. I. Wolf, and E. V. Koonin. (2010) The Tree and Net Components of Prokaryote Evolution. *Genome Biol Evol*2: 745–756.
- Putnam, N. H., M. Srivastava, U. Hellsten, B. Dirks, J. Chapman, A. Salamov, A. Terry, H. Shapiro, E. Lindquist, V. V. Kapitonov, J. Jurka, G. Genikhovich, I. V. Grigoriev, S. M. Lucas, R. E. Steele, J. R. Finnerty, U. Technau, M. Q. Martindale, and D. S. Rokhsar. (2007) Sea Anemone Genome Reveals Ancestral Eumetazoan Gene Repertoire and Genomic Organization. *Science*317: 86–94.
- Radman, M. (1975) SOS Repair Hypothesis: Phenomenology of an Inducible DNA Repair Which Is Accompanied by Mutagenesis. *Basic Life Sci*5A: 355–367.
- Radman, M., I. Matic, and F. Taddei. (1999) Evolution of Evolvability. *Ann N Y Acad Sci*870: 146–155.
- Ragan, M. A. (2001) Detection of Lateral Gene Transfer among Microbial Genomes. *Curr Opin Genet Dev*11: 620–626.
- Ragan, M. A., J. O. McInerney, and J. A. Lake, eds. (2009) Theme Issue: The Network of Life: Genome Beginnings and Evolution. *Phil Trans R Soc B*.
- Rao, V. B., and M. Feiss. (2008) The Bacteriophage DNA Packaging Motor. *Annu Rev Genet*42: 647–681.
- Raoult, D. (2010) The Post-Darwinist Rhizome of Life. *Lancet*375: 104–105.
- Raoult, D., S. Audic, C. Robert, C. Abergel, P. Renesto, H. Ogata, B. La Scola, M. Suzan, and J. M. Claverie. (2004) The 1.2-Megabase Genome Sequence of Mimivirus. *Science*306: 1,344—1,350.
- Raoult, D., and P. Forterre. (2008) Redefining Viruses: Lessons from Mimivirus. *Nat Rev Microbiol*6: 315–319.
- Ravasi, T., H. Suzuki, C. V. Cannistraci, S. Katayama, V. B. Bajic, K. Tan, A. Akalin, S. Schmeier, M. Kanamori-Katayama, N. Bertin, P. Carninci, C. O. Daub, A. R. Forrest, J. Gough, S. Grimmond, J. H. Han, T. Hashimoto, W. Hide, O. Hofmann, A. Kamburov, M. Kaur, H. Kawaji, A. Kubosaki, T. Lassmann, E. van Nimwegen, C. R. MacPherson, C. Ogawa, A. Radovanovic, A. Schwartz, R. D. Teasdale, J. Tegner, B. Lenhard, S. A. Teichmann, T. Arakawa, N. Ninomiya, K. Murakami, M. Tagami, S. Fukuda, K. Imamura, C. Kai, R. Ishihara, Y. Kitazume, J. Kawai, D. A. Hume, T. Ideker, and Y. Hayashizaki. (2010) An Atlas of Combinatorial Transcriptional Regulation in Mouse and Man. *Cell*140: 744–752.
- Raymond, J., O. Zhaxybayeva, J. P. Gogarten, S. Y. Gerdes, and R. E. Blankenship. (2002) Whole-Genome Analysis of Photosynthetic Prokaryotes. *Science*298: 1,616—1,620.
- Rees, M. (2001) *Our Cosmic Habitat*. Princeton: Princeton University Press.
- Richards, E. J. (2006) Inherited Epigenetic Variation – Revisiting Soft Inheritance. *Nat Rev Genet*7: 395–401.

- Richards, T. A., and T. Cavalier-Smith.(2005) Myosin Domain Evolution and the Primary Divergence of Eukaryotes. *Nature*436: 1,113—1,118.
- Ridley, M.(2006) *Genome*. New York: Harper Perennial.
- Rivera, M. C.,and J. A. Lake.(2004) The Ring of Life Provides Evidence for a Genome Fusion Origin of Eukaryotes. *Nature*431: 152–155.
- Robertson, M. P., S. M. Knudsen, and A. D. Ellington.(2004) In Vitro Selection of Ribozymes Dependent on Peptides for Activity. *RNA*10: 114–127.
- Roger, A. J.(1999) Reconstructing Early Events in Eukaryotic Evolution. *Am Nat*154: S146—163.
- Rogozin, I. B., M. K. Basu, M. Csuros, and E. V. Koonin. (2009) Analysis of Rare Genomic Changes Does Not Support the Unikont-Bikont Phylogeny and Suggests Cyanobacterial Symbiosis As the Point of Primary Radiation of Eukaryotes. *Genome Biol Evol*1: 99—113.
- Rogozin, I. B., K. S. Makarova, J. Murvai, E. Czabarka, Y. I. Wolf, R. L. Tatusov, L. A. Szekely, and E. V. Koonin. (2002) Connected Gene Neighborhoods in Prokaryotic Genomes. *Nucleic Acids Res*30: 2,212—2,223.
- Rogozin, I. B., Y. I. Wolf, A. V. Sorokin, B. G. Mirkin, and E. V. Koonin. (2003) Remarkable Interkingdom Conservation of Intron Positions and Massive, Lineage-Specific Intron Loss and Gain in Eukaryotic Evolution. *Curr Biol*13: 1,512—1,517.
- Rokas, A., and S. B. Carroll.(2006) Bushes in the Tree of Life. *PLoS Biol*4: e352. Rosenberg, S. M. (2001) Evolving Responsively: Adaptive Mutation. *Nat Rev Genet*2: 504–515.
- Russell, M. J.(2007) The Alkaline Solution to the Emergence of Life: Energy, Entropy, and Early Evolution. *Acta Biotheor*55: 133–179.
- Russell, M. J., and A. J. Hall.(1997) The Emergence of Life from Iron Monosulphide Bubbles at a Submarine Hydrothermal Redox and Ph Front. *J Geol Soc London*154: 377–402.
- Russell, R.(2008) RNA Misfolding and the Action of Chaperones. *Front Biosci*13: 1—20.
- Rutherford, S., Y. Hirate, and B. J. Swalla.(2007) The Hsp90 Capacitor, Developmental Remodeling, and Evolution: The Robustness of Gene Networks and the Curious Evolvability of Metamorphosis. *Crit Rev Biochem Mol Biol*42: 355–372.
- Sagan, L.(1967) On the Origin of Mitosing Cells. *J Theor Biol*14: 255–274.
- Salgado, H., G. Moreno-Hagelsieb, T. F. Smith, and J. Collado-Vides.(2000) Operons in Escherichia Coli: Genomic Analyses and Predictions. *Proc Natl Acad Sci USA*97: 6,652—6,657.
- Sapp, J.(2009) *The New Foundations of Evolution: On the Tree of Life*. Oxford: Oxford University Press.
- Saxinger, C., and C. Ponnampereuma.(1974) Interactions Between Amino Acids and Nucleotides in the Prebiotic Milieu. *Orig Life*5: 189–200.
- Schrimpf, S. P., M. Weiss, L. Reiter, C. H. Ahrens, M. Jovanovic, J. Malmstrom, E. Brunner, S. Mohanty, M. J. Lercher, P. E. Hunziker, R. Aebersold, C. von Mering, and M. O. Hengartner.(2009) Comparative Functional Analysis of the Caenorhabditis Elegans and Drosophila Melanogaster Proteomes. *PLoS Biol*7: e48.
- Schroedinger, E. *What Is Life?: With Mind and Matter and Autobiographical Sketches*. Cambridge: Cambridge University Press.
- Sella, G., and A. E. Hirsh.(2005) The Application of Statistical Physics to Evolutionary Biology. *Proc Natl Acad Sci USA*102: 9,541—9,546.
- Seshasayee, A. S., G. M. Fraser, and N. M. Luscombe.(2010) Comparative Genomics of Cyclic-Di-Gmp Signalling in Bacteria: Post-Translational Regulation and Catalytic Activity. *Nucleic Acids Res*38: 5,970—5,981.
- Shabalina, S. A., and E. V. Koonin. (2008) Origins and Evolution of Eukaryotic RNA Interference. *Trends Ecol Evol*23: 578–587.

- Shiflett, A. M., and P. J. Johnson. (2010) Mitochondrion-Related Organelles in Eukaryotic Protists. *Annu Rev Microbiol*64: 409–429.
- Shnol, S. E. (2001) Heroes, Villains, Conformists of Russian Science. Moscow: Kron-Press.
- Sicheritz-Ponten, T., and S. G. Andersson. (2001) A Phylogenomic Approach to Microbial Evolution. *Nucleic Acids Res*29: 545–552.
- Simpson, G. G. (1983) Tempo and Mode in Evolution. New York: Columbia University Press.
- Smith, M. W., D. F. Feng, and R. F. Doolittle. (1992) Evolution by Acquisition: The Case for Horizontal Gene Transfers. *Trends Biochem Sci*17: 489–493.
- Smolin, L. (1999) The Life of the Cosmos. Oxford: Oxford University Press.
- Sniegowski, P. D., P. J. Gerrish, and R. E. Lenski. (1997) Evolution of High Mutation Rates in Experimental Populations of *E. coli*. *Nature*387: 703–705.
- Snel, B., P. Bork, and M. A. Huynen. (2002) Genomes in Flux: The Evolution of Archaeal and Proteobacterial Gene Content. *Genome Res*12: 17–25.
- Soyfer, V. N. (1994) Lysenko and the Tragedy of Soviet Science. New Brunswick, NJ: Rutgers University Press.
- Soyfer, V. N. (2001) The Consequences of Political Dictatorship for Russian Science. *Nat Rev Genet*2: 723–729.
- Spiegelman, S. (1971) An Approach to the Experimental Analysis of Precellular Evolution. *Q Rev Biophys*4: 213–253.
- Srivastava, M., E. Begovic, J. Chapman, N. H. Putnam, U. Hellsten, T. Kawashima, A. Kuo, T. Mitros, A. Salamov, M. L. Carpenter, A. Y. Signorovitch, M. A. Moreno, K. Kamm, J. Grimwood, J. Schmutz, H. Shapiro, I. V. Grigoriev, L. W. Buss, B. Schierwater, S. L. Dellaporta, and D. S. Rokhsar. (2008) The Trichoplax Genome and the Nature of Placozoans. *Nature*454: 955–960.
- Srivastava, M., O. Simakov, J. Chapman, B. Fahey, M. E. Gauthier, T. Mitros, G. S. Richards, C. Conaco, M. Dacre, U. Hellsten, C. Larroux, N. H. Putnam, M. Stanke, M. Adamska, A. Darling, S. M. Degnan, T. H. Oakley, D. C. Plachetzki, Y. Zhai, M. Adamski, A. Calcino, S. F. Cummins, D. M. Goodstein, C. Harris, D. J. Jackson, S. P. Leys, S. Shu, B. J. Woodcroft, M. Vervoort, K. S. Kosik, G. Manning, B. M. Degnan, and D. S. Rokhsar. (2010) The Amphimedon *Queenslandica* Genome and the Evolution of Animal Complexity. *Nature*466: 720–726.
- Stalder, L., and O. Muhlemann. (2008) The Meaning of Nonsense. *Trends Cell Biol*18: 315–321.
- Stanier, R. Y., and C. B. Van Niel. (1962) The Concept of a Bacterium. *Arch Mikrobiol*42: 17–35.
- Stechmann, A., and T. Cavalier-Smith. (2003) The Root of the Eukaryote Tree Pinpointed. *Curr Biol*13: R665–666.
- Stern, A., L. Keren, O. Wurtzel, G. Amitai, and R. Sorek. (2010) Self-Targeting by CRISPR: Gene Regulation or Autoimmunity? *Trends Genet*26: 335–340.
- Stoltzfus, A. (1999) On the Possibility of Constructive Neutral Evolution. *J Mol Evol*49: 169–181.
- Stover, C. K., X. Q. Pham, A. L. Erwin, S. D. Mizoguchi, P. Warrener, M. J. Hickey, F. S. Brinkman, W. O. Hufnagle, D. J. Kowalik, M. Lagrou, R. L. Garber, L. Goltry, E. Tolentino, S. Westbrook-Wadman, Y. Yuan, L. L. Brody, S. N. Coulter, K. R. Folger, A. Kas, K. Larbig, R. Lim, K. Smith, D. Spencer, G. K. Wong, Z. Wu, I. T. Paulsen, J. Reizer, M. H. Saier, R. E. Hancock, S. Lory, and M. V. Olson. (2000) Complete Genome Sequence of *Pseudomonas Aeruginosa* Pao1, an Opportunistic Pathogen. *Nature*406: 959–964.
- Sullivan, M. B., D. Lindell, J. A. Lee, L. R. Thompson, J. P. Bielawski, and S. W. Chisholm. (2006) Prevalence and Evolution of Core Photosystem II Genes in Marine Cyanobacterial Viruses and Their Hosts. *PLoS Biol*4: e234.
- Susskind, L. (2003) The Anthropic Landscape of String Theory. *arXiv: hep-th/0302219*.
- Susskind, L. (2006a) The Cosmic Landscape. String Theory and the Illusion of Intelligent Design.

New York-Boston: Little, Brown and Company.

Susskind, L.(2006b) *The Cosmic Landscape: String Theory and the Illusion of Intelligent Design*. San Francisco: Back Bay Books.

Suttle, C. A.(2005) Viruses in the Sea. *Nature*437: 356–361.

Suttle, C. A.(2007) Marine Viruses – Major Players in the Global Ecosystem. *Nat Rev Microbiol*5: 801–812.

Sydow, J. F., and P. Cramer. (2009) RNA Polymerase Fidelity and Transcriptional Proofreading. *Curr Opin Struct Biol*19: 732–739.

Syvanen, M.(1994) Horizontal Gene Transfer: Evidence and Possible Consequences. *Annu Rev Genet*28: 237–261.

Szathmary, E.(1993) Coding Coenzyme Handles: A Hypothesis for the Origin of the Genetic Code. *Proc Natl Acad Sci USA*90: 9,916—9,920.

Szathmary, E.(1999) The Origin of the Genetic Code: Amino Acids as Cofactors in an RNA World. *Trends Genet*15: 223–229.

Szathmary, E.(2000) The Evolution of Replicators. *Philos Trans R Soc Lond B Biol Sci*355: 1,669—1,676.

Szathmary, E.,and L. Demeter.(1987) Group Selection of Early Replicators and the Origin of Life. *J Theor Biol*128: 463–486.

Szathmary, E.,and J. Maynard Smith.(1997) From Replicators to Reproducers: The First Major Transitions Leading to Life. *J Theor Biol*187: 555–571.

Takeuchi, N.,and P. Hogeweg. (2008) Evolution of Complexity in RNA-Like Replicator Systems. *Biol Direct*3: 11.

Takeuchi, N., P. Hogeweg, and E. V. Koonin. (2011) On the Origin of DNA Genomes: Evolution of the Division of Labor Between Template and Catalyst in Model Replicator Systems. *PLoS Comput Biol*,in press.

Tatusov, R. L., N. D. Fedorova, J. D. Jackson, A. R. Jacobs, B. Kiryutin, E. V. Koonin, D. M. Krylov, R. Mazumder, S. L. Mekhedov, A. N. Nikolskaya, B. S. Rao, S. Smirnov, A. V. Sverdlov, S. Vasudevan, Y. I. Wolf, J. J. Yin, and D. A. Natale.(2003) The Cog Database: An Updated Version Includes Eukaryotes. *BMC Bioinformatics*4: 41.

Tatusov, R. L., E. V. Koonin,and D. J. Lipman.(1997) A Genomic Perspective on Protein Families. *Science*278: 631–637.

Tejero, H., A. Marin, and F. Montero. (2011) The Relationship Between Error Catastrophe, Survival of the Flattest, and Natural Selection. *BMC Evol Biol*11: 2.

Theobald, D. L.(2010) A Formal Test of the Theory of Universal Common Ancestry. *Nature*465: 219–222.

Toor, N., K. S. Keating, S. D. Taylor, and A. M. Pyle.(2008) Crystal Structure of a Self-Spliced Group Ii Intron. *Science*320: 77–82.

Treangen, T. J., and E. P. Rocha. (2011) Horizontal Transfer, Not Duplication, Drives the Expansion of Protein Families in Prokaryotes. *PLoS Genet*7: e1001284.

Trifonov, E. N.(2004) The Triplet Code from First Principles. *J Biomol Struct Dyn*22: 1—11.

Trifonov, E. N., I. Gabdank, D. Barash, and Y. Sobolevsky. (2006) Primordia Vita. Deconvolution from Modern Sequences. *Orig Life Evol Biosph*36: 559–565.

Tuite, M. F., and T. R. Serio.(2010) The Prion Hypothesis: From Biological Anomaly to Basic Regulatory Mechanism. *Nat Rev Mol Cell Biol*11: 823–833.

Turk, R. M., N. V. Chumachenko, and M. Yarus.(2010) Multiple Translational Products from a Five-Nucleotide Ribozyme. *Proc Natl Acad Sci USA*107: 4,585—4,589.

Tyedmers, J., M. L. Madariaga, and S. Lindquist.(2008) Prion Switching in Response to

Environmental Stress. *PLoS Biol*6: e294.

Ulrich, L. E., E. V. Koonin, and I. B. Zhulin.(2005) One-Component Systems Dominate Signal Transduction in Prokaryotes. *Trends Microbiol*13: 52–56.

Vabulas, R. M., S. Raychaudhuri, M. Hayer-Hartl, and F. U. Hartl. (2010) Protein Folding in the Cytoplasm and the Heat Shock Response. *Cold Spring Harb Perspect Biol*2: a004390.

Van den Born, E., M. V. Omelchenko, A. Bekkelund, V. Leihne, E. V. Koonin, V. V. Dolja, and P. O. Falnes.(2008) Viral Alkb Proteins Repair RNA Damage by Oxidative Demethylation. *Nucleic Acids Res*36: 5,451—5,461.

Van der Giezen, M.(2009) Hydrogenosomes and Mitosomes: Conservation and Evolution of Functions. *J Eukaryot Microbiol*56: 221–231.

Van der Oost, J., M. M. Jore, E. R. Westra, M. Lundgren, and S. J. Brouns.(2009) CRISPR-Based Adaptive and Heritable Immunity in Prokaryotes. *Trends Biochem Sci*34: 401–407.

Van Etten, J. L., L. C. Lane, and D. D. Dunigan.(2010) DNA Viruses: The Really Big Ones (Giruses). *Annu Rev Microbiol*64: 83–99.

Van Melderen, L.(2010) Toxin – Antitoxin Systems: Why So Many, What For? *Curr Opin Microbiol*13: 781–785.

Van Melderen, L., and M. Saavedra De Bast.(2009) Bacterial Toxin – Antitoxin Systems: More Than Selfish Entities? *PLoS Genet*5: e1000437.

Van Niel, C. B.(1955) Natural Selection in the Microbial World. *J Gen Microbiol*13: 201–217.

Van Nimwegen, E.(2003) Scaling Laws in the Functional Content of Genomes. *Trends Genet*19: 479–484.

Van Valen, L.(1973) A New Evolutionary Law. *Evol. Tehory*1: 1—30.

Vargas, A. O.(2009) Did Paul Kammerer Discover Epigenetic Inheritance? A Modern Look at the Controversial Midwife Toad Experiments. *J Exp Zool B Mol Dev Evol*312: 667–678.

Venn, J.(1866) The Logic of Chance: An Essay on the Foundations and Province of the Theory of Probability, with Especial Reference to its Application to Moral and Socila Science. London and Cambridge: MacMillan and Co.

Venters, B. J., and B. F. Pugh. (2009) How Eukaryotic Genes Are Transcribed. *Crit Rev Biochem Mol Biol*44: 117–141.

Veretnik, S., C. Wills, P. Youkharibache, R. E. Valas, and P. E. Bourne. (2009) Sm/Lsm Genes Provide a Glimpse into the Early Evolution of the Spliceosome. *PLoS Comput Biol*5: e1000315.

Vetsigian, K., C. Woese, and N. Goldenfeld.(2006) Collective Evolution and the Genetic Code. *Proc Natl Acad Sci USA*103: 10,696—10,701.

Vilenkin, A.(1983) The Birth of Inflationary Universes. *Phys. Rev. D*27: 2,848.

Vilenkin, A.(2007) Many Worlds in One: The Search for Other Universes. Boston: Hill and Wang.

Vogel, G.(1997) Prusiner Recognized for Once-Heretical Prion Theory. *Science*278: 214.

Vogel, G., and E. Pennisi.(2009) Physiology Nobel. U.S. Researchers Recognized for Work on Telomeres. *Science*326: 212–213.

Volker, C., and A. N. Lupas.(2002) Molecular Evolution of Proteasomes. *Curr Top Microbiol Immunol*268: 1—22.

Von Dohlen, C. D., S. Kohler, S. T. Alsop, and W. R. McManus.(2001) Mealybug Beta-Proteobacterial Endosymbionts Contain Gamma-Proteobacterial Symbionts. *Nature*412: 433–436.

Von Neumann, J.(1966) Theory of Self-Reproducing Automata. Urbana, IL: University of Illinois Press.

Wachtershauser, G.(1997) The Origin of Life and Its Methodological Challenge. *J Theor Biol*187: 483–494.

Waddington, C. H., and E. Robertson.(1966) Selection for Developmental Canalisation. *Genet*

Wagner, A. (2008a) Neutralism and Selectionism: A Network-Based Reconciliation. *Nat Rev Genet*9: 965–974.

Wagner, A. (2008b) Robustness and Evolvability: A Paradox Resolved. *Proc Biol Sci*275: 91–100.

Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gay. (1982) Distantly Related Sequences in the Alpha – and Beta-Subunits of Atp Synthase, Myosin, Kinases, and Other Atp-Requiring Enzymes and a Common Nucleotide Binding Fold. *Embo J*1: 945–951.

Wang, E. T., R. Sandberg, S. Luo, I. Khrebtukova, L. Zhang, C. Mayr, S. F. Kingsmore, G. P. Schroth, and C. B. Burge. (2008) Alternative Isoform Regulation in Human Tissue Transcriptomes. *Nature*456: 470–476.

Wang, X., Y. Kim, Q. Ma, S. H. Hong, K. Pokusaeva, J. M. Sturino, and T. K. Wood. (2010) Cryptic Prophages Help Bacteria Cope with Adverse Environments. *Nat Commun*1: 147.

Wang, Z., and J. Zhang. (2009) Why Is the Correlation Between Gene Importance and Gene Evolutionary Rate So Weak? *PLoS Genet*5: e1000329.

Watson, J. D., and F. H. Crick. (1953a) Genetical Implications of the Structure of Deoxyribonucleic Acid. *Nature*171: 964–967.

Watson, J. D., and F. H. Crick. (1953b) Molecular Structure of Nucleic Acids; a Structure for Deoxyribose Nucleic Acid. *Nature*171: 737–738.

Watts, D. J. (2004) Six Degrees: The Science of a Connected Age. New York: W. W. Norton & Co.

Weinberg, S. (1994) Dreams of a Final Theory: The Scientist's Search for the Ultimate Laws of Nature. New York: Vintage.

Weinreich, D. M., N. F. Delaney, M. A. Depristo, and D. L. Hartl. (2006) Darwinian Evolution Can Follow Only Very Few Mutational Paths to Fitter Proteins. *Science*312: 111–114.

Weissmann, A. (1893) The Germ-Plasm. A Theory of Heredity. London: Charles Scribner's Sons.

Whitehead, D. J., C. O. Wilke, D. Vernazobres, and E. Bornberg-Bauer. (2008) The Look-Ahead Effect of Phenotypic Mutations. *Biol Direct*3: 18.

Wilke, C. O., J. L. Wang, C. Ofria, R. E. Lenski, and C. Adami. (2001) Evolution of Digital Organisms at High Mutation Rates Leads to Survival of the Flattest. *Nature*412: 331–333.

Wilson, A. C., S. S. Carlson, and T. J. White. (1977) Biochemical Evolution. *Annu Rev Biochem*46: 573–639.

Wochner, A., J. Attwater, A. Coulson, and P. Holliger. (2011) Ribozyme-Catalyzed Transcription of an Active Ribozyme. *Science*332: 209–212.

Woese, C. (1967) The Genetic Code. New York: Harper & Row.

Woese, C. (1998) The Universal Ancestor. *Proc Natl Acad Sci USA*95: 6,854–6,859.

Woese, C. R. (1987) Bacterial Evolution. *Microbiol Rev*51: 221–271.

Woese, C. R. (2002) On the Evolution of Cells. *Proc Natl Acad Sci USA*99: 8,742–8,747.

Woese, C. R., and G. E. Fox. (1977) The Concept of Cellular Evolution. *J Mol Evol*10: 1–6.

Woese, C. R., and N. Goldenfeld. (2009) How the Microbial World Saved Evolution from the Scylla of Molecular Biology and the Charybdis of the Modern Synthesis. *Microbiol Mol Biol Rev*73: 14–21.

Woese, C. R., O. Kandler, and M. L. Wheelis. (1990) Towards a Natural System of Organisms: Proposal for the Domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA*87: 4,576–4,579.

Wohlgemuth, I., M. Beringer, and M. V. Rodnina. (2006) Rapid Peptide Bond Formation on Isolated 50s Ribosomal Subunits. *EMBO Rep*7: 699–703.

Wolf, Y. I., L. Aravind, N. V. Grishin, and E. V. Koonin. (1999a) Evolution of Aminoacyl-tRNA

Synthetases – Analysis of Unique Domain Architectures and Phylogenetic Trees Reveals a Complex History of Horizontal Gene Transfer Events. *Genome Res*9: 689–710.

Wolf, Y. I., S. E. Brenner, P. A. Bash, and E. V. Koonin. (1999b) Distribution of Protein Folds in the Three Superkingdoms of Life. *Genome Res*9: 17–26.

Wolf, Y. I., L. Carmel, and E. V. Koonin. (2006) Unifying Measures of Gene Function and Evolution. *Proc Biol Sci*273: 1,507—1,515.

Wolf, Y. I., I. V. Gopich, D. J. Lipman, and E. V. Koonin. (2010) Relative Contributions of Intrinsic Structural-Functional Constraints and Translation Rate to the Evolution of Protein-Coding Genes. *Genome Biol Evol*2: 190–199.

Wolf, Y. I., and E. V. Koonin. (2007) On the Origin of the Translation System and the Genetic Code in the RNA World by Means of Natural Selection, Exaptation, and Subfunctionalization. *Biol Direct*2: 14.

Wolf, Y. I., P. S. Novichkov, G. P. Karev, E. V. Koonin, and D. J. Lipman. (2009) The Universal Distribution of Evolutionary Rates of Genes and Distinct Characteristics of Eukaryotic Genes of Different Apparent Ages. *Proc Natl Acad Sci USA*106: 7,273—7,280.

Wolf, Y. I., I. B. Rogozin, N. V. Grishin, and E. V. Koonin. (2002) Genome Trees and the Tree of Life. *Trends Genet*18: 472–479.

Wolf, Y. I., I. B. Rogozin, A. S. Kondrashov, and E. V. Koonin. (2001) Genome Alignment, Evolution of Prokaryotic Genome Organization, and Prediction of Gene Function Using Genomic Context. *Genome Res*11: 356–372.

Woods, R. J., J. E. Barrick, T. F. Cooper, U. Shrestha, M. R. Kauth, and R. E. Lenski. (2011) Second-Order Selection for Evolvability in a Large Escherichia Coli Population. *Science*331: 1,433—1,436.

Woodson, S. A. (2010) Taming Free Energy Landscapes with RNA Chaperones. *RNA Biol*7: 38–47.

Wozniak, R. A., and M. K. Waldor. (2010) Integrative and Conjugative Elements: Mosaic Mobile Genetic Elements Enabling Dynamic Lateral Gene Flow. *Nat Rev Microbiol*8: 552–563.

Wright, G. D. (2007) The Antibiotic Resistome: The Nexus of Chemical and Genetic Diversity. *Nat Rev Microbiol*5: 175–186.

Wu, D., S. C. Daugherty, S. E. Van Aken, G. H. Pai, K. L. Watkins, H. Khouri, L. J. Tallon, J. M. Zaborsky, H. E. Dunbar, P. L. Tran, N. A. Moran, and J. A. Eisen. (2006) Metabolic Complementarity and Genomics of the Dual Bacterial Symbiosis of Sharpshooters. *PLoS Biol*4: e188.

Wunderlich, Z., and L. A. Mirny. (2009) Different Gene Regulation Strategies Revealed by Analysis of Binding Motifs. *Trends Genet*25: 434–440.

Yang, D., Y. Oyaizu, H. Oyaizu, G. J. Olsen, and C. R. Woese. (1985) Mitochondrial Origins. *Proc Natl Acad Sci USA*82: 4,443—4,447.

Yang, Z. (2007) Paml 4: Phylogenetic Analysis by Maximum Likelihood. *Mol Biol Evol*24: 1,586—1,591.

Yarus, M. (1998) Amino Acids as RNA Ligands: A Direct-RNA-Template Theory for the Code's Origin. *J Mol Evol*47: 109–117.

Yarus, M., J. G. Caporaso, and R. Knight. (2005) Origins of the Genetic Code: The Escaped Triplet Theory. *Annu Rev Biochem*74: 179–198.

Yarus, M., J. J. Widmann, and R. Knight. (2009) RNA – Amino Acid Binding: A Stereochemical Era for the Genetic Code. *J Mol Evol*69: 406–429.

Yooseph, S., G. Sutton, D. B. Rusch, A. L. Halpern, S. J. Williamson, K. Remington, J. A. Eisen, K. B. Heidelberg, G. Manning, W. Li, L. Jaroszewski, P. Cieplak, C. S. Miller, H. Li, S. T. Mashiyama, M. P. Joachimiak, C. van Belle, J. M. Chandonia, D. A. Soergel, Y. Zhai, K. Natarajan, S. Lee, B. J. Raphael, V. Bafna, R. Friedman, S. E. Brenner, A. Godzik, D. Eisenberg, J. E. Dixon, S. Taylor, R. L.

Strausberg, M. Frazier, and J. C. Venter. (2007) The Sorcerer II Global Ocean Sampling Expedition: Expanding the Universe of Protein Families. *PLoS Biol*5: e16.

Yutin, N., K. S. Makarova, S. L. Mekhedov, Y. I. Wolf, and E. V. Koonin. (2008) The Deep Archaeal Roots of Eukaryotes. *Mol Biol Evol*25: 1,619—1,630.

Yutin, N., M. Y. Wolf, Y. I. Wolf, and E. V. Koonin. (2009) The Origins of Phagocytosis and Eukaryogenesis. *Biol Direct*4: 9.

Zhang, Y., C. Laing, M. Steele, K. Ziebell, R. Johnson, A. K. Benson, E. Taboada, and V. P. Gannon. (2007) Genome Evolution in Major Escherichia Coli O157:H7 Lineages. *BMC Genomics*8: 121.

Zuckerkandl, E., and L. Pauling. (1965). Evolutionary Divergence and Convergence of Proteins in Bryson, V., and H. J. Vogel (eds). *Evolving Genes and Proteins*. New York: Academic Press, pp. 97—166.

Благодарности

Раздел благодарностей – неременная часть любой книги, и ни одна стоящая научная идея, гипотеза или обобщение не родится в одиноком уме. Наука процветает лишь в сообществе ученых, и написать нечто достойное чтения возможно, лишь будучи подключенным к этому сообществу. С одной стороны, сочинение раздела благодарностей должно быть легким делом, потому что он в той или иной степени стандартен и формален; с другой же, эта задача не проста, ибо в работу и стиль мышления автора внесли свой разнообразнейший вклад столь многие коллеги и друзья, пусть даже и не помогая непосредственным образом в работе над этой книги. Я постараюсь упомянуть их всех, не вдаваясь при этом в чрезмерные тонкости.

Эта книга не могла бы быть завершена, если бы не великодушная помощь и поддержка Юрия Вольфа, моего давнего сотрудника и соавтора нескольких работ, основополагающих для развиваемой здесь концепции. Юрий не только прочитал текст целиком, предложил важные поправки и внес ценные предложения, но и подготовил многие иллюстрации. Трое других близких и давних сотрудников, Валерьян Доля, Билл Мартин и Татьяна Сенкевич (моя жена), прочитали всю рукопись и сделали многочисленные полезные замечания. Я хотел бы особо поблагодарить нескольких коллег и друзей, которые, на разных этапах моей научной работы, были не только наставниками или сотрудниками, но и источником интеллектуального вдохновения, повлияли на мое становление как исследователя и без чьей поддержки, мудрых советов и критики у меня не было бы ни стимула, ни возможности даже начать эту книгу. Это Вадим Израйлевич Агол (мой учитель в области вирусологии), Александр Евгеньевич Горбаленя (старший коллега по первой работе в сравнительной геномике), Л. Аравинд, Пеер Борк, Андрей Гудков, Валерьян Доля, Джим Каррингтон, Алексей Кондрашов, Дэвид Липман, Билл Мартин, покойный Александр Александрович Нейфах (младший), Павел Певзнер, Дидье Рауль, Патрик Фортер и Константин Чумаков. Беседы с У. Фордом Дулитлом и Майклом Линчем о эволюции геномов, Аланом Драммондом и Клаусом Вильке о эволюции белков, Сергеем Масловым, Эриком Ван Нимвегеном и Михаэлем Лассигом о физико-статистическом подходе к эволюции и Александром Виленкиным о космологии были неопределимо важны для всего моего образа мышления.

Я признателен бывшим и нынешним членам моей исследовательской группы за существенный вклад в те исследования, что были рассмотрены в этой книге, и многие другие проекты, которых у меня не было возможности коснуться: Лакшу Айяру, Вивеку Анантараману, Соне Васудеван, Максиму Вольфу, Михаилу Гальперину, Николаю Гришину, И. Кингу Джордану, Георгию Кареву, Лирану Кармелю, Федору Кондрашову, Дэвиду Кристенсену, Александру Лобковскому, Радже Мазумдеру, Кире Макаровой, Давиду Манагадзе, Бену Мансу, Сергею Мехедову, Аркадию Мушегяну, Анастасии Никольской, Павлу Новичкову (также за помощь с рис. 7–2), Артему Новожилову, Марине Омельченко, Перу Пуиджбо (также за большую помощь с рисунками к главе 6), Алисе Реш, Игорю Рогозину, Александру Свердлову, Алексею Спиридонову, Николаю Спиридонову, Нобуто Такэути, Роману Татузову, Рональду Уокеру, Панаиотису Цапарасу, Диане Черниковой, Светлане Шабалиной, Наталье Ютиной и Итаи Янаи. С не меньшей признательностью благодарю за ценный вклад сотрудников по коллективной работе и гостей моей группы Стивена Альтшуля, Одеда Беджу, Стивена Белла, Стэна Брунса, Тал Даган, Майкла Дели, Эда Де Лонга, Вишву Диксита, Михаила Гельфанда, Мартина Гюйнена, Детлефа Ляйпе, Леонардо Мартиньо-Рамиреса, Бориса Миркина, Андрея Миронова, Бернарда Мосса, Армена Мулкиджаняна, Джона ван дер Оста, Люку Пеллегрини, Криса Понтинга, Давида Прангишвили, Тересу Пшитицкую, Михаила Ройтберга, Кена Радда,

Алексея Савченко, Ласло Секея, Шамяля Сюняева, Дэна Хартла, Миклоша Чюроша (также за помощь с рис. 7–8), Эву Шабарку, Фреда Антсона, Александра Якунина и Вэй Ян.

Глубоко благодарен Данте за то, что он был и остается моим лучшим другом.

Первые черновики этой книги были написаны на весьма стимулирующей зимней конференции 2010 года в Аспенском Физическом центре (штат Колорадо), и мне хотелось бы выразить центру мою признательность. Последние штрихи были внесены во время столь же замечательной программы по микробной и вирусной эволюции в Институте теоретической физики Кавли университета Калифорнии в Санта-Барбаре, и мне приятно принести благодарность институту.

В завершение, но не в последнюю очередь, приношу свою благодарность редактору Кирку Дженсену за столь своевременное приглашение издать эту книгу и за его деликатное подталкивание, а также всему коллективу редакции «Пирсон Эдукейшн» за их первоклассную профессиональную работу.

Об авторе

Евгений Викторович Кунин окончил кафедру вирусологии биологического факультета МГУ в 1978 году и в 1983 году защитил там же под руководством профессора Вадима Израйлевича Агола кандидатскую диссертацию по репликации вирусной РНК. С 1983 по 1988 год он работал младшим, а затем старшим научным сотрудником Института полиомиелита и вирусных энцефалитов АМН СССР, а с 1988 по 1991 год – старшим научным сотрудником и руководителем лаборатории молекулярной систематики микробов Института микробиологии АН СССР. Начиная с 1985 года работал в области компьютерного анализа геномов вирусов, а затем и других организмов. В конце 1991 года Е. В. Кунин переехал в США, где работал вначале на временной, а затем, по настоящее время, на постоянной позиции ведущего исследователя Национального центра биотехнологической информации в Национальном институте здоровья (Бетесда, штат Мериленд). С 1997 года – руководитель группы эволюционной геномики, которая исследует широкий круг проблем, связанных с эволюцией геномов и белков, эволюционной вирусологией и микробиологией, а в последние годы – общими вопросами теории эволюции. Евгений Кунин – автор около 650 опубликованных научных статей (h-индекс 152) и книги *Sequence-Evolution-Function: Computational Approaches in Comparative Genomics* (Springer, New York, 2002, в соавторстве с Михаилом Гальпериным). Он также является основателем и главным редактором журнала *Biology Direct*, в котором реализуется принцип открытого рецензирования: статьи публикуются вместе с рецензиями, за подписью рецензента.

notes

Перевод заглавия этого введения представил серьезные трудности. В английском оригинале было towards a postmodern synthesis. Это конечно же игра слов: с одной стороны, postmodern означает просто «после Modern Synthesis» (то, что в русской литературе обычно называется синтетической теорией эволюции, СТЭ), а с другой – «постмодернистский». Как это передать по-русски, совершенно не очевидно. Хуже того, этот нехитрый каламбур неоднократно обыгрывается в дальнейших главах. Никакого способа справиться с этой трудностью, кроме написания этого примечания, ни переводчикам, ни автору в голову не пришло (примеч. авт. к русскому изданию здесь и далее курсивом).

Во многом эти представления опираются на публикации крупнейшего современного эволюциониста Форда Дулитла, которые цитируются в соответствующих главах.

По иронии судьбы, *magnum opus* Ламарка «Философия зоологии» был напечатан в год рождения Дарвина.

Удивительно точное и глубокое описание непосредственного воздействия и публичного признания книги Дарвина можно найти в романе Джона Фаулза «Любовница французского лейтенанта».

Естественно, сам Дарвин не использовал термин «дарвинизм»; этот не очень удачный и едва ли дальновидный неологизм введен последователем и защитником Дарвина, Томасом Генри Гексли, в отзыве на «Происхождение...» (*Huxley, T. H.* 1860. Darwin on the origin of Species. Westminster Review: 541–570). Это слово имеет некоторый оттенок догматичности, если не псевдонаучности, по созвучию с другими хорошо известными «-измами», например марксизмом, или фрейдизмом, или даже лысенкоизмом (по-русски – лысенковщиной; о лысенковщине см. гл. 9). Естественно, что никто не говорит о ньютонизме или эйнштейнизме, а слово «менделизм» (обычно в составе сложных слов «менделизм-вейсманизм» или «менделизм-морганизм») употреблялось исключительно в негативном контексте антинаучно мыслящими последователями Лысенко в Советском Союзе. Тем не менее термин, предложенный Гексли, твердо прижился и даже привлекает своей емкостью. В этой книге я использую его исключительно для описания «первой синтетической теории эволюции», которая была разработана Дарвином в «Происхождении...», а затем усовершенствована в последующих работах Гексли, Уоллеса, Вейсмана, Геккеля и других ранних последователей Дарвина.

Выражение «неупрощаемая сложность» было придумано Майклом Бихи, одним из основных сторонников антиэволюционной гипотезы разумного замысла (РЗ), в его (печально) известной книге «Черный ящик Дарвина» (*Behe, M. J. 2006. Darwin's Black Box: The Biochemical Challenge to Evolution. New York: Free Press*). Для Бихи и других сторонников РЗ «неупрощаемость» сложных биологических структур является будто бы подтверждением (даже доказательством) неизбежности РЗ. Конечно же РЗ – это злостная чушь, но термин «неупрощаемая сложность» вполне выразителен, хотя эволюционные биологи предпочли бы говорить о «видимой» или «кажущейся» неупрощаемости сложных структур.

Посетитель музея Менделя в Брно имеет возможность рассматривать экземпляр немецкого перевода «Происхождения видов...», густо испещренный пометками Менделя. Автор был там уже после публикации английского оригинала этой книги и остался под сильным впечатлением.

Сэр Рональд Фишер был настоящим гением. (*Fisher Box, J.1978. R. A. Fisher: The Life of a Scientist. New York: Wiley.*) Он был фактическим основателем не только популяционной генетики, но и, во многом, современной статистики и ввел математическое определение информации задолго до Клода Шеннона. В эту книгу мы включили также другие примеры его замечательных научных предвидений. В то же время сэр Рональд посвятил большую часть своей карьеры делу евгеники, области исследования, которая в настоящее время рассматривается как псевдонаука и граничит с преступлением. Следует избегать суждения о великих умах даже относительно недавнего прошлого по сегодняшним меркам.

В русской литературе используется термин «синтетическая теория эволюции» (СТЭ), который и употребляется далее в этой книге.

Это уже сделано во многих различного уровня учебниках и монографиях, как обзорных, так и глубоко специализированных. Взвешенное, умеренно специализированное представление данной темы можно найти в издании *D. L. Hartland A. G. Clark*(2006). *Principles of Population Genetics*, Sunderland, MA: Sinauer Associates.

В принципе если адаптивный ландшафт строится для гена, то число измерений будет равняться количеству нуклеотидных сайтов. Взаимодействие между сайтами (эпистаз) уменьшает размерность.

Формально теорема Фишера не запрещает всенисходящее движение, потому что она касается только той части адаптивных изменений, которые обусловлены отбором. Фишер, однако, считал, что на практике большинство, если не все популяции слишком велики, чтобы то явление, которое Райт обозначил как дрейф, могло иметь какое-либо влияние на них. Это было предметом ожесточенных споров между Фишером и Райтом. Окончательным победителем оказался, конечно, Райт.

Эта знаменитая фраза Добржанского является названием его эссе, опубликованного в журнале «Американский учитель биологии» (*Dobzhansky, T.1973. Nothing in biology makes sense except in the light of evolution // The American Biology Teacher 35: 125–129*). В целом это поразительный текст. Хотя большая часть эссе удивительно ясно объясняет идею эволюции, заключительные абзацы посвящены выразительной пропаганде совместимости эволюции и христианства, и трудно избавиться от впечатления, что именно это и было главной целью автора. Добржанский, как истинный прихожанин Русской православной церкви, верил, что Бог осуществил свой план сотворения мира как развертывающийся сценарий эволюции жизни. Более того, Добржанский изобретательно клеймит отрицание эволюции как кощунство, так как данная позиция подразумевает, что Бог – мошенник, который намеренно вводит человечество в заблуждение, предоставляя многочисленные доказательства эволюции. Я подозреваю, что не каждый, кто цитирует это изречение Добржанского в дискуссиях о преподавании эволюции, действительно читал это эссе.

Иногда та же самая аббревиатура расшифровывается как Last Universal Common Ancestor (последний универсальный общий предок). Однако представляется полезным подчеркнуть, что речь идет именно об общем предке всех клеточных форм жизни. Мы вернемся к этой теме в главе об эволюции вирусов.

Сами по себе эти принципы описываются в элементарных учебниках, но рассматриваемый здесь информационно-теоретический подход уже не так тривиален, так что я счел необходимым явно перечислить эти принципы.

Этот принцип может быть подвергнут достаточно обоснованному сомнению в свете открытия разнообразных явлений эпигенетической наследственности (в гл. 9 мы коснемся этих явлений, хотя и слишком кратко). В данной формулировке вся эпигенетика спрятана в слове «косвенно», но автор отдает себе отчет в спорности такого подхода.

Чаргафф не оценил вовремя исключительной важности своего открытия, и тот факт, что двум высокомерным молодым людям, не знавшим химии, удалось открыть тайну жизни, которой он, эксперт-химик, не понял, исполнил Чаргаффа горечи до конца его долгой жизни, вдохновляя его едкие, чтобы не сказать язвительные, книги. (*Chargaff E. Heraclitean Fire: Sketches from a Life Before Nature. New York: Rockefeller University Press, 1978.*)

Строго говоря, это верно лишь для идеализированных систем репликаторов, размножающихся без ограничений, то есть экспоненциально. Именно такие системы рассматривал Дарвин вслед за Мальтусом. В более реалистических системах с ограниченным (параболическим) размножением наблюдается «выживание всех», а не только наиболее приспособленных, и для селекции и дрейфа необходимы дополнительные условия, такие как компартментализация (Szathmary E. The origin of replicators and reproducers. Philos Trans R Soc Lond B Biol Sci . 2006 Oct 29;361(1474):1761-76).

Не совсем ясно, кто первым сформулировал принцип, который я здесь называю ПОР. Говорится о том, что в 1930-х годах эти идеи высказывали выдающиеся русские генетики Николай Константинович Кольцов и Николай Владимирович Тимофеев-Ресовский (*Шноль С. Э. Герои, злодеи, конформисты российской науки. 2-е изд., дополненное. М.: Крон-пресс, 2001*), но я ничего не знаю об их формальной публикации. Эрвин Шредингер приближается к этой идее в знаменитой книге (*Шредингер Э. Что такое жизнь? Физический аспект живой клетки / Пер. с англ. 2-е изд. Ижевск: РХД, 2002*), но все же не приходит к конкретной формулировке. В англоязычной литературе Ричард Докинз четко формулирует идею в 1976 году в ставшем классикой «Эгоистичном гене» (*Докинз Р. Эгоистичный ген / Пер. с англ. М.: Мир, 1993*). С другой, абстрактной точки зрения, теория самовоспроизводящихся автоматов, не привязанных ни к какому физическому воплощению, была разработана великим математиком Джоном фон Нейманом (*Нейман Дж. фон. Теория самовоспроизводящихся автоматов / Пер. с англ. М.: Мир, 1971*).

Известная русская поговорка «Не до жиру, быть бы живу» выражает ту же идею упрощенно, но достаточно точно.

Здесь уместно сказать несколько слов о метафоре в биологии, тем более что метафора «эгоистичного гена» была отмечена как особенно часто вводящая в заблуждение (Ball, 2011). Вне всякого сомнения, использовать и в особенности интерпретировать эту метафору следует с великой осторожностью, и даже легкого намека на очеловечивание генного «эгоизма» следует избегать всеми силами. Тем не менее я думаю, что метафоры необходимы для развития науки, и покуда научные результаты представляются посредством естественного языка (а не только математических выражений), метафоры неизбежны. Кроме того, хорошая метафора – краткая, запоминающаяся и запечатлевшая важные общие тенденции в широком поле наблюдения – способна сильно подтолкнуть свежее мышление и исследование. Я, например, считаю «эгоистичный ген», «мусорную ДНК» и «адаптивный ландшафт» прекрасными метафорами.

Во времена, когда еще не было полных геномных последовательностей, предвидение динамического генома было, вероятно, лучше всего представлено в исчерпывающей монографии выдающегося российского генетика Романа Бениаминовича Хесина (*Хесин Р. Б. Непостоянство генома*. М.: Наука, 1984). Эта пророческая книга была издана незадолго до безвременной кончины Хесина и стала легендарной среди русских биологов. К сожалению, она, по-видимому, не оказала большого влияния за пределами России.

Если уж мы обратили внимание на причастность сэра Рональда Фишера к евгенике, было бы несправедливо не упомянуть, что еще один патриарх популяционной генетики, Дж. Б. С. Холдейн, был многолетним членом Британской коммунистической партии и, вероятно, из партийной лояльности в течение многих лет поддерживал лысенковскую лженауку (или, по крайней мере, оставлял за ней презумпцию научности). Холдейн был ученым не меньших масштабов, чем Фишер, одним из последних великих универсалов в истории науки (*Dronamraju K. R. Haldane and Modern Biology*. Johns Hopkins University Press, 1968). Ему принадлежат не только многочисленные важные результаты в математической генетике (в том числе теория генетического груза), но и большой вклад в область кинетики ферментативных реакций и – вероятно, важнее всего – множество удивительно пророческих идей о разнообразнейших предметах, которые можно найти в его книгах и статьях (мы вернемся к некоторым из них в гл. 10). Холдейном также написаны сотни блестящих популярных очерков обо всех аспектах науки, многие из которых опубликованы в коммунистической газете «Дейли уоркер». В 1950 году Холдейн вышел из коммунистической партии, осознав степень разрушений, причиненных советской генетике Лысенко и его шайкой. Холдейн может послужить ярким примером того, что даже величайшие ученые не должны рассматриваться в отрыве от исторического контекста.

Эмиль Цукеркандль указал мне, что в его совместных с Лайнусом Полингом ранних работах по молекулярной эволюции проводится явное различие между гомологами, эволюционировавшими вертикальным путем, и теми, которые эволюционировали путем дупликации, хотя не вводилось никаких специальных терминов для обозначения этих разных классов гомологов. У каждого открытия или концептуального прорыва свои предшественники.

«Сказки просто так» – прекрасный сборник детских сказок Редьярда Киплинга (*Киплинг Р. Сказки просто так / Пер. с англ. К. Чуковского и Л. Хавкиной. М.: Росмэн-Пресс, 2011*). В них происхождение некоторых особенностей морфологии животных, таких как хобот слона и панцирь броненосца, возводится к различным своеобразным случаям. Киплинг, кажется, уже осознавал ошибочность панадаптационизма, хотя его выводы не всегда совпадали с выводами Гулда и Левонтина.

Здесь стоит сказать несколько слов о Карле Вёзе, создателе молекулярной филогенетики и первооткрывателе архей, который скончался уже после публикации оригинала этой книги, 30 декабря 2012 года, на 85-м году жизни (Goldenfeld N., Pace NR. Retrospective. Carl R. Woese (1928–2012) // Science . 2013 Feb 8;339(6120):661). Конечно, и при его жизни у коллег не было никаких сомнений в том, что он был крупнейшим ученым, подлинным революционером в микробиологии и эволюционной биологии. Однако, по точному выражению Анны Ахматовой, «когда человек умирает, изменяются его портреты». По крайней мере, для автора этой книги теперь очевидно, что в последней трети XX века Вёзе не было равных среди биологов. Его вклад в науку исключительно велик и разнообразен, и конечно же ни в коей мере не сводится к открытию архей. В этой книге его имя упоминается чаще, чем любое другое: во-первых, в связи с его пионерскими исследованиями по эволюции генетического кода, выполненными еще в 60-х годах прошлого века, затем, разумеется, при обсуждении молекулярной филогении и трех доменов клеточной жизни и, наконец, в контексте его глубоких концептуальных работ рубежа двух тысячелетий.

Я охотно признаю свое пристрастие к вирусам. Будучи второкурсником МГУ, я выбрал кафедру вирусологии в качестве специализации. Этот выбор был отчасти продиктован посторонними соображениями, такими как очевидный интерес к реальной науке и либеральная атмосфера на этой кафедре, в отличие от некоторых других. Это было важно в то время, и не было ошибкой. Но более фундаментальным побуждением было мое увлечение разнообразием генетических механизмов и организации генома вирусов, что приводит к идее о том, что вирусы могут иметь непосредственное значение для понимания ранних этапов эволюции жизни. Я все еще думаю, что эта идея совершенно верна, как обсуждается в гл. 10 и 11. Все эксперименты, которые я когда-либо провел, относились к области вирусологии; эта работа, хоть сама по себе и несущественная, была чрезвычайно поучительна для всех моих последующих исследований в области вычислительной биологии. Вероятно, еще важнее, что мои первые вылазки в сравнительную геномику, совпавшие по времени с рождением всей этой области исследования, были связаны с вирусными геномами. Эти небольшие геномы были идеальной стартовой площадкой: даже с тогдашними примитивными средствами вычислительной техники (и, конечно, со всем рвением новичка) в общем-то можно было проследить эволюцию каждой аминокислоты в вирусных белках.

Много хуже, чем Фишер или Холдейн, Мережковский публично выступал с чрезвычайно отвратительными взглядами явно фашистского толка. Тем не менее его работы по эндосимбиозу являют собою поразительный пример образцового исследования и по сей день (*W. Martinand K. V. Kowallik. Annotated English Translation of Mereschkowsky's 1905 Paper 'Über Natur und Ursprung der Chromatophoren im Pflanzenreiche' // European Journal of Psychology 34 [1999]: 287–296.*)

Обратите здесь внимание на сходство с адаптивным ландшафтом Райта. Насколько я знаю, это две независимые сошедшиеся вместе идеи.

Это нехитрое предсказание, разумеется, подтверждается: экспоненциальный, а возможно, даже более быстрый рост числа геномных последовательностей продолжается.

По-видимому, термин «геном» был впервые использован немецким ботаником Гансом Винклером в 1920 г. (*M. Ridley. Genome. New York: Harper Perennial, 2006*).

Типичный элемент американских квартир и домов, буквально поразивший воображение автора во время первого визита в США в 1991 году, а теперь, кажется, постепенно выходящий из моды.

Когда, сравнивая первые секвенированные бактериальные геномы, мы с Аркадием Мушегяном обнаружили, что порядок генов сохраняется столь слабо, это настолько нас поразило, что мы назвали краткую статью, описывающую это наблюдение, «Бактериальная эволюция не сохраняет порядка генов» (A. R. Mushegian and E. V. Koonin. Gene Order Is Not Conserved in Bacterial Evolution // *Trends in Genetics* 12 (1996a): 289–290). С точностью фактов, описанных в статье, кажется, все в порядке, но, если бы мне пришлось опубликовать ее сегодня, я бы постарался отыскать для нее более точное и осторожное название. Однако то название, которое мы дали статье, зафиксировало наше удивление контрастом между сохранностью генных последовательностей и изменчивостью порядка генов.

Разумеется, если биосфера выживает.

Статья, описывающая геном *M. genitalium*, называется «Минимальный набор генов бактерии *Mycoplasma genitalium*» (C. M. Fraser, J. D. Gocayne, O. White, M. D. Adams, R. A. Clayton, R. D. Fleischmann, C. J. Bult, A. R. Kerlavage, G. Sutton, J. M. Kelley et al. The Minimal Gene Complement of *Mycoplasma genitalium* // Science 270 (1995): 397–403). Однако, хоть и имея столь малое число генов, этот организм весьма специализирован и потому не может полагаться действительно «минимальным».

Уже после публикации оригинала этой книги такая модель в первом приближении была разработана (Lobkovsky AE, Wolf YI, Koonin EV. Gene frequency distributions reject a neutral model of genome evolution // Genome Biol Evol. 2013;5(1):233-42). Эта модель указывает на необходимость неравномерного распределения силы отбора, действующей на отдельные гены, для формирования характерной трехкомпонентной структуры геномной вселенной.

Эволюция выбора кодонов – сложная и интересная проблема, которую мы здесь не можем рассмотреть сколько-нибудь детально (см.: Plotkin JB, Kudla G. Synonymous but not the same: the causes and consequences of codon bias Nat Rev Genet. 2011 Jan;12(1):32–42). Отметим лишь, что, как правило, оптимальный кодон соответствует наиболее высоко экспрессируемой мРНК для данной кодонной серии.

В простейших моделях процесс укладки белка имитируется посредством подгонки модельной последовательности на ортогональной решетке. В более реалистичных (или, вернее, менее нереалистичных) моделях от решетки отказываются в пользу свободной укладки модельной полимерной цепи. Эта техника более приближена к реальному процессу укладки белка, но и требует больше вычислительных ресурсов. По ряду соображений точность решеточной модели показалась слишком низкой для обсуждаемого здесь вопроса, поэтому мы использовали безрешеточную модель.

Это предположение хорошо согласуется с результатами более ранней работы того же Джорджа Занга, в которой показано, что «мгновенная» скорость эволюции (то есть скорость, измеренная путем сравнения ортологов из близких видов) гораздо лучше коррелирует с эффектом нокаута (важностью гена), чем скорость, измеренная на более протяженных эволюционных интервалах (Zhang J, He X. Significant impact of protein dispensability on the instantaneous rate of protein evolution. *Mol Biol Evol.* 2005 Apr;22(4):1147-55). С другой стороны, конечно же не следует забывать, что предложенная гипотеза имеет статистический характер и не относится к группе наиболее важных генов (например, гены рибосомных белков остаются незаменимыми на протяжении всей эволюционной истории).

Возможно также рождение нового гена из некодирующей последовательности, особенно в больших геномах многоклеточных организмов.

Многие читатели вспомнят теорию шести рукопожатий. Еще более знакомая иллюстрация встречается в журналах, бесплатно предоставляемых во время рейса большинством авиалиний. В следующий раз во время перелета взгляните на неизменно печатаемую на тыльной стороне журнала сеть авиатрасс – это прекрасный пример масштабно-инвариантной сети, с авиационными хабами в Атланте, Чикаго или Денвере. И конечно, Интернет – это тоже масштабно-инвариантная сеть. Увлекательное, и достаточно точное теоретически, рассмотрение сетей во всех сферах жизни можно найти в популярных книгах Альберта-Ласло Барабаши, одного из пионеров сетевой биологии (*Linked. The New Science of Networks*. New York: Perseus Press, 2002), и Дункана Ваттса (*Six Degrees: The Science of a Connected Age*. New York: W.W. Norton & Co., 2004).

Точнее, конечно, будет сказать, что каждый ген вносит вклад в разные биологические функции, поскольку все гены обладают той или иной степенью плейотропии.

У каждого важного открытия есть свои предшественники – нужно только повнимательнее поискать. Я думаю, это один из «универсальных законов» истории науки. Похоже на то, что «закон Ван Нимвегена» был впервые описан в статье о геноме синегнойной палочки, но без акцента на нем и без сколько-нибудь серьезного анализа (*C. K. Stover, X. Q. Pham, A. L. Erwin, S. D. Mizoguchi, P. Warrenner, M. J. Hickey, F. S. Brinkman, W. O. Hufnagle, D. J. Kowalik, M. Lagrou, R. L. Garber, L. Goltry, E. Tolentino, S. Westbrook-Wadman, Y. Yuan, L. L. Brody, S. N. Coulter, K. R. Folger, A. Kas, K. Larbig, R. Lim, K. Smith, D. Spencer, G. K. Wong, Z. Wu, I. T. Paulsen, J. Reizer, M. H. Saier, R. E. Hancock, S. Lory, and M. V. Olson. Complete Genome Sequence of Pseudomonas Aeruginosa PAO1, an Opportunistic Pathogen // Nature 406 (2000): 959–964.*

По-видимому, многие из этих универсалий, во всяком случае те, которые математически представляются характерными функциями распределения, можно вывести из единого, чрезвычайно общего физического принципа, известного как принцип максимальной энтропии или принцип минимального производства информации (Jaynes, T. Probability Theory: The Logic of Science. Cambridge Univ Press, 2003). Вкратце и упрощая, можно сформулировать этот принцип так: любая макроскопическая величина, определяемая совокупностью многочисленных микроскопических процессов, принимает, с учетом необходимых ограничений, распределение с максимальной энтропией (то есть такое распределение, для получения которого требуется минимальная информация). Легко понять, что принцип максимальной энтропии тесно связан со вторым началом термодинамики. Применение этого принципа к биологической эволюции – глубокая и интересная тема, которую здесь, конечно, не раскрыть. Чрезвычайно содержательное обсуждение этих вопросов можно найти в статье Стивена Франка, с которой автор, к сожалению, познакомился уже после публикации оригинала этой книги (Frank SA. The common patterns of nature // J Evol Biol. 2009 Aug;22(8):1563-85).

Здесь надо вновь сделать упор на принцип максимальной энтропии.

К этому списку хочется добавить уже упомянутую статью Франка (Frank SA. The common patterns of nature. *J Evol Biol.* 2009 Aug;22(8):1563-85), которая представляет собой достаточно простое, удивительно четкое и интересное обсуждение универсальных распределений и принципа максимальной энтропии, а также статью автора этой книги, в которой обсуждается природа законов эволюционной геномики (Koonin EV. Are there laws of genome evolution? // *PLoS Comput Biol.* 2011 Aug;7(8):e1002173).

В действительности, хотя Дарвин не обсуждал микробов в печати, некоторые из его писем показывают значительный интерес к этой теме и ее понимание (*M. A. O'Malley, What Did Darwin Say About Microbes, and How Did Microbiology Respond? Trends in Microbiology*17 [2009]: 341–347).

По всей вероятности, это предвзятый взгляд, но для меня волнение, вызванное прочтением этих первых геномов, буквально ни с чем не сравнимо.

Бимодальное распределение какой-либо величины редко встречается в природе, его появление указывает, что наблюдается нечто интересное.

Здесь неперево́димая игра слов: по-английски такие гены называют ORFans. Это обозначение происходит, с одной стороны, от ORF, Open Reading Frame (открытая рамка считывания), а с другой стороны, от Orphans (сироты – то есть гены, у которых нет известных родственников).

Так в оригинале. Теперь автору представляется, что здесь было бы точнее сказать «еще не так давно были».

Первая публикация геномной последовательности гипертермофильной бактерии *Aquifex aeolicus* не смогла выявить дополнительные «архейные» гены, и в действительности даже однозначно заявляла об их отсутствии (*G. Deckert, P. V. Warren, T. Gaasterland, W. G. Young, A. L. Lenox, D. E. Graham, R. Overbeek, M. A. Snead, M. Keller, M. Aujay, R. Huber, R. A. Feldman, J. M. Short, G. J. Olsen, and R. V. Swanson*. The Complete Genome of the Hyperthermophilic Bacterium *Aquifex Aeolicus* // *Nature* 392 [1998]: 353–358). Идея о возможном обмене генами между археями и бактериями гипертермофилами определенно приходила авторам в голову. Единственная причина, по которой он не был обнаружен, заключалась в том, что геном *Aquifex* довольно маленький, но авторы не нормализовали число «архейных» генов к общему числу генов (или размеру генома). Как только подобная нормализация была сделана, она сразу же выявила поразительное преобладание «архейных» генов в сравнении с геномами мезофильных бактерий (*L. Aravind, R. L. Tatusov, Y. I. Wolf, D. R. Walker and E. V. Koonin*. Evidence for Massive Gene Exchange Between Archaeal and Bacterial Hyperthermophiles // *Trends in Genetics* 14 [1998]: 442–444).

А также и для разных функциональных групп генов.

Это представление несколько упрощенное, так как существенную роль играет и гомологичная рекомбинация. Однако основной вывод о необходимости замены мутантных генов интактными копиями не подлежит сомнению.

Это означает, что у прокариот трансляция начинается до завершения транскрипции, непосредственно на растущей молекуле мРНК. У эукариот такое сопряжение невозможно, так как транскрипция и трансляция пространственно разобщены ядерной мембраной (см. гл. 7). Любопытно, что в нашей заметке на эту тему мы с Биллом Мартином несколько поторопились, заявив, что сопряжение транскрипции и трансляции является определяющим признаком прокариот. Как выяснилось, в то время соответствующие экспериментальные данные об археях отсутствовали. Вообще говоря, такая неаккуратность в научной публикации никак не заслуживает поощрения. В этот раз, однако, все закончилось хорошо, поскольку наша заметка стимулировала эксперименты, которые убедительно продемонстрировали существование сопряжения у архей (French SL, Santangelo TJ, Beyer AL, Reeve JN. //Transcription and translation are coupled in Archaea. Mol Biol Evol. 2007 Apr;24(4):893-5).

К этому выводу следует относиться с долей скепсиса: вполне вероятно, мы недооцениваем частоту и значение такого использования мобильных элементов.

Мы здесь пользуемся терминами «листья» и «вершины» в математическом, а не в биологическом смысле.

Эта глава преднамеренно кратка, и в ней без лишних отступлений описываются последние исследования, которые, на мой взгляд, отражают текущее состояние проблемы ДЖ. Интересный исторический обзор можно найти в книге Саппа, также включающей в себя обсуждение современного состояния «древесного мышления» (*J. Sapp, The New Foundations of Evolution: On the Tree of Life Oxford. Oxford University Press, 2009*). Еще более подробную панораму современного взгляда на ДЖ можно найти в двух сериях статей, написанных биологами совместно с философами (*O'Malley, M. A., ed. (2010) Special Issue: The Tree of Life. Biology and Philosophy; Ragan, M. A., J. O. McInerney, and J. A. Lake, eds. (2009) Theme Issue: The Network of Life: Genome Beginnings and Evolution. Phil Trans R Soc B*).

*В оригинале NUTs (Nearly Universal Trees), весьма многозначное сокращение...
Предоставляем читателям самим в этом разобраться.*

Очевидная аналогия с космологической моделью Большого взрыва – моделью начала нашей Вселенной. В главе 12 и приложении II мы касаемся интерпретации Большого взрыва в современных космологических моделях.

Подчеркивая, таким образом, роль эндосимбиоза как эволюционного события, приводящего к возникновению принципиально новых групп организмов.

С тех пор как была написана эта глава, такое экспериментальное исследование было опубликовано – формирование филаментов архейными гомологами актина полностью подтвердилось (Ettema TJ, Lindes AC, Bernander R. An actin-based cytoskeleton in archaea // Mol Microbiol. 2011 May;80(4):1052-61).

То есть сплайсинг при помощи обратной транскриптазы, кодируемой другим интроном.

У архей этот белок участвует в процессинге малых РНК.

Уже после публикации английского оригинала этой книги на страницах солидных журналов разгорелись довольно жаркие дебаты о природе сходства внутриклеточной организации «ядерных» бактерий и эукариот (Devos DP, Reynaud EG. Evolution. Intermediate steps. Science. 2010 Nov 26;330(6008):1187-8; McInerney JO, Martin WF, Koonin EV, Allen JF, Galperin MY, Lane N, Archibald JM, Embley TM. Planctomycetes and eukaryotes: a case of analogy not homology // Bioessays. 2011 Nov;33(11):810-7; Reynaud EG, Devos DP. Transitional forms between the three domains of life and evolutionary implications // Proc Biol Sci. 2011 Nov 22;278(1723):3321-8; Devos DP. Regarding the presence of membrane coat proteins in bacteria: confusion? What confusion? Bioessays. 2012 Jan;34(1):38-9). Во второй из цитируемых статей мы с коллегами приводим ряд конкретных аргументов в пользу того, что это сходство представляет собой аналогию (то есть результат независимой эволюции внешне сходных черт), а никак не гомологию (то есть происхождение от общего предка). Мне эти аргументы и теперь кажутся вполне убедительными.

Лауреат Нобелевской премии по химии (вместе с Фредом Сэнгером) за разработку одного из двух первых эффективных методов секвенирования ДНК.

Сокращенные наименования видов: *Aureococcus anophagefferens* (Aano), *Aedes aegypti* (Aaeg), *Agaricus bisporus* (Abis), *Anopheles gambiae* (Agam), *Allomyces macrogynus* ATCC 38327 (Amac), *Apis mellifera* (Amel), *Aspergillus nidulans* FGSC A4 (Anid), *Acyrtosiphon pisum* (Apis), *Arabidopsis thaliana* (Atha), *Babesia bovis* (Bbov), *Batrachochytrium dendrobatidis* (Bden), *Branchiostoma floridae* (Bflo), *Botryotinia fuckeliana* B05.10 (Bfuc), *Brugia malayi* (Bmal), *Bombyx mori* (Bmor), *Coccomyxa* sp. C-169 (C169), *Chlorella* sp. NC64a (C64a), *Caenorhabditis briggsae* (Cbr), *Caenorhabditis elegans* (Cele), *Coprinopsis cinerea okayama7#130* (Ccin), *Cochliobolus heterostrophus* C5 (Chet), *Coccidioides immitis* RS (Cimm), *Ciona intestinalis* (Cint), *Cryptococcus neoformans* var. *neoformans* (Cneo), *Chlamydomonas reinhardtii* (Crei), *Capitella teleta* (Ctel), *Capsaspora owczarzaki* ATCC 30864 (Cowc), *Dictyostelium discoideum* (Ddis), *Dictyostelium purpureum* (Dpur), *Drosophila melanogaster* (Dmel), *Drosophila mojavensis* (Dmoj), *Daphnia pulex* (Dpul), *Danio rerio* (Drer), *Entamoeba dispar* (Edis), *Entamoeba histolytica* (Ehis), *Emiliana huxleyi* (Ehux), *Fragilariopsis cylindrus* (Fcyl), *Phanerochaete chrysosporium* (Fchr), *Phaeodactylum tricornutum* (Ftri), *Gallus gallus* (Ggal), *Gibberella zeae* PH-1 (Gzea), *Hydra magnipapillata* (Hmag), *Helobdella robusta* (Hrob), *Homo sapiens* (Hsap), *Ixodes scapularis* (Isca), *Laccaria bicolor* (Lbic), *Lottia gigantea* (Lgig), *Micromonas* sp. RCC299 (M299), *Monosiga brevicollis* (Mbre), *Mucor circinelloides* (Mcir), *Mycosphaerella fijiensis* (Mfij), *Mycosphaerella graminicola* (Mgra), *Magnaporthe grisea* 70-15 (Mgri), *Melampsora laricis-populina* (Mlar), *Micromonas pusilla* CCMP1545 (Mpus), *Neurospora crassa* OR74A (Ncra), *Nematostella vectensis* (Nvec), *Nasonia vitripennis* (Nvit), *Ostreococcus* sp. RCC809 (O809), *Ostreococcus lucimarinus* (Oluc), *Oryza sativa japonica* (Osat), *Ostreococcus taurii* (Otau), *Phytophthora capsici* (Pcap), *Plasmodium falciparum* (Pfal), *Puccinia graminis* (Pgra), *Pediculus humanus* (Phum), *Phaeosphaeria nodorum* SN15 (Pnod), *Physcomitrella patens* subsp. *patens* (Ppat), *Phytophthora ramorum* (Pram), *Pyrenophora tritici-repentis* Pt-1C-BFP (Prep), *Proterospongia* sp. ATCC 50818 (Prsp), *Phytophthora sojae* (Psoj), *Paramecium tetraurelia* (Ptet), *Plasmodium vivax* (Pviv), *Plasmodium yoelii yoelii* (Pyoe), *Rhizopus oryzae* (Rory), *Sorghum bicolor* (Sbic), *Saccharomyces cerevisiae* (Scer), *Schizosaccharomyces japonicus* yFS275 (Sjap), *Schistosoma mansoni* (Smans), *Selaginella moellendorffii* (Smoe), *Schizosaccharomyces pombe* (Spom), *Spizellomyces punctatus* DAOM BR1173 (Spun), *Strongylocentrotus purpuratus* (Spur), *Sporobolomyces roseus* (Sros), *Sclerotinia sclerotiorum* 1980 UF-70 (Sscl), *Trichoplax adhaerens* (Tadh), *Theileria annulata* (Tann), *Tribolium castaneum* (Tcas), *Toxoplasma gondii* (Tgon), *Taenopygia guttata* (Tgut), *Theileria parvum* (Tpar), *Thalassiosira pseudonana* (Tpse), *Tetrahymena thermophila* (Tthe), *Ustilago maydis* 521 (Umay), *Uncinocarpus reesii* 1704 (Uree), *Volvox carteri* (Vcar), *Vitis vinifera* (Vvin).

По различным аспектам сложности опубликовано огромное число книг, популярных и не очень, в том числе за авторством великого физика и автора теории кварков Мюррея Гелл-Манна (*M. Gell-Mann. The Quark and the Jaguar: Adventures in the Simple and the Complex. New York: St. Martin's Griffin, 1995*). Более специальная книга Стюарта Кауфмана предлагает множество оригинальных идей об эволюции сложности (*S. Kauffman. At Home in the Universe: The Search for the Laws of Self-Organization and Complexity. Oxford: Oxford University Press, 1996*). Еще один, современный краткий текст-введение – *N. Johnson. Simply Complexity: A Clear Guide to Complexity Theory. New York: Oneworld Publications, 2009*.

На всякий случай – это не просто фривольная шутка, а отсылка к знаменитой фразе члена Верховного суда США Поттера Стюарта по поводу того, как распознать порнографию: «Когда я это вижу, я это узнаю».

Наиболее глубоко этот вопрос рассматривается, по-видимому, в работах Дэна Макши, например: D.W. McShea. Functional complexity in organisms: parts as proxies. Biology and Philosophy 15 (2000): 641–668, и в недавней интересной книге: Daniel W. McShea and Robert N. Brandon (2010). Biology's First Law: The tendency for Diversity and Complexity to Increase in Evolutionary Systems. University of Chicago Press: Chicago.

Информация эквивалентна колмогоровской сложности только для строго случайных последовательностей с заданными частотами символов. Геномные последовательности, как правило, не таковы – они содержат зависимости между нуклеотидами в разных положениях. Несмотря на интуитивную понятность концепции колмогоровской сложности, не существует общей формулы для ее вычисления.

Эти вероятности не просто являются частотами, но теоретически должны задаваться непредвзятыми статистическими моделями для отдельных сайтов, которые, хоть и не известны в явной форме, аппроксимируются различными математическими моделями, точность которых растет с увеличением числа последовательностей в выравнивании.

Это, разумеется, статистическое утверждение. Невозможно исключить, что какая-то небольшая часть интронов выполняет важные функции и сохраняется в ходе эволюции именно по этой причине.

Макши и Брэндон в своей недавней книге формулируют «закон эволюции в условиях нулевой силы» (по аналогии с первым законом Ньютона), согласно которому эволюция популяции или экосистемы в отсутствие ограничивающего давления отбора ведет к увеличению сложности по чисто стохастическим причинам (Daniel W. McShea and Robert N. Brandon (2010). Biology's First Law: The tendency for Diversity and Complexity to Increase in Evolutionary Systems. University of Chicago Press: Chicago). По крайней мере формально сформулированное нами условие возникновения сложности согласуется с этим законом: сложность возникает только в условиях слабого, пусть и не нулевого, давления отбора.

Вероятно, не менее важна альтернативная транскрипция, которая также приводит к формированию разных белковых продуктов одного и того же гена.

Мы с коллегами недавно построили вероятностную модель, которая позволила экстраполировать имеющиеся данные и рассчитать, что геномы млекопитающих, скорее всего, кодируют примерно вдвое больше длинных некодирующих РНК, чем белков (Managadze D, Lobkovsky AE, Wolf YI, Shabalina SA, Rogozin IB, Koonin EV. The Vast, Conserved Mammalian lincRNome. PLoS Comput Biol. 2013 Feb;9(2):e1002917).

Вспомним здесь снова «закон нулевой силы», предложенный Макши и Брэндоном, согласно которому в отсутствие ограничений, накладываемых селекцией, эволюция идет именно в сторону усложнения (Daniel W. McShea and Robert N. Brandon (2010). *Biology's First Law: The tendency for Diversity and Complexity to Increase in Evolutionary Systems*. University of Chicago Press: Chicago). Справедливость этого «закона» сомнительна, учитывая закономерность, четко установленную для самых различных организмов, а именно высокую частоту делеций по сравнению со вставками в ходе эволюционного процесса (Kuo CH, Ochman H. Deletional bias across the three domains of life. *Genome Biol Evol.* 2009 Jun 27;1:145-52). Однако, если бы даже «закон нулевой силы» и выполнялся, приведенные здесь данные всего лишь означали бы, что в реальной эволюции ограничения перевешивают «нулевую» тенденцию.

Аналогичная реконструкция на основе более полной информации привела к еще более впечатляющим результатам, показывая, что общий предок архей, вероятно, имел даже более сложный геном, чем большинство современных свободноживущих форм (Wolf YI, Makarova KS, Yutin N, Koonin EV. Updated clusters of orthologous genes for Archaea: a complex ancestor of the Archaea and the byways of horizontal gene transfer. Biol Direct. 2012 Dec 14;7:46).

То есть задействования геномного мусора для выполнения биологических функций.

К моменту публикации русского перевода этой книги уже будет опубликована наша новая статья с Ю. И. Вульфом, в которой эта модель рассматривается более подробно, с использованием дополнительных данных (Wolf, Y. I., Koonin, E. V. (2013). Genome reduction as the dominant mode of evolution. BioEssays , in press).

Известный естественный теолог XVIII века Уильям Палей весьма ясно обрисовал проблему, разумно утверждая, что найденные во время прогулки по полям часы подразумевают существование часовщика. Этот ход мысли привел одновременно к современному движению «разумного дизайна» и знаменитому возражению Ричарда Докинза (*R. Dawkins. The Blind Watchmaker: Why the Evidence of Evolution Reveals a Universe without Design. London: W.W. Norton & Co., 1996*).

См. также недавнее обсуждение современного состояния концепции КНЭ: Stoltzfus A. Constructive neutral evolution: exploring evolutionary theory's curious disconnect. Biol Direct. 2012 Oct 13;7:35.

Эта метафора широко используется в математических исследованиях случайных процессов.

Не сводится ли к этой нехитрой идее «закон нулевой силы», предложенный Макши и Брэнденом и неоднократно упоминаемый выше?

Речь не идет ни о каком нелепом антропоморфном восприятии прогресса – имеется в виду лишь простая интуиция, приравнивающая прогресс к постепенно увеличивающейся сложности.

В 1880 году знаменитый немецкий биолог Август Вейсман, в контексте своей теории зародышевой плазмы и изоляции зародышевой линии от сомы, взялся непосредственно опровергнуть наследование приобретенных признаков в серии опытов, ставших не менее знаменитыми, чем жираф, который, согласно Ламарку, приобрел длинную шею, поскольку его предки из поколения в поколение тянулись к листьям в кроне высоких деревьев (*A. Weissmann. The Germ-Plasm. A Theory of Heredity. London: Charles Scribner's Sons, 1893*). Едва ли стоит упоминания, что у подопытных крыс с отрезанным Вейсманом хвостом не появилось не только бесхвостых крысят, но и вообще потомства со сколь-нибудь укороченным хвостом. Эксперименты Вейсмана нанесли серьезный удар по общественному восприятию идеи о наследовании приобретенных признаков, хотя, строго говоря, они не подпадают под концепцию Ламарка, которая, как уже говорилось, описывает наследование положительных изменений, в первую очередь вызванных постоянным использованием органа, но не требует наследования бессмысленных изменений. Ненаследование произвольных изменений, например обрезания у человека, было известно задолго до Вейсмана; утверждения о противном тем не менее были достаточно обычны во времена Вейсмана и, видимо, послужили непосредственным толчком к его опытам.

Вдохновленный идеями прогресса в биологической эволюции, яркий венский исследователь и популяризатор науки Пол Каммерер развернул в начале XX века серию исследований длиною в два десятилетия, чтобы продемонстрировать наследование приобретенных признаков (*S. Gliboff*. Protoplasm... Is Soft Wax in Our Hands?: Paul Kammerer and the Art of Biological Transformation. *Endeavour* 29 [2005]: 162–167; *E. Pennisi*. History of Science. The Case of the Midwife Toad: Fraud or Epigenetics? *Science* 325 [2009]: 1,194—1,195; *A. O. Vargas*. Did Paul Kammerer Discover Epigenetic Inheritance? A Modern Look at the Controversial Midwife Toad Experiments. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 312 [2009]: 667–678). Каммерер в основном экспериментировал с амфибиями, которые меняли цвет и брачные повадки в зависимости от факторов окружающей среды, таких как температура и влажность. Он настаивал, как ни удивительно, на том, что индуцированные изменения, которые он наблюдал, были полностью наследственными. опыты Каммерера вызывали критику в связи с небрежностью документации и подозрительными, вероятно, фальсифицированными рисунками и фотографиями. Каммерер активно защищал свои выводы, но в 1923 году его карьера подошла к печальному концу после того, как известный британский генетик Уильям Бейтсон обнаружил, что брачные подушечки на ногах образцовой жабы-повитухи Каммерера, признак якобы передаваемый по наследству, на самом деле были инъецированы черными чернилами. Каммерер покончил с собой через два года после этого позорного разоблачения. Остается неясным, был ли Каммерер мошенником в самом худшем смысле этого слова; считается, что он мог использовать чернила, чтобы «усилить» действительно наблюдавшееся изменение цвета, – научная практика, не одобряемая уже в то время, не говоря о нашем, но все же далекая от вопиющего обмана. Результаты Каммерера могли объясняться как наличием у подопытных животных скрытой вариативности признаков, которая незаметно для него самого стала предметом отбора, так и возможной эпигенетической наследственностью. Если мы примем наиболее благоприятное для Каммерера объяснение, его опытная техника была чрезвычайно небрежной, даже если он и обнаружил не осознанные им важные явления. Что бы там ни произошло на самом деле, преданный широкой огласке скандал с Каммерером вряд ли поднял репутацию ламарковской теории наследования. Самое худшее для Ламарка было еще, впрочем, впереди.

Следующее примечание наверняка покажется многим читателям русского перевода этой книги примитивным. Однако оно написано для англоязычных читателей, многие из которых совсем или почти совсем не знакомы с этой трагической страницей истории.

Жестокая ирония ситуации состояла в том, что большевистские вожди Советского Союза горячо приветствовали Каммерера и едва не открыли ему в СССР лабораторию (S. Gliboff. *Protoplasm... Is Soft Wax in Our Hands: Paul Kammerer and the Art of Biological Transformation. Endeavour* 29 [2005]: 162–167). Несмотря на поразительные успехи русской генетики в 1920-х годах (вспомним имена Сергея Четверикова и Николая Вавилова), лидеры партии лелеяли идеи быстрого, планового и простого улучшения природы, в том числе и природы человека. Когда общая ситуация в стране стала тяготеть к массовому террору и голоду в 1930 году, подходящая команда была найдена под руководством агронома Трофима Лысенко. Лысенко и его подручные были вообще не учеными, а совершенно бесстыдными преступниками, использовавшими ненормальную ситуацию в стране, чтобы получить власть над советскими научными кругами и за их пределами (V. N. Soyfer. *The Consequences of Political Dictatorship for Russian Science. Nature Reviews Genetics* 2 [2001]: 723–729; V. N. Soyfer. *Lysenko and the Tragedy of Soviet Science. New Brunswick, NJ: Rutgers University Press, 1994*). Ламарковское наследование, которое лысенковцами, не без порочного хитроумия (с отчетливым оруэлловским оттенком), объявлялось «истинным дарвиновским» механизмом эволюции, было краеугольным камнем их «теории». Они довели идею Ламарка до гротескных крайностей, утверждая, например, что кукушки неоднократно возникали заново из яиц мелких птиц, и приводя это как пример особо примечательной адаптации. В последние годы, будучи уже отстранен от власти, Лысенко сохранил экспериментальную лабораторию, где он, как сообщается, кормил коров сливочным маслом и шоколадом в попытке вывести породы, стабильно дающие жирное молоко. Главным образом лысенковская «наука истинного дарвинизма» не была даже мошенничеством, поскольку ее адепты отнюдь не утруждали себя подделкой каких бы то ни было «экспериментов», а просто рассказывали свои идеологически вдохновленные сказки. Все это было бы смешно, если бы не тот факт, что многие несогласные заплатили буквально своей жизнью, и почти все исследования в области биологии в Советском Союзе были затруднены в течение десятилетий. Нет никаких оснований далее обсуждать здесь Лысенко; о лысенковщине было опубликовано несколько подробных работ (Zh. A. Medvedev. *The Rise and Fall of T. D. Lysenko. New York: Columbia University Press, 1969*; V. N. Soyfer. *The Consequences of Political Dictatorship for Russian Science. Nature Reviews Genetics* 2 [2001]: 723–729; V. N. Soyfer. *Lysenko and the Tragedy of Soviet Science. New Brunswick, NJ: Rutgers University Press, 1994*). Материалы печально известной сессии ВАСХНИЛ 1948 года, официально покончившей с генетикой, – документ одновременно потрясающе интересный и душераздирающий (О положении в биологической науке, стенограмма сессии VI. Всесоюзная академия сельскохозяйственных наук, 31 июля по 7 августа. М.: Государственное издательство сельскохозяйственной литературы, 1948).

Как и почти любой запрет в биологии, центральная догма может оказаться не столь уж и незыблемой. Прямых путей передачи информации от белка к нуклеиновой кислоте, по-видимому, и правда нет, но способы обойти запрет могут существовать (Koonin EV. Does the central dogma still stand? Biol Direct. 2012 Aug 23;7:27). Мы еще вернемся к этой проблеме ниже.

История открытия *cas*-генов интересна и поучительна сама по себе, хоть и выходит за рамки основной темы этой книги. В нашем исследовании перекрывающихся цепочек генов в геномах прокариот, проделанном в 2002 году (см. гл. 5), группа *cas*-генов оказалась второй по величине связной геномной окрестностью после рибосомного супероперона (I. B. Rogozin, K. S. Makarova, J. Murvai, E. Czabarka, Y. I. Wolf, R. L. Tatusov, L. A. Szekely, and E. V. Koonin. Connected Gene Neighborhoods in Prokaryotic Genomes. *Nucleic Acids Research*30 [2002]: 2,212—2,223). После тщательного анализа последовательностей Cas-белков мы предсказали, что эти белки представляют собой неизвестную ранее систему репарации ДНК (K. S. Makarova, L. Aravind, N. V. Grishin, I. B. Rogozin, and E. V. Koonin. A DNA Repair System Specific for Thermophilic Archaea and Bacteria Predicted by Genomic Context Analysis. *Nucleic Acids Research*30 [2002]: 482–496). Такой прогноз, казалось, имеет смысл, если вспомнить различные роли нуклеаз, геликаз и полимераз в репарации. К сожалению, мы не исследовали соседних повторов. Только после независимого открытия фагоспецифичных спейсеров (A. Bolotin, B. Quinquis, A. Sorokin, and S. D. Ehrlich. Clustered Regularly Interspaced Short Palindrome Repeats (CRISPRs) Have Spacers of Extrachromosomal Origin. *Microbiology*151 [2005]: 2,551—2,561; F. J. Mojica, C. Diez-Villasenor, J. Garcia-Martinez, and E. Soria. Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *Journal of Molecular Evolution*60 [2005]: 174–182) все концы сошлись и появилась гипотеза о механизме антивирусного иммунитета, опосредованного CRISPR-Cas (K. S. Makarova, N. V. Grishin, S. A. Shabalina, Y. I. Wolf, and E. V. Koonin. A Putative RNA-Interference-Based Immune System in Prokaryotes: Computational Analysis of the Predicted Enzymatic Machinery, Functional Analogies with Eukaryotic RNAi, and Hypothetical Mechanisms of Action. *Biology Direct*1 [2006]: 7). Впоследствии ее основные положения были подтверждены опытным путем (R. Barrangou, C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero, and P. Horvath. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*315 [2007]: 1,709—1,712; F. V. Karginov and G. J. Hannon. The CRISPR System: Small RNA-Guided Defense in Bacteria and Archaea. *Molecular Cell*37 [2010]: 7—19). Из этого следует извлечь важный (и, в ретроспективе, самоочевидный) урок о том, что при интерпретации наблюдений следует принимать во внимание как можно больше фактов. Примечательный последний поворот в этой истории состоит в том, что по крайней мере один из Cas-белков, Cas1, присутствующий во всех системах CRISPR, действительно, по-видимому, способствует не только вставке спейсеров в CRISPR-кассеты, но и участвует в некоторых типах репарации (M. Babu, N. Beloglazova, R. Flick, C. Graham, T. Skarina, B. Nocek, A. Gagarinova, O. Pogoutse, G. Brown, A. Binkowski, S. Phanse, A. Joachimiak, E. V. Koonin, A. Savchenko, A. Emili, J. Greenblatt, A. M. Edwards, and A. F. Yakunin. A Dual Function of the CRISPR-Cas System in Bacterial Antivirus Immunity and DNA Repair. *Molecular Microbiology*79 [2011]: 484–502). В конце концов, оказывается, что первоначальный прогноз не был полностью ошибочным, хотя принципиальная новизна открытия и была упущена.

Системы CRISPR-Cas чрезвычайно разнообразны, в том числе и в том, что касается молекулярных механизмов активности; точная комплементарность спейсера и мишени важна лишь для некоторых из них, в то время как другие удовлетворяются частичной комплементарностью, как РНК у животных (Semenova E, Jore MM, Datsenko KA, Semenova A, Westra ER, Wanner B, van der Oost J, Brouns SJ, Severinov K. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc Natl Acad Sci USA. 2011 Jun 21;108(25):10098-103).

Последние исследования показывают, что дело обстоит сложнее: в механизме действия CRISPR-Cas-систем, наряду с ламарковской схемой, явно прослеживается и дарвиновский путь эволюции, включающий случайные изменения в геноме с последующей селекцией. Показано, что первоначально CRISPR-Cas-система интегрирует многочисленные фрагменты генома фага, но затем лишь очень небольшая доля исходных спейсеров фиксируется под действием отбора и обеспечивает устойчивость данному фагу (Paez-Espino D, Morovic W, Sun CL, Thomas BC, Ueda K, Stahl B, Barrangou R, Banfield JF. Strong bias in the bacterial CRISPR elements that confer immunity to phage. Nat Commun. 2013;4:1430).

Дальнейшее, еще более детальное изучение CRISPR-Cas- систем заставляет считать это объяснение весьма правдоподобным. Действительно, CRISPR-Cas локусы бактерий и архей включают токсин-антитоксинные элементы, которые, видимо, функционируют в тесном взаимодействии с собственно иммунным компонентом CRISPR-Cas (Makarova KS, Anantharaman V, Aravind L, Koonin EV. Live virus-free or die: coupling of antiviral immunity and programmed suicide or dormancy in prokaryotes. Biol Direct. 2012 Nov 14;7:40).

Интереснейшее явление капацитации, безусловно, еще недостаточно изучено, так что далеко идущие интерпретации экспериментов следует воспринимать максимально осторожно. Так, независимая серия экспериментов показывает, что эффект инактивации HSP90 обусловлен не демаскированием скрытой изменчивости, а активацией транспозонного мутагенеза (Specchia V, Piacentini L, Tritto P, Fanti L, D'Alessandro R, Palumbo G, Pimpinelli S, Bozzetti MP. Hsp90 prevents phenotypic variation by suppressing the mutagenic activity of transposons. Nature. 2010 Feb 4;463(7281):662-5). Примечательно, однако, что и в этих экспериментах HSP90 выступает как суппрессор изменчивости.

Прионы стали знамениты именно этим, когда было показано, что таинственные «вирусы» скрэпи овец и коровьего бешенства, а также некоторых редких фатальных неврологических заболеваний у людей на самом деле представляют собой «инфекционные белки» (*M. F. Tuite and T. R. Serio. The Prion Hypothesis: From Biological Anomaly to Basic Regulatory Mechanism. Nature Reviews Molecular Cell Biology*11 [2010]: 823–833). Выяснилось, что механизмы репликации прионов, лежащие в основе их инфекционности, включают в себя автокаталитическую пролиферацию белковых агрегатов (в 1997 году за это открытие Стенли Прузинеру была присуждена Нобелевская премия по физиологии и медицине; *S. B. Prusiner. Prions. Proceedings of the National Academy of Sciences USA*95 [1998]: 13,363—13,383; *G. Vogel. Prusiner Recognized for Once-Heretical Prion Theory. Science*278 [1997]: 214) и, как это обычно случается с поразительными открытиями в области биологии, хоть и новы и непредвиденны (хотя, в определенном смысле, Курт Воннегут и предсказал их в «Колыбели для кошки»), не нарушают центральную догму и принцип ПОР.

Отметим, не вдаваясь в детали, что эти выводы следует воспринимать с максимальной осторожностью, как сугубо предварительные. Имеются аргументы в пользу другой трактовки, согласно которой [PSI+] вовсе не имеет какой-либо функции, а представляет собой «эпигенетическую болезнь» (Wickner RB, Edskes HK, Bateman D, Kelly AC, Gorkovskiy A. The yeast prions [PSI+] and [URE3] are molecular degenerative diseases. *Prion*. 2011 Oct-Dec;5(4):258-62; Wickner RB, Fujimura T, Esteban R. Viruses and Prions of *Saccharomyces cerevisiae*. *Adv Virus Res*. 2013;86:1-36).

Непереводимая игра слов: survival of the fittest – survival of the flattest ; ясно, что «плоский» в этом контексте по сути означает наиболее устойчивый к мутациям.

Этот «обобщенный опережающий эффект» логично рассматривать как прямое нарушение центральной догмы (Koonin EV. Does the central dogma still stand? Biol Direct. 2012 Aug 23;7:27). Однако важно отметить, что это нарушение проявляется на феноменологическом уровне, но не на уровне фундаментальных молекулярных механизмов, в которых нет ничего необычного.

Здесь следует сделать еще одно замечание личного характера: мои собственные исследования в эволюционной геномике начались с вирусов в 1984 году. К тому времени небольшие геномы вирусов были единственными полными геномными последовательностями, и их, ожидающих сравнительного анализа, было уже тридцать или около того, от разных хозяев (животных, растений, бактерий). По-видимому, в то время лишь немногие исследователи осознавали ключевое различие между геномом, то есть полной генетической информацией организма (пусть такого, как вирус, который во многом зависит от хозяина), с его собственной эволюционной историей, и доступными тогда последовательностями фрагментов геномов клеточных форм жизни. В то время не существовало четкого разграничения между функциональной и эволюционной геномикой, так как большая часть последовательности каждого генома была *terra incognita*: попытки предсказания функций вирусных белков шли рука об руку с усилиями по воссозданию эволюционных взаимоотношений (не то чтобы сейчас такого взаимодействия не было, но часто оно не так очевидно). По большому счету, оба направления были плодотворны. Это было невероятно захватывающее время.

Многое здесь основывается на первой статье, касающейся гипотезы мира вирусов (*E. V. Koonin, T. G. Senkevich, and V. V. Dolja. The Ancient Virus World and Evolution of Cells. *Biology Direct*1 (2006): 29*), где можно найти множество ссылок. В этой главе я цитирую в основном публикации по вопросам эволюции вирусов, не вошедшие в статью 2006 года, или ключевые работы, появившиеся позже ее публикации.

В случае эгоистичных элементов не вполне очевидно, какую форму агент-специфической нуклеиновой кислоты следует обозначать как геномную. Для полноценных вирусов геном традиционно – и разумно – определен как инкапсидированная форма (включенная в капсид и вирион), служащая поэтому в качестве распространяемого генома. Для лишенных капсида генетических элементов определение дать труднее, но все еще имеет смысл предложить считать геномом инфекционную форму, в тех случаях, когда таковая есть.

Мой общий подход в этой книге заключается в том, чтобы переносить все мои личные воспоминания и ассоциации в эти примечания. Пусть это будет единственным исключением – его предмет слишком важен для меня лично и может представлять некоторый интерес и для других.

Геном (+)РНК-вируса, лишь отдаленно родственного известным вирусам этого класса, был обнаружен в горячих источниках, заселенных почти исключительно археями, в результате метагеномного анализа. Однако для идентификации природного хозяина этого вируса еще требуется большая экспериментальная работа.

Как нетрудно было ожидать, совсем недавно у нематод были обнаружены и вполне типичные вирусы (Félix MA, Ashe A, Piffaretti J, Wu G, Nuez I, Bélicard T, Jiang Y, Zhao G, Franz CJ, Goldstein LD, Sanroman M, Miska EA, Wang D. Natural and experimental infection of *Caenorhabditis* nematodes by novel viruses related to nodaviruses. *PLoS Biol.* 2011 Jan 25;9(1):e1000586; Franz CJ, Zhao G, Félix MA, Wang D. Complete genome sequence of Le Blanc virus, a third *Caenorhabditis* nematode-infecting virus. *J Virol.* 2012 Nov;86(21):1194).

За не столь долгое время, прошедшее с момента сдачи в набор оригинала этой книги, анализ метагеномов принес немало новых открытий, пожалуй, еще более удивительных, чем описанные здесь. За последний год или два стало рутинной собирать полные вирусные геномы из метагеномных последовательностей. Для некоторых семейств вирусов эти виртуальные геномы уже перекрывают разнообразие, открытое стандартными методами. Так, исключительно путем исследования метагеномов было обнаружено разнообразие геномов оцДНК бактериофагов, о котором до этого не подозревали (Roux S, Krupovic M, Poulet A, Debroas D, Enault F. Evolution and diversity of the Microviridae viral family through a collection of 81 new complete genomes assembled from virome reads. PLoS One. 2012;7(7):e40418). Сходным образом, к трем известным геномам вирофагов (см. ниже) было добавлено пять новых геномов, собранных из метагеномных библиотек (Zhou J, Zhang W, Yan S, Xiao J, Zhang Y, Li B, Pan Y, Wang Y. Diversity of virophages in metagenomic data sets. J Virol. 2013 Apr;87(8):4225-36). Сообщается и об открытии принципиально новых типов вирусных геномов, таких как, например, гибрид оцДНК и оцРНК вирусов (Diemer GS, Stedman KM. A novel virus genome discovered in an extreme environment suggests recombination between unrelated groups of RNA and DNA viruses. Biol Direct. 2012 Jun 11;7:13). Эти исследования не просто приводят к открытию новых групп вирусов, но меняют саму парадигму вирусологии: если традиционно вирус изолировали из организма хозяина, а затем определяли последовательность генома, то теперь накапливается все больше разнообразных геномов «в поисках вирусов и хозяев». Представляется вероятным, что в скором времени такие геномы будут составлять основу наших знаний о вирусах.

Уже упоминавшаяся игра слов: ORFan – производное от Open Reading Frame (открытая рамка считывания), Orphan – сирота; в английском произношении эти слова практически неразличимы.

Нобелевская премия по физиологии или медицине 2009 года была присуждена Элизабет Блекберн, Кэрол Грейдер и Джеку Шостаку за открытие теломеразы (*G. Vogel and E. Pennisi. Physiology Nobel. U.S. Researchers Recognized for Work on Telomeres. Science* 326 (2009): 212–213). Это, конечно, по любому счету замечательное открытие, но понятно, что отчасти побудительным мотивом для награждения служили последствия удлинения или укорочения теломер, вызываемые, соответственно, высоким или низким уровнем экспрессии теломеразы, для рака и старения. Поэтому особенно интересно, что такой важный и значимый для медицины фермент пришел прямиком из мира вирусов.

На ум приходит метафора корабля Тезея, которую неоднократно использовал в обсуждении эволюции генома Антуан Данчин: вопрос, который якобы задали Дельфийскому оракулу – если в корабле героя (о котором с любовью заботились благодарные граждане) постепенно заменить все доски на новые, возможно, из другого материала и даже чуть иной формы, но общие очертания останутся узнаваемыми, будет ли это все тот же корабль? «Правильного» научного определения тождественности не может существовать (см. прил. I), но внутри контекста и следуя логике нашего изучения мира вирусов разумным ответом кажется «да» (хотя и со множеством но).

Ли ван Вален предложил эту гипотезу в явной форме в 1973 году (*L. Van Valen. A New Evolutionary Law. Evol. Theory*1: 1—30). Ее название происходит от знаменитой сцены из «Алисы в Зазеркалье» Льюиса Кэрролла, где Красная Королева объясняет Алисе: «Ну а здесь, знаешь ли, приходится бежать со всех ног, чтобы только остаться на том же месте» (пер. Н. Демурова).

В совсем недавней работе мы с коллегами даем консервативные количественные оценки числа генов, кодирующих системы защиты у разнообразных бактерий и архей, и приводим данные, свидетельствующие о тесном и сложном взаимодействии между этими системами (Makarova KS, Wolf YI, Koonin EV. Comparative genomics of defense systems in archaea and bacteria. Nucleic Acids Res. 2013 Mar 6).

По-видимому, существует тесное сопряжение между системами ПКС и системами активной защиты, такими как CRISPR-Cas и рестрикция-модификация. Говоря в несколько антропоморфных терминах, когда защитные системы не справляются с инфекцией, они посылают сигнал бедствия системам ПКС, и клетка совершает самоубийство, что позволяет избежать заражения других клеток в той же колонии. По-видимому, эволюция альтруизма, в такой простейшей форме, возможна и у прокариот (Makarova KS, Anantharaman V, Aravind L, Koonin EV. Live virus-free or die: coupling of antiviral immunity and programmed suicide or dormancy in prokaryotes. *Biol Direct.* 2012 Nov 14;7:40; Makarova KS, Wolf YI, Koonin EV. Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Res.* 2013 Mar 6).

В последнее время обнаруживаются все более разнообразные системы взлома защиты, кодируемые бактериофагами. Пожалуй, наиболее поразительный пример – недавнее открытие у нескольких бактериофагов собственной системы CRISPR-Cas, которая содержит спейсеры, инактивирующие рестрикционные системы бактерии-хозяина (Seed KD, Lazinski DW, Calderwood SB, Camilli A. A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. Nature. 2013 Feb 28;494(7438):489-91).

Как уже отмечалось выше, иногда сокращение LUCA расшифровывается несколько по-другому: Last Universal Cellular Ancestor (*последний универсальный клеточный предок*). Это, по-видимому, точнее, так как обычно подразумевается именно общий предок всех существующих клеточных форм жизни.

Дуглас Теобальд опубликовал амбициозную статью, в которой заявил, что может привести формальное подтверждение существования LUCA, не зависящее от сходства последовательностей универсальных белков архей, бактерий и эукариот (*D. L. Theobald. A Formal Test of the Theory of Universal Common Ancestry. Nature* 465 [2010]: 219–222). Однако более тщательный анализ его подхода показывает, что его аргументы все равно содержат скрытое предположение о сохранении последовательностей (*E. V. Koonin and Y. I. Wolf . The Common Ancestry of Life. Biology Direct* 5 [2010a]: 64).

Мембранные фосфолипиды архей – простые терпеноидные эфиры глицерол-1-фосфата, в то время как бактериальные фосфолипиды – сложные эфиры жирных кислот и глицерол-3-фосфата, таким образом липиды в этих двух царствах различаются не только химическим составом, но и хиральностью.

Аналогии с историей человеческой цивилизации (включая, само собой, науку) очевидны и, следует думать, поучительны: существование *lingua franca* заметно ускоряет прогресс, тогда как изолированные общества останавливаются в своем развитии и в конце концов обречены на вымирание.

Имеются и существенно иные варианты этой гипотезы. Так, в нашей недавней работе с Арменом Мулкиджаняном и другими коллегами приводятся достаточно серьезные геохимические и биохимические аргументы в пользу того, что благоприятные условия для ранней эволюции и возникновения первых клеток существовали вблизи наземных, а не морских геотермальных источников и что сети неорганических ячеек, в которых могла происходить эволюция, скорее состояли из сульфида цинка, а не сульфида железа (Mulkiđjanian AY, Vuchkov AY, Dibrova DV, Galperin MY, Koonin EV. Origin of first cells at terrestrial, anoxic geothermal fields. Proc Natl Acad Sci USA . 2012 Apr 3;109(14):E821-30). Однако главное здесь – принципиальная идея о существовании абиогенной компартментализации, которая только и могла обеспечить условия для эволюции.

Совсем недавние исследования, возможно, проливают свет на эту проблему. Так называемые L-формы бактерий (в частности, бацилл) лишены клеточной стенки и делятся очень простым способом, без участия сложного бактериального аппарата клеточного деления. Говоря кратко, деление L-форм происходит чисто механически, за счет напряжения в мембране, возникающего при росте клетки. Вполне вероятно, что похожим способом делились первые протоклетки (Koonin EV, Mulkidjanian AY. Evolution of cell division: from shear mechanics to complex molecular machineries. *Cell*. 2013 Feb 28;152(5):942-4; Briers Y, Walde P, Schuppler M, Loessner MJ. How did bacterial ancestors reproduce? Lessons from L-form cells and giant lipid vesicles: multiplication similarities between lipid vesicles and L-form bacteria. *Bioessays*. 2012 Dec;34(12):1078-84).

Важно заметить, что такой переход с одного уровня селекции на другой, по-видимому, лежит в основе многих принципиальных инноваций в эволюции жизни, в частности возникновения многоклеточных организмов, как обсуждается в знаменитой книге Мейнард-Смита и Саммары (Maynard-Smith, J. and Szathmary, E. (1998) The Major Transitions in Evolution. Oxford University Press, Oxford). Оглядываясь назад, в данной книге этому важнейшему аспекту эволюции уделяется недостаточное внимание. Если автору когда-нибудь доведется работать над новым изданием, это будет исправлено.

Согласно недавно предложенной гипотезе Черной Королевы (названной по очевидной аналогии со знаменитой концепцией Красной Королевы – см. гл. 10), такая кооперация путем разделения функций является одним из главных факторов всей эволюции микробных сообществ (Morris JJ, Lenski RE, Zinser ER. The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. *MBio*. 2012 May 2;3(2)).

Напомним, что репликация и экспрессия вирусного генома в простейшем случае происходят вне вириона.

Выше мы уже отмечали новые данные, указывающие на возможность существования примитивных мембран на стадии LUCAS, это может частично снять отмеченный парадокс (Koonin EV, Mulkijanian AY. Evolution of cell division: from shear mechanics to complex molecular machineries. Cell. 2013 Feb 28;152(5):942-4).

Те, кто вынужден был в свое время изучать «диалектический материализм», странный коктейль, который якобы должен был формировать философские основы марксизма в Советском Союзе и других странах социалистического лагеря, никогда не забудут определения Фридриха Энгельса: «Жизнь есть способ существования белковых тел». Если мы отодвинем в сторону отвращение от безжалостного заколачивания этой формулы, вместе с другими драгоценностями марксистской мудрости, в наши бедные мозги, то заметим, что она не так уж плохо звучит и в наше время, хоть и тривиальна и в значительной степени не относится к делу.

Недавно опубликован интересный сравнительный обзор определений жизни (Trifonov EN. Vocabulary of definitions of life suggests a definition. J Biomol Struct Dyn. 2011 Oct;29(2):259-66), с многочисленными комментариями, в основном весьма критическими, написанными в том числе Джеком Шостаком (напомним, одним из лауреатов Нобелевской премии за открытие теломеразы) (Szostak JW. Attempts to define life do not help to understand the origin of life. J Biomol Struct Dyn. 2012;29(4):599-600) и автором этой книги (Koonin EV. Defining life: an exercise in semantics or a route to biological insights? J Biomol Struct Dyn. 2012;29(4):603-5).

Примечательным исключением являются самовоспроизводящиеся прионы, которые мы кратко упоминали в гл. 9. Несмотря на то что прионы формально представляют собой высшую форму сильной эпигенетической наследственности (вплоть до того, что они могут выступать как инфекционные агенты), основанную исключительно на белках, синтез прионных белков по-прежнему полностью зависит от обычной, основанной на нуклеиновых кислотах клеточной системы передачи информации.

Это два класса неродственных, широко распространенных нуклеотид-связывающих доменов.

Джон Уолкер (будущий лауреат Нобелевской премии за структуру мембраны АТФазы) и его коллеги впервые описали фосфат-связывающую петлю в 1982 г. (*J. E. Walker, M. Saraste, M. J. Runswick, and N. J. Gay. Distantly Related Sequences in the Alpha- and Beta-Subunits of ATP Synthase, Myosin, Kinases, and other ATP-Requiring Enzymes and a Common Nucleotide Binding Fold. EMBO Journal 1 [1982]: 945–951*) как мотив, присутствующий в двух субъединицах H⁺-АТФазы и ряде других АТФ-связывающих белков, имеющих мало иного структурного сходства между собой или вовсе несходных. Неожиданная консервация этого мотива в самых разнообразных вирусных белках была предметом моей самой первой работы в вычислительной биологии (*Горбаленя А. Е., Блинов В. М., Донченко А. П., Кунин Е. В. Молекулярная генетика, микробиология и вирусология 11 [1985]: 30–36*). Я полагаю, что точное попадание в наиболее консервативный белковый мотив в столь ранней работе было сочетанием счастливого случая и «предпочтительного присоединения».

Весьма поучительно просто взглянуть на опубликованную модель пространственной структуры рибосомы, основанную на результатах рентгеноструктурного анализа (работа, которая принесла Нобелевскую премию Аде Йонат, Венки Рамакришнану и Томасу Стейтцу) (Steitz TA. A structural understanding of the dynamic ribosome machine. Nat Rev Mol Cell Biol. 2008 Mar;9(3):242-53). Хотя рибосома включает только три молекулы РНК и около 50 молекул белков, РНК намного больше по размеру, так что рибосома выглядит как гигантский конгломерат шпилек и петель РНК с небольшими вкраплениями белков – эта картина сама по себе напоминает нам об РНК-мире.

К сожалению, это примечание не было сделано в английском оригинале, пусть оно будет хотя бы в русском переводе: эта секция представляет интерес только для тех читателей, кто, с одной стороны, хочет разобраться в деталях существующих представлений об эволюции трансляции, а с другой стороны, знаком с материалом на уровне курса молекулярной биологии. Остальные читатели потеряют не слишком много, перейдя сразу к следующему разделу.

«Теория» Опарина была отчасти ответом на требования философии «диалектического материализма», внушавшейся в Советском Союзе как часть общего марксистского мировоззрения и призывавшей к простому материалистическому (на практике часто механистическому и до смешного упрощенному) объяснению всех природных явлений. Пожалуй, не случайно, что первая короткая книга Опарина была напечатана издательством «Московский рабочий» (*Опарин А. И. Происхождение жизни. М.: Московский рабочий, 1924*). Александр Иванович Опарин был весьма одиозной личностью; он успешно подстраивался под переменчивую линию коммунистической партии на протяжении всей своей долгой и очень успешной карьеры, в ходе которой добрался до вершин советской научной иерархии. Его поведение в эпоху лысенковщины, и особенно после разгрома генетики в 1948 г., было столь позорным, насколько вообще возможно. Конечно, все это было чистым стремлением к выживанию, и Опарин прекрасно понимал важность генетики и начал недвусмысленно об этом говорить, как только Лысенко оказался не у власти. У меня была возможность встретиться с Опариным лично в 1971 г., когда я принимал из его рук приз за второе место в биохимической олимпиаде для средних классов. Он производил впечатление довольно отрешенного (скорее всего, из-за глухоты, от которой мало помогали тогдашние примитивные слуховые аппараты), но добродушного старого профессора.

Как уже отмечалось выше, в нашей недавней совместной работе с Арменом Мулкиджаняном и другими коллегами представлена серьезная альтернатива гипотезе Рассела о происхождении жизни в окрестностях глубоководных горячих источников (Mulkidjanian AY, Bychkov AY, Dibrova DV, Galperin MY, Koonin EV. Origin of first cells at terrestrial, anoxic geothermal fields. Proc Natl Acad Sci USA . 2012 Apr 3;109(14):E821-30; Mulkidjanian AY, Bychkov AY, Dibrova DV, Galperin MY, Koonin EV. Open questions on the origin of life at anoxic geothermal fields. Orig Life Evol Biosph. 2012 Oct;42(5):507-16). Вкратце дело сводится к фундаментальному обобщению, которое мы назвали «принципом химической консервативности». Этот принцип состоит в том, что соотношения концентраций основных ионов (в первую очередь калия, натрия, магния, цинка и фосфата) во всех клетках постоянны и, соответственно, по-видимому, не менялись на протяжении всей истории жизни на Земле. Эти соотношения более консервативны, чем любые клеточные или молекулярные структуры. В клетках современных организмов оптимальные концентрации ионов поддерживаются с помощью энергозависимых мембранных насосов, и, соответственно, организмы могут существовать в самых разнообразных условиях, в частности в морской воде, где соотношения концентраций ионов совсем другие, чем в клетках. Однако первые протоциты, естественно, не могли иметь мембран современного типа, снабженных такими насосами (Mulkidjanian AY, Galperin MY, Koonin EV. Co-evolution of primordial membranes and membrane proteins. Trends Biochem Sci. 2009 Apr;34(4):206-15).

Соответственно, первые клетки могли возникнуть только в среде с «правильными» концентрациями ионов. А раз так, можно попытаться выяснить, где на Земле существуют или существовали в прошлом подходящие условия. Морская вода не подходит совершенно, в ней на порядки больше натрия, чем калия, а внутри клеток – наоборот; к тому же в морской воде почти нет фосфата, а в клетках его концентрация высока. Замечательным образом ионный состав, весьма напоминающий внутриклеточную среду, обнаруживается в конденсатах испарений наземных геотермальных источников, которые, как мы полагаем, и могли быть «колыбелями жизни». Эта гипотеза хорошо совместима с результатами сравнения структур древнейших паралогичных белков, в которых отмечается жесткая консервация сайтов связывания калия и фосфата.

Автор отдает себе полный отчет в том, что для некоторых, возможно, многих читателей содержание этого раздела не только противоречит здравому смыслу и устоявшемуся представлению о научном методе, но, что куда хуже, может подорвать доверие ко всей книге. Если у автора и были какие-то иллюзии по этому поводу, то довольно многочисленные комментарии, в том числе на сайте amazon.com, полностью их развеивают. По этому поводу хочется сделать несколько существенных замечаний. Во-первых, важно отдавать себе отчет в том, что теория бесконечной инфляции и гипотеза мультивселенной, на которых основан этот раздел, – это ни в коей мере не философские, а физические концепции. Очень важно, что, хотя прямое наблюдение других вселенных, видимо, невозможно в принципе, наша собственная Вселенная могла сохранить следы инфляционных процессов, которые с неизбежностью вели к мультивселенной (и первые такие следы, похоже, уже обнаружены – см. примечание в конце этого раздела). Вполне вероятно, что теория бесконечной инфляции научна в самом жестком смысле, то есть фальсифицируема (Adrienne L. Erickcek, Marc Kamionkowski, Sean M. Carroll. A Hemispherical Power Asymmetry from Inflation <http://arxiv.org/abs/0806.0377>). Во-вторых, следует подчеркнуть, что эти теории ни в коей мере не были созданы для объяснения происхождения жизни. Их авторы не имели в виду ничего подобного, их интересовало лишь объяснение некоторых важных особенностей структуры нашей Вселенной, которые трудно объяснить иначе (см. приложение II). В более субъективном плане хочется отметить, что этот раздел ни в коей мере не является главным в книге, цель которой – обсуждение новых представлений о всем эволюционном процессе, многие из которых уже хорошо обоснованы, без специального упора на гипотезы происхождения жизни. Для автора крайне огорчительно, когда читатели обращаются в первую очередь к этому разделу (было бы, однако, ошибкой предупреждать против этого в предисловии – несомненно, эффект был бы тем же самым, что и в случае Венедикта Ерофеева, предупреждения которого по поводу состоявшейся полностью из ныне законодательно запрещенных в России выражений главы «Серп и Молот – Карачарово» в его бессмертной поэме привели к тому, что юные читательницы только ее и читали). Именно во избежание этого «эффекта Ерофеева» автор рассматривал возможность перенести весь этот материал в приложение. Однако это не было сделано по соображениям интеллектуальной честности – автор считает представленные здесь идеи важными, хотя и никак не центральными в книге.

Имеется в виду начало эволюции путем естественного отбора (и дрейфа), а не формирование видов, которое понимал под «дарвиновским порогом» Карл Вёзе.

Конечно, это формулировка «слабого» антропного принципа как единственно разумной научной интерпретации антропного рассуждения. Так называемый «сильный» антропный принцип – это телеологическое представление о том, что наше человеческое существование является, в некоем загадочном смысле, «целью» эволюции Вселенной. Как таковая, эта идея не принадлежит к области науки (*J. D. Barrow and F. J. Tipler. The Anthropic Cosmological Principle. Oxford: Oxford University Press, 1988*).

Интересно и важно отметить, что, в то время как готовился русский перевод этой книги, анализ наблюдений радиоастрономического зонда «Планк» выявил гетерогенность реликтового микроволнового излучения, которая предсказывается теорией инфляции (Planck Collaboration: P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Bandy, R. B. Barreiro, M. Bartelmann, J. G. Bartlett, E. Battaner, K. Benabed, A. Benoit, A. Benoit-Lèvy, J.-P. Bernard, M. Bersanelli, P. Bielewicz, J. Bobin, J. J. Bock, A. Bonaldi, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, J. W. Bowyer, M. Bridges, M. Bucher, C. Burigana, R. C. Butler, B. Cappellini, J.-F. Cardoso, R. Carr, M. Casale, A. Catalano, A. Challinor, A. Chamballu, R.-R. Chary, X. Chen, L.-Y. Chiang, H. C. Chiang, P. R. Christensen, S. Church, D. L. Clements, S. Colombi, L. P. L. Colombo, F. Couchot, A. Coulais, B. P. Crill, A. Curto, F. Cuttaia, L. Danese, R. D. Davies, R. J. Davis, P. de Bernardis, A. de Rosa, G. de Zotti, et al. (218 additional authors not shown). Planck 2013 results. I. Overview of products and scientific results. arXiv :1303.5062 [astro-ph. CO]; Planck Collaboration. Planck 2013 results. XXII. Constraints on inflation arXiv :1303.5082 [astro-ph.CO]). И хотя это предварительные данные, от которых далеко до прямой поддержки МММ, вероятность того, что рассуждения в этом разделе книги вообще ни на чем не основаны, по-видимому, сильно упала.

Хотя и это не стоит сбрасывать со счета, поскольку ситуация с фондированием фундаментальных исследований, не имеющих немедленного практического выхода, мягко говоря, не улучшается.

Используя термин «ризома», я следую авторитетнейшему микробиологу Дидье Раулю (*D. Raoult*. The Post-Darwinist Rhizome of Life. *Lancet* 375 [2010]: 104–105) и цитирую знаменитый постмодернистский текст Делёза и Гваттари, которым «ризома» служила метафорой сложности мира в целом (*G. Deleuze and F. Guattari*. *Thousand Plateaus: Capitalism and Schizophrenia*. Minneapolis, MN: University of Minnesota Press, 1987; перевод: Делёз Ж., Гваттари Ф. Тысяча плато. Капитализм и шизофрения / Пер. с франц. Я. Свирский. Екатеринбург: У-Фактория, 2010). Замена одной биологической метафоры (дерева) другой (ризомой) весьма уместна.

Подобно взаимодействию между частицами в жидкости или твердом теле, в отличие от идеального газа.

Исследование весьма общих моделей эволюции на пересеченных ландшафтах показывает, что во многих случаях дело обстоит именно так (Lobkovsky AE, Koonin EV. Replaying the tape of life: quantification of the predictability of evolution. Front Genet. 2012;3:246).

Совсем недавно этот вездесущий транскрипционный шум послужил поводом для беспрецедентной шумихи и чуть ли не скандала вокруг заявлений проекта ENCODE о функциональности 80 % последовательностей в геноме человека, к тому же не совсем точно истолкованных журналистами (*ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57–74*). С точки зрения эволюционной биологии подобные утверждения смехотворны. С одной стороны, в небольших популяциях, как у млекопитающих, очищающий отбор в принципе не может контролировать эволюцию таких гигантских количеств генетического материала, а с другой стороны, совершенно бессмысленная транскрипция большей части генома неизбежна по той же самой причине: слабый отбор не в состоянии ей эффективно препятствовать. Эти и подобные аргументы приводятся в нескольких интересных статьях, направленных на опровержение утверждения авторов ENCODE (*Graur D, Zheng Y, Price N, Azevedo RB, Zufall RA, Elhaik E. On the Immortality of Television Sets: «Function» in the Human Genome According to the Evolution-Free Gospel of ENCODE. Genome Biol Evol. 2013 Jan;5(3):578-90; Niu DK, Jiang L. Can ENCODE tell us how much junk DNA we carry in our genome? Biochem Biophys Res Commun. 2013 Jan 25;430(4):1340-3. Doolittle WF. Is junk DNA bunk? A critique of ENCODE. Proc Natl Acad Sci USA. 2013 Apr 2;110(14):5294-300*). Эта громкая история чрезвычайно поучительна: она показывает, что понимание простых принципов эволюционной теории незаменимо для самых актуальных и практически важных исследований в современной биологии.

Подчеркнем, что связь между геномом и фенотипом не только сложна. По-видимому, многие ее аспекты не детерминированы в принципе. Далеко идущие следствия этого положения еще необходимо изучить и понять.

Уже после публикации оригинала этой книги такое исследование было проведено и выявило существенные различия в свойствах разных типов ландшафтов (Lobkovsky AE, Wolf YI, Koonin EV. Quantifying the similarity of monotonic trajectories in rough and smooth fitness landscapes. Mol Biosyst. 2013 Mar 4).

Слово «открытие» здесь употребляется вполне сознательно – имея в виду открытие именно вирусного мира, а не вирусов как таковых (они были открыты более ста лет назад). До последнего времени мы не имели сколько-нибудь разумного представления о масштабах, а значит, по сути, и о существовании этого мира.

Сам термин «империя» идет от интереснейшего дебата между двумя великими эволюционистами, Эрнстом Майром и Карлом Вёзе. Однако они спорили лишь о классификации клеточных форм жизни: Вёзе настаивал на трехдоменной схеме, а Майр, на основании чисто биологических соображений, считал, что «правильные» империи – прокариоты и эукариоты (Mayr E. Two empires or three? Proc Natl Acad Sci USA . 1998 Aug 18;95(17):9720-3; Woese CR. Default taxonomy: Ernst Mayr's view of the microbial world. Proc Natl Acad Sci USA . 1998 Sep 15;95(19):11043-6). Автор этой книги имел честь консультировать Эрнста Майра по поводу эволюции архей и удостоился благодарности в статье. Это, похоже, редчайший повод для «законной гордости».

Насколько нам известно в настоящее время. Гипотеза Черной Королевы, уже упоминавшаяся выше, предполагает, что даже для этих организмов некоторые вещества поставляются другими членами микробного сообщества (Morris JJ, Lenski RE, Zinser ER. The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. MBio. 2012 May 2;3(2)).

Здесь мы говорим о прогрессе в исследованиях, который, несомненно, имеет место, – в отличие от прогресса в эволюции.

В недавней популярной книге «Высший замысел» Стивен Хокинг и Леонард Млодинов пишут: «Что касается законов, управляющих Вселенной, скажем следующее: непохоже, чтобы существовала единая математическая модель или теория, которая могла бы описать всю Вселенную. Вместо этого... возникает целая сеть теорий, называемая М-теорией. Каждая из теорий в сети М-теории хорошо описывает явления в определенных рамках. Там, где их границы перекрываются, различные теории в этой сети согласуются, так что все они могут считаться частями одной и той же теории. Но нет ни одной теории, которая была бы в состоянии описать все аспекты Вселенной: все силы природы, частицы, подверженные этим силам, и структуру пространства-времени, в котором разворачивается их действие. Хотя в этой ситуации и не сбывается мечта традиционной физики о теории великого объединения, она приемлема в рамках модельно-зависимого реализма» (*S. W. Hawking and L. Mlodinow. The Grand Design. London: Bantam, 2010. P. 58*). Вероятно, трезвая оценка, весьма далекая от надежды на то, что «мы читаем мысли Бога», выражением которой завершается ставшая классической книга Хокинга «Краткая история времени» (*S. W. Hawking. A Brief History of Time: From the Big Bang to Black Holes. London: Bantam, 1988*; перевод: Хокинг С. Краткая история времени / Пер. с англ. Н. Смородинской. М.: Амфора, 2010).

По знаменитому выражению Эрнеста Резерфорда.

А с другой точки зрения, только предсказательная сила.

Приложение представляет собой переработанную статью *E. V. Koonin*. The Cosmological Model of Eternal Inflation and the Transition from Chance to Biological Evolution in the History of Life. *Biology Direct*2 (2007): 15.

Существование связи между ландшафтом в теории суперструн и инфляционной мультивселенной не бесспорно.