

УДК 575

## ВЗАИМОСВЯЗЬ РАЗМЕРА ГЕНОМА И СЛОЖНОСТИ ОРГАНИЗМА В ЭВОЛЮЦИОННОМ РЯДУ ОТ ПРОКАРИОТ К МЛЕКОПИТАЮЩИМ

© 2010 г. А. В. Марков\*, В. А. Анисимов\*\*, А. В. Коротаев\*\*\*

\*Палеонтологический институт им. А.А. Борисяка РАН  
e-mail: markov\_a@inbox.ru

\*\*A2iA: Искусственный интеллект и анализ изображений  
e-mail: va@a2ia.com

\*\*\*Российский государственный гуманитарный университет  
e-mail: akorotayev@mail.ru

Поступила в редакцию 21.04.2009 г.

Принята к печати 19.06.2009 г.

Отсутствие строгой зависимости между уровнем организации (сложностью) организмов и размером генома связано с вариабельностью размеров “факультативной” части генома. В эволюционной линии, ведущей от прокариот к млекопитающим, в которой тенденция к усложнению проявилась наиболее сильно, выявлена прямая зависимость между уровнем организации и минимальным размером генома (МРГ). Динамика МРГ в этой линии хорошо описывается моделью гиперэкспоненциального роста, что позволяет предполагать наличие нелинейных положительных обратных связей, обеспечивающих ускорение роста МРГ. Рассмотрена природа этих связей, включая формирование новых генов путем перекомбинирования фрагментов имеющихся генов и формирование “ниш” для новых генов в ходе эволюции генных сетей, а также усложнение регуляторных участков. Гиперэкспоненциальный рост показателей, отражающих уровень организации биосферы и социума (биоразнообразие, МРГ, размер и сложность организмов, народонаселение, уровень технологического развития, урбанизации и т.д.) позволяет предположить, что эволюция биосферы и социума по пути к общему усложнению является самоускоряющимся (автокаталитическим) процессом.

Вопрос о возможной связи между размером генома и уровнем морфофизиологической организации (или сложностью строения) организма давно привлекает внимание биологов. Предположение о существовании положительной корреляции между этими показателями представляется логичным, поскольку для построения в ходе онтогенеза более сложного организма, казалось бы, требуется более объемная “программа развития”. Однако до последнего времени выявить сколько-нибудь строгую зависимость такого рода не удавалось (Gregory, 2005). Попытки выявить закономерный рост генома в отдельных эволюционных линиях, как правило, не приводили к успеху (Thomas, 1971; Gregory, Hebert, 1999; Gregory, 2005, 2008). С другой стороны, отмечается значительное увеличение генома при переходе от прокариот к эукариотам и от одноклеточных эукариот к многоклеточным; имеются указания на наличие положительной связи между размером уникальных геномных последовательностей и уровнем организации; кроме того, отмечается общий рост генома в наиболее “прогрессивной” эволюционной линии, ведущей от прокариот к млекопитающим (в этой линии наиболее четко выражена тенденция к морфофизиологическому прогрессу) (Рэфф, Кофман, 1986; Patthy, 1999; Sharov, 2006). Целью настоящей работы является анализ взаимо-

связи между сложностью генома и организма в этой эволюционной линии. В качестве “меры сложности” генома нами используется минимальный размер генома (МРГ) в пределах крупного таксона. Ниже мы покажем, что в данной эволюционной линии наблюдается рост МРГ, который хорошо описывается моделью двухэкспоненциальной разновидности гиперэкспоненциального роста. Ранее модели гиперэкспоненциального роста (прежде всего гиперболические) с успехом применялись для описания динамики ряда демографических и макросоциологических показателей (народонаселения, уровней технологического и экономического развития, урбанизации, грамотности и др.) (Коротаев и др., 2007а, б; Гринин, Коротаев, 2009). Кроме того, показана приложимость гиперболической модели к динамике таксономического разнообразия фанерозойской биоты (Марков, Коротаев, 2008, 2009). Гиперэкспоненциальный рост обычно свидетельствует о наличии сложных нелинейных положительных обратных связей, способствующих ускорению роста изучаемого показателя.

Работа выполнена при поддержке Программ Президиума РАН “Биоразнообразие и динамика генофондов” и “Происхождение биосферы и эволюция гео-биологических систем, Подпрограмма II” и РФФИ (проект 09-05-01009). Авторы благодарны

А.Д. Панову, А.С. Раутиану и А.А. Фомину за обсуждение идей, рассматриваемых в данной статье, и ценные рекомендации.

### МИНИМАЛЬНЫЙ РАЗМЕР ГЕНОМА КАК МЕРА НЕОБХОДИМОГО (НЕИЗБЫТОЧНОГО) КОЛИЧЕСТВА ГЕНЕТИЧЕСКОЙ ИНФОРМАЦИИ

Мы рассматриваем именно минимальный (а не средний и не максимальный) размер генома в каждой группе, потому что, как известно, геномы большинства организмов содержат много так называемой “мусорной ДНК” (*junk DNA*), которая не является необходимой для жизни организма. Ее количество может сильно варьировать даже в пределах одного класса или отряда. Варибельность объема “мусорной ДНК” во многом определяется случайными факторами, не связанными с жизнедеятельностью и адаптивными характеристиками организма. Значительная часть “мусорной ДНК” представлена мобильными генетическими элементами (МГЭ). Так, у приматов на долю МГЭ приходится около 50% всего генома (Xing et al., 2007). Присутствие “лишней” ДНК в геноме возлагает дополнительную нагрузку на организм, которому приходится тратить больше ресурсов на репликацию ДНК; разрастание генома может влиять на размер клетки, скорость деления и др. (Gregory, 2005; Gregory, Hebert, 1999). С другой стороны, представление о бесполезности МГЭ, интронов и других некодирующих последовательностей является весьма условным. МГЭ служат источником генетической изменчивости, а отдельные фрагменты этих эгоистических элементов могут быть адаптированы для выполнения полезных функций (Miller et al., 1999; Volf, 2006). Показана значительная роль МГЭ в эволюции эукариот, в том числе в эволюции по пути усложнения (Bowen, Jordan, 2002; Muotri et al., 2007).

В ходе эволюции устанавливается динамическое равновесие между тенденциями к разрастанию и сокращению “мусорной ДНК”. Первая тенденция обусловлена, прежде всего, самопроизвольным размножением МГЭ, вторая — их утратой в результате делеций (Gregory, 2004b). Утрата МГЭ в общем случае выгодна хозяину, поскольку обеспечивает экономию ресурсов, расходуемых на синтез и содержание в каждой клетке “лишней” ДНК. Если одна из двух тенденций начинает преобладать, происходит “разбухание” или, наоборот, сокращение генома. Преобладание первой тенденции может быть связано с появлением новой формы МГЭ, обладающей повышенной скоростью размножения, или с ослаблением клеточных систем контроля МГЭ. Вторая тенденция может преобладать, если утрата лишних фрагментов генома дает значительное адаптивное преимущество. Типичный пример — сокращение генома у птиц и рукокрылых по сравнению с неле-

тающими тетраподами. Летающим позвоночным уменьшение веса тела дает адаптивное преимущество. “Лишняя” ДНК присутствует в каждой клетке и требует для своего содержания и воспроизводства множества “лишних” белков, что проявляется в положительной корреляцией между размером генома и размером клетки (Organ et al., 2007). Поэтому отбор благоприятствовал утрате “мусорной ДНК” у летающих позвоночных, что привело к сокращению генома (Hughes, Hughes, 1995).

Подчеркнем, что ни рост сложности, ни рост размера генома не являются всеобщими эволюционными тенденциями. Так, в эволюции прокариот доминирующей тенденцией является не увеличение, а сокращение генома (Ochman, 2005). Аналогичная картина, по-видимому, наблюдалась и в эволюции ящеротазовых динозавров и птиц (Organ et al., 2007).

Физиологические особенности организма могут накладывать ограничения на размер генома. Так, у прокариот, с их несовершенными системами репарации и распределения ДНК в ходе деления (отсутствии митоза), геном не может разрастаться сверх некой предельной величины (Sharov, 2006): максимальный размер генома у бактерий — 13.03 млн. пар нуклеотидов (п.н.) (*Sorangium cellulosum*), у архей — 5.75 млн. п.н. (*Methanosarcina acetivorans*) (Complete microbial genomes, 2008). Ограничителем роста генома у прокариот может служить и громадная численность популяций, что замедляет генетический дрейф (Lynch, Conery, 2003). С другой стороны, имеющийся уровень сложности организма предполагает, что его геном не может и сократиться ниже определенного минимального уровня.

Мы полагаем, что МРГ в большой группе организмов может быть использован для приблизительной оценки объема необходимой (неизбыточной) генетической информации, обеспечивающей существование представителей данного таксона. Более точную оценку получить трудно, поскольку нет абсолютно надежных способов отличить действительно ненужные участки ДНК от функционально значимых (например, от некодирующих последовательностей, выполняющих регуляторные функции).

### МАТЕРИАЛЫ И МЕТОДЫ

Сравнивались МРГ в девяти последовательных “вложенных” друг в друга группах организмов (табл. 1). Последовательность групп соответствует эволюционной линии, ведущей от первых прокариот к млекопитающим. Выбор организма с минимальным геномом в каждой группе осуществлялся без учета внутриклеточных симбионтов и паразитов, которые часто подвергаются генетическому упрощению и в определенном смысле теряют право именоваться самостоятельными организмами (см.,

**Таблица 1.** Минимальный (≈неизбыточный) размер генома (МРГ) в девяти “вложенных” группах организмов

Группа	МРГ (млн. п.н.)	Примерное время появления (млн. лет назад)	Вид – обладатель самого маленького генома (кроме внутриклеточных паразитов)
1. Все живое	<b>1.3</b>	<b>4000</b>	Морская свободноживущая бактерия <i>Pelagibacter ubique</i> , штамм HTCC1062
2. Эукариоты	<b>9.2</b>	<b>2000</b>	Сумчатый гриб <i>Ashbya gossypii</i> – обладатель самого маленького генома среди свободноживущих эукариот
3. Животные (Metazoa)	<b>19.6</b>	<b>1250</b>	Нематода <i>Pratylenchus coffeae</i>
4. Хордовые	<b>68.6</b>	<b>575</b>	<i>Oikopleura dioica</i> , аппендикулярия, представитель подтипа Tunicata (оболочники) типа хордовых
5. Позвоночные	<b>342</b>	<b>540</b>	Костная рыба <i>Tetraodon fluviatilis</i>
6. Тетраподы	<b>931</b>	<b>375</b>	Лягушка <i>Limnodynastes ornatus</i>
7. Амниоты	<b>951</b>	<b>315</b>	Фазан <i>Phasianus colchicus</i>
8. Млекопитающие	<b>1695</b>	<b>220</b>	Летучая мышь <i>Miniopterus schreibersi</i>
9. Приматы	<b>2215</b>	<b>65</b>	Обезьяна-прыгун <i>Callicebus torquatus</i>

**Примечание.** Данные о размерах минимальных геномов в группах основаны на следующих источниках: Pellicciari et al., 1982; Gregory, 2004a, 2008; Dietrich et al., 2004; Complete Microbial Genomes, 2008; Eukaryotic genome sequencing projects, 2008. Приблизительные датировки времени появления групп основаны на молекулярных и палеонтологических данных из следующих источников: Marshall, Schultze, 1992; Shu et al., 1999; Heges, Kumar, 2003; Battistuzzi et al., 2004; Федонкин, 2006; Falcon-Lang et al., 2007. Ряд данных указывает на более раннее появление эукариот – возможно, 2,7 млрд. лет назад (Розанов, 2003), однако один из важнейших фактов такого рода (наличие эукариотических биомаркеров в архейских породах) недавно был подвергнут серьезным сомнениям (Rasmussen, 2008). В связи с этим мы используем более консервативную оценку 2 млрд. лет, согласующуюся с мнением многих авторов, а также с молекулярными данными.

например: Nakabachi et al., 2006). Внутриклеточные симбионты эксплуатируют гены хозяина взамен утраченных собственных генов, что и позволяет им обходиться без многих генов, которые для свободноживущих представителей данной группы являются абсолютно необходимыми. Гены симбионта могут переноситься в геном хозяина, где и продолжают функционировать к общему благу симбионта и хозяина (как это было при становлении симбиоза ранних эукариот с будущими митохондриями и пластидами) (Stegemann et al., 2003; Марков, Куликов, 2005).

Каждая последующая из рассматриваемых групп организмов является “подмножеством” предыдущей, причем именно тем подмножеством, в пределах которого организмы достигли наибольшего уровня сложности. Характерно, что в пределах каждой группы наименьший размер генома отмечен для тех представителей группы, которые не входят в состав следующей вложенной группы. Например, в группе 6 (тетраподы) наименьший геном характерен для представителя амфибий, то есть “низших” тетрапод, а не для кого-либо из представителей “высшей” вложенной группы 7 (амниоты), хотя амниоты входят в состав тетрапод. Уже из этого можно

видеть, что между МРГ и сложностью организма существует положительная связь.

Нам пришлось исключить из рассмотрения некоторые промежуточные уровни, по которым нет достаточно надежных данных. Следуя использованной логике отбора групп, можно было между группами 5 (позвоночные) и 6 (тетраподы) поместить “промежуточную” группу челюстноротых. Это не было сделано из-за недостаточности данных по бесчелюстным, то есть по тем позвоночным, которые не входят в состав челюстноротых. Бесчелюстные были разнообразны в палеозое, но в современной биоте они представлены только двумя крайне специализированными группами – миногами и миксинами, среди которых пока не обнаружены формы с маленькими геномами (Gregory, 2008). Однако это не дает оснований утверждать, что у палеозойских бесчелюстных геномы были также велики, как у современных миног и миксин (на основании данных, изложенных в настоящей статье, мы предполагаем, что не были). Тем не менее, следует признать, что учет подобных реликтовых или слабо изученных групп мог бы в известной мере замаскировать обсуждаемые в данной статье закономерности.

**Таблица 2.** Данные, использованные ранее для обоснования экспоненциального характера роста неизбыточного размера генома (Sharov, 2006)

Группа	Неизбыточный размер генома (млн. п.н.)	Время появления (млн. лет назад)
1. Прокариоты	0.5	3500
2. Эукариоты	2.9	2000
3. Черви	72.8	1000
4. Рыбы	133.3	500
5. Млекопитающие	480	125

Важный вопрос, который неизбежно встает при любом исследовании с применением регрессионного анализа – выбор адекватных моделей. Слишком простая модель может не позволить выявить существенные детали в исследуемых закономерностях, в то время как слишком сложная грозит опасностью излишней настройки на шумовую (случайную) составляющую экспериментальных данных. Учитывая это, мы исследовали два класса моделей: двух- и трехпараметрические. Двухпараметрические модели в общем случае более надежны и менее подвержены влиянию погрешностей измерений, а трехпараметрические – более информативны (при условии достаточной точности исходных данных).

Мы исследовали два семейства двухпараметрических моделей – экспоненциальные и гиперболические. Экспоненциальные модели в общем случае могут быть представлены в виде:  $L = Ae^{(-BT)}$ , где  $L$  – МРГ таксона (в п.н.),  $A$ ,  $B$  – “подгоночные” параметры,  $T$  – период времени, прошедшего от предполагаемого появления таксона (в млн. лет). Гиперболические модели описываются формулой  $L = A/(B + T)$ . Поскольку значение  $L$  в ряду от прокариот к млекопитающим изменяется более, чем на три порядка, представляется адекватным использование логарифмической шкалы (мы пользовались натуральными логарифмами МРГ).

Были также рассмотрены три семейства трехпараметрических моделей – степенная экспоненциальная, степенная гиперболическая и двухэкспоненциальная. Степенная экспоненциальная модель описывается формулой  $L = Ae^{(-BT^N)}$ , где  $N$  – третий подбираемый параметр. Степенной гиперболической модели соответствует обобщенная формула  $L = A/(B + T)^N$ . Двойная экспоненциальная модель задается в форме  $L = Ce^{Ae^{(-BT)}}$ , причем роль третьего параметра играет нормировочный коэффициент  $C$ . Оптимальные параметры моделей подбирались прямым численным перебором при помощи метода наименьших квадратов.

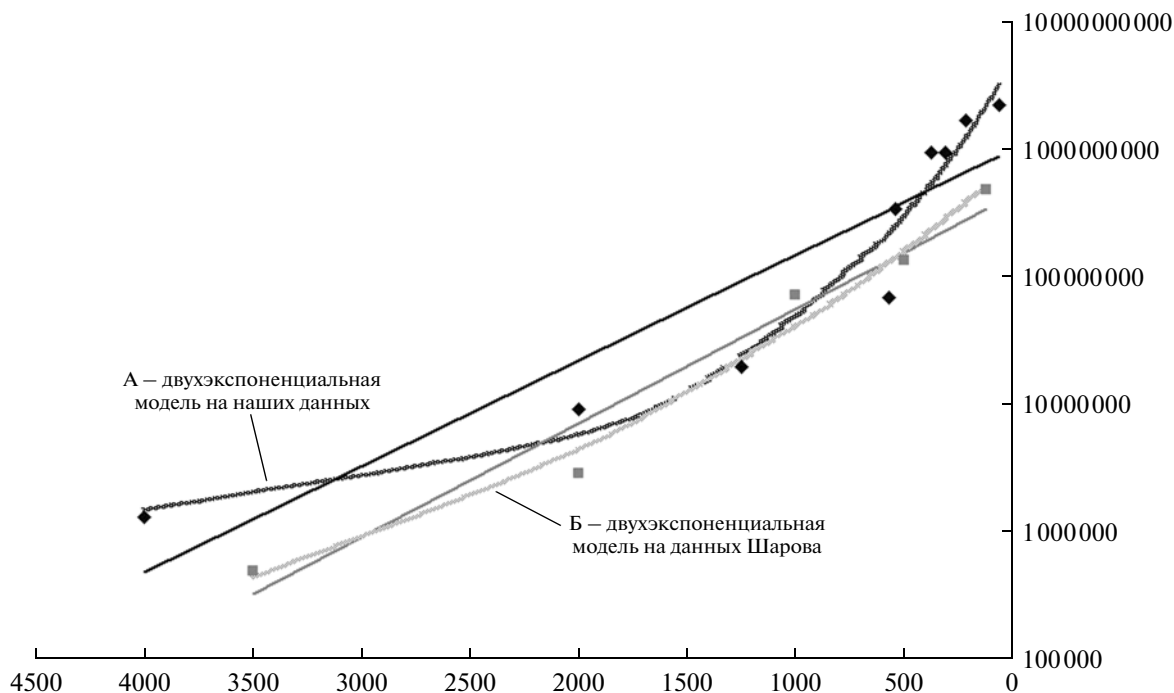
## РЕЗУЛЬТАТЫ

Из рассмотренных двухпараметрических моделей имеющимся данным (табл. 1) наилучшим образом соответствует экспоненциальная модель со значениями параметров  $A = 1.002 \times 10^9$ ,  $B = 0.00191$ ; коэффициент корреляции Пирсона достаточно высок ( $R^2 = 0.914$ ). Из трехпараметрических моделей наилучшее приближение дает двухэкспоненциальная модель со значениями параметров:  $A = 8.41$ ,  $B = 7.98 \times 10^{-4}$ ,  $C = 1.069 \times 10^8$ . Корреляция с наблюдаемыми данными оказывается выше, чем для простой экспоненциальной модели ( $R^2 = 0.979$ ) (что неудивительно, поскольку трехпараметрические модели в общем случае позволяют точнее аппроксимировать фактические данные, чем двухпараметрические). Для контроля были также вычислены параметры экспоненциальной и двухэкспоненциальной моделей по данным из статьи А.А. Шарова (Sharov, 2006; табл. 2; см. Обсуждение). Для экспоненциальной модели получились следующие оптимальные значения параметров:  $A = 4.37 \times 10^8$ ,  $B = 0.00206$ , для двухэкспоненциальной:  $A = 11.75$ ,  $B = 2.88 \times 10^{-4}$ ,  $C = 6.077 \times 10^3$ . Отметим, что сам Шаров в рамках экспоненциальной модели пришел к аналогичным результатам.

Динамика роста МРГ в рассматриваемом эволюционном ряду показана на рис. 1 (черные ромбы). Там же показаны данные Шарова (серые квадраты). Прямые, соответствующие оптимальным двухпараметрическим моделям, показаны сплошными линиями черного (для наших данных) и серого (для данных Шарова) цветов, а кривые, соответствующие трехпараметрическим моделям – в виде муаровых черной (наши данные) и серой (данные Шарова) линий.

## ОБСУЖДЕНИЕ

**Экспоненциальный или гиперэкспоненциальный рост?** Анализ динамики размера “неизбыточного” генома в данной эволюционной линии, проведенный Шаровым (Sharov, 2006), привел его к выводу об экспоненциальном росте этого показателя. Наш анализ отчасти повторяет работу Шарова, однако мы иначе оцениваем неизбыточный геном, а также используем более детальные данные. Шаров построил свой экспоненциальный график на основе пяти точек, которые по своему биологическому смыслу соответствуют рассматриваемым нами группам 1, 2, 3, 5 и 8 (табл. 1, 2). Почему же Шаров сделал вывод об экспоненциальном характере роста, а у нас с разумной степенью уверенности получается более ускоренный, гиперэкспоненциальный рост? Это объясняется, во-первых, тем, что Шаров в своей работе вообще не рассматривал возможность существования иного, чем экспоненциальный, закона роста МРГ, во-вторых, разницей в подходах к оценке неизбыточного генома. Для групп 1 и 2 Ша-



**Рис. 1.** Двухэкспоненциальный рост минимального размера генома (МРГ) в ряду от прокариот к млекопитающим. По горизонтальной оси — время появления группы, млн. лет назад. По вертикальной оси — в логарифмической шкале МРГ (п.н.). Маркеры соответствуют группам, указанным в табл. 1. А — двухэкспоненциальная модель (муаровая черная линия) на наших данных описывает наблюдаемую динамику лучше, чем экспоненциальная (сплошная черная линия). Б — двухэкспоненциальная модель (муаровая серая линия) по данным Шарова (Sharov, 2006) тоже описывает наблюдаемую динамику несколько лучше, чем экспоненциальная (сплошная серая линия).

ров использовал размер геномов паразитических микроорганизмов, тогда как мы используем минимальные геномы свободноживущих форм. Однако эти различия слабо влияют на корреляцию данных с экспоненциальной и двухэкспоненциальной моделями. Необходимый размер генома в группе 3 у Шарова равен 72.8 (это размер генома нематоды *Caenorhabditis elegans* за вычетом 25%, которые предположительно считаются нефункциональными). Наша оценка (19.6) более реалистична, поскольку таков размер генома другого круглого червя (*Pratylenchus coffeae*). Очевидно, в геноме *C. elegans*, даже за вычетом “нефункциональной” части, присутствует много участков, без которых другие круглые черви успешно обходятся. Для рыб и млекопитающих оценки Шарова занижены по сравнению с нашими, потому что Шаров исключил из рассмотрения те части генома, которые предположительно являются нефункциональными (это 2/3 генома рыб и 85% генома млекопитающих). Однако критерии, применяемые Шаровым для оценки размера “нефункциональной” части генома, по-видимому, не учитывают того обстоятельства, что многие некодирующие участки с неустановленной функцией в действительности может выполнять важные регуляторные функции или кодировать функциональные РНК. Так, в геномах млекопитающих подвергаются транскрипции многие участки генома, считав-

шиеся нефункциональными; кроме того, среди этих участков недавно обнаружен целый класс не известных ранее генов, кодирующих большие молекулы РНК с регуляторными функциями (Guttman et al., 2009). Многие факты указывают и на важнейшую роль МГЭ в эволюции эукариот, в том числе в эволюции по пути усложнения (Miller et al., 1999; Bowen, Jordan, 2002; Muotri et al., 2007). На наш взгляд, адекватное представление о минимально необходимом геноме у позвоночных дают те организмы, для которых уменьшение генома является адаптивно выгодным. По-видимому, как раз в таком положении находятся летающие позвоночные — птицы и рукокрылые (Hughes, Hughes, 1995; Organ et al., 2007). Именно их геномы и рассматривались нами в качестве “минимально необходимых” для групп 7 (амниоты) и 8 (млекопитающие). Последнее отличие данных Шарова состоит в том, что по Шарову, млекопитающие появились около 125 млн. лет назад (ранний мел). Мы не можем согласиться с такой оценкой, поскольку млекопитающие известны уже из позднего триаса (около 220 млн. лет назад), тогда как 125 млн. лет — это возраст древнейшей находки *плацентарных* млекопитающих (Ji et al., 2002).

**Сравнение полученной нами двухпараметрической модели с результатами Шарова.** Несмотря на отмеченные расхождения в исходных данных, итоговый

наклон экспоненциальных прямых на логарифмическом графике (рис. 1) получился очень сходным. Это косвенно подтверждает адекватность тех грубых оценок, которые дают экспоненциальные модели. Более простые двухпараметрические модели, по-видимому, дают хотя локально и более грубую, но глобально более надежную оценку исследуемых закономерностей. Поэтому не исключено, что и к прикидкам сроков появления жизни во Вселенной (не менее 7 млрд. лет назад), полученным Шаровым на основе экстраполяции модели в прошлое, стоит относиться достаточно серьезно (но только в том случае, если признать принципиальную допустимость подобной экстраполяции).

#### **Размер генома и сложность строения организма.**

Утверждение об отсутствии прямой корреляции между сложностью организма и размером генома имеет под собой веские основания. Так, среди одноклеточных эукариот есть формы, геном которых превосходит по размеру все изученные геномы многоклеточных; среди земноводных есть виды с геномами большего размера, чем у любых млекопитающих и т.д. (Thomas, 1971; Gregory, Hebert, 1999; Gregory, 2005). Некоторые причины широкой вариативности размера генома упомянуты выше. Однако в эволюционной линии от прокариот к млекопитающим, в которой рост сложности организма был выражен наиболее сильно, положительная корреляция между избыточным размером генома и сложностью организма отчетливо прослеживается (Patthy, 1999; Sharov, 2006). Это соответствует теоретическим ожиданиям, основанным на представлении о геноме как о “программе” развития и функционирования организма. Естественно ожидать, что чем сложнее организм, тем сложнее (а следовательно и больше) должна быть соответствующая “программа”. Так, предложено рассматривать размер функциональной избыточной части генома в качестве меры биологической сложности организма (Adami et al., 2000). Мы полагаем, что МРГ в пределах большой группы организмов является хорошим приближением к трудновычисляемому размеру “неизбыточного” генома. Таким образом, предположение о тесной связи МРГ со сложностью организма не противоречит общепринятым взглядам.

**Вендско-кембрийское ускорение роста МРГ.** Несмотря на в целом хорошую аппроксимацию динамики МРГ двойной экспоненциальной формулой, на графике видно существенное расхождение фактических данных с аппроксимирующей кривой, приходящееся примерно на конец протерозоя (венд) и кембрий. Резкое ускорение роста МРГ в венде — кембрии совпадает с адаптивной радиацией Metazoa (вендско-кембрийский “взрыв”). К этому времени относится, по-видимому, появление первых хордовых (точка 4 на графике), а затем и первых позвоночных (точка 5). Ближайшей очевидной причиной резкого роста МРГ при появлении по-

звоночных являются две полногеномные дубликации (учетверение генома) — ключевое событие ранней эволюции позвоночных, во многом предопределившее возможности дальнейшей прогрессивной эволюции этого подтипа (Putnam et al., 2008).

Следует также учитывать, что в рассматриваемый период происходила коренная перестройка структуры биосферы. Резкий скачок МРГ, возможно, отражает переход МРГ с одной устойчивой (экспоненциальной?) траектории роста на другую, что в какой-то мере может быть следствием (или проявлением) этой глобальной перестройки. Опыт изучения сложных систем с нелинейными положительными обратными связями, развивающихся в больших масштабах времени в соответствии с гиперэкспоненциальным (в том числе гиперболическим) законом, показывает, что во многих случаях такое глобальное эволюционное движение системы при рассмотрении в меньших масштабах распадается на ряд этапов, разделенных “фазовыми переходами”, последовательность которых, собственно, и формирует крупномасштабную гиперэкспоненциальную динамику. При этом у системы обычно существуют так называемые “точки притяжения” — аттракторы, вблизи которых она попадает в условия локального оптимума и может находиться в нем довольно долго, пока внешние воздействия или изменения, подспудно накапливающиеся в самой системе, не вытолкнут ее из очередного аттрактора. После этого, как правило, система снова начинает быстро эволюционировать до тех пор, пока после очередного фазового перехода не попадет в ловушку очередной точки притяжения. Примеры подобного поведения социальной Мир-Системы, эволюционировавшей вплоть до самого последнего времени в направлении гиперэкспоненциального (в том числе гиперболического) увеличения своих базовых макросоциологических параметров (численность народонаселения, уровни технологического, экономического и социокультурного развития, степень урбанизированности и политической сложности и др.) приведены в работах одного из авторов данной статьи и его коллег (Коротаяев, 2006, 2007; Коротаяев, Гринин, 2007; Коротаяев и др., 2007а, б; Гринин, Коротаяев, 2007, 2009). Можно предположить, что и резкое изменение динамики МРГ в венде — кембрии было связано с тем, что в это время система (биосфера) покинула очередной аттрактор (“мир прокариот и одноклеточных эукариот”) и начала перестраиваться, превращаясь в процессе своеобразного фазового перехода в “мир многоклеточных”. Характерно, что в конце неопротерозоя — начале палеозоя отмечается и скачкообразный (на много порядков) рост такого показателя, как максимальный размер живого организма, что связывается, в частности, с резким ростом концентрации свободного кислорода (Payne et al., 2009). Выше мы упоминали, что между размером организма и раз-

мером генома также прослеживается положительная связь.

**О положительных обратных связях, обеспечивающих гиперэкспоненциальный рост МРГ.** По аналогии с динамикой народонаселения и других макросоциологических параметров, а также с динамикой биоразнообразия (Марков, Коротаев, 2009), можно предположить, что гиперэкспоненциальный рост генома обеспечивается сложными положительными обратными связями. Идея о том, что рост генома в ходе эволюции может идти под управлением положительных обратных связей, обсуждается в работе Шарова (Sharov, 2006), где указаны следующие механизмы роста генома, основанные на положительных обратных связях:

1) Геном можно рассматривать как сообщество симбиотических самореплицирующихся элементов, или гиперцикл (Eigen, Schuster, 1977). Например, ген, обеспечивающий более высокую точность репликации ДНК, способствует более точному копированию всех остальных генов в клетке, причем это преимущество распространяется не только на уже существующие гены, но и на те, которые появятся в будущем. Таким образом, существующие гены могут способствовать внедрению в геном и закреплению в нем новых генов.

2) Новые гены часто появляются в результате дупликации или рекомбинации имеющихся генов. Поэтому большой геном предоставляет больше исходного материала для формирования новых генов.

3) Большие геномы поддерживают большее разнообразие метаболических сетей и морфологических элементов, обеспечивая тем самым большее число потенциальных “ниш” для новых генов (Sharov, 2006).

Чтобы понять, какие из предполагаемых положительных обратных связей могут обеспечить гиперэкспоненциальный рост МРГ, следует рассмотреть их более детально. Прежде всего отметим, что экспоненциальная модель соответствует случаю, когда исследуемая переменная растет со скоростью, прямо пропорциональной самому значению переменной ( $dL/dT \sim L$ ), тогда как в двухэкспоненциальной модели скорость роста пропорциональна произведению ее текущего значения на его логарифм ( $dL/dT \sim L \ln(L)$ ). Именно наличие второго сомножителя и определяет гиперэкспоненциальный рост исследуемой величины. Еще большее ускорение роста переменной наблюдается при другом варианте гиперэкспоненциальной динамики, а именно при гиперболическом росте. В этом случае скорость роста переменной пропорциональна квадрату ее значения:  $dL/dT \sim L^2$  (поскольку решением данного дифференциального уравнения является гиперболическая функция, см., например: Капица, 1992, 1999; Korotayev et al., 2006, с. 119–120). Можно заметить, что двухэкспоненциальная динамика

представляет собой промежуточный вариант между экспоненциальной и гиперболической.

Анализ имеющихся данных о механизмах роста генома позволяет предположить, что одни положительные обратные связи, управляющие ростом МРГ, в тенденции могли бы порождать экспоненциальную динамику, другие — гиперболическую, а их совместное действие приводит к генерации промежуточной, двухэкспоненциальной динамики. Рассмотрим один из главных механизмов роста генома — дупликацию фрагментов ДНК. В простейшем случае можно допустить, что вероятность дупликации некоторого участка ДНК фиксированной длины является постоянной величиной. Поскольку количество таких участков пропорционально длине генома, скорость его роста за счет случайных дупликаций тоже должна быть пропорциональна его длине. Это в тенденции должно вести к экспоненциальному росту генома. Однако трехпараметрическая модель указывает на существенное отклонение от экспоненциального закона, особенно начиная с появления многоклеточных.

Гиперэкспоненциальный рост может быть связан с формированием новых функциональных участков ДНК (кодирующих и регуляторных) на основе комбинаторного принципа. Новые гены часто формируются за счет перекомбинирования фрагментов имеющихся генов (Patthy, 1999). Нетрудно показать, что число потенциально возможных новых комбинаций фрагментов (то есть новых генов, которые потенциально могут быть сформированы таким путем) примерно пропорционально квадрату числа имеющихся генов. Представим себе следующую идеализированную ситуацию: каждый ген состоит из двух доменов (функциональных блоков); новые гены формируются путем объединения копий двух доменов, происходящих от двух разных генов. В этой ситуации каждая пара генов потенциально может “породить” четыре новых гена (если не учитывать порядок расположения доменов в новом гене). Следовательно, общее число потенциально возможных новых генов можно вычислить по формуле  $2(N^2 - N)$ , где  $N$  — число генов в геноме. С учетом того обстоятельства, что в геномах свободноживущих организмов число генов довольно велико (от первых тысяч до первых десятков тысяч), формула может быть упрощена:  $2(N^2 - N) \approx 2N^2$ . Если сделать правдоподобное допущение, что скорость роста числа генов прямо пропорциональна числу потенциально возможных новых генов, получаем выражение:  $dN/dt = kN^2$ , что соответствует гиперболическому росту числа генов. Такое же рассуждение можно применить и к некодирующим регуляторным последовательностям, которые тоже, по видимому, могут образовываться путем перекомбинирования фрагментов имеющихся регуляторных последовательностей. Как уже отмечалось, совместное действие факторов, одни из которых способствуют экспоненциальному, а другие — гипербо-

лическому росту МРГ, может вести к промежуточному типу динамики — например, к двойному экспоненциальному росту.

Важно, что описанный механизм формирования новых генов путем перекомбинирования доменов (или экзонов) старых генов распространен у Metazoa гораздо шире, чем у других живых организмов. Основная масса новых генов животных, сформировавшихся таким путем (то есть модульных многодоменных генов), появилась на ранних этапах эволюции Metazoa (Patthy, 1999), что совпадает с периодом резкого ускорения роста МРГ в конце неопротерозоя — начале палеозоя. Отмечено, что большинство этих новых генов имеет непосредственное отношение к специфическим особенностям животных, таким как дифференцировка клеток и тканей, межклеточные взаимодействия, другие аспекты функционирования многоклеточного животного организма как целого. Иными словами, эти гены непосредственно связаны с “кодированием сложности организации” именно в животном царстве (Patthy, 1999). Преобладание данного механизма, способного обеспечить гиперэкспоненциальный рост генома, именно у животных (а не у других многоклеточных), хорошо объясняет, почему этот рост отчетливо выявляется именно в рассматриваемой нами эволюционной линии.

Гиперэкспоненциальный рост МРГ может быть также связан с закономерностями эволюции генных сетей (Колчанов и др., 2000, 2003, 2004). В соответствии с принципом комбинаторного кодирования сложности, усложнение генных сетей идет не только за счет увеличения числа вовлеченных в их работу генов, но и за счет усложнения механизмов их взаимодействия (взаимной регуляции). Очевидно, что потенциальное число межгенных взаимодействий — прямых или опосредованных — должно расти пропорционально квадрату числа генов (поскольку потенциально любой ген может взаимодействовать с любым другим геном в геноме). Для реализации межгенных взаимодействий используются регуляторные участки ДНК (в частности, разнообразные энхансеры, содержащие сайты связывания белков — регуляторов транскрипции). Известно, что рост сложности многоклеточных животных обеспечивался во многом именно появлением новых регуляторных последовательностей (Mikkelsen et al., 2007). Таким образом, рост числа генов должен создавать новые “ниши” для регуляторных последовательностей, причем число ниш должно расти пропорционально квадрату числа генов. Картина еще более усложнится, если учесть, что и сами регуляторные последовательности могут вступать друг с другом в различные взаимоотношения, например, благодаря конкуренции за одни и те же регуляторные белки, причем один и тот же ген может регулироваться целым комплексом разных регуляторных последовательностей и транскрипционных факторов, действующих согласованно (Колчанов и др.,

2000; Takahashi et al., 2008). Важно, что новые регуляторные последовательности часто возникают путем “одомашнивания” МГЭ, которые, таким образом, интенсивно превращаются из “мусорной ДНК” в функционально важные компоненты генома (Miller et al., 1999; Mikkelsen et al., 2007). Более того, новые регуляторные белки — транскрипционные факторы — тоже могут формироваться путем “одомашнивания” генов МГЭ (Lin et al., 2007). Этот механизм роста функциональной части генома за счет превращения “мусорной ДНК” в функциональную, очевидно, должен приводить к тому, что разрастание нефункциональной части генома в долгосрочной перспективе будет способствовать ускорению роста МРГ.

Дополнительно отметим, что рост размера организмов (отчасти сопряженный с ростом МРГ) ведет к сокращению численности популяций, и, следовательно, к усилению генетического дрейфа (то есть к росту вероятности случайной фиксации нейтральных и слабавредных генетических изменений). Поэтому новые копии размножающихся МГЭ имеют больше шансов закрепиться в маленькой популяции, чем в большой. Это может приводить к пассивному разрастанию “мусорной” части генома у организмов с низкой численностью, для которых характерны крупные размеры и большие геномы (Lynch, Conery, 2003). Это отчасти подтверждается тем хорошо известным обстоятельством, что по мере роста размера генома (и организма) растет и доля некодирующих участков. Так, в рассматриваемой линии от прокариот к млекопитающим размер генома возрастает примерно на три порядка (от первых миллионов до первых миллиардов п.н.), тогда как число белок-кодирующих генов возрастает лишь на один порядок (от первых тысяч до первых десятков тысяч). С другой стороны, известно, что у сложных организмов в геноме много функциональных некодирующих участков, которые играют важную роль в эволюции сложности (Miller et al., 1999; Bowen, Jordan, 2002; Volff, 2006; Mikkelsen et al., 2007; Muotri et al., 2007). Иными словами, рост сложности и размера организма может вести к разрастанию как “мусорных”, так и функциональных некодирующих участков.

По-видимому, рост сложности организма иногда может предшествовать росту генома и стимулировать его за счет механизма “ухода от адаптивного конфликта” при удвоении генов. При усложнении организма многие гены могут приобретать дополнительные функции, то есть становиться многофункциональными. Такие гены находятся в состоянии адаптивного конфликта: отбор не может эффективно оптимизировать их для выполнения одной из функций, так как это привело бы к снижению эффективности других функций, выполняемых геном. Когда происходят генные дубликации, появившиеся новые копии генов получают возможность специализироваться для выполнения разных



функций. Это существенно снижает вероятность того, что появившиеся в результате дупликаций новые копии окажутся “лишними” и будут потеряны (Des Marais, Rausher, 2008). Данный механизм играл важную роль в ранней эволюции позвоночных, когда после двух полногеномных дупликаций многие образовавшиеся “лишние” копии генов сохранились у потомков и приобрели новые функции (Putnam et al., 2008).

Можно также предположить существование связи между ростом биоразнообразия и ростом размера генома у самых высокоорганизованных представителей биоты. Рост биоразнообразия ведет к росту сложности и гетерогенности биотической среды. Это создает предпосылки для формирования сложных адаптаций и, соответственно, сложных организмов. Рост сложности организма, в свою очередь, создает предпосылки для дальнейшего роста генома. В компьютерных экспериментах было показано, что у организмов, эволюционирующих в “информационно-насыщенной” (сложной и гетерогенной) среде, происходит рост генома за счет того, что геном “вбирает” в себя информацию о среде и о том, как в ней оптимально функционировать. У организмов, эволюционирующих в “информационно-бедной” среде, размер генома, напротив, со временем сокращается (Adami et al., 2000; Ofria et al., 2003).

Однако следует отметить, что динамика гиперэкспоненциального роста у биоразнообразия и у МРГ в ряду прокариоты — млекопитающих существенно различна. В первом случае период взрывного роста приходится на последние 100–150 млн. лет (конец мезозоя — кайнозой). Первые признаки начала выхода из режима с обострением намечаются во второй половине кайнозоя (Марков, Коротаев, 2009). В динамике МРГ период взрывного роста соответствует палеозою и фактически завершается в триасе с появлением млекопитающих. Таким образом, в динамике роста МРГ выход из режима с гиперэкспоненциальным ускорением произошел уже в начале — середине мезозоя, на 100–200 млн. лет раньше, чем в динамике роста биоразнообразия. Не исключено, что рост генома в какой-то мере способствовал росту биоразнообразия, предоставив дополнительные степени свободы для генетических преобразований (хотя необходимо помнить, что взрывной рост разнообразия в мезозое и кайнозое происходил в основном за счет сравнительно просто устроенных животных, таких как моллюски и насекомые, и в меньшей степени за счет костистых рыб, птиц и млекопитающих).

Морфологическая сложность организмов в рассмотренном ряду, вероятно, тоже росла по крайней мере экспоненциально (или даже гиперэкспоненциально), однако проверить это трудно из-за отсутствия достаточно надежных количественных оце-

нок уровня морфологической сложности в рассмотренных группах организмов.

Замедление роста МРГ после появления млекопитающих не означает, что столь же сильно замедлился и рост сложности организмов. Если понимать фенотип в расширенном смысле, включая сюда не только строение организма, но и поведение, а также экстрасоматические адаптации (бобровые плотины, птичьи гнезда и т.п.) (Dawkins, 1982), то представляется очевидным, что темпы роста сложности фенотипа отнюдь не замедлились в мезозое и кайнозое. Рост коэффициента энцефализации у млекопитающих в кайнозое сопровождался ростом способности к обучению и усложнением поведения (в том числе социального). Более того, этот рост постепенно создавал предпосылки для перерастания биологической эволюции в культурно-социальную (Гринин и др., 2008). Эти явления можно рассматривать как этапы единого ускоряющегося глобального процесса извлечения биотой информации из окружающей среды и сохранения ее на “внешнем носителе”. Роль носителя вначале выполняла ДНК, но после появления более эффективных средств обработки, передачи и хранения информации в виде сложной нервной системы (особенно мозга млекопитающих с развитым неокортексом), а затем речи и письменности, эволюционное давление в направлении увеличения информационной емкости генома, по-видимому, ослабло. Хотя последующие прогрессивные изменения уже не сопровождался ускоренным ростом МРГ, некоторые из них, по-видимому, требовали развития более сложных механизмов регуляции работы генов (Mikkelsen et al., 2007).

В качестве примера механизма, позволяющего увеличить “содержательную” сложность и информационную емкость генома без увеличения его размера, упомянем альтернативный сплайсинг — механизм редактирования молекул матричных РНК, благодаря которому на основе одного и того же гена клетка может синтезировать не один, а несколько разных белков. Показано, что около 94% человеческих генов подвергается альтернативному сплайсингу, в то время как у низших животных альтернативный сплайсинг характерен для меньшинства генов (например, у *C. elegans* — около 15%) (Wang et al., 2008; Pan et al., 2008). Это открытие дает ответ на интригующий вопрос о том, почему в геноме человека примерно столько же генов, что и у гораздо более просто устроенного червя *C. elegans* (около 20000). Как выяснилось, разнообразие белков в человеческом организме в действительности гораздо выше, чем у червя (как и следовало ожидать), однако это разнообразие достигается не за счет разрастания генома, а за счет развития альтернативного сплайсинга.

Имеющиеся данные о механизмах и темпах роста генома пока недостаточны для построения адек-

ватных математических моделей этого процесса. Главная трудность состоит в отсутствии строгих количественных оценок относительного вклада разных механизмов роста генома в общую динамику этого роста.

За рамками настоящей статьи осталось обсуждение возможности экстраполяции полученных модельных кривых в прошлое. Результаты подобной процедуры Шаров ранее интерпретировал как довод в пользу внеземного происхождения жизни (Sharov, 2006). Правомерность и методика подобной экстраполяции, а также выводы, которые можно сделать на ее основе, требуют подробного обсуждения, которому будет посвящена отдельная публикация. Отметим лишь, что имеются доводы как “за”, так и “против” утверждения о том, что резкий “обрыв” модельных кривых в левой части графика (рис. 1), то есть в момент предполагаемого появления прокариотической клетки, может быть истолкован как аргумент в пользу гипотезы внеземного абиогенеза.

## ЗАКЛЮЧЕНИЕ

В ходе развития биосферы, как и в эволюции социума, ряд параметров, отражающих общий уровень развития или сложность системы, изменялись во времени в соответствии с гиперэкспоненциальным (а часто даже гиперболическим) законом. В биосфере к таким параметрам относятся биоразнообразие и избыточный размер генома у самых сложных организмов, а также, вероятно, максимальный размер организмов и максимальный уровень сложности их организации. Гиперэкспоненциальный рост этих показателей позволяет предположить, что эволюция биосферы по пути к общему усложнению, как и социокультурное развитие человечества, регулируется нелинейными положительными обратными связями и является самоускоряющимся процессом. Иными словами, возможно, что причиной прогрессивного роста сложности биологических и социальных систем является сама их сложность.

## СПИСОК ЛИТЕРАТУРЫ

- Гринин Л.Е., Коротаев А.В. Политическое развитие Мир-Системы: формальный и количественный анализ // История и математика: макроисторическая динамика общества и государства / Ред. С.Ю. Малков, Л.Е. Гринин, А.В. Коротаев. М.: КомКнига/URSS, 2007. С. 49–101.
- Гринин Л.Е., Коротаев А.В. Социальная макроэволюция: генезис и трансформации Мир-Системы. М.: ЛИБРОКОМ/URSS, 2009. 568 с.
- Гринин Л.Е., Марков А.В., Коротаев А.В. Макроэволюция в живой природе и обществе. М.: ЛКИ/URSS, 2008. 248 с.
- Капица С.П. Математическая модель роста населения мира // Мат. моделир. 1992. Т. 4. № 6. С. 65–79.
- Капица С.П. Общая теория роста человечества. М.: Наука, 1999. 190 с.
- Колчанов Н.А., Ананько Е.А., Колпаков Ф.А. и др. Генные сети // Мол. биол. 2000. Т. 34. С. 533–544.
- Колчанов Н.А., Суслов В. В., Шумный В.К. Молекулярная эволюция генетических систем // Палеонтол. журн. 2003. № 6. С. 58–71.
- Колчанов Н.А., Суслов В.В., Гунбин К.В. Моделирование биологической эволюции: регулятивные генетические системы и кодирование сложности биологической организации // Вестн. ВОГиС. 2004. Т. 8. № 2. С. 86–89.
- Коротаев А.В. Периодизация истории Мир-Системы и математические макромодели социально-исторических процессов // История и математика. Проблемы периодизации исторических макропроцессов / Ред. Л.Е. Гринин, А.В. Коротаев, С.Ю. Малков. М.: КомКнига/URSS, 2006. С. 116–167.
- Коротаев А.В. Макродинамика урбанизации Мир-Системы: количественный анализ // История и математика: макроисторическая динамика общества и государства / Ред. С.Ю. Малков, Л.Е. Гринин, А.В. Коротаев. М.: КомКнига/URSS, 2007. С. 21–39.
- Коротаев А.В., Гринин Л.Е. Урбанизация и политическое развитие Мир-Системы: сравнительный количественный анализ // История и математика: макроисторическая динамика общества и государства / Ред. С.Ю. Малков, Л.Е. Гринин, А.В. Коротаев. М.: КомКнига/URSS, 2007. с. 102–141.
- Коротаев А.В., Комарова Н.Л., Халтурина Д.А. Законы истории. Вековые циклы и тысячелетние тренды. Демография. Экономика. Войны. М.: УРСС/КомКнига, 2007а. 256 с.
- Коротаев А.В., Малков А.С., Халтурина Д.А. Законы истории. Математическое моделирование развития Мир-Системы. Демография. Экономика. Культура. М.: УРСС/КомКнига, 2007б. 224 с.
- Марков А.В., Коротаев А.В. Гиперболический рост разнообразия морской и континентальной биот фанерозоя и эволюция сообществ // Журн. общ. биол. 2008. Т. 69. С. 175–194.
- Марков А.В., Коротаев А.В. Гиперболический рост в живой природе и обществе. М.: УРСС, 2009. 200 с.
- Марков А.В., Куликов А.М. Происхождение эвкарриот: выводы из анализа белковых гомологий в трех надцарствах живой природы // Палеонтол. журн. 2005. № 4. С. 3–18.
- Розанов А.Ю. Ископаемые бактерии, седиментогенез и ранние стадии эволюции биосферы // Палеонтол. журн. 2003. № 6. С. 41–49.
- Рэфф Р., Кофман Т. Эмбрионы, гены и эволюция. М.: Мир, 1986. 402 с.
- Федонкин М.А. Две летописи жизни: опыт сопоставления (палеобиология и геномика о ранних этапах эволюции биосферы) // Проблемы геологии и минералогии / Ред. А.М. Пыстин. Сыктывкар: Геопринт, 2006. С. 331–350.
- Adami C., Ofria C., Collier T.C. Evolution of biological complexity // Proc. Nat. Acad. Sci. USA. 2000. V. 97. № 9. P. 4463–4468.
- Battistuzzi F.U., Feijao A., Hedges S.B. A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land // BMC Evol. Biology. 2004. V. 4. P. 44.

- Bowen N.J., Jordan I.K.* Transposable elements and the evolution of eukaryotic complexity // *Curr. Iss. Mol. Biol.* 2002. V. 4. № 3. P. 65–76.
- Complete microbial genomes. 2008. <http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi> (accessed 11 Oct 2008).
- Dawkins R.* The extended phenotype: the long reach of the gene. Oxford: Univ. Press, 1982. 310 p.
- Des Marais D.L., Rausher M.D.* Escape from adaptive conflict after duplication in an anthocyanin pathway gene // *Nature*. 2008. V. 454. № 7205. P. 762–765.
- Dietrich F.S., Voegeli S., Brachat S. et al.* The *Aschysia gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome // *Science*. 2004. V. 304. № 5668. P. 304–307.
- Eigen M., Schuster P.* The hypercycle. A principle of natural self-organization. Part A: emergence of the hypercycle // *Naturwiss.* 1977. V. 64. № 11. P. 541–565.
- Eukaryotic genome sequencing projects. 2008. <http://www.ncbi.nlm.nih.gov/genomes/leuks.cgi> (accessed 11 Oct 2008).
- Falcon-Lang H.J., Benton M.J., Stimson M.* Ecology of earliest reptiles inferred from basal Pennsylvanian trackways // *J. Geol. Soc.* 2007. V. 164. № 6. P. 1113–1118.
- Gregory T.R.* Macroevolution, hierarchy theory, and the C-value enigma // *Paleobiol.* 2004a. V. 30. № 2. P. 179–202.
- Gregory T.R.* Insertion–deletion biases and the evolution of genome size // *Gene*. 2004b. V. 324. P. 15–34.
- Gregory T.R.* The C-value enigma in plants and animals: a review of parallels and an appeal for partnership // *Ann. Bot.* 2005. V. 95. № 1. P. 133–146.
- Gregory T.R.* Animal genome size database. 2008. <http://www.genomesize.com>.
- Gregory T.R., Hebert P.D.N.* The modulation of DNA content: proximate causes and ultimate consequences // *Genome Res.* 1999. V. 9. № 4. P. 317–324.
- Guttman M., Amit I., Garber M. et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals // *Nature*. 2009. V. 458. № 7235. P. 253–227.
- Heges S.B., Kumar S.* Genomic clocks and evolutionary timescales // *Trends in Genetics*. 2003. V. 19. № 4. P. 200–206.
- Hughes A.L., Hughes M.K.* Small genomes for better flyers // *Nature*. 1995. V. 377. № 6548. P. 391.
- Ji Q., Luo Z.-X., Yuan C.-X. et al.* The earliest known eutherian mammal // *Nature*. 2002. V. 416. № 6883. P. 816–822.
- Korotayev A., Malkov A., Khalitourina D.* Introduction to social macrodynamics: compact macromodels of the World System Growth. Moscow: URSS, 2006. 128 p.
- Lin R., Ding L., Casola C. et al.* Transposase-derived transcription factors regulate light signaling in *Arabidopsis* // *Science*. 2007. V. 318. № 5854. P. 1302–1305.
- Lynch M., Conery J.S.* The origins of genome complexity // *Science*. 2003. V. 302. № 5649. P. 1401–1404.
- Marshall C., Schultze H.-P.* Relative importance of molecular, neontological, and paleontological data in understanding the biology of the vertebrate invasion of land // *J. Mol. Evol.* 1992. V. 35. № 2. P. 93–101.
- Mikkelsen T.S., Wakefield M.J., Aken B. et al.* Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences // *Nature*. 2007. V. 447. № 7141. P. 167–177.
- Miller W.J., McDonald J.F., Nouaud D., Anxolabehère D.* Molecular domestication – more than a sporadic episode in evolution // *Genetica*. 1999. V. 107. V 1–3. P. 197–207.
- Muotri A.R., Marchetto M.C., Coufal N.G., Gage F.H.* The necessary junk: new functions for transposable elements // *Hum. Mol. Genet.* 2007. V. 16. Spec. № 2. P. R159–R167.
- Nakabachi A., Yamashita A., Toh H. et al.* The 160-kilobase genome of the bacterial endosymbiont *Carsonella* // *Science*. 2006. V. 314. № 5797. P. 267.
- Ochman H.* Genomes on the shrink // *Proc. Nat. Acad. Sci. USA*. 2005. V. 102. № 34. P. 11959–11960.
- Ofria C., Adami C., Collier T.C.* Selective pressures on genomes in molecular evolution // *J. Theor. Biol.* 2003. V. 222. № 4. P. 477–483.
- Organ C.L., Shedlock A.M., Meade A. et al.* Origin of avian genome size and structure in non-avian dinosaurs // *Nature*. 2007. V. 446. № 7132. P. 180–184.
- Pan Q., Shai O., Lee L.J. et al.* Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing // *Nature Genet.* 2008. V. 40. № 12. P. 1413–1415.
- Patthy L.* Genome evolution and the evolution of exon-shuffling – a review // *Gene*. 1999. V. 238. № 1. P. 103–114.
- Payne J.L., Boyer A.G., Brown J.H. et al.* Two-phase increase in the maximum size of life over 3.5 billion years reflects biological innovation and environmental opportunity // *Proc. Nat. Acad. Sci. USA*. 2009. V. 106. № 1. P. 24–27.
- Pellicciari C., Formenti D., Redi C.A., Manfredi Romanini M.G.* DNA content variability in primates // *J. Hum. Evol.* 1982. V. 11. № 2. P. 131–141.
- Putnam N.H., Butts T., Ferrier D.E.K. et al.* The amphioxus genome and the evolution of the chordate karyotype // *Nature*. 2008. V. 453. № 7198. P. 1064–1071.
- Rasmussen B., Fletcher I.R., Brocks J.J., Kilburn M.R.* Reassessing the first appearance of eukaryotes and cyanobacteria // *Nature*. 2008. V. 455. № 7216. P. 1101–1104.
- Sharov A.A.* Genome increase as a clock for the origin and evolution of life // *Biology Direct*. 2006. V. 1. P. 17.
- Shu D.G., Luo H.-L., Conway Morris S. et al.* Lower Cambrian vertebrates from south China // *Nature*. 1999. V. 402. № 6757. P. 42–46.
- Stegemann S., Hartmann S., Ruf S., Bock R.* High-frequency gene transfer from the chloroplast genome to the nucleus // *Proc. Nat. Acad. Sci. USA*. 2003. V. 100. № 15. P. 8828–8833.
- Takahashi K., Hayashi N., Shimokawa T. et al.* Cooperative regulation of Fc receptor gamma-chain gene expression by multiple transcription factors, including Sp1, GABP, and Elf-1 // *J. Biol. Chem.* 2008. V. 283. № 22. P. 15134–15141.
- Thomas C.A.J.* The genetic organization of chromosomes // *Ann. Rev. Genet.* 1971. V. 5. P. 237–256.
- Volff J.N.* Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes // *Bioessays*. 2006. V. 28. № 9. P. 913–922.
- Wang E.T., Sandberg R., Luo S. et al.* Alternative isoform regulation in human tissue transcriptomes // *Nature*. 2008. V. 456. № 7221. P. 470–476.
- Xing J., Witherspoon D.J., Ray D.A. et al.* Mobile DNA elements in primate and human evolution // *Amer. J. Phys. Anthropol.* 2007. Suppl. 45. P. 2–19.

## **Relationship Between Genome Size and Organismal Complexity in the Lineage Leading from Prokaryotes to Mammals**

**A. V. Markov, V. A. Anisimov, A. V. Korotayev**

The lack of a strict relationship between genome size and organismal complexity (level of organization) is largely due to size variability of the facultative part of the genome. However, there is a direct relationship between the level of organization and the minimal genome size (MGS) in the lineage leading from prokaryotes to mammals, in which the tendency towards increasing complexity is especially clear. The dynamics of MGS in this lineage can be adequately described by the model of hyperexponential growth. This implies the existence of nonlinear positive feedbacks that account for the acceleration of MGS growth. The nature of these feedbacks is discussed, including the formation of new genes by means of recombination of the fragments of existing genes, formation of “niches” for new genes in the course of evolution of gene networks, and the expansion of regulatory regions. Hyperexponential growth of different variables related to the level of organization of the biosphere and society (biodiversity, MGS, size and complexity of organisms, world population, technological development, urbanization, etc.) suggests that the evolution of the biosphere and humanity in the direction of increasing complexity is a self-accelerating (autocatalytic) process.

*Key words:* genome size, evolution, level of organization, complexity.