

Модель ранней эволюции геномов и формирования видов¹

С.А. Королев, В.А. Любецкий, А.В. Селиверстов

Институт проблем передачи информации им. А.А. Харкевича, Российская академия наук, Москва,
Россия

Поступила в редколлегию 23.10.2013

Аннотация—Описана упрощённая модель эволюции и формирования устойчивых геномов в среде генов, которая не учитывает события с геном как последовательностью. Обсуждается её приложение к предсказанию появления лекарственно устойчивых штаммов.

КЛЮЧЕВЫЕ СЛОВА: эволюция, кластер, модель.

Формирование устойчивых кластеров в наперёд заданной среде из некоторых элементов, которые в дальнейшем и составляют эти кластеры, — одна из общих задач, связанных с эволюцией систем. Ситуация конкретизируется, если рассмотреть одну из самых фундаментальных задач, относящихся к биологической эволюции: формирование устойчивых геномов в среде генов. Задача исходит из идеи: имеется среда, представленная также геномами, но с активными горизонтальными переносами между ними, что делает их “высоко прозрачными”, т. е. фактически имеется среда из генов, которые также меняются. Среда испытывала динамику, приводящую её к “мало прозрачным” геномам, которые распределяются по кластерам (каждый состоит из похожих по характеристикам геномов); так сформировавшиеся кластеры (“виды”) наблюдаются в настоящее время. Динамика характеризуется одновременно протекающими событиями *популяционного типа*, т. е. относящимся к геномам как цельным элементам, и *геномного типа*, т. е. относящимися к генам, которые разделяются на события с геном как цельным элементом (потери, переносы и т. д.) и с геном как изменяющейся последовательностью в фиксированном алфавите (точечные мутации). Хотя здесь употребляются биологические слова, по существу речь идёт об общей задаче эволюции системы, приводящей к кластеризации.

Опишем упрощённую модель такой эволюции, которая не учитывает события с геном как последовательностью. В ней принимаются следующие положения.

1. Случайно (согласно заданной функции приспособленности f) выбирается самый приспособленный геном, который удваивается и заменяет наименее приспособленный геном. Приспособленность действует в окрестности радиуса r с учётом состояния среды, т. е. степени напряжённости в популяции.
2. В геноме имеются “мобильные” и “ферментные” гены; последние могут в большей или меньшей степени соответствовать заданной среде. Гены переносятся в свой или иной геном (с сохранением в источнике) в зависимости от состояния среды. Переносы и потери генов повышают или понижают приспособленность к среде.

Опишем модель формально. Рассмотрим множество $A \subseteq \mathbb{R}^3$ геномов, привязанных к точкам пространства. Среда задаётся функцией $s(t, x) \in [0, 1]$, которая имеет предел при $t \rightarrow \infty$. Каждой точке $x \in A \subseteq \mathbb{R}^3$ сопоставлена пара чисел (натуральное и рациональное): $m(x, t)$ и $q(x, t) \in [0, 1]$. Здесь $m(x, t)$ — число мобильных генов (представляющих мобильные элементы

¹ Работа выполнена при поддержке грантами Министерства образования и науки РФ, № 14.740.11.1053 и соглашение № 8481.

и т. п.), $q(x, t)$ — доля “ферментных” генов (кодирующих ферменты, транспортёры и т. д.), идеально приспособленных к среде, определяемой тождеством $s \equiv 1$. Поле $\langle m(x, t) \rangle$ обозначим $h(x, t)$. Исходное состояние системы, т. е. это поле в момент $t = 0$, задано. Число “ферментных” генов не меняется в динамике; меняется только доля генов, приспособленных к изменяющейся среде. *Приспособленность* генома к среде описывается функцией, например,

$$f(m, q, s) = 2 - (s - q)^2 - \frac{2}{\pi} \operatorname{arctg}(m) > 0,$$

где $h(x, t) = (m, q)$. В каждой точке $x \in A$ *интенсивность замены генома* $h(x, t)$ на геном $(m', q') \neq (m, q)$ в той же точке x описывается функцией, например,

$$\frac{1}{1 + f(h(x, t), s(x, t))} \left(\sum_{z \in A, z \neq x, h(z, t) = (m', q')} m(z) \frac{f(h(z, t), s(z, t))}{r(x, z)} \right).$$

Вместо $m(z)$ можно написать другую монотонную функцию.

В упрощённом варианте этой модели (рассматриваются только “длинные” и “короткие” геномы, которые превращаются друг в друга, типа модели Изинга) с евклидовым расстоянием в квазистатистическом приближении возникает доменная структура, характерная для ферромагнетиков при понижении температуры ниже точки Кюри.

Полученная модель может применяться для предсказания поведения бактериальных сообществ в условиях, способствующих быстрому накоплению генетических изменений, в том числе при загрязнении окружающей среды радиоактивными материалами или под воздействием антибиотиков. Вызванные нарушения могут приводить к нарушению установившегося в природных экосистемах баланса между организмами с большими и маленькими геномами, что описывается в нашей модели.

Модель можно использовать для предсказания времени формирования новых видов и штаммов, в том числе — появления лекарственно-устойчивых штаммов. Возникновение устойчивости к антибиотику или бактериофагу обычно связано не с изменением генома в целом, а только небольшой части его генов. Такая ситуация моделируется, если вместо больших и маленьких геномов рассмотреть геномы, имеющие большое или маленькое количество копий (паралогов) некоторых генов.

A model of ancestral genome evolution and speciation

S.A. Korolev, V.A. Lyubetsky, A.V. Seliverstov

A simplistic model is presented that describes the evolution and stabilization of genotypes in a gene pool irrespective of the gene sequence evolution. Its application to predict drug-resistant lineages are discussed.

KEYWORDS: evolution, cluster, model.