

УДК 573.55

МОДЕЛИ ПРЕДБИОЛОГИЧЕСКОЙ ЭВОЛЮЦИИ

© 2020 г. В. Г. Редько*

Научно-исследовательский институт системных исследований РАН
Нахимовский просп., 36, к. 1, Москва, 117218 Россия

*E-mail: vgreddko@gmail.com

Поступила в редакцию 08.02.2019 г.

После доработки 08.11.2019 г.

Принята к публикации 18.12.2019 г.

Характеризуются модели предбиологической эволюции, в исследовании которых принимал участие автор настоящей статьи: модель квазивидов, модель гиперциклов, модель сайзеров. Модель квазивидов была предложена Эйгеном в начале 1970-х годов. Эта модель описывает эволюцию популяции макромолекул РНК, кодирующих наследственную информацию. При размножении молекул РНК происходит копирование наследуемой информации. Ошибки в процессе копирования приводят к мутациям РНК. Эволюция популяции приводит к отбору квазивида — такого распределения цепочек РНК, в которое входит как “наилучшая РНК”, размножающаяся с максимальной скоростью, так и близкие к ней цепочки, отличающиеся от этой наилучшей мутационными заменами. Модель гиперциклов была предложена в конце 1970-х годов Эйгеном и Шустером. В гиперцикле к цепочкам РНК добавляются цепочки аминокислот, которые выполняют определенные каталитические функции и вместе с цепочками РНК формируют целостную систему кооперативно взаимодействующих макромолекул. Еще более близкая к простейшим организмам модель — модель сайзеров — была предложена В.А. Ратнером и В.В. Шаминам в начале 1980-х годов. Термин *сайзер* (*syser*) происходит от слов SYstem of SElf-Reproduction¹ (самовоспроизводящаяся система). В сайзер входят полинуклеотидная цепочка, фермент репликации, фермент трансляции и другие ферменты и белки. Также отмечена модель стохастического корректора, которая, как и модель сайзеров, характеризует определенное приближение к реальным биологическим клеткам по сравнению с гиперциклами. Проводится сопоставление модели сайзеров и модели стохастического корректора. В настоящей работе кратко характеризуются все перечисленные модели и излагаются результаты оригинальных расчетов по этим моделям.

DOI: 10.31857/S0044459620020074

В настоящей работе охарактеризованы три формальные модели, относящиеся к проблеме происхождения жизни: модель квазивидов (Eigen, 1971; Эйген, 1973; Eigen, Schuster, 1979; Эйген, Шустер, 1982; Eigen et al., 1989), модель гиперциклов (Eigen, Schuster, 1979; Эйген, Шустер, 1982), модель сайзеров (Ратнер, Шамин, 1980, 1983; White, 1980; Feistel, 1983). Все эти модели формальны и описывают только гипотетические процессы, которые могли идти в процессе происхождения жизни. Но четкость этих моделей может помочь пониманию реальных предбиологических процессов. Более того, в рамках моделей возникает возможность анализа близких эволюционных процессов. В частности, в рамках модели квазивидов, как это кратко представлено ниже, можно получить достаточно общие оценки ско-

рости эволюции для нескольких близких вариантов эволюционных процессов.

В рамках модели сайзеров можно построить модель адаптивного сайзера, характеризующего макромолекулярную систему приспособления к переменной внешней среде.

Модель сайзеров сопоставляется также с идейно близкой к ней моделью стохастического корректора (stochastic corrector model). Последняя была предложена и исследована венгерским теоретиком-эволюционистом Шатмари с сотрудниками (Szathmáry, 1986a, b; Szathmáry, Demeter, 1987).

Статья во многом опирается на оригинальные работы автора по перечисленным моделям.

ИСПОЛЬЗУЕМЫЕ МЕТОДЫ

При исследовании скорости и эффективности эволюционных процессов в модели квазивидов использовались: 1) аналитические математиче-

¹ Если следовать оригинальному звучанию составных частей термина, то получается “сисер”, однако устоялась предложенная ранее транскрипция (*Примечание редактора*).

Таблица 1. Алгоритм эволюции в модели квазивидов

Шаг 0. <i>Формирование начальной популяции</i> $\{S_k(0)\}$. Для каждого $k = 1, \dots, n$ и для каждого $i = 1, \dots, L$ выбираем случайно значение S_{ki} , полагая его равным 0 либо 1.
Шаг 1. <i>Отбор</i> .
Подшаг 1.1. <i>Расчет приспособленностей</i> . Для каждого $k = 1, \dots, n$ вычисляем величину $f(S_k)$.
Подшаг 1.2. <i>Формирование новой популяции</i> $\{S_k(G+1)\}$. Отбор n особей в новую популяцию $\{S_k(G+1)\}$ с вероятностями, пропорциональными $f(S_k)$.
Шаг 2. <i>Мутации особей в новой популяции</i> . Для каждого $k = 1, \dots, n$, для каждого $i = 1, \dots, L$ с вероятностью P_M заменяем символ $S_{ki}(G+1)$ на произвольный символ (0 либо 1). Параметр P_M характеризует интенсивность мутаций.
<i>Организация последовательности поколений</i> . Повторяем шаги 1, 2 для $G = 1, 2, \dots$

ские оценки в определенных приближениях (используемые приближения представлены в соответствующих местах статьи), 2) компьютерные расчеты по моделям, 3) проверка аналитических оценок с помощью компьютерных расчетов.

При исследовании модели сайзеров использовались методы анализа нелинейных динамических систем, методы анализа нелинейных обыкновенных дифференциальных уравнений.

МОДЕЛЬ КВАЗИВИДОВ

Описание основной модели квазивидов с мерой близости Хемминга

Содержательно в модели *квазивидов* рассматривается эволюция популяции цепочек РНК, которые при наличии притока энергии в виде богатых энергией молекул типа АТФ могут реплицироваться, т.е. размножаться (путем саморепликации или репликации с помощью простейших ферментных систем). Ошибки в процессе репликации приводят к мутациям РНК. Эволюция популяции приводит к отбору квазивида — распределения цепочек РНК, в которое входит как “наилучшая РНК” (размножающаяся с максимальной скоростью), так и близкие к ней цепочки, отличающиеся от этой наилучшей мутационными заменами.

Перейдем к формальному описанию модели квазивидов. Рассматриваем эволюцию популяции особей. Считаем, что особи популяции — информационные цепочки, которые могут размножаться и мутировать. Отдельная особь задается своим “генотипом” S_k . Популяция есть множество $\{S_k\}$, состоящее из n особей S_k , $k = 1, \dots, n$, где n — численность популяции. Генотип S_k представляет собой цепочку из L символов, S_{ki} , $i = 1, \dots, L$, где L — длина цепочки. Предполагается, что длина цепочек L и численность популяции n велики ($L, n \gg 1$) и не меняются в ходе эволюции. В основной модели предполагаем, что символы цепочек S_{ki} при-

нимают два значения: $S_{ki} = 0$ либо 1. Генотип S определяет приспособленность $f(S)$ модельных “организмов”. Приспособленности неотрицательны: $f(S) \geq 0$.

В основной формальной модели будем рассматривать простой случай, когда имеется одна оптимальная цепочка (the master sequence) S_M , имеющая максимальную приспособленность. Приспособленность $f(S)$ произвольной особи S определяется расстоянием по Хеммингу $\rho(S, S_M)$ между S и S_M (числом несовпадающих компонент в соответствующих позициях этих цепочек), причем $f(S)$ экспоненциально уменьшается с ростом $\rho(S, S_M)$. Иными словами, чем менее данная цепочка похожа на оптимальную, тем меньше ее приспособленность. Формально считаем:

$$f(S) = \exp[-\beta \rho(S, S_M)], \quad (1)$$

где β — положительный параметр интенсивности отбора.

Эволюционный процесс состоит из последовательности поколений. Новое поколение $\{S_k(G+1)\}$ получается из старого $\{S_k(G)\}$ путем отбора и мутаций цепочек $S_k(G)$; G — номер поколения. Отбор представляет собой формирование нового поколения в соответствии с приспособленностями особей $f(S_k)$: чем больше приспособленность особи, тем больше у нее будет потомков в новом поколении. Мутации состоят в случайной замене символов S_{ki} у особей нового поколения. Начальная популяция ($G = 0$) формируется из n случайных особей. Формальный алгоритм эволюции представлен в табл. 1.

Формирование новой популяции (табл. 1, подшаг 1.2) организуем следующим образом. Представим, что у нас есть игровая рулетка (упрощенная, с закрепленным указателем-стрелкой вместо шарика). Для каждого поколения отмечаем на рулетке n секторов, долю k -го сектора (отнесенную ко всей площади круга) считаем равной $q_k = f_k / \sum_{i=1}^n f_i$

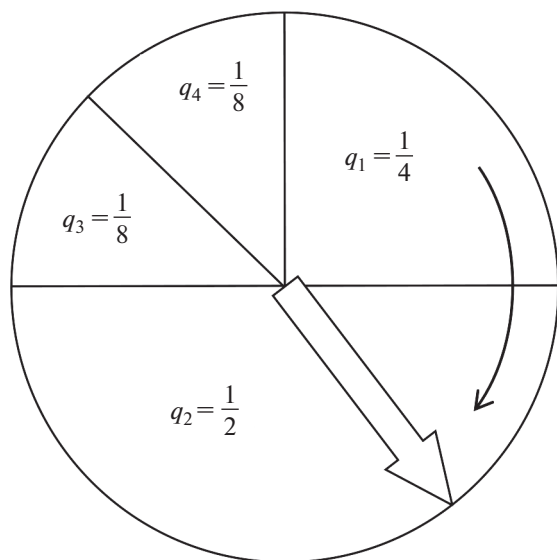


Рис. 1. Пример схемы отбора, согласно которой особи выбираются в популяцию нового поколения с вероятностями q_k , пропорциональными их приспособленностям f_k : $n = 4, f_1 = 2, f_2 = 4, f_3 = 1, f_4 = 1$.

(рис. 1), здесь $f_k = f(S_k)$. Далее n раз крутим рулетку, каждый раз определяем номер сектора, на котором останавливается стрелка, и соответствующую этому номеру особь выбираем в популяцию следующего поколения. Таким образом, в следующее поколение будет отобрано ровно n особей. При этом для каждого вращения рулетки вероятность k -й особи попасть в следующее поколение пропорциональна ее приспособленности f_k . Некоторые особи будут отобраны в новое поколение несколько раз — мы интерпретируем это как появление в новой популяции нескольких потомков данной особи. Такой метод отбора хорошо известен в теории эволюционных алгоритмов. Он называется *рулеточным* (*roulette wheel selection*) или отбором с вероятностями, *пропорциональными приспособленностям* (*fitness proportionate selection*).

Результаты численного моделирования для основной модели

Характер эволюции существенно зависит от численности популяции n . Если n очень велико ($n \gg 2^L$), то численности особей каждого вида можно рассматривать как большие числа, а эволюцию — как детерминированный процесс (2^L — общее число различных цепочек длины L с двоичными символами). В этом случае эволюционная динамика популяции может быть описана системой обыкновенных дифференциальных уравнений, методы исследования которых хорошо известны. Показано, что в отсутствие мутаций происходит отбор особи, имеющей наибольшую приспособленность, а при наличии мутаций в результате эволюции

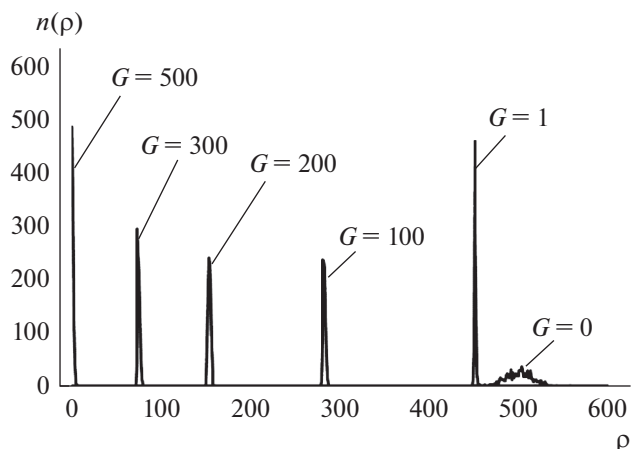


Рис. 2. Эволюция распределения особей при $n = L = 1000, \beta = 1, P_M = 1/L = 0.001$. По оси ординат отложена доля $n(\rho)$ особей, имеющих рассматриваемое значение ρ ; G — номер поколения.

формируется квазивид (подробнее об этом детерминированном случае см. Eigen, 1971; Эйген, 1973; Eigen, Schuster, 1979; Эйген, Шустер, 1982; Eigen et al., 1989; Редько, 2005).

Детерминированный случай может иметь место только при малых длинах цепочек S . При $L > 100$ для любой разумной численности популяции эволюция в модели квазивидов должна рассматриваться как стохастический процесс. Поэтому необходим анализ стохастического случая, для которого $2^L \gg n$. Этот случай и рассмотрен ниже. В этом случае для характеристики основных особенностей эволюции целесообразно использовать количественные оценки и методы компьютерного моделирования.

Далее между параметрами модели предполагаем следующие естественные соотношения: $L, n \gg 1, 2^L \gg n, \beta \gg P_M L, P_M L \ll 1$. Неравенство $2^L \gg n$ означает, что эволюционный процесс существенно стохастичен, число видов цепочек генотипов в конкретной популяции сравнительно невелико, а отдельные виды цепочек вообще отсутствуют в популяции. Соотношение $\beta \gg P_M L$ означает, что интенсивность отбора достаточно велика, а $P_M L \ll 1$ — что интенсивность мутаций сравнительно невелика. В работах автора (Редько, 1986а; Редько, Цой, 2005) был проведен цикл компьютерных расчетов в рамках рассматриваемой модели квазивидов с хемминговой мерой близости между цепочками генотипов. Результаты расчетов демонстрируют, что для выбранных соотношений между параметрами эволюцию можно охарактеризовать следующим образом (пример результатов расчета представлен на рис. 2):

- распределение по ρ в начальной случайной популяции близко к нормальному распределе-

нию со средним $\langle \rho \rangle = L/2$ и дисперсией $L/4$ ($\langle \rho \rangle$ — среднее по популяции расстояние по Хеммингу между цепочками генотипов и оптимальной цепочкой S_M);

- процесс эволюции можно характеризовать двумя стадиями: первой — быстрой и второй — медленной;

- на первой стадии происходит отбор особей с достаточно малыми значениями ρ , и распределение $n(\rho)$ сжимается;

- на второй стадии появление новых особей в популяции ограничено мутациями, из-за этого ограничения скорость уменьшения ρ значительно меньше, чем на первой стадии;

- окончательное распределение характеризует *квазивид* — распределение в окрестности оптимальной цепочки S_M ;

- при малых интенсивностях отбора и мутаций ($1 \gg \beta \gg P_M L$) распределение в квазивиде близко к распределению Пуассона со средним $\langle \rho \rangle = P_M L / \beta$.

Разделение эволюционного процесса на две стадии тесно связано с неравенством $2^L \gg n$. Это неравенство обуславливает то, что новые виды цепочек генотипов в популяции могут появиться только в результате мутаций. А так как интенсивность мутаций сравнительно невелика, вторая стадия длится достаточно долго. Для расчета, представленного на рис. 2, указанное неравенство выполняется с большим запасом: $2^L = 2^{1000} \sim 10^{300}$, $n = 1000$, поэтому вторая стадия занимает практически весь эволюционный процесс.

Для представленного на рис. 2 расчета первая стадия за счет достаточно интенсивного отбора длится всего лишь одно поколение эволюции. Основное время эволюции занимает вторая стадия. Уменьшение величин ρ со временем на второй стадии происходит неравномерно: в начале стадии скорость уменьшения ρ выше, чем в дальнейшем. Это обусловлено тем, что число благоприятных мутаций, приводящих к понижению ρ , постепенно уменьшается. Окончательное распределение (квазивид), близкое к показанному при $G = 500$, формируется в окрестности оптимальной цепочки S_M , для которой $\rho = 0$. Характерное время (число поколений) сходимости всего эволюционного процесса порядка длины гено-типа L .

Так как $2^L \gg n$, рассматриваемые эволюционные процессы имеют стохастический характер. В частности, необходимо учитывать нейтральный отбор, т.е. фиксацию особей, не зависящую от их приспособленностей. Характерное время нейтрального отбора порядка численности популяции: $G_n \sim n$ (Кимура, 1985; Редько, 2018).

Оценка скорости эволюции для основной модели квазивидов

Учитывая качественный характер эволюции (рис. 2) и роль нейтрального отбора, оценим скорость и эффективность рассматриваемого эволюционного процесса в модели квазивидов с хемминговой мерой близости.

Считаем, что численность популяции n достаточно велика:

$$G_n \gg G_T, \quad [1 - (1 - P_M)^L]^n \ll 1, \quad (2)$$

где G_n — характерное время (число поколений) нейтрального отбора, G_T — характерное время сходимости всего эволюционного процесса. Первое неравенство в (2) означает, что мы считаем не слишком сильным влияние нейтрального отбора, так как характерное время всей эволюции меньше или порядка G_n , т.е. эволюция не менее эффективна, чем нейтральный отбор. Второе неравенство соответствует пренебрежению мутационными потерями уже найденных удачных особей в популяции.

При больших L характерное число поколений всего эволюционного процесса, т.е. величина G_T , определяется второй (медленной) стадией эволюции, на которой эволюционный поиск происходит следующим образом. Для того чтобы появились новые особи с меньшими значениями ρ , должно произойти достаточное количество мутаций, а затем — фиксация этих новых особей в результате отбора. Оценим характерное число поколений эволюции G_{-1} , за которое $\langle \rho \rangle$ уменьшается на 1 ($\langle \rho \rangle$ — среднее по популяции расстояние по Хеммингу между цепочками генотипов и оптимальной цепочкой S_M). Оно составляет $G_{-1} \sim G_M + G_S$, где $G_M \sim (LP_M)^{-1}$ — характерное число поколений, за которое особи популяции промутируют, $G_S \sim \beta^{-1}$ — характерное число поколений, за которое особи, для которых $\rho = \langle \rho \rangle - 1$, в результате отбора вытеснят из популяции особи, для которых $\rho = \langle \rho \rangle$. Учитывая, что за время всего эволюционного процесса $\langle \rho \rangle$ должно уменьшиться на величину, приближенно равную $L/2$, полагаем $G_T \sim G_{-1}L$ и имеем:

$$G_T \sim (P_M)^{-1} + L\beta^{-1}. \quad (3)$$

Общее число особей, участвующих в эволюции, составляет $n_{\text{total}} = nG_T$. Оценим величины G_T и n_{total} для заданного L при достаточно разумно выбранных остальных параметрах β , P_M , n . Эти параметры выбираем таким образом, чтобы, по возможности, уменьшить величины G_T и n_{total} . Интенсивность отбора считаем достаточно большой: $\beta \gg P_M L$, — и тогда можно пренебречь вторым слагаемым в (3). Интенсивность мутаций не должна быть слишком высокой, чтобы не было мутационных потерь у найденных уже в процессе

эволюции удачных особей, но и не слишком мала, чтобы мутационный поиск происходил достаточно быстро. Полагаем, что интенсивность мутаций “оптимальна”, т.е. P_M соответствует примерно одной мутации в каждой цепочке в каждом поколении, $P_M \sim L^{-1}$. При такой частоте мутаций, с одной стороны, появление новых генотипов особей в популяции в результате мутаций происходит достаточно быстро и, с другой стороны, можно пренебречь мутационными потерями (выполняется второе неравенство в (2)). Тогда имеем $G_T \sim (P_M)^{-1} \sim L$. Также полагаем, что первое из неравенств (2) выполняется “на пределе” ($G_n \sim G_T$), т.е. для численности популяции имеем оценку:

$$n \sim G_n \sim G_T \sim L. \quad (4)$$

Оценки (4) предполагают минимальную допустимую численность популяции, при которой еще не очень существенны потери особей в результате нейтрального отбора. С учетом сделанных предположений получаем оценки характерного времени (числа поколений) эволюции G_T и общего числа особей, участвующих в эволюции n_{total} :

$$G_T \sim L, \quad n_{\text{total}} \sim L^2. \quad (5)$$

Выражения (5) определяют оценки скорости сходимости всего эволюционного процесса (значение G_T) и эффективности эволюционного поиска оптимальной цепочки S_M (значение n_{total}). Оценки (5) были получены при использовании ряда допущений. Поэтому они были проверены путем компьютерных расчетов (Редько, Цой, 2005; Редько, 2018). Эти расчеты подтвердили справедливость оценок (5). Сходятся с оценками и результаты, приведенные на рис. 2: согласно рисунку, число поколений для сходимости всего эволюционного процесса имеет порядок длины цепочек генотипов L : $G_T \sim L$.

Ближкие эволюционные модели

Хотя оценки (5) получены для основной модели квазивидов с хемминговой мерой близости, эта же методика получения оценок скорости и эффективности эволюции применима к нескольким близким эволюционным моделям. Кратко охарактеризуем эти модели.

Несколько символов в цепочках генотипов

Аналогичную методику В.Г. Редько и Ю.Р. Цой применили к модели квазивидов, в которой символы цепочек принимали λ различных значений (λ — небольшое целое число, например, $\lambda = 4$; Red'ko, Tsoy, 2006). Как и в изложенной выше модели с $\lambda = 2$, предполагаем существование оптимальной цепочки S_M и экспоненциальное убыва-

ние приспособленности произвольной цепочки-особи S с ростом расстояния по Хеммингу $\rho(S, S_M)$ между S и S_M (см. выражение (1)). Изложенная схема получения оценок полностью применима и к этой модели. Установлено, что характерное число поколений эволюции G_T и общее число n_{total} особей, участвующих в эволюции, определяются выражениями:

$$G_T \sim (\lambda - 1)L, \quad n_{\text{total}} \sim [(\lambda - 1)L]^2. \quad (6)$$

Отметим, что (6) обобщают выражения (5): при $\lambda = 2$ (6) переходят в (5).

Модель узкого канала

Еще одна близкая модель — модель узкого канала, в которой рассматривается эволюция популяции цепочек-особей S_k . Как и в модели с хемминговой мерой близости, символы цепочек S_{ki} принимают два значения: $S_{ki} = 0$ либо 1 ; $i = 1, 2, \dots, L$; $k = 1, 2, \dots, n$; L — длина цепочек; n — численность популяции. Также имеется оптимальная цепочка S_M , приспособленность которой максимальна. Однако число цепочек S_k , приспособленность которых отлична от нуля, ограничено. Имеется только $L + 1$ таких цепочек S_k , $k = 0, 1, 2, \dots, L$. Упорядочим нумерацию цепочек: S_0 — оптимальная цепочка, $S_0 = S_M$, число несовпадающих символов в S_k и S_M равно k , а приспособленность цепочки S_k равна $f(S_k) = \exp(-\beta k)$, $k \leq L$. Приспособленности всех остальных цепочек-особей равны 0 . При этом для того, чтобы перейти от цепочки S_k к цепочке S_{k-1} , нужно изменить только один определенный символ. Следовательно, чтобы перейти от цепочки S_L к цепочке S_0 нужно поочередно изменить L символов в строго заданном порядке, т.е. пройти по “узкому каналу”. И в процессе эволюции при поиске S_0 мутации должны произойти именно в этом порядке. При “неправильных” мутациях будут появляться нежизнеспособные цепочки-особи с нулевой приспособленностью. Для наглядности приведем простой пример таких особей ($L = 4$): $S_4 = \{0, 0, 0, 0\}$; $S_3 = \{0, 0, 0, 1\}$; $S_2 = \{0, 0, 1, 1\}$; $S_1 = \{0, 1, 1, 1\}$; $S_0 = \{1, 1, 1, 1\}$. Все остальные цепочки имеют нулевую приспособленность. Видно, что для того, чтобы пройти от S_4 к S_0 , нужно идти по “узкому каналу”, поочередно меняя символы в нужном порядке.

Для модели узкого канала справедливы все рассуждения, сделанные при оценке скорости и эффективности эволюции в модели с хемминговой мерой близости, поэтому оценки (5) применимы и для модели узкого канала. Единственная разница состоит в том, что в основной модели с хемминговой мерой близости число благоприятных мутаций, ведущих к S_M , значительно больше, чем в модели узкого канала, и потому в модели уз-

кого канала скорость эволюции несколько замедляется по сравнению с основной моделью. Такое замедление подтверждается компьютерными расчетами.

Модель эволюционного поиска минимумов энергии спинового стекла

Рассмотренные эволюционные модели соответствуют случаю единственного максимума приспособленности, что является их ограничением. Но с биологической точки зрения важно рассмотреть модели с множеством различных максимумов приспособленности. Кратко охарактеризуем “спин-стекольную” модель эволюции (Редько, 1990а, 2018), в которой число различных максимумов приспособленности велико, оно экспоненциально растет с ростом длины цепочки-особи.

В физике хорошо известна модель спинового стекла, которая описывает систему попарно взаимодействующих спинов (Sherrington, Kirkpatrick, 1975; Kirkpatrick, Sherrington, 1978). Взаимодействия между спинами предполагаются случайными. Формально модель спинового стекла сводится к следующему.

Имеется система S , состоящая из L спинов: $S = \{S_1, S_2, \dots, S_L\}$; число спинов полагается большим, $L \gg 1$. Спины принимают значения $+1$ либо -1 : $S_i = +1, -1$. Взаимодействия между спинами случайны. Энергия $E(S)$ спиновой системы есть

$$E(S) = - \sum_{i,j=1; i < j}^L J_{ij} S_i S_j, \quad (7)$$

где J_{ij} — элементы матрицы случайных взаимодействий между спинами. Величины J_{ij} нормально распределены около нуля, плотность распределения $P(J_{ij})$ задается выражением

$$P(J_{ij}) = (2\pi)^{-1/2} (L-1)^{1/2} \exp[-J_{ij}^2 (L-1)/2]. \quad (8)$$

Для модели (7)–(8) известно, что число M локальных минимумов энергии спинового стекла экспоненциально растет с увеличением L (Tanaka, Edwards, 1980):

$$M \sim \exp(\alpha L), \quad \alpha \approx 0.2. \quad (9)$$

Локальный минимум энергии есть такое состояние системы спинов S_L , для которого смена знака любого одного спина ($S_i \rightarrow -S_i$) приводит к повышению энергии.

Глобальный минимум энергии E_0 приближенно составляет величину $-0.8L$ (Young, Kirkpatrick, 1982):

$$E_0 \approx -0.8L. \quad (10)$$

Согласно (7)–(8) среднее вариации энергии ΔE при смене знака одного спина ($S_i \rightarrow -S_i$) имеет порядок 1 (Редько, 1990а):

$$\langle \Delta E \rangle \sim 1, \quad (11)$$

а среднее значение энергии при случайной ориентации спинов равно 0:

$$\langle E \rangle = 0. \quad (12)$$

Рассмотрим эволюцию популяции, в которой генотипы модельных особей S представляют собой цепочки L символов: $S = (S_1, S_2, \dots, S_L)$, $S_i = +1$ либо -1 . Приспособленность особи S_k определяем как

$$f(S_k) = \exp[-\beta E(S_k)], \quad (13)$$

где $E(S_k)$ — энергия спинового стекла (7), задаваемого генотипом S_k , β — положительный параметр интенсивности отбора.

Рассматриваем следующий эволюционный процесс. В начальном поколении особи (цепочки спинов) случайны. Имеется отбор и мутации особей, отбор производится описанным выше рулеточным методом, мутации представляют собой случайные замены символов $S_i \rightarrow -S_i$. Подчеркнем, что удаление средней энергии в начальном поколении ($G = 0$) от глобального минимума согласно (10), (12) имеет порядок L . Вариация энергии при мутации $S_i \rightarrow -S_i$ согласно (11) порядка 1. Таким образом, эти параметры, определяющие скорость и эффективность эволюционных процессов, близки к таковым для модели квазивидов с хемминговой мерой близости. Поэтому и для модели эволюции в спин-стекольной модели также справедливы оценки скорости и эффективности эволюции (5).

Итак, на основе понятия спиновых стекол можно построить модель эволюции для “организмов”, генотипы которых состоят из множества случайно попарно взаимодействующих между собой элементов. Эволюция может рассматриваться как процесс поиска такой комбинации элементов, которая обеспечивает наиболее эффективную кооперацию элементов генотипа. Качественно “спин-стекольная” модель эволюции аналогична модели с хемминговой мерой близости. Основное отличие состоит в том, что в эволюционном процессе квазивид формируется не в окрестности единственного максимума приспособленности, а в окрестности одного из многочисленных локальных максимумов приспособленности.

Выводы по оценкам скорости и эффективности эволюции

Количественные оценки скорости сходимости и эффективности нескольких сходных эволюционных алгоритмов подтверждены результатами компьютерных расчетов. Наиболее полный анализ проведен для модели квазивидов с хемминговой мерой близости. Для этой модели формулы

(5) оценивают характерное число поколений сходимости к квазивиду и общее число особей, участвующих в эволюции. Эти же оценки приближенно применимы и к остальным рассмотренным моделям.

Оценки получены при следующем выборе параметров:

- 1) $L \gg 1$ — длина цепочек генотипов велика;
- 2) $\beta \gg P_M L$ — интенсивность отбора достаточно велика;
- 3) $P_M \sim L^{-1}$ — мутации “оптимальны” (порядка одной замены символа в генотипе за одно поколение);
- 4) $n \sim L$ — численность популяции порядка длины генотипа; при этом условие пренебрежения нейтральным отбором выполняется “на пределе”: рассматривается минимальная допустимая численность популяции n .

Результат оценок таков:

- 1) характерное число поколений эволюции, обеспечивающей сходимость к квазивиду, порядка длины цепочки генотипа L : $G_T \sim L$;
- 2) общее число особей, участвующих в эволюционном процессе, порядка L^2 : $n_{\text{total}} = n G_T \sim L^2$.

Ясно, что эти оценки для реальной биологической эволюции очень грубые. Тем не менее должна существовать четкая опорная модель, количественно характеризующая процесс биологической эволюции. Модель квазивидов играет роль такой опорной модели. Разные варианты модели показывают, что модель квазивидов охватывает достаточно большое количество схем эволюционных процессов.

МОДЕЛЬ ГИПЕРЦИКЛОВ

Так как модель гиперциклов хорошо известна (Eigen, Schuster, 1979; Эйген, Шустер, 1982), здесь мы только кратко охарактеризуем ее основные черты.

Одна из содержательных интерпретаций рассмотренной выше модели квазивидов — эволюция полинуклеотидных цепочек РНК. Такая эволюция могла иметь место на самых начальных этапах происхождения жизни. Цепочки РНК могли самореплицироваться, однако точность копирования примитивных полинуклеотидов была мала, поэтому длина таких цепочек могла быть только небольшой. Эйген и Шустер предложили модель гиперциклов, в которой к цепочкам РНК добавляются ферменты, которые выполняют определенные каталитические функции и вместе с цепочками РНК формируют целостную систему кооперативно взаимодействующих макромолекул. Ферменты могли способствовать повышению точности копирования, в результате количество информации, которое такие прими-

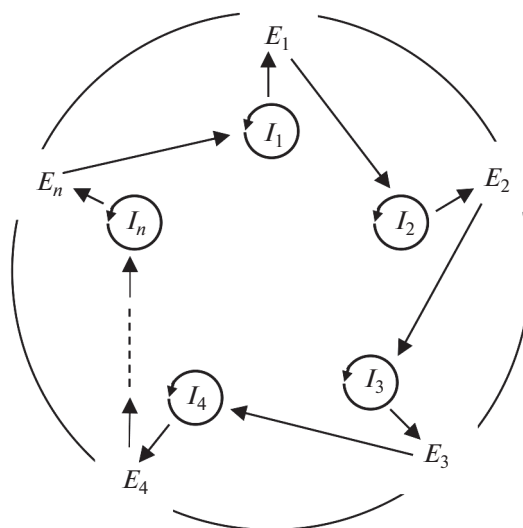


Рис. 3. Структура гиперцикла. I_i — РНК-цепочки, E_i — ферменты репликации ($i = 1, 2, \dots, n$).

тивные “особи” могли передавать потомкам, возросло. Итак, модель гиперциклов интерпретирует гипотетическую стадию эволюции, которая могла следовать за квазивидом. Схема гиперцикла представлена на рис. 3. В гиперцикле РНК и ферменты кооперируются следующим образом: имеются РНК-цепочки (I_i); i -я РНК кодирует i -й фермент E_i ($i = 1, 2, \dots, n$); ферменты циклически катализируют репликацию РНК, а именно E_1 способствует репликации I_2 , E_2 способствует репликации I_3 , ..., E_n способствует репликации I_1 . Кроме того, упомянутые макромолекулы кооперативно обеспечивают примитивную трансляцию, так что информация, закодированная РНК-цепочками, транслируется в структуру ферментов, аналогично обычному механизму трансляции в биологических клетках.

Отметим, что для эффективной конкуренции различные гиперциклы должны быть помещены в отдельные компартменты, т.е. должно быть пространственное разделение гиперциклов разного вида (Файстель и др., 1980). Это условие необходимо для того, чтобы в процессе эволюции отбирались гиперциклы, имеющие лучшие каталитические способности входящих в них ферментов. Если не размещать гиперциклы по компартментам, то селекция определяется начальными условиями, а не каталитическими свойствами макромолекул гиперцикла.

Отметим также и слабую сторону модели гиперциклов. Как было показано в ряде работ (Szathmáry, 1986b; Szathmáry, Demeter, 1987; Zintzaras et al., 2002), гиперциклы слабо устойчивы по отношению к мутациям: при мутациях отдельные РНК-цепочки могут выходить из строя, что будет приводить к разрушению структуры гиперцикла.

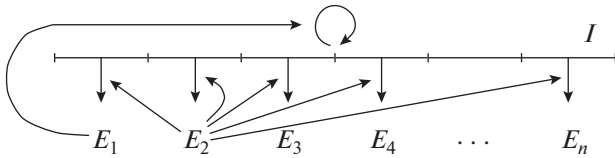


Рис. 4. Общая схема сайзера. I — полинуклеотидная матрица, E_i — ферменты/белки. Круговая стрелка над матрицей символизирует процесс репликации. Стрелки, направленные вертикально вниз, соответствуют процессам трансляции. Стрелки от ферментов E_1 и E_2 поясняют, что эти ферменты катализируют процессы репликации и трансляции соответственно.

МОДЕЛЬ САЙЗЕРОВ

В начале 1980-х годов несколько авторов предложили фактически одну и ту же модель (Ратнер, Шамин, 1980, 1983; White, 1980; Feistel, 1983), которую В.А. Ратнер и В.В. Шамин назвали “Сайзер” (Syser).

Модель *сайзеров*, так же как и модель гиперциклов, может рассматриваться как разумная модель возникновения кооперативных макромолекулярных самовоспроизводящихся систем, содержащих как полинуклеотиды, так и ферменты. Однако сайзеры более сходны с биологическими организмами, чем гиперциклы. Кроме того, сайзеры — достаточно универсальная модель самовоспроизводящейся системы с кибернетической точки зрения. В частности, общая архитектура сайзеров подобна структуре самовоспроизводящихся автоматов, предложенных и исследованных фон Нейманом на заре современной компьютерной эры (Нейман, 1971).

Модель сайзеров может служить основой для интерпретации этапов эволюции от мини-сайзеров (содержащих только необходимые для самовоспроизведения макромолекулы) к простейшим одноклеточным организмам. Например, можно построить модель простейшей молекулярной системы адаптации к переменной внешней среде, которая могла бы возникнуть на предбиологическом уровне (такая модель адаптивного сайзера будет рассмотрена ниже). Далее в этом разделе мы рассмотрим общую схему сайзеров, проанализируем динамику макромолекул в сайзерах и рассмотрим конкуренцию сайзеров разных типов.

Общая схема сайзеров

Сайзер включает в себя (рис. 4): полинуклеотидную матрицу (полинуклеотидную цепочку) I , фермент репликации E_1 , фермент трансляции E_2 и другие возможные белки $E_3, \dots, E_n$. Полинуклеотидная матрица I кодирует белки, фермент репликации E_1 обеспечивает репликацию матрицы I , фермент трансляции E_2 обеспечивает синтез

белков в соответствии с информацией, хранящейся в матрице I .

Сайзеры имеют определенное сходство с биологическими клетками. Матрица I хранит “генетическую” информацию; ферменты E_1 и E_2 представляют собой простые аналоги довольно сложных систем репликации и трансляции биологических клеток. Архитектура сайзеров уже приближается к схеме самовоспроизводящейся системы управления живой клетки (Ратнер, 2001).

Под ферментом репликации E_1 можно подразумевать целую систему ферментов (несколько различных ферментов, которые в совокупности выполняют функцию репликации) — такая замена не изменит существенно приведенное ниже математическое описание сайзеров. То же самое справедливо и для фермента трансляции E_2 .

Предполагается, что как ферменты репликации, так и ферменты трансляции функционируют универсально, т.е. все они “работают” с одинаковой эффективностью, независимо от свойств матриц, которые они реплицируют или с которых они “считывают” белки.

Отметим, что имеет смысл рассматривать динамику макромолекул сайзеров не в гомогенной среде, а предполагать, что каждый вид сайзеров располагается в своем компартменте, в прото-клетке (Chen, 2006; Chen, Walde, 2010). Анализ множества различных сайзеров в гомогенной среде (Редько, 1986б, 2005) показывает, что в этом случае концентрация полинуклеотидных матриц вообще не меняется, а концентрации ферментов при больших временах стремятся к величинам, пропорциональным исходным концентрациям матриц. Это имеет простой смысл: так как ферменты репликации всех видов сайзеров универсальны, относительные концентрации полинуклеотидных матриц разных видов не меняются, а так как ферменты трансляции также универсальны, синтезируемые ферменты “подстраиваются” под матрицы, и их концентрации в конце концов становятся пропорциональными концентрациям матриц. Иными словами, несмотря на разнообразие сайзеров, в гомогенной среде нет существенного их отбора в соответствии с эффективностью ферментов — все определяется только начальными концентрациями матриц. Чтобы выйти из такого “эволюционного застоя” и перейти к естественному отбору лучших “особей”, поместим сайзеры в протоклетки.

Понятие протоклетки активно используется исследователями проблемы происхождения жизни. Под *протоклеткой* понимается капля, содержащая макромолекулы — предшественники макромолекул биологических клеток. Протоклетка окружена мембранной оболочкой. В определенном смысле протоклетки подобны коацерватным каплям (Опарин 1957, 1968). Подробнее о прото-

клетках можно прочитать в работах Чена (Chen, 2006; Chen, Walde, 2010).

Математическое описание сайзеров

Динамика макромолекул

Для описания динамики сайзеров в прото- клетках используем метод, предложенный в рабо- те Р. Файстеля с соавт. (1980) для описания кон- куренции гиперциклов. Предполагаем, что сайзе- ры помещены в протоклетки, каждая из которых есть капля, имеющая полупроницаемую оболоч- ку: она проницаема для малых молекул (таких как энергетически богатые АТФ), но непроницаема для макромолекул (полинуклеотидов и фермен- тов). Предполагаем, что 1) сайзеры помещены в протоклетки; 2) каждая протоклетка включает только один тип сайзеров; 3) объем любой прото- клетки пропорционален числу макромолекул внутри данной протоклетки.

Используя эти предположения, получаем сле- дующие уравнения, характеризующие динамику макромолекул в отдельной протоклетке:

$$\frac{dN_i}{dt} = Vf_i, \quad V = \frac{1}{c} \sum_{l=0}^n N_l, \quad x_i = \frac{N_i}{V}, \quad (14)$$

$$i = 0, 1, 2, \dots, n,$$

где N_i и x_i — число и концентрация макромолекул i -го типа в рассматриваемой протоклетке соот- ветственно; t — время; V — объем протоклетки; c — суммарная концентрация макромолекул в данной протоклетке ($c = \sum_{i=0}^n x_i$); f_i — скорость синтеза макромолекул i -го типа. Индекс $i = 0$ относится к матрице I , остальные индексы i ($= 1, \dots, n$) отно- сятся к ферментам $E_1, \dots, E_n$.

Считаем, что имеется баланс между осмоти- ческим давлением, которое обусловлено избы- точной концентрацией макромолекул внутри протоклетки, и давлением поверхностного натя- жения оболочки протоклетки, так что суммар- ная концентрация макромолекул c постоянна ($\sum_{i=0}^n x_i = c = \text{const}$) (Файстель и др., 1980).

Скорости синтеза макромолекул определяем следующим образом:

$$f_0 = a_0 x_0 x_1, \quad f_j = a_j x_0 x_2, \quad j = 1, \dots, n, \quad (15)$$

где a_i — неотрицательные константы. Выражения (15) подразумевают, что скорость синтеза мат- риц/ферментов пропорциональна концентрации матриц и концентрации фермента реплика- ции/трансляции.

Из уравнений (14) с учетом постоянства сум- марной концентрации макромолекул имеем:

$$\frac{dx_i}{dt} = f_i - x_i \sum_{l=0}^n f_l, \quad i = 0, 1, 2, \dots, n, \quad (16)$$

$$\frac{dV}{dt} = \frac{V}{c} \sum_{l=0}^n f_l. \quad (17)$$

Согласно (15), (16) динамика концентраций макромолекул описывается системой обыкновен- ных нелинейных дифференциальных уравнений. Эти уравнения проанализированы качественными методами в работах В.Г. Редько (2005, 2018). Проведенный анализ показывает, что при доста- точно больших временах концентрации макро- молекул сходятся к устойчивой особой точке $\mathbf{x}^0 = \{x_i^0\}$. Величины x_i^0 определяются выражениями:

$$x_0^0 = \frac{ca_0 a_1}{a_0 a_1 + a_2(a_1 + \dots + a_n)}, \quad (18)$$

$$x_i^0 = \frac{ca_i a_2}{a_0 a_1 + a_2(a_1 + \dots + a_n)}, \quad i = 1, \dots, n.$$

Характерное время τ сходимости к точке $\mathbf{x}^0$ со- ставляет

$$\tau = \left[\frac{1}{c} \sum_{l=0}^n f_l(\mathbf{x}^0) \right]^{-1}. \quad (19)$$

Селективные ценности сайзеров.

Конкуренция сайзеров разного типа

Согласно уравнению (17) скорость роста объе- мов протоклеток для определенного сайзера про- порциональна величине

$$W = \frac{1}{c} \sum_{l=0}^n f_l. \quad (20)$$

Комбинируя формулы (15), (18), (20), находим величины W для равновесной точки $\mathbf{x}^0$:

$$W^0 = \frac{ca_0 a_1 a_2}{a_0 a_1 + a_2(a_1 + \dots + a_n)}. \quad (21)$$

Если есть несколько типов сайзеров, которые имеют различные параметры a_i , то величины W , характеризующие суммарную скорость макромо- лекулярного синтеза сайзеров, можно рассматри- вать как естественную меру приспособленности, определяющую конкуренцию сайзеров разных типов. Величины W (20), будем называть *селек- тивными ценностями* сайзеров. Здесь мы считаем, что времена τ сходимости к устойчивому состоя- нию много меньше других времен, характеризую- щих рассматриваемые процессы. Это означает, что селективные ценности сайзеров W определя- ются формулой (21).

Для анализа конкуренции сайзеров в явном виде будем считать, что 1) каждая протоклетка де- лится пополам, когда ее объем превышает неко-

торую критическую величину, и 2) полный объем всех протоклеток постоянен: $V_T = \sum_k V_k = \text{const}$, где V_k – суммарный объем протоклеток, содержащих сайзеры k -го типа. Тогда вместо выражения (17) имеем:

$$\frac{dV_k}{dt} = W_k V_k - E V_k, \quad (22a)$$

$$E = \frac{1}{V_T} \sum_k W_k V_k, \quad (22b)$$

где селективные ценности W_k определяются параметрами скорости синтеза a_i согласно выражению (21). Селективные ценности различны для различных сайзеров. Параметр E в уравнениях (22a) характеризует однородное разбавление популяции протоклеток (Файстель и др., 1980). Формула (22b) есть следствие условия $dV_T/dt = 0$.

Уравнения (22) описывают конкуренцию сайзеров. Согласно общему решению этих уравнений (Eigen, 1971; Эйген, 1973; Eigen, Schuster, 1979; Эйген, Шустер, 1982; Редько, 2005), в результате конкуренции происходит отбор такого типа сайзеров, который имеет максимальную селективную ценность W .

Итак, в протоклетках происходит отбор сайзеров, обладающих наибольшей суммарной скоростью синтеза макромолекул. Выживают сайзеры, обладающие максимальной “производительностью труда”.

Рассмотренная модель конкуренции сайзеров, располагающихся в протоклетках, показывает, что такая конкуренция обеспечивает условия для *эволюционного прогресса*: если в популяции в результате мутаций появляется новый сайзер, селективная ценность которого выше по сравнению с другими сайзерами, то такой сайзер в результате отбора вытеснит остальные.

Сопоставление модели сайзеров и модели стохастического корректора

Модель *стохастического корректора* (*stochastic corrector model*) была предложена и исследована в работах Шатмари с соавт. (Szathmáry, 1986a, b; Szathmáry, Demeter, 1987; Zintzaras et al., 2002). Суть модели стохастического корректора состоит в следующем:

А) Имеется несколько типов репликаторов, которые входят в протоклетку и конкурируют между собой, но тем не менее взаимно дополняют друг друга. В работах Шатмари (Szathmáry, 1986a, b; Szathmáry, Demeter, 1987) рассматриваются два типа репликаторов. Клетке для существования нужны два типа репликаторов, и, хотя они конкурируют между собой, оба типа нужны клетке рассматриваемого протоорганизма. Репликаторы совместно способствуют формированию моно-

меров, используемых в репликации (Könnyű, Czárán, 2013). Как и в модели сайзеров предполагается, что репликаторы помещены в определенный компартмент – протоклетку.

Б) Число молекул репликаторов в клетке невелико, несколько штук, в процессе жизни клетки число молекул репликаторов увеличивается.

В) Имеется механизм дупликации клеток. Оба репликатора участвуют в производстве определенного полезного для клетки вещества S , когда число молекул этого вещества S превосходит определенный порог, клетка делится пополам на две дочерние клетки, получившиеся при этом репликаторы случайно распределяются по клеткам-потомкам.

Г) Потомки с “правильной” пропорцией между числом репликаторов разных типов живут дальше, потомки с нарушенной пропорцией между репликаторами вымирают. “Правильная” пропорция может состоять в том, что число репликаторов двух входящих в клетку типов примерно равно друг другу.

Процесс деления протоклеток в модели стохастического корректора хорошо проиллюстрирован на сайте (Tyler, 2019).

Важные моменты модели стохастического корректора: 1) небольшое число молекул репликаторов в протоклетке, 2) конкуренция между разными типами репликаторов, аналогичная конкуренции между органеллами в клетках эукариот (Szathmáry, 1986a), но разные типы репликаторов нужны “протоорганизму” так же, как нужны разные органеллы в клетке эукариот.

Фактически модель стохастического корректора и модель сайзеров взаимно дополняют друг друга. Модель сайзеров более четко воспроизводит общую схему репликации хромосомы и трансляции белков по информации, закодированной в хромосоме. Но число макромолекул в разработанных моделях сайзеров велико. Модель стохастического корректора более четко характеризует процесс использования небольшого числа молекул взаимно конкурирующих и дополняющих друг друга репликаторов.

Отметим, что как в модели сайзеров, так и в модели стохастического корректора возможно не только помещение макромолекул в компартмент, но адсорбция макромолекул на минеральной подложке, что может способствовать более совершенной эволюции рассматриваемого “протоорганизма” (Könnyű, Czárán, 2013).

Сайзеры и самовоспроизводящиеся автоматы фон Неймана

На заре современной компьютерной эры фон Нейман предложил и исследовал модель *самовоспроизводящихся автоматов* (Neumann, Burks,

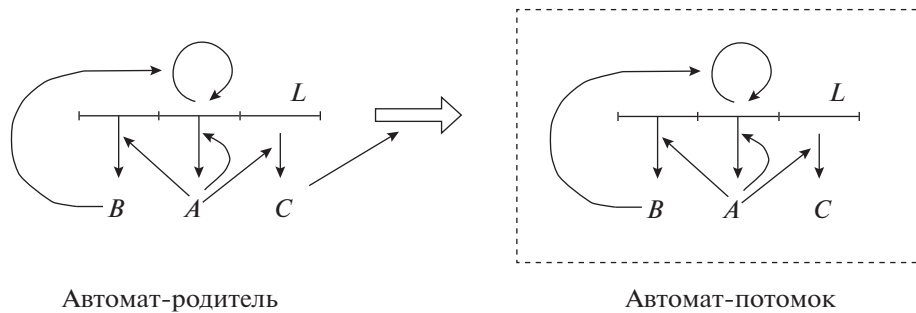


Рис. 5. Схема самовоспроизводящихся автоматов по фон Нейману (1971).

1966; Нейман, 1971). Они состоят из следующих основных компонент (рис. 5): 1) лента L , хранящая информацию, 2) автомат A , предназначенный для изготовления произвольного автомата согласно информации, закодированной в ленте L , 3) автомат B , предназначенный для копирования ленты L , 4) автомат C , координирующий процесс отделения нового изготовленного автомата-потомка от автомата-родителя. Кроме этих необходимых компонент, самовоспроизводящиеся автоматы могут включать и дополнительные автоматы, конструкция которых кодируется лентой L .

Общие архитектура и принципы функционирования сайзеров подобны таковым для самовоспроизводящихся автоматов. Компоненты самовоспроизводящихся автоматов и соответствующие им аналоги сайзеров представлены в табл. 2.

Таким образом, модель сайзеров характеризует достаточно общую структуру самовоспроизводящихся систем.

Модель адаптивного сайзера

Модель адаптивного сайзера характеризует схему адаптации к внешней среде, которая могла возникнуть на предбиологическом уровне или на очень ранних этапах биологической эволюции. Ниже дано краткое изложение модели адаптивного сайзера (Редько, 1990б, 2018), имеющего простую макромолекулярную систему управления, которая включает и выключает синтез определенного фермента в ответ на изменения внеш-

ней среды; схема этой молекулярной адаптации подобна классической схеме регулирования синтеза белков в бактериях по Жакобу и Моно (Jacob, Monod, 1961).

Общая схема и принцип функционирования адаптивного сайзера

Сконструируем схему *адаптивного сайзера*, добавляя необходимые ферменты к схеме минимально возможного сайзера (мини-сайзера). Естественно, конструируем все по подобию с биологическим прототипом (Jacob, Monod, 1961). Схемы мини-сайзера и адаптивного сайзера представлены на рис. 6.

Мини-сайзер (рис. 6а) — это простейший сайзер, содержащий только те макромолекулы, которые необходимы и достаточны для самовоспроизведения, а именно полинуклеотидную матрицу I , фермент репликации E_1 и фермент трансляции E_2 .

В адаптивный сайзер включим два дополнительных фермента: фермент-регулятор E_3 и фермент-адаптер E_4 (рис. 6б). Фермент-регулятор E_3 отслеживает состояние внешней среды и включает или выключает синтез фермента-адаптера E_4 в соответствии с изменениями окружающей среды.

Рассмотрим оба этих сайзера и сравним их селективные свойства, делая естественные предположения о приспособлении адаптивного сайзера к внешней среде. Предполагаем, что имеется два возможных состояния внешней среды: среда A и

Таблица 2. Сравнение архитектур автоматов фон Неймана и сайзеров

Самовоспроизводящиеся автоматы фон Неймана	Сайзеры
Запоминающая лента L	Полинуклеотидная матрица I
Автомат A , предназначенный для изготовления произвольного автомата согласно информации, закодированной в ленте L	Фермент трансляции E_2
Автомат B , копирующий ленту L	Фермент репликации E_1
Автомат C , координирующий процесс отделения автомата-потомка от автомата-родителя	Деление протоклеток в процессе роста сайзеров

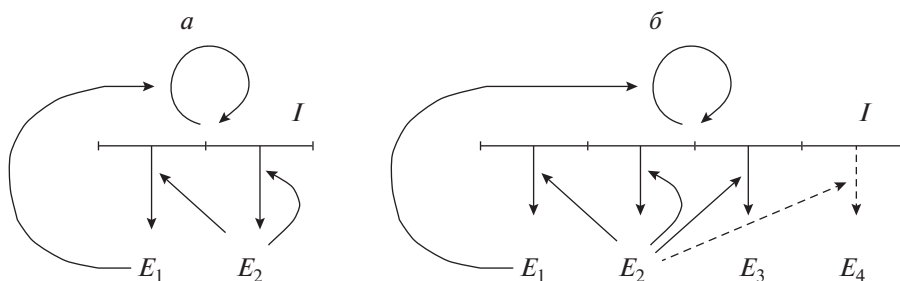


Рис. 6. Схемы мини-сайзера (а) и адаптивного сайзера (б). I — полинуклеотидная матрица, E_1 — фермент репликации, E_2 — фермент трансляции, E_3 — фермент-регулятор, E_4 — фермент-адаптер.

среда B . Среда A — обычная, в которой оба сайзера обладают способностью самовоспроизведения. Среда B — измененное состояние, в котором макромолекулярный синтез имеет место только в адаптивном сайзере и только при наличии фермента-адаптера E_4 . Фермент-регулятор E_3 в адаптивном сайзере синтезируется постоянно с малой интенсивностью в обеих средах и включает синтез адаптера E_4 в среде B и выключает этот синтез в среде A .

В качестве примера двух сред можно рассматривать ситуацию, когда среда A соответствует обычному субстрату S_A , которым могут питаться оба сайзера, а среда B соответствует обычно “несъедобному” химическому субстрату S_B , который в адаптивном сайзере может быть преобразован в обычный “съедобный субстрат” S_A посредством адаптера E_4 . Естественно, что для экономии жизненных ресурсов желательно синтезировать адаптер E_4 только тогда, когда это действительно необходимо, т.е. только в среде B . Чтобы отслеживать состояние окружающей среды, в адаптивный сайзер включен фермент-регулятор E_3 , который синтезируется всегда, но “экономно”, с малой интенсивностью.

Какой сайзер будет выживать в переменной среде? Ответ на этот вопрос далеко не очевиден. Действительно, ясно, что адаптивный сайзер имеет селективные преимущества по сравнению с мини-сайзером в окружающей среде B . Но для переменной среды ситуация сложнее, так как адаптивный сайзер должен синтезировать дополнительный фермент E_3 , и это — дополнительная нагрузка для адаптивного сайзера по сравнению с мини-сайзером. Более того, адаптивному сайзеру необходимо определенное время для перехода от режима одной среды к другой, так что конкуренция сайзеров должна зависеть и от частоты смены сред. Поэтому необходим анализ конкуренции сайзеров, который и представлен ниже.

Математическое описание мини-сайзера и адаптивного сайзера

Проанализируем динамику макромолекул в мини-сайзере и адаптивном сайзере, используя изложенную выше общую схему расчета конкуренции сайзеров. В общем виде эта динамика описывается уравнениями (16), в которых нужно конкретизировать формулы для скорости синтеза макромолекул f_i у рассматриваемых сайзеров в обеих средах. Определим скорости синтеза следующим образом.

Для мини-сайзера полагаем:

$$f_0 = a_0 x_0 x_1, \quad f_i = a_i x_0 x_2, \quad i = 1, 2, \quad \text{в среде } A, \quad (23a)$$

$$f_0 = f_1 = f_2 = 0, \quad \text{в среде } B. \quad (23b)$$

Для адаптивного сайзера полагаем:

$$f_0 = a_0 x_0 x_1, \quad f_i = a_i x_0 x_2, \quad i = 1, 2, 3, \quad \text{в среде } A, \quad (24a)$$

$$f_0 = b_0 x_0 x_1, \quad f_i = b_i x_0 x_2, \quad i = 1, 2, 3, 4, \quad \text{в среде } B, \quad (24b)$$

где b_i — параметры скорости синтеза в среде B . Здесь, как и выше, x_0 — концентрации матриц, x_i — концентрации ферментов E_i ($i = 1, 2, \dots$).

Аналогично (15) эти формулы подразумевают, что скорость синтеза матриц/ферментов пропорциональна концентрации матриц и концентрации фермента репликации/трансляции. Выражение (23b) означает, что в мини-сайзере в среде B синтез макромолекул отсутствует.

Как показано выше, в фиксированной окружающей среде концентрации макромолекул в сайзерах сходятся к устойчивой особой точке x^0 . В переменной среде конкуренция мини- и адаптивного сайзеров зависит от частоты чередования сред. В работах В.Г. Редько (1990б, 2018) проанализированы два предельных случая: 1) редкая смена сред — длительности сред T_A, T_B намного больше характерных времен τ_A, τ_B сходимости к устойчивой точке x^0 : $T_A \gg \tau_A, T_B \gg \tau_B$; 2) частая смена сред — длительности сред T_A, T_B намного

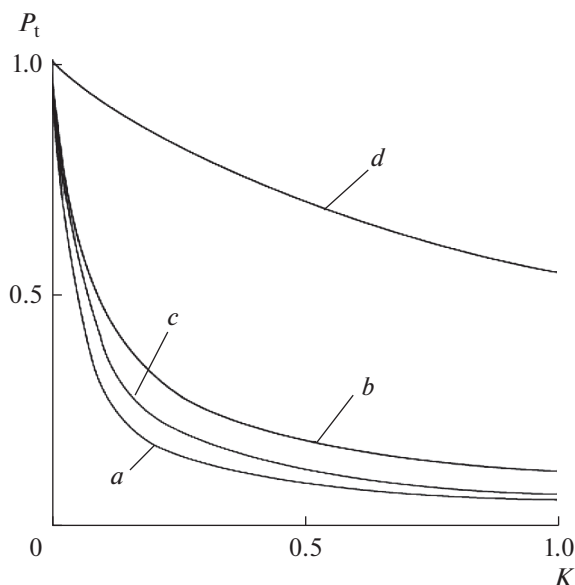


Рис. 7. Зависимость пороговой относительной частоты P_i среды B от отношения макромолекулярных скоростей синтеза K для редкой (a, b) и частой (c, d) смены сред: $\varepsilon_A = 0.033$, $\varepsilon_B = 0.5$ (a, c) или $\varepsilon_B = 2$ (b, d). Адаптивный сайзер имеет селективное преимущество в области выше соответствующей кривой.

меньше характерных времен τ_A , τ_B сходимости к устойчивой точке x^0 : $T_A \ll \tau_A$, $T_B \ll \tau_B$. Был проведен параметрический анализ конкуренции мини-сайзера и адаптивного сайзера. Основные результаты этого анализа показаны на рис. 7. Предполагалось, что процесс смены сред периодический и что относительные скорости макромолекулярного синтеза в адаптивном сайзере (см. формулы (24)) не зависят от окружающей среды: $b_i = Ka_i$, $i = 0, \dots, 3$, где K — константа пропорциональности.

Рис. 7 показывает, при какой относительной частоте P среды B выживает тот или иной сайзер. Если $P > P_i$, то выживает адаптивный сайзер, если $P < P_i$, то выживает мини-сайзер, P_i — пороговое значение относительной частоты среды B . Примеры зависимостей $P_i(K)$ для случаев редкой и частой смены сред для нескольких значений параметров ε_A и ε_B показаны на рис. 7. Эти параметры определяются скоростями синтеза макромолекул (см. выражения (23), (24)) следующим образом:

$$\varepsilon_A = \frac{a_2 a_3}{a_0 a_1 + a_2(a_1 + a_2)}, \quad (25)$$

$$\varepsilon_B = \frac{b_2 b_4}{b_0 b_1 + b_2(b_1 + b_2 + b_3)}.$$

Видно, что для больших ε_B (т.е. для больших скоростей синтеза фермента-адаптера E_4) область выживания адаптивного сайзера для частой сме-

ны сред намного меньше, чем для редкой смены сред.

Таким образом, разработанная модель описывает конкуренцию адаптивного сайзера и мини-сайзера. Адаптивный сайзер имеет селективные преимущества по сравнению с мини-сайзером, но не всегда, а лишь тогда, когда “расходы” (необходимость постоянного синтеза фермента-регулятора E_3), требуемые для обеспечения функционирования молекулярной системы переключения, достаточно малы. Тем самым данная модель демонстрирует, что новое “изобретение” эволюции может быть полезно, но не всегда, а только тогда, когда расходы на поддержание работы этого “изобретения” меньше, чем получаемая польза.

Рассмотренная система регулирования синтеза ферментов могла бы быть первой, возникшей в процессе эволюции системой адаптации к переменной внешней среде.

ЗАКЛЮЧЕНИЕ

Рассмотренные модели характеризуют постепенно усложняющиеся макромолекулярные предшественники биологических организмов.

Квазивиды характеризуют эволюцию простых информационных единиц — РНК-цепочек небольшой длины.

Гиперциклы — это макромолекулярные системы, учитывающие кооперацию цепочек РНК с простейшими ферментами, кодируемыми этими цепочками.

Сайзеры — это молекулярные “конструкции”, достаточно близкие к самовоспроизводящейся молекулярно-генетической системе живой клетки.

Представленные модели — это только возможные элементы эволюции предбиологических систем. Тем не менее они позволяют четко и на серьезном математическом уровне представить гипотетические этапы предбиологической эволюции.

Настоящая работа выполнена в рамках государственного задания по проведению фундаментальных научных исследований по теме “Исследование нейроморфных систем обработки больших данных и технологии их изготовления” (№ 0065-2019-0003). Автор благодарен рецензенту и научному редактору за полезные замечания и рекомендации, способствовавшие улучшению статьи.

СПИСОК ЛИТЕРАТУРЫ

Кимура М., 1985. Молекулярная эволюция: теория нейтральности. М.: Мир. 394 с.

- Опарин А.И., 1957. Возникновение жизни на Земле. 3-е изд., перераб. М.: Изд-во АН СССР. 459 с.
- Опарин А.И., 1968. Жизнь, ее природа, происхождение и развитие. 2-е доп. изд. М.: Наука. 173 с.
- Нейман Дж., фон, 1971. Теория самовоспроизводящихся автоматов. М.: Мир. 382 с.
- Ратнер В.А., 2001. Молекулярно-генетическая система управления // Природа. № 3. С. 16–22.
- Ратнер В.А., Шамин В.В., 1980. Сайзеры: моделирование фундаментальных особенностей молекулярно-биологической организации // Математические модели эволюционной генетики. Новосибирск: ИЦИГ. С. 66–126.
- Ратнер В.А., Шамин В.В., 1983. Сайзеры: моделирование фундаментальных особенностей молекулярно-биологической организации. Соответствие общих свойств и конструктивных особенностей коллективов макромолекул // Журн. общ. биологии. Т. 44. № 1. С. 51–61.
- Редько В.Г., 1986а. Оценка скорости эволюции в моделях Эйгена и Куна // Биофизика. Т. 31. № 3. С. 511–516.
- Редько В.Г., 1986б. Поведение каталитически взаимодействующих макромолекул (сайзеров) в коацерватах // Биофизика. Т. 31. № 4. С. 705–707.
- Редько В.Г., 1990а. Спиновые стекла и эволюция // Биофизика. Т. 35. № 5. С. 831–834.
- Редько В.Г., 1990б. Адаптивный сайзер // Биофизика. Т. 35. № 6. С. 1007–1011.
- Редько В.Г., 2005. Эволюция, нейронные сети, интеллект. Модели и концепции эволюционной кибернетики. М.: КомКнига/URSS. 224 с.
- Редько В.Г., 2018. Моделирование когнитивной эволюции: На пути к теории эволюционного происхождения мышления. Изд. 2-е, испр. и доп. М.: ЛЕНАНД/URSS. 264 с.
- Редько В.Г., Цой Ю.Р., 2005. Оценка эффективности эволюционных алгоритмов // ДАН. Т. 404. № 3. С. 312–315.
- Файстель Р., Романовский Ю.М., Васильев В.А., 1980. Эволюция гиперциклов Эйгена, протекающих в коацерватах // Биофизика. Т. 25. № 5. С. 882–887.
- Эйген М., 1973. Самоорганизация материи и эволюция биологических макромолекул. М.: Мир. 224 с.
- Эйген М., Шустер П., 1982. Гиперцикл. Принципы самоорганизации макромолекул. М.: Мир. 270 с.
- Chen I.A., 2006. The emergence of cells during the origin of life // Science. V. 314. № 5805. P. 1558–1559.
- Chen I.A., Walde P., 2010. From self-assembled vesicles to protocells // Cold Spring Harb. Perspect. Biol. V. 2. № 7. P. a002170.
- Eigen M., 1971. Selforganization of matter and the evolution of biological macromolecules // Naturwissenschaften. V. 58. № 10. P. 465–523.
- Eigen M., Schuster P., 1979. The Hypercycle: A Principle of Natural Self-Organization. Berlin: Springer Verlag. 92 p.
- Eigen M., McCaskill J., Schuster P., 1989. The molecular quasi-species // Adv. Chem. Phys. V. 75. P. 149–263.
- Feistel R., 1983. On the evolution of biological macromolecules. Precellular organization. 4. Holobiotic competition // Studia Biophys. V. 93. № 2. P. 113–128.
- Jacob F., Monod J., 1961. Genetic regulatory mechanisms in the synthesis of proteins // J. Mol. Biol. V. 3. № 3. P. 318–356.
- Kirkpatrick S., Sherrington D., 1978. Infinite range model of spin-glass // Phys. Rev. B. V. 17. № 11. P. 4384–4403.
- Könnnyű B., Czárán T., 2013. Spatial aspects of prebiotic replicator coexistence and community stability in a surface-bound RNA world model // BMC Evol. Biol. V. 13. P. 204.
- Neumann J., von, Burks A.W., 1966. Theory of Self-Reproducing Automata. Urbana: Univ. of Illinois Press. 418 p.
- Red'ko V.G., Tsoy Yu.R., 2006. Estimation of the evolution speed for the quasispecies model: Arbitrary alphabet case // Artificial Intelligence and Soft Computing – ICAISC 2006. 8th Int. Conf. / Eds Rutkowski L., Tadeusiewicz R., Zadeh L., Zurada J. Zakopane, Poland: Springer. P. 460–469.
- Sherrington D., Kirkpatrick S., 1975. Solvable model of spin-glass // Phys. Rev. Lett. V. 35. № 26. P. 1792–1796.
- Szathmáry E., 1986а. The eukaryotic cell as an information integrator // Endocytobiosis Cell Res. V. 3. № 2. P. 113–132.
- Szathmáry E., 1986b. Some remarks on hypercycles and the stochastic corrector model // Endocytobiosis Cell Res. V. 3. № 3. P. 337–339.
- Szathmáry E., Demeter L., 1987. Group selection of early replicators and the origin of life // J. Theor. Biol. V. 128. № 4. P. 463–486.
- Tanaka F., Edwards S.F., 1980. Analytic theory of the ground state of a spin glass: 1. Ising spin glass // J. Phys. F. Metal Phys. V. 10. № 12. P. 2769–2778.
- Tyler T., 2019. The Origin of Life. Stochastic corrector. www.originoflife.net/stochastic_corrector.
- White D.H., 1980. A theory for the origin of a self-replicating chemical system. I: Natural selection of the autogen from short random oligomers // J. Mol. Evol. V. 16. № 2. P. 121–147.
- Young A.P., Kirkpatrick S., 1982. Low-temperature behavior of the infinite-range Ising spin-glass: Exact statistical mechanics for small samples // Phys. Rev. B. V. 25. № 1. P. 440–451.
- Zintzaras E., Mauro S., Szathmáry E., 2002. “Living” under the challenge of information decay: The stochastic corrector model versus hypercycles // J. Theor. Biol. V. 217. № 2. P. 167–181.

Models of prebiotic evolution

V. G. Red'ko*

*Scientific Research Institute for System Analysis, RAS
Nakhimovskiy pr., 36-1, Moscow, 117218 Russia*

**e-mail: vgreddko@gmail.com*

Models of prebiotic evolution, investigated by the author, are characterized, namely 1) the quasispecies model, 2) the hypercycle model, and 3) the syser model. The quasispecies model was proposed by Manfred Eigen, this model describes an evolution of population of the RNA macromolecules that could encode hereditary information. During the reproduction of RNA molecules, the inherited information is copied. Errors in the copying process lead to RNA mutations. The evolution of such a population leads to the selection of a quasispecies. The quasispecies is such a distribution of RNA chains that include both the “best RNA” (i.e., reproducing at maximal speed) and similar RNA chains that differ slightly from this best RNA by mutational substitutions. The hypercycle model was proposed in the late 1970s by Manfred Eigen and Peter Schuster. The hypercycle includes both the RNA chains and chains of amino acids, which perform certain catalytic functions. Both types of chains form a system of cooperatively interacting macromolecules. The syser model was proposed by Vadim Ratner and Vladimir Shamin. The term syser is abbreviation from the SYstem of SElf-Reproduction. Sysers are closer to simplest organisms as compared with hypercycles. A syser includes the polynucleotide chain, replication enzyme, translation enzyme, and other enzymes/proteins. A stochastic correlator model is also outlined, which, like the syser model, characterizes a certain approach to real biological cells. The syser model and the stochastic correlator model are compared. The current study characterizes all the noted models. The brief mathematical description of the models and the main results of computer simulations corresponding to the models are presented.