

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ  
ЭВОЛЮЦИИ МИКРОБНОЙ ПОПУЛЯЦИИ: ПРЕОДОЛЕНИЕ  
ЛОКАЛЬНЫХ МИНИМУМОВ ПРИ ДОСТИЖЕНИИ ПИКА  
НА ЛАНДШАФТЕ ПРИСПОСОБЛЕННОСТИ

© 2020 г. С. А. Лашин<sup>1,2,\*</sup>, З. С. Мустафин<sup>1</sup>, А. И. Клименко<sup>1</sup>,  
Д. А. Афонников<sup>1,2</sup>, Ю. Г. Матушкин<sup>1,2</sup>

<sup>1</sup>Федеральный исследовательский центр Институт цитологии и генетики  
Сибирского отделения Российской академии наук, Новосибирск, 630090 Россия

<sup>2</sup>Новосибирский национальный исследовательский государственный университет,  
кафедра информационной биологии, Новосибирск, 630090 Россия

\*e-mail: lashin@bionet.nsc.ru

Поступила в редакцию 11.02.2019 г.

После доработки 12.04.2019 г.

Принята к публикации 22.08.2019 г.

В данном исследовании с помощью компьютерного моделирования анализируются механизмы пересечения долин пониженной приспособленности популяцией гаплоидных микроорганизмов. Приспособленность этих микроорганизмов зависит от аллельных значений в двух различных локусах и определяется сложным ландшафтом, форма которого соответствует паттерну “окруженная ровом гора в поле”. Изучено влияние различных биологических факторов на эволюционную перспективу колоний микроорганизмов, скорость размножения которых определяется некоторым белком, состоящим из двух субъединиц, кодируемых в разных локусах. Рассматриваются факторы как молекулярно-генетической (частота мутаций, афинность субъединиц), популяционной (ландшафт функции приспособленности и плотность популяции), так и экологической (концентрация доступного субстрата в среде обитания, скорость протока) природы. Наши результаты демонстрируют, что разница в приспособленности для различных аллельных комбинаций, определяя форму ландшафта приспособленности, задает оптимальную для преодоления его долин частоту мутаций, открывая окно возможностей для эволюции популяции в сторону состояния наибольшей средней приспособленности. Более того, в зависимости от характера ландшафта приспособленности оптимальными для достижения пика оказываются либо градуальный, либо сальтационный эволюционные режимы.

*Ключевые слова:* ландшафт приспособленности, долина приспособленности, микробное сообщество, экологическое моделирование, эволюция, частота мутаций.

DOI: 10.31857/S0016675820020071

Большинство фенотипических признаков живых организмов определяются аллельными комбинациями более чем одного локуса. В связи с этим для адекватного описания эффекта мутаций при моделировании эволюционной динамики необходимо учитывать такие важные факторы как межлокусные взаимодействия (эпистаз), экспрессию признаков и их влияние на приспособленность особей [1]. В результате наблюдаются такие интересные эволюционные эффекты как компенсаторные нейтрализующие мутации, впервые рассмотренные Кимурой в работе [2]. Математическая модель, предложенная в этой работе, описывает эволюцию двух генов, мутации в каждом из которых поодиночке вредоносны, но в случае если му-

тированы оба гена сохраняется приспособленность на уровне дикого типа. Существенным затруднением, с которым сталкиваются модели такого типа, оказывается способность популяции преодолевать промежуточные эволюционные стадии низкой приспособленности после приобретения отдельных вредных мутаций. В более широких терминах эта проблема может быть сформулирована как проблема смещающегося баланса (shifting balance) в эволюционной модели, описываемой ландшафтом приспособленности [3–5]. С практической точки зрения такие модели оказываются полезны для направленной эволюции ферментов в сторону улучшенной стабильности и активности [6], терапии онкологических заболе-

ваний [7], процессов видообразования [8], а также для понимания механизмов устойчивости и выживаемости вирусов [9]. Интерес к данной проблеме постоянно растет на фоне использования различных методов математического и компьютерного моделирования, позволяющих исследователям подробно изучать различные аспекты таких моделей [10]. В настоящее время этот интерес также подогревается попытками с помощью моделей объяснить большие объемы данных, собранных недавно в результате экспериментов по высокопроизводительному секвенированию геномов [11, 12], что, в свою очередь, позволяет пересмотреть классические и относительно новые модели генетики и эволюции популяций.

Эти эволюционные модели предназначены для ответа на следующий основной вопрос: какие факторы позволяют популяциям пересекать “долины”, соответствующие пониженной приспособленности на ландшафте приспособленности, при их переходе от одного пика к другому? Ранее было показано, что наиболее важными факторами, ответственными за успешное продвижение популяции по ландшафтными долинам, являются: размер популяции [13–15], скорость фиксации мутаций [13, 16], частота рекомбинации в популяциях, использующих половое размножение [17], а также форма ландшафта приспособленности [18]. Большинство из приведенных выше исследований рассматривают генотип особи как набор аллелей из дискретного пространства аллелей (см., например, [19]).

В настоящей работе мы проанализировали механизмы пересечения долин ландшафта приспособленности популяцией гаплоидных микроорганизмов, приспособленность которых зависела от аллельных значений в двух локусах и определялась сложным ландшафтом, форма которого может быть описана как “окруженная ровом гора в поле” (см. рис. 2, в дальнейшем мы будем считать, что ров – это частный случай более общего термина “долина”). Эта форма является расширением ландшафта “Гора Фудзи” [20], использованного для аппроксимации приспособленности мутантных ферментов. На молекулярном уровне эта двухлокусная модель описывает взаимодействие двух разных субъединиц, образующих фермент, способность которого расщеплять неспецифический субстрат, являющийся источником энергии микроорганизма, определяет его репродуктивную эффективность. Ширина рва между областями с высокой приспособленностью на ландшафте достаточно велика: требуется, как минимум, несколько последовательных мутаций, приводящих к сниженной приспособленности, чтобы преодолеть этот ров.

Мы рассмотрели влияние различных генетических и экологических факторов на эволюцион-

ную перспективу микробной популяции, состоящей из клеток, чей генотип определяется двумя взаимодействующими локусами. Эти факторы включали в себя как молекулярно-генетическую (частота мутаций, афинность субъединиц), популяционную (ландшафт функции приспособленности), так и экологическую (концентрация доступного субстрата в среде обитания) составляющие.

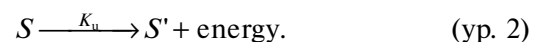
Наши результаты показывают, что успех в преодолении долин приспособленности определяется как различной приспособленностью разных аллельных комбинаций, так и частотой мутаций. Для каждого ландшафта приспособленности существует своя оптимальная частота мутаций. Более того, в зависимости от характера ландшафта приспособленности оптимальными для достижения пика оказываются либо градуальный (постепенная дивергенция между родственными таксонами путем равномерного накопления небольших изменений – “природа не делает скачков” [21, 22]), либо сальтационный (быстрые и значительные преобразования предковых форм [21, 22]) эволюционные режимы.

## ОПИСАНИЕ МОДЕЛИ

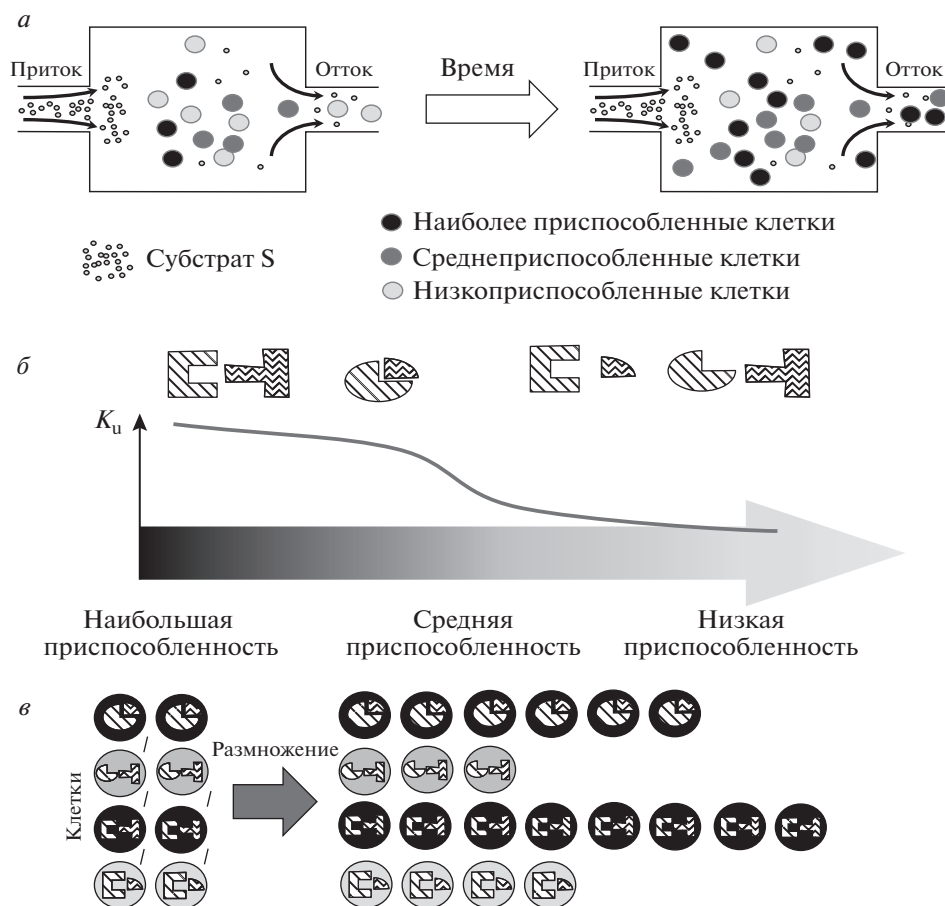
Мы построили модель эволюционной динамики в популяции гаплоидных микроорганизмов, состоящей из клеток одного вида, населяющих резервуар с постоянным притоком неспецифического субстрата  $S$ , являющегося источником энергии для живущих в нем клеток (рис. 1,а). Приспособленность клеток определяется размером прироста популяции в единицу времени и зависит от эффективности утилизации субстрата, доступности субстрата и смертности. Популяционная динамика ( $\Delta P$ ) описывается уравнением:

$$\Delta P = PK_u S_{\text{uptaken}} - k_{\text{flow}} P - k_{\text{dens}} P^2, \quad (\text{ур. 1})$$

где  $P$  – размер популяции,  $S_{\text{uptaken}}$  – количество субстрата, поглощенного клетками из окружающей среды;  $k_{\text{flow}}$  и  $k_{\text{dens}}$  – коэффициенты протока и плотности соответственно. В данном уравнении лишь один член зависит от генотипа, а именно способность клеток утилизировать субстраты, добывая энергию, что определяется параметром утилизации  $K_u$ :



Эффективность утилизации субстрата  $K_u$  определяется некоторым ферментом  $E$ , состоящим из двух субъединиц  $E_1$  и  $E_2$ , кодируемых генами, расположенными в двух различных локусах  $E_1$  и  $E_2$  (см. рис. 1,б, в). Таким образом,  $K_u$  зависит от аллельных вариантов обоих локусов; обозначим эти варианты  $e_1$  и  $e_2$ :



**Рис. 1.** Общая схема модели. *a* – схема моделируемой системы: проточный резервуар, на вход которого подается неспецифический субстрат – источник энергии. Резервуар населяет гетерогенная микробная популяция (клетки разных генотипов показаны разным цветом). С течением времени меняется как размер популяции, так и соотношения между представленностями клеток различных генотипов; *б* – схема функции  $K_u$ , демонстрирующая изменение приспособленности в зависимости от комбинации аллельных вариантов  $E_1$  (косая штриховка) и  $E_2$  (волнистая штриховка); *в* – более приспособленные клетки дают больше потомства за единицу времени и в долгосрочной перспективе вытесняют менее приспособленных.

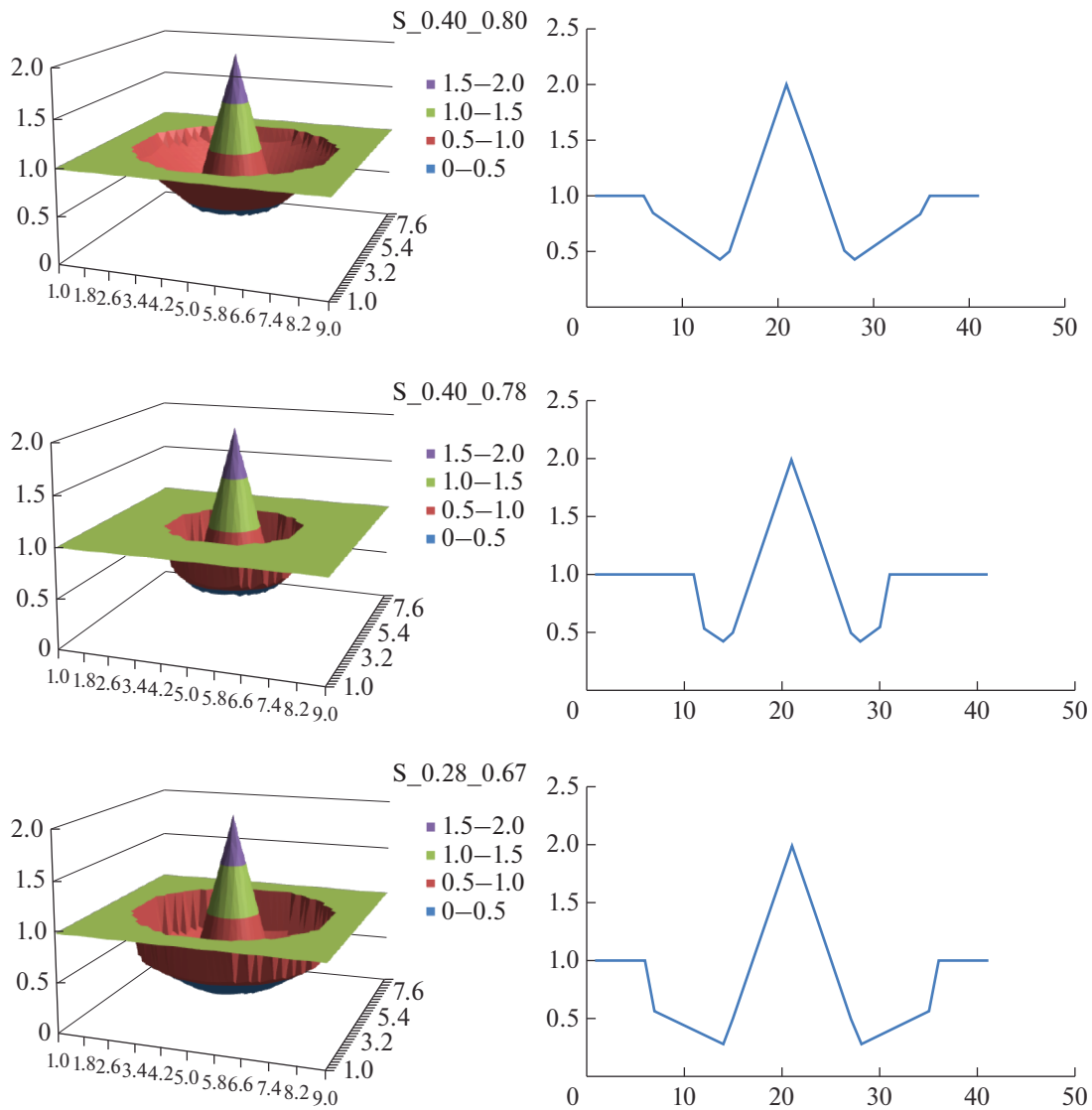
$$K_u = F(e_1, e_2). \quad (\text{ур. 3})$$

В нашей модели мы берем значения аллелей из заданных отрезков  $[M_{i, \min}; M_{i, \max}]$  ( $i = 1, 2$ ). На молекулярном уровне это соответствует различным значениям физико-химических характеристик аминокислотных остатков, участвующих во взаимодействиях между белковыми субъединицами (например, заряд, объем или полярность). Мы можем рассматривать генетические последовательности, кодирующие аминокислоты в тех позициях белка, которые формируют поверхность взаимодействия субъединиц, как генетические локусы и представить их вышеописанным образом. С другой стороны, эти локусы можно рассматривать и более общо – как множества позиций аминокислотных последовательностей в субъединицах фермента, которые образуют поверхность взаимодействия субъединиц. В послед-

нем случае аллельное значение может соответствовать определенной интегральной характеристике аминокислотных остатков, формирующих поверхность взаимодействия, например их совокупный заряд или объем [23].

#### Ландшафты приспособленности

Функция  $F$  описывает зависимость эффективности утилизации субстрата от значений аллелей (ур. 3). Хотя ее можно определить параметрически, в данном исследовании мы используем таблично-заданные функции для значений  $K_u$ . Мы задаем квадратную сетку соответствующих значений аллелей,  $e_1, e_2$  ( $1 \leq e_1 \leq 9; 1 \leq e_2 \leq 9$ ), с шагом сетки, равным 0.2. Общий размер сетки равен  $41 \times 41 = 1681$  – числу возможных аллельных комбинаций. На данной сетке задается поверхность, высота которой определяется значениями из таб-



**Рис. 2.** Варианты ландшафта, задаваемые функцией  $K_u$ . Слева приведены пространственные изображения ландшафтов (оси  $O_x$ ,  $O_y$  соответствуют аллельным состояниям локусов  $E_1$  и  $E_2$ ), справа приведены сечения плоскостью, перпендикулярной плоскости  $xy$  и параллельной оси  $O_x$ . Мнемонические обозначения ландшафтов вида  $S\_A\_B$  расшифровываются в табл. 1.

личной функции  $K_u$ , аппроксимирующей  $F$  – поверхность произвольной формы. Однако следует заметить, что мы не закладываем в нашу модель никаких предзаданных отношений между значениями  $K_u$  для каждой пары аллелей и приспособленностью организма, обладающего данным генотипом. Реальная приспособленность дополнительно обуславливается состоянием популяции (находится она в фазе роста или в стационарной фазе), ее плотностью, а также доступностью субстрата и влиянием скорости протока.

В данном исследовании мы рассмотрели множество  $F(e_1, e_2)$  поверхностей с пиком и долиной различной формы, соответствующих паттерну “окруженная рвом гора в поле” (см. пример на

рис. 2). Эти сложные поверхности были сгенерированы как комбинация нескольких элементарных фрагментов поверхностей. Первым фрагментом выступает горизонтальная плоскость  $K_u = 1$  (константный базовый уровень утилизации). Второй – это фрагмент поверхности конуса, вершина которого соответствует значению  $K_u = 2$ , а фокальный радиус ( $r_c$ ) равен 1. Область пика соответствует комбинациям аллелей, приводящим к наиболее эффективной утилизации субстрата. Она отделена от остальной поверхности окольцовывающей долиной. Обозначим ее внутренний радиус как  $r_{inner}$ , а внешний радиус – как  $r_{outer}$ . Этот центральный пик похож на ландшафт типа “Гора Фудзи” в модели, предложенной Aita и Ну-

**Таблица 1.** Типы и характеристики поверхностей  $K_u$ , использованных в работе

Тип поверхности $K_u$	Ширина долины, $w$	Максимальная глубина долины, $K_{u, \min}$	$\bar{K}_u$	Описание
S_0.44_0.86	2	0.44	0.86	Широкая мелкая долина
S_0.40_0.80	2	0.4	0.8	Широкая неглубокая долина
S_0.34_0.73	2	0.34	0.73	Широкая глубокая долина
S_0.28_0.67	2	0.28	0.67	Широкая глубочайшая долина
S_0.44_0.81	1	0.44	0.81	Узкая мелкая долина
S_0.40_0.78	1	0.4	0.78	Узкая неглубокая долина
S_0.34_0.75	1	0.34	0.75	Узкая глубокая долина
S_0.28_0.72	1	0.28	0.72	Узкая глубочайшая долина

simi [20]. Область долины (“ров”) соответствует пониженным значениям параметра  $K_u$  ( $K_u < 1$ ). В данной работе мы проанализировали влияние формы ландшафта приспособленности на способность популяции достигать максимальных значений  $K_u$ . Мы не меняли положение пика приспособленности и его форму, а также внутренний радиус долины,  $r_{inner}$ . Только внешний радиус долины (параметр  $r_{outer}$ ), определяющий ее ширину  $w = r_{outer} - r_{inner}$ , и глубина долины варьировались в нашем исследовании. Мы проанализировали несколько типов поверхностей ландшафта приспособленности, описание которых можно найти на рис. 2 и в табл. 1. Три примера таких поверхностей можно найти на рис. 2.

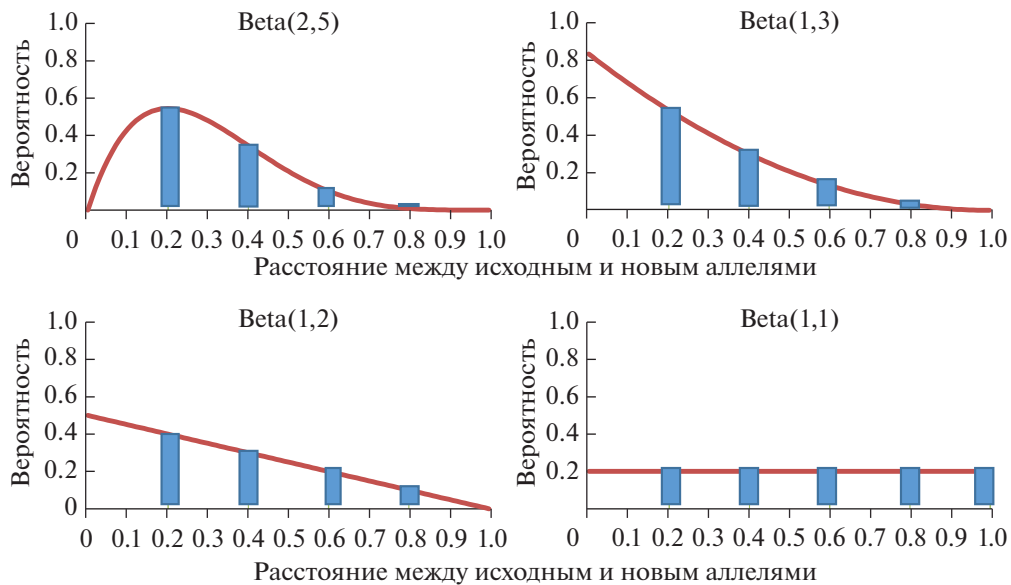
Чтобы сгенерировать различные поверхности  $K_u$  мы использовали два значения ширины долины,  $w = 1$  и  $w = 2$ . Кроме того, варьировали минимальные значения  $K_u$  ( $K_{u, \min}$ ) для долины и наклона ее поверхности. Чтобы охарактеризовать соотношение между высокими и низкими значениями  $K_u$ , соответствующими пику и долине, мы посчитали среднее  $\bar{K}_u$  для узлов сетки в области между пиком и долиной (центр окружности радиуса  $r_{outer}$  совпадает с пиком). Чтобы упростить обозначение поверхностей ввели следующую нотацию, отражающую значения параметров, использованных для их генерации: S\_A\_B, где A – это значение  $K_{u, \min}$ , а B – значение  $\bar{K}_u$ . Таким образом, мы изучили ландшафты с узкими долинами (S\_0.44\_0.81), широкими и неглубокими (S\_0.44\_0.86), а также широкими и глубокими (S\_0.28\_0.67) (см. табл. 1).

#### Изменение аллельных значений

Известно, что частоты замен аминокислот в белках распределены неравномерно: чем меньше физико-химические различия, тем более часто одна аминокислота заменяется на другую (полярная на полярную, заряженная на заряженную и т.д.) [24]. Мы учли этот эффект в своей модели, когда описывали изменение значений аллелей в локусах. Вероятность изменения аллельного значения описывается с помощью дискретной аппроксимации бета-распределения, учитывая четыре возможные комбинации параметров (рис. 3): в первом, втором и третьем случаях малые изменения значений аллелей более вероятны, чем существенные; в четвертом случае все изменения равновероятны. Затем мы дискретизируем случайные числа, полученные из этих распределений, шаг за шагом согласно модели конечных сайтов [25]. Форма распределения определяет размер шага для изменения аллельного значения вдоль поверхности  $K_u$ . Следует отметить, что параметры были выбраны таким образом, чтобы долина могла быть преодолена лишь спустя, как минимум, несколько последовательных шагов.

#### Частота мутаций

Для того чтобы иметь возможность варьировать вероятность мутаций в модели, мы ввели параметр *частоты мутаций*. Этот параметр определяет вероятность того, что аллель материнской клетки изменит свое значение у дочерних клеток. Мы оценили этот параметр на основе имеющихся литературных данных: прежде всего известно, что



**Рис. 3.** Функции плотности распределения вероятности изменения аллельного состояния в результате мутации. Столбцы – дискретные аппроксимации бета-распределения  $\text{Beta}(a,b)$ . В каждом случае возможны переходы аллеля в одно из пяти новых состояний (шаг дискретизации равен 0.2).

для прокариот он составляет порядка 0.003–0.005 мутаций на геном на поколение (МГП) [26, 27], затем размер генома (РГ) прокариот, как правило, находится в диапазоне  $\sim 0.5$ –5 млн пн, в то время как типичная длина фермента составляет порядка 300–400 аминокислот (т.е. соответствующие гены имеют длину (ДГ – длина гена) около 900–1200 нуклеотидов). Таким образом, вероятность того, что такой локус окажется мутирован, можно оценить с помощью произведения  $\text{МГП} \times (\text{ДГ}/\text{РГ})$ :  $5.4 \times 10^{-7}$ – $1.2 \times 10^{-5}$  на локус на клетку на поколение. Мутации в разных локусах моделируются независимо и равновероятно. В литературе отмечался тот факт, что частота мутаций и вероятность рекомбинации могут существенно повышаться в условиях стресса [28–30]. Так, Корогодин с коллегами [31, 32] показали, что стрессовые условия (в частности, у ауксотрофных по аденину дрожжей) могут приводить к тысячекратному повышению мутабельности генов, относящихся к метаболизму аденина. Поэтому в данной работе мы рассматриваем частоту мутаций в пределах типичного диапазона в  $0.000001$ – $0.15$  на клетку на поколение.

#### Схема численного эксперимента

Начальная популяция состояла из однородной группы клеток, “расположенных” в нижнем левом углу ландшафта приспособленности, соответствующего аллельному значению  $(e_1, e_2) = (1, 1)$ . Начальный размер популяции составлял  $1 \times 10^7$  клеток. Дальнейший рост популяции обеспечивал

постепенное достижение максимального размера, равного  $1.5 \times 10^7$  (для стартового генотипа) и обусловленного параметрами среды обитания. В ходе дальнейшей эволюции и движения по ландшафту приспособленности происходит рост средней приспособленности клеток популяции, что приводит к дальнейшему росту размера популяции, достигая соответствующего пика максимума в  $2.1 \times 10^7$  клеток.

Мы проанализировали данную модель на протяжении 10000 поколений (см. сценарий гаплоидного эволюционного конструктора (ГЭК) по ссылке [http://evol-constructor.bionet.nsc.ru/?page\\_id=313](http://evol-constructor.bionet.nsc.ru/?page_id=313)), причем за каждое поколение выполняется стандартная итерация ГЭК, состоящая из следующих шагов:

- 1) клетки поглощают субстраты из окружающей среды;
- 2) клетки метаболизируют поглощенные субстраты, используя их для:
  - а) синтеза и секреции некоторых метаболитов,
  - б) размножения, в ходе которого могут произойти мутации;
- 3) состояние окружающей среды обновляется, в том числе обрабатываются приток и отток субстрата, а также отток клеток из системы.

В результате мутаций могут появиться новые аллельные варианты. Таким образом, мутации с каждым поколением повышают генетическое разнообразие популяции и приводят к “движению” потомков клетки по ландшафту приспособ-

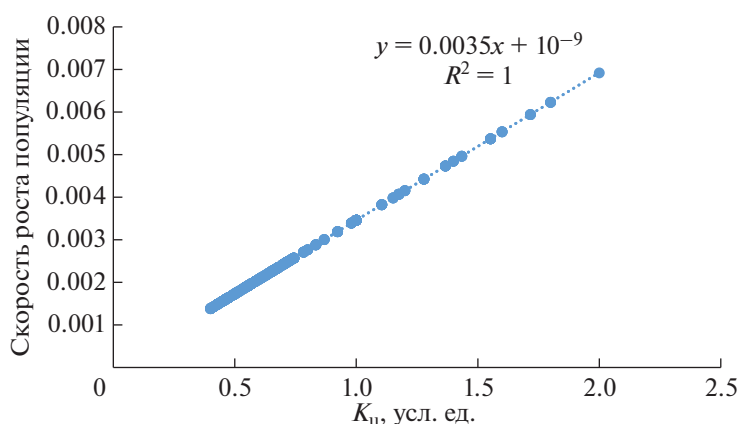


Рис. 4. Зависимость скорости роста популяции (число потомков на клетку за единицу времени) от значения  $K_u$ .

ленности с течением времени, измеряемом в поколениях. Некоторые аллельные комбинации могут оказаться более адаптивными, чем другие, поэтому несущие их мутантные клетки будут расти быстрее и постепенно вытеснят менее адаптивные клетки из популяции. В зависимости от глубины “рва”, окружающего центральный пик на ландшафте приспособленности (рис. 2), клеткам потребуется более или менее интенсивный поток генов, “поставляемых” мутационным процессом, чтобы достичь наиболее приспособленного состояния.

Для того чтобы построить и исследовать описанную выше многоуровневую модель, был использован программный комплекс “Гаплоидный эволюционный конструктор” (ГЭК) [33, 34]. Ранее ГЭК использовался для моделирования различных аспектов функционирования и эволюции микробных сообществ: баланса между приспособляемостью и биологическим разнообразием в симбиотических сообществах [35], эволюционных трендов в системе “бактериальное сообщество–бактериофаг” [36, 37], изменений сложности генома в прокариотических популяциях [38] и влияния пространственной организации сообщества на эволюцию составляющих его микроорганизмов [37, 39].

## РЕЗУЛЬТАТЫ

### *Корреляции между скоростью утилизации субстрата и приспособленностью клеток*

Как мы отметили ранее, в нашей модели нет явной зависимости между приспособленностью клеток и их генотипом. Эта взаимосвязь определяется *a posteriori* через моделирование действия естественного отбора на популяционном уровне. Она зависит не только от эффективности утилизации субстрата, но также и от доступности субстрата

в среде (которая может быть пространственно неоднородна), а также от скорости гибели клеток. Проверка этого соответствия с использованием данных моделирования представляла особый интерес. Ради этой цели мы посчитали среднюю скорость размножения клеток, т.е. число потомков в единицу времени для представителей каждого генотипа, определяемого как аллельная комбинация. Затем мы построили зависимость этой скорости от значений  $K_u = F(e_{1,i}, e_{2,j})$  для всех узлов поверхности. Результаты представлены на рис. 4 и демонстрируют линейную зависимость между значением  $K_u$  и средней скоростью размножения клеток:  $Fitness = aK_u + b$ , где  $a = 0.0035$ ;  $b = 10^9$ , коэффициент корреляции Пирсона равен 0.9999 ( $p < 10^{-9}$ ). Следовательно, можно сделать вывод, что в нашей модели поверхность приспособленности изоморфна поверхности, задаваемой функцией  $F(e_1, e_2)$ , т.е. ее можно использовать в качестве заместителя для исследования ландшафта приспособленности. Схожие результаты были получены для всех вариантов поверхности. Таким образом, было показано, что в рассматриваемых условиях поверхность  $K_u$  полностью определяет относительную приспособленность клеток как функцию их генотипа в нашей модели.

### *Достижение максимума средней приспособленности популяции на разных ландшафтах*

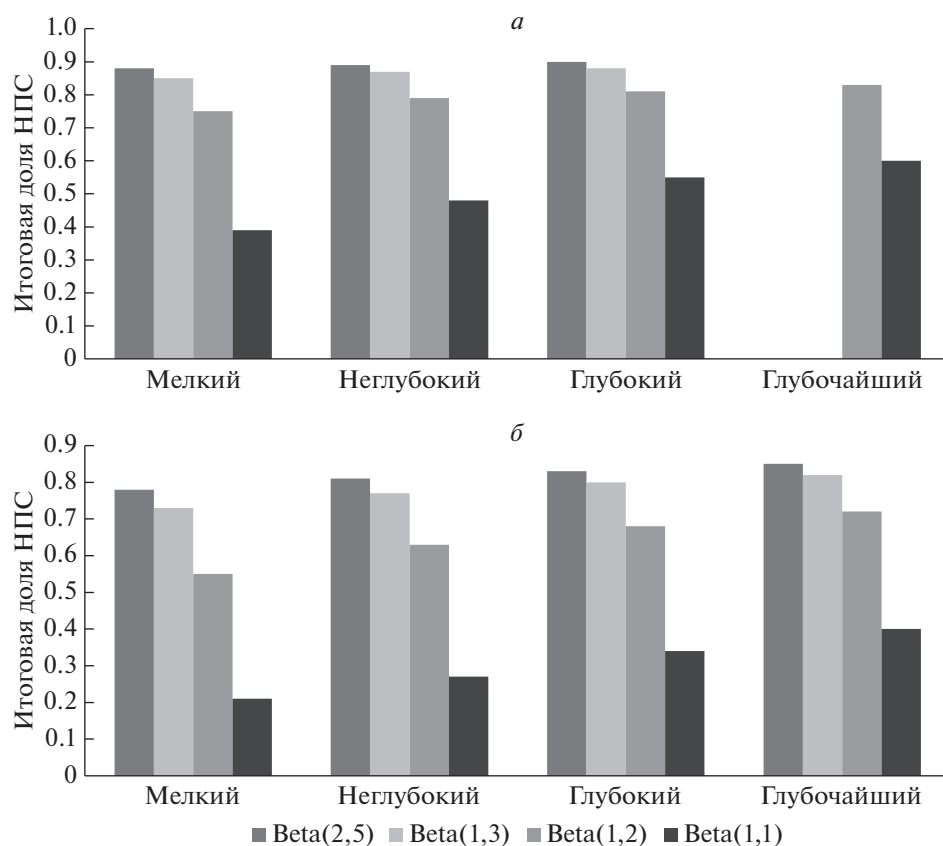
Главный вопрос, на который мы стремились ответить, звучит следующим образом: как взаимосвязаны ландшафт приспособленности, частота мутаций и распределение аллелей среди мутантов? Как эти факторы способствуют или препятствуют появлению наиболее приспособленных клеток? Размер наиболее приспособленной субпопуляции (НПС) напрямую зависит от частоты

**Таблица 2.** Итоговая доля НПС наиболее приспособленной субпопуляции в случае средней частоты мутаций

Распределение	S_0.44_0.86	S_0.40_0.80	S_0.34_0.73	S_0.28_0.67
Beta(2,5)	0.88	0.89	0.9	0
Beta(1,3)	0.85	0.87	0.88	0
Beta(1,2)	0.75	0.79	0.81	0.83
Beta(1,1)	0.39	0.48	0.55	0.6

**Таблица 3.** Итоговая доля НПС наиболее приспособленной субпопуляции в случае высокой частоты мутаций

Распределение	S_0.44_0.86	S_0.40_0.80	S_0.34_0.73	S_0.28_0.67
Beta(2,5)	0.78	0.81	0.83	0.85
Beta(1,3)	0.73	0.77	0.8	0.82
Beta(1,2)	0.55	0.63	0.68	0.72
Beta(1,1)	0.21	0.27	0.34	0.4

**Рис. 5.** Итоговая доля НПС – клеток наиболее приспособленной субпопуляции, полученная по окончании процесса моделирования для разных вариантов ландшафтов приспособленности и распределений мутантов. *а* – средняя частота мутаций; *б* – высокая частота мутаций.



мутаций. Когда она низкая, в системе появляется недостаточно мутантов для того, чтобы популяция в целом могла преодолеть долину приспособленности. Когда частота мутаций высока, то даже наиболее приспособленные клетки продолжают мутировать, и НПС не может достигнуть критической доли во всей популяции. Если мутационный шаг мал, то маловероятно, что даже интенсивный мутационный процесс произведет наиболее приспособленные клетки.

В табл. 2 и на рис. 5,а представлена итоговая доля НПС (т.е. клеток, генотип которых соответствует области пика приспособленности). В случае малого мутационного шага при использовании распределений Beta(2,5) и Beta(1,3) клетки гибнут в глубине долины, не доходя до области наиболее приспособленного генотипа. Мы повышали параметр частоты мутации, чтобы найти размер получившейся НПС. В случаях распределений Beta(2,5) и Beta(1,3) это дает возможность клеткам достигать зоны наиболее приспособленных генотипов. Однако размер НПС снизился во всех прочих вычислительных экспериментах из-за постоянной потери клеток в результате приобретения ими вредных мутаций (табл. 3, рис. 5,б).

Из рис. 5 видно, что если изменения значений аллелей малы (Beta(2,5) и Beta(1,3) согласно рис. 3), то вероятность преодоления долины низкой приспособленности теряется при средних частотах мутаций. Мы также установили зависимость между шириной долины и распределением генотипов в том случае, когда частота мутаций была низкой. В табл. 4 показано существование наиболее приспособленной популяции в разных вариантах расчета моделей.

#### *Влияние популяционных параметров и параметров окружающей среды*

Мы рассмотрели влияние других параметров модели на достижимость пика приспособленности. Коэффициент  $k_{\text{dens}}$  из уравнения 1 отражает долю особей, которые погибают во время процесса изменения численности популяции по квадратичному закону (плотностно-зависимым образом). Коэффициент протока  $k_{\text{flow}}$  из уравнения 1 описывает линейное вымывание части клеток из популяции. Наконец, объем окружающей среды (*volume*) оказывает прямое влияние на размер популяции путем сдерживания ее роста (методические детали, связанные с описанием объема окружающей среды в моделях ГЭК приведены в нашей предыдущей работе [34]). Если в окружающей среде достаточно субстрата и в популяции не действуют никакие ингибирующие законы, уменьшающие размер популяции при ее росте, то увеличение

**Таблица 4.** Существование НПС в конце расчетов модели в зависимости от ширины долины ландшафта приспособленности и распределения мутаций

Распределение	Ширина долины				
	0	0.2	0.4	0.6	0.8
Beta(2,5)	+	+	–	–	–
Beta(1,3)	+	+	–	–	–
Beta(1,2)	+	+	+	–	–
Beta(1,1)	+	+	+	+	–

Примечание. Знаком “+” обозначено существование НПС, “–” – отсутствие НПС.

**Таблица 5.** Номер поколения фиксации 99%-ной доли НПС в зависимости от значений  $k_{\text{dens}}$  и *volume* (при  $k_{\text{flow}} = 0.01$ )

Log10 ( <i>volume</i> )	Log10 ( $k_{\text{dens}}$ )			
	–16	–17	–18	–19
4	×	×	×	×
5	×	×	453276	445499
6	×	418906	313627	284643
7	×	399870	265096	204690

**Таблица 6.** Номер поколения фиксации 99%-ной доли НПС в зависимости от значений  $k_{\text{dens}}$  и  $k_{\text{flow}}$  (при *volume* = 10<sup>6</sup> мл)

$k_{\text{flow}}$	Log10( $k_{\text{dens}}$ )			
	–16	–17	–18	–19
0.001	×	×	454093	446415
0.01	×	418906	313627	284643
0.1	×	399824	264252	202600

*volume* гарантированно дает увеличение максимально возможного размера популяции.

Сочетание трех перечисленных параметров определяет, будет ли преодолена долина заданного ландшафта при заданной скорости мутационного процесса и заданном распределении значений но-

вых мутантных аллелей. В табл. 5 и 6 можно видеть на каком поколении доля НПС достигает 99%.

В табл. 5 и 6 указан номер поколения, на котором достигнуто 99%-ное присутствие у особей сочетания аллелей, соответствующего вершине пика на ландшафте приспособленности. Обозначение крестиком (×) означает, что такого сочетания за время моделирования (500000 поколений) получено не было. То есть возможны два варианта: либо в популяции образовалась другая стационарная точка (как правило, формируется пик недалеко от начала долины) и долина не смогла быть преодолена, либо долина могла бы быть преодолена позже, но 500000 поколений оказалось недостаточно для этого. Приведены результаты для высокой частоты мутаций.

### ОБСУЖДЕНИЕ

В данной работе мы использовали методы компьютерного моделирования для исследования траекторий движения популяции микроорганизмов по различным ландшафтам приспособленности. Мы проанализировали влияние таких параметров как форма функции приспособленности, частота возникновения мутаций и распределение допустимых отклонений при мутации. Было показано, что для различных ландшафтов приспособленности существуют свои оптимальные частоты возникновения мутаций, благоприятствующие успешному преодолению зон сниженной приспособленности — так называемых долин. Такое различие в частотах мутаций может быть интерпретировано как соответствие различным эволюционным режимам — градуальному и сальтационному.

В данной работе мы использовали программный комплекс ГЭК [34] для моделирования эволюции популяции гаплоидных организмов, учитывая ландшафты приспособленности различной формы. Необходимо отметить, что в основе нашей модели лежит представление о физико-химических взаимодействиях между субъединицами фермента, отвечающего за утилизацию субстрата. Эти субъединицы представлены двумя различными локусами в геноме. Таким образом, различные аллельные значения не определяют приспособленность организма напрямую, как в большинстве подходов, использующихся в этой области [16–18]. Другие факторы, такие как пространственная доступность субстрата, например, могут также оказывать влияние на размножение клеток. Влияние этих дополнительных факторов стохастично по своей природе и варьирует от клетки к клетке, однако мы показали, что в нашей

модели активность фермента хорошо коррелирует со способностью клетки произвести потомка (т.е. с приспособленностью) (рис. 4). Эти результаты показывают применимость моделей, опирающихся на молекулярный фенотип как предиктор приспособленности при описании эволюционных процессов в прокариотической популяции.

В основе нашей модели лежит статический ландшафт приспособленности. Эта концепция не может отражать динамические изменения параметров окружающей среды. Однако она применима в тех случаях, когда, например, окружающая среда меняется медленно по сравнению со временем, которое требуется для того, чтобы популяция достигла пика приспособленности. С другой стороны, если изменения в ландшафте характеристик происходят быстро, то эволюционный процесс можно представить как движение по поверхности приспособленности, усредненной по времени [40]. В этих случаях статический ландшафт может выступать в качестве оправданного приближения.

Другим результатом нашего исследования является то, что форма распределения вероятности приобретения аллельных значений в результате мутационного процесса имеет важное значение для способности популяции преодолевать долину. Этот параметр может служить в качестве посредника при рассмотрении изменчивости физико-химических свойств аминокислот при мутациях. Он отражает сложность соответствия между мутациями нуклеотидов в кодоне и изменениями в физико-химических свойствах аминокислоты, что очень важно для ферментной активности. Значения этих свойств распределены не равномерно между 20 аминокислотами, входящими в состав белка, но напротив, для большинства свойств аминокислоты кластеризуются в независимые группы (полярные/неполярные, малые/большие, заряженные/незаряженные) [41]. Как мы показали в своей модели, это соответствие может быть важно для эволюции популяции.

В целом наши результаты раскрывают характер взаимосвязей между мутационными изменениями значений аллелей (заместителей физико-химических свойств аминокислот), частотой мутаций и формой ландшафта приспособленности, зависящей от параметров активности фермента, при эволюционном сценарии, когда популяция гаплоидных организмов пересекает долину низкой приспособленности на ландшафте приспособленности.

На основании данных моделирования можно предложить следующий ряд биологических интерпретаций:

1) чем более однородным является распределение аллелей, тем более вероятно, что глобальный максимум будет достигнут (см. результаты моделирования для распределения Beta(1,1));

2) чем менее однородным является распределение аллелей, тем выше вероятность достичь локальных максимумов (см. результаты моделирования для распределения Beta(2,5));

3) локальные максимумы формируются в областях с наибольшим числом высокоприспособленных соседей (зоны, располагающиеся вблизи долины и начинающиеся в каждом из углов ландшафта приспособленности);

4) чем выше частота мутаций, тем выше шанс достичь глобального максимума, но ниже число клеток, относящихся к наиболее приспособленной субпопуляции (НПС);

5) чем ниже частота мутаций, тем ниже шанс достичь глобального максимума, но выше число клеток НПС (зона локального максимума или глобального, если он существует);

6) чем меньше глубина долины, тем выше вероятность достичь глобального максимума.

Более реалистичные модели должны учитывать дополнительные факторы, влияющие на эволюцию микробных популяций и сообществ: динамические экологические условия, изменяющие ландшафт приспособленности и частоты мутаций, внутри- и межвидовые взаимодействия, приводящие к сложной популяционной динамике, факторы пространственной организации сообщества. Создание и детальное исследование таких сложных многопараметрических моделей помогут более глубокому пониманию эволюционных механизмов.

Работа была поддержана бюджетным проектом № 0324-2019-0040.

Настоящая статья не содержит каких-либо исследований с использованием в качестве объекта животных.

Настоящая статья не содержит каких-либо исследований с участием в качестве объекта людей.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Epistasis / Eds Moore J.H., Williams S.M. N.Y.: Springer, 2015. V. 1253.  
<https://doi.org/10.1007/978-1-4939-2155-3>

2. Kimura M. The role of compensatory neutral mutations in molecular evolution // J. Genet. 1985. V. 64. № 1. P. 7–19.  
<https://doi.org/10.1007/BF02923549>
3. Wright S. Evolution and the Genetics of Populations. V. 3: Experimental Results and Evolutionary Deductions. Univ. Chicago Press, 1984.
4. Whitlock M.C., Phillips P.C., Moore F.B.-G., Tonsor S.J. Multiple fitness peaks and epistasis // Annu. Rev. Ecol. Syst. 1995. V. 26. P. 601–629.
5. Агмухов Ю.П. Генетические процессы в популяциях. М.: Наука, 1983. 280 с.
6. Kuchner O., Arnold F.H. Directed evolution of enzyme catalysts // Trends Biotechnol. 1997. V. 15. № 12. P. 523–530.  
[https://doi.org/10.1016/S0167-7799\(97\)01138-4](https://doi.org/10.1016/S0167-7799(97)01138-4)
7. Gatenby R.A., Vincent T.L. Application of quantitative models from population biology and evolutionary game theory to tumor therapeutic strategies // Mol. Cancer Ther. 2003. V. 2. № 9. P. 919–927.
8. Turelli M., Barton N.H., Coyne J.A. Theory and speciation // Trends Ecol. Evol. 2001. V. 16. № 7. P. 330–343.  
[https://doi.org/10.1016/S0169-5347\(01\)02177-2](https://doi.org/10.1016/S0169-5347(01)02177-2)
9. Domingo E., Holland J.J. RNA virus mutations and fitness for survival // Annu. Rev. Microbiol. 1997. V. 51. № 1. P. 151–178.  
<https://doi.org/10.1146/annurev.micro.51.1.151>
10. Szendro I.G., Schenk M.F., Franke J. et al. Quantitative analyses of empirical fitness landscapes // J. Stat. Mech. Theory Exp. 2013. V. 2013. № 1. P. P01005.  
<https://doi.org/10.1088/1742-5468/2013/01/P01005>
11. Meer M.V., Kondrashov A.S., Artzy-Randrup Y., Kondrashov F.A. Compensatory evolution in mitochondrial tRNAs navigates valleys of low fitness // Nature. 2010. V. 464. № 7286. P. 279–282.  
<https://doi.org/10.1038/nature08691>
12. Mackay T.F.C. Epistasis and quantitative traits: using model organisms to study gene–gene interactions // Nat. Rev. Genet. 2013. V. 15. № 1. P. 22–33.  
<https://doi.org/10.1038/nrg3627>
13. Handel A., Rozen D.E. The impact of population size on the evolution of asexual microbes on smooth versus rugged fitness landscapes // BMC Evol. Biol. 2009. V. 9. № 1. P. 236.  
<https://doi.org/10.1186/1471-2148-9-236>
14. Whitlock M.C., Griswold C.K., Peters A.D. Compensating for the meltdown: The critical effective size of a population with deleterious and compensatory mutations // Ann. Zool. Fennici. 2003. V. 40. № 2. P. 169–183.
15. Szendro I.G., Franke J., de Visser J.A.G.M., Krug J. Predictability of evolution depends nonmonotonically on population size // Proc. Natl Acad. Sci. USA. 2013. V. 110. № 2. P. 571–576.  
<https://doi.org/10.1073/pnas.1213613110>

16. *Gokhale C.S., Iwasa Y., Nowak M.A., Traulsen A.* The pace of evolution across fitness valleys // *J. Theor. Biol.* 2009. V. 259. № 3. P. 613–620.  
<https://doi.org/10.1016/j.jtbi.2009.04.011>
17. *Weissman D.B., Feldman M.W., Fisher D.S.* The rate of fitness-valley crossing in sexual populations // *Genetics*. 2010. V. 186. № 4. P. 1389–1410.  
<https://doi.org/10.1534/genetics.110.123240>
18. *Weissman D.B., Desai M.M., Fisher D., Feldman M.W.* The rate at which asexual populations cross fitness valleys // *Theor. Popul. Biol.* 2009. V. 75. № 4. P. 286–300.  
<https://doi.org/10.1016/j.tpb.2009.02.006>
19. *Haarsma L.L., Nelesen S., VanAndel E. et al.* Simulating evolution of protein complexes through gene duplication and co-option // *J. Theor. Biol.* 2016. V. 399. P. 22–32.  
<https://doi.org/10.1016/j.jtbi.2016.03.028>
20. *Aita T., Husimi Y.* Fitness spectrum among random mutants on Mt. Fuji-type fitness landscape // *J. Theor. Biol.* 1996. V. 182. № 4. P. 469–485.  
<https://doi.org/10.1006/jtbi.1996.0189>
21. *Иорданский Н.* Эволюция жизни. М.: Академия, 2001. 425 с.
22. *Лашин С.А., Суслов В.В., Матушкин Ю.Г.* Теории биологической эволюции с позиций современного развития системной биологии // *Генетика*. 2012. Т. 48. № 5. С. 573–589.
23. *Afonnikov D.A., Oshchepkov D.Y., Kolchanov N.A.* Detection of conserved physico-chemical characteristics of proteins by analyzing clusters of positions with coordinated substitutions // *Bioinformatics*. 2001. V. 17. № 11. P. 1035–1046.  
<https://doi.org/10.1093/bioinformatics/17.11.1035>
24. *Dayhoff M.O., Schwartz R.M., Orcutt B.C.* A model of evolutionary change in proteins // *Atlas of Protein Sequence and Structure*, vol. 5, suppl. 3 / Ed. Dayhoff M.O. Washington, DC: Natl. Biomed. Res. Found. 1978. P. 345–352.
25. *Ewens W.* *Mathematical population genetics: I. Theoretical Introduction*. N.Y.: Springer, 2004. 418 p.  
<https://doi.org/10.1007/978-0-387-21822-9>
26. *Drake J.W., Charlesworth B., Charlesworth D., Crow J.F.* Rates of spontaneous mutation // *Genetics*. 1998. V. 148. № 4. P. 1667–1686.
27. *Lynch M.* Evolution of the mutation rate // *Trends Genet.* 2010. V. 26. № 8. P. 345–352.  
<https://doi.org/10.1016/j.tig.2010.05.003>
28. *Markel A.L.* Stress and evolution // *Russ. J. Genet. Appl. Res.* 2008. V. 12. № 1–2. P. 206–215.
29. *Foster P.L.* Stress responses and genetic variation in bacteria // *Mutat. Res. Mol. Mech. Mutagen.* 2005. V. 569. № 1–2. P. 3–11.  
<https://doi.org/10.1016/j.mrfmmm.2004.07.017>
30. *Hoffmann A.A., Hercus M.J.* Environmental stress as an evolutionary force // *Bioscience*. 2000. V. 50. № 3. P. 217.  
[https://doi.org/10.1641/0006-3568\(2000\)050\[0217:ES-AAEF\]2.3.CO;2](https://doi.org/10.1641/0006-3568(2000)050[0217:ES-AAEF]2.3.CO;2)
31. *Korogodin V.I., Korogodina V.L., Fajsz C. et al.* On the dependence of spontaneous mutation rates on the functional state of genes // *Yeast*. 1991. V. 7. № 2. P. 105–117.  
<https://doi.org/10.1002/yea.320070204>
32. *Ильина В.Л., Корогодин В.И., Файси Ч.* Зависимость частоты образования реверсов разных типов у ауксотрофных по аденину дрожжей от содержания аденина в среде // *Генетика*. 1987. Т. 23. № 4. С. 637–642.
33. *Lashin S.A., Matushkin Y.G.* Haploid evolutionary constructor: new features and further challenges // *In Silico Biol.* 2012. V. 11. № 3–4. P. 125–135.  
<https://doi.org/10.3233/ISB-2012-0447>
34. *Lashin S.A., Klimenko A.I., Mustafin Z.S. et al.* HEC 2.0: improved simulation of the evolution of prokaryotic communities // *Math. Biol. Bioinforma.* 2014. V. 9. № 2. P. 585–596.  
<https://doi.org/10.17537/2014.9.585>
35. *Lashin S.A., Suslov V.V., Matushkin Y.G.* Comparative modeling of coevolution in communities of unicellular organisms: adaptability and biodiversity // *J. Bioinform. Comput. Biol.* 2010. V. 08. № 03. P. 627–643.  
<https://doi.org/10.1142/S0219720010004653>
36. *Lashin S.A., Matushkin Yu.G., Suslov V.V., Kolchanov N.A.* Evolutionary trends in the prokaryotic community and prokaryotic community-phage systems // *Russ. J. Genet.* 2011. V. 47. № 12. P. 1487–1495.
37. *Klimenko A.I., Matushkin Yu.G., Kolchanov N.A., Lashin S.A.* Bacteriophages affect evolution of bacterial communities in spatially distributed habitats: a simulation study // *BMC Microbiol.* 2016. V. 16. № S1. P. S10.  
<https://doi.org/10.1186/s12866-015-0620-4>
38. *Lashin S.A., Matushkin Yu.G., Suslov V.V., Kolchanov N.A.* Computer modeling of genome complexity variation trends in prokaryotic communities under varying habitat conditions // *Ecol. Modell. Elsevier B.V.* 2012. V. 224. № 1. P. 124–129.  
<https://doi.org/10.1016/j.ecolmodel.2011.11.004>
39. *Klimenko A.I., Matushkin Yu.G., Kolchanov N.A., Lashin S.A.* Modeling evolution of spatially distributed bacterial communities: a simulation with the haploid evolutionary constructor // *BMC Evol. Biol.* 2015. V. 15. P. S3.  
<https://doi.org/10.1186/1471-2148-15-S1-S3>
40. *Wilke C.* Dynamic fitness landscapes in molecular evolution // *Phys. Rep.* 2001. V. 349. № 5. P. 395–446.  
[https://doi.org/10.1016/S0370-1573\(00\)00118-6](https://doi.org/10.1016/S0370-1573(00)00118-6)
41. *Kawashima S., Pokarowski P., Pokarowska M. et al.* AAindex: amino acid index database, progress report 2008 // *Nucl. Acids Res.* 2007. V. 36. № Database. P. D202–D205.  
<https://doi.org/10.1093/nar/gkm998>

## Computer Simulation of the Evolution of Microbial Population: Overcoming Local Minima When Reaching a Peak on the Fitness Landscape

S. A. Lashin<sup>a, b, \*</sup>, Z. S. Mustafin<sup>a</sup>, A. I. Klimenko<sup>a</sup>, D. A. Afonnikov<sup>a, b</sup>, and Yu. G. Matushkin<sup>a, b</sup>

<sup>a</sup>*Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Science, Novosibirsk, 630090 Russia*

<sup>b</sup>*Department of Information Biology, Novosibirsk State University, Novosibirsk, 630090 Russia*

*\*e-mail: lashin@bionet.nsc.ru*

The study focuses on the mechanisms of crossing the valleys of fitness by a population of haploid microorganisms, whose fitness depends on allelic values at two different loci and is determined by a complex landscape, the shape of which corresponds to the pattern “a mountain surrounded by a trench in the valley”, that are analyzed using computer modelling. We have studied the influence of various biological factors on the evolutionary perspective of microbial colonies, the reproduction rate of which is controlled by a protein consisting of two subunits encoded at different loci. Both molecular genetic (mutation rate, affinity of subunits), population (fitness function landscape and population density) and ecological (concentration of available substrate in the habitat, flow intensity) factors were considered. Our results demonstrate that the difference in fitness for various allelic combinations, while determining the shape of the fitness landscape, sets the optimal mutation rates to overcome its valleys and opens a window of opportunity for the evolution of the population towards the state of the highest average fitness. Moreover, depending on the fitnesslandscape type, either gradual or saltational evolutionary regimes are optimal for reaching the peak.

**Keywords:** fitness landscape, fitness valley, microbial community, ecological modelling, evolution, mutation rate.