

УДК 575.174.4

## К АНАЛИЗУ СЛАБОГО ДВУЛОКУСНОГО ОТБОРА ПО ЖИЗНЕСПОСОБНОСТИ И КВАЗИРАВНОВЕСИЯ ПО СЦЕПЛЕНИЮ

В. П. Пасеков

Представлено академиком РАН Е. К. Гинтером 17.01.2018 г.

Поступило 05.10.2018 г.

Анализируется модель слабого отбора по жизнеспособности по двум локусам, заданная обыкновенными дифференциальными уравнениями. Слабый отбор рассматривается как возмущение модели без отбора. С помощью теории возмущений найдена оценка коэффициента  $D$  неравновесности по сцеплению в терминах привычных концепций средних эффектов в количественной генетике и в терминах теории сингулярных возмущений в математике. Полученные результаты обобщают выводы М. Кимуры для двух диаллельных локусов на случай множественных аллелей и позволяют стандартизировать анализ слабого действия систематических факторов микроэволюции в популяционной генетике.

*Ключевые слова:* теоретическая популяционная генетика, математические модели, слабый двулокусный отбор по жизнеспособности, множественные аллели, квазиравновесие по сцеплению, сингулярные возмущения.

DOI: <https://doi.org/10.31857/S0869-56524846781-785>

Проблемы теоретического анализа генетической детерминации количественных признаков привлекали внимание и разрабатывались со времени возникновения популяционной генетики [1, 2]. Тогда же формализовались понятия аддитивности и эпистаза эффектов генов, детерминирующих величину признака. С точки зрения микроэволюции особый интерес представляет собой такой признак, как жизнеспособность особей, поскольку она самым очевидным образом влияет на динамику генетической структуры популяции.

На заре становления теоретической популяционной генетики были определены основные простые свойства генетической структуры при случайном скрещивании и отсутствии отбора в диплоидной популяции — закон Харди—Вайнберга в отношении одного локуса и предельная с течением времени независимость комбинирования генов в генотипах для нескольких локусов (равновесие по сцеплению). В дальнейшем выяснилась возможность её нарушения, например, из-за отбора, что резко усложняет анализ структуры.

Для модели с непрерывным временем отбора по жизнеспособности в популяции с неперекрывающимися поколениями без возрастной структуры

обоснование использования равновесия по сцеплению было представлено автором [3] на основе выделения быстрых и медленных переменных. Позже появилось доказательство и для модели с дискретным временем [4]. В цитируемых работах случайное комбинирование генов рассматривалось как лишь приближённое, и отклонение от него было изучено в меньшей степени.

В 1965 г. вышла работа М. Кимуры [5], исследовавшего нарушение от свободного сочетания генов в генотипах в случае двух диаллельных локусов при слабом отборе по жизнеспособности. М. Кимура показал, что нарушение такой свободы невелико (т.е. достигается квазиравновесие по сцеплению), и разработал оригинальное обоснование этих результатов. С тех пор к проблеме квазиравновесия привлекались теоретические соображения, в основном качественные, на базе геометрической теории сингулярных возмущений. Такие соображения [6] обосновывают существование квазиравновесия по сцеплению, близкое к почти независимому сочетанию генов в генотипах, но всё-таки не оценивают его количественно явным образом.

Цель настоящей работы — обобщение анализа слабого отбора по жизнеспособности по двум диаллельным локусам на случай множественных аллелей с использованием теории возмущений [7], стандартизация этого анализа, а также анализ оценки квазиравновесия для коэффициента  $D$  неравновесности по сцеплению и получение результатов в терминах

Вычислительный центр им. А.А. Дородницына  
Федерального исследовательского центра  
“Информатика и управление”  
Российской Академии наук, Москва  
E-mail: [pass40@mail.ru](mailto:pass40@mail.ru)

привычных концепций средних эффектов в количественной генетике и в терминах теории сингулярных возмущений в математике.

Основной вывод настоящей работы сформулирован как Результат 4.

**ДВУЛОКУСНАЯ МОДЕЛЬ  
С ДИСКРЕТНЫМ ВРЕМЕНЕМ  
ПОПУЛЯЦИИ  
С НЕПЕРЕКРЫВАЮЩИМИСЯ  
ПОКОЛЕНИЯМИ**

Пусть жизненный цикл исследуемой популяции имеет вид: *замещение родительской популяции потомками* → *отбор по жизнеспособности* → *рекомбинация* → *появление потомков (нового поколения)*.

Эффектами редких мутаций будем пренебрегать в предположении анализа динамики на ограниченном промежутке времени. Напомним основные обозначения и известные результаты для двулокусной модели динамики под влиянием рекомбинаций (см., например, [8]). Рассмотрим модель случайно скрещивающейся диплоидной (гаплоидной) популяции с неперекрывающимися поколениями и дискретным временем в отношении концентраций  $p_i \equiv p_{i_k i_m}$  гаплоидных генотипов двух аутосомных локусов  $k$  и  $m$ , несущих аллели с номерами  $i_k, i_m$  соответственно, и с вероятностью рекомбинации  $r > 0$ . Уравнения динамики частот  $p_{i_k i_m}$  в модели давления рекомбинаций имеют вид

$$p'_{i_k i_m} = p_{i_k i_m} - r(p_{i_k i_m} - x_{i_k} y_{i_m}) \equiv p_{i_k i_m} - r D_{i_k i_m},$$

$$D'_{i_k i_m} = (1 - r) D_{i_k i_m}; \tag{1}$$

$$x_{i_k} \equiv \sum_{i_m} p_{i_k i_m}, \quad y_{i_m} \equiv \sum_{i_k} p_{i_k i_m},$$

$$D_{i_k i_m} \equiv p_{i_k i_m} - x_{i_k} y_{i_m} (p_{i_k i_m}(t) = x_{i_k}(t) y_{i_m}(t) + D_{i_k i_m}(t)), \tag{2}$$

$x_{i_k}, y_{i_m}, D_{i_k i_m}$  — частоты аллелей  $k$ -го и  $m$ -го локусов и коэффициент (гаметической) неравновесности по сцеплению соответственно,  $t$  — время в поколениях, штрих символизирует следующее поколение. На траекториях (1) частоты аллелей  $x_{i_k}, y_{i_m}$  постоянны. При  $t \rightarrow \infty$   $D_{i_k i_m}(t) \rightarrow 0, p_{i_k i_m}(t) \rightarrow x_{i_k} y_{i_m}$ .

**АППРОКСИМАЦИЯ МОДЕЛИ  
СЛАБОГО ОТБОРА**

Уравнения динамики для стадии отбора по жизнеспособности не зависят от того, рассматривается одлокусный или многолокусный случай, и имеют вид

$$p'_i - p_i \equiv \Delta p_i = \frac{v_i(\mathbf{p}, t)}{v(\mathbf{p}, t)} p_i - p_i = \frac{v_i(\mathbf{p}, t) - v(\mathbf{p}, t)}{v(\mathbf{p}, t)} p_i,$$

$$v(\mathbf{p}, t) = \sum_j v_j(\mathbf{p}, t) p_j,$$

где  $\mathbf{p}$  — вектор частот гаплоидных генотипов,  $v_i(\mathbf{p}, t)$  — жизнеспособность  $i$ -го из них.

Будем называть отбор слабым, если вызываемые им изменения генетической структуры за поколение малы при любом генетическом состоянии популяции  $\mathbf{p}$ , скажем, имеют порядок  $O(\mu), 0 < \mu \ll 1$ . Напомним, что  $O(\mu)$  означает, что изменения по модулю меньше либо равны  $C|\mu|$  для некоторой положительной константы  $C$ .

**Результат 1.** *Для того чтобы эффекты стадии отбора по жизнеспособности в популяции с зависящими от генетического состояния  $\mathbf{p}$  относительными жизнеспособностями  $v_i = v_i(\mathbf{p})$  генотипов были слабыми, достаточно выполнения условия  $|v_i(\mathbf{p}) - v_j(\mathbf{p})| = O(\mu), 0 < \mu \ll 1$  для любых  $i, j, \mathbf{p}$ .*

*Необходимым условием слабого отбора в популяции с постоянными относительными жизнеспособностями  $v_i$  является равенство  $v_{\max} - v_{\min} = O(\mu)$ , где  $v_{\max}$  и  $v_{\min}$  обозначают максимальную и минимальную ( $v_{\max} > v_{\min}$ ) относительные жизнеспособности.*

**Результат 2 [8].** *Пусть рассматривается класс моделей, описываемых системой разностных уравнений  $\mathbf{p}' = \mathbf{p} + \Delta \mathbf{p}(\mathbf{p}, t)$ , где величина координат вектора  $\Delta$  порядка  $O(\mu), 0 < \mu \ll 1$ . Тогда решение системы данного класса аппроксимируется решением системы обыкновенных дифференциальных уравнений*

$$d\mathbf{p}/dt = \Delta \mathbf{p}(\mathbf{p}, t).$$

*Если в модели этого класса диплоидной (гаплоидной) популяции с неперекрывающимися поколениями изменения концентраций гамет  $\Delta \mathbf{p}$  в результате давления  $i$ -го из нескольких последовательно действующих в течение жизненного цикла факторов микроэволюции за поколение малы, то решение модели аппроксимируется решением системы обыкновенных дифференциальных уравнений*

$$dp_i/dt = \sum_i \Delta_i \mathbf{p}(\mathbf{p}, t).$$

**СЛАБЫЙ ОТБОР  
ПО ЖИЗНЕСПОСОБНОСТИ  
КАК ВОЗМУЩЕНИЕ**

Очевидно, переход к дифференциальным уравнениям возможен лишь при слабом отборе, причём на стадии отбора в гаплоидной популяции

$$dp_i/dt = \mu(v_i(\mathbf{p}, t) - v(\mathbf{p}, t)) p_i;$$

$$v(\mathbf{p}, t) \equiv \sum_i v_i(\mathbf{p}, t) p_i, \quad v_i(\mathbf{p}, t) = O(1).$$

Если принимать в расчёт отбор и рекомбинации, то

$$dp_{i_1 i_2} / dt = \mu(v_{i_1 i_2}(\mathbf{p}) - v(\mathbf{p}))p_{i_1 i_2} - rD_{i_1 i_2};$$

$$v(\mathbf{p}) \equiv \sum_{i_1, i_2} v_{i_1 i_2}(\mathbf{p})p_{i_1 i_2}. \quad (3)$$

Перейдём от частот гамет  $p_i$  к коэффициенту неравновесности по сцеплению  $D$  и частотам аллелей:

$$D \equiv p_{11} - xy, \quad x \equiv p_{11} + p_{12} \quad (x_{i_1} = p_{i_1 1} + p_{i_1 2}),$$

$$y \equiv p_{11} + p_{21} \quad (y_{i_2} = p_{1 i_2} + p_{2 i_2}); \quad (4)$$

$$p_{11} = xy + D, \quad p_{i_1 i_2} = x_{i_1} y_{i_2} + (-1)^{i_1 + i_2} D.$$

С помощью (3) и (4) можно составить уравнения динамики для переменных  $D, x, y$ , но они чрезмерно громоздки для эффективного анализа решения. Обратимся к асимптотическим методам теории возмущений. При слабом по сравнению с рекомбинациями отборе скорость динамики из-за последнего имеет порядок  $O(\mu)$ ,  $0 < \mu \ll r, r = O(1)$ . Тогда эффекты отбора можно рассматривать как малое возмущение модели без отбора. Из определения концентрации аллеля  $x$  в (4) следует, что

$$dx/dt = dp_{11}/dt + dp_{12}/dt.$$

Здесь члены из-за рекомбинаций сокращаются, а слагаемые из-за отбора имеют порядок  $O(\mu)$ , откуда  $dx/dt = O(\mu)$ . Аналогично  $dy/dt = O(\mu)$ . Что касается порядка  $dD/dt$ , то сюда входят эффекты и отбора, и рекомбинаций, которые запишем как  $dD/dt = -rD + \mu F$ . Для новых переменных

$$\frac{d}{dt} \begin{bmatrix} D \\ x \\ y \end{bmatrix} = \begin{bmatrix} -rD + \mu F \\ \mu O(1) \\ \mu O(1) \end{bmatrix} = -r \begin{bmatrix} D \\ 0 \\ 0 \end{bmatrix} + \mu \begin{bmatrix} F \\ O(1) \\ O(1) \end{bmatrix}, \quad (5)$$

$$0 < \mu \ll r = O(1).$$

Здесь  $\mu F$  соответствует слабому давлению отбора и/или других систематических факторов микроэволюции. Данная система является возмущённой,  $\mu$  — малый параметр ( $r$  не зависит от генотипа и фиксирован).

**Результат 3.** Пусть рассматривается модель динамики генетической структуры гаплоидной популяции (5) в отношении двух диаллельных локусов. Тогда приближённое значение квазиравновесия для  $D$  представимо в виде

$$D^* = \frac{\mu}{r} F(0, x, y), \quad (6)$$

где  $\mu F(D, x, y)$  определяет вклад в скорость динамики  $D$  слабого давления систематических факторов микроэволюции. Если  $\mu F$  соответствует слабому отбору в гаплоидной популяции с жизнеспособностями генотипов  $\{v_{i_1 i_2}(\mathbf{p})\}$ , то значением (6) будет

$$D^* = \varepsilon(\mathbf{p}) \frac{\mu}{r} x(1-x)y(1-y), \quad (7)$$

$$\varepsilon(\mathbf{p}) \equiv v_{11}(\mathbf{p}) - v_{12}(\mathbf{p}) - v_{21}(\mathbf{p}) + v_{22}(\mathbf{p}).$$

Квазиравновесие по сцеплению  $D^*$  в (7) представимо через генные эффекты в виде, допускающем обобщение на случай множественных аллелей:

$$D^* = (\delta_{11}(\mathbf{p}) - \delta_{1\cdot}^*(\mathbf{p}) - \delta_{\cdot 1}^*(\mathbf{p})) \frac{\mu}{r} x_1 y_1 \equiv \varepsilon_{11}(\mathbf{p}) \frac{\mu}{r} x_1 y_1.$$

Здесь  $\delta_{11} = v_{11} - v^*$  — эффект генотипа,  $\delta_{1\cdot}^* = y_{1\cdot}^* - v^*$ ,  $\delta_{\cdot 1}^* = v_{\cdot 1}^* - v^*$  — эффекты его аллелей, жирная точка в нижнем индексе означает усреднение (суммирование по индексу на этом месте с весами, равными концентрациям аллелей соответствующего локуса),  $v^*$  — популяционное среднее для  $v_{i_1 i_2}$  (звёздочка указывает на вычисление при равновесии по сцеплению),  $\varepsilon_{11}(\mathbf{p})$  интерпретируется как коэффициент эпистаза. Кроме того, величина  $D^*$  имеет порядок  $O(\mu^2)$ .

Уравнение для отыскания  $D^*$  мы получили за счёт использования первых членов разложения в ряд Тейлора функций модели. Оно совпадает с соответствующей частью системы первого приближения асимптотического ряда для решения сингулярно возмущённых уравнений.

### СЛАБЫЙ ОТБОР КАК СИНГУЛЯРНОЕ ВОЗМУЩЕНИЕ

Обратим внимание на то, что в возмущённой модели (5) скорость изменения неравновесности  $D$  много больше скорости медленно эволюционирующих частот аллелей. Такие системы обыкновенных дифференциальных уравнений с быстрыми и медленными переменными называются сингулярно возмущёнными. Для них на первом этапе с большой скоростью идут изменения быстрых переменных в силу системы, получающейся из исходной (возмущённой) при  $\mu = 0$ , тогда как медленные переменные практически фиксированы. У нас на этом этапе быстрые переменные сходятся к некоторому устойчивому по первому приближению равновесию, которое в итоге будет функцией фиксированных текущих медленных переменных: неравновесность по сцеплению  $D$  сходится к квазиравновесному нулевому значению. На следующем этапе происходит эволюция медленных переменных при условии квазиравновесия быстрых: эволюция частот аллелей идёт под действием отбора при условии равновесия по сцеплению. Здесь неравновесное состояние популяции является одновременно как бы итогом для быстрых переменных и этапом эволюции для медленных. Для решения сингулярно возмущённых

систем можно построить асимптотический ряд (см., например, [9]), в терминах которого приведённые выражения  $D^*$  дают первое приближение точного решения.

### МНОЖЕСТВЕННЫЕ АЛЛЕЛИ

При множественных аллелях анализируемых локусов генетическое состояние популяции можно описывать либо частотами двулокусных гамет, либо частотами аллелей  $x_{i_1}$  и  $y_{i_2}$  и неравновесностями по сцеплению  $D_{i_1 i_2}$ :

$$x_{i_1} \equiv \sum_{i_2} p_{i_1 i_2}, \quad y_{i_2} \equiv \sum_{i_1} p_{i_1 i_2}, \\ D_{i_1 i_2} \equiv p_{i_1 i_2} - x_{i_1} y_{i_2}, \quad p_{i_1 i_2} = D_{i_1 i_2} + x_{i_1} y_{i_2}.$$

Подобно диаллельному случаю частоты аллелей при слабом отборе будут медленными переменными, а неравновесности  $D_{i_1 i_2}$  — быстрыми. Динамика в случае множественных аллелей качественно не отличается от наблюдаемой при диаллельных локусах. При слабом давлении  $\{\mu F_{i_1 i_2}\}$  систематических факторов уравнения динамики генетического состояния популяции в терминах (уже не одной, а ряда) неравновесностей по сцеплению  $D_{i_1 i_2}$  и частот аллелей по-прежнему имеют вид (5), где теперь  $\mathbf{D}$ ,  $\mathbf{x}$ ,  $\mathbf{y}$  будут векторами.

**Результат 4.** Пусть рассматривается модель динамики генетической структуры гаплоидной популяции в отношении двух локусов с множественными аллелями при слабом давлении систематических факторов микроэволюции. Тогда приближённым квазиравновесным значением коэффициента  $D_{i_1 i_2}$  будет  $D_{i_1 i_2}^* = \frac{\mu}{r} F_{i_1 i_2}(\mathbf{0}, \mathbf{x}, \mathbf{y})$ , а при слабом отборе по жизнеспособности

$$D_{i_1 i_2}^* = \varepsilon_{i_1 i_2}(\mathbf{p}) \frac{\mu}{r} x_{i_1} y_{i_2}, \quad \varepsilon_{i_1 i_2}(\mathbf{p}) \equiv (\delta_{i_1 i_2} - \delta_{i_1}^* - \delta_{i_2}^*).$$

### ДИПЛОИДНЫЕ ПОПУЛЯЦИИ

В диплоидной популяции со случайным скрещиванием уравнения динамики генетической струк-

туры (на уровне гамет), рассматриваемой в отношении двух диаллельных аутосомных локусов, формально совпадают с уравнениями (3) для гаплоидного случая, но теперь жизнеспособность  $i$ -го гаплоидного генотипа (гаметы) будет средней при условии учёта лишь генотипов, порождённых слиянием  $i$ -й и какими-либо ещё гаметами.

Тем не менее выводы для гаплоидного случая остаются верными, только в них жизнеспособности гамет следует понимать как указанные условные средние, а зависимости от  $v_i$  (например в выражении для  $\varepsilon$ ) — как зависимости от их усреднённых значений, поскольку все манипуляции над (3) для получения результатов допускали зависимость жизнеспособностей от  $\mathbf{p}$ . Первое приближение для неравновесности по сцеплению, например, для двух диаллельных локусов записывается как

$$D^* = \bar{\varepsilon} \frac{\mu}{r} x(1-x)y(1-y), \quad \bar{\varepsilon} \equiv \overline{v_{11}} - \overline{v_{12}} - \overline{v_{21}} + \overline{v_{22}},$$

где надчёркивание означает усреднение по отсутствующим в индексе символам второй гаметы диплоидного генотипа.

### СПИСОК ЛИТЕРАТУРЫ

1. Fisher R.A. // Trans. Roy. Soc. Edinburgh. 1918. V. 52. Pt 2. P. 399–433.
2. Hill W.G. // Genome. 1989. V. 31. № 1. P. 190–195.
3. Пасекоев В.П. // ДАН. 1984. Т. 277. № 6. С. 1338–1341.
4. Nagylaki T. // Genetics. 1993. V. 134. № 6. P. 624–647.
5. Kimura M. // Genetics. 1965. V. 52. № 5. P. 875–890.
6. Shahshahani S. // Mem. Amer. Math. Soc. 1979. V. 7. № 211. 34 p.
7. Banasiak J., Lachowicz M. Methods of Small Parameter in Mathematical Biology. B.: Springer, 2014. 285 p.
8. Свирежев Ю.М., Пасекоев В.П. Основы математической генетики. М.: Наука, 1982. 511 с.
9. Новожилов И.В. Методы разделения движений: Конспект лекций. М.: МЭИ, 1981. 48 с.

## TO THE ANALYSIS OF WEAK TWO-LOCUS VIABILITY SELECTION AND QUASI LINKAGE EQUILIBRIUM

V. P. Pasekov

Presented by Academician of the RAS E.K. Ginter January 17, 2018

Received October 5, 2018

A model of weak viability selection at two multi-allele loci with standardization of approaches through the use of perturbation theory is examined. The estimate of the quasi-equilibrium value for the linkage disequilibrium coefficient  $D$  is analyzed, and results in terms of average effects in quantitative genetics and in terms of the theory of singular perturbations in mathematics are obtained. The approximation of a discrete-time model of a random mating population with non-overlapping generations under weak selection by ordinary differential equations is considered. Weak selection is considered as a perturbation of the model without selection. The resulting model is singularly perturbed; that is, fast ( $D$ ) and slow (allele frequencies) variables can be distinguished. The first approximation equation for quasi-equilibrium of  $D$  is obtained using the first terms of the Taylor series expansion of the model functions. It coincides with the corresponding part of the system of the first approximation of the asymptotic series for solving singularly perturbed equations.

*Keywords:* theoretical population genetics, mathematical models, weak two-locus viability selection, multiple alleles, quasi linkage equilibrium, singular perturbations.