

МАТЕМАТИЧЕСКИЕ
МОДЕЛИ И МЕТОДЫ

УДК 575.174.4

КВАЗИРАВНОВЕСИЕ ПО СЦЕПЛЕНИЮ ПРИ СЛАБОМ
ДВУЛОКУСНОМ ОТБОРЕ ПО ЖИЗНЕСПОСОБНОСТИ.
II. ЛОКУСЫ С МНОЖЕСТВЕННЫМИ АЛЛЕЛЯМИ

© 2019 г. В. П. Пасеков*

*Вычислительный центр им. А.А. Дородницына Федерального исследовательского центра
“Информатика и управление” Российской академии наук, Москва, 119991 Россия***e-mail: pass40@mail.ru*

Поступила в редакцию 24.04.2018 г.

После доработки 17.07.2018 г.

Принята к публикации 23.07.2018 г.

Рассматривается модель слабого отбора по жизнеспособности по двум локусам с множественными аллелями в гаплоидной популяции. Предполагается, что скрещивание в популяции случайно, поколения не перекрываются, динамика генетической структуры моделируется обыкновенными дифференциальными уравнениями. Слабый отбор рассматривается как возмущение модели без отбора. Полученные ранее при диаллельных локусах оценки квазиравновесия для коэффициента неравновесности по сцеплению D обобщаются на случай множественных аллелей. При этом интерпретация оценок в терминах средних эффектов в количественной генетике оказалась независимой от числа аллелей. Первым приближением для квазиравновесия $D_{i_1 i_2}$ будет $D_{i_1 i_2}^* = \varepsilon_{i_1 i_2}(\mathbf{p}) \frac{\mu}{r} x_{i_1} y_{i_2}$, где μ — интенсивность отбора; r — коэффициент рекомбинации; $\varepsilon_{i_1 i_2}(\mathbf{p})$ — коэффициент неаддитивности эпистаза для жизнеспособностей $v_{i_1 i_2}$ гаплоидных генотипов $i_1 i_2$; x_{i_1}, y_{i_2} — частоты составляющих генотип аллелей с номерами i_1, i_2 первого и второго локуса соответственно. Специфика возмущенной модели в том, что в ней выделяются быстрые (неравновесности $D_{i_1 i_2}$) и медленные (частоты аллелей) переменные, что характеризует модель как сингулярно возмущенную. Это позволяет без решения уравнений выделить в динамике два этапа. На первом быстрые переменные сходятся к квазиравновесным значениям, а на втором идет эволюция медленных переменных при условии квазиравновесия быстрых. Полученные дельта-методом уравнения относительно квазиравновесных значений $D_{i_1 i_2}$ дают корректный результат с точки зрения разложения по степеням малого параметра решения сингулярно возмущенных уравнений. Обсуждаются проблемы измерения эпистаза и разложения коэффициента эпистаза на элементарные составляющие, а также анализируются двулокусные коэффициенты эпистаза как характеристики многолокусной модели. Кратко указано обобщение полученных результатов для гаплоидной популяции на диплоидный случай.

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

DOI: 10.1134/S001667581904012X

Рассмотрим модель динамики генетической структуры гаплоидной популяции с неперекрывающимися поколениями и случайным скрещиванием в непрерывном времени при учете отбора по жизнеспособности по двум локусам k и m (см. [1]). Напомним, что предполагается следующая последовательность стадий жизненного цикла:

замещение родительской популяции потомками →
→ отбор по жизнеспособности → рекомбинации →
→ появление потомков (нового поколения).

Генетическое состояние (структура) популяции описывается вектором $\mathbf{p}$ концентраций двулокусных гаплоидных генотипов:

$$\mathbf{p} = \{p_i = p_{i_k i_m}\}, \quad \sum_i p_i = 1, \quad 1 \geq p_i \geq 0, \quad \mathbf{i} = (i_k, i_m).$$

Здесь i_k — номер аллеля k -го локуса, входящего в $\mathbf{i}$ -й генотип (короче, i_k обозначает соответствующий аллель); $p_{i_k i_m}$ — частота двулокусного гаплоидного генотипа с аллелями i_k, i_m . В дальнейшем

состояние популяции будем относить к началу поколения (к “новорожденным”, заменяющим поколение родителей).

Покоординатно уравнения динамики генетического состояния популяции в модели имеют вид

$$dp_{ikim}/dt = \mu(v_{ikim}(\mathbf{p}) - v(\mathbf{p}))p_{ikim} - rD_{ikim}.$$

Здесь t , $v_{ikim}(\mathbf{p})$ дают время и жизнеспособность генотипа i_{ikim} соответственно, $v(\mathbf{p}) \equiv \sum_{i_{ikim}} v_{ikim} p_{ikim}$ — средняя жизнеспособность, μ и r обозначают параметры давления отбора и независимый от генотипа коэффициент рекомбинации.

Генетическое состояние популяции можно описывать не только частотами двулокусных гамет, но и частотами аллелей x_{ik} и y_{im} и неравновесностями по сцеплению D_{ikim} :

$$x_{ik} \equiv \sum_{i_m} p_{ikim}, \quad y_{im} \equiv \sum_{i_k} p_{ikim}, \\ D_{ikim} \equiv p_{ikim} - x_{ik}y_{im} \quad (p_{ikim} = D_{ikim} + x_{ik}y_{im}).$$

Здесь допускаются множественные аллели.

Традиционной задачей исследования модели является анализ ее решения и в идеале получение явного выражения состояния популяции в зависимости от времени, начального состояния, жизнеспособностей и коэффициента рекомбинации. Когда такого явного выражения получить не удастся, желательно определить качественную картину динамики: типы поведения, равновесия, циклы, их устойчивость и т.д. При этом важную роль могут играть методы приближенного анализа, в частности асимптотического анализа при различных соотношениях между параметрами.

При отсутствии отбора генетическое состояние популяции с течением времени приходит в результате процессов рекомбинации к *равновесию по сцеплению*, когда сочетания аллелей в генотипах особей независимы (и все коэффициенты неравновесности D_{ikim} равны нулю). Процесс перехода при произвольном числе аутосомных локусов для диплоидной популяции со случайным скрещиванием детально исследован Ю.И. Любичем [2, 3], получившим явное решение с использованием техники генетических алгебр. Качественно картина динамики перехода охарактеризована более простыми методами в [4]. Отбор способен препятствовать достижению такой независимости, и возникает проблема оценивания квазиравновесных значений коэффициентов D_{ikim} .

Случай слабого давления отбора по жизнеспособности, зависящей от двух диаллельных локусов, исследовал М. Кимура [5], использовавший показатель неравновесности по сцеплению $z \equiv p_{11}p_{22}/(p_{12}p_{21})$. Предложенный им подход трудно перенести на другие сходные ситуации слабого

давления систематических факторов микроэволюции, в частности на случай отбора по локусам с множественными аллелями. Здесь кажется привлекательным рассматривать слабый отбор как возмущение нейтральной модели без отбора и воспользоваться теорией возмущений в соответствии с логикой работы [1].

Изложение материала в настоящем сообщении придерживается следующей схемы. Сначала продемонстрировано, что слабый по сравнению с рекомбинациями отбор можно рассматривать как возмущение нейтральной модели. Далее дельта-методом находятся зависящие от отношения μ/r квазиравновесные значения коэффициентов D_{ikim} неравновесностей по сцеплению. Затем показано, что уравнения модели в терминах частот аллелей x_{ik} и y_{im} и неравновесностей по сцеплению D_{ikim} имеют вид сингулярно возмущенных, где D_{ikim} являются быстрыми, а частоты аллелей x_{ik} и y_{im} медленными переменными. Далее обосновывается вывод уравнения относительно квазиравновесного значения D_{ikim} и соответствующего коэффициента эпистаза с точки зрения разложения в ряд по степеням малого параметра решения сингулярно возмущенных систем уравнений. Затем обсуждаются проблемы измерения эпистаза и разложения коэффициента эпистаза на элементарные составляющие, а также анализируются двулокусные коэффициенты эпистаза как характеристики многолокусной модели. Кратко рассматривается обобщение полученных для гаплоидной популяции результатов на диплоидный случай. В заключение дано обсуждение результатов. В тексте векторы и матрицы набраны полужирным шрифтом. Получаемые выводы сформулированы как результаты и ремарки, конец доказательств отмечен символом ◀.

ЛОКУСЫ С МНОЖЕСТВЕННЫМИ АЛЛЕЛЯМИ

При изучении двух первых локусов ($k = 1, m = 2$) уравнения динамики частот гаплоидных генотипов $p_{i_1i_2}$ с учетом слабого отбора по жизнеспособности можно записать как

$$dp_{i_1i_2}/dt = \mu p_{i_1i_2} (v_{i_1i_2}(\mathbf{p}) - v(\mathbf{p})) - rD_{i_1i_2}, \\ v(\mathbf{p}) \equiv \sum_{i_1, i_2} v_{i_1i_2}(\mathbf{p}) (x_{i_1}y_{i_2} + D_{i_1i_2}). \quad (1)$$

Напомним, что здесь r — коэффициент рекомбинации, $v_{i_1i_2}(\mathbf{p})$ — жизнеспособность гаплоидной особи с генотипом i_1i_2 при генетическом состоянии популяции $\mathbf{p}$.

Как и ранее [1], при диаллельности эти уравнения можно переписать в терминах неравновесно-

стей по сцеплению и частот аллелей в векторном виде

$$\frac{d}{dt} \begin{bmatrix} \mathbf{D} \\ \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} -r\mathbf{D} + \mu\mathbf{F} \\ \mu\mathbf{O}(1) \\ \mu\mathbf{O}(1) \end{bmatrix} = -r \begin{bmatrix} \mathbf{D} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} + \mu \begin{bmatrix} \mathbf{F} \\ \mathbf{O}(1) \\ \mathbf{O}(1) \end{bmatrix}, \quad (2)$$

$$0 < \mu \ll 1,$$

где $\mathbf{D}$, $\mathbf{x}$, $\mathbf{y}$ обозначают векторы коэффициентов неравновесностей и частот аллелей первого и второго локуса соответственно, вектор-функция $\mathbf{F}$ описывает давление систематических факторов (в частности, отбора). Первое слагаемое в правой части соответствует нейтральной модели, а второе — ее возмущению давлением систематических факторов. При $\mu = 0$ получаем невозмущенную модель, где в равновесии, очевидно, $\mathbf{D} = \mathbf{0}$.

Результат 1. Пусть рассматривается модель (2) динамики генетической структуры гаплоидной популяции в отношении двух локусов с множественными аллелями при слабом давлении систематических факторов микроэволюции $\mu\mathbf{F}$.

Тогда приближенным квазиравновесным значением коэффициента $D_{i i_2}$ будет

$$D_{i i_2}^* = \frac{\mu}{r} F_{i i_2}(\mathbf{0}, \mathbf{x}, \mathbf{y}). \quad (3)$$

Доказательство следует схеме диаллельного случая в [1]. Будем анализировать модель в терминах неравновесностей и частот аллелей. Уточненная оценка квазиравновесного значения $D_{i i_2}$ отличается от нулевого приближения $D_{i i_2} = 0$ приращением $\Delta D_{i i_2}$, и задача редуцируется к поиску $\Delta D_{i i_2}$.

А. Приращение $\Delta D_{i i_2}$ вызвано отклонением $\Delta\mu$ малого параметра μ (новым значением которого станет $\mu = \Delta\mu$) и векторами малых приращений $\Delta\mathbf{x}$ и $\Delta\mathbf{y}$. Запишем $dD_{i i_2}/dt$ как $-rD_{i i_2} + \mu F_{i i_2}(\mathbf{D}, \mathbf{x}, \mathbf{y})$, откуда в невозмущенном равновесии $(-rD_{i i_2} + \mu|_{\mu=0} F_{i i_2}(\mathbf{D}, \mathbf{x}, \mathbf{y}))|_{\mathbf{D}=\mathbf{0}} = 0$. Аналогично в уточненном квазиравновесии возмущенной модели

$$-rD_{i i_2}|_{\mathbf{D}=\mathbf{0}+\Delta\mathbf{D}} + \mu|_{\mu=0+\Delta\mu} F_{i i_2}(\mathbf{D}, \mathbf{x} + \Delta\mathbf{x}, \mathbf{y} + \Delta\mathbf{y})|_{\mathbf{D}=\mathbf{0}+\Delta\mathbf{D}} = 0.$$

При указанных приращениях аргументов функция $-rD_{i i_2} + \mu F_{i i_2}$ не изменилась, т.е. приближенно ее приращение Δ можно записать как

$$\Delta = \left(\frac{\partial}{\partial \mathbf{D}} (-rD_{i i_2} + \mu F_{i i_2}), \Delta\mathbf{D} \right) + \left(\frac{\partial}{\partial \mathbf{x}} (-rD_{i i_2} + \mu F_{i i_2}), \Delta\mathbf{x} \right) + \left(\frac{\partial}{\partial \mathbf{y}} (-rD_{i i_2} + \mu F_{i i_2}), \Delta\mathbf{y} \right) + \frac{\partial}{\partial \mu} (-rD_{i i_2} + \mu F_{i i_2}) \Delta\mu = 0,$$

где операция $(,)$ означает скалярное умножение, производные $F_{i i_2}$ вычисляются при $\mathbf{D} = \mathbf{0}$, при текущих значениях $\mathbf{x}$, $\mathbf{y}$ и при $\mu = 0$. Это уравнение отличается от полученного в диаллельном случае тем, что обычное умножение производных на приращения заменено на скалярное, производные по вектору понимаем как вектор-градиент.

Б. В результате приближенное значение $D_{i i_2}^*$ удовлетворяет уравнению

$$-r\Delta D_{i i_2} + \mu|_{\mu=0} \left(\left(\frac{\partial}{\partial \mathbf{D}} F_{i i_2}, \Delta\mathbf{D} \right) + \left(\frac{\partial}{\partial \mathbf{x}} F_{i i_2}, \Delta\mathbf{x} \right) + \left(\frac{\partial}{\partial \mathbf{y}} F_{i i_2}, \Delta\mathbf{y} \right) \right) + F_{i i_2}(\mathbf{0}, \mathbf{x}, \mathbf{y}) \Delta\mu = -rD_{i i_2}^* + F_{i i_2}(\mathbf{0}, \mathbf{x}, \mathbf{y}) \Delta\mu = 0. \quad (4)$$

Отсюда видно, что значения отличных от $D_{i i_2}$ неравновесностей (с другими индексами) и их приращения, а также $\Delta\mathbf{x}$, $\Delta\mathbf{y}$ не важны для вычисления $D_{i i_2}^*$ ($D_{i i_2}^*$ обозначает уточненное значение квазиравновесия $D_{i i_2}$), поскольку от $\Delta\mathbf{x}$, $\Delta\mathbf{y}$ зависит лишь коэффициент при нулевом μ . Кроме того, $\Delta D_{i i_2} = D_{i i_2}^*$, так как приращение $\Delta D_{i i_2}$ отсчитывается от нуля до $D_{i i_2}^*$, и аналогично $\Delta\mu = \mu$. В итоге искомым значением $D_{i i_2}^*$ будет $D_{i i_2}^* = \frac{\mu}{r} F_{i i_2}(\mathbf{0}, \mathbf{x}, \mathbf{y})$. Тем самым и часть Б доказана. ◀

Следующая задача состоит в вычислении $D_{i i_2}^*$, когда слабое давление систематических факторов микроэволюции соответствует слабому давлению отбора с соответствующим значением $F_{i i_2}(\mathbf{D}, \mathbf{x}, \mathbf{y})$ при $\mathbf{D} = \mathbf{0}$.

Результат 2. Пусть рассматривается модель гаплоидной популяции (1) со слабым отбором по жизнеспособности, зависящей от двух локусов с множественными аллелями.

Тогда приближенным значением квазиравновесия $D_{i i_2}$ для модели (1) будет

$$D_{i_1 i_2}^* = \varepsilon_{i_1 i_2}(\mathbf{p}) \frac{\mu}{r} x_{i_1} y_{i_2} = \left(\sum_{j_1, j_2} \varepsilon_{i_1 i_2 j_1 j_2} x_{j_1} y_{j_2} \right) \frac{\mu}{r} x_{i_1} y_{i_2};$$

$$\varepsilon_{i_1 i_2}(\mathbf{p}) \equiv (v_{i_1 i_2} - v_{i_1}^* - v_{i_2}^* + v^*) = \sum_{j_1, j_2} \varepsilon_{i_1 i_2 j_1 j_2} x_{j_1} y_{j_2}, \quad (5)$$

$$\varepsilon_{i_1 i_2 j_1 j_2} \equiv v_{i_1 i_2} - v_{i_1 j_2} - v_{j_1 i_2} + v_{j_1 j_2}.$$

Доказательство. Воспользуемся уравнением (1) динамики $p_{i_1 i_2}$ (для краткости записи зависимость жизнеспособностей от генетической структуры $\mathbf{p}$ популяции опустим). Из определения частот аллелей следует, что

$$dx_{i_1}/dt \equiv \sum_{i_2} dp_{i_1 i_2}/dt =$$

$$= \sum_{i_2} (\mu p_{i_1 i_2} (v_{i_1 i_2}(\mathbf{p}) - v(\mathbf{p})) - r D_{i_1 i_2}) =$$

$$= \sum_{i_2} \mu (x_{i_1} y_{i_2} + D_{i_1 i_2}) (v_{i_1 i_2} - v) =$$

$$= \mu x_{i_1} \left(\sum_{i_2} v_{i_1 i_2} y_{i_2} - v \right) + \mu G \equiv \mu x_{i_1} (v_{i_1}^* - v) + \mu G,$$

где мы учли, что из определения неравновесности следует равенство $\sum_{i_2} r D_{i_1 i_2} = 0$, и остаются только члены, содержащие произведение неравновесностей и жизнеспособностей. Их суммарное значение обозначено как G . Точный вид G для дальнейшего не имеет значения, важно лишь, что G обращается в нуль при равновесии по сцеплению, т.е. при $\mathbf{D} = \mathbf{0}$ (будем сохранять обозначение G за разными функциями с данным свойством по аналогии с определением функции O). Жирной точкой обозначается индекс у $v_{i_1 i_2}$, по которому произведено суммирование с весами, равными частотам аллелей соответствующего локуса. Значение $v_{i_1}^*$ дает среднюю (маргинальную) жизнеспособность аллеля первого локуса при равновесии по сцеплению. Аналогично находим, что

$$dy_{i_2}/dt = \mu y_{i_2} (v - v) + \mu G.$$

Как и в диаллельном случае, $\mu F_{i_1 i_2}(\mathbf{0}, \mathbf{x}, \mathbf{y}) = dD_{i_1 i_2}/dt$ при $\mathbf{D} = \mathbf{0}$, а $dD_{i_1 i_2}/dt = dp_{i_1 i_2}/dt - d(x_{i_1} y_{i_2})/dt$ согласно определению $D_{i_1 i_2}$. Вспомним, что при $\mathbf{D} = \mathbf{0}$ (т.е. при $G = 0$ и $v = v^*$) $p_{i_1 i_2} = x_{i_1} y_{i_2}$ (откуда $dp_{i_1 i_2}/dt = \mu x_{i_1} y_{i_2} (v_{i_1 i_2} - v^*)$), и найдем $d(x_{i_1} y_{i_2})/dt$:

$$d(x_{i_1} y_{i_2})/dt = \mu x_{i_1} (v_{i_1}^* - v^*) y_{i_2} +$$

$$+ \mu x_{i_1} y_{i_2} (v_{i_2}^* - v^*) = \mu x_{i_1} y_{i_2} (v_{i_1}^* + v_{i_2}^* - 2v^*).$$

Подставим в $dD_{i_1 i_2}/dt|_{\mathbf{D}=\mathbf{0}}$ эти выражения. В результате после сокращения μ получим

$$F_{i_1 i_2}(\mathbf{0}, \mathbf{x}, \mathbf{y}) = (dp_{i_1 i_2}/dt - d(x_{i_1} y_{i_2})/dt)|_{\mathbf{D}=\mathbf{0}} =$$

$$= x_{i_1} y_{i_2} (v_{i_1 i_2} - v^*) - x_{i_1} y_{i_2} (v_{i_1}^* + v_{i_2}^* - 2v^*) =$$

$$= x_{i_1} y_{i_2} (v_{i_1 i_2} - v_{i_1}^* - v_{i_2}^* + v^*) =$$

$$= x_{i_1} y_{i_2} \sum_{j_1, j_2} (v_{i_1 i_2} - v_{i_1 j_2} - v_{j_1 i_2} + v_{j_1 j_2}) x_{j_1} y_{j_2} \equiv$$

$$\equiv x_{i_1} y_{i_2} \sum_{j_1, j_2} \varepsilon_{i_1 i_2 j_1 j_2} x_{j_1} y_{j_2} \equiv \varepsilon_{i_1 i_2} x_{i_1} y_{i_2}.$$

Здесь коэффициент $\varepsilon_{i_1 i_2 j_1 j_2} \equiv v_{i_1 i_2} - v_{i_1 j_2} - v_{j_1 i_2} + v_{j_1 j_2}$ будем называть коэффициентом неаддитивности эпистаза для жизнеспособностей гаплоидных генотипов, образованных аллелями элементарного “генофонда”, состоящего из рассматриваемого генотипа $i_1 i_2$ и случайно выбранного из популяции генотипа $j_1 j_2$. При этом $\varepsilon_{i_1 i_2}$ интерпретируется как средний (маргинальный) эпистатический эффект генотипа $i_1 i_2$. Итак, искомое значение

$$F_{i_1 i_2}(\mathbf{0}, \mathbf{x}, \mathbf{y}) \text{ будет равно } \varepsilon_{i_1 i_2} x_{i_1} y_{i_2}, \text{ а } D_{i_1 i_2}^* \text{ имеет вид (5):}$$

$$D_{i_1 i_2}^* = \varepsilon_{i_1 i_2}(\mathbf{p}) \frac{\mu}{r} x_{i_1} y_{i_2}. \blacktriangleleft$$

Отметим, что $\varepsilon_{i_1 i_2}$ с точностью до множителя $x_{i_1} y_{i_2}$ редуцируется к коэффициенту неаддитивности эпистаза в случае двух диаллельных локусов, рассмотренном в [1]. В форме (5) $D_{i_1 i_2}^*$ формально совпадает с квазиравновесностью при диаллельных локусах, а коэффициенты эпистаза в диаллельном случае и при множественных аллелях одинаково интерпретируются в терминах средних эффектов аллелей:

$$\varepsilon_{i_1 i_2}(\mathbf{p}) \equiv v_{i_1 i_2} - v_{i_1}^* - v_{i_2}^* + v^* =$$

$$= (v_{i_1 i_2} - v^*) - (v_{i_1}^* - v^*) - (v_{i_2}^* - v^*) =$$

$$= \delta_{i_1 i_2} - \delta_{i_1}^* - \delta_{i_2}^*.$$

Здесь $\delta_{i_1}^*$ имеет смысл (среднего) эффекта аллеля i_1 первого локуса, $\delta_{i_2}^*$ — эффект аллеля i_2 второго локуса, $\delta_{i_1 i_2}$ — эффект генотипа $i_1 i_2$, а $\varepsilon_{i_1 i_2}(\mathbf{p})$ — эффект взаимодействия аллелей i_1, i_2 , или эпистатическое отклонение (см., например, [6]). Такая интерпретация позволяет наглядно представить коэффициент $\varepsilon_{i_1 i_2}$ как отклонение жизнеспособности $v_{i_1 i_2}$ особи с генотипом $i_1 i_2$ от средней жизнеспособности популяции v за вычетом вклада в отклонение аддитивного генотипа $\delta_{i_1}^* + \delta_{i_2}^*$.

Как и в диаллельном случае, при аддитивной детерминации жизнеспособности коэффициент $\varepsilon_{i_1 i_2}$ равен нулю, а при мультипликативной

$$\varepsilon_{i_1 i_2}(\mathbf{p}) = (u_{i_1} - u_i^*)(w_{i_2} - w_{i_2}^*),$$

$$u_i^* \equiv \sum_{i_1} u_{i_1} x_{i_1}, \quad w_{i_2}^* \equiv \sum_{i_2} w_{i_2} y_{i_2},$$

где u_{i_1} и w_{i_2} — мультипликативные компоненты жизнеспособности, соответствующие первому и второму локусам.

СЛАБЫЙ ОТБОР КАК СИНГУЛЯРНОЕ ВОЗМУЩЕНИЕ

Согласно изучению модели без отбора (см., например, [4]) частоты аллелей в ряду поколений остаются постоянными при динамике генетического состояния популяции. В ней уравнения, описывающие поведение во времени неравновесностей $D_{i_k i_m}$, отделяются, их можно изучать независимо. Каждая из неравновесностей сходится к нулю с относительной скоростью r . При слабом отборе частоты аллелей подобно диаллельному случаю будут медленными переменными (dx_{i_1}/dt и dy_{i_2}/dt порядка $O(\mu)$ согласно доказательству предыдущего результата), а неравновесности $D_{i_k i_m}$ — быстрыми ($dD_{i_k i_2}/dt = O(r)$ согласно (2)), т.е. уравнения модели являются *сингулярно возмущенными* (см., например, [7]). В результате качественно динамика в случае множественных аллелей не отличается от наблюдаемой при диаллельных локусах. Согласно теории сингулярных возмущений на первом этапе все неравновесности по сцеплению быстро затухают с одной и той же относительной скоростью r ($\mathbf{D} = \mathbf{0}$ является асимптотически устойчивым и устойчивым по первому приближению равновесием невозмущенной системы). После этого эволюция медленных частот аллелей приближенно идет при условии данного равновесия быстрых переменных — при многолокусных соотношениях Харди–Вайнберга. В результате, например, при постоянных жизнеспособностях эволюция протекает в направлении, соответствующем увеличению средней жизнеспособности популяции [8].

Разумеется, это лишь приближенное описание динамики. В точном решении уравнений динамики соотношения Харди–Вайнберга нарушаются в слабой степени и желательно оценить степень нарушения (величину отклонений от нуля неравновесностей $D_{i_k i_m}$). Такая оценка найдена выше путем, аналогичным предпринятому при изучении диаллельных локусов [1]. Обоснованием корректности полученных оценок квазиравновесия служит тот факт, что уравнения относительно $D_{i_1 i_2}^*$ совпадают с соответствующей частью системы первого приближения асимптотическо-

го ряда для решения сингулярно возмущенных уравнений (см., например, [7]).

Ремарка 3. Построение модели динамики генетического состояния популяции под действием отбора и рекомбинаций мы начали с уравнений в терминах частот гамет (1). В ней сингулярный характер возмущений из-за слабого отбора не очевиден, и нужна некоторая изобретательность для отыскания замены переменных, приводящей уравнения модели к стандартному виду сингулярно возмущенной системы. Эта задача облегчается при наличии констант движения у невозмущенной системы. В [9] и, например, в [10] указано, что в таком случае при замене, в которой константы движения будут новыми переменными, последние являются медленными переменными.

В частности, у нас частоты аллелей являются константами движения невозмущенной модели динамики генетического состояния популяции под действием рекомбинаций при случайном скрещивании (см., например, [4]). Они же оказались медленными переменными при возмущении селективно нейтральной модели слабым отбором. Аналогична ситуация в многолокусном случае [8].

ОБ ОПРЕДЕЛЕНИИ ЭПИСТАЗА

Коснемся проблемы эпистаза подробнее. Общий случай отсутствия эпистаза в широком смысле при детерминации рассматриваемого признака многими локусами можно определить как такую ситуацию, когда значение признака находится каким-то образом по однолокусным характеристикам генотипов. Например, для такого признака как жизнеспособность эпистаза не будет в случае аддитивного (мультипликативного) отбора, когда для каждого многолокусного генотипа она определяется как сумма (произведение) однолокусных компонентов — характеристик жизнеспособности на уровне отдельных локусов. Аддитивная детерминация особо важна и выделяется среди прочих случаев эпистаза тем, что при наличии отклонений от нее именно аддитивная составляющая эффектов генов (аддитивный генотип) определяет среднее значение рассматриваемого признака у потомков. В связи с этим мы из различных определений эпистаза [11] предпочитаем вслед за Фишером [12] определять эпистаз как нарушения аддитивности вкладов генов в величину признака.

При наличии эпистаза возникает проблема его количественной характеристики. Поскольку в качестве точки отсчета мы выбрали простейший случай аддитивной по локусам детерминации признака, попытаемся измерить отклонение от аддитивности. При этом отклонения от аддитивности вкладов отдельных генов в признак разли-

чаются между генотипами, и появляется задача построения показателя, характеризующего такие отклонения в целом для всего рассматриваемого генофонда. Первое, что приходит в голову, это как-то усреднить отклонения по генотипам. Видимо, при этом следует принимать во внимание их распространенность (частоту) в популяции и определять подобные характеристики как соответствующие средние. Заметим, что общее усреднение дает неинформативное нулевое значение для отклонений, как показано ниже.

ЭЛЕМЕНТАРНЫЙ ГЕНОФОНД И КОЭФФИЦИЕНТ НЕАДДИТИВНОСТИ ЭПИСТАЗА

Очевидно, об эпистазе невозможно говорить по данным о величине признака только для одного генотипа. Показатель отклонения от аддитивности эпистаза для признака должен базироваться на генофонде, детерминирующем изменчивость величины признака v . В простейшей двулокусной диаллельной гаплоидной ситуации элементарный генофонд содержит по два аллеля в каждом из двух локусов с номерами i_1, i_2, j_1, j_2 , нижний индекс дает номер локуса, а сама индексированная переменная — номер аллеля данного локуса. В этой ситуации имеются четыре гаплоидных генотипа, которые можно формально выписать как результат раскрытия скобок выражения

$$(i_1 + j_1)(i_2 + j_2) = i_1i_2 + i_1j_2 + j_1i_2 + j_1j_2,$$

где сложение символизирует включение в спектр возможных аллелей (и спектр генотипов, порожаемых данным набором аллелей для соответствующего генофонда).

Чтобы судить о нарушении аддитивности, надо сравнить значения признака $v_{i_1i_2}$ (для особи с генотипом i_1i_2) у разных генотипов. Поскольку соответствие $i_1i_2 \rightarrow v_{i_1i_2}$ между генотипами и средними (генотипическими) значениями признака однозначно, спектру генотипов отвечает спектр генотипических значений признаков $v_{i_1i_2}, v_{i_1j_2}, v_{j_1i_2}, v_{j_1j_2}$. Займемся построением показателя неаддитивности двулокусной детерминации признака v в гаплоидной популяции. Напомним, что в случае мультипликативной детерминации $\varepsilon = (u_1 - u_2)(w_1 - w_2)$, где u_i и w_i — мультипликативные компоненты жизнеспособности аллелей (с номерами i_1 и i_2 , принимающими значения 1 и 2 как для первого, так и для второго локуса соответственно).

Рассмотрим произведение $(i_1 - j_1)(i_2 - j_2)$. При формальном раскрытии скобок и замене получаемых произведений (символизирующих генотипы) соответствующими жизнеспособностями v_{ij} получим коэффициенты $\varepsilon_{i_1i_2j_1j_2}$ неаддитивности

эпистаза для элементарного генофонда (содержащего по два аллеля каждого из двух локусов с номерами i_1, i_2, j_1, j_2). Коэффициент

$$\varepsilon_{i_1i_2j_1j_2} \equiv v_{i_1i_2} - v_{i_1j_2} - v_{j_1i_2} + v_{j_1j_2}$$

можно понимать как функцию гаплоидных генотипов i_1i_2 и j_1j_2 . При аддитивной детерминации признака $\varepsilon_{i_1i_2j_1j_2}$ обращается в нуль. Причина этого именно в аддитивности, а не в отсутствии эпистаза в широком смысле, так как при мультипликативной детерминации данный коэффициент отличен от нуля, хотя эпистаза нет.

Ремарка 4. Основными легко проверяемыми подстановкой в определение $\varepsilon_{i_1i_2j_1j_2}$ свойствами коэффициента неаддитивности эпистаза при равновесии по сцеплению будут следующие.

1. При постоянных жизнеспособностях все коэффициенты $\varepsilon_{i_1i_2j_1j_2}$ являются константами, характеризующими свойства генетической детерминации признака на организменном уровне (тонкие свойства эпистаза), тогда как коэффициент эпистаза $\varepsilon_{i_1i_2}$ дает более грубый усредненный показатель.

2. Коэффициенты $\varepsilon_{i_1i_2j_1j_2}$ не изменятся, если к всевозможным значениям признака v прибавить на одну и ту же константу или функцию состояния популяции.

3. При слабом отборе величина коэффициентов $\varepsilon_{i_1i_2j_1j_2}$ как сумма разностей (относительных) жизнеспособностей имеет порядок $O(\mu)$, где μ характеризует давление отбора.

4. В случае аддитивной детерминации признака v все коэффициенты $\varepsilon_{i_1i_2j_1j_2}$ равны нулю.

5. При мультипликативной детерминации признака v коэффициент $\varepsilon_{i_1i_2j_1j_2}$ равен $(u_{i_1} - u_{j_1}) \times (w_{i_2} - w_{j_2})$, где u_{j_1} и w_{i_2} — соответствующие мультипликативные компоненты признака v , определяемого как $v_{i_1i_2} \equiv u_{i_1} w_{i_2}$.

6. Перестановка соответствующих одному и тому же локусу индексов у коэффициента неаддитивности не изменяет значения $\varepsilon_{i_1i_2j_1j_2}$.

7. При совпадении у коэффициента $\varepsilon_{i_1i_2j_1j_2}$ индексов, относящихся к одному и тому же локусу, значение коэффициента будет равно нулю. Это означает, что равенство значений индексов на первом и третьем и/или втором и четвертом местах определяет нулевое значение коэффициента.

8. В случае двух диаллельных локусов единственным ненулевым из коэффициентов $\varepsilon_{i_1i_2j_1j_2}$ будет ε_{1122} — только у него нет совпадающих значений индексов, соответствующих одному и тому же локусу. Этот коэффициент равен показателю неаддитивности ε в диаллельном случае.

9. Сумма коэффициентов $\varepsilon_{i_1 i_2}$ по любому из пары индексов (относящихся к одному и тому же локусу) с весами, равными концентрациям соответствующих аллелей, равна нулю, так как получаемые в итоге маргинальные жизнеспособности сокращаются.

Отсюда следует, что взвешенная сумма коэффициентов $\varepsilon_{i_1 i_2}$ по любому из индексов i_1, i_2 равна нулю. Действительно, $\varepsilon_{i_1 i_2}$ получается взвешенным суммированием $\varepsilon_{i_1 i_2 j_1 j_2}$ по паре индексов разных локусов, а сумма еще по одному индексу означает, что произошло суммирование по двум индексам одного из локусов, сумма по которым даст нуль. Тем самым сумма по любому из индексов квазиравновесия $D_{i_1 i_2}^* = \left(\sum_{j_1, j_2} \varepsilon_{i_1 i_2 j_1 j_2} x_{j_1} y_{j_2} \right) \frac{1}{r} x_{i_1} y_{i_2}$ равна нулю, как и в случае обычного коэффициента неравновесности $D_{i_1 i_2}$.

10. Коэффициент неаддитивности $\varepsilon_{i_1 i_2 j_1 j_2}$ также интерпретируется как оценка отклонения от аддитивности, получаемая при сравнении значений признака, порождаемых генотипами $i_1 i_2$ и $j_1 j_2$ двух особей (аллелями элементарного генофонда).

Согласно последнему свойству коэффициента $\varepsilon_{i_1 i_2 j_1 j_2}$ его можно интерпретировать как отвечающую генотипам двух особей характеристику, где генотип одной особи фиксирован, а вторая особь случайно выбрана из популяции. Такая постановка позволяет говорить о средней неаддитивности отдельного генотипа гаплоидной особи, скажем, $i_1 i_2$ при усреднении по случайному выбору второй особи.

Пример. Продолжим измерение эпистаза в рамках концепций количественной генетики. Как уже говорилось, *эффекты генотипов (и аллелей)* находятся как средние отклонения величины признака у особей, несущих рассматриваемый аллель (соответственно генотип), от популяционного среднего. Средняя величина признака у таких особей вычисляется при условии, что учитываются только особи данного генотипа (или генотипа, содержащего данный аллель). Обозначим для элементарного генофонда со спектром значений признака $v_{AB}, v_{Ab}, v_{aB}, v_{ab}$ средние эффекты аллелей A и B (первого и второго локуса соответственно) как δA и δB . При их вычислении учтем, что здесь частота каждого из аллелей (и генотипов) равна $1/2$ (и $1/4$), эффекты δ рассматриваем как условные средние.

Например, δA является отклонением средней жизнеспособности генотипа при условии, что он содержит аллель A . Поэтому при вычислении δA используются условные вероятности $P(g|A) \equiv P(gA)/P(A)$, где g – событие наличия у особи генотипа g , gA обозначает, что данный генотип g содержит аллель A , $P(A)$ – вероятность события об-

ладания аллелем A . В рассматриваемой ситуации $P(gA) = 1/4$ как вероятность одного из четырех равновероятных генотипов, $P(A) = 1/2$ как вероятность одного из двух аллелей первого локуса

(аналогично $P(B) = 1/2$). Отсюда $P(g|A) = \frac{1/4}{1/2} =$

$\frac{1}{2}$, и δA (ожидаемое отклонение жизнеспособности особей-носителей A от популяционного среднего v) определяется как

$$\delta A \equiv \sum_{gA} v_{gA} \frac{P(gA)}{P(A)} - v = \left(v_{AB} \frac{1}{2} + v_{Ab} \frac{1}{2} \right) - v = (v_{AB} + v_{Ab} - 2v) \frac{1}{2}, \quad \delta B \equiv (v_{AB} + v_{aB} - 2v) \frac{1}{2}. \quad (6)$$

При аддитивной детерминации признака v или при нарушении ε_{AB} аддитивности значения δ_{AB} записываются соответственно как

$$\delta_{AB} \equiv \delta A + \delta B, \quad \delta_{AB} \equiv \delta A + \delta B + \varepsilon_{AB},$$

где ε_{AB} – коэффициент взаимодействия аллелей разных локусов (эпистаза). Таким образом, показатель неаддитивности можно определить как

$$\varepsilon_{AB} \equiv \delta_{AB} - (\delta A + \delta B). \quad (7)$$

Коэффициент ε_{AB} имеет ясный смысл отклонения значения признака от средней величины v ($v = \frac{1}{4} (v_{AB} + v_{Ab} + v_{aB} + v_{ab})$), вызванного эпистатическим взаимодействием аллелей. Таким образом, это отклонение дополняет отклонение аддитивного генотипа $\delta A + \delta B$. Вычислим ε_{AB} :

$$\begin{aligned} \varepsilon_{AB} &= \delta_{AB} - (\delta A + \delta B) = v_{AB} - v - (\delta A + \delta B) = \\ &= v_{AB} - v - \frac{1}{2} ((v_{AB} + v_{Ab} - 2v) - \\ &\quad - \frac{1}{2} (v_{AB} + v_{aB} - 2v)) = \\ &= \frac{1}{2} (2v_{AB} - 2v - 2v_{AB} - v_{Ab} - v_{aB} + 4v) = \\ &= \frac{1}{2} \left(-v_{Ab} - v_{aB} + 2 \frac{1}{4} (v_{AB} + v_{Ab} + v_{aB} + v_{ab}) \right) = \\ &= \frac{1}{4} (v_{AB} - v_{Ab} - v_{aB} + v_{ab}) = \frac{1}{4} \varepsilon. \end{aligned}$$

Значение ε_{AB} можно также записать как $\frac{1}{4} \varepsilon_{ABab}$.

Множитель $\frac{1}{4}$ равен произведению частот аллелей разных локусов генотипа AB .

КОЭФФИЦИЕНТ НЕАДДИТИВНОСТИ ЭПИСТАЗА НА ПОПУЛЯЦИОННОМ УРОВНЕ

Если в случае элементарного генофонда или двух диаллельных локусов можно при постоянных жизнеспособностях говорить о коэффициенте эпистаза $\varepsilon_{i_1 i_2 j_1 j_2}$ как о константе, характеризующей детерминацию признака на уровне организма, то средний коэффициент $\varepsilon_{i_1 i_2}$ для популяции с множественными аллелями будет не константой, а одновременной характеристикой как организмов, так и популяции. Он служит общим и грубым показателем генетической детерминации жизнеспособности, тогда как элементарные коэффициенты $\varepsilon_{i_1 i_2 j_1 j_2}$ дают ее более детализированную структуру. При усреднении элементарных коэффициентов неаддитивности приходится учитывать концентрации генотипов в популяции, и не имеет смысла обсуждать средний эффект, скажем, генотипа без указания к какой популяции он относится.

Например, при условии многолокусных соотношений Харди–Вайнберга в гаплоидной популяции (при $\mathbf{D} = \mathbf{0}$) такой средней неаддитивностью эпистаза для жизнеспособности генотипа $i_1 i_2$ будет знакомый нам показатель $\varepsilon_{i_1 i_2} \equiv \sum_{j_1, j_2} \varepsilon_{i_1 i_2 j_1 j_2} x_{j_1} y_{j_2}$. Здесь учтена вероятность $x_{j_1} y_{j_2}$ случайно выбранной для сравнения жизнеспособностей особи иметь генотип $j_1 j_2$. Значит, в найденное в (5) квазиравновесное значение $D_{i_1 i_2}^*$ входит коэффициент эпистаза $\varepsilon_{i_1 i_2}$, являющийся характеристикой популяционного уровня и имеющий в рамках стандартных концепций количественной генетики одну и ту же интерпретацию среднего эффекта неаддитивности и в случае двух, и множественных аллелей у пары локусов.

В простейшем случае аддитивной детерминации жизнеспособности коэффициент $\varepsilon_{i_1 i_2}$ равен нулю. При ее мультипликативной детерминации он равен $(u_{i_1} - u_2)(w_{i_2} - w_2)$:

$$\begin{aligned} \varepsilon_{i_1 i_2} &= \sum_{j_1, j_2} \varepsilon_{i_1 i_2 j_1 j_2} x_{j_1} y_{j_2} = \\ &= \sum_{j_1, j_2} (u_{i_1} - u_{j_1})(w_{i_2} - w_{j_2}) x_{j_1} y_{j_2} = \\ &= (u_{i_1} - u_1)(w_{i_2} - w_2), \end{aligned}$$

где u_1 и w_2 — взвешенные средние значения мультипликативных компонентов жизнеспособности первого и второго локусов.

Напомним, что общее среднее значение коэффициента неаддитивности $\varepsilon_{i_1 i_2 j_1 j_2}$ в популяции равно нулю, поскольку среднее для каждого члена в определении $\varepsilon_{i_1 i_2 j_1 j_2} \equiv v_{i_1 i_2} - v_{i_1 j_2} - v_{j_1 i_2} + v_{j_1 j_2}$ равно v .

Следовательно, данная характеристика эпистаза неинформативна. Чтобы избежать подобного, можно использовать средний квадрат коэффициента неаддитивности (в силу нулевой средней для коэффициента он совпадает с его дисперсией). На этом пути можно двигаться далее и разложить дисперсию на компоненты, соответствующие типам эффектов.

ДВУЛОКУСНЫЕ КВАЗИРАВНОВЕСИЯ И МНОГОЛОКУСНЫЕ МОДЕЛИ

В реальных исследованиях природных популяций характер генетической детерминации жизнеспособности особей неизвестен в полном объеме, и нельзя гарантировать, что изучены все влияющие на нее локусы. Какой смысл в этом случае можно придать коэффициентам эпистаза, полученным при изучении лишь двулокусных генотипов и выражающимся в терминах жизнеспособностей и эффектов аллелей данных локусов? Очевидно, при оценивании жизнеспособности генотипов с помощью случайной выборки получаемые результаты характеризуют маргинальные жизнеспособности, усредненные по (пусть и неизвестным) многолокусным генотипам, содержащим в качестве своей части генотипы двух рассматриваемых локусов. Такое усреднение происходит неявно и автоматически при обработке случайной выборки данных по жизнеспособности.

Подобное неявное (без расшифровки генотипов особей в полном объеме) усреднение не означает чего-либо катастрофического. Фактически использованные при вычислении коэффициентов эпистаза эффекты аллелей являются средними для особей, генотипы которых содержат соответствующий аллель, и никак не используют информацию о прочем составе генотипа [13]. Значит, оценивание средних эффектов аллелей не требует дополнительных сведений о генотипе особей. Поэтому полученные оценки эпистаза для двух рассматриваемых локусов не опираются на дополнительные данные по другим локусам, а основываются на маргинальных жизнеспособностях, для вычисления которых нужно использовать сведения только по генотипам двух рассматриваемых локусов. При этом используемая ранее фраза о зависящей от двух локусов жизнеспособности не запрещает зависимости и от других локусов, лишь бы она не противоречила условию слабого отбора. Таким образом, двулокусные коэффициенты эпистаза следует понимать как маргинальные двулокусные характеристики многолокусных жизнеспособностей.

Это, например, подтверждается совпадением формул для коэффициентов квазиравновесия для заданного двулокусного генотипа при диаллельных локусах и при множественных (5) аллелях, поскольку они одинаково выражаются в терми-

нах эффектов жизнеспособностей (генотипа и аллелей). Вычисление данных характеристик требует знания только соответствующей части генотипа особи и классификации особей лишь по аллелям этой части. В результате найденные коэффициенты эпистатического взаимодействия на уровне двух локусов можно использовать при оценивании эпистаза для трех локусов и т.д., последовательно повышая уровень по числу локусов. Как показали В. Юэнс и Г. Томсон [13] для модели с дискретным временем, все подсистемы полного множества локусов удовлетворяют соответствующим стандартным уравнениям динамики (и условиям равновесия), в которых используются маргинальные жизнеспособности. На этом пути рассмотренное оценивание эпистаза можно попытаться обобщить на случай нескольких локусов.

ДИПЛОИДНЫЕ ПОПУЛЯЦИИ

В диплоидной популяции со случайным скрещиванием уравнения динамики генетической структуры (на уровне гамет), рассматриваемой в отношении двух диаллельных аутосомных локусов, формально совпадают с уравнениями (1) для гаплоидного случая, но теперь жизнеспособность i -го гаплоидного генотипа (гаметы) будет условной средней:

$$\begin{aligned} dp_{i_1 i_2} / dt &= p_{i_1 i_2} (v_{i_1 i_2}(\mathbf{p}) - v(\mathbf{p})) - r(-1)^{i_1 + i_2} D = \\ &= p_i (v_i(\mathbf{p}) - v(\mathbf{p})) - r(-1)^{i_1 + i_2} D; \\ v_i(\mathbf{p}) &= \sum_j v_{ij}(\mathbf{p}) p_j, \\ v(\mathbf{p}) &\equiv \sum_i v_i(\mathbf{p}) p_i = \sum_{i,j} v_{ij}(\mathbf{p}) p_i p_j. \end{aligned} \quad (8)$$

Здесь $\mathbf{p}$ — вектор частот гаплоидных гамет, произведения которых (p_i на p_j) дают концентрации p_{ij} диплоидных генотипов ij при случайном скрещивании, (маргинальная) жизнеспособность i -й гаметы $v_i(\mathbf{p}) \equiv v_{i i_2}(\mathbf{p})$ определяется как средняя жизнеспособность диплоидных особей при условии, что их генотип порожден случайным слиянием в зиготе i -й гаметы и какой-либо еще. Напомним вычисление жизнеспособности $v_i(\mathbf{p})$ i -й гаметы как условной средней при жизнеспособностях $v_{ij}(\mathbf{p}) \equiv v_{i_1 i_2 j_1 j_2}(\mathbf{p})$ генотипов ij :

$$\begin{aligned} v_i(\mathbf{p}) &= \sum_j v_{ij}(\mathbf{p}) P(i|j) = \sum_j v_{ij}(\mathbf{p}) \frac{p_{ij}}{p_i} = \\ &= \sum_j v_{ij}(\mathbf{p}) \frac{p_{ij}}{p_i} = \sum_j v_{ij}(\mathbf{p}) p_j; \\ i &= (i_1, i_2), \quad j = (j_1, j_2). \end{aligned}$$

Итак, уравнение динамики генетической структуры диплоидной популяции (8) согласует-

ся с уравнением (1) для гаплоидного случая, однако теперь жизнеспособности гамет являются условными популяционными средними, т.е. даже при постоянных жизнеспособностях диплоидных особей они будут функциями генетического состояния $\mathbf{p}$.

Тем не менее независимо от числа аллелей рассматриваемых локусов 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 $\bar{\epsilon}$) — как зависимости от их определяемых генетическим состоянием $\mathbf{p}$ усредненных значений, поскольку все манипуляции над (1) для получения результатов допускали зависимость жизнеспособностей от $\mathbf{p}$. Теперь при постоянных жизнеспособностях и диаллельных локусах в уравнениях динамики медленных частот аллелей в квазиградиентном виде (см., например, [1]) потенциалом будет не средняя жизнеспособность v , а $\frac{1}{2}v$.

Уравнения динамики при слабом отборе по-прежнему будут сингулярно возмущенными, качественная картина динамики сохраняется с теми же быстрыми и медленными переменными. Первое приближение для неравновесности по сцеплению, например, для двух диаллельных локусов записывается как

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

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

При этом медленная динамика частот аллелей (подчеркнем, что ее анализ лежит за рамками настоящей работы) теперь допускается при постоянных жизнеспособностях существование полиморфного равновесия — нового качественного свойства, отсутствующего в модели гаплоидной популяции.

ОБСУЖДЕНИЕ

В [8] показано, что качественно сходный характер динамики генетической структуры при слабом отборе присущ диплоидной популяции, рассматриваемой в отношении не двух, а произвольного числа аутосомных локусов с множественными аллелями. Причина сходства в том, что уравнения слабого отбора и здесь являются сингулярно возмущенными. При этом приближенно медленная эволюция частот аллелей происходит так, что средняя жизнеспособность v возрастает, если жизнеспособности генотипов постоянны и генетическая структура неравновесна. Более того, средняя жизнеспособность является

потенциалом градиентного движения, что впервые показал для однолокусного случая Ю.М. Свирежев (см., например, [4, 14, 15]). И здесь средняя жизнеспособность определяется при условии квазиравновесного в нулевом приближении значения быстрых переменных, т.е. при нулевой неравновесности по сцеплению, соответствующей независимому комбинированию аллелей в гаметах.

Полученное уточненное значение квазиравновесия в модели с двумя диаллельными локусами означает малое отклонение от равновесия по сцеплению. На эволюционном этапе при текущих значениях частот аллелей x и y приближенным значением отклонения (коэффициента неравновесности) будет в нулевом приближении $D_{i_1 i_2} = 0$ с точностью $O(\mu)$. Уточненным значением с точностью $O(\mu^2)$ будет первое приближение

$$D_{i_1 i_2}^* = \varepsilon_{i_1 i_2} \frac{\mu}{r} x_{i_1} y_{i_2}.$$

Настоящее исследование идейно мотивировано работой М. Кимуры [5], но отличается подходом и непрерывным временем модели. М. Кимура исследовал динамику генетического состояния популяции в отношении двух диаллельных локусов при слабом отборе в дискретном времени при неперекрывающихся поколениях. Проведенное им компьютерное моделирование, во многом подкрепленное аналитически, выявило качественную картину поведения, совпадающую с описанной для случая сингулярных возмущений. Из результатов изучения М. Кимурой [5, формула (14)] динамики показателя неравновесности по сцеплению z ,

$$z \equiv p_{11}p_{22}/(p_{12}p_{21}) = 1 + D/(p_{12}p_{21}),$$

можно вывести сходную с полученной нами ранее [1] количественную оценку квазиравновесия D^* . Поскольку D и z выражаются одно через другое, между их динамикой должны были наблюдаться параллели. Неудивительно, что изученная в [5] динамика показателя неравновесности z оказалась соответствующей приведенной нами качественной картине сингулярного возмущения отбором модели динамики генетической структуры популяции в результате рекомбинаций и случайного скрещивания.

Отличие нашего подхода состоит в применении стандартных методов анализа, не требующих изобретения в конкретных случаях специальных приемов, а также в использовании в явном виде коэффициента степени давления отбора μ (показывающего будет отбор слабым или сильным). Коэффициент μ рассматривается в качестве малого параметра, r предполагается фиксированным и независимым от генотипа, откуда отношение μ к коэффициенту рекомбинации r также является малым параметром. Именно оно играет

ведущую роль в возможности охарактеризовать динамику генетического состояния популяции как сингулярно возмущенную, а не отношение ε/r . Использование медленных и быстрых переменных позволяет без решения уравнений динамики представить качественные особенности поведения траекторий при слабом отборе.

Ремарка 5. Если μ и отношение μ/r малы, то мало и ε , но данное отношение может быть велико даже при нулевом значении ε . Например, при аддитивном по локусам отборе, когда эффекты отдельных локусов суммируются в общей жизнеспособности, коэффициент неаддитивности эпистаза ε равен нулю. Это выполняется как при слабой, так и при сильной (с большей величиной различий между жизнеспособностями) степени отбора μ , при которой μ/r уже невозможно рассматривать как малый параметр, а уравнения — как сингулярно возмущенные. То же самое остается верным и при слабом нарушении аддитивности жизнеспособности, когда скорость изменений частот аллелей может быть велика из-за давления отбора, на которое не наложено ограничений, а ε остается малым в силу почти аддитивной жизнеспособности. Таким образом ε может быть малым как при сильном, так и слабом отборе, но при условии слабого отбора ε может быть только малым. Действительно, при таком условии разности (относительных) жизнеспособностей генотипов малы, а отсюда следует малость ε как суммы двух таких разностей.

Разработанный М. Кимурой специфический прием анализа данного конкретного случая трудно обобщить для применений в других ситуациях и оценить точность получаемых результатов. Все это послужило дополнительным стимулом для применения в настоящей статье хорошо разработанной к настоящему времени теории малых возмущений с использованием стандартных уравнений для получения приближенного описания поведения модели генетической системы с заданной погрешностью. Кроме того, при этом достигается интерпретация результатов в терминах классической количественной генетики.

Обобщение анализа двулокусной модели на случай множественных аллелей показывает, что выражения для полученных показателей квазиравновесности в терминах эффектов генов идентичны. При записи коэффициентов квазиравновесности по сцеплению через средние эффекты аллелей и их взаимодействия на уровне популяции видно, что она не зависит от числа аллелей.

СПИСОК ЛИТЕРАТУРЫ

1. Пасек В.П. Квазиравновесие по сцеплению при слабом двулокусном отборе по жизнеспособности. I. Гаплоидная популяция с диаллельными локусами // Генетика. 2019. Т. 55. № 4. С. 468–484.

2. Любич Ю.И. Основные понятия и теоремы эволюционной генетики свободных популяций // Успехи матем. наук. 1971. Т. 26. № 5. С. 51–116.
3. Любич Ю.И. Математические структуры в популяционной генетике. Киев: Наук. думка, 1983. 296 с.
4. Свирежев Ю.М., Пасеков В.П. Основы математической генетики. М.: Наука, 1982. 511 с.
5. Kimura M. Attainment of quasi linkage equilibrium when gene frequencies are changing by natural selection // Genetics. 1965. V. 52. № 5. P. 875–890.
6. Фолкнер Д.С. Введение в генетику количественных признаков. М.: Агропромиздат, 1985. 485 с.
7. Новожиллов И.В. Методы разделения движений. Конспект лекций. М.: МЭИ, 1981. 48 с.
8. Пасеков В.П. Асимптотический анализ отбора в полилокусной полиаллельной популяции // Докл. АН СССР. Кибернетика и теория регулирования. 1984. Т. 277. № 6. С. 1338–1341.
9. Пасеков В.П. Математические модели эколого-генетических взаимодействий / Итоги науки и техники. Серия “Математическая биология и медицина”. Т. 2. М.: ВИНТИ, 1988. С. 4–80.
10. Auger P.M. Hierarchically organized systems: The response to complexity // J. Sci. Industrial Res. 1992. V. 51. P. 725–735.
11. Cordell H.J. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans // Hum. Mol. Genet. 2002. V. 11. Issue 20. P. 2463–2468. doi 10.1093-hmg-11.20.2463
12. Fisher R.A. The correlation between relatives on the supposition of Mendelian inheritance // Trans. Roy. Soc. Edinburgh. 1918. V. 52. Part 2. P. 399–433.
13. Ewens W.J., Thomson G. Properties of equilibria in multi-locus genetic systems // Genetics. 1977. V. 87. № 4. P. 807–819.
14. Shahshahani S. A new mathematical framework for the study of linkage and selection // Memoirs Amer. Math. Soc. 1979. V. 7. № 211. 34 p.
15. Akin E. The Geometry of Population Genetics. Berlin e.a.: Springer, 1979. 205 p.

Quasi Linkage Equilibrium under Weak Two-Locus Viability Selection. II. Loci with Multiples Alleles

V. P. Passekov*

*Dorodnicyn Computing Centre, Federal Research Center “Computer Science and Control” of Russian Academy of Sciences,
Moscow, 119991 Russia*

*e-mail: pass40@mail.ru

The model of weak viability selection at two loci with multiple alleles in a haploid population is considered. Random mating and non-overlapping generations are assumed. The dynamics of the genetic structure is modeled by ordinary differential equations. Weak selection is considered as a perturbation of the model without selection. The estimates of quasi-equilibrium value of the coefficient of linkage disequilibrium D (which are obtained for diallelic loci in the first part of the paper) are generalized to the case of multiple alleles. The interpretation of the estimates in terms of average effects in quantitative genetics turned out to be independent

of the number of alleles. The first approximation for the quasi-equilibrium $D_{i_1 i_2}$ is $D_{i_1 i_2}^* = \varepsilon_{i_1 i_2}(\mathbf{p}) \frac{\mu}{r} x_{i_1} y_{i_2}$. Here μ is the selection intensity, r is the coefficient of recombination, $\varepsilon_{i_1 i_2}(\mathbf{p})$ is the index of the epistasis non-additivity for the viabilities $v_{i_1 i_2}$ of the haploid genotypes $i_1 i_2$, x_{i_1} and y_{i_2} are the frequencies of alleles constituting the genotype $i_1 i_2$ (allele numbers are i_1, i_2 for the first and the second locus, respectively). The specificity of the perturbed model is that it have fast (the coefficients of disequilibrium $D_{i_1 i_2}$) and slow (allele frequencies) variables. This characterizes the model as singularly perturbed and allows us to separate two stages in the dynamics without solving the equations. On the first stage the fast variables converge to quasi-equilibrium values, and on the second one the slow variables evolve under the condition of quasi-equilibrium values of the fast variables. The equations in the quasi-equilibrium values of $D_{i_1 i_2}$ are obtained by the delta-method. They give a correct result from the point of view of expansion in powers of the small parameter of the solution of singularly perturbed equations. The problems of measuring the epistasis and decomposition of the coefficient of epistasis into elementary components are discussed and the two-locus coefficients of the epistasis are considered as characteristics of the multi-locus model. The generalization of the results obtained for the haploid population to a diploid case is briefly indicated.

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