

КВАЗИРАВНОВЕСИЕ ПО СЦЕПЛЕНИЮ ПРИ СЛАБОМ ДВУЛОКУСНОМ ОТБОРЕ ПО ЖИЗНЕСПОСОБНОСТИ. I. ГАПЛОИДНАЯ ПОПУЛЯЦИЯ С ДИАЛЛЕЛЬНЫМИ ЛОКУСАМИ

© 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.

Изучена модель слабого отбора по жизнеспособности по двум диаллельным локусам со стандартизацией подходов за счет использования теории возмущений; дан анализ оценки квазиравновесия для коэффициента D неравновесности по сцеплению; результаты получены в терминах средних эффектов в количественной генетике и в терминах теории сингулярных возмущений в математике. Рассматривается аппроксимация обыкновенными дифференциальными уравнениями модели с дискретным временем случайно скрещивающейся популяции с неперекрывающимися поколениями при слабом отборе. Слабый отбор рассматривается как возмущение модели без отбора. Получаемая модель является сингулярно возмущенной, т.е. в ней выделяются быстрые (D) и медленные (частоты аллелей) переменные. Уравнение первого приближения для квазиравновесия D получено с использованием первых членов разложения в ряд Тейлора функций модели. Оно совпадает с соответствующей частью системы первого приближения асимптотического ряда для решения сингулярно возмущенных уравнений. Первым приближением для квазиравновесия D будет $D^* = \varepsilon(\mathbf{p}) \frac{\mu}{r} x(1-x)y(1-y)$, $\varepsilon(\mathbf{p}) \equiv v_{11}(\mathbf{p}) - v_{12}(\mathbf{p}) - v_{21}(\mathbf{p}) + v_{22}(\mathbf{p})$, где μ – интенсивность отбора; r – коэффициент рекомбинации; $\varepsilon(\mathbf{p})$ – коэффициент неаддитивности эпистаза для жизнеспособностей v_{ij} ; x, y – частоты аллелей первого и второго локусов соответственно. Эволюция неравновесного состояния популяции фактически протекает при квазиравновесии (при итоге этапа динамики быстрых переменных D) и текущем значении медленных x, y .

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

DOI: 10.1134/S0016675819040118

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 количественного признака. С точки зрения микроэволюции генетической структуры популяции особый интерес представляет собой такой количественный признак как жизнеспособность особей, поскольку она самым очевидным образом влияет на динамику генетической структуры популяции.

На заре становления теоретической популяционной генетики были установлены основные простые свойства генетической структуры при случайном скрещивании и отсутствии отбора – закон Харди–Вайнберга в ситуации одного локуса и предельная с течением времени независимость комбинирования генов в генотипах для нескольких локусов (*многолокусные соотношения Харди–Вайнберга, или равновесие по сцеплению*). В двулокусном случае доказательство сходимости к равновесию по сцеплению в модели давления рекомбинаций элементарно. Для любого числа аутосомных локусов диплоидной популяции со случайным скрещиванием это будет непростой

задачей, которая доведена до явного решения с использованием генетических алгебр Ю.И. Любичем [3, 4]. Картина динамики качественно охарактеризована более простыми методами в [5]. При дальнейшем развитии теории и практики генетических исследований выяснилась возможность нарушения предельной независимости сочетаний генов в генотипах, например из-за отбора. Данное нарушение весьма обременительно для анализа генетической структуры, так как резко усложняет его.

В частности, анализ такого важного фактора микроэволюции как отбор становится еще более затруднительным. Для нелинейной модели отбора по жизнеспособности неизвестно точное решение даже в однолокусном диплоидном случае. Анализ качественных характеристик моделей и их линейной аппроксимации вблизи состояний равновесия продвинуто в большей степени (см., например, [6, 7]). Наиболее ярко качественные свойства динамики генетической структуры популяции в случае одного локуса и постоянных жизнеспособностей характеризуются фундаментальной теоремой естественного отбора Р.А. Фишера [8], согласно которой в результате отбора неравновесное состояние изменяется в направлении увеличения средней жизнеспособности популяции.

Нарушение свободного сочетания генов в генотипах и гаметах приводит к невозможности использования теоремы Фишера при анализе динамики многолокусной генетической структуры. В связи с этим вызывает особый интерес поиск условий, при которых можно использовать в моделях, хотя бы приближенно, равновесие по сцеплению. Так, С. Райт [9] давно указывал, что в случае слабого по сравнению с рекомбинациями отбора можно пользоваться предположением о независимости комбинирования генов в многолокусных генотипах. Однако строгого обоснования этому не было дано.

Для модели с непрерывным временем отбора по жизнеспособности в популяции с перекрывающимися поколениями без возрастной структуры такое обоснование было представлено автором [10] на основе выделения быстрых и медленных переменных. Позже появилось доказательство и для модели с дискретным временем [11]. Во всех подобных обоснованиях случайное комбинирование генов рассматривалось лишь как приближенное и отклонение от него изучено недостаточно.

В 1965 г. вышла работа М. Кимуры [12], 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 обоснова-

ние этих результатов. С тех пор к проблеме квазиравновесия привлекались теоретические соображения в основном качественные на базе известных в математике концепций геометрической теории сингулярных возмущений. Такие соображения [13] обосновывают существование квазиравновесия по сцеплению, близкое к почти независимому сочетанию генов в генотипах, но все-таки не оценивают его количественно явным образом.

Статья М. Кимуры привлекла к себе внимание, ее цитируют более чем в полутора сотнях работ, но до сих пор она мало развивалась в отношении количественных характеристик квазиравновесия. Она также не обобщалась на случай множественных аллелей и локусов, не получила интерпретации результатов в терминах стандартных концепций количественной генетики, а также не была определена методологическая координата анализа исследуемой проблемы в русле развитых математических методов и подходов. Среди таких подходов, в частности, привлекательны методы анализа моделей сложных систем как возмущений более простых и легче поддающихся изучению.

К настоящему времени уже немало исследователей прилагали усилия к получению точных решений многих ранее сформулированных моделей, и пока не часто получается говорить об увеличении шансов добиться успеха на этом пути. Здесь кажется перспективным отыскание приближенных решений с заданной точностью. Существуют различные методы приближенного анализа популяционно-генетических моделей, среди которых рассмотрим метод разделения переменных на быстрые и медленные. Его эффективность при изучении слабого отбора по отношению к рекомбинациям и/или миграциям продемонстрирована в ряде работ [10, 14–17]. Биологические приложения рассматривались также в [18, 19], где приведена обширная библиография. Однако в ней немногочислены приложения к проблемам популяционной генетики. В настоящей работе мы применим метод разделения переменных по скорости к анализу квазиравновесия по сцеплению в случае слабого отбора по жизнеспособности, зависящей от двух локусов.

Цель настоящей серии из двух сообщений состоит в изучении квазиравновесия по сцеплению для неравновесности D при слабом отборе по жизнеспособности по двум диаллельным локусам с использованием теории возмущений, в обобщении анализа результатов для диаллельных локусов на случай множественных аллелей, в стандартизации подходов к изучению слабого влияния систематических факторов микроэволюции. Попутно решаются задачи получения оценки квазиравновесия в более удобном виде как в плане использования стандартного коэффициента D неравновесности по сцеплению, так и интерпретации результатов в

терминах привычных концепций средних эффектов в количественной генетике и в терминах теории сингулярных возмущений в математике. Данное (первое) сообщение целиком посвящено анализу двух диаллельных локусов.

Статья построена по следующему плану. Сначала рассматривается заданная разностными уравнениями модель динамики генетической структуры двулокусной популяции с неперекрывающимися поколениями. В следующем разделе обосновывается ее аппроксимация обыкновенными дифференциальными уравнениями. Далее описывается влияние на генетическую структуру отбора по жизнеспособности и рекомбинации и детально анализируется гаплоидная популяция с двумя диаллельными локусами. Затем в двух центральных разделах рассматривается модель слабого отбора как возмущение нейтральной модели без отбора. Для нее найдена приближенная квазиравновесная величина стандартного коэффициента неравновесности по сцеплению D в виде, допускающем интерпретацию в терминах дисперсий встречаемости аллелей в популяции и коэффициента взаимодействия аллелей (эпистаза) при детерминации такого количественного признака как жизнеспособность, а также в терминах эффектов генов и их комплексов.

Обсуждаются качественные свойства микроэволюции двулокусной генетической структуры, дающие наглядную картину динамики без точного решения уравнений модели и опирающиеся только на разномасштабность скоростей изменений переменных (выделение быстро меняющегося коэффициента неравновесности по сцеплению D и медленно эволюционирующих частот аллелей). Приблизительно динамика генетической структуры популяции происходит практически при равновесии по сцеплению и в направлении увеличения жизнеспособности популяции, если жизнеспособности постоянны. Отклонения от равновесия по сцеплению (от нулевого приближения, когда $D = 0$) при слабом отборе имеют порядок $O(\mu)$, где μ – малый параметр, характеризующий давление отбора. Напомним, что $O(\mu)$ означает, что отклонения по модулю меньше либо равны $C|\mu|$ для некоторой положительной константы C . Иначе говоря, при уменьшении μ отклонения если и уменьшаются, то не быстрее μ (например, когда существует предел модуля отношения отклонения к μ , он конечен и больше нуля). Полученный далее коэффициент квазиравновесности D^* оценивает величину отклонений и дает следующее приближение с точностью $O(\mu^2)$. Он удовлетворяет уравнениям первого приближения теории сингулярных возмущений, хотя фактически получен дельта-методом.

Во второй части [20] настоящей серии результаты обобщаются на ситуацию, когда локусы с

множественными аллелями. Качественная картина динамики сохраняется и в случае множественных аллелей. Показано, что запись коэффициентов квазиравновесности в терминах эффектов аллелей и их взаимодействий одинаково проста и прозрачна как при диаллельности, так и при множественных аллелях. Указано, что обобщение результатов на диплоидные популяции в сущности не вносит ничего принципиально нового. Обсуждаются свойства коэффициентов неаддитивности эпистаза для элементарных генофондов, соответствующих генотипам пар особей. Коэффициенты неаддитивности эпистаза для эффектов отдельных генотипов, определяющие величину квазиравновесности на популяционном уровне, представляют собой средние значения для элементарных коэффициентов.

В данном сообщении векторы и матрицы набраны полужирным шрифтом. Получаемые выводы сформулированы как результаты, замечания и ремарки, конец соответствующего доказательства отмечается символом ◀.

ДВУЛОКУСНАЯ МОДЕЛЬ С ДИСКРЕТНЫМ ВРЕМЕНЕМ ПОПУЛЯЦИИ И НЕПЕРЕКрываЮЩИМИСЯ ПОКОЛЕНИЯМИ

Использование популяционно-генетических моделей необходимо для понимания и объяснения, предсказания и ретроспективного анализа генетического состояния популяции. Такие модели могут характеризовать динамику генетического состояния, например в терминах концентраций p_i многолокусных гамет (концентраций гаплоидных особей). Далее под гаплоидными генотипами подразумеваем в соответствии с контекстом генотипы гаплоидных особей или генотипы гамет в диплоидной популяции. В настоящем разделе будем рассматривать модели динамики генетической структуры в отношении двух аутосомных локусов преимущественно диплоидных популяций, поскольку модели для гаплоидной ситуации здесь можно рассматривать как их частный случай.

Пусть жизненный цикл исследуемой популяции с неперекрывающимися поколениями имеет вид

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

Эффектами редких мутаций будем пренебрегать в предположении анализа динамики на ограниченном промежутке времени. Начнем с напоминания результатов изучения отправной точки для дальнейших обобщений – простейшего случая популяции со случайным скрещиванием при отсутствии давления систематических факторов (в приведенной схеме опустим отбор по жизнеспособности) и отсутствии эффектов случайного

генного дрейфа (из-за величины популяции). Будем говорить о моделях как селективно *нейтральных*, когда в них не учитывается действие отбора. Поскольку при *случайном скрещивании* происходит независимое слияние гамет в диплоидных генотипах особей (случайный выбор второго гаплоидного генотипа при рекомбинации у гаплоидов), после рекомбинационной перетасовки генов концентрации гаплоидных генотипов соответствуют генетической структуре следующего поколения.

Отсюда ясно, что в простейшем случае динамика частот многолокусных гаплоидных генотипов в ряду поколений обусловлена лишь процессами сегрегации и рекомбинации, т.е. рекомбинациями в широком смысле. Так как в силу независимого слияния гамет в генотипах диплоидных особей частоты генотипов находятся как произведения концентраций соответствующих гаплоидных гамет, генетическое состояние популяции можно задать вектором \mathbf{p} частот многолокусных гаплоидных генотипов как в диплоидной, так и в гаплоидной популяции:

$$\mathbf{p} = \{p_i = p_{i_1 i_2 \dots i_l}\}, \sum_i p_i = 1, 1 \geq p_i \geq 0, \\ \mathbf{i} = (i_1, i_2, \dots, i_l).$$

Здесь координата i_k мультииндекса \mathbf{i} равна номеру аллеля k -го локуса, входящего в \mathbf{i} -ю гамету, l — количество локусов в рассматриваемом множестве локусов L . Условия $\sum_i p_i = 1, 1 \geq p_i \geq 0$ задают симплекс возможных генетических состояний популяции \mathbf{p} . В дальнейшем состояние популяции будем относить к началу поколения (к “новорожденным”, заменяющим поколение родителей).

В наиболее простой модели динамики генетического состояния диплоидной популяции рассматривается множество из одного аутосомного локуса, и состоянием будет вектор частот его аллелей. Согласно закону Харди–Вайнберга при случайном скрещивании частоты аллелей (однолокусных гамет) постоянны в ряду поколений и не зависят от поведения многолокусной системы в целом, куда входит рассматриваемый локус (“динамика” частот однолокусных гамет замкнута, частоты постоянны, так как генетические преобразования сводятся лишь к перераспределению аллелей в генотипах при скрещивании).

Более сложная генетическая система состоит уже из двух аутосомных локусов, скажем, с номерами k и m и с вероятностью рекомбинации (обмена локусами между двумя гаплоидными генотипами (гаметами при диплоидности)) $r > 0$. Приведем для замкнутости изложения основные обозначения и главные известные результаты для двулокусной модели динамики под влиянием рекомбинаций.

Результат 1. Пусть рассматривается случайно скрещивающаяся диплоидная (гаплоидная) популяция с неперекрывающимися поколениями и дискретным временем в отношении концентраций $p_{i_k i_m}$ гаплоидных генотипов двух аутосомных локусов k и m , несущих аллели с номерами i_k, i_m соответственно (в остальном на состав генотипов нет ограничений) и с вероятностью рекомбинации $r > 0$. Тогда

А. Уравнения динамики частот гаплоидных генотипов гамет (особей) $p_{i_k i_m}$ в модели давления рекомбинаций имеют вид

$$p'_{i_k i_m} = p_{i_k i_m} - r(p_{i_k i_m} - x_{i_k} y_{i_m}) \equiv \\ \equiv p_{i_k i_m} - r D_{i_k i_m}(\mathbf{p}), \quad D'_{i_k i_m} = (1 - r) D_{i_k i_m}, \quad (1)$$

где

$$x_{i_k} \equiv \sum_{i_m} p_{i_k i_m}, \quad y_{i_m} \equiv \sum_{i_k} p_{i_k i_m}, \\ D_{i_k i_m}(\mathbf{p}) \equiv p_{i_k i_m} - x_{i_k} y_{i_m} (p_{i_k i_m}(t) = \\ = x_{i_k}(t) y_{i_m}(t) + D_{i_k i_m}(t)), \quad (2)$$

x_{i_k} и y_{i_m} — частоты аллелей k -го и m -го локусов и $D_{i_k i_m}$ — коэффициент (гаметической) неравновесности по сцеплению соответственно, t — время в поколениях, штрих символизирует значения после этапа рекомбинаций.

Б. На траекториях (1) частоты аллелей x_{i_k}, y_{i_m} постоянны (являются константами движения).

В. Решением (1) будет

$$p_{i_k i_m}(t) = x_{i_k}(0) y_{i_m}(0) + (1 - r)^t D_{i_k i_m}(0), \\ D_{i_k i_m}(t) = (1 - r)^t D_{i_k i_m}(\mathbf{p}(0)). \quad (3)$$

В пределе при времени $t \rightarrow \infty D_{i_k i_m}(t) \rightarrow 0, p_{i_k i_m}(t) \rightarrow x_{i_k} y_{i_m}$ и аллели разных локусов сочетаются в генотипах независимо.

Доказательство этих известных фактов (см., например, [5]) опустим.

Укажем следующее свойство неравновесностей. Сумма по любому их двух индексов $D_{i_k i_m}$ равна нулю (оба случая доказываются аналогично):

$$\sum_{i_m} D_{i_k i_m} \equiv \sum_{i_m} p_{i_k i_m} - \sum_{i_m} x_{i_k} y_{i_m} = \\ = x_{i_k} - x_{i_k} = 0 \quad (\text{аналогично } \sum_{i_k} D_{i_k i_m} = 0). \quad (4)$$

ОПИСАНИЕ СЛАБОГО ОТБОРА ПО ЖИЗНЕСПОСОБНОСТИ В ПОПУЛЯЦИИ С НЕПЕРЕКРЫВАЮЩИМИСЯ ПОКОЛЕНИЯМИ

Анализ начнем с описания его изолированно-го давления на выживаемость (на стадии отбора по жизнеспособности). *Жизнеспособность* $v_i(\mathbf{p}, t)$

особей с i -м гаплоидным генотипом понимаем как вероятность особи дожить до стадии размножения. Очевидно, изменение частоты гаплоидного генотипа на этапе отбора определяется вероятностями $v_i(\mathbf{p}, t)$, зависящими в общем случае от генетического состояния популяции \mathbf{p} и времени t .

На стадии отбора по жизнеспособности частоты генотипов (независимо от пloidности) изменяются пропорционально их жизнеспособностям, а чтобы новые значения к концу этапа оставались частотами (чтобы их сумма равнялась единице), надо ввести нормировку, т.е. $p_i' = p_i v_i(\mathbf{p}, t) / \sum_j v_j(\mathbf{p}, t) p_j$. Когда жизнеспособности заданы, изменения частот генотипов за рассматриваемую стадию отбора не зависят от того, рассматривается однолокусный или многолокусный случай и имеют вид

$$\begin{aligned} p_i' - p_i &\equiv \Delta p_i = \frac{v_i(\mathbf{p}, t)}{v(\mathbf{p}, t)} p_i - p_i = \\ &= \frac{v_i(\mathbf{p}, t) - v(\mathbf{p}, t)}{v(\mathbf{p}, t)} p_i, \quad v(\mathbf{p}, t) = \sum_j v_j(\mathbf{p}, t) p_j. \end{aligned} \quad (5)$$

Здесь \mathbf{p} — вектор частот генотипов; $v_i(\mathbf{p}, t)$ — жизнеспособность i -го генотипа. Из (5) видно, что сами по себе величины жизнеспособностей не влияют на темп динамики в том смысле, что если умножить всех их на как угодно малый (большой) положительный множитель (на константу или одну и ту же функцию s состояния и экологических параметров), то приращения Δp_i и динамика останутся прежними. Действительно, при этом числитель и знаменатель (5) получают множитель s , сокращающийся в итоге. Поэтому интерес представляет не сама величина жизнеспособностей, а величина вызываемых ими изменений *генетической структуры* \mathbf{p} , понимаемой как распределение концентраций генотипов в популяции. В последующем будем пользоваться относительными жизнеспособностями, более удобными для характеристики генетической динамики. Под *относительной жизнеспособностью* понимают некоторую нормировку абсолютной жизнеспособности, получаемую обычно путем деления на ее максимальное значение [21].

Далее будем рассматривать достаточно типичную природную ситуацию медленного изменения (микроэволюции) генетической структуры в результате отбора по жизнеспособности. Будем называть отбор *слабым*, если вызываемые изменения генетической структуры за поколение малы при любом генетическом состоянии популяции \mathbf{p} , скажем, имеют порядок $O(\mu)$, $0 < \mu \ll 1$. Поскольку при одинаковых жизнеспособностях отбор в рамках рассматриваемой модели отсутствует, интуитивно понятно, что при малых различиях между ними отбор будет слабым.

Однако различия абсолютных жизнеспособностей можно сделать как угодно малыми при умножении всех их на достаточно малую положительную константу, хотя изменение генетической структуры из-за отбора при этом останется прежним, как уже говорилось ранее. Для относительных жизнеспособностей такой проблемы не существует: умножение на константу не меняет не только приращения концентраций генотипов (степень давления отбора), но сохраняются сами относительные жизнеспособности (и различия между ними).

Ремарка 2. Приращения концентраций генотипов в общем случае индуцируют изменение частот аллелей, гамет и прочих подмножеств генов всего генотипа, но для дальнейшего точный вид таких изменений нам пока не нужен.

В важном частном случае диплоидности степень давления отбора одинакова на уровнях генотипов и гамет: на стадии действия отбора по жизнеспособности изменения концентраций диплоидных генотипов малы, если и только если индуцированные ими изменения частот гаплоидных гамет малы. Так как (маргинальные) частоты гамет выражаются через частоты генотипов путем соответствующего суммирования последних, то когда приращения из-за отбора концентраций диплоидных генотипов малы, того же порядка будут и изменения частот гамет. Аналогично при малости приращений концентраций гаплоидных гамет, через которые гладким образом можно выразить изменения частот диплоидных генотипов (например, по закону Харди–Вайнберга), последние также будут того же порядка.

Значит, для диплоидных популяций отбор является слабым в смысле изменения гаметического состояния \mathbf{p} за поколение тогда и только тогда, когда изменения концентраций диплоидных генотипов малы. Поэтому в дальнейшем доказательстве при упоминании о концентрациях генотипов не требуется конкретизировать их пloidность.

Сфокусируемся на стадии отбора по жизнеспособности. Как отмечено выше, здесь формулы изменения частот генотипов не зависят от количества локусов. Естественно, нет зависимости и от системы скрещивания.

Результат 3. 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 мини-

мальную ($v_{\max} > v_{\min}$) относительные жизнеспособности.

Доказательство. Докажем достаточное условие слабого отбора, т.е. если для любых i, j и \mathbf{p} выполняется неравенство $|v_i(\mathbf{p}) - v_j(\mathbf{p})| \leq C\mu$, то $|\Delta p_i| = O(\mu)$. Так как формально можно считать отбор однолокусным, индексы-номера генотипов будем писать как скалярные. Фиксируем произвольные i, \mathbf{p} и оценим $|\Delta p_i|$:

$$\begin{aligned} |\Delta p_i| &= \left| p_i \frac{v_i(\mathbf{p}, t) - v(\mathbf{p}, t)}{v(\mathbf{p})} \right| = \\ &= \frac{p_i}{v(\mathbf{p})} \left| v_i(\mathbf{p}) - \sum_{j=1}^n v_j(\mathbf{p}) p_j \right| = \\ &= \frac{p_i}{v(\mathbf{p})} \left| \sum_{j=1}^n (v_i(\mathbf{p}) - v_j(\mathbf{p})) p_j \right| \leq \\ &\leq \frac{p_i}{v(\mathbf{p})} \sum_{j=1}^n |v_i(\mathbf{p}) - v_j(\mathbf{p})| p_j \leq \\ &\leq \frac{p_i}{v(\mathbf{p})} C\mu \sum_{j=1}^n p_j \leq C_1 C\mu = O(\mu). \end{aligned}$$

Здесь $\sum_{j=1}^n p_j = 1$, функция $p_i/v(\mathbf{p})$ непрерывна по \mathbf{p} на компактном симплексе, следовательно достигает на нем максимума при некотором \mathbf{p} , и C_1 обозначает его наибольшее по i значение. Таким образом, $|\Delta p_i| = O(\mu)$, т.е. отбор является слабым.

Теперь докажем необходимое условие слабого отбора при постоянных жизнеспособностях генотипов v_i . Оценим сверху константу $v_{\max} - v_{\min}$. Выберем такую генетическую структуру \mathbf{p} , когда оценивание наиболее просто. Пусть в популяции только два генотипа i и j . Тогда $p_j = 1 - p_i$, $v = v_i p_i + v_j p_j = v_i p_i + v_j(1 - p_i)$ и согласно (5)

$$\begin{aligned} |\Delta p_i| &= \frac{p_i}{v(\mathbf{p})} |v_i - v(\mathbf{p})| = \frac{p_i}{v(\mathbf{p})} |v_i - (v_i p_i + v_j p_j)| = \\ &= \frac{p_i}{v(\mathbf{p})} |v_i - (v_i p_i + v_j(1 - p_i))| = \\ &= \frac{p_i}{v(\mathbf{p})} |(v_i - v_j)(1 - p_i)| = \frac{p_i(1 - p_i)}{v(\mathbf{p})} |v_i - v_j|. \end{aligned}$$

При слабом отборе $C\mu \geq \max |\Delta p_i| = |v_i - v_j| C_2 \max \frac{p_i(1 - p_i)}{v(\mathbf{p})}$, откуда $|v_i - v_j| \leq (C/C_2)C\mu$. Полученное неравенство для $|v_i - v_j|$ верно для произвольно выбранных генотипов i и j с любыми частотами, в частности для генотипов с максимальной и минимальной жизнеспособностями и с частотами, доставляющими максимум $p_i(1 - p_i)/v(\mathbf{p}) \equiv C_2$. Значит, $v_{\max} - v_{\min} \leq (C/C_2)\mu$, т.е. $v_{\max} - v_{\min} = O(\mu)$. ◀

Пример. Итак, необходимым и достаточным условием слабого отбора является малость различий между относительными жизнеспособностями генотипов. Приведем иллюстрацию явной зависимости жизнеспособностей от малого параметра μ , когда их абсолютные значения могут быть велики, а различия относительных жизнеспособностей малы. Пусть жизнеспособности $w_{ij}(\mathbf{p}, t)$ имеют вид $w_{ij}(\mathbf{p}, t) = c(\mathbf{p}, t) + \mu v_{ij}(\mathbf{p}, t)$, где функция $c(\mathbf{p}, t)$ положительна, ограничена сверху и снизу константами. В этом случае максимальное различие относительных жизнеспособностей генотипов не превосходит $\mu (\max_{i,j} v_{ij}(\mathbf{p}, t) - \min_{i,j} v_{ij}(\mathbf{p}, t)) = O(\mu)$ и

$$\begin{aligned} \Delta p_i &= \frac{w_i(\mathbf{p}, t) - w(\mathbf{p}, t)}{w(\mathbf{p}, t)} p_i = \mu \frac{v_i(\mathbf{p}, t) - v(\mathbf{p}, t)}{c(\mathbf{p}, t) + \mu v(\mathbf{p}, t)} p_i = \\ &= \mu \frac{v_i(\mathbf{p}, t) - v(\mathbf{p}, t)}{c(\mathbf{p}, t)} p_i + O(\mu^2) = \\ &= \frac{\mu}{c(\mathbf{p}, t)} (v_i(\mathbf{p}, t) - v(\mathbf{p}, t)) p_i + O(\mu^2); \quad 0 < \mu \ll 1. \end{aligned}$$

Далее для определенности будем характеризовать генотипы относительными жизнеспособностями и при слабом отборе записывать их как $\mu v_{ij}(\mathbf{p}, t)$, а приращение частоты генотипа из-за отбора будем писать как

$$\Delta p_i = \mu (v_i(\mathbf{p}, t) - v(\mathbf{p}, t)) p_i + O(\mu^2). \quad (6)$$

Замечание 4 [5]. Пусть в популяции с непрерывно сменяющимися поколениями в течение жизненного цикла последовательно и слабо действуют два (и более) фактора микроэволюции (у нас в многолокусном случае отбор и рекомбинации), т.е. изменение $\Delta \mathbf{p}$ из-за влияния i -го фактора имеет порядок $O(\mu)$, скажем, в нем можно выделить сомножитель μ .

Тогда в модели с разностными уравнениями результат их совместного влияния $\Delta \mathbf{p}$ можно аппроксимировать с точностью $O(\mu^2)$ суммой соответствующих изолированному действию каждого фактора по отдельности вкладов $\Delta_i \mathbf{p}(\mathbf{p}, t)$ как функций одного и того же текущего состояния популяции:

$$\Delta \mathbf{p} = \sum_i \Delta_i \mathbf{p}(\mathbf{p}, t) + O(\mu^2).$$

Доказательство. Представим приближенно приращение $\Delta \mathbf{p}$ как сумму первых членов ряда Тейлора для $\Delta \mathbf{p}$ — сумму вкладов каждого отдельного фактора, малых по предположению (равных $O(\mu)$). Отсюда дополнительное изменение малого значения $\Delta \mathbf{p}$ из-за отдельного фактора за счет приращения другого фактора будет порядка $O(\mu^2)$. ◀

АППРОКСИМАЦИЯ РАЗНОСТНЫХ УРАВНЕНИЙ МОДЕЛИ ОБЫКНОВЕННЫМИ ДИФФЕРЕНЦИАЛЬНЫМИ

Рассмотрим модели динамики генетической структуры популяции типа $\mathbf{p}' = \mathbf{p} + \Delta\mathbf{p}(\mathbf{p}, t)$. В целях упрощения и сокращения доказательств далее перейдем к аппроксимации таких разностных уравнений обыкновенными дифференциальными.

Замечание 5 [5]. Пусть рассматривается класс моделей, описываемых системой разностных уравнений $\mathbf{p}' = \mathbf{p} + \Delta\mathbf{p}(\mathbf{p}, t)$, где величина координат вектора Δ 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 диплоидной (гаплоидной) популяции с неперекрывающимися поколениями этого класса 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 $d\mathbf{p}/dt = \sum_i \Delta_i \mathbf{p}(\mathbf{p}, t)$.

Доказательство. Пусть изменения концентраций гамет $\Delta\mathbf{p}$ за поколение малы (порядка μ , что удобно записать как $\mu\Delta\mathbf{p}$, $0 < \mu \ll 1$). Когда состояние можно описать единственной переменной, динамика во времени геометрически иллюстрируется дискретным графиком, где по оси абсцисс время, а по оси ординат — состояние. Чем меньше изменения состояния \mathbf{p} за поколение, тем меньше отстоят друг от друга соседние точки-состояния по вертикали. При переходе к новому медленно (“эволюционному”) времени s , в котором длительность поколения (единица прежнего времени t) занимает очень небольшой промежуток $\Delta s = \mu$, соседние точки будут мало отстоять друг от друга и по горизонтали и все теснее ложиться друг возле друга, все более напоминая непрерывную кривую.

Если дискретный процесс изменения генетической структуры \mathbf{p} по поколениям описывать в непрерывном времени ломаной, соединяющей значения \mathbf{p} в точках смены поколений, то эта ломаная будет ломаной Эйлера для гладкой интегральной кривой обыкновенного дифференциального уравнения $d\mathbf{p}/ds = \Delta\mathbf{p}(\mathbf{p}, s)$. Чем меньше μ , тем ближе располагается ломаная к интегральной кривой и при $\mu \rightarrow 0$ сходится к последней. В этом смысле рассматриваемая модель с дискретным временем аппроксимируется дифференциальным уравнением, которое можно анализировать в прежнем масштабе с единицей времени t , равной длительности поколения [5]. Для многомерного состояния \mathbf{p} результат остается верным, но не допускает графической иллюстрации.

Теперь обратимся к модели динамики состояния популяции под влиянием некоторой последовательности факторов микроэволюции. Согласно замечанию 4 представим приближенно приращение $\Delta\mathbf{p}$ как сумму первых членов его ряда Тейлора — сумму вкладов каждого отдельного фактора, малых по предположению (равных $O(\mu)$). Дополнительное изменение малого значения $\Delta\mathbf{p}$ из-за отдельного фактора за счет приращения другого фактора будет порядка $O(\mu^2)$, и при переходе к дифференциальным уравнениям при μ , стремящемся к нулю, получаем аппроксимацию требуемого вида. ◀

Для получения условий корректной аппроксимации разностных уравнений модели с неперекрывающимися поколениями дифференциальными нужно оценить влияние различных факторов на темп $\Delta\mathbf{p}$ изменения генетической структуры \mathbf{p} за поколение. Очевидно, что переход к дифференциальным уравнениям возможен лишь при слабом отборе, т.е. когда изменения концентраций генотипов за поколение малы (полагаем, порядка $O(\mu)$). На стадии отбора по жизнеспособности в гаплоидной популяции согласно (6)

$$dp_i/dt = \mu(v_i(\mathbf{p}, t) - v(\mathbf{p}, t))p_i; \\ v(\mathbf{p}, t) \equiv \sum_i v_i(\mathbf{p}, t)p_i, \quad v_i(\mathbf{p}, t) = O(1). \quad (7)$$

При отборе по жизнеспособности в многолокусном случае правая часть предельной модели, описываемой обыкновенными дифференциальными уравнениями, равна сумме компонента, соответствующего (формально однолокусной) модели чистого отбора, и компонента, соответствующего (формально нейтральной) многолокусной модели процессов рекомбинации в соответствии с замечанием 5. В самом простом случае многолокусная генетическая система состоит из двух диаллельных локусов, скажем, с номерами k и m и с вероятностью рекомбинации локусов $r > 0$. При (корректной при малом коэффициенте r) аппроксимации разностных уравнений (1) модели давления рекомбинаций дифференциальными уравнениями получаем

$$dp_{i_{k|m}}/dt = -r(p_{i_{k|m}} - x_{i_k}y_{i_m}) = -rD_{i_{k|m}}(\mathbf{p}),$$

а в случае слабого давления отбора получаем (7). Если в модели принимаются в расчет и влияние отбора по жизнеспособности (напомним, что соответствующий компонент модели выглядит одинаково для однолокусного и многолокусного случаев отбора), и влияние процессов рекомбинации—сегрегации, то в уравнения динамики входят оба компонента, и в двулокусном случае

$$dp_{i_{k|m}}/dt = \mu(v_{i_{k|m}}(\mathbf{p}) - v(\mathbf{p}))p_{i_{k|m}} - rD_{i_{k|m}}(\mathbf{p}).$$

Ремарка 6. В модели популяции с непрерывным временем и перекрывающимися поколениями

ми без возрастной структуры уравнения динамики при интенсивной плодовитости будут такими же без обременительного требования малости коэффициента рекомбинации [5]. В этом отношении существуют ясно видимые параллели со сходностью к соотношениям Харди–Вайнберга в случае одного локуса в модели с непрерывным временем популяции с перекрывающимися поколениями без возрастной структуры [14].

МОДЕЛЬ ОТБОРА ПО ЗАВИСЯЩЕЙ ОТ ДВУХ ДИАЛЛЕЛЬНЫХ ЛОКУСОВ ЖИЗНЕСПОСОБНОСТИ В ГАПЛОИДНОЙ ПОПУЛЯЦИИ

В моделях гаплоидной (и диплоидной) популяции, рассматриваемой в отношении двух диаллельных локусов (скажем, первого и второго) с коэффициентом рекомбинации r , в качестве независимых переменных можно взять частоты x и y первых аллелей локусов. Тогда для обозначений частот всех аллелей имеем для первого локуса $x_1 \equiv x$ и $x_2 \equiv 1 - x$, а для второго $y_1 \equiv y$ и $y_2 \equiv 1 - y$, концентрации двулокусных гамет по-прежнему обозначаем $p_i = p_{i_1 i_2}$. Как известно, в данном случае коэффициенты неравновесностей для различных типов гамет могут отличаться лишь знаком (см., например, [5]):

$$D \equiv D(\mathbf{p}) \equiv p_{11} - xy, \quad D_{11} \equiv D, \\ D_{12} = D_{21} = -D, \quad D_{22} = D.$$

В терминах частот всех аллелей и коэффициента неравновесности D имеем

$$D_{i_1 i_2} = (-1)^{i_1+i_2} D; \quad p_{i_1 i_2} = x_{i_1} y_{i_2} + (-1)^{i_1+i_2} D, \\ x_{i_1} \equiv \sum_{i_2} p_{i_1 i_2}, \quad y_{i_2} \equiv \sum_{i_1} p_{i_1 i_2}.$$

В случае диаллельных локусов дифференциальные уравнения динамики частот гамет и коэффициента неравновесности D из-за влияния рекомбинаций записываются как

$$dp_{i_1 i_2} / dt = -(-1)^{i_1+i_2} rD, \quad dD / dt = -rD, \\ dD_{i_1 i_2} / dt = -rD_{i_1 i_2}, \quad D_{i_1 i_2} = (-1)^{i_1+i_2} D. \quad (8)$$

Далее предполагаем, что жизнеспособности не зависят от t и $\{v_i = v_i(\mathbf{p})\}$. При учете влияний факторов рекомбинаций и отбора получаем

$$dp_{i_1 i_2} / dt = \mu(v_{i_1 i_2}(\mathbf{p}) - v(\mathbf{p}))p_{i_1 i_2} - \\ - (-1)^{i_1+i_2} rD(\mathbf{p}); \quad v(\mathbf{p}) \equiv \sum_{i_1, i_2} v_{i_1 i_2}(\mathbf{p})p_{i_1 i_2}. \quad (9)$$

Перейдем от частот гамет p_i к новым переменным, задающим генетическое состояние популяции в терминах коэффициента неравновесности

по сцеплению D и частот аллелей. Прямая и обратная замены имеют вид

$$D \equiv p_{11} - xy, \quad x \equiv p_{11} + p_{12} \quad (x_{i_1} = p_{i_1 1} + p_{i_1 2}), \\ y \equiv p_{11} + p_{21} \quad (y_{i_2} = p_{1 i_2} + p_{2 i_2}); \quad (10) \\ p_{11} = xy + D, \quad p_{i_1 i_2} = x_{i_1} y_{i_2} + (-1)^{i_1+i_2} D.$$

Таким образом, рассматриваемые модели можно изучать в терминах независимых частот аллелей x и y (без индексов) и в терминах всех частот аллелей (которые пишем с нижними индексами) и по ходу переключаться между ними, где это доставляет удобства. Изучение можно вести и в терминах частот двулокусных гаплоидных генотипов $p_{i_1 i_2}$, где также можно исключить одну из них как зависимую переменную, поскольку сумма частот равна единице. Для перехода между переменными применяются формулы (10). Какой именно из вариантов используется в дальнейших доказательствах видно из их деталей.

При отборе по жизнеспособности в уравнениях динамики частот двулокусных генотипов в гаплоидной популяции и в уравнениях относительно однолокусных гамет в диплоидной фигурируют формально одинаковые жизнеспособности $v_i = v_{i_1 i_2}$, см., например, [5, 6]. Однако существенным различием служит то, что в случае аутосомного локуса при отсутствии материнского эффекта однолокусные жизнеспособности диплоидов симметричны (не изменяются при перестановке индексов, поскольку генотип остается прежним), тогда как перестановка индексов у жизнеспособности в гаплоидном случае приводит к изменению генотипа, откуда *жизнеспособности двулокусных гаплоидов несимметричны*. Например, при равновесии по сцеплению уравнения динамики частот двулокусных генотипов в гаплоидной популяции и однолокусных гамет в диплоидной формально одинаковы, но матрица жизнеспособности для гаплоидов несимметрична, хотя при этом остается верной фундаментальная теорема отбора Фишера, если жизнеспособности постоянны и достигнуто равновесие по сцеплению.

С помощью (9)–(10) можно выписать уравнения динамики для переменных D, x, y в явном виде, но они оказываются чрезмерно громоздкими для эффективного анализа решения. Поэтому далее обратимся к более простым асимптотическим методам получения приближенных результатов согласно теории возмущений. В соответствии с представлениями Чарлза Дарвина об эволюции как медленном постепенном процессе под влиянием отбора будем рассматривать достаточно типичную ситуацию слабого отбора.

СЛАБЫЙ ОТБОР КАК ВОЗМУЩЕНИЕ НЕЙТРАЛЬНОЙ МОДЕЛИ И КВАЗИРАВНОВЕСИЕ ПО СЦЕПЛЕНИЮ

Займемся оценкой скоростей динамики в предположении слабого по сравнению с рекомбинациями отбора, когда скорости из-за этого фактора имеют порядок $O(\mu)$, $0 < \mu \ll r$, $r = O(1)$. Тогда эффекты отбора можно рассматривать как малое возмущение модели без отбора. Например, из определения концентрации аллеля x (10) следует, что $dx/dt = dp_{11}/dt + dp_{12}/dt$. В этой сумме члены из-за рекомбинаций сокращаются согласно (9), а оставшиеся слагаемые из-за отбора имеют порядок $O(\mu)$, откуда $dx/dt = O(\mu)$. Аналогично получаем, что $dy/dt = O(\mu)$. Что касается порядка dD/dt , то сюда входят эффекты и отбора и рекомбинаций, которые запишем как $dD/dt = -rD + \mu F$. Если представить члены $O(\mu)$ как $\mu O(1)$, то уравнения динамики (9) в новых переменных принимают вид

$$\begin{aligned} \frac{d}{dt} \begin{bmatrix} D \\ x \\ y \end{bmatrix} &= \begin{bmatrix} -rD + \mu F \\ \mu O(1) \\ \mu O(1) \end{bmatrix} = -r \begin{bmatrix} D \\ 0 \\ 0 \end{bmatrix} + \\ &+ \mu \begin{bmatrix} F \\ O(1) \\ O(1) \end{bmatrix}, \quad 0 < \mu \ll r = O(1). \end{aligned} \quad (11)$$

Здесь μF может соответствовать слабому давлению отбора и/или других систематических факторов микроэволюции. Отсюда видно, что данную систему можно рассматривать как возмущенную, μ — как малый параметр (r считаем фиксированным и независимым от генотипа). Пусть, например, при отсутствии возмущений $dD/dt = g(D, x, y)$, а с их учетом $dD/dt = g + \mu F$. Уравнению $g = 0$ удовлетворяет невозмущенное равновесие по D при равновесности прочих переменных. Аналогично в возмущенном равновесии $g + \mu F = 0$. Если g зависит от D линейно (и g можно записать как $g_0(x, y, \mu) + cD$), а F не зависит от D , то в возмущенном равновесии $D = -(g_0 + \mu F)/c$. Конечно, в общем случае зависимости нелинейны, но если искать равновесие приближенно дельта-методом, в котором используются только первые члены разложения Тейлора рассматриваемых функций, то линейность выполняется. В результате в общих чертах задача приближенного отыскания равновесия D (квазиравновесия D^*) редуцируется к вычислению g_0 (у нас g_0 описывает давление рекомбинаций), μF (давление систематических факторов) и c . Детально этот подход реализуется в доказательствах соответствующих утверждений далее.

В (11) невозмущенная часть получается при $\mu = 0$; очевидно, она определяет нейтральную модель динамики в результате рекомбинаций, а воз-

мущение соответствует эффектам систематических факторов. В этой системе μ , x , y на начальном этапе динамики можно рассматривать как фиксированные параметры, определяющие соответствующее равновесие D невозмущенного уравнения $dD/dt = -rD$, которым, очевидно, будет $D = 0$. В результате малого возмущения μ , x , y получат приращения $\Delta\mu$, Δx , Δy , благодаря чему у D будет новое равновесие, мало отличающееся от нулевого исходной нейтральной модели. Называть его равновесием можно лишь условно, поскольку “параметры” x , y сами медленно меняются согласно точным уравнениям возмущенной динамики, поэтому за данным показателем закрепилось название *квазиравновесие*. Таким образом, нулевое значение D приближенно оценивает квазиравновесную величину D (нулевое приближение). Далее мы приступим к отысканию более точной оценки D , получаемой с учетом вида возмущений (первое приближение).

Результат 7. Пусть рассматривается модель динамики генетической структуры гаплоидной популяции (11) в отношении двух диаллельных локусов.

Тогда приближенное значение квазиравновесия для D представимо в виде

$$D^* = \frac{\mu}{r} F(0, x, y), \quad (12)$$

где функция $\mu F(D, x, y)$ определяет вклад в скорость динамики D слабого давления систематических факторов микроэволюции.

Доказательство. Положим, что уточненная оценка квазиравновесного значения D отличается от нулевого приближения $D = 0$ приращением ΔD , т.е. задача состоит в отыскании ΔD .

А. Координатами равновесия при $\mu = 0$ будут $D = 0$ и текущие значения x , y , а уточненное равновесие соответствует приращению $\Delta\mu$ малого параметра. Переменные x и y также получают какие-то малые приращения Δx и Δy . Напомним, что правая часть dD/dt как функция от D , x , y , μ имеет вид $-rD + \mu F(D, x, y)$. В любом равновесии эта часть равна нулю, т.е. при отсутствии возмущений $(-rD + \mu F(D, x, y))|_{D=0, \mu=0} = 0$, и аналогично в уточненном квазиравновесии

$$\begin{aligned} &-rD|_{D=0+\Delta D} + \mu|_{\mu=0+\Delta\mu} \times \\ &\times F(D, x + \Delta x, y + \Delta y)|_{D=0+\Delta D} = 0. \end{aligned}$$

Таким образом, при указанных приращениях аргументов выражение $-rD + \mu F$ получило нуле-

вое изменение, которое обозначим Δ . Значит приближенно

$$\begin{aligned} \Delta &= \frac{\partial}{\partial D}(-rD + \mu F)\Delta D + \frac{\partial}{\partial x}(-rD + \mu F)\Delta x + \\ &+ \frac{\partial}{\partial y}(-rD + \mu F)\Delta y + \frac{\partial}{\partial \mu}(-rD + \mu F)\Delta \mu = \\ &= -r\Delta D + \mu|_{\mu=0} \left(\frac{\partial}{\partial D} F\Delta D + \frac{\partial}{\partial x} F\Delta x + \frac{\partial}{\partial y} F\Delta y \right) + \\ &+ F(0, x, y)\Delta \mu = -r\Delta D + F(0, x, y)\Delta \mu = 0, \end{aligned}$$

где производные вычисляются при $\Delta D = \Delta x = \Delta y = \Delta \mu = 0$. Впрочем, видно, что значения Δx , Δy не влияют на отыскание ΔD , так как при них стоит равный нулю множитель μ .

Б. Удачность последнего обстоятельства в том, что благодаря ему не потребуется решать систему уравнений относительно приращений всех переменных, и искомое приращение ΔD (приближенное значение D^*) удовлетворяет уравнению

$$-rD^* + \mu F(0, x, y) = 0. \tag{13}$$

Здесь мы учли, что $\Delta D = D^*$ (D^* обозначает уточненное значение квазиравновесия D), поскольку приращение ΔD до D^* отсчитывается от нуля, и аналогично $\Delta \mu = \mu$. Отсюда D^* имеет вид $D^* = \frac{\mu}{r} F(0, x, y)$. ◀

Значит при отыскании уточненного приближенного значения квазиравновесия D^* нужно только вычислить $\mu F(D, x, y)$ при $D = 0$, т.е. значение F при достижении равновесия по сцеплению из-за доминирующего влияния процессов рекомбинации.

Результат 8. При слабом отборе в гаплоидной популяции с жизнеспособностями генотипов $\{v_{i_2}(\mathbf{p})\}$, зависящими от двух диаллельных локусов, приближенным значением квазиравновесия (12) будет

$$\begin{aligned} D^* &= \varepsilon(\mathbf{p}) \frac{\mu}{r} x(1-x)y(1-y), \\ \varepsilon(\mathbf{p}) &\equiv v_{11}(\mathbf{p}) - v_{12}(\mathbf{p}) - v_{21}(\mathbf{p}) + v_{22}(\mathbf{p}). \end{aligned} \tag{14}$$

Доказательство. Будем анализировать модель (9) в терминах переменных D, x, y . Поскольку для возмущенной модели (9) $dD/dt = -rD + \mu F(D, x, y)$, то искомое значение $\mu F(0, x, y)$ равно правой части dD/dt при одновременном выполнении условий $dD/dt = 0$ и $D = 0$. В соответствии с определением $D \equiv p_{11} - xy$ производная dD/dt равна $dp_{11}/dt - d(xy)/dt$. Напомним, что согласно (9) $dp_{11}/dt = \mu p_{11}(v_{11}(\mathbf{p}) - v(\mathbf{p}))$ при $D = 0$. Остается вычислить $d(xy)/dt$. Так как сейчас в (9) правая часть вычисляется при условии $D = 0$ (т.е. при многоло-

кусных соотношениях Харди–Вайнберга, когда $p_{11} = xy, p_{12} = x(1-y)$), то из $x \equiv p_{11} + p_{12}$ получаем

$$\begin{aligned} \frac{1}{\mu} dx/dt &= xy(v_{11}(\mathbf{p}) - v^*(\mathbf{p})) + \\ &+ x(1-y)(v_{12}(\mathbf{p}) - v^*(\mathbf{p})) = \\ &= xy(v_{11}(\mathbf{p}) - v_{12}(\mathbf{p})) + x(v_{12}(\mathbf{p}) - v^*(\mathbf{p})), \\ \frac{1}{\mu} dy/dt &= xy(v_{11}(\mathbf{p}) - v_{21}(\mathbf{p})) + \\ &+ y(v_{21}(\mathbf{p}) - v^*(\mathbf{p})). \end{aligned}$$

Используем полученные выражения для вычисления $d(xy)/dt$:

$$\begin{aligned} \frac{1}{\mu} d(xy)/dt &= xy(y(v_{11}(\mathbf{p}) - v_{12}(\mathbf{p})) + v_{12}(\mathbf{p}) - \\ &- v^*(\mathbf{p})) + xy(x(v_{11}(\mathbf{p}) - v_{21}(\mathbf{p})) + v_{21}(\mathbf{p}) - v^*(\mathbf{p})) = \\ &= xy(x(v_{11}(\mathbf{p}) - v_{21}(\mathbf{p})) + y(v_{11}(\mathbf{p}) - v_{12}(\mathbf{p})) + \\ &+ v_{12}(\mathbf{p}) + v_{21}(\mathbf{p}) - 2v^*(\mathbf{p})). \end{aligned}$$

Подставим в dD/dt найденное выше выражение с учетом того, что $dp_{11}/dt = \mu xy(v_{11}(\mathbf{p}) - v^*(\mathbf{p}))$ при равном нулю D , и вынесем xy за скобки. В результате получаем

$$\begin{aligned} F(0, x, y) &= \frac{1}{\mu} dD/dt|_{D=0} = dp_{11}/dt - d(xy)/dt = \\ &= xy(v_{11}(\mathbf{p}) - v^*(\mathbf{p})) - (xy(x(v_{11}(\mathbf{p}) - v_{21}(\mathbf{p})) + \\ &+ y(v_{11}(\mathbf{p}) - v_{12}(\mathbf{p})) + v_{12}(\mathbf{p}) + v_{21}(\mathbf{p}) - 2v^*(\mathbf{p}))) = \\ &= xy(v_{11}(\mathbf{p}) - x(v_{11}(\mathbf{p}) - v_{21}(\mathbf{p})) - \\ &- y(v_{11}(\mathbf{p}) - v_{12}(\mathbf{p})) - v_{12}(\mathbf{p}) - v_{21}(\mathbf{p}) + v^*(\mathbf{p})). \end{aligned}$$

Здесь средней жизнеспособностью $v^*(\mathbf{p})$ при условии $D = 0$ будет

$$\begin{aligned} \frac{1}{\mu} v^*(\mathbf{p}) &= \varepsilon(\mathbf{p}) xy + x(v_{12}(\mathbf{p}) - v_{22}(\mathbf{p})) + \\ &+ y(v_{21}(\mathbf{p}) - v_{22}(\mathbf{p})) + v_{22}(\mathbf{p}); \\ \varepsilon(\mathbf{p}) &\equiv v_{11}(\mathbf{p}) - v_{12}(\mathbf{p}) - v_{21}(\mathbf{p}) + v_{22}(\mathbf{p}). \end{aligned}$$

Подставим $v^*(\mathbf{p})$ в правую часть полученного выражения для $F(0, x, y)$ и приведем подобные члены. В итоге получим, что все слагаемые содержат множитель ε (называемый *коэффициентом неаддитивности эпистаза*, поскольку он обращается в нуль при аддитивной жизнеспособности):

$$\begin{aligned} \mu F(D, x, y)|_{D=0} &= dD/dt|_{D=0} = \\ &= \mu xy(-\varepsilon x - \varepsilon y + \varepsilon x y + \varepsilon) = \varepsilon \mu xy(1-y)(1-x). \end{aligned}$$

В итоге уточненным квазиравновесным значением D согласно (12) будет

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

Следовательно, на эволюционном этапе динамики при текущих значениях частот аллелей x и y

приближенной величиной коэффициента неравновесности по сцеплению будет $\epsilon(\mathbf{p}) \frac{\mu}{r} x(1-x)y(1-y)$. Эта величина мультипликативно определяется дисперсиями встречаемости аллелей первого и второго локусов у особей в популяции $x(1-x)$ и $y(1-y)$ соответственно, отношением давления отбора μ к коэффициенту рекомбинации r и показателем неаддитивности эпистаза ϵ . Эпистаз характеризует взаимодействие неаллельных генов при детерминации признака. Показателем неаддитивности эпистаза служит отклонение значения признака от соответствующего сумме отклонений, вызванных отдельными аллелями (см. операционное определение $\epsilon_{11}(\mathbf{p})$ в (15), редуцирующееся к данному в (14) для двулокусного диаллельного случая). При постоянных жизнеспособностях коэффициент эпистаза ϵ будет независимой от генетического состояния популяции константой, и неравновесность D^* достигает своего равного $\epsilon \frac{\mu}{r} \frac{1}{4}$ максимума на траектории, проходящей через точку $(x, y) = (1/2, 1/2)$. Квазиравновесность изменяется с эволюцией частот аллелей и может как увеличиваться, так и уменьшаться в зависимости от текущего состояния популяции. Приведенное приближенное вычисление D^* в сущности основано на дельта-методе, а используемое для отыскания D^* уравнение совпадает с уравнением первого приближения теории сингулярных возмущений. Далее мы допустим некоторую вольность в формулах, используя для краткости записи одновременно переменные \mathbf{p} и x, y , относящиеся к разным системам координат.

Результат 9. Квазиравновесие по сцеплению D^* в (14) можно представить через генные эффекты в виде, допускающем обобщение на случай множественных аллелей:

$$D^* = (\delta_{11}(\mathbf{p}) - \delta_{i_1}^*(\mathbf{p}) - \delta_{i_1}^*(\mathbf{p})) \frac{\mu}{r} x_1 y_1 \equiv \epsilon_{11}(\mathbf{p}) \frac{\mu}{r} x_1 y_1. \quad (15)$$

Здесь $\delta_{11} = v_{11} - v^*$, $\delta_{i_1}^* = v_{i_1}^* - v^*$, $\delta_{i_1}^* = v_{i_1}^* - v^*$, жирная точка означает усреднение (суммирование по индексу на этом месте с весами, равными концентрациям аллелей соответствующего локуса), $\delta_{i_1}^*$ имеет смысл среднего эффекта первого аллеля первого локуса, аналогично $\delta_{i_1}^*$ — средний эффект первого аллеля второго локуса, а δ_{11} — отклонение жизнеспособности v_{11} от популяционного среднего v^* (звездочка указывает, что вычисление соответствующего показателя производится при равновесии по сцеплению). Разность $\delta_{11}(\mathbf{p}) - \delta_{i_1}^*(\mathbf{p}) - \delta_{i_1}^*(\mathbf{p})$

интерпретируется как средний эффект взаимодействия первых аллелей разных локусов (как коэффициент эпистаза $\epsilon_{11}(\mathbf{p})$).

Кроме того, величина D^* имеет при слабом отборе порядок $O(\mu^2)$.

Доказательство. Напомним, что по (12) $D^* = \frac{\mu}{r} F(0, x, y)$, значит для вычисления D^* надо найти $F(0, x, y)$. Для возмущенной модели (9) при $D = 0$ согласно (11) $dD/dt|_{D=0} = (-rD + \mu F(D, x, y))|_{D=0}$, т.е. $\mu F(0, x, y)$ находится как $dD/dt|_{D=0}$. Как и в предыдущем случае, займемся вычислением dD/dt . Из определения D получаем $dD/dt = dp_{11}/dt - d(xy)/dt$. При $D = 0$

$$\begin{aligned} p_{11} &= xy = x_1 y_1, & dp_{11}/dt &= \mu x_1 y_1 (v_{11} - v^*), \\ dp_{12}/dt &= \mu x_1 y_2 (v_{12} - v^*), \\ dx_1/dt &= d(p_{11} + p_{12})/dt = & (16) \\ &= \mu x_1 ((v_{11} y_1 + v_{12} y_2) - \\ &- (y_1 + y_2) v^*) \equiv \mu x_1 (v_{i_1}^* - v^*). \end{aligned}$$

При выводе учтено, что $y_1 + y_2 = 1$, $v_{i_1}^*(\mathbf{p})$ обозначает среднюю жизнеспособность генотипа при условии, что он содержит первый аллель первого локуса (коротко, жизнеспособность этого аллеля). Здесь точкой обозначен индекс, по которому произведено суммирование с весами, равными частотам аллелей данного локуса. В этих обозначениях среднюю жизнеспособность первого аллеля второго локуса можно записать как $v_{i_1}^*$. Аналогично получаем, что $dy_1/dt = y_1 (v_{i_1}^* - v^*)$. В итоге при $D = 0$

$$\begin{aligned} \mu F(0, x, y) &= dD/dt|_{D=0} = \mu x_1 y_1 (v_{11} - v^*) - \\ &- \mu x_1 (v_{i_1}^* - v^*) y_1 - \mu x_1 y_1 (v_{i_1}^* - v^*) = \\ &= \mu x_1 y_1 (\delta_{11} - \delta_{i_1}^* - \delta_{i_1}^*) \equiv \mu \epsilon_{11}(\mathbf{p}) x_1 y_1. \end{aligned}$$

Значит согласно (12)

$$\begin{aligned} D^* &= (\delta_{11}(\mathbf{p}) - \delta_{i_1}^*(\mathbf{p}) - \delta_{i_1}^*(\mathbf{p})) \frac{\mu}{r} x_1 y_1 \equiv \\ &\equiv \epsilon_{11}(\mathbf{p}) \frac{\mu}{r} x_1 y_1. \end{aligned}$$

Отметим, что при слабом отборе D^* имеет порядок $O(\mu^2)$, так как

$$\begin{aligned} |D^*| &= |\epsilon_{11}(\mathbf{p})| O(\mu) = |\delta_{11}(\mathbf{p}) - \delta_{i_1}^*(\mathbf{p}) - \\ &- \delta_{i_1}^*(\mathbf{p})| O(\mu) = |v_{11} - v_{i_1}^* - v_{i_1}^* + v^*| O(\mu) \leq \\ &\leq |v_{11} - v_{i_1}^*| \mu + |v_{i_1}^* - v^*| O(\mu) = O(\mu^2), \end{aligned}$$

поскольку различия между жизнеспособностями при слабом отборе порядка μ . ◀

Замечание 10. Коэффициент эпистаза $\varepsilon_{11}(\mathbf{p})$ генотипа (11) в (15) является средним по элементарным генофондам, соответствующим аллелям данного генотипа и генотипа особи, случайно выбранной из популяции, в которой достигнуто равновесие по сцеплению.

$$\begin{aligned} \varepsilon_{11}(\mathbf{p}) &\equiv \delta_{11}(\mathbf{p}) - \delta_{i_1}^*(\mathbf{p}) - \delta_{i_2}^*(\mathbf{p}) = \\ &= \sum_{k,m=1}^2 (v_{11}(\mathbf{p}) - v_{1m}(\mathbf{p}) - v_{k1}(\mathbf{p}) + v_{km}(\mathbf{p}))x_k y_m \equiv \\ &\equiv \sum_{k,m=1}^2 \varepsilon_{11km} x_k y_m = \varepsilon_{1122} x_2 y_2 = \varepsilon x_2 y_2; \\ \varepsilon_{11km} &\equiv v_{11}(\mathbf{p}) - v_{1m}(\mathbf{p}) - v_{k1}(\mathbf{p}) + v_{km}(\mathbf{p}), \end{aligned} \quad (17)$$

где $\varepsilon_{11}(\mathbf{p})$ интерпретируется как средний неаддитивный эффект генотипа (11).

Доказательство. Очевидно, надо доказать, что $\varepsilon_{11}(\mathbf{p}) = \varepsilon x_2 y_2$. Рассмотрим по отдельности коэффициенты ε_{11km} , входящие в сумму в (17) с весами $x_k y_m$. При расписывании ε_{11km} через компоненты v_{km} обнаруживается, что для индексов $km = (11), (12), (21)$ коэффициенты $\varepsilon_{11km} = v_{11}(\mathbf{p}) - v_{1m}(\mathbf{p}) - v_{k1}(\mathbf{p}) + v_{km}(\mathbf{p})$ обращаются в нуль. В итоге в (17), т.е. в ε_{11} , останется только одно слагаемое, и

$$\begin{aligned} \varepsilon_{11}(\mathbf{p}) &= \varepsilon_{1122} x_2 y_2 = \\ &= (v_{11}(\mathbf{p}) - v_{12}(\mathbf{p}) - v_{21}(\mathbf{p}) + v_{22}(\mathbf{p})) x_2 y_2 \equiv \varepsilon x_2 y_2. \end{aligned}$$

Доказательство закончено. ◀

В результате

$$\begin{aligned} \varepsilon_{11} &= v_{11}(\mathbf{p}) - v_{i_1}^*(\mathbf{p}) - v_{i_2}^*(\mathbf{p}) + v^*(\mathbf{p}) = \\ &= \delta_{11}(\mathbf{p}) - \delta_{i_1}^*(\mathbf{p}) - \delta_{i_2}^*(\mathbf{p}) = \varepsilon_{1122} x_2 y_2 = \varepsilon x_2 y_2 \end{aligned}$$

в диаллельном случае. Кроме того, из (12)–(17) получаем различные варианты выражений для D^* :

$$\begin{aligned} D^* &= \varepsilon(\mathbf{p}) \frac{\mu}{r} x(1-x)y(1-y) \equiv \\ &\equiv (v_{11}(\mathbf{p}) - v_{12}(\mathbf{p}) - v_{21}(\mathbf{p}) + v_{22}(\mathbf{p})) \frac{\mu}{r} \times \\ &\times xy(1-x)(1-y) = (\delta_{11}(\mathbf{p}) - \delta_{i_1}^*(\mathbf{p}) - \\ &- \delta_{i_2}^*(\mathbf{p})) \frac{\mu}{r} xy \equiv \varepsilon_{11}(\mathbf{p}) \frac{\mu}{r} xy = \varepsilon_{1122} x_2 y_2 \frac{\mu}{r} x_1 y_1. \end{aligned}$$

Ремарка 11. В доказательстве не предполагалось, что жизнеспособности являются константами, они могут зависеть от генетического состояния \mathbf{p} , но у них нет явной зависимости от t .

Ремарка 12. Представление (17) коэффициента эпистаза обобщается на случай множественных аллелей [20].

Ремарка 13. Коэффициенты ε_{ijkm} можно назвать показателями (коэффициентами) неаддитивности эпистаза для элементарного генофонда, содержащего аллели i, k первого локуса и j, m второго. Легко видеть, что эти коэффициенты будут нулевыми при аддитивной по локусам жизнеспособности. Однако аддитивность не является необходимым условием для существования нулевых коэффициентов ε_{ijkm} — некоторые из них могут быть нулевыми и при неаддитивной жизнеспособности, как мы уже видели в приведенном выше доказательстве. Жизнеспособности генотипов при мультипликативной детерминации имеют вид $v_{i_1 i_2} \equiv u_{i_1} w_{i_2}$. В этом случае $\varepsilon = (u_1 - u_2)(w_1 - w_2)$, где u_{i_1} и w_{i_2} — мультипликативные компоненты жизнеспособности для первого и второго локусов, i_1 и i_2 принимают значения 1 или 2.

Таким образом, эпистатические эффекты в формулах для квазинеравновесности D^* допускают различную интерпретацию. Это коэффициент $\varepsilon = v_{11} - v_{12} - v_{21} + v_{22}$, который при постоянных жизнеспособностях не зависит от генетической структуры, а является константой, описывающей генетическую детерминацию жизнеспособности на организменном уровне; это стандартно определяемая в терминах эффектов аллелей и их взаимодействий популяционная (в общем случае) характеристика неаддитивности эпистаза ε_{11} и это ε_{11} как усредненный коэффициент отклонения от аддитивности в детерминации жизнеспособности, явно выражаемый через коэффициенты эпистаза ε_{11km} элементарных генофондов. Некоторые из перечисленных свойств видны более ярко в случае множественных аллелей [20].

Далее обратимся к динамической интерпретации полученного квазиравновесия по сцеплению D^* в рамках анализа возмущения отбором как сингулярного.

СЛАБЫЙ ОТБОР КАК СИНГУЛЯРНОЕ ВОЗМУЩЕНИЕ

Обратим внимание, что в возмущенной модели (11) при условии $0 < \mu/r \ll r$ скорость изменения неравновесности D много больше скорости медленно эволюционирующих частот аллелей. Такие системы обыкновенных дифференциальных уравнений с быстрыми и медленными переменными называются *сингулярно возмущенными*. Для них можно без решения уравнений охарактеризовать качественную картину динамики следующим образом. На первом этапе с большой скоростью идут изменения быстрых переменных в силу системы, получающейся из исходной (возмущенной) при $\mu = 0$, тогда как медленные пере-

менные практически постоянны. Их здесь можно рассматривать как фиксированные параметры. Часто такое поведение быстрых переменных много проще, чем для возмущенной модели, и понятно на качественном уровне.

Пусть на этом этапе *быстрые переменные сходятся к некоторому устойчивому по первому приближению равновесию*, которое в итоге будет функцией медленных переменных. У нас быстро изменяется неравновесность по сцеплению D , которая согласно (8) и вышеизложенному сойдется к нулевому значению (в данном случае независимо от текущих значений не фиксированных на нуле или единице концентраций аллелей).

На втором этапе происходит эволюция медленных переменных при условии квазиравновесия быстрых, т.е. в силу системы, получающейся при подстановке в уравнения для медленных переменных вместо быстрых их равновесие (точнее, квазиравновесие), полученное на первом этапе как функция текущих значений медленных. В нашем случае получается, что эволюция частот аллелей будет идти под действием отбора при условии многолокусных соотношений Харди–Вайнберга (равновесии по сцеплению).

Таким образом происходит *редукция первоначальной размерности задачи* на обоих этапах: на первом рассматривается динамика только быстрых переменных, а на втором, эволюционном, только медленных (вместо частот двулокусных гаплоидных генотипов анализируется меньшее количество частот аллелей). Кроме того, частоты аллелей представляют собой агрегированные переменные — они являются частотами меньшего количества объединенных в блоки двулокусных генотипов по принципу включения в свой состав того или иного аллеля. Значит теория сингулярных возмущений обосновывает *агрегацию переменных*, которая может очень существенно уменьшить размерность задачи, что подчеркивается во многих работах (например, в [17]).

Приведенная выше картина динамики, конечно, лишь приближенная, и нулевое значение коэффициента неравновесности по сцеплению несколько отличается от точного. Очевидно, что давление рекомбинаций направлено на устранение неравновесности, тогда как давление отбора может этому препятствовать. Баланс указанных факторов способен привести к отклонению от нулевого значения неравновесности при динамике медленных частот аллелей, и появляется задача оценивания величины возникающего отклонения. В данном разделе эта задача анализируется с применением стандартных методов теории сингулярных возмущений, согласно которой на пер-

вом этапе быстрой динамики нужно анализировать невозмущенную систему, получающуюся при $\mu = 0$.

Замечание 14. При $\mu = 0$ возмущенная система (9), (11) редуцируется к уравнениям (8), описывающим динамику генетического состояния популяции из-за давления процессов рекомбинации.

При фиксированных частотах аллелей единственным асимптотически устойчивым (и устойчивым по первому приближению) равновесием (8) будет $D = 0$.

Доказательство. Невозмущенная система уравнений имеет вид

$$\frac{d}{dt} \begin{bmatrix} D \\ x \\ y \end{bmatrix} = -r \begin{bmatrix} D \\ 0 \\ 0 \end{bmatrix}, \text{ или } dD/dt = -rD.$$

Ее очевидное решение показывает, что D стремится к нулю независимо от начального значения, и нуль будет асимптотически устойчивым (и устойчивым по первому приближению) равновесием для быстрой переменной D . Хотя в данной выше записи формально фигурируют значения медленных частот аллелей, они фиксированы на первом этапе, т.е. рассматриваются как параметры. ◀

Результат 15. Асимптотически при $\mu/r \rightarrow 0$ на траекториях возмущенной системы (11) в промежутке $0 < t_1 < t < \infty$ выполняется:

А. Концентрации гамет равны произведениям частот входящих в них аллелей ($D = 0$).

Б. Уравнениями для частот аллелей (медленных переменных) будут

$$\begin{aligned} dx/dt &= \mu x \left(v_{i_1}^* - v^* \right), \quad dy/dt = \mu y \left(v_{i_1}^* - v^* \right); \\ v_{i_1}^* &\equiv \sum_{i_2} v_{i_1 i_2}(\mathbf{p}) y_{i_2}, \quad v_{i_2}^* \equiv \sum_{i_1} v_{i_1 i_2}(\mathbf{p}) x_{i_1}, \\ v^* &= v^*(x_1, x_2, y_1, y_2) = \sum_{i_1, i_2} v_{i_1 i_2} x_{i_1} y_{i_2} = \\ &= \sum_{i_1} v_{i_1}^*(\mathbf{p}) x_{i_1} = \sum_{i_2} v_{i_2}^*(\mathbf{p}) y_{i_2}, \end{aligned} \quad (18)$$

где $v_{i_1}^*$ и $v_{i_2}^*$ — маргинальные жизнеспособности аллелей первого и второго локусов соответственно при $D = 0$ (напомним, что точкой обозначен индекс, по которому произведено усреднение), а v^* — средняя жизнеспособность при равновесии по сцеплению; $x_1 = x, x_2 = 1 - x, y_1 = y, y_2 = 1 - y$.

Доказательство. А. Согласно (11) уравнения возмущенной динамики можно записать как

$$\begin{aligned} dD/dt &= -rD + \mu F, & dx/dt &= \mu f_1, \\ dy/dt &= \mu f_2, & 0 < \mu &\ll 1 \end{aligned}$$

с некоторыми функциями F, f_1, f_2 нужной гладкости. Перейдем к медленному (“эволюционному”) времени $s = \mu t$, единице которого соответствует большой промежуток быстрого времени t . В результате такой замены уравнения записываются в стандартном виде для сингулярно возмущенной системы (с малым параметром при производной)

$$\begin{aligned} \mu dD/ds &= -rD + \mu F(D, x, y), \\ dx/ds &= f_1(D, x, y), & dy/ds &= f_2(D, x, y), \quad (19) \\ s &= \mu t. \end{aligned}$$

Согласно теореме Тихонова (см., например, [22]) нулевое приближение $x^0(s), y^0(s)$ точного решения (19) находится из вырожденной системы

$$\begin{aligned} (-rD + \mu F(D, x^0, y^0))\Big|_{\mu=0} &= 0, \\ dx^0/ds &= f_1(\varphi(x^0, y^0), x^0, y^0), \\ dy^0/ds &= f_2(\varphi(x^0, y^0), x^0, y^0), \end{aligned}$$

где $\varphi(x^0, y^0)$ – устойчивое по первому приближению равновесие присоединенной системы (для быстрых переменных) при используемых начальных данных

$$dD/ds = (-rD + \mu F(D, x^0, y^0))\Big|_{\mu=0} = -rD,$$

в которой x^0, y^0 рассматриваются как параметры. При $\mu \rightarrow 0$ нулевое приближение (x^0, y^0) стремится к точному решению при $0 < s < \infty$.

У нас, очевидно, нетипичный случай, когда $\varphi(x, y)$ – равновесное для присоединенной системы значение D не зависит от медленных переменных (оно равно нулю, откуда концентрации гамет будут равны произведениям частот входящих в них аллелей, и устойчиво по первому приближению). Этим часть А доказана.

Б. Вернемся к традиционной в популяционной генетике записи модели в быстром времени t , в котором x и y рассматриваем как нулевое приближение для точного решения (9), (11). Согласно (16) для dx/dt (и аналогично для dy/dt) при $D = 0$ имеем

$$\begin{aligned} dx/dt &= \mu x \left(v_{1.}^*(\mathbf{p}) - v^*(\mathbf{p}) \right), \\ dy/dt &= \mu y \left(v_{.1}^*(\mathbf{p}) - v^*(\mathbf{p}) \right), \end{aligned}$$

что доказывает часть Б.

Остается рассмотреть в каком диапазоне времени можно использовать уравнения медленной эволюции частот аллелей. Обозначим через t_1 интервал быстрого времени, требуемый для достижения D квазинулевого значения с заданной точностью. С этого момента приблизительно медленные переменные эволюционируют при квазиравновесии быстрой. Интервалу $(0, t_1)$ быстрого времени t соответствует интервал $(0, \mu t_1)$ медленного s , длина которого стремится к нулю при $\mu/r \rightarrow 0$. При этом сходимость приближенного решения (вырожденной системы) к точному верна при $0 < s < \infty$. ◀

Результат 16. В гаплоидной популяции при постоянных жизнеспособностях, зависящих от двух диалельных локусов, и при равновесии по сцеплению:

А. Уравнения динамики частот аллелей можно записать как

$$\begin{aligned} dx/dt &= \mu x(1-x) \partial v^*/\partial x, \\ dy/dt &= \mu y(1-y) \partial v^*/\partial y. \end{aligned} \quad (20)$$

Здесь $v^* = v^*(x, y)$ – средняя жизнеспособность популяции при условии $D = 0$:

$$\begin{aligned} v^*(x, y) &= \varepsilon x + (v_{12} - v_{22})x + \\ &+ (v_{21} - v_{22})y + v_{22}, \quad \varepsilon \equiv v_{11} - v_{12} - v_{21} + v_{22}. \end{aligned}$$

Б. Средняя жизнеспособность популяции v^* возрастает вдоль неравновесных траекторий приведенной выше системы уравнений (20) медленной эволюции частот аллелей.

Доказательство. Уравнения (18) напоминают уравнения модели отбора по жизнеспособности по одному локусу в диплоидной популяции, которые (см., например, [5]) формально совпадают с (18) при $x = y$, но теперь жизнеспособности несимметричны по индексам.

А. Будем анализировать модель в терминах частот всех аллелей, когда

$$x_1 = x, \quad x_2 = 1 - x, \quad y_1 = y, \quad y_2 = 1 - y.$$

Напомним, что при многолокусных соотношениях Харди–Вайнберга (при $D = 0$) средняя двулокусная жизнеспособность для гаплоидной популяции в терминах всех переменных записывается как

$$\begin{aligned} v^* &= v^*(x, y) = v^*(x_1, x_2, y_1, y_2) = \\ &= \sum_{i_1, i_2=1}^2 v_{i_1 i_2}^* x_{i_1} y_{i_2} = \sum_{i_1} v_{i_1.}^* x_{i_1} = \sum_{i_2} v_{.i_2}^* y_{i_2}. \end{aligned}$$

Заметим, что, с одной стороны, производные по времени x и y выражаются в соответствии с (16) через маргинальные жизнеспособности в (18). С другой стороны, при постоянных жизнеспособ-

ностях маргинальные жизнеспособности выражаются через производные v^* :

$$v_{1.}^* = \partial v^* / \partial x_1, \quad v_{.1}^* = \partial v^* / \partial y_1;$$

$$v^* = \sum_{i=1}^2 (\partial v^* / \partial x_i) x_i,$$

откуда

$$\begin{aligned} dx/dt &= \mu x_1 (v_{1.}^* - v^*) = \\ &= \mu x_1 \left(\partial v^* / \partial x_1 - \sum_{i=1}^2 (\partial v^* / \partial x_i) x_i \right) = \\ &= \mu x_1 (\partial v^* / \partial x_1 (1 - x_1) - \\ &- (\partial v^* / \partial x_2) x_2) = \mu x_1 (1 - x_1) \partial v^* / \partial x_1 - \\ &- \mu x_1 x_2 \partial v^* / \partial x_2 = \mu x_1 (1 - x_1) (\partial v^* / \partial x_1 - \\ &- \partial v^* / \partial x_2). \end{aligned}$$

Поскольку $x_1 = x$, $x_2 = 1 - x_1$, а $\partial v^* / \partial x = \partial v^* / \partial x_1 - \partial v^* / \partial x_2$, то

$$\begin{aligned} dx/dt &= \mu x_1 (1 - x_1) (\partial v^* / \partial x_1 - \partial v^* / \partial x_2) = \\ &= \mu x (1 - x) \partial v^* / \partial x. \end{aligned}$$

Аналогично находим, что при постоянных жизнеспособностях

$$\begin{aligned} dy/dt &= \mu y_1 (1 - y_1) (\partial v^* / \partial y_1 - \partial v^* / \partial y_2) = \\ &= \mu y (1 - y) \partial v^* / \partial y. \end{aligned}$$

Таким образом, при нулевом значении D эволюция медленных переменных, частот аллелей, будет квазиградиентной, уравнения динамики записываются как было предложено еще С. Райтом [9], в виде $dx/dt = \mu x(1-x) \partial v^* / \partial x$, $dy/dt = \mu y(1-y) \partial v^* / \partial y$. Тем самым часть А доказана.

Б. Эта часть следует из непосредственного вычисления производной $dv^*(x, y)/dt$ при условии $D = 0$ на траектории медленного движения:

$$\begin{aligned} \frac{d}{dt} v^*(x, y) &= \frac{d}{dx} v^*(x, y) \frac{d}{dt} x + \frac{d}{dy} v^*(x, y) \frac{d}{dt} y = \\ &= \mu x (1 - x) \left(\frac{d}{dx} v^*(x, y) \right)^2 + \\ &+ \mu y (1 - y) \left(\frac{d}{dy} v^*(x, y) \right)^2 \geq 0. \end{aligned}$$

Данное неравенство обращается в равенство только в равновесии, где либо x и y обращаются в нуль или единицу (граничное равновесие), либо одновременно $\frac{d}{dx} v^*(x, y)$ и $\frac{d}{dy} v^*(x, y)$ равны нулю (внутреннее полиморфное равновесие, невозможное в случае гаплоидности). ◀

Ремарка 17. Для сингулярно возмущенных систем разработано (см., например, [22]) построе-

ние асимптотики решения в виде суммы регулярного и погранслоного асимптотических рядов по степеням μ , чья частичная сумма до членов с μ^n включительно аппроксимирует точное решение с погрешностью $O(\mu^{n+1})$. Значение $D = 0$ дает нулевое приближение с точностью $O(\mu)$ для быстрой координаты равновесного решения (9). Обоснование результатов использования для отыскания D^* дельта-метода (приведшего к формуле (13)) в том, что он дал уравнение, совпадающее с соответствующей частью системы первого приближения для асимптотического ряда. Отличие найденного квазиравновесия D^* от точного значения на траектории медленного движения частот аллелей имеет порядок $O(\mu^2)$.

ОБСУЖДЕНИЕ

Изучение коэффициента неравновесности по сцеплению D в модели с непрерывным временем слабого по сравнению с рекомбинациями отбора по жизнеспособности по двум диаллельным локусам показало, что D является быстрой переменной, а частоты аллелей — медленными переменными. Это позволяет без явного решения обыкновенных дифференциальных уравнений модели представить качественную картину генетической динамики популяции. На первом этапе коэффициент D быстро сходится к практически нулевому значению, означающему равновесия по сцеплению и являющемуся итогом быстрой динамики. После этого медленная эволюция частот аллелей происходит при условии равновесия по сцеплению, что существенно упрощает анализ. Таким образом, текущее генетическое состояние популяции является одновременно итогом и этапом микроэволюции. Приведенная картина подтверждает то положение С. Райта [9], что условие равновесия по сцеплению можно использовать, если давление отбора существенно меньше давления рекомбинаций.

Описанная картина дает лишь приближенное описание динамики, и возникает проблема оценивания отклонений от равновесия по сцеплению (отличие D от нуля). Данное отклонение можно охарактеризовать различным образом. Пионером изучения нарушения равновесия по сцеплению при слабом отборе является М. Кимура [12], использовавший в качестве показателя нарушения $z \equiv p_{11}p_{22}/(p_{12}p_{21})$. Преимуществом предложенного нами подхода является оценивание популярного показателя такого нарушения D с использованием коэффициента эпистаза ϵ в терминах привычных концепций средних эф-

фектов δ в количественной генетике: $D^* = (\delta_{11}(\mathbf{p}) - \delta_{11}^*(\mathbf{p}) - \delta_{11}^*(\mathbf{p})) \frac{\mu}{r} x_1 y_1 \equiv \varepsilon_{11}(\mathbf{p}) \frac{\mu}{r} x_1 y_1$.

Это показывает важность отношения давления отбора μ к давлению рекомбинаций r для характеристики квазиравновесия D^* (и динамики генетического состояния популяции как сингулярно возмущенной, а не использование отношения ε/r). Кроме того, использование теории сингулярных возмущений позволило провести оценивание D^* на твердой теоретической основе. Данный подход не требует изобретения специальных приемов для проблемы оценивания D^* и позволяет стандартизировать ее решение. Полученные результаты приложимы и к диплоидной популяции при соответствующем вычислении коэффициента эпистаза [20].

СПИСОК ЛИТЕРАТУРЫ

1. 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.
2. Hill W.G. Sewall Wright and quantitative genetics // Genome. 1989. V. 31. № 1. P. 190–195.
3. Любич Ю.И. Основные понятия и теоремы эволюционной генетики свободных популяций // Успехи матем. наук. 1971. Т. 26. № 5. С. 51–116.
4. Любич Ю.И. Математические структуры в популяционной генетике. Киев: Наук. думка, 1983. 296 с.
5. Свирежев Ю.М., Пасеков В.П. Основы математической генетики. М.: Наука, 1982. 511 с.
6. Kingman J. A mathematical problem in population genetics // Proc. Cambr. Philos. Soc. 1961. V. 57. № 3. P. 574–582.
7. Кун Л.А., Лупичев Л.Н. Устойчивые полилокусные полиаллельные популяции при аддитивном отборе // Пробл. передачи информ. 1983. Т. 20. Вып. 1. С. 86–96.
8. Fisher R.A. The Genetical Theory of Natural Selection. Oxford: Clarendon Press, 1930. 272 p.
9. Wright S. The distribution of gene frequencies in population // Proc. Natl Acad. Sci. USA. 1937. V. 23. № 6. P. 307–320.
10. Пасеков В.П. Асимптотический анализ отбора в полилокусной полиаллельной популяции // Докл. АН СССР. Кибернетика и теория регулирования. 1984. Т. 277. № 6. С. 1338–1341.
11. Nagylaki T. The evolution of multilocus systems under weak selection // Genetics. 1993. V. 134. № 6. P. 624–647.
12. Kimura M. Attainment of quasi linkage equilibrium when gene frequencies are changing by natural selection // Genetics. 1965. V. 52. № 5. P. 875–890.
13. Shahshahani S. A new mathematical framework for the study of linkage and selection // Memoirs Amer. Math. Soc. 1979. V. 7. № 211. 34 p.
14. Hoppensteadt F.C. Mathematical Theories of Populations: Demographics, Genetics and Epidemics. SIAM: Philadelphia, 1975. 72 p. doi 10.1137/1.9781611970487
15. Пасеков В.П. Математические модели эколого-генетических взаимодействий / Итоги науки и техники. Сер. “Математическая биология и медицина”. Т. 2. М.: ВИНТИ, 1988. С. 4–80.
16. Пасеков В.П. О фундаментальной теореме естественного отбора Фишера для пространственно структурированных популяций при интенсивных миграциях // ДАН. Сер. Общая биология. 2007. Т. 412. № 1. С. 137–141.
17. Auger P., Bravo de la Parra R. et al. Aggregation methods in dynamical systems and applications in population and community dynamics // Physics Life Rev. 2008. V. 5. P. 79–105.
18. Banasiak J., Lachowicz M. Methods of Small Parameter in Mathematical Biology. Berlin: Springer, 2014. 285 p.
19. Hek G. Geometrical singular perturbation theory in biological practice // J. Math. Biol. 2010. V. 60. P. 347–386.
20. Пасеков В.П. Квазиравновесие по сцеплению при слабом двулокусном отборе по жизнеспособности. II. Лocusы с множественными аллелями // Генетика. 2019. Т. 55. № 5.
21. Orr H.A. Fitness and its role in evolutionary genetics // Nat. Rev. Genet. 2009. V. 10. P. 531–539.
22. Новожилов И.В. Методы разделения движений. Конспект лекций. М.: МЭИ, 1981. 48 с.

Quasi Linkage Equilibrium under Weak Two-Locus Viability Selection. I. Haploid Population with Diallelic Loci

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 aim of the work is to study a weak viability selection at two diallelic loci with standardization of approaches due to the use of perturbation theory, to analyze the estimate of quasi-equilibrium value for linkage dis-

equilibrium coefficient D and to obtain results in terms of average effects in quantitative genetics and in terms of the theory of singular perturbations in mathematics. Approximation by ordinary differential equations of a model with a discrete time of a random mating population with nonoverlapping generations under weak selection is considered. Weak selection is regarded as a perturbation of the model without selection. The resulting model is singularly perturbed, that is, fast (D) and slow (allele frequencies) variables are separated in it. The equation in the first approximation for quasi-equilibrium D is obtained using the first terms in the Taylor expansion of the model functions. It coincides with the corresponding part of the first approximation system of the asymptotic series for the solution of singularly perturbed equations. The first approximation for the quasi-equilibrium of D is $D^* = \varepsilon(\mathbf{p}) \frac{\mu}{r} x(1-x)y(1-y)$, $\varepsilon(\mathbf{p}) \equiv v_{11}(\mathbf{p}) - v_{12}(\mathbf{p}) - v_{21}(\mathbf{p}) + v_{22}(\mathbf{p})$. Here μ is the intensity of selection, r is the recombination coefficient, $\varepsilon_{i_1 i_2}(\mathbf{p})$ 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 and y are the frequencies of alleles constituting the genotype $i_1 i_2$. Evolution of the nonequilibrium state of the population actually takes place under quasi-equilibrium D (as the result of the stage of fast variable D dynamics) and the current values of slow x , y .

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