

А. В. ПЕТРОСЯН, Ш. М. КОЧАРЯН, А. С. ЗУРАБЯН

УСЛОВИЕ ДОСТИЖЕНИЯ РАВНОВЕСИЯ В БОЛЬШОЙ ПАНМИКТИЧЕСКОЙ ПОПУЛЯЦИИ

В 1908 г. Гарди [5] опубликовал работу, в которой для идеальной (гардниевской) популяции установил условие равновесия в панмиктической популяции для случая одного диаллельного локуса. Холден [4] показал для случая двух диаллельных локусов, что при независимости комбинирования генов в гаметах популяция достигает равновесия. Применяя различные математические подходы, Гейрингер [2, 3] и Эландт-Джонсон [1] доказали достаточность этого условия для более общего случая, т. е. для произвольного числа локусов с произвольным числом аллелей в каждом. Однако во всех указанных выше работах не была доказана необходимость условия независимости комбинирования генов в гаметах или зиготах.

В настоящей работе нами, используя метод математической индукции, доказывается необходимость и достаточность этого условия для достижения равновесия в большой панмиктической популяции.

Модель гаметы и генотипа. Согласно (1), рассматривается диплоидная перекрестнооплодотворяющаяся популяция Г. Каждая особь популяции имеет n локусов, которые обозначаются:

$$\alpha_1, \alpha_2, \dots, \alpha_n.$$

В h -том локусе ($h=1, 2, \dots, n$) может присутствовать либо аллель α_{ph} , либо — α_{qh} , т. е. рассматривается диаллельный случай. Тогда модель генетического строения гаметы можно представить в виде формального произведения:

$$\gamma_k = \prod_{h=1}^n \alpha_{ph}^{u_{ph}^k} \cdot \alpha_{qh}^{u_{qh}^k}, \quad (1)$$

где u_{ph}^k и u_{qh}^k — числа, принимающие значения 0 или 1 и связанные условием:

$$u_{ph}^k + u_{qh}^k = 1 \quad 1 \leq h \leq n. \quad (1a)$$

Символом γ_k обозначается k -тая гамета. Очевидно, что число таких всевозможных типов гамет равно 2^n . А так как зигота образуется слиянием гамет, то модель генотипа можно представить в виде:

$$\begin{aligned}\Gamma_{i,j} &= \gamma_i \cdot \gamma_j = \prod_{h=1}^n \alpha_{ph}^{u_{ph}^i} \cdot \alpha_{qh}^{u_{qh}^i} \cdot \prod_{h=1}^n \alpha_{ph}^{u_{ph}^j} \cdot \alpha_{qh}^{u_{qh}^j} = \\ &= \prod_{h=1}^n \alpha_{ph}^{u_{ph}^i + u_{ph}^j} \cdot \alpha_{qh}^{u_{qh}^i + u_{qh}^j} = \prod_{h=1}^n \alpha_{ph}^{w_{ph}^{i,j}} \cdot \alpha_{qh}^{w_{qh}^{i,j}},\end{aligned}\quad (2)$$

где числа $w_{ph}^{i,j}$ и $w_{qh}^{i,j}$ принимают значения 0, 1 и 2. Число таких генотипов будет $2^n \times 2^n$. При этом $\Gamma_{i,j}$ совпадает с $\Gamma_{j,i}$. Кроме того, существуют такие k, m и k', m' , что даже при $k \neq k'$ и $m \neq m'$ $\Gamma_{k,m}$ совпадает с $\Gamma_{k',m'}$. Учитывая это, наряду с моделью генотипа (2), введем в рассмотрение также и другую модель:

$$\Gamma_s = \prod_{h=1}^n \alpha_{ph}^{w_{ph}^s} \alpha_{qh}^{w_{qh}^s}, \quad (3)$$

где числа w_{ph}^s и w_{qh}^s принимают значения 0, 1 и 2 и связаны соотношением:

$$w_{ph}^s + w_{qh}^s = 2. \quad (3a)$$

Число всевозможных генотипов будет 3^n , поскольку при фиксированном n возможны 3 различных исхода:

$$\alpha_{ph} \alpha_{ph}; \alpha_{ph} \alpha_{qh} \text{ и } \alpha_{qh} \alpha_{qh}.$$

В дальнейшем мы будем оперировать в основном этой моделью (3).

Вероятности гамет и генотипов. Обозначим через g_k вероятность гаметы γ_k . Тогда

$$\sum_{k=1}^{2^n} g_k = 1.$$

Вероятность генотипа $\Gamma_{i,j}$ обозначим через $z_{i,j}$:

$$\sum_{i=1}^{2^n} \sum_{j=1}^{2^n} z_{i,j} = 1,$$

а вероятность генотипа Γ_s в случае модели (3) через z_s :

$$\sum_{s=1}^{3^n} z_s = 1.$$

Кроме того, обозначим через $c_k(s)$ условную вероятность продуцирования k -той гаметы s -тым генотипом. Тогда вероятность k -той гаметы в следующем поколении g_k^* будет:

$$g_k^* = \sum_{s=1}^{3^n} c_k(s) z_s. \quad (4)$$

Постановка задачи и ограничения.

Определение. 1. Мы будем считать, что популяция Γ достигла равновесия, если выполняется условие $z_s^{(t)} = z_s^{(t+1)}$,

$$\text{где} \quad s = 1, 2, \dots, 3^n, \quad (5)$$

а верхний индекс показывает номер поколения.

Нас интересует, при каких условиях популяция достигает равновесия.

Примем следующие ограничения: в популяции осуществляется панмиксия, поколения не перекрываются, в популяции отсутствуют давление миграции, мутации и отбора, признак не сцеплен с полом и вероятность рекомбинации для любой пары локусов равна $1/2$. Популяцию, удовлетворяющую всем поставленным требованиям, обозначим через Γ° . Для такой популяции можно легко показать, что равенство (5) эквивалентно следующему:

$$g_k^{(t)} = g_k^{(t+1)} \quad k = 1, 2, \dots, 2^n. \quad (6)$$

Вероятности генов.

Определение 2. Под вероятностью гена α_{ph} (или его аллеля α_{qh}) в множестве гамет мы будем понимать вероятность события, заключающегося в том, что у случайно выбранной гаметы в h -том локусе будет ген α_{ph} (α_{qh}). Эту вероятность обозначим a'_{ph} (a'_{qh}). Тогда по определению 2 имеем:

$$\begin{aligned} a'_{ph} &= \sum_{k=1}^{2^n} u_{ph}^k g_k, \\ a'_{qh} &= \sum_{k=1}^{2^n} u_{qh}^k g_k. \end{aligned} \quad (7)$$

Определение 3. Под вероятностью гена α_{ph} (или его аллеля α_{qh}) в множестве зигот мы будем понимать вероятность события, заключающегося в том, что ген, случайно выбранный из двух генов h -того локуса случайно выбранной зиготы будет α_{ph} (α_{qh}). Эту вероятность обозначим a''_{ph} (или a''_{qh}). Тогда по определению 3

$$\begin{aligned} a''_{ph} &= \frac{1}{2} \sum_{s=1}^{3^n} w_{ph}^s z_s \\ a''_{qh} &= \frac{1}{2} \sum_{s=1}^{3^n} w_{qh}^s z_s. \end{aligned} \quad h = 1, 2, \dots, n \quad (8)$$

Покажем эквивалентность определений 2 и 3:

$$\begin{aligned} a''_{ph} &= \frac{1}{2} \sum_{s=1}^{3^n} w_{ph}^s z_s = \frac{1}{2} \sum_{i=1}^{3^n} \sum_{j=1}^{2^n} w_{ph}^{i,j} z_{i,j} = \frac{1}{2} \sum_{i=1}^{2^n} \sum_{j=1}^{2^n} (u_{ph}^i + u_{qh}^j) g_i g_j = \\ &= \frac{1}{2} \sum_{i=1}^{2^n} g_i (u_{ph}^i \sum_{j=1}^{2^n} g_j + \sum_{j=1}^{2^n} g_j u_{ph}^i). \end{aligned}$$

И теперь с учетом (7) получим:

$$\begin{aligned} a_{ph}'' &= \frac{1}{2} \sum_{i=1}^{2^n} g_i (u_{ph}^i + a_{ph}') = \frac{1}{2} \left(\sum_{i=1}^{2^n} u_{ph}^i g_i + a_{ph}' \sum_{i=1}^{2^n} g_i \right) = \\ &= \frac{1}{2} (a_{ph}' + a_{ph}') = a_{ph}'. \end{aligned}$$

Таким образом, $a_{ph}'' = a_{ph}' = a_{ph}$.

Аналогично доказывается: $a_{qh}'' = a_{qh}' = a_{qh}$.

Очевидно, что $a_{ph} + a_{qh} = 1$.

Докажем лемму о вероятности генов, принимая во внимание, что вероятности генов для популяции Γ° одинаковы в последующих поколениях [1].

Л е м м а 1. Вероятность гена α_{ph} ($1 \leq h \leq n$) относительно n и $n-1$ пары локусов одинакова.

Д о к а з а т е л ь с т в о.

Если в гамете с n локусами рассматриваются $n-1$ локус, то такая гамета γ_k^{n-1} по формуле (1) будет иметь вид:

$$\gamma_k^{n-1} = \prod_{h=1}^{n-1} \alpha_{ph}^{u_{ph}^k} \alpha_{qh}^{u_{qh}^k} \quad k = 1, 2, \dots, 2^{n-1}.$$

Если теперь ввести в рассмотрение n -тый локус, то число типов гамет, естественно, удвоится:

$$\gamma_k^{n1} = \alpha_{ph} \gamma_k^{n-1},$$

$$\gamma_k^{n2} = \alpha_{qh} \gamma_k^{n-1}.$$

При этом вероятности гамет будут связаны соотношением:

$$P(\gamma_k^{n-1}) = P(\gamma_k^{n1}) + P(\gamma_k^{n2}). \quad (9)$$

Вероятность гена α_{ph} , если исключить из рассмотрения n -тый локус, по формуле (7) будет:

$$a_{ph}^{n-1} = \sum_{k=1}^{2^{n-1}} u_{ph}^k P(\gamma_k^{n-1}), \quad (10)$$

а вероятность гена α_{ph} с учетом n -того локуса по определению 2

$$a_{ph}^n = \sum_{k=1}^{2^{n-1}} u_{ph}^k \cdot P(\gamma_k^{n1}) + \sum_{k=1}^{2^{n-1}} u_{ph}^k P(\gamma_k^{n2}) = \sum_{k=1}^{2^{n-1}} u_{ph}^k [P(\gamma_k^{n1}) + P(\gamma_k^{n2})].$$

С учетом равенств (9) и (10) получим: $a_{ph}^n = a_{ph}^{n-1} = a_{ph}$. Аналогично доказывается лемма для a_{qh} .

Теорема о равновесии.

Предварительно докажем лемму 2.

Лемма 2. Если для популяции Γ° равновесие относительно n локусов наступило, то оно наступило и для меньшего числа локусов.

Доказательство:

Доказательство непосредственно вытекает из равенств (6) и (9).

Теорема о равновесии. Необходимым и достаточным условием для того, чтобы популяция Γ° достигла равновесия, является независимость комбинирования генов в гаметах.

Доказательство:

Докажем сначала достаточность условия. Из условия независимости комбинирования генов в гаметах следует:

$$g_k = \prod_{h=1}^n a_{ph}^{u_k} a_{qh}^{u_k}, \quad k = 1, 2, \dots, 2^n, \quad (11)$$

а вероятность генотипа

$$z_s = 2^n \prod_{h=1}^n \frac{a_{ph}^{w_s} a_{qh}^{w_s}}{(w_{ph}^s)! (w_{qh}^s)!}, \quad s = 1, 2, \dots, 3. \quad (12)$$

Вероятность $c_k(s)$ может быть выражена согласно (2) в виде

$$\begin{aligned} c_k(s) &= \prod_{h=1}^n \frac{1}{2} \binom{w_{ph}^s}{u_{ph}^k} \binom{w_{qh}^s}{u_{qh}^k} = \frac{1}{2^n} \prod_{h=1}^n \frac{(w_{ph}^k)!}{(u_{ph}^k)! (w_{ph}^s - u_{ph}^k)!} \frac{(w_{qh}^s)!}{(u_{qh}^k)! (w_{qh}^s - u_{qh}^k)!} = \\ &= \frac{1}{2^n} \prod_{h=1}^n \frac{(w_{ph}^s)! (w_{qh}^s)!}{(w_{ph}^s - u_{ph}^k)! (w_{qh}^s - u_{qh}^k)!}. \end{aligned}$$

При этом $c_k(s)$ равняется 0, если

$$(w_{ph}^s - u_{ph}^k)! (w_{qh}^s - u_{qh}^k)! = 2. \quad (13)$$

Так как при выполнении (13) из равенства (1а) и (3а) вытекает, что либо $w_{ph}^s = 0$, а $u_{ph}^k = 1$, либо $w_{qh}^s = 0$, а $u_{qh}^k = 1$, в обоих случаях число сочетаний из 0 по 1 равно 0.

Таким образом, в зависимости от k и s

$$c_k(s) = \begin{cases} 0 \\ \frac{1}{2^n} \prod_{h=1}^n (w_{ph}^s)! (w_{qh}^s)! \end{cases}. \quad (14)$$

Подставляя (12) и (14) в (4), получим:

$$g_k^* = \sum_{c_k(s) \neq 0} \prod_{h=1}^n a_{ph}^{w_{ph}^s} a_{qh}^{w_{qh}^s}. \quad (15)$$

Среди членов произведения в выражении (15) у нас остались только те, которые несут в себе гамету $\prod_{h=1}^n a_{ph}^{u_{ph}^k} a_{qh}^{u_{qh}^k}$, поскольку для остальных

ных $c_k(s) = 0$. Число оставшихся таких произведений равно числу всевозможных типов гамет, то есть 2^n . Следовательно:

$$\begin{aligned} g_k^* &= \sum_{m=1}^{2^n} \prod_{h=1}^n a_{ph}^{u_{ph}^m + u_{ph}^k} a_{qh}^{u_{qh}^m + u_{qh}^k} = \prod_{h=1}^n a_{ph}^{u_{ph}^k} a_{qh}^{u_{qh}^k} \sum_{m=1}^{2^n} \prod_{h=1}^n a_{ph}^{u_{ph}^m} a_{qh}^{u_{qh}^m} = \\ &= \prod_{h=1}^n a_{ph}^{u_{ph}^k} a_{qh}^{u_{qh}^k} \sum_{m=1}^{2^n} g_m = \prod_{h=1}^n a_{ph}^{u_{ph}^k} a_{qh}^{u_{qh}^k} = g_k. \end{aligned}$$

Таким образом, нами новым способом доказана достаточность условия независимости комбинирования генов в гаметах (зиготах) для достижения равновесия популяции Γ° .

Докажем теперь необходимость условия методом математической индукции.

Базис индукции. Для числа локусов $n=2$ имеем 2 пары генов: $\alpha_{p1}; \alpha_{q1}$ и $\alpha_{p2}; \alpha_{q2}$. Число всевозможных гамет равно 4: $\gamma_1 = \alpha_{p1} \alpha_{p2}$; $\gamma_2 = \alpha_{p1} \alpha_{q2}$; $\gamma_3 = \alpha_{q1} \alpha_{p2}$; $\gamma_4 = \alpha_{q1} \alpha_{q2}$.

Вероятность i -той гаметы $P(\gamma_i)$, как и раньше, будем обозначать g_i , причем $\sum_{i=1}^4 g_i = 1$.

Число всевозможных генотипов равно 9:

$$\begin{aligned} \Gamma_1 &= \alpha_{p1} \alpha_{p1} \alpha_{p2} \alpha_{p2} & \Gamma_4 &= \alpha_{p1} \alpha_{q1} \alpha_{p2} \alpha_{p2} & \Gamma_7 &= \alpha_{q1} \alpha_{q1} \alpha_{p2} \alpha_{p2} \\ \Gamma_2 &= \alpha_{p1} \alpha_{p1} \alpha_{p2} \alpha_{q2} & \Gamma_5 &= \alpha_{p1} \alpha_{q1} \alpha_{p2} \alpha_{q2} & \Gamma_8 &= \alpha_{q1} \alpha_{q1} \alpha_{p2} \alpha_{q2} \\ \Gamma_3 &= \alpha_{p1} \alpha_{p1} \alpha_{q2} \alpha_{q2} & \Gamma_6 &= \alpha_{p1} \alpha_{q1} \alpha_{q2} \alpha_{q2} & \Gamma_9 &= \alpha_{q1} \alpha_{q1} \alpha_{q2} \alpha_{q2}. \end{aligned}$$

Вероятность генотипа Γ_s обозначим z_s ($s = 1 \div 9$) и $\sum_{s=1}^9 z_s = 1$.

По формуле (4), $g_1^* = \sum_{s=1}^9 c_k(s) z_s = 1 z_1 + 1/2 z_2 + 1/2 z_4 + 1/4 z_5$.

С учетом панмиксии $g^* = g_1 g_1 + g_1 g_2 + g_1 g_3 + \frac{1}{2} (g_1 g_4 + g_2 g_3)$.

Для выполнения условия равновесия для первого типа гамет необходимо выполнение равенства:

$$g_1 g_4 = g_2 g_3. \quad (16)$$

Применяя аналогичные рассуждения к g_2^* , g_3^* и g_4^* , всякий раз приходим к необходимости равенства (16). Поскольку считаем, что равновесие относительно каждого из этих двух локусов наступило, то, применяя леммы 2 и 1, получим:

$$\begin{aligned} g_1 &= a_{p1} P(\alpha_{p2}/\alpha_{p1}); & g_2 &= a_{p1} P(\alpha_{q2}/\alpha_{p1}); \\ g_3 &= a_{q1} (P \alpha_{p2}/\alpha_{q1}); & g_4 &= a_{q1} P(\alpha_{q2}/\alpha_{q1}). \end{aligned} \quad (17)$$

Подставляя (17) в (16) и используя условия

$$P(\alpha_{p2}/\alpha_{p1}) + P(\alpha_{q2}/\alpha_{q1}) = 1,$$

$$P(\alpha_{p2}/\alpha_{q1}) + P(\alpha_{q2}/\alpha_{q1}) = 1,$$

получим:

$$P(\alpha_{p2}/\alpha_{p1}) = P(\alpha_{p2}/\alpha_{q1}) = P(\alpha_{p2}) = a_{p2},$$

$$P(\alpha_{q2}/\alpha_{p1}) = P(\alpha_{q2}/\alpha_{q1}) = P(\alpha_{q2}) = a_{q2}.$$

Аналогично получаем:

$$P(\alpha_{p1}/\alpha_{p2}) = a_{p1},$$

$$P(\alpha_{q1}/\alpha_{p2}) = a_{q1}.$$

Подставляя полученные выражения в (17) будем иметь

$$g_k = \prod_{h=1}^2 a_{ph}^{u_k} a_{qh}^{u_k} \quad k = 1, 2, 3, 4.$$

На этом завершается доказательство базиса.

Прежде чем перейти к индукционному шагу, докажем базис индукции в более общем виде.

Лемма 3. Чтобы наступило равновесие относительно двух локусов, из которых один имеет два аллеля, а другой l аллелей, где $l > 2$, необходима независимость комбинирования генов в гаметах.

Доказательство:

Число всевозможных типов гамет будет $2l$. Число всевозможных генотипов — $N_l = \frac{3}{2}(l+1)l$. По формуле (4) имеем

$$g_1 = \sum_{s=1}^{N_l} c_k(s) z_s.$$

Далее, проводя выкладки, аналогичные базису индукции, приходим к необходимости выполнения равенств:

$$g_1 g_4 = g_2 g_3$$

$$g_1 g_5 = g_2 g_4$$

$$g_1 g_{2l} = g_2 g_{2l-1}.$$

Доказательство леммы завершается приведением рассуждений, аналогичных базису индукции.

Индукционный шаг. Предположение индукции следующее: в случае $n-1$ локуса для установления равновесия необходима независимость комбинирования генов в гаметах. Мы хотим доказать это утверждение для n локусов.

По условию равновесие для n локусов наступило. Тогда по лемме 2 оно наступило и для $n-1$ локусов. В этом случае по предположению индукции

$$g_k^{n-1} = \prod_{h=1}^{n-1} a_{ph}^{u_k} a_{qh}^{u_k} \quad k = 1, 2, \dots, 2^{n-1},$$

где верхний индекс показывает число локусов.

Отметим, что по лемме 1 частоты генов α_{ph} и α_{qh} относительно n и $n-1$ локусов равны.

Поскольку равновесие относительно $n-1$ локусов наступило, то вероятности g_k^{n-1} будут постоянны во всех поколениях. Следовательно, можно каждую гамету, состоящую из $n-1$ генов, рассматривать, как один «ген», вероятность которого равна произведению вероятностей генов, составляющих этот условный ген.

Таким образом, задача сведена к двум локусам, из которых один обладает двумя аллелями α_{ph} и α_{qh} , а другой — $l = 2^{n-1}$.

Тогда по лемме 3 имеем:

$$g_k^n = a_{pn}^{uk} a_{qn}^{uk} \cdot \prod_{h=1}^{n-1} a_{ph}^{uk} a_{qh}^{uk} = \prod_{h=1}^n a_{ph}^{uk} a_{qh}^{uk}.$$

Таким образом доказана необходимость условия независимого комбинирования генов.

Доказанная теорема позволяет определить распределение вероятностей генотипов относительно n генов ($n=1,2,\dots$) в равновесной популяции по частоте генов. Это может найти применение в исследованиях по генетике популяций человека и медицинской генетике при установлении типа наследования и определения количества генов, вовлеченных в определение того или иного признака.

ЕрНИИММ,

Чаренцаванский филиал ВНИИ генетики,
Институт кардиологии МЗ АрмССР

Поступило 6.VII 1972 г.

Ա. Վ. ՊԵՏՐՈՍՅԱՆ, Շ. Մ. ԳՈՉԱՐՅԱՆ, Ա. Ս. ԶՈՒՐԱՅԱՆ

ՀԱՎԱՍԱՐԱԿՇՈՒԹՅԱՆ ՀԱՍՆԵԼՈՒ ՊԱՅՄԱՆՐ ԵՆԾ ՊԱՆՄԻՆԻԿԱԿԱՆ ՊՈՊՈՒԼՅԱՅԻԱՅՈՒՄ

Ա մ փ ո փ ո լ մ

Կառուցված է մեծ պանմիկտիկ պոպուլյացիայի մաթեմատիկական մոդելը, որի անհատները օժտված են դիալել լոկուսների կամայական թվով: Նոր մեթոդով ապացուցվել է գամետներում գենների կոմբինացման անկախության պայմանի բավարարությունը հավասարակշռության հասնելու համար, ինչպես նաև առաջին անգամ ապացուցվել է այդ պայմանի անհրաժեշտությունը: Հստակացված են այնպիսի բանաձևեր, որոնք թույլ են տալիս հավասարակշիռ պոպուլյացիայի համար ըստ գենների հաճախության հաշվարկել բոլոր հնարավոր գենոտիպների հաճախությունները:

Л И Т Е Р А Т У Р А

1. Elandt-Johnson, Regina C. Bull. Math. Biophysics., 29, 437, 1967.
2. Geiringer, Hilda. J. Math. and Phys., 26, 246, 1948.
3. Geiringer, Hilda. Bull. Math. Biophysics., 11, 197, 1949.
4. Haldane J. B. S. J. Genet., 22, 359, 1930.
5. Hardy G. H. Science, 28, 45, 1908.