

Глава 3

КОМПОНЕНТЫ ДИСПЕРСИЙ И ОТБОР

3.1. Цель разложения дисперсий

Если генотипическая ценность растений по определенному признаку зависит только от аддитивных эффектов генов a_i и влияние средовых модификаций e мало, то отбор в популяции растений с максимальной фенотипической ценностью P соответствует отбору лучших генотипов. Это означает, что эффективность отбора равна 100 %, т.е. сдвиг R среднего значения признака в популяции потомков отобранных растений равен сдвигу S , вызванному отбором по фенотипу в популяции родителей. Например, если среди прочих отобран один генотип $A_1A_1B_1B_1$ с фенотипической ценностью $m + a_1 + a_2$, то его потомство будет лучше потомств других отобранных генотипов. Эффекты доминирования и эпистаза в отобранной части популяции могут меняться при дальнейшем скрещивании и расщеплении. Поэтому, если на изменчивость фенотипа, кроме аддитивных эффектов, влияет доминирование или различные типы эпистаза, эффект селекции будет ниже 100 % даже при отсутствии ошибок, вызванных средовыми отклонениями. Ясно, что существенные средовые отклонения или низкий вклад генотипической изменчивости в фенотипическую изменчивость признака не позволит надежно отобрать нужные генотипы. Подобные рассуждения нуждаются в более детальной и конкретной формализации.

В этой главе подробно рассмотрена зависимость результативности отбора в поколениях от различных соотношений компонент генотипической и средовой изменчивости в популяциях самоопылителей и перекрестников. В отличие от главы 2 далее изменчивость эффектов генов и среды будем определять через их дисперсии. Эффективность отбора выразим через соотношения этих дисперсий в исходной популяции, т.е. переходим от модели эффектов и взаимодействий, отраженных для случая дигенного наследования признака в табл. 1, к аддитивной модели соответствующих дисперсий

$$\sigma_G^2 = \sigma_a^2 + \sigma_d^2 + \sigma_i^2 + \sigma_j^2 + \sigma_l^2,$$

здесь σ_G^2 — общая дисперсия генотипической изменчивости признака в популяции (генотипическая дисперсия); σ_a^2, σ_d^2 — дисперсии аддитивных и доминантных эффектов генов. Три последние дисперсии связаны с изменчивостью эпистатических эффектов в популяции (см. раздел 1.2).

Для упрощения формул будем полагать, что эти три дисперсии невелики по сравнению с остальными. Наконец, в соответствии с формулой (1), дисперсия σ_P^2 общей фенотипической изменчивости признака в популяции (фенотипическая дисперсия) разлагается на σ_G^2 и σ_e^2 — дисперсию средовой изменчивости (средовая дисперсия). Будет использована также σ_V^2 — дисперсия селекционных ценностей генотипов, составляющих популяцию.

Поскольку эффективность отбора оказывается зависимой от соотношения дисперсий, возникает задача разделения σ_P^2 на компоненты, связанные с генетическими эффектами различных типов и средовыми σ_e^2 . Биометрико-генетические модели и методы, основанные на специальных схемах опытов, обычно включающих скрещивания, позволяют провести такое разделение. Модели дают также возможность изучить эффективность различных схем отбора.

3.2. Оценка наследуемости

На рис. 7 показан эффект отбора R , т.е. сдвиг признака X у потомства в ответ на сдвиг S среднего значения признака, вызванного отбором в родительской популяции. *Селекционным дифференциалом* S называется разность между средним фенотипическим значением P' признака у растений той части популяции, которая отбирается для образования следующего поколения, и средним P исходной популяции до отбора. Эффект отбора (R) равен разности среднего значения (P_1) признака в популяции потомков отобранных растений и исходного значения P . Если $R < S$ ($R/S < 1$), то эффективность отбора меньше 100 %.

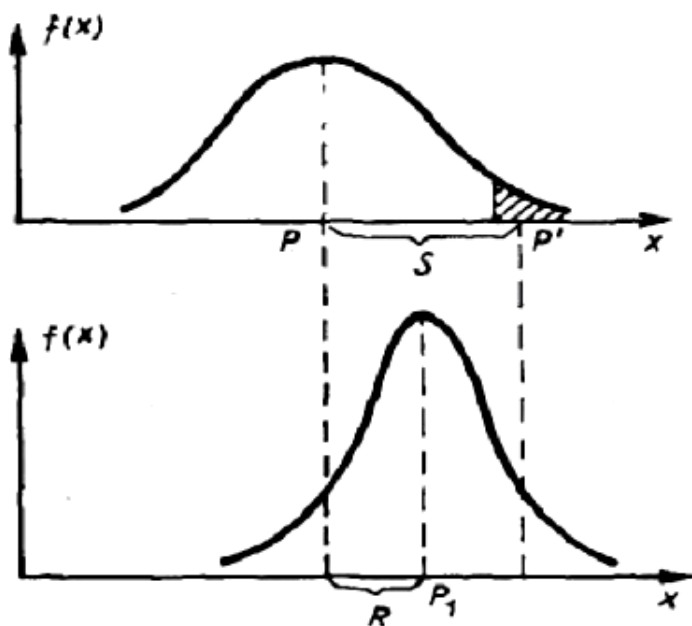


Рис 7. Схема связи селекционного дифференциала (S) в исходной популяции и эффекта отбора (R) в популяции потомков отобранных растений. Заштрихована доля отобранных родительских растений ($s\%$)

Селекционный дифференциал S удобнее выражать через квадратный корень фенотипической дисперсии (σ_P), т.е. $S = k \cdot \sigma_P$. Коэффициент k связан с долей растений, отбираемых в исходной популяции. Например, если отбирается 20 % ($s = 20\%$), то $k = 1,4$; $s = 10\%$ — $k = 1,76$; $s = 5\%$ — $k = 2,06$; если оставляют 2% растений, то $k = 2,42$. Величину k называют стандартным селекционным дифференциалом.

Между S и R можно установить регрессионную связь:

$$R = h^2 S = h^2 k \sigma_P, \quad (13)$$

где h^2 — коэффициент регрессии R на S называется коэффициентом наследуемости. Он указывает, какую часть селекционного дифференциала составляет эффект отбора.

Доказано, что в зависимости от способа опыления растений, схемы скрещивания, используемой в селекционной программе, и других факторов h^2 по-разному выражается через дисперсии генетических и средовых эффектов в исходной популяции. Например, предполагается провести отбор в популяции перекрестников. Тогда h^2 оценивают по отношению дисперсии селекционных ценностей σ_V^2 и дисперсии общей фенотипической ценности (фенотипическая дисперсия) для генотипов популяции до отбора:

$$h^2 = \sigma_V^2 / \sigma_P^2 \quad (14)$$

σ_V^2 при моногенном наследовании — это обычная дисперсия величин, представленных в формуле (12).

Применяемый в подобных ситуациях со скрещиванием и расщеплением h^2 называется *коэффициентом наследуемости в узком смысле*. Если же отбор будет проведен среди константных в поколениях генотипов (например, среди линий самоопылителей, клонов), то в формуле (13) используют *коэффициент наследуемости в широком смысле*.

Для отличия от h^2 его обычно обозначают H^2 :

$$H^2 = \sigma_G^2 / \sigma_P^2, \quad (15)$$

здесь σ_G^2 — дисперсия генотипической изменчивости исходного набора генотипов до отбора (генотипическая дисперсия).

Следует подчеркнуть, что речь всюду идет об эффектах однократного отбора. Перед вторым циклом отбора все генетические параметры селекционного материала надежнее оценить заново, так как они могли измениться именно вследствие проведенного отбора. P и G были определены в модели (1). Там же пояснялось, что все эффекты генов, а значит, и общая генотипическая ценность может меняться даже без отбора — от условий выращивания. Следовательно, коэффициент наследуемости также величина изменчивая. В зависимости от среды коэффициент наследуемости в популяции по данному признаку будет больше или меньше. По коэффициенту можно подбирать условия, в которых эффективность отбора будет выше. Среда, где четко проявляются генетические различия растений, где коэффициент наследуемости велик и R близко к S , называется *дифференцирующей средой*.

Например, для выращивания томатов в открытом грунте желательнее быстрое, дружное прорастание после посева. М. Эль-Сайед и С. Джон определили наследуемость (h^2) признака «скорость прорастания» при разной температуре воздуха:

близкая к среднепогодной		0,336
13°C		0,549
днем (16 ч) 27°C, ночью (8 ч)	17°C	0,644
10°C		0,676

Видим, что хотя h^2 существенно зависит от условий выращивания, в этих опытах не удалось подобрать среду, которую можно назвать дифференцирующей.

Отметим, что массовый отбор в популяции ожидается достаточно эффективным, если $h^2 > 0,4$. Если это условие не выполнено, следует применять более сложные схемы, чем массовый отбор в популяции, где оценивался h^2 .

Для признаков с пренебрежимо малым взаимодействием генотип-среда (см. раздел 1.2), например, для устойчивости овса к ржавчине, можно пользоваться одной оптимальной средой отбора. Но чаще это взаимодействие велико, как, например, для урожайности сортов 3-го и 4-го на рис. 5. Среда E здесь наиболее соответствует понятию дифференцирующей. Но селекция в этой среде может привести к отбору генотипа (сорт 3), который в обычных условиях выращивания (диапазон A — B) не даст прибавки урожайности по сравнению с отбракованным генотипом (сорт 4).

Взаимодействия генотип-среда порождает проблему выбора оптимальной для отбора нормы высева, сроков сева, подбора агрофона и т.п. Селекционеры Велакарденас и Фрей установили, что среда, дающая оптимальные результаты селекции овса, зависит от признака: высокая норма высева позволяла вести успешный отбор по высоте растений и урожайности, низкая норма благоприятствовала отбору по числу колосков в метелке, для селекции на крупность зерна были пригодны обе нормы высева.

Коэффициенты наследуемости одного признака могут существенно зависеть от состава популяции, климатических условий года, места испытаний и т.д. (табл. 12). Кроме того, формулы (14) и (15) основаны на ряде предположений. Например, в формуле (14)

предполагается отсутствие эпистатических эффектов, равновесность исходной популяции и популяции потомков и т.д. Все это требует осторожности при использовании формулы (13) для прогнозирования результатов отбора. Но все же коэффициенты наследуемости дают возможность хотя бы приблизительно оценить, стоит ли проводить отбор в конкретном селекционном материале по простым схемам при определенных условиях выращивания.

12. Пределы изменений величины H^2 для некоторых признаков культурных растений [1]

Вид	Признак	H^2
Мягкая озимая пшеница	Продуктивность растения	0,58...0,97
	Продуктивная кустистость	0,21...0,93
	Масса 1000 зерен	0,62...0,98
Томаты	Урожайность	0,27...0,54
	Средняя масса плода	0,06...0,95
	pH сока плодов	0,34...0,79
	Содержание сухих веществ в них	0,34...0,85
	Скороспелость	0,29...0,56

Отметим, что оценка коэффициентов наследуемости непосредственно в процессе или после отбора прорабатываемого селекционного материала имеет лишь научное значение. Для практической селекции может быть более полезно произвести небольшую по объему случайную выборку генотипов из прорабатываемого материала и оценить коэффициенты наследуемости по ней. При этом следует учесть, что на такую оценку в реальных полевых условиях может уйти сезон или два сезона: выращивание родительских генотипов, отбор, выращивание и оценка отобранных потомков, вычисление h^2 по формуле (13). Кроме того, напомним, что простейшие генетические параметры, типа коэффициентов наследуемости, зависят от условий выращивания в год оценки. Результаты отбора во всем материале, проведенного на следующий, несходный по условиям год, могут быть неэффективными.

Рассмотрим частные случаи и формулы для оценок коэффициентов наследуемости без отбора по ограниченной выборке данных.

Дисперсия фенотипической ценности растений в поколении F_2 (скрещивание двух чистых линий $P_1 \times P_2$ и расщепление F_1) или растений двойного гибрида (например, $AB \times CD$, где A, B, C, D — гомозиготные родительские инцухт-линии) оценивается по обычной формуле:

$$\sigma_p^2 = [\sum x_i^2 - (\sum x_i)^2/n]/(n-1), \quad (16)$$

здесь n — число растений, на которых было проведено измерение x_i признака.

По общей формуле (16) оценивают σ_p^2 у совместно высаженных растений популяции или смеси чистых линий, клонов и т.д.

Дисперсия средовой изменчивости (или проще, средовая дисперсия σ_e^2), вызванной случайными различиями условий выращивания совместно высаженных растений, оценивается через дисперсии признака у растений отдельно высаженных поколений. Для случая популяции F_2 :

$$\sigma_e^2 = 1/3[\sigma^2(P_1) + \sigma^2(P_2) + \sigma^2(F_1)]. \quad (17)$$

Три последние дисперсии — это дисперсии изменчивости фенотипической ценности для растений P_1, P_2, F_1 , каждая из которых оценена по формуле (16). Ясно, что в этих поколениях генотипической изменчивости нет, следовательно, фенотипическая изменчивость совпадает со

средовой. Отметим, что из-за различий генетической структуры уровень средовой изменчивости может существенно отличаться в чистых линиях P_1 , P_2 и F_1 . Поэтому правильнее усреднять все три дисперсии фенотипической изменчивости по формуле (17).

Для случая двойного гибрида $(A \times B) \times (C \times D)$ или, проще, $AB \times CD$ дисперсия σ_e^2 также оценивается усреднением σ_p^2 тех отдельно высаженных поколений, где фенотипическая дисперсия совпадает со средовой

$$\sigma_e^2 = 1/6[\sigma^2(A) + \sigma^2(B) + \sigma^2(C) + \sigma^2(D) + \sigma^2(AB) + \sigma^2(CD)].$$

После этого дисперсия генотипической изменчивости оценивается разностью фенотипической и средовой

$$\sigma_G^2 = \sigma_p^2 - \sigma_e^2, \quad (18)$$

а коэффициент наследуемости — по формуле (15).

Рассмотрим пример вычисления H^2 по признаку «высота растения от земли до мужского соцветия» для популяции — двойного гибрида кукурузы (табл. 13).

Оценка средовой и генотипической дисперсии растений двойного гибрида: $\sigma_e^2 = 1/6(185,0 + \dots + 240,3) - 246,9$, $\sigma_G^2 = 475,3 - 246,9 = 228,4$. $H^2 = 228,4/475,3 = 0,48$.

13. Средняя высота (X) растений и фенотипическая дисперсия (σ_p^2) в инцухт-линиях, у простых гибридов и двойного гибрида, высаженных отдельно (по И. Бос)

	A	B	C	D	AB	CD	AB × CD
X	103,8	121,1	80,5	111,6	177,6	141,2	188,2
σ_p^2	185,0	256,0	90,3	285,6	424,4	240,3	475,3

Отметим, что такой способ расчета для представленных результатов дает заниженную оценку σ_e^2 . Дело в том, что заметна положительная статистическая связь средней высоты и σ_e^2 в поколениях. Поскольку X у двойного гибрида выше, чем в других поколениях, то σ_e^2 , по-видимому, больше 246,9. Значит H^2 меньше 0,48. Подобный недостаток оценок иногда удается компенсировать логарифмическим преобразованием исходных результатов измерений.

Коэффициент наследуемости в широком смысле выше, чем в узком смысле. Ведь в числителе формулы (15) стоит дисперсия, включающая также изменчивость доминирования и всех типов эпистаза. Эти эффекты не полностью сохраняются в поколениях при отборе, особенно у самоопылителей. h^2 в основном отражает лишь сохраняющуюся в поколениях аддитивную часть генотипической изменчивости. Поэтому полученная завышенная $H^2 = 0,48$ говорит о том, что эффективность массового отбора в популяции $AB \times CD$ будет низка. Отметим, что сопоставление величин H^2 и h^2 , оцененных в одной популяции, дает информацию о доле неаддитивной наследственной изменчивости признака.

Дополнительные трудности, снижающие реальную эффективность отбора, связаны с конкуренцией совместно высеянных растений, если у них разные генотипы. Этот и другие источники изменчивости, снижающие эффективность отбора, будут рассмотрены отдельно, в главах 6 и 10.

Рассмотрим другой способ оценки H^2 по изучаемому признаку (x) для набора константных генотипов. В большом «исходном» наборе клонов, чистых линий или гибридов F_1 для оценки H^2 случайным образом отбирают m генотипов. Семена каждого из них высевают (возможно, после размножения) на n делянках-повторениях опыта, которые для каждого генотипа размещены на

поле случайно (схема полной рандомизации). Формулы дисперсионного анализа данных x_{ij} (i — номер генотипа, j — повторения) представлены в табл. 14.

14. Дисперсионный анализ результатов испытания m генотипов

Источники изменчивости	SS_k , т.е. сумма квадратов	Степени свободы	Структура MS , т.е. средних квадратов
Различия:			
между генотипами	$n(\sum_i^m x_{i.}^2 - mx.^2)$	$df_1 = m - 1$	$\sigma_e^2 + n\sigma_G^2$
внутри генотипов	$\sum_i^m \sum_j^n x_{ij}^2 - n\sum_i^m x_{i.}^2$	$df_2 = m(n - 1)$	σ_e^2
Сумма	$\sum_i^m \sum_j^n x_{ij}^2 - nm\bar{x}^2$	$mn - 1$	

Примечание. Точка вместо индекса ($x_{i.}$, $x_{.}$) означает усреднение по нему.

Таким образом, с учетом формул из табл. 14, имеем

$$\sigma_e^2 = MS_2 = SS_2 / [m(n - 1)]; \quad \sigma_G^2 = (MS_1 - MS_2) / n;$$

$$H^2 = \sigma_G^2 / (\sigma_G^2 + \sigma_e^2).$$

Здесь предполагается, что единственный источник средней изменчивости — случайные колебания условий выращивания по повторениям. Отсюда становится ясно, что все коэффициенты наследуемости, характеризующие эффективность отбора, зависят от схемы и тщательности проведения полевых опытов.

Все три параметра, оцененные таким способом, относятся ко всему «исходному» набору генотипов, так как m генотипов отбирается из него случайным образом. Это соответствует так называемой модели II дисперсионного анализа — со случайными факторами. Поэтому H^2 оценивает долю генотипической изменчивости признака в фенотипической для всего «исходного» набора.

Для проверки значимости генотипической дисперсии применяют F -критерий Фишера. Если $MS_1/MS_2 = F_{\text{расч}}$ больше табличного значения $F_{0,05}(df_1, df_2)$, то σ_G^2 значимо больше нуля — среди генотипов «исходного» набора есть существенно различные. Таблицы значений $F_{0,05}$ и $F_{0,01}$ даны в приложениях 3 и 4. Рассмотрим упрощенный числовой пример (табл. 15).

15. Продуктивность растений четырех линий ($m = 4$) яровой пшеницы в опыте с тремя повторениями ($n = 3$) [2]

i	$j = 1$	$j = 2$	$j = 3$	$x_{i.}$
1	2,30	1,85	2,40	2,18
2	1,43	1,27	1,35	1,35
3	2,20	2,25	2,40	2,28
4	1,44	2,54	2,24	2,07

Примечание. $x_{i.}$ — средние по повторениям для линий с номерами i .

Проследим порядок вычислений для H^2 .

$$\begin{aligned}
 x_{..} &= 1,97. \quad MS_1 = 3(2,18^2 + 1,35^2 + \dots - 4 \cdot 1,97^2) / 3 = 0,54; \\
 MS_2 &= [2,30^2 + \dots + 2,24^2 - 3(2,18^2 + 1,35^2 + 2,28^2 + 2,07^2)] / 8 = 0,11; \\
 \sigma_G^2 &= (0,54 - 0,11) / 3 = 0,14; \quad H^2 = 0,14 / (0,14 + 0,11) = 0,55. \\
 F_{расч} &= 0,54 / 0,11 = 5,05 > F_{0,05} = 4,07.
 \end{aligned}$$

Следовательно, в выборке есть существенно различные генотипы, что подтверждает прогноз эффективности отбора генотипов по H^2 .

3.3 Структура изменчивости признака для самоопылителей

Для оценки h^2 необходимо вначале выяснить структуру генотипической дисперсии. Только после этого можно провести выделение ее компонент с целью формирования оценки дисперсии, стоящей в числителе формулы (15).

В результате алгебраических преобразований была получена формула генотипической дисперсии признака для произвольной популяции с частотами $p(A_1)$ и $q(A_2)$, эффектами a , d и коэффициентом инбридинга F :

$$\begin{aligned}
 \sigma_G^2 &= 2pq(1+F)[a - (p-q)d(1-F)/(1+F)]^2 + \\
 &+ 4pqd^2[F + pq(1-F)^2](1-F)/(1+F). \quad (19)
 \end{aligned}$$

Ввиду громоздкости формулы рассмотрим простейший случай — популяцию самоопылителей, полученную в результате скрещивания двух чистых линий P_1 и P_2 , отличающихся по аллелям одного локуса. У такой популяции в отсутствие отбора $p=q=1/2$ для любого поколения. Напомним, что F зависит от номера поколения самоопыления: для F_2 — $F=0$, для F_3 — $0,5$, для F_4 — $0,75$, ... для F_∞ — коэффициент инбридинга $F=1$.

Подставляя эти параметры в формулу (19), получаем следующую формулу, например, для F_2 : $\sigma_G^2 = 1/2a^2 + 1/4d$, для F_∞ $\sigma_G^2 = a^2$.

В общем случае расщепления по нескольким несцепленным генам без эпистаза вклады в генотипическую изменчивость их аддитивных a_i^2 и отдельно доминантных d_i^2 эффектов суммируются. В формулах генотипической дисперсии вместо a^2 появится Σa_i^2 (принято обозначать через D), а вместо d^2 — Σd_i^2 (обозначают H). Тогда формулы принимают вид, удобный для анализа (табл. 16).

16. Структура генотипической дисперсии (σ_G^2) в поколениях самоопыления без отбора при расщеплении по нескольким полиморфным локусам

F_1	F_2	F_3	F_4	... F_∞
0	$1/2D + 1/4H$	$3/4D + 3/16H$	$7/8D + 7/64H$... D

Примечание. D — это дисперсия аддитивных эффектов генов,
 H — дисперсия доминантных эффектов.

Фенотипическая дисперсия в каждом поколении состоит, естественно, из σ_G^2 , представленной в табл. 16, и средовой (σ_e^2). Далее предполагается, что эффекты генов и σ_e^2 мало меняются от условий.

Пусть цель отбора в поколениях самоопылителей состоит в выделении лучших чистых линий. Тогда успех отбора, его эффективность будет прежде всего зависеть от доли аддитивной изменчивости генов (D) в общей генотипической изменчивости (эффекты доминирования постепенно исчезают). Именно поэтому, с учетом табл. 16, в F_2 имеем:

$$h^2 = \frac{1/2D}{(1/2D + 1/4H + \sigma_e^2)} ; \quad H^2 = \frac{(1/2D + 1/4H)}{(1/2D + 1/4H + \sigma_e^2)} , \quad (20)$$

а в поколении для F_∞ : $h^2 = H^2 = D/(D + \sigma_e^2)$

Видно, как важно оценить D уже в первых поколениях, т.е. на ранних этапах селекционной программы. Для этой цели разработана серия сходных биометрико-генетических методов. Они отличаются необходимым набором популяций, данные по которым привлекаются для вычислений.

Например, достаточно знать величины σ_G^2 в двух поколениях — F_2 и F_3 , полученных самоопылением без отбора. Приравнивая эти величины к структурам σ_G^2 в F_2 и F_3 (табл. 16), получаем простую систему из двух линейных уравнений с двумя неизвестными — H и D . Решение не составляет труда. Но для оценки исходных σ_G^2 необходимо знать σ_p^2 в нескольких поколениях: P_1, P_2, F_1, F_2, F_3 . Тогда средовую дисперсию (σ_e^2) оцениваем по формуле (17). После этого, применяя формулу (18) для F_2 и F_3 , находим σ_G^2 в этих поколениях. Однако такая оценка σ_e^2 из-за различия генетической структуры поколений часто отличается от σ_e^2 в F_2, F_3 . Поэтому обычно применяют более простой и надежный метод.

Сумму σ_p^2 для двух беккроссов определяют по формуле

$$2\sigma_p^2(B_1) + \sigma_p^2(B_2) = \sigma_G^2(B_1) + \sigma_e^2 + \sigma_G^2(B_2) + \sigma_e^2 = 1/2D + 1/2H + 2\sigma_e^2.$$

Достаточно знать только эту величину и σ_e^2 в поколении F_2 . Тогда, с учетом структуры σ_G^2 ; для F_2 из табл. 16, имеем:

$$2\sigma_p^2(F_2) - [\sigma_p^2(B_1) + \sigma_p^2(B_2)] = 1/2D. \quad (21)$$

Например, по данным Х. Смита при скрещивании двух гомозиготных растений, относящихся к разным видам табака, были получены следующие оценки σ_p^2 для признака «высота венчика»: F_2 — 130,5; B_1 — 85,5; B_2 — 98,5. Из формулы (21) получаем: $1/2D = 77$, $D = 154$. Из формулы (20) в поколении F_2 : $h^2 = 77/130,5 = 0,59$. Отбор в F_2 ожидается эффективным.

Различия коэффициентов наследуемости в широком и узком смыслах наглядно видны в числовом примере, представленном в табл. 17.

17. Коэффициенты наследуемости h^2 и H^2 в поколениях F_2 и F_∞ ($h^2 = H^2$) при восьми соотношениях D, H и σ_e^2

Номер	D	H	σ_e^2	$h^2(F_2)$	$H^2(F_2)$	$h^2(F_\infty)$
1	2	0	8	0,11	0,11	0,20
2	2	2	8	0,10	0,16	0,20
3	4	0	8	0,20	0,20	0,33
4	4	4	8	0,18	0,27	0,33
5	8	0	8	0,33	0,33	0,50
6	8	8	8	0,29	0,43	0,50
7	12	0	8	0,43	0,43	0,60
8	12	12	8	0,35	0,53	0,60

Коэффициент h^2 для поколения F_2 следует понимать как отношение ожидаемого сдвига R_a среднего значения аддитивных эффектов генов для признака в F_3 к селекционному дифференциалу S , характеризующему отбор по фенотипическим проявлениям признака в F_2 . Коэффициент H^2 — как отношение сдвига R_G средней генотипической ценности растений в отобранной части популяции F_2 к величине S . Сдвиг R_G , естественно, не реализуется полностью в F_3 из-за расщепления, происходящего в отобранной части генотипов-самоопылителей. В тех случаях, когда требуется отобрать лучшие в F_2 генотипы и далее размножить их клонированием, эффективность отбора следует оценивать по величине H^2 .

Поскольку в F_∞ расщепления уже нет, то $H^2 = h^2$, а сдвиг R_a совпадает с R_G и со сдвигом R из формулы (13). Следует подчеркнуть, что эти сдвиги связаны с изменением частот генов и генотипов популяции, вызванным отбором.

Из табл. 17 видно, что отбор в F_2 будет эффективным только при седьмом соотношении дисперсий, когда $h^2(F_2) > 0,4$. В последних трех-четыре случаях есть смысл получить более поздние поколения и начинать отбор с них (в поздних поколениях вклад D в σ_G^2 увеличен, а вклад H снижен). При первых четырех соотношениях дисперсий отбор не будет эффективным и в F_∞ . Тогда следует повысить долю аддитивной изменчивости признака в популяции (применить более сложные схемы скрещиваний и отборов, возможно, с вовлечением других сортов) или попытаться снизить σ_e^2 (повысить точность опытов по оценке коэффициентов наследуемости и по самому отбору в исходной популяции).

При селекции по количественному признаку у самоопылителей весьма важно оценить, хотя бы приблизительно, число полиморфных локусов в популяции. Например, при дигенном наследовании число различных гомозигот в F_∞ не более $2^2 = 4$, а при шести, полиморфных локусах оно может достигнуть $2^6 = 64$. Объем селекционных опытов для выделения лучших чистых линий в этих двух ситуациях существенно различен.

Пусть P_1 и P_2 различаются по аллелям k независимо расщепляющихся локусов без эпистаза: P_1 содержит все плюс-аллели (повышающие величину признака), а P_2 все минус-аллели. Если аддитивные эффекты всех аллелей равны ($a_1 = a_2 = \dots = a_k = a$), то различие генотипических ценностей P_1 и P_2 будет $V = g(P_1) - g(P_2) = 2ka$, $V^2 = 4k^2a^2$.

$$\text{Так как } D = \sum_{i=1}^k a_i^2 = ka^2, \text{ то}$$

$$k = V^2 / 4D. \quad (22)$$

Этот способ оценки дает лишь грубое приближение к истинному числу полиморфных локусов, так как он основан на жестких ограничительных предположениях об эффектах генов.

3.4. Перекрестноопыляемые культуры

Если, по принятому предположению, целью отбора в расщепляющейся популяции самоопылителей являются чистые линии с максимальным генотипическим выражением признака, то у перекрестников задачи отбора обычно другие. В качестве такой задачи упрощенно будем рассматривать повышение генотипической ценности популяции (Q). Напомним, что эта характеристика связана с оценкой селекционной ценности генотипов (раздел 2.5.). Именно поэтому в числителе формулы (14), характеризующей эффективность отбора у перекрестников, стоит дисперсия селекционных ценностей генотипов, составляющих популяцию до отбора. Рассмотрим проблему оценки коэффициента наследуемости в узком смысле для перекрестников. Она связана с разделением общей фенотипической дисперсии признака на генотипическую и средовую, а также с выделением из генотипической дисперсии части, обусловленной изменчивостью селекционных ценностей генотипов.

Оценка генетической дисперсии. Структура генотипической дисперсии признака в равновесной популяции перекрестников при одном полиморфном локусе выводится из общей формулы (19) при $F = 0$ и конкретных значениях частот p и q

$$\sigma_G^2 = 2pq[a - (p - q)d]^2 + (2pqd)^2 \quad (23)$$

При перекрестном опылении без отбора частоты генотипов в поколениях постоянны, следовательно постоянна и σ_G^2 в отличие от σ_G^2 у самоопылителей (табл.16). Две части σ_G^2 в литературе имеют специальные названия: *генетическая дисперсия* – $\sigma_a^2 = 2pq[a - (p - q)d]^2$; *дисперсия доминирования* – $\sigma_d^2 = (2pqd)^2$.

Генетическую дисперсию не следует путать с дисперсией аддитивных эффектов генов. На σ_a^2 для перекрестников при $p \neq q$ влияет также эффект d . Из формулы (12) несложно вывести, что σ_a^2 равна дисперсии селекционной ценности генотипов популяции при $p' = p$, т.е. когда тестером служит сама родительская популяция.

Рассмотрим частные случаи соотношений a , d и изменение параметров популяции в результате отбора.

При аддитивной схеме наследования: $d = 0 = \sigma_d^2$; $\sigma_G^2 = \sigma_a^2 = 2pqa^2$; $h^2 = H^2$. В этих предположениях отбор далее будет невозможен ($h^2=0$), когда p достигнет 1 (Q достигнет максимума, равного a). Из этого примера ясно, что максимум Q не всегда можно считать целью отборов для перекрестников: $p=1$ соответствует полной элиминации гетерозигот, что часто приводит к нежелательным последствиям.

При $a=d$ (доминирование) $\sigma_a^2 = 8pq^3a^2$, $\sigma_d^2 = (2pqa)^2$. Отбор на повышение Q далее будет невозможен ($h^2=0$) так же, как в предыдущем случае, при $p=1$.

При $d>a$ (сверхдоминирование) σ_a^2 и, следовательно, h^2 обращается в 0 при $p = (a + d)/2d$, т.е. Q достигает максимума (см. раздел 2.2.). При сверхдоминировании хотя бы в одном локусе максимум Q не связан с элиминацией гетерозигот.

Приведем несложные способы оценки h^2 для перекрестников из опытов.

Успех селекции всегда выше при большем сходстве фенотипических величин среднего значения признака у пары родителей (X) и у их потомства (Y). Это сходство можно измерять с помощью ковариации между X и Y или коэффициента b регрессии Y на X . В частности, несложно доказать, что ковариация между средними $\bar{P} = \frac{1}{2}(P_1 + P_2)$ у многих пар родителей, случайно выбранных из родительской популяции, и средними у потомков (FS) таких пар дает оценку половины генетической дисперсии родительской популяции:

$$COV(\bar{P}, FS) = \frac{1}{2}\sigma_a^2.$$

Следует подчеркнуть, что эта ковариация, построенная по оценкам фенотипических ценностей родителей и потомков, не включает средовую дисперсию. Дело в том, что ковариация строится на основании достаточно большой выборки данных, т.е. средовые_ошибки усредняются.

Дисперсия средних \bar{P} в исходной популяции равна половине фенотипической дисперсии признака в ней:

$$\sigma^2(\bar{P}) = \frac{1}{2}\sigma_P^2.$$

Тогда ясно, что отношение ковариации к дисперсии, т.е. b — коэффициент регрессии FS на \bar{P} (рис. 8) равен коэффициенту наследуемости в узком смысле, представленному в формуле (14):

$$b(FS, \bar{P}) = COV(\bar{P}, FS) / \sigma^2(\bar{P}) = [\frac{1}{2}\sigma_a^2] / [\frac{1}{2}\sigma_P^2] = h^2. \quad (24)$$

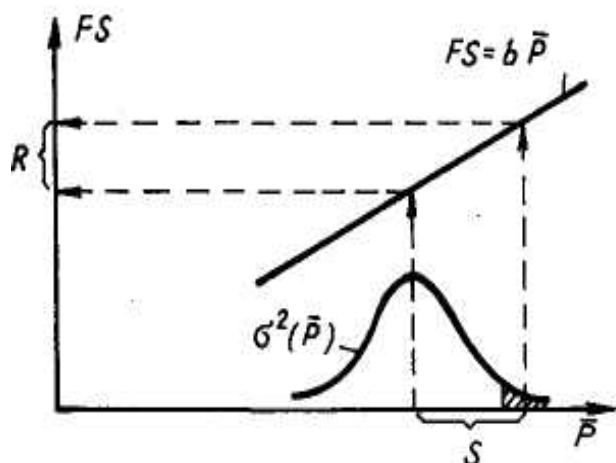


Рис. 8. Регрессия $FS = b \times \bar{P}$ средних значений признака у sibсов на средние их родителей: S - селекционный дифференциал при отборе пар по величине P в популяции родителей; R — ожидаемый эффект отбора, т.е. сдвиг среднего значения признака у потомков

То есть $b(FS, \bar{P}) = h^2$ можно получить непосредственно из опыта. Достаточно измерить средние \bar{P} у возможно большего числа пар случайно выбранных родителей, провести попарное скрещивание, получить потомство каждого такого скрещивания и у каждой семьи оценить среднее значение признака sibсов FS . Полученные результаты подставляют в формулу (24).

Другой способ основан на сопоставлении величин признака у материнских растений (M) и у их потомков (HS) при поликроссе, т.е. когда материнские формы опыляются случайной смесью пыльцы отцовской популяции (p' в отцовской популяции равно p). Можно доказать, что

$$COV(M, HS) = \frac{1}{2}\sigma_a^2; \quad b(HS, M) = \frac{[\frac{1}{2}\sigma_a^2]}{\sigma^2(M)} = \frac{1}{2}h^2.$$

Если предполагается вести массовый отбор материнских растений, то следует сделать замечание. В отличие от предположений, на которых ранее строились модели отбора, при описанной схеме скрещивания материнские растения опыляются отцовскими растениями, не подвергшимися отбору. Отбор материнских растений ведется после опыления. Именно поэтому его эффективность оценивается половиной коэффициента наследуемости, т.е. снижается в два раза. Во всех подобных ситуациях следует пользоваться не формулой (13), а другой: $R = 0,5h^2S$.

Описанные несложные способы оценки σ_a^2 и h^2 пригодны для практического применения и при полигенном наследовании признака, но без эпистаза.

Естественно, при выводе формул предполагались одинаковые условия выращивания родителей и потомков, что не всегда выполнимо. В связи с этим разработаны и испытаны специальные схемы скрещиваний («Северная Каролина I» и «Северная Каролина II»). Они дают возможность оценить генетическую дисперсию и дисперсию доминирования на основании измерений признака у растений за один год [3].

Разработаны также специальные схемы скрещиваний, которые позволяют получить оценки компонент генотипической дисперсии для гермафродитных растений. В случайной выборке растений (сортов, клонов и т.п.) выполняют все возможные комбинации скрещиваний, в том числе реципрокные и самоопыление. Это так называемая полная диаллельная схема, которая, наряду с ее модификациями, будет рассмотрена в главе 5.

Средняя степень доминирования. Знание степени доминирования, т.е. отношения d/a , имеет большое значение для селекции перекрестноопыляемых растений. Дело в том, что в популяции перекрестников часть растений обычно гетерозиготна. Для одного полиморфного локуса вклад гетерозигот в генотипическую ценность Q свободно скрещивающейся популяции равен $2pqd$ (см. раздел 2.2.). При $p = q = 1/2$ (доминирование) этот вклад определяющий так как

$Q = (p - q)a + 2pqd = d/2$. Если степень доминирования обозначить через $d' = d/a$, то могут реализоваться следующие ситуации:

- $d' < -1$ – сверхдоминирование минус-аллеля;
- $d' = -1$ – полное доминирование минус-аллеля;
- $-1 < d' < 0$ – неполное доминирование минус-аллеля;
- $d' = 0$ – аддитивная схема наследования;
- $0 < d' < 1$ – неполное доминирование плюс-аллеля;
- $d' = 1$ – полное доминирование плюс-аллеля;
- $d' > 1$ – сверхдоминирование плюс-аллеля.

Для признака с полигенной схемой наследования представляет интерес \bar{d}' — средняя (по полиморфным локусам) степень доминирования.

Ее модуль можно оценить по формуле

$$|\bar{d}'| = \sqrt{\left(\frac{2\sigma_d^2}{\sigma_a^2}\right)} \quad (25)$$

Эта оценка особенно удобна для применения, если справедливы следующие допущения:

- локусы расщепляются независимо;
- эпистаз отсутствует;
- для каждого из k локусов справедливо $p_i = q_i = 1/2$ (это случай расщепления потомства, полученного от скрещивания чистых линии P_1 и P_2);
- аддитивные эффекты в различных локусах близки ($a_1 = a_2 = \dots = a_k = a$) так же, как и доминантные эффекты ($d_1 = d_2 = \dots = d_k = d$)

Тогда справедливо:

$$\sigma_a^2 = 2 \sum p_i q_i [a_i - (p_i - q_i) d_i]^2 = \sum a_i^2 / 2 = D / 2,$$

$$\sigma_d^2 = 4 \sum p_i^2 q_i^2 d_i^2 = \sum d_i^2 / 4 = H / 4.$$

Таким образом,

$$2\sigma_d^2 / \sigma_a^2 = \sum d_i^2 / \sum a_i^2 = H / D.$$

Последнее допущение приводит к дальнейшему упрощению:

$$\sqrt{H/D} = \sqrt{\sum d_i^2 / \sum a_i^2} = \sqrt{k d^2 / k a^2} = \sqrt{(d/a)^2} = d' = |\bar{d}'|. \quad (26)$$

Другими словами, если получены оценки σ_a^2 и σ_d^2 то $\sqrt{2\sigma_d^2 / \sigma_a^2}$ дает оценку $|\bar{d}'|$. При невыполнении последнего из допущений эта оценка завышает истинное число \bar{d}' . Например, при

$k=2$, если $a_1 = a_2 = a$, $d_1 = a_1$, $d_2 = 3a_2$, то $|\bar{d}'| = \frac{(1+3)}{2} = 2$. Однако оценка по формуле (25) составит

$$\sqrt{(a_1^2 + 9a_2^2) / (a_1^2 + a_2^2)} = \sqrt{10a^2 / 2a^2} = \sqrt{5} = 2,24.$$

При $a_1 = a_2$, $d_1 = a_1$ и $d_2 = -a_2$ получаем $|\bar{d}'| = 0$, но $\sqrt{(a_1^2 + a_2^2) / (a_1^2 + a_2^2)} = 1$

Влияние различных схем отбора. До сих пор мы оценивали эффективность массового отбора в популяции. Во многих случаях такой отбор малоэффективен. Поэтому желательно проанализировать с помощью биометрико-генетических моделей более сложные схемы отбора. При этом будем исходить из ранее принятых предположений: отсутствие эпистаза в схеме наследования селективируемого признака, равновесное состояние популяции, отсутствие взаимодействия генотип-среда.

В зависимости от схемы селекции меняется структура фенотипической дисперсии (σ_p^2) признака между растениями или семьями, подвергающимися отбору, и доля генетической дисперсии ($\sigma_a^2 = \sigma_v^2$) влияющей на эффективность отбора. Соответственно различаются коэффициенты наследуемости $h^2 = \sigma_v^2 / \sigma_p^2$ и эффекты отбора R (табл. 18).

Простейшая схема — отбор лучших растений в популяции по фенотипической ценности признака. Как уже отмечалось, он может проводиться до или после опыления. Примером массового отбора до опыления может служить селекция двухлетних растений по признаку первого года жизни (масса корнеплода свеклы, кочана капусты и т.д.). После опыления отбор ведут по таким признакам однолетних растений, как масса зерна с растения у злаков и др. Преимущества массового отбора — простота и короткий цикл селекции. Недостаток — отбор основан на выделении отдельных растений по фенотипу, на который могут сильно влиять средовые отклонения.

18. Эффект отбора за один цикл селекции

для различных схем отбора перекрестноопыляющихся культур

Схема отбора	Эффект отбора R [см. формулу (13)]
Массовый отбор до опыления	$k\sigma_a^2/(\sigma_a^2 + \sigma_d^2 + \sigma_e^2)^{0,5}$
Массовый отбор после опыления	$0,5k\sigma_a^2/(\sigma_a^2 + \sigma_d^2 + \sigma_e^2)^{0,5}$
Семейный отбор полусибсов	$0,25k\sigma_a^2/(0,25\sigma_a^2 + \sigma_e^2)^{0,5}$
Семейный отбор полных сибсов	$0,5k\sigma_a^2/(0,5\sigma_a^2 + 0,25\sigma_d^2 + \sigma_e^2)^{0,5}$
Инбридинг $S1$	$k\sigma_a^2/(\sigma_a^2 + 0,25\sigma_d^2 + \sigma_e^2)^{0,5}$
Инбридинг $S2$	$1,5k\sigma_a^2/(1,5\sigma_a^2 + 0,19\sigma_d^2 + \sigma_e^2)^{0,5}$

Примечание. k — стандартный селекционный дифференциал.

Эффективность отбора можно повысить, используя семьи растений как селекционные единицы (вместо отдельных растений). Примером служит отбор семей полусибсов или полных сибсов. Эти методы отбора обычно требуют больше времени на цикл, чем массовый отбор (2—3 поколения). При отборе семей полусибсов используют семена от свободного опыления отдельных растений популяции. Семьи полных сибсов получают от скрещивания между парами растений в пределах популяции. Испытывают высеянные отдельно семьи полу- или полных сибсов.

Возможны различные модификации семейного отбора. Например, в селекции перекрестников часто используют инбридинг для получения константных линий при создании гетерозисных гибридов. Формулы для прогноза эффективности отбора между потомками $S1$ и $S2$ (1 и 2 поколение самоопыления) также приведены в табл. 18. Отметим, что в данных формулах средовая дисперсия неодинакова. При массовом отборе σ_e^2 относится к величине признака одного растения, а при отборе семей — к среднему по всем растениям семьи σ_e^2 . В последнем случае дисперсия меньше.

Сравнение эффективности методов отбора с целью выбора оптимального для конкретной культуры, популяции, размера опытного поля и т.п. является сложной задачей. Д.С.Фалконер считает, что семейному отбору следует отдавать предпочтение при низкой наследуемости отбираемого признака, низкой изменчивости, обусловленной общей для семьи средой, при большом числе особей в семье. В этом случае средовые отклонения признака у особей при усреднении значений признака в семье взаимно погашаются и фенотипическое среднее для семьи оказывается близким к генотипической ценности.

В качестве примера сравнения различных схем отбора приведем оценки эффективности селекции кукурузы тремя методами:

- массовый отбор (M) после опыления, цикл — один год;
- семейный отбор полусибсов (HS), цикл — два года;
- семейный отбор полных сибсов (FS), цикл — два года.

Пусть компоненты дисперсии количественного признака, по которому ведут отбор, имеют следующие значения: $\sigma_a^2 = 1$, $\sigma_d^2 = 0,5$, $\sigma_e^2 = 0,5$, $\sigma_e'^2 = 0,1$. Стандартный селекционный дифференциал одинаков ($k=2,06$).

Эффект отбора за один цикл

$$R_M = 0,5 \cdot 2,06 \cdot 1 / (1 + 0,5 + 0,5)^{0,5} = 0,728;$$

$$R_{HS} = 0,25 \cdot 2,06 \cdot 1 / (0,25 \cdot 1 + 0,1)^{0,5} = 0,873;$$

$$R_{FS} = 0,5 \cdot 2,06 \cdot 1 / (0,5 \cdot 1 + 0,25 \cdot 0,5 + 0,1)^{0,5} = 1,210.$$

Рассматриваемые схемы отличаются по длительности цикла отбора. Поэтому естественно сравнить не R , а R' — эффекты отбора на одно поколение:

$$R'_M = R_M = 0,728; R'_{HS} = R_{HS}/2 = 0,436; R'_{FS} = R_{FS}/2 = 0,605.$$

Таким образом, при заданном значении дисперсий количественного признака в родительской популяции кукурузы наиболее эффективным из трех схем будет массовый отбор.

КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ

1. Чем отличаются коэффициенты наследуемости в широком (H^2) и узком смыслах (h^2)?
2. Почему эффективность отбора в расщепляющейся популяции оценивают по h^2 , а при отборе лучших константных генотипов по H^2 ?
3. Каковы недостатки этих параметров при прогнозировании эффективности отбора?
4. Доказать, что если часть минус-аллелей содержит родитель P_1 (при сохранении остальных предположений), то формула (22) дает заниженную оценку числа полиморфных локусов в популяции.
5. Какие факторы влияют на величину генетической, генотипической и средовой дисперсий признака в популяции?
6. Доказать, что генетическая дисперсия в популяции перекрестников равна дисперсии селекционных ценностей [формула (12)] ее генотипов при выборе в качестве тестера самой родительской популяции.
7. Оценить генотипическую и генетическую дисперсии для популяции с параметрами, приведенными в задании 9 к главе 2.
8. Охарактеризовать влияние различных схем отбора у перекрестноопыляемых культур на его эффективность.