

А.В. Смиряев

*Посвящается 150-летию
Российского государственного аграрного университета
- МСХА имени К.А. Тимирязева*

**ГЕНЕТИЧЕСКАЯ ДИВЕРГЕНЦИЯ
РОДИТЕЛЬСКИХ ФОРМ И
НАСЛЕДСТВЕННАЯ ИЗМЕНЧИВОСТЬ
ПОТОМСТВА
Биометрико-генетический анализ**

Монография

Москва
Издательство РГАУ-МСХА
2016

УДК 631.523.4
ББК 41.31
С 507

Рецензенты: д.б.н., профессор, член-корреспондент НАН Беларуси **А.В. Кильчевский** (Президиум НАН Беларуси); д.с.-х.н., профессор **А.В. Исачкин** (РГАУ-МСХА имени К.А. Тимирязева)

Смирязев А.В. Генетическая дивергенция родительских форм и наследственная изменчивость потомства. Биометрико-генетический анализ. Издание 2-е исправленное и дополненное / А.В. Смирязев. М.: Издательство РГАУ-МСХА, 2016. 134 с.

В монографии изложены итоги многолетних разработок биометрико-генетических методик для прогноза наследственной изменчивости потомства на основе оценок генетического несходства родительских форм. Проанализирована адекватность различных моделей для подбора родительских пар в селекции на гетерозис. Предложены новые метрики и способы оценки генетического несходства гомозиготных родительских форм. Изучена изменчивость схем наследования расчетных количественных признаков в связи с эффективностью отбора в популяциях потомства. Биометрико-генетические методики проиллюстрированы на данных полевых экспериментов с озимой и яровой пшеницей, кукурузой и ячменем.

Предназначено для научных сотрудников, аспирантов, преподавателей ВУЗов биологического и сельскохозяйственного профиля.

© Смирязев А.В., 2016
© Издательство РГАУ– МСХА, 2016

СОДЕРЖАНИЕ

Введение	5
Глава 1. Изменчивость биометрико-генетических параметров и прогноз количественных признаков у гибридов F_1	6
1.1. Результаты анализа изменчивости биометрико-генетических параметров в диаллельных скрещиваниях.....	7
1.2. Регрессионные модели для прогноза признаков у гибридов F_1	12
Результаты сравнения адекватности регрессионных моделей.....	13
1.3. Причины адекватности регрессии на ОКС: гипотеза об аллельном несходстве родительских форм.....	21
Оценка влияния отклонений от предположений и случайных ошибок опытов.....	26
1.4. Прогноз гетерозиса с помощью оценок генетической дивергенции родителей.....	32
Заключение.....	35
Глава 2. Анализ сходства генотипов по эколого-онтогенетической изменчивости ростового признака	37
2.1. Гипотеза о сходстве относительной изменчивости признака.....	40
2.2. Параметры и количественный анализ степени сходства генотипов.....	46
2.3. Проверка гипотезы – анализ сходства сортов озимой пшеницы.....	52
Методика получения данных.....	53
Регрессионный и кластерный анализ сходства сортов.....	54
Сходство формы реакции в онтогенезе и по конечным значениям высоты растений.....	61
Сходство аллельного состава и относительной изменчивости в онтогенезе.....	62
Синхронность прохождения фенофаз и сходство относительной изменчивости.....	64
Заключение.....	66
Глава 3. Наследование несходства в диаллельных скрещиваниях	68
3.1. Модель и параметры наследования парного несходства.....	69
Графический анализ парного несходства.....	73
Дополнительные генетико-статистические параметры и связи.....	76
Оценка гетерозиготности гибридов в диаллельном скрещивании.....	77
3.2. Биометрико-генетический анализ несходства генотипов в диаллельных скрещиваниях пшеницы.....	80

Материалы и параметры: несходство по форме реакции.....	80
Результаты анализа по несходству формы реакции.....	81
Пример анализа несходства по модулю разности признака.....	89
Заключение.....	92
Глава 4. Прогноз эффективности отбора в популяциях потомства по оценкам дивергенции родителей.....	94
4.1. Проблема количественного сравнения родительских пар по генетическому несходству.....	94
Метрики для экспериментальной оценки дивергенции.....	96
Экспериментальный материал.....	97
Результаты и обсуждение: изменчивость корреляций.....	98
Причины изменчивости первого типа.....	101
Причины изменчивости второго типа.....	103
Причины изменчивости третьего типа.....	108
4.2. Моделирование наследственной изменчивости расчетных признаков.....	109
Структура и генетический полиморфизм исходных и расчетных признаков.....	110
Перестройка схем наследования признаков.....	114
Один исходный – один расчетный признак.....	116
Несколько исходных – один расчетный признак.....	117
Несколько расчетных признаков.....	120
Наследственная изменчивость популяционных параметров.....	122
Заключение.....	125
Библиографический список.....	127

Введение

Интенсификация растениеводства и селекционной работы повышают роль исследований генетического потенциала растений в онтогенезе и при их выращивании в изменчивых агроэкологических условиях (в экологических градиентах). В последние десятилетия для подбора родительских сортов, нацеленного на увеличения полезной изменчивости по количественным признакам у потомства, используют различные оценки генетической дивергенции родителей, т.е. несходства их нуклеотидного или аллельного состава. При этом возникает ряд важных вопросов.

Какие из многочисленных метрик – оценок несходства обеспечат требуемую наследственную изменчивость потомства? Чем вызвана низкая адекватность большинства моделей, предназначенных для прогноза наследственной изменчивости гибридов F_1 по биометрико-генетическим параметрам родительских форм? Действительно ли после скрещивания, например, сортов-самоопылителей с большими значениями метрики парного несходства увеличивается вероятность получить гетерозисные гибриды, а в поздних поколениях расщепления – больше трансгрессивных форм по многим количественным признакам?

Современные методы подбора пар для скрещивания с использованием молекулярно-генетические маркеров пока не гарантируют объективную оценку несходства по генам, определяющим хозяйственно ценные признаки. Поэтому необходимо также совершенствовать методы оценки несходства родительских форм по их количественным признакам. Изложенные в монографии биометрико-генетические методики являются результатом многолетних исследований автора в этом направлении.

Для понимания изложенного материала читатель должен быть знаком с традиционными биометрико-генетическими моделями (олигогенными, полигенными и феноменологическими), а также с основанными на них методами, используемыми, в частности, для повышения эффективности селекции растений (Мазер, Джинкс, 1985; Смирязев и др., 1992).

Автор выражает благодарность сотрудникам и аспирантам РГАУ-МСХА им. К.А. Тимирязева за участие в получении полевых экспериментальных данных, необходимых для проверки эффективности биометрико-генетических методик, изложенных в монографии. Автор также признателен сотруднице ИГЦ НАН Беларуси Л.А. Тарутиной за данные диаллельного скрещивания кукурузы, предоставленные для биометрико-генетического анализа.

Глава 1. Изменчивость биометрико-генетических параметров и прогноз количественных признаков у гибридов F_1

На данных диаллельных скрещиваний ячменя, кукурузы, а также озимой пшеницы выявлено сильное варьирование параметров и схем наследования количественных признаков в зависимости от условий выращивания. Сравняются по адекватности различные регрессионные модели для прогнозирования величин признаков у гибридов F_1 в диаллельных скрещиваниях. Проведен анализ особенностей схем наследования признаков – условий, обеспечивающих адекватность некоторых моделей в широком диапазоне изменчивости биометрико-генетических параметров. С помощью аналитического и компьютерного моделирования оценена чувствительность этих моделей к отклонениям от условий адекватности и случайным ошибкам полевых экспериментов. Исследована возможность прогнозирования эффектов гетерозиса в зависимости от генетического несходства родительских форм.

В последние десятилетия в России и за рубежом интенсивно изучается изменчивость количественных признаков и биометрико-генетических параметров растений в зависимости от экологических условий, а также в процессе онтогенеза (Драгавцев, Литун и др., 1984; Драгавцев, 2003). Эта изменчивость значительно сильнее проявляется у растений по сравнению с животными. В отличие от животных растения отвечают на стрессы резким изменением фенотипа, т.к. они имеют значительно меньше возможностей и механизмов гомеостаза. Можно предложить две основные не альтернативные гипотезы для объяснения генетических механизмов этой изменчивости.

Первая – в разных экологических условиях и на разных этапах роста на любой количественный признак действуют разные полиморфные гены, т.е. меняется спектр генов, влияющих на изменчивость признака, происходит переопределение спектра генов (Драгавцев, 1983). Говоря упрощенно, если возникла эпифитотия какого-либо заболевания, то различие, например, урожайности сортов набора в основном определяет полиморфизм по генам устойчивости к этой болезни. Если же проявилась засуха – по генам засухоустойчивости. Вторая гипотеза состоит в том, что полное переопределение спектра – это редкое явление. Обычно меняется лишь активность (экспрессия) одних и тех же полиморфных генов (Смиряев, 1987).

Следует отметить, что гены могут нелинейно трансформировать свое влияние на величины признаков через сложные физиологические

механизмы формирования признака, причем даже после полного переопределения спектра генов в онтогенезе накопленная ранее изменчивость признака в наборе генотипов сохраняется.

Все это затрудняет биометрико-генетический анализ и прогнозирование изменчивости признаков в потомстве при подборе родительских форм.

1.1. Результаты анализа изменчивости биометрико-генетических параметров в диаллельных скрещиваниях

Несмотря на трудоемкость диаллельные скрещивания многие десятилетия используются в генетических и селекционных исследованиях. Одна из причин – возможность получать достаточно подробные сведения о схемах наследования количественных признаков на основе данных лишь о родительских генотипах и их гибридах F_1 (Смиряев, 1987). Для иллюстрации высокой степени изменчивости признаков и биометрико-генетических параметров приведем примеры анализа результатов специальных экспериментов с наборами сортообразцов ячменя (Пыльнев, Смиряев, 1995), кукурузы (Турбин и др., 1974), а также озимой пшеницы (Силис, Шмакова, 1986). Отметим, что в научной литературе довольно широко освещены результаты диаллельного анализа количественных признаков в экологических градиентах (Горшкова, 1988; Федулова, Анисков, 1990), но не в онтогенезе. Используя методы Гриффинга и Хеймана, можно получить новую информацию об изменчивости генетического контроля количественных признаков.

В эксперименте с ячменем анализировали 6 родительских форм и их гибридов F_1 . Среди родительских форм представлены районированные сорта (Московский 3, Зазерский 85), а также линии и образцы из каталога ВИР им. Н.И. Вавилова (К-28662, К-8695, Л-60/91, Л-131/91) разного происхождения. Родительские формы существенно различаются по морфологическим, физиологическим и количественным признакам.

Опыты проводили по диаллельной схеме 6×6 без реципрокных скрещиваний. В 1993 г. этот опыт был заложен в НПО «Подмосковье» в 3 повторностях сотрудниками ИОГен под руководством проф. В.А. Пухальского. Анализ изменчивости ростовых признаков у родительских форм и гибридов F_1 в онтогенезе проводились сотрудниками РГАУ-МСХА. В 1994 г. опыт с теми же родительскими формами и гибридами проводился на селекционной станции РГАУ-МСХА в 2 повторностях с двумя сроками сева. Первый срок (IV декада апреля) оптимален для

ячменя, второй (10 мая) – поздний. Метеорологические условия в годы исследований существенно различались.

Анализировали такие ростовые признаки, которые можно измерять несколько раз за период вегетации, не повреждая растения: высоту растений и площадь флагового листа. Измерения высоты проводили каждые 6 дней, флагового листа – каждые 5 дней. В результате появилась возможность анализировать изменчивость параметров и схем наследования ростовых признаков по двум сочетаниям год-место и двум срокам сева, а также на разных этапах роста и развития ячменя. Математическая обработка данных по каждому признаку и сроку измерения осуществлялась методом Гриффинга (2-й метод модель 1), а также параметрическим и графическим анализом по Хейману (1958) и Джонсу (1965).

До сих пор нет прямого доказательства тесной связи спектра генов, определяющих высоту растений и площадь флагового листа, с фазами онтогенеза у ячменя. Поэтому, учитывая возрастание их ростовых признаков у родительских форм и гибридов F_1 в процессе онтогенеза, изучалась динамика биометрико-генетических параметров в процессе онтогенеза в зависимости именно от X_c – средних значений соответствующего признака в каждом опыте.

Анализ признака «высота растений» показывает, что характер изменения и ранги общей комбинационной способности ОКС одного и того же сорта в разные годы и в разные сроки не имеет качественных различий (рис. 1.1). Однако величины ОКС по сочетаниям год-место и срокам сева для некоторых родительских форм различаются в 1,5-2 раза. Существенно различаются ранги ОКС по сортообразцам. Так, сорт Зазерский 85 и линия К-8695 обладают максимальными по модулю эффектом ОКС, которые увеличиваются параллельно с ростом растений. Но у линии К-8695 ОКС положительна, а у сорта Зазерский 85 – отрицательна. Остальные родительские формы характеризуются средней ОКС, практически не меняющейся на протяжении развития растений.

Установлено также, что у ячменя в значительно большей степени, чем эффекты ОКС, меняются по сочетаниям год-место и срокам сева значения и ранги специфической комбинационной способности (СКС). Это типичный вывод о несходной изменчивости ОКС и СКС различных культур в экологических градиентах.

В качестве примера сопоставим изменчивость ОКС и СКС признака «длина початка» (см) у семи инцухт-линий кукурузы. 21 прямой гибрид F_1 этих линий испытывали в опытах с тремя рандомизированными повторениями (блоками), проведенных в двух пунктах (Турбин и др., 1974). Ранги значений ОКС (рис. 1.2а)

меняются по 2 пунктам испытания значительно меньше, чем СКС (рис. 1.2б).

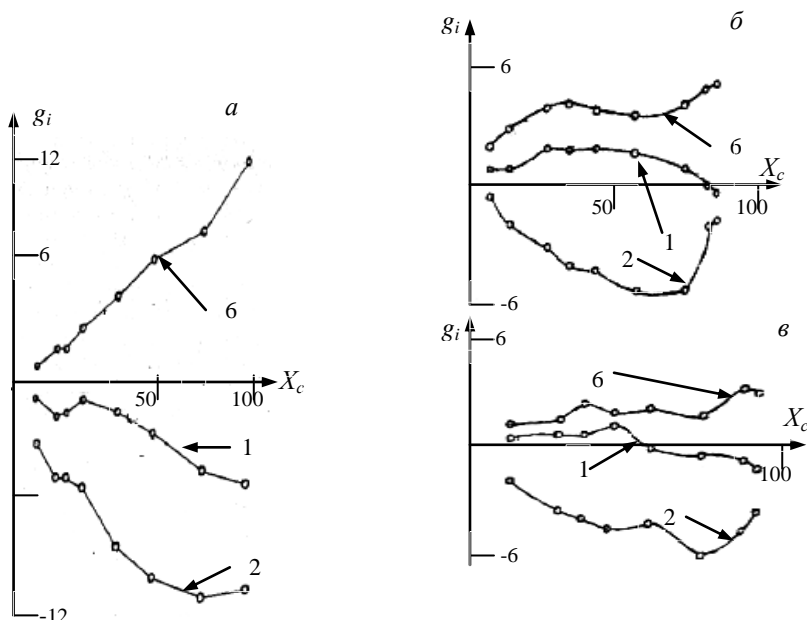


Рисунок 1.1. Эффекты ОКС по высоте растения у трех из шести формообразцов (1 – Московский 3, 2 – Зазерский 85, 6 – К-8695) в онтогенезе. g_i – эффекты ОКС (см); X_c (см) – среднее значение признака по всем сортам (линиям).

a – динамика 1993 г., *б* и *в* – динамика 1994 г. соответственно по 1-му и 2-му срокам сева.

Отметим, что эффекты СКС, как считается, зависят от доминирования и эпистаза в схеме наследования признака, а ОКС – от аддитивных эффектов генов (Турбин и др., 1974). Представление о большей стабильности аддитивных эффектов генов (ОКС) по сравнению с доминированием и эпистазом (СКС) важно не только в плане теоретической генетики количественных признаков, но и для практической селекции. Дело в том, что в линейной селекции важны именно аддитивные эффекты, а в гетерозисной – неаддитивные (Смиряев и др., 1992).

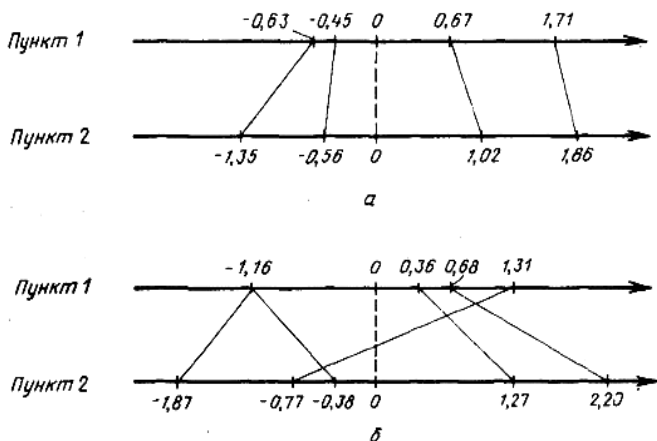


Рисунок 1.2. Варьирование значений некоторых эффектов ОКС (а) и СКС (б) для признака «длина початка» при изменении условий испытания (пункты 1 и 2)

При анализе по Хейману данных по высоте растений у ячменя обнаружилось, что в первой половине проведения измерений каждого из двух лет проявляется эпистаз – т.е. неаллельные взаимодействия генов, что не дает возможности корректно применить этот метод. Если же убрать из диаллельной схемы родительскую форму К-286621 и ее гибриды F_1 , то для оставшихся форм коэффициенты регрессии между ковариансами гибридов (W_r) и дисперсиями признака в потомстве (V_r) значимо отличаются от нуля и не отличаются от единицы при всех сроках измерения. Это доказывает отсутствие неаллельных взаимодействий в схеме наследования высоты растений для остальной диаллельной схемы и повышает объективность анализа по Хейману.

По высоте растений у ячменя соотношение между аддитивными и доминантными эффектами генов сильно варьирует по датам измерений, срокам сева и сочетаниям год-место.

Аналогичное варьирование схемы наследования наблюдалось, например, при анализе данных диаллельных скрещиваний семи сортов озимой мягкой пшеницы в трех условиях выращивания (Силис, Шмакова, 1986). Изменчивость генетической детерминации признака «масса 1000 зерен» как для отдельных сортов, так и в среднем по набору отражают графики регрессии W_r на V_r в трех средах испытания (рис. 1.3 а, б, в). Угол наклона прямых линий, т.е. коэффициент b на графиках Хеймана, в каждой среде незначимо отличается от единицы и значимо – от нуля. Следовательно, схема

наследования аддитивно-доминантная. В условиях *a* и *б* доминирование близко к полному (график пересекает ось W_r вблизи нуля), а в условиях *в* проявилось сверхдоминирование (пересечение ниже нуля).

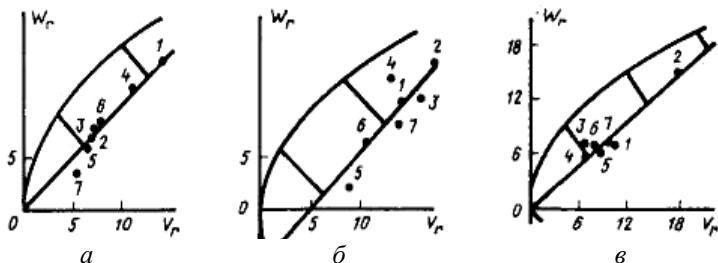


Рисунок 1.3. Графики регрессий W_r на V_r по признаку «масса 1000 зерен» для трех условий выращивания семи сортов озимой мягкой пшеницы:

a – Москва, 1983 г.; *б* – Москва, 1984 г.; *в* – Дербент, 1982 г.

Сорта: 1 – Абе, 2 – Одесская полукарликовая, 3 – Сейм, 4 – Прикульская 36, 5 – Заря, 6 – Чайка, 7 – Одинцовская (по Д.Я. Силис и Т.В. Шмакова, 1986)

В отличие от высоты растения эффекты ОКС по площади флагового листа ячменя мало меняются на протяжении роста и развития ячменя. Значения ОКС некоторых сортов (линий) контрастны, но их соотношение у сортов довольно стабильно. У гибридов от скрещивания 6-й родительской формы с другими отмечено сравнительно большее значение ОКС по площади флагового листа. Наблюдается изменение ОКС в зависимости от места, года и срока сева, причем у некоторых родительских форм это проявляется в большей мере, чем у других.

При анализе данных по методу Хеймана почти на всех стадиях измерений и в оба года выявлено неполное доминирование. Исключением были первый срок измерения при 1-м сроке сева (сверхдоминирование) и второй срок измерения при 2-м сроке сева (полное доминирование).

Объяснить и, тем более, прогнозировать варьирование параметров и схем наследования признаков, обнаруженных в рассмотренных диаллельных скрещиваниях ячменя, кукурузы и озимой пшеницы, довольно сложно из-за отсутствия общей теории изменчивости детерминации количественных признаков.

1.2. Регрессионные модели для прогноза признаков у гибридов F₁

Проблема подбора родительских форм и условий выращивания для получения высокогетерозисных гибридов F₁ в настоящее время не имеет строгого решения. Основная причина – сложный и изменчивый характер средовой и генетической детерминации количественных признаков растений. Это вызывает необходимость проводить большой объем скрещиваний, испытывать сорта вместе с гибридами в различных условиях. Существуют разрозненные сведения о возможности прогнозирования уровней хозяйственно ценных признаков гибридов F₁ по эколого-генетическим параметрам родительских форм (Драгавцев, 1983; Хасан, 1989). В частности, проверены различные регрессионные модели и корреляционные связи величин признака у гибридов с родителями, их комбинационными способностями или другими характеристиками (Kang, Board, et. al, 2005; Kazunori, et. al, 2006; Han, et. al, 2006; Hashemi, et. al, 2008; Quero-Garcna, et. al, 2009). Большинство таких моделей недостаточно адекватны для получения надежных теоретических выводов и селекционных рекомендаций.

Автором была предпринята попытка выявить и использовать для сокращения объемов испытаний статистические регрессионные связи величин признака у гибридов F₁ с эффектами общей комбинационной способности родительских форм или непосредственно с величинами того же признака у родителей. Предполагается, что родительские формы и гибриды выращиваются в серии опытов (годы, места, агроприемы и т.п.).

Первая модель:

$$x_{ijk}=M+l_k+C_i\times g_{jk}+C_j\times g_{ik}+e \quad (1.1),$$

где x_{ijk} – величина признака у гибрида F₁, полученного от скрещивания i -го и j -го родителей (если $i=j$, то x_{iik} – признак i -й родительской формы) в k -й среде – опыте совместного испытания родителей и гибридов; g_{ik} , g_{jk} – эффекты общей комбинационной способности (ОКС) i -й и j -й родительской формы в опыте с номером k ; M – общее среднее значение признака по всей серии опытов, l_k – отклонение значения признака по всей диаллельной схеме в опыте с номером k от M ; e – случайные ошибки опытов в повторностях с дисперсией, обозначаемой далее σ_e^2/n , где n – число повторений в опытах.

Константы C_i , C_j – влияние i -й и j -й родительских форм, подбираемые в каждой регрессионной модели методом наименьших квадратов (МНК). То есть эти константы – коэффициенты регрессии

обеспечивают минимум среднего квадрата отклонений оценок признака по регрессионной модели для гибридов и родительских форм от реальных данных (x_{ijk}), полученных во всей серии опытов.

Вторая модель:

$$x_{ijk} = M + (1 + 0,5 C_i) g_j + (1 + 0,5 C_j) g_i + (1 + C_i)(1 + C_j) l_k + e \quad (1.2).$$

В отличие от модели (1.1) в модели (1.2) участвуют только средние по опытам серии (по k) эффекты ОКС (g_i, g_j).

Каждая из этих моделей далее рассматривается в двух вариантах – с включением и без включения данных по родительским формам (только данные по гибридам F_1). Для простоты предполагают отсутствие реципрокных эффектов.

Испытывали также ряд моделей, где величины признаков гибридов связывали с признаками у их родительских форм. Лучшими при анализе опытных данных оказались следующие две.

Первая:

$$x_{ijk} = M + l_k + (C_i + C_j)(A_{ijk} - x_k) + e \quad (1.3),$$

где $A_{ijk} = (x_{iik} + x_{jjk})/2$, x_{iik} – величина признака i -го родителя, x_{jjk} – j -го; x_k – средняя величина признака всех родителей в k -м опыте серии.

Вторая модель:

$$x_{ijk} = (C_i + C_j) [D \times \min(x_{iik}, x_{jjk}) + (1-D) \times \max(x_{iik}, x_{jjk})] + e \quad (1.4),$$

где $\min(x_{iik}, x_{jjk})$ и $\max(x_{iik}, x_{jjk})$ – минимальное и максимальное значения признака у двух родителей; константе D придавали различные значения, в частности 1; 0,5; 0.

Проверка адекватности каждой модели велась с использованием результаты опытов – диаллельных скрещиваний кукурузы (Смиряев, Таругина, 1991) и озимой пшеницы (Смиряев, 1990) с помощью критерия Фишера, что ниже будет объяснено подробно.

Результаты сравнения адекватности регрессионных моделей

Для проверки степени адекватности моделей по кукурузе использовали данные скрещиваний шести инцухт-линий (Л21, Л20, 024а, S27, Л15 и О18а), полученные сотрудниками Института генетики и цитологии Белоруссии. Опыты проводились в пяти повторностях четыре года (1974–1977 гг.) в Белоруссии, и три года (1974–1976 гг.) в

Чехословакии. Высевались родительские линии и все их прямые гибриды F_1 .

Использованы данные по пяти признакам, приведенным в табл. 1.1. Схемы наследования рассмотренных количественных признаков были существенно неаддитивными, т.е. кроме аддитивных эффектов генов в схемах наследования присутствует доминирование и, в некоторых случаях, эпистаз. Величины каждого признака и эффекты ОКС одних и тех же форм существенно менялись от места и года выращивания. В рамках каждой регрессионной модели обобщались наборы данных по одному признаку за все годы либо только из Белоруссии, либо из Чехословакии, либо из этих двух мест одновременно. Последнее требует пояснения.

Оценки коэффициентов C в регрессиях (1.1) и (1.2), показавших лучшие результаты, получали по каждому набору данных взвешенным методом наименьших квадратов (МНК). То есть при вычислении проводилось упрощенное «взвешивание» наименьшего квадрата, т. е. учитывалась только различная точность опытов по годам и местам. Поэтому, если в рамках одной модели обобщались опыты с существенно различными дисперсиями случайных ошибок в повторностях (σ_e^2/n), то данные со значительно большей ошибкой практически «терялись» на фоне остальных и не оказывали влияния на результаты анализа.

Был выбран предел, соответствующий десятикратному различию дисперсий σ_e^2/n обобщаемых опытов. Если в обобщаемом наборе данных из Белоруссии и Чехословакии оказывались хотя бы два опыта с дисперсиями σ_e^2/n , различающиеся более чем в десять раз, то обработка велась только по отдельности – для Белоруссии (за четыре года) и для Чехословакии (за три года). В результате, для признаков «масса верхнего развитого початка», «длина початка» и «ширина початка у основания» анализ моделей велся только по отдельности, для остальных двух признаков – как по отдельности, так и вместе для Белоруссии и Чехословакии. При проверке адекватности моделей учитывались также ошибки оценок $l_k, g_{ik}, g_{jk}, g_i, g_j$, стоящих в правых частях регрессий.

Для проверки адекватности любой модели ее взвешенный (с учетом разных ошибок опытов для разных лет и мест) наименьший квадрат отклонений (ms_0) сравнивали по критерию Фишера со средним квадратом случайных ошибок (ms_e): $F_p = ms_0/ms_e$. Затем, как обычно, сравнивали F_p с F_τ при уровне значимости 0,05.

Для модели (1.1) в варианте «с родительскими линиями» число степеней свободы отклонений ms_0 будет $df_1 = p(p+1)L/2 - Lp - p$, «без линий» – $df_1 = p(p-1)L/2 - Lp - p$. Для модели (1.2) «с линиями» – $df_1 = p(p-$

1) $L/2-2p$ «без линий» – $df_1=p(p+1)L/2-2p$. Здесь L – число опытов серии, обобщаемых в одной модели. «Взвешивание» ошибок с учетом различной точности отдельных опытов касалось не только параметра ms_e , но и его степеней свободы df_2 (Welch, 1974). Если $F_p < F_T$ (т.е. $F_p/F_T < 1$), то модель считалась адекватной данным, т.е. отклонения оценок по модели от реальных данных обусловлены только случайными ошибками опытов.

Таблица 1.1. F_p/F_T для различных наборов данных, обобщаемых в регрессионных моделях

Признак	Модель (1.1)		Модель (1.2)	
	с линиями	без линий	с линиями	без линий
Белоруссия, серия из 4-х опытов: 1974, 1975, 1976, 1977 гг.				
масса верхнего развитого початка	3,86	0,89	5,65	1,54
длина початка	4,62	0,84	6,81	1,82
ширина початка у основания	4,13	0,83	5,96	1,17
число рядков зерен на початке	1,50	0,99	2,29	2,16
число зерен в ряду початка	5,41	1,08	8,74	2,03
Чехословакия, три опыта – 1974–1976 гг.				
масса верхнего развитого початка	4,30	1,56	12,48	7,01
длина початка	5,40	1,12	28,08	8,55
ширина початка у основания	9,28	1,76	13,52	2,10
число рядков зерен на початке	2,50	1,04	3,89	2,76
число зерен в ряду початка	12,93	0,94	21,72	3,07
Серия из семи опытов: Белоруссия, 1974–1977 гг. и Чехословакия, 1974–1976 гг.				
число рядков зерен на початке	2,28	1,11	3,78	2,75
число зерен в ряду початка	10,44	1,3	15,64	2,99

Результаты проверки адекватности моделей (1.1) и (1.2) в вариантах «с линиями» и «без линий» по различным наборам данных

приведены в табл. 1.1. Модель (1.1) «без линий» в большинстве расчетов показывает близость к единице величин F_p/F_T . Остальные варианты расчетов, отраженные в табл. 1.1, нельзя признать адекватными. Худшее соответствие моделям показали оба варианта «с линиями», т.е. включение в модель наряду с гибридами данных по родительским линиям значительно снижает соответствие модели данным.

Модель (1.2) оказалась значительно хуже более простой по структуре модели (1.1). В двух-трех случаях отношение F_p/F_T для модели (1.1) «без линий» больше единицы. Однако следует напомнить, что во всех опытах высевались только прямые гибриды F_1 . Следовательно, во всех моделях в равной степени не учтены неизвестные ошибки, связанные с реципрочными эффектами. Без учета этих эффектов величины ms_e по существу занижены, что вызывает завышение значений F_p/F_T .

Эти отношения в модели (1.2) заметно больше единицы для наборов данных, куда входят опыты, проведенные в Чехословакии. Анализ величин случайных опытов из Чехословакии показывает, что дисперсия σ_e^2/n там была значительно ниже, чем в опытах из Белоруссии. В результате, вклад неучтенных реципрочных различий гибридов F_1 в общие ошибки моделей (ms_0) становится значительно более заметным, чем по данным только из Белоруссии. По-видимому, это вызывает завышение величин F_p/F_T при анализе опытов из Чехословакии.

Таблица 1.2. F_p/F_T для различных наборов данных в модели (1.3)

Признак	Белоруссия 1974–1977 гг.	Чехословакия 1974–1976 гг.	Белоруссия и Чехословакия
масса верхнего развитого початка	1,14	3,18	1,20
длина початка	1,27	2,67	1,52
ширина початка у основания	1,50	1,88	1,80,
число рядков зерен на початке	1,67	1,75	1,79
число зерен в ряду початка	1,25	1,75	1,57

Большинство лучших вариантов моделей (1.3) и (1.4), где в правых частях – данные по родительским линиям, не показали адекватности. F_p/F_T приблизительно в 1,1-3 раза выше, чем для модели (1.1) в варианте «без линий». Пример для лучшей из двух модели (1.3)

приведен в табл. 1.2. Подобные модели можно использовать лишь для приблизительного прогноза гетерозиса у гибрида F_1 .

Следует отметить, однако, что модель (1.1) сложнее в применении, так как для ее построения кроме данных о родительских формах необходимо иметь оценки их ОКС.

По озимой пшенице биометрико-генетическому анализу подвергали данные диаллельных скрещиваний, полученные специалистами лаборатории гетерозиса ВНИИ сельскохозяйственной биотехнологии за ряд лет в различных условиях выращивания (Смиряев, 1990). Использовали следующие наборы данных (серии опытов):

I – по шести сортам: Мироновская юбилейная, Мироновская 808, ППГ-186, Карстенс VIII, Гейнес, Кавказ, их прямым обратным гибридам F_1 (опыты в двух повторностях проводили в 1973, 1974 и 1975 гг. в Москве, РГАУ-МСХА);

II – по семи сортам: Абе, Одесская полукарликовая, Сейм, Прикульская 36, Заря, Чайка, Одинцовская и их прямым гибридам F_1 (опыты в трех повторностях проводили в 1982, 1983 и 1984 гг. в Москве, РГАУ-МСХА);

III – по предыдущим семи сортам и их прямым гибридам F_1 (опыты в трех повторностях проводили в 1982 и 1983 гг. в Дербенте – ДОС ВИР им. Н.И. Вавилова и Измаиле – на Измаильском ГСУ при орошении).

В ходе исследований измеряли пять признаков (табл. 1.3). Величины признаков и биометрико-генетических параметров значительно изменялись по опытам в пределах каждого набора данных. В опытах проявлялась существенная неаддитивность схем их наследования. Результаты оценки степени адекватности моделей (1.1) и (1.2) с включением и без включения показателей родительских сортов по трем наборам данных приведены в табл. 1.3. Регрессия (1.1) без включения данных о родительских сортах, как и для данных по кукурузе, адекватна в большинстве случаев ($F_p/F_T < 1$ для 12 из 15).

В этом варианте расчета для I набора данных величины F_p/F_T меньше единицы по всем признакам. Предположение об отсутствии реципрокных эффектов в I наборе более естественно, так как все расчеты проводили после усреднения данных по реципрокным гибридам. Во II и III наборах данных по «обратным» гибридам F_1 не было. В результате m_s_e для них несколько заниженные, и небольшие превышения единицы величин F_p/F_T (в трех случаях на 14, 7 и 13%), по крайней мере, отчасти объясняются влиянием неучтенных реципрокных эффектов. Для наиболее адекватной модели (1.1) «без сортов», возможно, проявляется еще одна причина повышения F_p/F_T , которая будет обсуждаться в следующем разделе.

Таблица 1.3. F_p/F_T для трех наборов данных по озимой пшенице.

Признак	Модель (1.1)		Модель (1.2)	
	с сортами	без сортов	с сортами	без сортов
I набор данных				
Масса зерна главного колоса	2,07	0,91	3,62	2,39
Масса 1000 зерен	1,79	0,89	2,01	1,90
Продуктивная кустиность	1,20	0,55	1,59	1,12
Число зерен главного колоса	1,86	0,90	2,80	1,91
Масса зерна растения	2,24	0,64	3,22	1,09
II набор данных				
Масса зерна главного колоса	1,08	0,65	1,93	1,49
Масса 1000 зерен	3,55	0,86	4,24	1,44
Продуктивная кустиность	0,76	0,48	0,91	0,59
Число зерен главного колоса	1,91	1,14	2,09	2,01
Масса зерна растения	0,98	0,79	1,40	1,34
III набор данных				
Масса зерна главного колоса	1,28	0,81	2,43	1,71
Масса 1000 зерен	3,24	1,07	4,87	1,66
Продуктивная кустиность	0,85	0,67	1,13	0,90
Число зерен главного колоса	1,02	0,72	3,82	1,76
Масса зерна растения	1,20	1,13	1,47	1,35

Таблица 1.4. Отношение F_p/F_T для модели (1.4) в зависимости от величины параметра D . Пояснение в тексте

Признаки	$D=0,5$	$D=1$	$D_{\text{опт}}$ (в скобках)
I набор данных			
Масса зерна главного колоса	1,53	1,85	1,49 (0,44)
Масса 1000 зерен	1,60	2,32	1,41 (0,60)
Продуктивная кустистость	1,87	1,20	1,20 (1,07)
Число зерен главного колоса	2,29	2,57	2,22 (0,63)
Масса зерна растения	1,62	1,36	1,38 (0,93)
II набор данных			
Масса зерна главного колоса	1,16	1,32	1,17 (0,38)
Масса 1000 зерен	1,63	2,38	1,61 (0,37)
Продуктивная кустистость	0,72	0,67	0,65 (1,16)
Число зерен главного колоса	3,04	2,89	2,94 (0,95)
Масса зерна растения	1,20	1,25	1,19 (0,21)
III набор данных			
Масса зерна главного колоса	1,48	1,31	1,24 (0,82)
Масса 1000 зерен	1,81	2,30	1,80 (0,60)
Продуктивная кустистость	0,94	0,87	0,88 (1,00)
Число зерен главного колоса	1,62	1,62	1,51 (0,76)
Масса зерна растения	1,47	1,31	1,32 (1,12)

Модель (1.2), судя по величинам F_p/F_T из табл. 1.3, хуже модели (1.1). Включение данных по родительским формам делает неадекватными модель (1.1) в 12 случаях из 15, а модель (1.2) – в 14 из 15.

Также как и для кукурузы, модели (1.3) и (1.4), где в правых

частях – величины признаков сортов, для пшеницы оказались хуже модели (1.1) в варианте «без сортов». Степень неадекватности большинства лучших вариантов модели (1.4) в 1,5–2 раза выше (табл. 1.4).

Эта модель при $D=0$ в 14 из 15 сочетаний «признак–набор данных» оказалась хуже, чем в вариантах $D=1$ и $D=0,5$. Поэтому в табл. 1.4 вариант $D=0$ не отражен. Отметим, что при $D=0,5$ часть выражения (1.4), стоящая в квадратных скобках, есть среднее арифметическое значений признака у двух родительских форм, а при $D=1$ – меньшее значение признака из двух родительских форм. Таким образом, эти характеристики родительских сортов оказываются более информативными индексами сред для описания изменчивости признака у гибридов F_1 , чем вариант $D=0$. Из табл. 1.4 следует, что иногда небольшой выигрыш по степени адекватности (снижение отношения F_p/F_T) можно получить, если в модели (1.4) подбирать оптимальное значение D ($D_{\text{опт}}$ в скобках) методом наименьших квадратов одновременно с коэффициентами C . Эту модель, как и модель (1.3), можно использовать лишь для приблизительного прогноза значений признаков у гибрида F_1 .

Описанные модели позволяют сократить объем скрещиваний и испытаний с целью выявления высокогетерозисных гибридов. Для прогнозирования признака у гибрида F_1 по модели (1.1) в определенном диапазоне внешних условий необходимо оценить эффекты ОКС. Для этого не обязательно проводить испытания полных диаллельных схем во многих средах. Эффекты ОКС можно оценить, используя различные схемы неполных скрещиваний: топкроссы, сетевые пробные скрещивания, частичные, фракциональные и парциальные диаллельные кроссы (Драгавцев, Цильке, 1984; Савченко, 1984; Смирязев и др., 1992).

Следует отметить, что модель (1.1) сложнее в применении по сравнению с (1.3) и (1.4), хотя последние менее адекватны. Предварительно, как и для использования модели (1.1), необходимо провести испытания диаллельных схем и получить МНК-оценки коэффициентов C в регрессиях (1.3) или (1.4). Но после этого для прогноза по ним значений признаков их гибридов F_1 в новых неконтрастных условиях выращивания достаточно иметь данные в этих условиях лишь о признаках тех же родительских форм. Оценки ОКС в новых условиях не нужны.

1.3. Причины адекватности регрессии на ОКС: гипотеза об аллельном несходстве родительских форм

Желательно понять причины значительно более частой

адекватности регрессионной модели (1.1) для варианта «без сортов» по сравнению с остальными моделями. Для этой цели были использованы методы аналитического и компьютерного моделирования (Смиряев, 2011).

Прежде всего, вызывает удивление, что при обобщении любого набора данных по кукурузе или озимой пшенице за 2-4 несходных по условиям года и, иногда, из разных мест выращивания в рамках одной модели (1.1) достаточно единой МНК-оценки C_i для каждой i -й родительской формы. Ведь, как отмечалось выше, значение любого признака и селекционно-генетического параметра родителей и их гибридов F_1 существенно меняются по опытам, обобщаемым в модели. С другой стороны проверка модели (1.1) при анализе данных каждого опыта по отдельности (одно сочетание год-место) также выявила ее адекватность. Но МНК-оценки C_i для каждой родительской формы менялись от опыта к опыту одного набора данных. Попытки установить причины такого противоречия не привели к успеху. Связь величин C_i с параметрами i -х родительских форм, их гибридов и т.п. не проявлялась. Этот факт может означать, что вариабельность коэффициентов C_i вызвана в основном случайными ошибками опытов.

Для проверки был проведен подбор единой по родительским формам МНК-оценки коэффициента C для каждого опыта с пшеницей и кукурузой по отдельности. В результате установлено, что независимо от культуры, признака, опыта или целого набора данных МНК-оценка C близка к $\frac{p-1}{p-2}$, где p – число родителей в схеме скрещивания.

Например, для набора с шестью родительскими формами ($p=6$) МНК-оценка близка к $C=5/4$, для $p=7$ – $C=6/5$. Поэтому далее провели анализ адекватности модели (1.1) в варианте «без сортов» при едином заранее выбранном значении $C = \frac{p-1}{p-2}$.

Напомним, что средний квадрат отклонений от регрессионной модели имеет структуру $ms_0 = SS_0/df_1$, где SS_0 – сумма квадратов отклонений экспериментальных данных от регрессии, а расчетный критерий Фишера: $F_p = ms_0/ms_e$. Значение SS_0 для модели (1.1), т.е. при подборе различающихся C_i по родительским формам в каждом расчете, естественно, было меньше, чем SS_0 при едином заранее выбранном

значении $C = \frac{p-1}{p-2}$: за счет побора МНК-оценок C_i , для каждого родителя по отдельности, всегда удается сократить SS_0 . Но числа степеней свободы (df_1) также отличаются в этих двух вариантах расчета.

При априорном выборе единого значения $C = \frac{p-1}{p-2}$ число df_1 больше.

Например, для схем только с прямыми гибридами и при расчете «без сортов» в одних условиях испытания диаллельной схемы (одно сочетание год-место) при $C = \frac{p-1}{p-2}$ имеем $df_1 = p(p-1)/2 - p$. При подборе

дополнительно p штук МНК-оценок C_i число степеней свободы становится меньше: $df_1 = p(p-1)/2 - 2p$. Значение ms_e случайных ошибок, естественно одинаково в этих двух вариантах расчета. Из-за различий df_1 величины F_T – табличных значения критерия Фишера также несколько различаются, что следует учесть при сравнении отношений F_p/F_T в двух вариантах расчета.

Итак, при едином значении $C = \frac{p-1}{p-2}$ регрессионная модель (1.1)

упрощается:

$$x_{ijk} = M + l_k + \frac{p-1}{p-2} (g_{ik} + g_{jk}) + e \quad (1.5).$$

Установлено, что для большинства опытов с кукурузой и пшеницей модель (1.5) обычно остается адекватной. Причина – большее число df_1 : параметры SS_0 увеличиваются по сравнению с исходной моделью (1.1), где используются МНК-оценки C_i , но значения $ms_0 = SS_0/df_1$ иногда даже снижаются. В подобных случаях оправдан переход от исходной модели (1.1) к более простому варианту регрессии (1.5).

Расчет для одного опыта набора данных означает, что в модели (1.5) номер k – фиксирован. Если для каждого опыта набора по отдельности адекватна модель (1.5), то, естественно, она адекватна для всего набора данных.

В качестве примера в табл. 1.5 приведены оценки F_p/F_T , полученные при проверке адекватности моделей (1.1) и (1.5) по данным II набора озимой пшеницы отдельно по каждому из 3-х лет испытания диаллельной схемы. Модель (1.1) адекватна в 12 расчетах из 15, модель (1.5) также в 12 из 15. Для признака «продуктивная кустистость» отношения F_p/F_T , особенно в опыте 1983 г., явно занижены. Причина, возможно, в том, что плотность распределения этого признака не соответствует нормальному.

Причины значительного превосходства отношения F_p/F_T в модели (1.5) над моделью (1.1) для признака «число зерен главного колоса», будут проанализированы далее.

Таблица 1.5. Параметры F_p/F_T для моделей (1.1) и (1.5) по пяти признакам набора данных II для трех лет испытания по отдельности.

Признак	1982 г.		1983 г.		1984 г.	
	(1.1)	(1.5)	(1.1)	(1.5)	(1.1)	(1.5)
Масса зерна главного колоса	0,73	0,75	0,53	0,82	0,83	1,07
Масса 1000 зерен	0,71	0,54	1,07	0,91	0,79	0,64
Продуктивная кустистость	0,45	0,29	0,17	0,20	0,58	0,64
Число зерен главного колоса	0,65	1,17	1,37	2,66	0,95	2,81
Масса зерна растения	0,66	0,63	0,77	0,81	1,29	0,84

Возможная причина частой адекватности модели (1.5) – это особенность аллельного несходства набора из p гомозиготных родительских форм, по локусам, определяющим изменчивость конкретного признака и опыта в диаллельной схеме. А именно, каждая родительская форма из p по некоторым полиморфным локусам имеет свою уникальную группу аллелей. Остальные $(p-1)$ родители по этим локусам несут другой, но одинаковый набор «общих» аллелей. Для пояснения этой гипотезы на рис. 1.4 представлена упрощенная схема такого распределения групп уникальных и «общих» аллелей среди p родительских форм.

$$\begin{array}{l}
 i=1: U_1 G_2 G_3 \dots G_p \\
 i=2: G_1 U_2 G_3 \dots G_p \\
 i=3: G_1 G_2 U_3 \dots G_p \\
 \vdots \\
 \vdots \\
 i=p: G_1 G_2 G_3 \dots U_p
 \end{array}$$

Рисунок 1.4. Гипотетическое распределение групп аллелей у p родителей диаллельной схемы скрещиваний. Пояснение в тексте

Поскольку сцепление локусов не влияет на выражение количественного признака у гибридов F_1 диаллельной схемы, уникальную группу аллелей каждого родителя (i), независимо от расположения их в геноме, можно рассматривать как один аллель. Он обозначен на рис.1.4 одним значком U_i . Аналогично каждая группа «общих» аллелей представлена значком G_i .

Для проверки гипотезы об аллельном несходстве родительских форм проведено аналитическое моделирование генотипической

ценности полученных гибридов F_1 диаллельной схемы и ОКС гомозиготных родителей в одних условиях выращивания. Кроме гомозиготности родительских форм с особым распределением аллелей (рис. 1.4) в этом анализе предполагалась аддитивно-доминантная схема наследования (эпистаза нет).

Для пояснения последнего предположения в табл. 1.6 представлена схема наследования признака в рамках F_{∞} -метрики для двух полиморфных локусов (Смиряев и др., 1992). Эпистаза в схеме наследования признака нет, если $i=j'=l=0$.

Таблица 1.6. Структура генотипических ценностей признака при двух полиморфных локусах (A и B). Пояснения в тексте

	B_1B_1	B_1B_2	B_2B_2
A_1A_1	$m+a_1+a_2+i$	$m+a_1+d_2+j$	$m+a_1-a_2-i$
A_1A_2	$m+d_1+a_2+j'$	$m+d_1+d_2+l$	$m+d_1-a_2-j'$
A_2A_2	$m-a_1+a_2-i$	$m-a_1+d_2-j$	$m-a_1-a_2+i$

Здесь m – среднее значение признака по всем четырем гомозиготам; a_1, a_2, d_1, d_2 – аддитивные и доминантные эффекты аллелей A_1, B_1, A_2, B_2 локусов A и B .

i – эффект межлокусного взаимодействия (эпистаза) типа «гомозигота (локус A) - гомозигота (локус B)»;

j – взаимодействие «гомозигота-гетерозигота»;

j' – взаимодействие «гетерозигота-гомозигота»;

l – взаимодействие «гетерозигота-гетерозигота».

Через a_i далее обозначен суммарный аддитивный эффект группы аллелей G_i – тех «общих» аллелей, которых нет у i -го родителя, но есть у остальных (рис. 1.4). При нескольких полиморфных локусах для любого гибрида F_1 с участием i -го и j -го родителей через d_i обозначен суммарный доминантный эффект группы его гетерозиготных локусов, несущих у i -го родителя аллели группы U_i , а у остальных – G_i . Аналогично, d_j – доминантный эффект группы локусов, имеющих у его j -го родителя аллели группы U_j , а у остальных – G_j . Через m обозначим среднюю генотипическую ценность всех возможных гомозигот, которые можно составить из p полиморфных локусов, содержащихся в наборе родителей. Тогда в рамках F_{∞} -метрики без учета случайной ошибки опыта (e) и реципрокного эффекта генотипическая ценность гибрида F_1 i -го и j -го родителей равна:

$$x_{ij} = m + \sum_{l=1}^p a_l - a_i - a_j + d_i + d_j \quad (1.6).$$

В таких обозначениях после преобразований получаем выражение средней генотипической ценности всех гибридов F₁ диаллельной схемы (без родителей) в одних условиях выращивания:

$$M = \frac{1}{p(p-1)} \sum_{i=1}^p \sum_{j \neq i}^p x_{ij} = m + 2d + (p-2)a,$$

где $a = \frac{1}{p} \sum_{i=1}^p a_i$, $d = \frac{1}{p} \sum_{i=1}^p d_i$.

После алгебраических преобразований получаем значение ОКС *i*-го родителя, оцененное по его гибридам F₁:

$$g_i = \frac{1}{p-1} \sum_{j \neq i}^p x_{ij} - M = \frac{p-1}{p-2} \cdot [(d_i - d) - (a_i - a)].$$

После подстановки полученных выражений *M*, *g_i*, *g_j* в модель (1.5) для одних условий выращивания (год-место) без учета *e* получаем:

$$x_{ij} = M + \frac{p-1}{p-2} \cdot (g_j + g_i) = m + \sum_{l=1}^p a_l - a_i - a_j + d_i + d_j,$$

что совпадает с (1.6).

Таким образом, отсутствия эпистаза и распределения аллелей, представленного на рис.1.4, достаточно для адекватности модели (1.5) при обобщении данных одного опыта (сочетание год-место). Если от опыта к опыту меняется экспрессия генов (*m*, *a₁*, *a₂*, *d₁*, *d₂*), или даже происходит переопределение спектра полиморфных генов, влияющих на изменчивость признака, адекватность моделей (1.5) по отдельным опытам означает адекватность этой модели по всему набору опытов.

В рамках названных предположений генотипические ценности гомозиготных родительских форм равны:

$$x_{ii} = m + \sum_{l=1}^p a_l - 2a_i.$$

Если эти выражения признака у родителей вместе с генотипическими ценностями (1.6) для гибридов F₁ попытаться подставить в модель (1.5), то полученные по ней прогнозы не совпадут с этими генотипическими ценностями. Модель (1.5) утратит адекватность, как и модель (1.1). Именно поэтому модели в вариантах «с родительскими формами», как правило, неадекватны (табл. 1.1 и 1.3).

С учетом рис. 1.4 в предположении отсутствия эпистаза адекватность модели (1.5) «без родительских форм» может служить аналогом известного теста на аллелизм для генов количественного признака, который применяется при анализе качественных признаков (Лобашов, 1967). Отметим также, что вопреки традиционному представлению о более существенном влиянии на ОКС (g_i) аддитивных эффектов генов в представленной модели (1.5) на изменчивость g_i в равной степени влияют и доминантные эффекты:

$$g_i = \frac{p-1}{p-2} \cdot [(d_i - d) - (a_i - a)].$$

Селекционеры нередко составляют наборы неродственных или несходных по фенотипу родительских форм, в надежде получить при их скрещивании гетерозисные гибриды или полиморфные популяции с высоким выходом трансгрессивных рекомбинантов. По-видимому, при таком подборе родителей уникальность их олигогенов или блоков (рис. 1.4) – довольно типичное явление в наборе. Этим и объясняется частая адекватность модели (1.5) при анализе случайно выбранных нами селекционно-генетических опытов с пшеницей и кукурузой.

Отметим, что для частичной проверки выполнения условий адекватности модели (1.5) применим метод Хеймана. С помощью этого метода проверяется отсутствие эпистаза в схеме наследования признака, а также условие т.н. независимого распределения генов у родительских форм. Последнее необходимо для выполнения второго предположения – о распределении аллелей (рис. 1.4).

Если регрессионный анализ выявит неадекватность модели (1.5), то ее можно попытаться компенсировать методом перебора. Для этого следует последовательно исключать из данных диаллельного скрещивания гибриды F_1 одного-двух родителей и проводить регрессионный анализ для оставшихся данных. Возможно, только исключенный материал не соответствуют предположениям модели (1.5).

Оценка влияния отклонений от предположений и случайных ошибок опытов

Два предположения – аддитивно-доминантная схема наследования признака (нет эпистаза) и специальное распределение аллелей, представленное на рис.1.4, являются лишь достаточным условием для адекватности моделей (1.1) и (1.5) в вариантах «без родительских форм». Возможно, существуют другие совокупности предположений, обеспечивающих адекватность. Кроме того, названные предположения, обычно, не выполняются в точности, а на результаты

экспериментов всегда накладываются случайные ошибки. Следует оценить чувствительность моделей к воздействию названных факторов и попытаться объяснить неадекватность моделей для некоторых опытных данных. С этой целью был использован метод компьютерного моделирования, реализованный в специально разработанных программах.

Для набора из шести гипотетических гомозиготных родительских генотипов ($p=6$) случайно выбирались величины a_i , d_i , m . При подстановки их в выражение (1.6) составляли значения x_{ij} для всех 15 диаллельных гибридов F_1 набора. С помощью метода Монте-Карло к полученным значениям x_{ij} добавляли «случайные ошибки опыта» с заданным уровнем дисперсии σ_e^2/n .

Проводился регрессионный анализ полученных значений x_{ij} в рамках моделей (1.1) и (1.5), рассчитывались их остаточные средние квадраты ms_01 и ms_05 . Поскольку, кроме прочих, рассматривались варианты, где $\sigma_e^2/n=0$ (F_p/F_T получить невозможно), для общности получаемых выводов достаточно сравнивать ms_01 и ms_05 . При одном значении σ_e^2/n расчеты повторялся несколько раз, после чего полученные оценки ms_01 и ms_05 усреднялись для каждой из двух моделей по отдельности. Затем дисперсии ошибок σ_e^2/n придавалось новое значение в диапазоне от 0 до $0,5\sigma_G^2$ (т.е. до половины генотипической дисперсия признака у гибридов F_1 всей диаллельной схемы) и расчеты ms_01 , ms_05 многократно повторялись.

Моделировались различные варианты отклонений от предположений, т.е. от генотипической ценности гибридов F_1 , соответствующей выражению (1.6). Во-первых, это отклонения от распределения аллелей у родительских форм, представленного на рис.1.4: множественный аллелизм хотя бы по одной группе аллелей, совпадение хотя бы одной группы уникальных аллелей у двух родителей, полное совпадение генотипов у двух родителей и т.п.

Во-вторых, моделировались стандартные типы эпистаза, представленные в табл. 1.6 (т.е. эффектам i, j, j', l придавались значения, отличные от нуля), а также особый тип взаимодействия генов, рассмотренный в работах Rasmusson (1933), Forkmann и Seyffert (1977). В последнем случае предполагалось, что при увеличении суммарных эффектов генов, влияющих на «внутренний признак» x_{ij} , величина измеряемого в опыте признака выражается гладкой функцией $f(x_{ij})$ – нелинейной трансформацией x_{ij} . Функция $f(x_{ij})$ возрастает по x_{ij} , но при этом ее первая производная по x_{ij} убывает, возможно, вплоть до 0

(выход на плато).

Подобная ситуация возникает, когда реализация признака, измеряемого в опыте, сопровождается физиологическими лимитирующими эффектами, ослабляющими влияние на признак эффектов генов по мере роста их суммарного значения x_{ij} (рис. 1.5). Например, если два локуса влияют на изменчивость «внутреннего признака» x_{ij} – концентрацию вещества, например, ауксина, стимулирующего рост растения. Влияния отдельных локусов на x_{ij} просто суммируются – эпистаза нет. При генетически обусловленном повышении значений x_{ij} возникает эффект «насыщения»: прибавка ауксина не вызывает прямо пропорционального роста измеряемого признака – высоты растения. Более того, начиная с определенной концентрации ауксина, его прибавка может даже угнетать растяжение клеток побега (Полевой, 1984). В результате прибавка высоты, вызванная увеличением вклада одного локуса, влияющего на изменчивость концентрации ауксина, зависит от вклада второго локуса. В схеме наследования высоты, которая анализируется в рамках моделей (1.1) и (1.5), проявится эпистаз. Такой тип взаимодействия генов можно назвать «эпистазом-трансформацией».

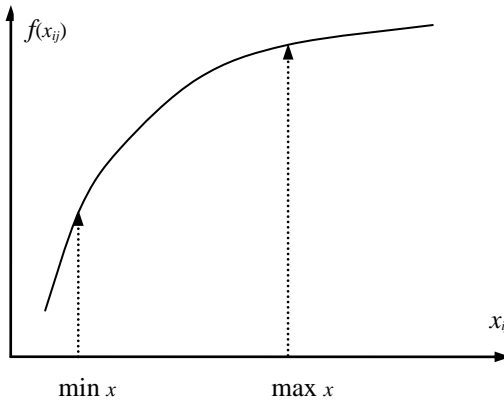


Рисунок 1.5. Характер нелинейной трансформации x_{ij} – суммарного эффекта генов. $\min x$ – $\max x$: диапазон изменений моделируемых величин x_{ij} «внутреннего признака», влияющего на $f(x_{ij})$ – измеряемый признак гибридов F_1 .

Выводы по результатам компьютерного моделирования формировались на основе оценок остаточных средних квадратов ms_{01} и ms_{05} . При необходимости результаты сопоставлялись с оценками,

полученными с помощью регрессионного анализа реальных данных диаллельных скрещиваний.

Компьютерное моделирование позволило прийти к следующим выводам:

1) Если оба предположения выполнялись (аддитивно-доминантная схема наследования, распределение аллелей соответствует рис.1.4), то при наложении на x_{ij} ошибок опытов с дисперсией σ_e^2/n , как и ожидалось, параметр ms_05 обычно был меньше ms_01 и оба незначимо отличались от σ_e^2/n . Причины связаны с разным числом степеней свободы и обсуждались выше. При фиксированных a_i , d_i , m по мере возрастания дисперсии σ_e^2/n от 0 до $0,5\sigma_G^2$ коэффициенты регрессии

(1.1), как и ожидалось, менялись от единого $C = \frac{p-1}{p-2}$ до МНК-оценок

C_i , случайных, но все более различающихся по родителям (i).

2) При $\sigma_e^2/n=0$ любые вышеназванные варианты нарушений предположений о схеме наследования признака вызывали отклонения от обеих регрессионных моделей ($ms_0>0$), т.е. потерю адекватности. Но неадекватность проявлялась в разной степени. Так, при $\sigma_e^2/n=0$ в любом моделируемом варианте отклонения от распределения аллелей, представленного на рис. 1.4, параметр ms_01 был в 1,4-2,5 раза больше ms_05 . Стандартные типы эпистаза (гомозигота-гомозигота, гомозигота-гетерозигота и т.п. – см. табл. 1.6) или их сочетания также вызывали потерю адекватности моделей, но как и в предыдущих случаях параметр ms_05 был несколько меньше ms_01 . Возрастание σ_e^2/n от 0 до $0,5\sigma_G^2$, естественно, вызывало увеличение отклонений от моделей, но превышение ms_01 над ms_05 немного сокращалось – до 1,2-2 раз. То есть модель (1.1) более чувствительна к нарушениям «стандартных» предположений, чем модель (1.5).

3) Противоположная ситуация наблюдалась при оценке эффекта эпистаза-трансформации $f(x_{ij})$, вызванной физиологическим эффектом насыщения (рис. 1.5). В компьютерном моделировании использовали функции двух видов: $f(x_{ij})=\ln(x_{ij})$ или $f(x_{ij})=(x_{ij})^2$, где величина Z принимала значения 0,8, 0,6 или 0,4. Это позволило менять вид трансформации, а также ее степень в достаточно большом диапазоне.

Установлено, что при $\sigma_e^2/n=0$ оба вида трансформации приводили к потере адекватности моделей (1.1) и (1.5), но последняя модель более чувствительна к эпистазу такого типа: ms_05 неизменно больше ms_01 . Значительная компенсация неадекватности для модели

(1.1) происходила за счет подбора МНК-оценок C_i . Превосходство ms_05 над ms_01 существенно зависело от схемы наследования в группах полиморфных локусов (например, $a_i > d_i$ – неполное доминирование, $a_i < d_i$ – свехдоминирование, их комбинации по локусам), а также от диапазона изменений моделируемых величин x_{ij} в одном расчете. Так, если x_{ij} менялись от 10 до 100 (диапазон $\min x - \max x$ на рис. 1.5), то в зависимости от схемы наследования значение ms_05 было больше ms_01 приблизительно в 1,5-4 раза. Причем конкретный вид функции $f(x_{ij})$ и Z (степень трансформации) влияли на это слабо. Отмечено также, что в модели (1.1) различия C_i по родительским формам не являются случайными, и корреляция по родителям между g_i и C_i для разных вариантов таких расчетов обычно отрицательна. То есть, значения C_i не были случайными, в отличие от влияния на C_i ошибок опытов в повторениях.

Для всех вариантов эпистаза-трансформации при наложении случайных ошибок опытов по мере возрастания σ_e^2/n оба значения ms_0 , естественно, также растут, но ms_01 приближается к ms_05 . Динамика зависит в основном от соотношения трех параметров: исходного ms_05 (т.е. при $\sigma_e^2/n=0$), σ_G^2 и σ_e^2/n . При малых значениях исходного ms_05 (например, 5% от σ_G^2), увеличение σ_e^2/n вызывало быстрое приближение ms_01 к ms_05 и к σ_e^2/n . Модели (1.1) и (1.5) по критерию Фишера становились почти адекватными уже при $\sigma_e^2/n=0,05 \sigma_G^2$: случайная ошибка «перекрывает» влияние слабой трансформации. При большем исходном ms_05 (например, 30% от σ_G^2), даже если $\sigma_e^2/n=0,3 \sigma_G^2$, параметр ms_05 оставался больше ms_01 приблизительно в 2 раза. Отметим, что $\sigma_e^2/n=0,05 \sigma_G^2$ соответствует коэффициенту наследуемости H_{F1}^2 в наборе гибридов диаллельной схемы, равному 0,95, а $\sigma_e^2/n=0,3 \sigma_G^2 - H_{F1}^2=0,77$.

В табл.1.5 для признака «число зерен главного колоса» по трем годам полевого эксперимента было отмечено превышение ms_05 над ms_01 в 2-3 раза. При этом H_{F1}^2 варьировал по 3-м годам от 0,93 до 0,98. Возможно, это свидетельствует о влиянии эпистаза-трансформации рассмотренного типа.

Для проверки такого предположения экспериментальные данные набора II для каждого года по этому измеряемому признаку были подвергнуты «обратной трансформации» $f^{-1}(x_{ij})$ с производной, возрастающей по x_{ij} . Цель – попытаться избавиться от нелинейной

зависимости, подобной рис. 1.5, вернуть схему наследования преобразованного признака «число зерен главного колоса» к аддитивно-доминантной. Учитывая, что вид возможной трансформации $f(x_{ij})$ неизвестен, в качестве $f^{-1}(x_{ij})$ использовали простой вариант:

$$f^{-1}(x_{ij}) = (a + b x_{ij})^c, \text{ где } 1 < c \leq 2 \quad (1.7).$$

Коэффициенты a , b , c подбирались по критерию $\min(ms_05)$ для модели (1.5) в каждом из трех лет эксперимента отдельно. Результаты регрессионного анализа данных эксперимента, подвергнутых «обратной трансформации» (1.7) с подобранными коэффициентами a , b , c , в сравнении с ms_01 и ms_05 до преобразования представлены в табл.1.7. Отметим, что коэффициенты a , b , c , подобранные для каждого года по отдельности, близки по значениям.

После обратного преобразования (1.7) параметры ms_05 снизились до уровня, соответствующего адекватности этой модели, а параметры ms_01 для всех 3-х лет эксперимента наоборот стали выше ms_05 . Для 3-го года модель (1.1) сохранила адекватность.

Таблица 1.7. Параметр ms_01 и ms_05 до и после преобразования (1.7) данных по признаку «число зерен главного колоса». Звездочкой отмечены параметры ms_0 , значимо большие ms_e ($P < 0,05$).

	1982 г.		1983 г.		1984 г.	
	(1.1)	(1.5)	(1.1)	(1.5)	(1.1)	(1.5)
Для исходных данных	2,88	4,43*	5,08*	8,46*	3,00	7,66*
После преобразования (1.7)	5,06*	3,87	9,08*	6,76	4,14	3,94
	$a=-61,0; b=3,3; c=2$		$a=-54,5; b=3,2; c=2$		$a=-59,0; b=3,4; c=2$	

Таким образом, результаты табл. 1.7 подтверждают возможность присутствия эпистаза-трансформации (рис. 1.5) в схеме наследования признака «число зерен главного колоса». Если при анализе экспериментальных данных модель (1.1) проявила адекватность, а модель (1.5) – нет, то с помощью преобразований, подобных (1.7), можно попытаться выявить этот тип эпистаза.

В заключение отметим, что во всех вариантах как компьютерного моделирования, так и анализа экспериментальных данных, при попытке заменить модель (1.1) на следующую:

$$x_{ijk} = M + I_k + C_i \times g_{ik} + C_j \times g_{jk} + e$$

существенно увеличивался параметр ms_01 : адекватность резко снижалась. То есть «перекрестный» характер взаимной реакции генотипов родителей в их гибриде $F_1 (C_i \times g_{jk} + C_j \times g_{ik})$ в структуре модели (1.1) важен для ее адекватности.

1.4. Прогноз гетерозиса с помощью оценок генетической дивергенции родителей.

В литературе представлено значительное количество работ, где пытаются установить связь различных показателей генетического несходства родительских форм не с величинами признака их гибридов F_1 , а со степенью гетерозиса этих гибридов: для пшеницы (Barbose, 1996; Benin et al., 2012), гороха (Sarawat et. al, 1994; Esposito et. al, 2014), кукурузы (Dubreuil et. al, 1996; Yang et al., 2011), рапса (Diers, 1996; Ahmad, Quiros, 2011), риса (Saghai et. al, 1997; Cho et. al, 2004), перца (Шаптуренко и др., 2014). Обычно в таких работах генетическое несходство родителей, т.е. несходство аллельного состава, оценивается косвенно с помощью разнообразных биометрико-генетических показателей. Используют сравнительно простые оценки (разности величин признака или ОКС родительских форм, коэффициенты их родства, многомерные статистики типа евклидова расстояния, несходство главных компонент и т. д.) и более современных, основанных на сравнении генетических маркеров, результатов RFLP и т.п. Кроме эффектов гетерозиса, в результате скрещивания генетически дивергентных родителей ожидают повышение генетической вариации количественных признаков в гибридных популяциях и увеличение вероятности трансгрессивного расщепления. Генетическое несходство родителей вызывает увеличение доли гетерозиготных локусов у гибридов F_1 .

Нами также была проведена попытка установить взаимосвязь двух биометрико-генетических оценок генетического расстояния между сортами озимой пшеницы с эффектами гетерозиса их гибридов F_1 (Смиряев и др, 1999). В качестве первой оценки генетического расстояния, т.е. аллельного несходства родительских сортов, использованы парные коэффициенты родства, которые определяются по их родословным (Смиряев и др, 1992). Эти коэффициенты были вычислены для семи родительских сортов озимой пшеницы, одинаковых в II и III наборах данных, доктором биологических наук С.П. Мартыновым. Чем ниже коэффициент родства двух сортов, тем менее сходны они по аллельному составу, а их гибрид F_1 , имеет более высокую долю гетерозиготных локусов.

Вторая оценка – парные евклидовы расстояния семи сортов по количественным признакам (Смиряев и др., 1992). Вначале оценки получили по отдельности для каждого из семи условий испытания диаллельной схемы набора из 7 сортов:

$$E_{ij} = \sqrt{\sum_{l=1}^N (x_{il} - x_{jl})^2 / \sigma_l^2},$$

где x_{il} , x_{jl} – значения l -го признака соответственно у i -го и j -го родительских сортов; σ_l^2 – дисперсия величин l -го признака семи сортов. N – число признаков, которое при расчете было равно четырем, а не пяти (табл. 1.3): признак «продуктивная кустистость» не соответствует требованию евклидова расстояния о нормальном распределении всех N признаков.

Полученные оценки E_{ij} из 7 условий выращивания усреднялись по всем этим опытам, но для каждой из 21 пары сортов (i, j) по отдельности. В результате получена 21 парная оценка \bar{E}_{ij} – евклидовых расстояний семи сортов. Чем выше \bar{E}_{ij} для пары (i, j) родительских сортов, тем менее сходен, как предполагается, их аллельный состав.

Для каждого из 21 гибрида F_1 и каждого из четырех признаков для 7 условий испытания отдельно вычисляли показатель гипотетического гетерозиса H – отклонение признака у гибрида от полусуммы значений этого признака у двух родительских сортов. Затем получали \bar{H} – отклонение, усредненное по 7 условиям испытания II и III наборов данных. Показатели гетерозиса 21 гибрида для каждого признака по отдельности сопоставляли с двумя оценками генетического несходства пары их родительских сортов: коэффициентами родства и евклидова расстояния.

Совместный анализ величин H и парных коэффициентов родства для пшеницы не выявил достоверных статистических связей между ними. Иная картина наблюдалась при анализе корреляций \bar{E} с эффектами гетерозиса H по четырем признакам у 21 гибрида F_1 , в каждом из условий выращивания, а также с их средними по 7 опытам (\bar{H}).

Хотя $\text{cor}(\bar{H}, \bar{E})$ были выше $\text{cor}(H, \bar{E})$, но сами корреляции $\text{cor}(\bar{H}, \bar{E})$ были невелики (1-я строка табл. 1.8), причем 3 коэффициента корреляции из 4 отрицательные. Однако для 3 из 4 признаков была выявлена более четкая статистическая связь \bar{E} не с \bar{H} , а с разбросом \bar{H} вокруг линии регрессии \bar{H} на \bar{E} .

Таблица 1.8. Коэффициенты корреляции между эффектами гетерозиса (\bar{H} или $|d\bar{H}|$) 21 гибрида F_1 пшеницы и оценками генетического несходства \bar{E} их родительских сортов.

Признак	масса 1000 зерен	масса зерна главного колоса	число зерен главного колоса	масса зерна растения
$\text{cor}(\bar{H}, \bar{E})$	0,58**	-0,51*	-0,52*	-0,35
$\text{cor}(d\bar{H} , \bar{E})$	0,34	0,78**	0,84**	0,69**

* – значимо на уровне $P < 0,05$, ** – $P < 0,01$

Например, по признаку «масса зерна главного колоса» (гр.) на рис. 1.6 представлена 21 точка с координатами (\bar{H}, \bar{E}) для 21 гибрида F_1 , и нанесена линия регрессии. Видно, что проявляется слабая тенденция снижения средних значений \bar{H} от \bar{E} , но разброс значений гетерозисных эффектов гибридов вокруг линии регрессии увеличивается при возрастании несходства родительских сортов (\bar{E}).

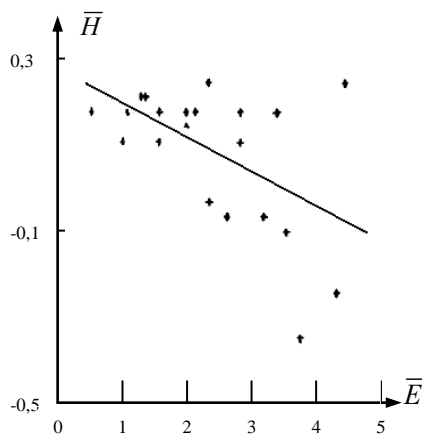


Рисунок 1.6. Двумерное распределение значений (\bar{H}, \bar{E}) для признака «масса зерна главного колоса» (гр.) у 21 гибрида пшеницы. Прямая – линия регрессии \bar{H} на \bar{E} .

Если подобный разброс для каждого анализируемого признака оценить через $|d\bar{H}|$ (абсолютные отклонения \bar{H} от соответствующей линии регрессии \bar{H} на \bar{E}), то для 3 признаков из 4-х коэффициенты корреляции $\text{cor}(|d\bar{H}|, \bar{E})$ оказываются положительными и высокозначимыми (2-я строка табл. 1.8). Можно сказать, что в таких

ситуациях по мере увеличения \bar{E} возрастает вероятность получить гибриды с различными, чаще отрицательными, отклонениями признака от его среднего у 2 родителей. При низких значениях \bar{E} разброс отклонений \bar{H} значительно меньше. Для признака «масса 1000 зерен» более четкая корреляционная связь проявилась между \bar{H} и \bar{E} (табл. 1.8).

Заключение

Объяснить и, тем более, прогнозировать сильное варьирование параметров и схем наследования признаков, обнаруженных в рассмотренных диаллельных скрещиваниях ячменя, кукурузы и озимой пшеницы, довольно сложно из-за отсутствия общей теории изменчивости детерминации количественных признаков. Несмотря на такое варьирование предложенные модели позволяют сократить объем скрещиваний и испытаний с целью выявления высокогетерозисных гибридов.

В частности, модели (1.1) и (1.5) в варианте «без родительских форм» оказалась высокоадекватными. С помощью аналитического моделирования показано, что возможная причина – аддитивно-доминантная схема наследования количественных признаков и особенность распределение аллелей полиморфных локусов. Установлено также, почему модель теряет адекватность, если в нее, наряду с гибридами F_1 , подставить данные по родительским формам.

Адекватность подобных моделей сохраняется, даже если происходит переопределение спектра полиморфных генов количественного признака в разных условиях выращивания. Причина подобной универсальности, по-видимому, именно в выявленном частном варианте распределения аллелей полиморфных локусов. Последнее может быть вызвано подбором несходных родительских форм при селекции на гетерозис.

Компьютерное моделирование позволило оценить чувствительность моделей к отклонениям от названных условий адекватности и к случайным ошибкам опытов. В частности, показана большая, по сравнению с моделью (1.1), чувствительность модели (1.5) к эпистазу, возникающему из-за физиологического эффекта насыщения.

Вопреки традиционному представлению о более существенном влиянии на эффекты ОКС аддитивных эффектов генов родителей, адекватность модели (1.5) предполагает, что на ОКС в равной степени влияют и доминантные эффекты. Модель (1.5) может служить аналогом известного теста на аллелизм для генов количественного признака, который применяется при анализе качественных признаков.

На основе экспериментальных данных исследована вероятность получения гетерозисных гибридов F_1 в зависимости от генетического несходства родительских форм, оцененного с помощью косвенных метрик несходства. Показано, что при увеличении евклидова расстояния между родительскими сортами озимой пшеницы возрастает вероятность получить гибриды F_1 с различными отклонениями некоторых признаков от их среднего у 2 родителей.

Глава 2. Анализ сходства генотипов по эколого-онтогенетической изменчивости ростового признака

Ростовой признак рассматривается как индикатор онтогенеза растений с определенным генотипом (сорт, гибрид F₂, клон и т.п.). Предложен новый подход для биометрико-генетического анализа сходства генотипов: анализируется сходство по форме реакции их ростового признака на изменение условий совместного выращивания в серии полевых опытов. В основе методики – регрессионные модели, аппроксимирующие кривые взаимной, относительной изменчивости роста одного признака у пары генотипов. Для количественной оценки сходства каждой пары генотипов вводятся параметры «негладкости» таких кривых и их расхождения в разных условиях совместного выращивания (годы, места и т.п.). Намечаются способы оценки взаимосвязи этих парных онтогенетических параметров сходства в наборе генотипов со степенью их близкородственности, а также с их парной изменчивостью в экологических градиентах, т.е. по конечным значениям того же признака. Выдвинута гипотеза о сходстве относительной изменчивости признака в экологических градиентах и онтогенезе, а также о связи степени такого сходства генотипов со сходством их аллельного состава. Проведена проверка этой гипотезы на трехлетних данных, полученных при многократном измерении трех количественных признаков (высота растений, длина колоса, площадь флагового листа) в период их роста для 22 совместно испытываемых сортов мягкой озимой пшеницы, с привлечением семилетних данных о высоте растений этих сортов при уборке. Используются коэффициенты родства сортов, показатели сходства аллельного состава генов, кодирующих запасные белки, а также различия дат наступления фенофаз.

Несмотря на огромный объем накопленных данных о росте и развитии высших организмов, генетика онтогенеза остается одним из наименее развитых направлений биологии. Важность этого направления в селекционно-генетических исследованиях растений обусловлена сильной вариабельностью количественных хозяйственно ценных признаков в зависимости от условий выращивания (в экологических градиентах).

Многие десятилетия генетика количественных признаков сельскохозяйственных растений ориентировалась в основном на анализ генотипической и средовой изменчивости количественных значений ростовых признаков, измеренных на моменты окончания роста или уборки урожая. Классические биометрико-генетические подходы

(оценка числа полиморфных генов или блоков, схемы наследования, доли средовой изменчивости и т.п.) для анализа признака одного набора генотипов или популяции в разных условиях выращивания демонстрируют, как правило, пеструю лабильную картину (Драгавцев, Цильке и др., 1984; Мазер, Джинкс, 1985). В результате не удается выработать четкие рекомендации селекционеру.

Лишь в последние десятилетия формируется новое направление генетики количественных признаков. Оно предполагает, что эффективное управление изменчивостью растений на организменном, популяционном, ценотическом, экологическом уровнях невозможно без выявления сходства и различия процессов индивидуального роста и развития, без изучения эколого-генетической детерминации их изменчивости (Драгавцев, Утемишева, 1975; Драгавцев, Аверьянова, 1983; Смиряев, 1985; Шевелуха, 1992).

Как уже отмечалось, предложена модель эколого-генетического контроля признаков продуктивности растений, в которой вместо предположения о полигенном контроле вводится понятие лабильного спектра генов, влияющего на изменчивость признака и его составляющих в онтогенезе и экологических градиентах (Драгавцев, Утемишева, 1975). Модель позволяет объяснять и прогнозировать, хотя и в качественной форме, характеристики стабильности генотипов по количественным признакам, коэффициенты их генотипических корреляций, изменчивость генетико-статистических параметров и т.п. Однако для прогнозирования требуется достаточно полная информация о генетической специфике исследуемых генотипов и о конкретных лимитирующих факторах в полевом опыте по этапам онтогенеза (Лопатина и др., 1986). Подобная информация обычно недоступна. Кроме того, в этой модели упрощенная идея полигенной детерминации изменчивости количественного признака заменяется представлением о качественном переопределении признака формулой. Причем нормой представляется переключение спектра локусов, определяющих изменчивость признаков, в ответ на смену одного, главного лимитирующего фактора.

По моему мнению, следует учитывать не только главные лимиты среды, но и количественную реакцию экспрессии генов на изменения всей совокупности внешних условий. Кроме того, при построении генетико-статистических моделей и методов приходится принимать во внимание ограниченность современных знаний об изменчивости путей от генов до количественного признака (Корочкин, 2002). Поэтому остается ориентироваться в первую очередь на анализ варьирования признака не у каждого генотипа, а на исследование относительной изменчивости, т.е. на поиск сходства и различия по

норме, форме или хотя бы по направлению реакции изучаемого признака у совместно испытываемых генотипов в экологических градиентах и в онтогенезе.

Не отрицая возможностей традиционных генетико-статистических методов и моделей, в качестве дополнительного предлагается новый подход – биометрико-генетический анализ относительной онтогенетической изменчивости кривых роста количественного признака у группы генотипов в серии совместных полевых испытаний. Этот подход свободен от некоторых обременительных предположений о генетических механизмах и требований к исходной информации. В рамках такого подхода кривые роста признака рассматриваются как индикаторы для количественного анализа сходства и различия генетической детерминации процесса онтогенеза сравниваемых генотипов с учетом всей совокупности внешних условий, реализующихся в серии опытов.

Процесс или просто кривую роста трудно описать одним числом, как при измерении конечной величины признака. Известно, что в постоянных условиях выращивания, если не учитывать внутрисуточные колебания скорости роста признака, изменение (тренд) значений ростового признака во времени достаточно точно аппроксимирует S-образная кривая Сакса. Каждая кривая описывается несколькими константами (Шмидт, 1984). История изучения кривых роста признаков у биологических объектов насчитывает многие десятилетия (Kovallski, 1974), но пока не разработан биометрико-генетический анализ сходства и различия кривых у растений с учетом изменчивости условий полевого опыта за период их развития.

В реальных полевых условиях рост не всегда удается описать гладкой S-образной кривой. Существенные колебания условий за период вегетации могут неоднократно ускорять или замедлять рост признака, причем по-разному в зависимости от генотипа. Закономерности изменчивости кривой для каждого генотипа весьма сложны, так как на нее влияют колебания условий не только за период его роста, но и на начальных этапах онтогенеза, когда признака еще не существует. В результате проявляется «негладкость» кривой роста, а в серии испытаний генотипа в экологических градиентах – существенные ее изменения.

Однако можно выделить генотипы со сходной изменчивостью формы кривых роста признака в ответ на изменения совместных условий выращивания в полевых условиях: кривые роста у них меняются хотя и значительно, но более «синхронно», чем у генотипов, несходных в этом смысле. Степень сходства можно оценить количественно, вводя и анализируя специальные параметры

относительной изменчивости ростового признака для каждой пары генотипов. Различия и сходство таких реакций по конкретному признаку обусловлены проявлением в фенотипе отличающихся генетических программ детерминации роста и развития организма, что в свою очередь связано со степенью родства генотипов. Иначе говоря, ростовой признак можно рассматривать как индикатор не только вариабельности внешних условий, но и особенностей онтогенеза конкретного генотипа, а, значит, и генетического сходства совместно испытываемых генотипов.

Для развития этого направления генетики количественных признаков необходимо разработать четкие подходы к оценке различия и сходства онтогенетической изменчивости генотипов, а также соответствующие методические приемы их сравнительного биометрико-генетического анализа. Анализ должен включать количественные модели онтогенетической изменчивости признаков, просто и адекватно отражающие экологическую и генотипическую изменчивость для группы совместно испытываемых форм, числовые параметры сходства и различия онтогенетической изменчивости этих генотипов по отдельным ростовым признакам в разных средах с учетом родства генотипов, сходства аллельного состава и т.д. Методика должна быть пригодна для реализации на данных, которые могут быть получены в типичных полевых экспериментах сравнительного испытания группы форм с константными генотипами (сортами, гибридами, клонами и т.п.).

2.1. Гипотеза о сходстве относительной изменчивости признака

Сходство реакций сравниваемых генотипов по изменчивости ростового количественного признака в онтогенезе и сходство по конечным значениям этого признака в экологических градиентах (для разных лет, мест совместного выращивания, предшественников и т.п. на моменты уборки), по-видимому, взаимосвязано. Это, в частности, нашло отражение в гипотезе о сходстве (Смиряев, 1985, 1997). Она была сформулирована на основании результатов обработки обширных данных государственных испытания наборов новых сортов озимой пшеницы – по 8-10 и более в наборе (Смиряев, Гохман, 1985).

Было обнаружено, что у 75-80% и более сортов каждого набора форма (но не норма) реакций урожайностей на изменение условий по средам (2-3 года совместного испытания в 4-6 близких областях) сходна. Это проявилось в высокой степени адекватности так называемых регрессий на средние по Eberhart, Russell (1966), которые

используют для изучения взаимодействия «генотип-среда».

То есть строились линейные регрессии Y_{ik} – урожайностей каждого (i -го) сорта набора в 7-8 и более опытах серии на значения V_k – индекса условий совместного испытания набора. $V_k = (\sum_{i=1}^p Y_{ik})/p$ – результат простого усреднение p сортов набора по их урожайности в k -х условиях серии опытов совместного испытания набора.

$$\hat{Y}_{ik} = a_i + b_i V_k,$$

где $k=1, 2, \dots, E$, E – общее число серии опытов, обобщаемых в этой регрессиях Y_{ik} на V_k .

Адекватность линейной регрессии для i -го сорта означает, что отклонения Y_{ik} – реальных урожайностей сорта в серии опытов от \hat{Y}_{ik} – их оценок по такой регрессионной модели незначимо отличались от случайных ошибок опытов в повторениях.

В качестве примера приведем результаты госиспытания одного набора. Восемь сортов озимой пшеницы испытывали совместно на четырех госсортоучастках четырех областей (табл. 2.1): Волынской (по предшественнику клевер за оба года), Черниговской (по гороху в I год и по однолетним травам во II год), Ивано-Франковской (по занятому пару за оба года), Тернопольской (по гороху за оба года). Поскольку в этом варианте обобщают данные испытания из четырех областей за два года, то число значений индекса условий их совместного испытания (V_k – последняя строка в табл. 2.1) для построения регрессии: $E=2 \times 4=8$. Например, $V_5 = (26,0 + \dots + 29,2)/8 = 27,7$.

Далее проводили обычный расчет МНК-оценок коэффициентов линейных регрессий Y_{ik} на V_k . Результаты приведены в трех последних столбцах табл. 2.1. При 5%-й точности опытов в повторениях все восемь регрессий адекватны: табличный критерий Фишера равен 2,51, расчетный F_p – из табл. 2.1.

Высокий процент адекватных линейных регрессий на средние в десятках наборов не мог быть случайностью, так как обнаруживался стабильно, несмотря на то, что величина исследуемого признака сортов могла меняться по опытам каждой серии в 2-3 и более раз. Причем уровни отзывчивости сравниваемых сортов (норма реакции) на общие изменения условий совместного выращивания (V_k) в каждой серии опытов различались существенно. Последнее проявилось в значительном несходстве b_i – коэффициентов регрессий на средние у сортов (пример в столбце b_i табл. 2.1 и на рис. 2.1). Сопоставление

Таблица 2.1. Урожайность (ц/га) и параметры регрессий восьми сортов озимой пшеницы (Смиряев, Гохман, 1985)

Сорт	I. Волынская область		II. Черниговская область		III. Ивано-Франковская область		IV. Тернопольская область		a_i	b_i	F_{pi}
	I год	II год	I год	II год	I год	II год	I год	II год			
Белоцерковская 18	56,7	48,6	53,9	27,7	26,0	53,1	52,6	65,6	-0,54	1,05	2,50
Запорожская остистая	53,7	39,0	53,0	29,1	29,1	39,3	43,6	63,0	1,29	0,92	1,49
Ильичевка	54,6	45,6	60,3	28,9	24,9	45,7	50,2	62,9	-1,81	1,05	1,30
Ивановская 12	55,5	42,0	54,6	18,5	28,2	42,2	45,2	67,3	-10,60	1,19	2,09
Киянка	65,0	44,4	61,2	27,7	32,1	44,1	51,3	63,1	-0,63	1,07	2,28
Мироновская 808	53,1	46,8	48,4	35,1	28,3	48,1	40,2	54,4	14,10	0,66	2,37
Мироновская 25	58,6	44,5	58,8	32,4	23,9	49,4	45,5	61,7	-0,09	1,02	1,66
Полесская 70	59,8	44,2	53,9	26,7	29,2	43,2	46,4	64,4	-1,77	1,04	0,63
V_k	57,1	44,4	55,5	28,3	27,7	45,6	46,9	62,8			

родословных новых сортов, поступавших в госиспытание, обычно выявляло высокую степень их родства. То есть сходство формы реакции, отразившееся в частой адекватности регрессий на средние, а также родство сортов не были связаны со сходством по норме реакции на изменения общих условий выращивания.

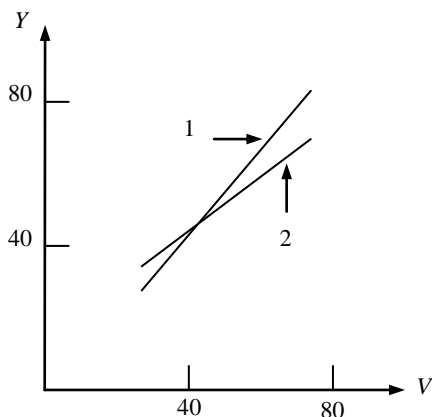


Рисунок 2.1. Линейные регрессии на средние для урожайности (ц/га) двух сортов контрастных по норме реакции (см. b_i из табл. 2.1): Ивановская 12 (1) и Мироновская 808 (2).

Можно показать (Смиряев, Гохман, 1985), что адекватность регрессий на средние (не обязательно линейных) для группы сортов означает, что для каждой их пары (X, Y) в серии опытов адекватна модель:

$$Y_k = f(X_k) + e,$$

где f — гладкая кривая, описывающая относительную изменчивость признака в экологических градиентах (k — года, места испытания и пр.) и специфичная для данной пары сортов; e — случайные ошибки опытов в повторениях.

Обратное также справедливо: адекватность регрессий f (например, линейных), аппроксимирующих зависимости для каждой пары из группы сортов, означает адекватность линейных регрессии на их средние по данным той же серии совместных испытаний.

Гипотеза состояла в том, что причина адекватности регрессий f в экологических градиентах, т.е. по конечным значениям признака, для пары сортов заключается в сходстве относительной изменчивости этого

признака у такой пары в онтогенезе. По-видимому, для пары таких сортов кривые (тренд) роста признака у одного сорта $[X(t)]$ относительно второго $[Y(t)]$ подобны в серии опытов. То есть эти кривые относительного роста у двух сортов в онтогенезе, во-первых, достаточно гладкие, без «изломов», хотя каждая кривая роста во времени $[X(t)]$ и $[Y(t)]$ по отдельности, как уже отмечалось, возможно, негладкая. Во-вторых, относительные кривые роста одной пары сортов «накладываются» друг на друга для разных опытов серии сравнительных испытаний, обобщаемых в регрессии f . Упрощенная иллюстрация для 3-х лет совместного испытания на рис. 2.2 а.

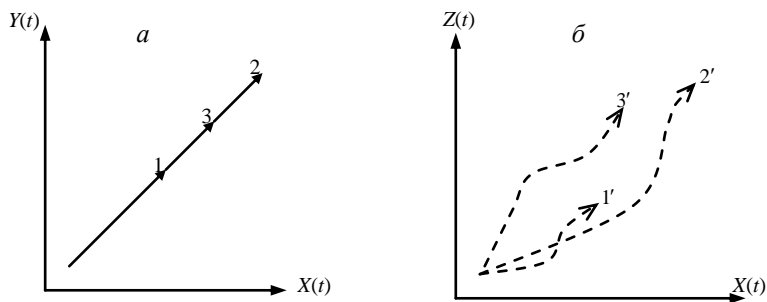


Рисунок 2.2. Гипотеза о сходстве относительной изменчивости. $Y(t)=f[X(t)]$ на рис. 2.2 а – три кривые относительной онтогенетической изменчивости признака для пары сходных по форме реакции сортов: их кривые «наложились» за три года совместного испытания. Стрелки на рис. – направление роста признака. 1-3 – значения ростового признака этих сортов на моменты уборки за эти три года. Пунктиры и точки 1' - 3' на рис. 2.2 б – то же для двух несходных сортов.

Условия совместного выращивания меняют лишь скорость «скольжения» координат $[X(t), Y(t)]$ признака по f – единой эколого-онтогенетической кривой относительной изменчивости признака, а также точку «остановки» роста признака на ней (1, 2, 3 на рис. 2.2 а) в моменты завершения роста или уборки урожая в разные годы серии испытаний. Только в подобных случаях по данным, полученным в разные годы и в разных местах совместного испытания, можно строить адекватные парные регрессии f , а также адекватные регрессии на средние ($\hat{Y}_{ik} = a_i + b_i V_k$), для целой группы сходных сортов набора.

Напротив, для двух несходных сортов невозможно объяснить регулярное «возвращение» на моменты уборки в разные годы на какую-либо одну гладкую кривую f «негладких» в онтогенезе и расходящихся по годам кривых (пунктиры на рис. 2.2. б) относительного роста пары

сортов. Для такой пары получим неадекватную регрессию f , построенную по конечным значениям признака в экологических градиентах. При попытке обобщить в моделях регрессии на средние данных по достаточно большой группе сходных сортов и, например, по одному несходному с ними проявляется неадекватность регрессии на средние для последнего. То есть координаты «остановки» роста признака такого сорта на моменты уборки в серии совместных испытаний формируют совокупность точек, по которым не удастся построить адекватную регрессию на средние.

Важно подчеркнуть, что гипотеза сформулирована для серии совместных полевых испытаний. В каждом испытании условия постоянно и в известной степени случайно меняются во времени, что влияет на ход кривых роста признака во времени, формируя в каждом испытании по годам (местам) новые кривые. В контролируемых условиях, гипотеза не всегда справедлива. Например, в фитотроне, если по опытам серии совместных испытаний лишь один параметр среды меняется с определенным шагом, то адекватность регрессий, построенных по данным о конечных значениях признака, не обязательно связана со сходством формы онтогенетической изменчивости, определенным выше.

Гладкость и наложение кривых относительной изменчивости для пары сортов – два взаимосвязанных фенотипических проявления их сходства по форме онтогенетической изменчивости ростового признака – индикатора сходства сортов. В соответствии с гипотезой именно это отражается в определенной выше адекватности регрессионных моделей, построенных по конечным значениям признака в экологических градиентах (годы, места совместных испытаний и т.п.). Естественно, правильнее говорить не о точном «наложении» и абсолютной гладкости, а о статистически проявляющейся тенденции сходства, приводящей к адекватности регрессий f на фоне случайных ошибок опытов.

Для серий полевых опытов сходные в этом смысле сорта, видимому, чаще близкородственны. Родственность в свою очередь связана со сходством их аллельного состава, что, однако, не соответствует прямо сходству изменчивости сортов по конкретному количественному признаку.

Если гипотеза справедлива, то, например, по наклону кривой f , полученной по 2-3-летним данным для близкородственных сортов можно предварительно сравнивать их стабильность (норму реакции) в экологических градиентах.

2.2. Параметры и количественный анализ степени сходства генотипов

Необходима формализация и проверка описанных подходов. Для оценки степени сходства двух генотипов по онтогенетической изменчивости ростового признака наиболее естественными являются модели регрессионного анализа. Так, гладкость относительной парной кривой роста $Y_k(t) = f[X_k(t)] + e$ в одних условиях испытания (при фиксированном значении k) означает, что ее удобно аппроксимировать полиномом степени n :

$$\hat{Y}(t) = a + \sum_{i=1}^n b_i X^i(t) \quad (2.1).$$

Константы a , b_i ($i=1, 2, \dots, n$) оцениваются методом взвешенных наименьших квадратов по опытным данным – m измерениям признака $X(t_j)$, $Y(t_j)$ у двух генотипов в одни моменты времени t_j ($j=1, 2, \dots, m$).

Однако адекватность модели (2.1) не всегда означает гладкость кривой f . Первая причина такого несоответствия – статистическая ограниченность объема выборки пар данных $\{X(t_j), Y(t_j)\}$, полученных в период роста признака в одном опыте серии. Известно, что любую совокупность точек $\{X(t_j), Y(t_j)\}$ объема m на плоскости можно соединить «без отклонений» гладкой кривой, описываемой полиномом (2.1) степени $n=m-1$. Полиномы степени, близкой к $(m-1)$, несут мало информации о гладкости самой кривой f .

Вторая причина – биологическая. Особенности реакций в онтогенезе на одинаковые изменения внешних условий у двух генотипов с несходными программами генетической детерминации роста признака могут вызвать во времени постепенный «плавный» отход парной кривой от простейшей линейной зависимости:

$$Y(t) = a + bX(t) \quad (2.2).$$

Это в свою очередь отразится в высокой адекватности гладкой аппроксимации (2.1), например, второго или третьего порядка ($n = 2, 3$). Конечно, чем более несходны генетические программы 2-х генотипов, реализовавшиеся в данных условиях выращивания, тем, при достаточном объеме выборки (m), выше минимальная степень n адекватной регрессионной зависимости (2.1). Однако многое зависит от этих условий конкретного полевого опыта. В отдельные годы они могут сформировать гладкие кривые у пары генотипов, т.е. не вызвать существенного несходства реакций по форме роста признака у двух генетически несходных генотипов.

Вероятность первой ошибки в выводах о гладкости кривых f можно было бы снизить, значительно увеличив объем выборки измерений в онтогенезе (m), что затруднительно, а также ограничившись для реального объема (m от 6 до 20) аппроксимацией кривой f полиномом (2.1) второго порядка:

$$Y(t) = a + bX(t) + cX^2(t) \quad (2.3).$$

Ошибка по второй причине можно компенсировать сравнением кривых f , получив их аппроксимации по данным из разных условия совместного выращивания двух генотипов, например, за несколько лет и (или) в нескольких местах одной агроклиматической зоны. Сопоставление таких кривых в общей системе координат $X(t)$, $Y(t)$ позволит оценить, кроме гладкости, степень их расхождения по годам, что в соответствие с предложенным подходом является вторым проявлением несходства генотипов по форме реакции в онтогенезе. Меняя годы и места проведения опытов совместного испытания, обобщаемых в одной регрессионной модели, можно подойти к оценке стабильности проявления сходства аллельного состава пары генотипов в конкретном признаке – индикаторе сходства.

Для решения поставленных задач необходим статистический анализ, как степени гладкости, так и расхождения. При этом не следует ограничиваться выводом об адекватности или неадекватности единой регрессии (2.1), если построить ее по всем данным о росте признака пары генотипов при их совместном испытании в нескольких средах. Необходим количественный анализ степени несходства генотипов в рамках статистических моделей.

Применимы следующие аппроксимирующие модели регрессионного анализа для того, чтобы ввести и анализировать количественные параметры несходства реакций пары генотипов. Первая – с коэффициентами, меняющимися по k – опытам серии:

$$Y(t_{jk}) = \alpha_k + \beta_k X(t_{jk}) + \gamma_k X^2(t_{jk}) + \delta_{jk} + e \quad (2.4).$$

Здесь $Y(t_{jk})$, $X(t_{jk})$ – величины признака у двух генотипов при j -м их измерении в онтогенезе в k -х условиях совместного выращивания (всего измерений m); α_k , β_k , γ_k – коэффициенты квадратичной регрессии $Y(t)$ на $X(t)$ для k -х условий. Оценки этих трех коэффициентов с помощью метода взвешенных наименьших квадратов для каждой (k -х) условий совместного испытания далее обозначены a_k , b_k , c_k . Эффекты δ_{jk} – отклонения от регрессионной модели, связанные с «негладкостью» парной кривой f в k -х условиях выращивания, e - ошибки опыта в

повторениях. Дисперсии последних могут существенно различаться по условиям испытания и датам измерений, а также в зависимости от самой величины признака (X и Y) у двух генотипов. Именно поэтому для анализа следует использовать метод взвешенных наименьших квадратов. Всего число моделей (2.4) для одной пары генотипов равно числу условий их совместного выращивания E ($k = 1, 2, \dots, E$).

Дополнительно вводится еще одна модель относительного роста признака 2-х генотипов, единая по условиям выращивания, реализовавшимся в серии опытов:

$$Y(t_{jk}) = \alpha + \beta X(t_{jk}) + \gamma X^2(t_{jk}) + \delta'_{jk} + e \quad (2.5).$$

Отклонения δ'_{jk} обычно больше δ_{jk} из модели (2.4), так как в модели (2.5) α, β, γ одинаковы для всех E условий выращивания.

Можно разложить SS_0 – сумму квадратов отклонений δ'_{jk} (число степеней свободы SS_0 равно $df_0 = mE - 3$) на две составляющие. Первая $SS_{нт}$ с $df_1 = E(m - 3)$ – сумма квадратов отклонений δ_{jk} моделей (2.4), обусловленная «негладкостью» кривых $Y(t) = f[X(t)]$, суммарной по всем E условиям выращивания. Вторая составляющая SS_p с $df_2 = df_0 - df_1 = 3(E - 1)$ – обусловлена расхождением этих кривых по E условиям выращивания. Если для каждого опыта (k) объема выборки измерений в онтогенезе (m) недостаточно для построения квадратичных моделей (2.4), например $m < 6$, то можно ограничиться линейной регрессией, т.е. принять $\gamma_k = 0, \gamma = 0$. Тогда $df_0 = mE - 2, df_1 = E(m - 2), df_2 = 2(E - 1)$.

Сопоставление соответствующих средних квадратов $ms_0, ms_{нт}, ms_p$ для одной пары сортов между собой и со средними квадратами эффектов e – ошибок опытов в повторениях опытов позволяет проверять различные гипотезы об адекватности моделей (2.4) и (2.5). Сравнение однотипных ms (или SS) для разных пар сортов позволяет количественно оценить количественно степень парного сходства сортов по форме онтогенетической изменчивости ростового признака-индикатора.

Возвращаясь к гипотезе о сходстве относительной изменчивости, можно предложить измененную формулировку гипотезы, пригодную для ее прямой проверки на основании данных по группе сортов, совместно испытанных в M условиях выращивания. Предполагается, что ростовой признак измерен во всех M условиях выращивания на моменты окончания роста (в экологических градиентах) и, дополнительно, хотя бы в некоторых $E < M$ условиях по несколько раз в онтогенезе одновременно у всех сортов набора. Если гипотеза верна, то

пары сортов с более сходной реакцией по конечным значениям признака (по данным в экологических градиентах), т.е. с меньшими SS , – суммами квадратов отклонений δ_i , от парных регрессий:

$$Y_l = \alpha + \beta X_l + \gamma X_l^2 + \delta_l + e \quad (l=1, 2, \dots, M) \quad (2.6)$$

отличаются и большим сходством онтогенетических реакций, т.е. меньшими отклонениями (меньшими SS_p) от парных регрессий по данным о росте признака в онтогенезе, полученным в E условиях. Последнее означает большую гладкость относительных парных кривых (меньшие $SS_{нг}$ в моделях (2.4)) и меньшее расхождение этих кривых по E условиям выращивания. Как отмечалось выше, SS_p – степень расхождения можно оценить путем вычитания: $SS_0 - SS_{нг}$. То есть следует вычесть суммарные по E средам значения $SS_{нг}$ отклонений от регрессий (2.4) этой пары сортов из SS_0 для их единой регрессии (2.5), построенной по всем измерениям признака в онтогенезе у пары сортов, полученным в E средах. Всего для набора из p совместно испытанных сортов получим три ряда SS_0 , $SS_{нг}$, SS_p по $p(p-1)/2$ чисел в каждом ряду, что соответствует всем возможным парам сортов в этом наборе.

Для прямой проверки гипотезы о сходстве следует оценить коэффициенты корреляции (r) между тремя названными рядами SS . При этом следует учесть, что вычисленные таким образом $r(нг, \varepsilon)$, $r(p, \varepsilon)$ и $r(нг, p)$ – фенотипические оценки с ошибками, обусловленными ограниченностью выборки измерений признака в период онтогенеза и выборки сред совместного испытания. Поэтому, например, все параметры $ms=SS/df$ существенно отличаются от истинных генотипических значения дисперсий (σ_G^2) отклонений от регрессии каждой пары сортов. σ_G^2 могут быть получены практически без ошибок лишь при большой выборке M – условий совместного выращивания, E , а также числа измерений признака в онтогенезе для каждого из E условий. Желательно учесть эти ошибки, чтобы хотя бы приблизительно оценить r_G – генотипические коэффициенты корреляции любых двух рядов чисел SS (обозначим SS_1, SS_2). Для этого можно воспользоваться некоторыми фактами из теории регрессионного анализа.

Известно, что случайная величина SS/σ_G^2 имеет χ^2 -распределение со степенями свободы суммы квадратов отклонений SS (Демиденко, 1981). Тогда несложно показать, во-первых, что cov_p – фенотипическая ковариация между двумя рядами значений SS_1 и SS_2 даст несмещенную

оценку cov_G – генотипической ковариации между двумя рядами величин $df_1 \sigma_{G1}^2$ и $df_2 \sigma_{G2}^2$. То есть можно принять:

$$cov_p(SS_1, SS_2) = cov_G(df_1 \sigma_{G1}^2, df_2 \sigma_{G2}^2).$$

Учитываем далее, что фенотипический коэффициент корреляции двух рядов SS_1 и SS_2 равен $r_p = cov_p / \sqrt{\sigma_1^2 \sigma_2^2}$, где σ_1^2, σ_2^2 – фенотипические дисперсии рядов SS_1 и SS_2 . Отсюда ясно, что оценку r_G можно получить по следующей формуле:

$$r_G = r_p / \sqrt{H_1^2 H_2^2},$$

где H_1^2 – отношение дисперсии первого ряда значений $df_1 \sigma_{G1}^2$ для всех пар сортов к σ_1^2 – дисперсии значений SS_1 этих пар. Аналогичен смысл H_2^2 для второго ряда.

Во-вторых, с учетом свойств распределения χ^2 можно показать, что для первого и второго ряда дисперсии значений $df \sigma_G^2$ (по всем парам сортов) меньше соответствующих дисперсий значений SS . Поэтому H_1^2 и H_2^2 меньше единицы, причем величины H^2 в основном зависят от df для соответствующих значений SS .

Оценки показывают, что при $df=7$ $H^2 \approx 0,63$, при $df=6$ $H^2 \approx 0,59$, при $df=5$ $H^2 \approx 0,52$, при $df=4$ $H^2 \approx 0,42$, при $df=3$ $H^2 \approx 0,34$, при $df=2$ $H^2 \approx 0,26$. Эти величины и последняя формула перехода от r_p к r_G позволяют оценивать доверительные интервалы оценок r , что необходимо для проверки гипотезы о сходстве относительной изменчивости признака.

При построении и использовании регрессионных моделей (2.4), (2.5) следует учитывать два осложнения. Первое связано со скоррелированностью e – случайных ошибок опыта для разных дат измерения признака каждого сорта, а значит и их пары, в одних условиях выращивания (т.н. автокорреляции e). Второе – с тем, что случайные ошибки есть у измерений обоих сортов пары (X и Y). В результате способы оценки ms и коэффициентов регрессий (2.4), (2.5) значительно сложнее обычных ситуаций, когда ошибки не коррелируют между собой и считаются равными нулю у так называемых независимых переменных, стоящих в правых частях регрессионных моделей.

Следует отдельно пояснить связь отклонений от моделей (2.4), (2.5) для пар сортов с адекватностью регрессии на средние для их группы по тому же признаку. Если для любой пары из группы, включающей D сортов, адекватна модель (2.5), т.е. в результате проверки принимается гипотеза $\delta'_{jk}=0$ (значит и $\delta_{jk}=0$) для всех j и k , то для каждого q -го сорта этой группы адекватна модель регрессии на средние для всей серии опытов по данным, полученным в онтогенезе:

$$Z_q(t_{jk}) = \tau_q + \omega_q V(t_{jk}) + \varepsilon_q V^2(t_{jk}) + \delta''_{qjk} + e \quad (2.7).$$

Здесь $Z_q(t_{jk})$ – величина ростового признака у q -го сорта группы, полученная при j -м его измерении в k -х совместных условиях выращивания. $V(t_{jk}) = \sum_{q=1}^D Z_q(t_{jk}) / D$ – средние значения признака всех сортов группы при j -м измерении в k -х условиях; τ_q , ω_q , ε_q – коэффициенты, подбираемые методом взвешенных наименьших квадратов в предположении малых ошибок у средних V , но с учетом автокорреляций e для одних условий испытания набора сортов. Адекватность модели (2.7) означает, что принимается $\delta''_{qjk}=0$ для всех вариантов j, k серии обобщаемых опытов.

Обратное также справедливо: если в модели (2.7) для группы сортов все $\delta''_{qjk}=0$, то δ'_{jk} и δ_{jk} в моделях (2.4) и (2.5) также равны нулю для каждой пары сортов группы. Конечно, средние квадраты (ms) отклонений δ''_{qjk} , δ'_{jk} и δ_{jk} не бывают в точности равны нулю. Причины этого – всегда имеющиеся, хотя бы небольшие, «негладкости» и расхождение по условиям испытаний кривых относительной изменчивости, т.е. $X(t)$ относительно $Y(t)$ или $Z(t)$ относительно $V(t)$. В последнем случае причина неадекватности в том, что общие средние значения $V(t)$ в модели (2.7) не для каждого сорта набора удовлетворительно отражают форму реакции ростового признака на изменение условий совместного выращивания в серии опытов.

В рамках модели (2.7) оценивают не парные показатели сходства (SS_{ij}) сортов i и j , а SS_q – сходство каждого q -го сорта с единым по группе сортов показателем $V(t)$ – индикатором их усредненной реакции. Величина SS_q для сорта, несходного с большинством остальных, сходных между собой по форме реакции, значительно превосходит аналогичные показатели для большинства. На графиках роста признака это должно проявляться в «негладкости» кривых $Z_q(t)$ относительно $V(t)$

для отдельных опытов серии и(или) расхождением кривых по этим опытам.

Именно SS – суммы квадратов (или ms – средние квадраты) отклонений предлагается использовать в качестве основных количественных параметров при анализе сходства сортов по форме их реакции как в онтогенезе, так и с учетом экологической изменчивости признака по условиям совместных полевых испытаний серии.

2.3. Проверка гипотезы – анализ сходства сортов озимой пшеницы

Для проверки гипотезы о сходстве относительной изменчивости признака в экологических градиентах и онтогенезе использованы данные, полученные в многолетнем опыте с 22 сортами мягкой озимой пшеницы (Смиряев и др., 1997). Сорта были районированы в разное время в Центральных районах Нечерноземной зоны России. Они были отнесены к трем группам по периодам сортосмены (табл. 2.2).

Таблица 2.2. Названия и годы районирования сортов озимой пшеницы

I группа			II группа			III группа		
№	сорт	год	№	сорт	год	№	сорт	год
1*	Эритроспермум 917	1929	10	Лютесценс 116	1944	16*	Мироновская 808	1963
2	Дюрабль	1929	11*	ППГ-599	1948	17*	Ахтырчанка	1978
3	Московская 2453	1929	12*	Лютесценс 266	1950	18*	Заря	1978
4	Ферругинеум 1239	1933	13*	ППГ-186	1953	19*	Нива	1981
5*	Сандомирки местные	1934	14	Мильтурум 513	1955	20*	Звезда	1992
6	Вязниковская местная	1938	15*	Кунцевская 45	1960	21*	Московская низкостебельная	1992
7*	Мильтурум перерод	1939				22*	Инна	1992
8*	Глебовская местная	1940						
9	Ульяновка	1942						

* Сорта, участвовавшие в оценках коэффициентов родства.

В частности, сопоставляли трехлетние данные об изменчивости высоты растений этих сортов в онтогенезе и семилетние результаты измерения того же признака на моменты уборки. При сопоставлении этих данных появилась возможность прямой проверки гипотезы. Попытались также установить взаимосвязь параметров, предложенных

для оценки сходства формы онтогенетической изменчивости сортов, по высоте растений, длине колоса и площади флагового листа, с известными характеристиками генетического сходства сортов. Для этого использованы коэффициенты родства, оцененные по родословным сортам, и сведения об аллельном составе локусов, кодирующих запасные белки зерна, а также о синхронности наступления фенофаз сортов по 6-летним данным.

Методика получения данных

Многолетний эксперимент с 22 сортами проводился на опытном поле РГАУ-МСХА им. К.А. Тимирязева. Опыт закладывался в трех повторностях методом рандомизированных блоков. Площадь делянок составляла 5 м^2 (5×1). В ходе эксперимента измеряли три ростовых признака.

В 1992-1994 гг. с момента вступления в фазу кушения в каждой повторности каждые пять дней одновременно измерялась средняя высота 20 случайно взятых растений каждого сорта. В 1992 г. проведено 10 измерений, в 1993 и 1994 гг. – по 9 измерений. С момента начала фазы выхода в трубку каждые пять дней измерялась длина конуса нарастания (20 растений в каждом повторении), а с фазы колошения – длина колоса. В 1992 г. удалось получить 7, а в 1994 г. – 8 измерений этого признака, который далее именуется «длина колоса». В 1993 г. из-за плохих условий перезимовки не было достаточного количества материала для получения полной выборки измерений длины конуса нарастания. Поэтому данные по длине колоса за 1993 г. при анализе не использовались. С фазы колошения каждые пять дней в каждой повторности измерялась длина (d) и ширина (b) флагового листа у 20 растений. После чего производился расчет площади флагового листа по формуле: $S=0,67d \times b$. В 1993 г. получено 6 измерений, в 1994 г. – 5 измерений. В 1992 г. из-за опоздания с началом измерений данные по нарастанию площади флагового листа не были получены. Для каждого признака, измерения и сорта использовали данные в каждой повторности, усредненные по 20 растениям.

Для проверки гипотезы о сходстве относительной изменчивости были использованы 7-летние данные (1989-1995 гг.) по первым 20 сортам по конечной высоте растений, т.е. на моменты уборки. Из-за различий погодных условий за три года конечные величины изученных количественных признаков существенно различались – для отдельных сортов максимум на 20-30%, а за семь лет по высоте растений максимум на 40%.

Для оценки связи парного сходства изменчивости сортов по признакам-индикаторам и сходства их по срокам прохождения фенофаз использованы данные за 1990-1995 гг. по датам кущения, выхода в трубку, колошения, цветения и молочной спелости каждого сорта по отдельности. Эти данные за каждый год и по каждому сорту были преобразованы в число дней – продолжительность от посева до наступления соответствующей фенофазы. Полученные 13 чисел для каждого сорта (в некоторые годы данные снимались не по всем фенофазам) можно рассматривать как измерения количественного признака, связанного с особенностями онтогенеза этого сорта, так как сроки наступления одной фенофазы различались по сортам иногда на 5-7 суток. Поэтому для общей количественной оценки синхронности прохождения фенофаз у пары сортов (i, j) по 13 парам чисел определялась FS_{ij} – сумма квадратов их разностей (по продолжительностям).

Регрессионный и кластерный анализ сходства сортов

Данные об изменчивости трех ростовых признаков обрабатывались с помощью специальных модификаций регрессионного и корреляционного анализа, изложенных выше, а также с использованием кластерного анализа. Кроме того, ряды величин SS_{ij} сопоставлялись с рядами FS_{ij} и с двумя рядами генетико-статистических показателей: R_{ij} – коэффициентов родства сортов, вычисленных по их родословным (Смиряев и др., 1992), и с I_{ij} – показателей Нея (Nei, 1972), характеризующих парное сходство сортов по аллельному составу локусов, кодирующих запасные белки зерна.

Адекватность регрессионной модели проверяли по F -критерию Фишера. Расчетное значение F_p равно отношению ms отклонений от модели к ms_e случайных ошибок опытов. Результаты построения квадратичных регрессий на средние для 22 сортов представлены в табл. 2.3а – для высоты растений по трехлетним данным и в табл. 2.3б – для длины колоса и площади флагового листа по двухлетним данным. Для последнего признака здесь и далее использованы линейные регрессии ($\varepsilon_q=0$ в уравнении (2.7), $\gamma_k=0$ в (2.4), $\gamma=0$ в (2.5)), так как за эти два года получено всего по 5-6 измерений в период роста площади флагового листа.

В табл. 2.3а приведены a, b, c – оценки коэффициентов $\tau, \omega, \varepsilon$ регрессии (2.7) и F_p – значения расчетного критерия Фишера, полученные без учета автокорреляций ошибок e , а F'_p – с учетом автокорреляций по Rao (1959). В последнем случае оценки a, b, c приведены в скобках и мало отличаются от оценок коэффициентов

Таблица 2.3. Регрессионный и кластерный анализ 22 сортов озимой пшеницы (Смиряев и др., 1997).

а) высота растений по трехлетним данным (28 измерений за три года по каждому сорту)

Сорт	<i>a</i>	<i>b</i>	<i>c</i> ×1000	F_p	F'_p	$F'_{гр}$	$F'_{кл}$	№
1	4,17 (4,42)	0,80 (0,81)	2,08 (2,07)	1,13	1,67	1,47	1,58	(3)
2	4,27 (4,28)	0,71 (0,70)	3,13 (3,08)	1,33	1,59	1,47	1,85	(3)
3	5,14 (5,14)	0,71 (0,71)	3,21 (3,20)	1,09	1,64	1,36	-	-
4	3,41 (3,41)	0,85 (0,85)	2,55 (2,54)	0,92	0,97	0,99	1,09	(3)
5	-0,20 (-0,20)	1,12 (1,11)	-0,17 (-0,17)	3,01	3,43	3,26	-	-
6	5,27 (6,00)	0,76 (0,76)	3,03 (3,06)	0,93	1,25	1,22	1,39	(3)
7	1,58 (1,34)	0,95 (0,95)	1,65 (1,68)	0,71	1,78	1,41	1,61	(3)
8	3,57 (4,55)	0,74 (0,68)	3,02 (3,59)	1,78	2,28	2,08	0,76	(1)
9	-1,35 (-1,59)	0,99 (0,99)	0,41 (0,40)	1,33	0,90	0,83	0,97	(3)
10	-2,11 (-2,85)	1,09 (1,08)	-0,30 (-0,24)	0,98	1,99	1,67	0,44	(1)
11	-0,63 (-0,63)	0,87 (0,87)	1,81 (1,81)	1,93	2,87	2,51	0,77	(1)
12	-2,98 (-2,74)	1,17 (1,17)	-0,67 (-0,59)	1,69	2,69	1,54	-	-
13	-5,48 (-5,48)	1,30 (1,30)	-2,54 (-2,54)	2,33	3,11	2,86	0,59	(2)
14	-2,34 (-2,35)	1,09 (1,08)	-0,22 (-0,28)	1,24	2,46	1,81	-	-
15	-5,11 (-4,86)	1,24 (1,24)	-2,04 (-2,11)	1,41	1,96	1,04	0,42	(2)
16	-3,26 (-3,50)	1,16 (1,11)	-1,80 (-1,94)	1,28	2,37	2,45	0,46	(2)
17	-4,93 (-5,18)	1,27 (1,26)	-3,39 (-3,43)	0,88	2,24	1,17	1,89	(3)
18	-3,17 (-2,92)	1,19 (1,19)	-2,52 (-2,46)	0,50	1,91	1,08	1,41	(3)
19	-0,60 (-0,38)	1,11 (1,12)	-2,25 (-2,19)	1,42	1,27	0,96	1,04	(3)
20	-3,41 (-3,41)	0,95 (0,95)	-1,21 (-1,21)	0,68	1,11	0,68	0,90	(3)

Продолжение табл. 2.3а

Сорт	a	b	$c \times 1000$	F_p	F'_p	$F'_{гр}$	$F'_{кл}$	№
21	2,05 (2,29)	0,98 (0,98)	-2,05 (-2,08)	1,37	1,06	0,91	0,89	(3)
22	4,90 (5,15)	0,94 (0,93)	-1,69 (-1,43)	0,99	1,11	0,78	0,74	(3)
$F_{т,5\%}$				1,74	2,15	2,15	2,15	

б) длина колоса (15 измерений) и площадь флагового листа (11 измерений) по двухлетним данным

Сорт	Длина колоса				Площадь флагового листа			
	F'_p	$F'_{гр}$	$F'_{кл}$	№	F'_p	$F'_{гр}$	$F'_{кл}$	№
1	4,31	2,53	1,01	(1)	6,41	5,46	-	
2	6,29	5,56	-		6,89	5,20	-	
3	2,08	1,57	0,78	(3)	13,3	10,5	-	
4	5,41	5,21	-		5,17	5,86	-	
5	4,30	4,70	-		4,91	5,81	0,79	(2)
6	1,59	1,63	0,47	(3)	1,37	2,95	1,05	(1)
7	1,74	1,58	1,43	(3)	2,18	2,04	1,09	(1)
8	0,85	0,81	1,16	(3)	3,18	2,58	-	
9	1,49	1,19	0,92	(1)	0,99	1,01	1,37	(3)
10	4,88	3,96	-		2,06	1,17	0,99	(3)
12	2,71	2,23	1,55	(3)	3,11	3,52	-	
13	9,50	8,20	-		3,84	2,56	1,65	(3)
14	2,35	2,19	-		8,34	7,58	-	
15	3,81	4,42	1,50	(2)	2,68	1,97	2,18	(3)
16	2,72	1,34	1,23	(2)	0,96	0,97	0,53	(3)
17	5,94	3,96	-		3,03	2,52	-	
18	3,31	3,29	-		0,59	0,52	0,49	(3)
19	1,71	0,39	0,24	(2)	1,81	1,47	0,76	(1)
20	3,69	4,44	1,44		1,30	0,99	1,02	(2)
21	1,80	0,73	0,68		1,28	0,92	0,68	(2)
22	1,94	2,33	1,80		2,82	1,74	1,58	(3)
$F_{т,5\%}$	2,09	2,09	2,09		2,21	2,21	2,21	

Примечание. a , b , c – оценки параметров τ , ω , ε регрессии (2.7) для каждого сорта. Номера сортов из табл. 2.2. № – номера трех наибольших кластеров сортов (для высоты растений – см. рис. 2.4). $F_{т,5\%}$ – табличное значение критерия Фишера для $P < 0,05$.

регрессий (2.7) без учета автокорреляций ошибок e . Однако, оценки коэффициентов автокорреляций, полученные по данным опытов в

повторениях, были больше нуля. Поэтому параметры F'_p для высоты в среднем по сортам на 44% больше F_p , для длины колоса на 55% больше, для площади флагового листа – на 26%.

В результате расчета выяснилось, что различия F_p и F'_p при двух способах построения регрессий (2.7) особенно сильны по однолетним данным. Так, в регрессиях (2.7) по высоте растений для десяти измерений 1992 г. среднее по сортам значение критерия F_p равно 0,33, а среднее по сортам значение $F'_p=1,56$. Величины F_p явно занижены, так как по смыслу расчетного критерия Фишера среднее его значение для набора сортов не может быть меньше 1. По трехлетним данным (табл. 2.3а) контрасты невелики из-за сильного увеличения F_p для многих сортов, связанного с расхождением их онтогенетических траекторий высоты по годам. Таким образом, при оценке средних квадратов отклонений и расчетных критериев Фишера следует учитывать автокорреляции.

В качестве примера исходных данных для построения регрессий на средние (2.7) на рис. 2.3 приведены значения (усредненные по трем

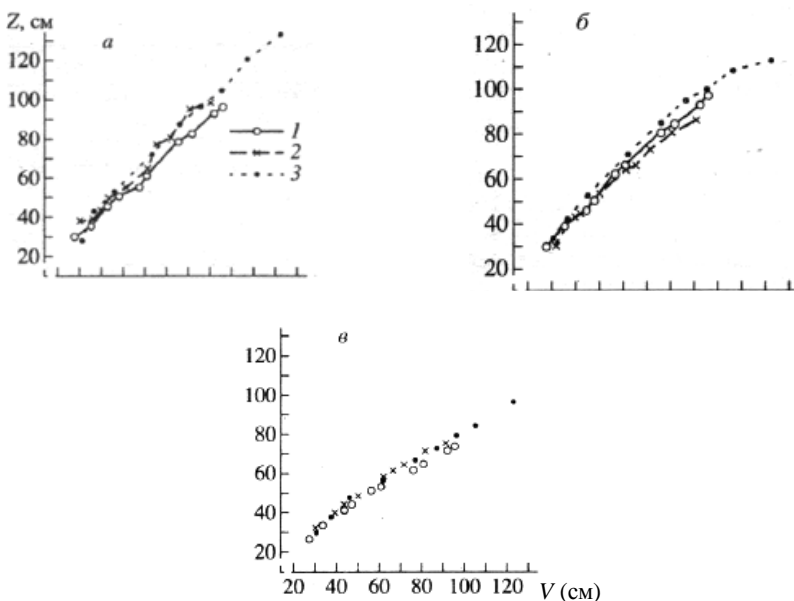


Рисунок 2.3. Кривые роста признака «высота растений» (ось ординат, Z) у сортов: *а* – сорт № 5; *б* – сорт № 13; *в* – сорт № 22 за 3 года в зависимости от величины V – среднего значения высоты во всем наборе 22 сортов. 1 – 1992 г., 2 – 1993 г., 3 – 1994 г

повторениям) десяти измерений высоты растений в 1992 и по девять измерений в 1993 и 1994 гг. для сортов с номерами 5, 13 и 22 (номера сортов – из табл. 2.2).

Если три внутригодовые кривые роста сорта 13 (относительно средних V по всем 22 сортам) заметно расходятся, то для сорта 22 такие расхождения по годам значительно меньше. Кривая роста высоты относительно V для сорта 5 на рис. 2.3а за 1993 г. заметно изломана – негладкая. Следует напомнить, что для оценки параметров регрессий использован метод взвешенных наименьших квадратов. Поэтому большее увеличение ms – средних квадратов отклонений вызывают отличия признака от оценки по регрессии в диапазоне с наименьшими случайными ошибками опытов в повторениях. Такие ошибки обычно проявляются при малых средних значениях индекса условий выращивания (V), в том числе по высоте растений. В результате изломы и некоторое расхождение траекторий на первых этапах роста высоты, как у сорта № 5, резко увеличивают F'_p .

Разная степень негладкости и расхождения отражается в различиях SS – сумм квадратов отклонений ($\delta_{ijk}'' + e$) в модели (2.7) у трех сортов. Например, рост средних значений V в онтогенезе лучше соответствует форме реакции по росту этого признака для сорта № 22, чем для сортов № 13 и 5. Поэтому у двух последних сортов F'_p значительно больше, чем у первого (табл. 2.3 а).

Можно подсчитать, что в табл. 2.3 расчетное значение F'_p больше табличного $F_{t,5\%}$ по высоте для 36% сортов, по длине колоса для 59%, по площади флагового листа – также для 59% сортов. Другими словами, для каждого признака среди 22 сортов есть несколько с несходными реакциями в онтогенезе. В результате регрессия (2.7) неадекватна для многих сортов набора – средние V не являются удачным, общим для всех сортов регрессором по форме реакции признака в онтогенезе и годам выращивания.

Попытка построения регрессий (2.7) для сортов каждой группы районирования по отдельности (V - среднее только по сортам, относящимся к одной группе) не достаточно снижает процент неадекватных моделей, особенно для двух последних признаков. Это видно из табл. 2.3 при сравнении F'_p и $F'_{гр}$ рассчитанных для сортов каждой группы по отдельности. Следовательно, реакции сортов в онтогенезе и по годам, особенно для двух последних признаков, в пределах группы не намного более сходны по форме, чем реакции сортов из разных групп районирования.

Такая ситуация с данными, полученными в онтогенезе, принципиально отличается от высокого процента адекватных регрессий

на средние для наборов сортов, построенных на основе многолетних данных госсортоиспытания по урожайности озимой пшеницы (Смиряев, Гохман, 1985; Смиряев и др., 1992).

Возможно, причина в отличии такого анализа конечных величин количественных признаков (по 7 годам) и данных, полученных в онтогенезе. Однако попытка применения регрессий на средние к данным только о конечной высоте растений первых 20 сортов набора, имеющимся за 7 лет, показала неадекватность квадратичной регрессии на средние для 50% сортов. Можно сделать вывод, что причина неадекватности в несходстве реакции некоторых сортов всего набора как в онтогенезе, так и по годам, т.е. на основе только конечных значений высоты. Возможно, есть группы сортов, сходных по форме реакций, но при попытке связать сорта их разных групп в рамках регрессии (2.7) с усредненной их реакцией, отраженной в изменчивости индекса V , сходство практически не проявляется.

Естественным продолжением анализа сходства по 2-3-летним данным в онтогенезе является переход от регрессий на средние к оценкам SS отклонений от регрессий (2.5) для каждой пары сортов (всего 231 пара для 22 сортов). После этого можно провести кластеризацию – объединение сортов в группы с минимальными отклонениями от этих парных регрессий. Дендрограмма – результат кластеризации по минимуму SS отклонений (точнее по минимуму F'_p регрессий (2.5)) для высоты растений представлена на рис. 2.4. Уровень сходства кластеров $K_{ij} = 1/F'_p$ отложен на оси ординат.

Далее оценивали адекватность регрессий (2.7) для группы сортов, объединенных в один кластер на очередном этапе кластеризации. Если останавливать процесс кластеризации до момента появления первой неадекватной регрессии (2.7) хотя бы в одном кластере, то по каждому из трех ростовых признаков можно выделить по три кластера, охватывающих 82% сортов для высоты растений, 64% сортов для длины колоса и 59% сортов для площади флагового листа. Соответствующие $F'_{кл}$ – расчетные значения критерия Фишера для регрессий (2.7) в пределах каждого кластера с указанием в скобках его номера для сорта приведены в табл. 2.3 а и б.

Для каждого из трех ростовых признаков сформировались два небольших кластера по 2-3 сорта и один большой – по 6-12 сортов. Отметим, что состав кластеров мало соответствует группам по срокам районирования. Сходство по форме реакции для высоты растений в онтогенезе проявляется у некоторых сортов из первой и третьей групп районирования, по остальным двум признакам - из всех трех групп. Такие сорта обычно попадают в большие кластеры. При этом $F'_{кл}$ для некоторых таких сортов даже немного больше $F'_{гр}$, оцененных для них

же в рамках регрессии на средние одной группы районирования (табл. 2.3).

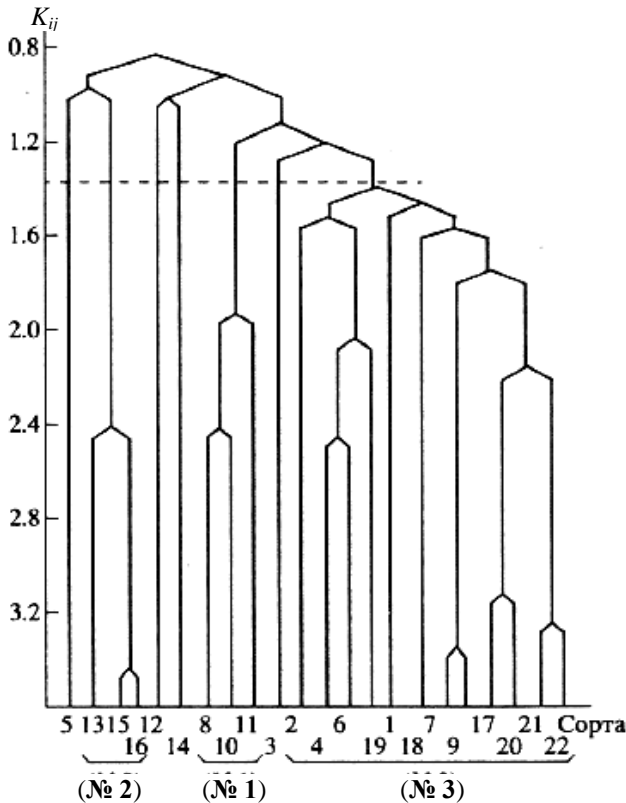


Рисунок 2.4. Дендрограмма – результат кластеризации 22 сортов на основе сходства их формы реакции в онтогенезе по высоте растений в ответ на изменения условий выращивания за три года. Пунктир – момент остановки процедуры кластеризации. В скобках – номера кластеров.

Следует подчеркнуть отсутствие достоверной связи между параметром SS_{ij} (парное несходство формы реакции сортов на изменения общих условий выращивания) и парным их несходством по степени (норме) реакции в онтогенезе. Последнее для любой пары сортов можно оценить модулем разности их коэффициентов b – оценок параметров ω в линейной варианте модели (2.7), т.е. при $\varepsilon=0$. Коэффициент корреляции между SS_{ij} и $|b_i - b_j|$ по данным обо всех парах

сортов, полученным в онтогенезе за три года для высоты растений, составил всего -0,16, за два года для длины колоса -0,10, за два года для площади флагового листа -0,11.

Сходство формы реакции в онтогенезе и по конечным значениям высоты растений

Доступные семилетние экспериментальные данные о высоте растений первых 20 сортов из табл. 2.2 на моменты уборки (в экологических градиентах) и тех же сортов в онтогенезе за 3 года позволяют проверить гипотезу о сходстве относительной изменчивости признака в экологических градиентах и онтогенезе. Сопоставляются ряды показателей парного сродства сортов $SS_{нг}$, SS_p , и SS_s . Всего для 20 сортов таких пар 190. Для проверки гипотезы следует оценить $r(нг, э)$ – коэффициент генотипической корреляции между 190 числами $SS_{нг}$ и 190 числами SS_s , $r(p, э)$ – между SS_p и SS_s , а также дополнительно $r(нг, p)$ – между $SS_{нг}$ и SS_p . Были получены следующие оценки, значимо большие нуля при $P < 0,01$: $r(нг, э) = 0,27 \pm 0,10^{**}$, $r(p, э) = 0,62 \pm 0,13^{**}$ и $r(нг, p) = 0,48 \pm 0,08^{**}$, т.е. $r(нг, p) \approx [r(нг, э) + r(p, э)]/2$.

Следовательно, по высоте растений степень расхождения парных кривых относительной онтогенетической изменчивости за три года довольно тесно связана с несходством формы реакций тех же сортов в экологических градиентах (по данным за 7 лет в рамках моделей (2.6)). Однако степень гладкости парных кривых роста сортов в рамках модели (2.4) за три года (10+9+9 измерений) слабее связана с парным сродством формы реакций за 7 лет. Последнее, возможно, объясняется проявлением эффекта сглаживания кривых квадратичными парными регрессиями (2.4), что отмечалось при обсуждении теории аппроксимации кривых роста. Действительно, пересчет $SS_{нг}$ и SS_p для трехлетних данных в рамках моделей (2.4), (2.5) при $\gamma = 0$, $\gamma_k = 0$ (парные линейные регрессии) дал значительно более высокую оценку $r(нг, э) = 0,82 \pm 0,08^{**}$, но при этом оценка $r(p, э)$ немного снизилась – до $0,57 \pm 0,16^{**}$.

Таким образом, гипотезу о сходстве можно считать подтвержденной по рассматриваемым данным. При этом ожидаемое несходство сортов по форме реакции конечных значений признака, т.е. в экологических градиентах, надежнее прогнозировать на основе отклонений от линейных регрессий, аппроксимирующих парные кривые для этих сортов в онтогенезе, а также на основе расхождений этих кривых, но в рамках квадратичных парных регрессий.

При попытке проверять сходство формы реакций по 7-летним данным о конечной высоте растений и по 1-2-летним данным об их

росте выяснилось, что оценки r неустойчивы. Как уже отмечалось, такое явление можно объяснить тем, что условия выращивания одного или даже двух лет не всегда выявляют различие реакций в онтогенезе по конкретному ростовому признаку. Следует или увеличивать число лет сравнительного испытания сортов, или оценить общее сходство реакций по нескольким ростовым признакам. В последнем случае следует учитывать, что различные ростовые признаки являются проявлением разных сторон единой генетической программы онтогенеза, которая, возможно, в разной степени отличается у сортов.

Сходство аллельного состава и относительной изменчивости в онтогенезе

В качестве показателя общего сходства аллельного состава сортов в работе использованы R_{ij} – парные коэффициенты родства сортов, оцененные по известной методике на основе их родословных (Смиряев и др., 1992). Упрощенно, R_{ij} оценивает долю одинаковых аллелей во всех локусах i -го и j -го сортов, полученных ими от общих предков в процессе селекции. $R_{ij}=0,5$, если сорта i и j имеют одинаковую родословную, $R_{ij}=0$, если в их родословных нет ни одной общей предковой формы. Значения R_{ij} были рассчитаны для 105 пар 15 (из 22) сортов, помеченных звездочкой в табл. 2.2. Только для них имелись полные сведения о родословных.

Использован также показатель Нея I_{ij} (Nei, 1972). В данном случае он характеризовал частное парное сходство тех же 15 сортов по аллельному составу пяти локусов, кодирующих запасные белки зерна (Пыльнев, 1994). I_{ij} может принимать значения от 1, когда сорта i и j имеют одни и те же аллели по каждому из пяти локусов, до нуля, когда все аллели различны.

Ряды R и I для 105 пар 15 сортов сопоставляются с рядами из 105 значений $SS_{нг}$ и SS_p по онтогенетической изменчивости высоты растений, длины колоса и площади флагового листа. Результаты оценки коэффициентов корреляций этих рядов представлены в табл. 2.4.

Как и следовало ожидать, все коэффициенты корреляции отрицательны: чем более сходны сорта по общему (R) и частному (I) аллельному составу, тем меньше несходство их формы реакции в онтогенезе, т.е. более гладки онтогенетические парные кривые ($SS_{нг}$ меньше), а также меньше SS_p , характеризующие их расхождение за годы испытания. При этом линейные регрессии (2.4) для всех случаев, кроме одного, дают более точные сведения о гладкости парных траекторий.

Хотя коэффициенты корреляции стабильно отрицательны, но все они невелики по абсолютной величине, что можно объяснить двумя причинами.

Таблица 2.4. Коэффициенты корреляции, отражающие связи несходства родословных и генов запасных белков с несходством онтогенетической изменчивости количественных признаков 15 сортов за 3 года

	Высота растений		Длина колоса		Площадь флагового листа	
	$SS_{нг}$	SS_p	$SS_{нг}$	SS_p	$SS_{нг}$	SS_p
<i>R</i>	-0,18 (-0,31**)	-0,17	-0,11 (-0,23**)	-0,46**	(-0,32**)	(-0,19)
<i>I</i>	-0,37** (-0,47**)	-0,32**	-0,35** (-0,26**)	-0,21	(-0,29**)	(-0,47**)

Примечание. Без скобок – коэффициенты, рассчитанные на основе квадратичных регрессий (2.4), (2.5); в скобках – на основе линейных, т. е. при $\gamma = \gamma_k = 0$. ** – значимо меньше нуля при $P < 0,01$.

Во-первых, показатель *R*, учитывая сходство родословных, опирается на ряд упрощенных предположений. Например, что после скрещивания двух родительских форм на очередном этапе выведения сорта и после расщеплений в потомстве такого скрещивания формы, отобранные из полученной расщепляющейся популяции для дальнейших скрещиваний, несут по 50% аллелей родительских форм (Смиряев и др., 1992). То есть, несмотря на жесткий отбор в популяции, отобранная ее часть всегда несет равные доли аллелей родительских форм.

Во-вторых, возможно, что онтогенетическая изменчивость количественного признака, например высоты, в наборе сортов определяется лишь частью гетерогенных локусов. Парное несходство аллелей в этой части локусов не обязательно соответствует общему несходству аллельного состава, оцениваемому по *R*, и частному несходству генов запасных белков, отраженному в показателе *I*.

Коэффициенты корреляции указывают на определенное гомологичное сходство изменчивости формы реакций по каждому из трех признаков у пар сортов с большими значениями *R* и *I* (табл. 2.4). В частности, корреляционные связи $SS_{нг}$, SS_p с *I*, возможно, вызваны плейотропными влияниями генов запасных белков на онтогенетическую изменчивость ростовых признаков и(или) сцепление с генами, определяющими эту изменчивость.

Выводы о причинах низких абсолютных значений коэффициентов корреляции табл. 2.4. подтверждаются также и тем, что между R и I коэффициент корреляции тоже слабые ($0,3 \pm 0,09^{**}$). Возможно также, что три рассмотренных ростовых признака сходны по набору полиморфных локусов и(или) блоков, определяющих их изменчивость в онтогенезе. Это также могло снизить корреляции между параметрами сходства формы их реакций и показателями R , I , но повысить – между $SS_{нг}$ и SS_p . Такое предположение соответствует в среднем более высоким, чем в табл. 2.4, оценкам коэффициентов генотипических корреляций между $SS_{нг}$ и SS_p для трех признаков по 105 парам 15 сортов (табл. 2.5). Все эти коэффициенты положительны, некоторые указывают на достоверную статистическую связь между однотипными параметрами несходства (негладкостью или расхождением парных кривых роста) для трех количественных признаков в онтогенезе.

Таблица 2.5. Матрица коэффициентов генотипических корреляций r между $SS_{нг}$ трех признаков (над диагональю), между SS_p (под диагональю), а также между $SS_{нг}$ и SS_p для каждого из трех признаков (на диагонали – подчеркнуто)

	1. Высота растений	2. Длина колоса	3. Площадь флагового листа
1	<u>0,56**</u>	0,42**	(0,24*)
2	0,26	<u>0,73**</u>	(0,34*)
3	(0,85**)	(0,48*)	<u>(0,20)</u>

Примечание. Без скобок – оценки, полученные на основе квадратичных регрессий (2.4), (2.5), в скобках – линейных, т.е. при $\gamma = \gamma_k = 0$.

* – значимо больше нуля при $P < 0,05$, ** – при $P < 0,01$.

По-видимому, косвенным проявлением генетического несходства родительских форм для озимой пшеницы является характер разброса гетерозисных эффектов их гибридов F_1 (рис. 1.6). Чем сильнее различаются родители пары (увеличивается \bar{E} – их евклидово расстояние), тем больше в онтогенезе вероятность несходства кривой роста признака у гибрида F_1 и средней кривой его роста для двух родителей. Этот эффект и проявляется на моменты уборки в наблюдаемой тенденции – увеличении $|d\bar{H}|$ т.е. «разброса» показателя гипотетического гетерозиса \bar{H} для генетически несходных родителей (рис. 1.6 и высоко значимые коэффициенты корреляции для трех признаков во 2-й строке табл. 1.8).

Синхронность прохождения фенофаз и сходство относительной изменчивости

Различие сроков наступления одинаковых фаз развития у двух генотипов естественно рассматривать как проявление несходства их аллельного состава. Последнее, как показано выше, вызывает определенное несходство по форме реакции ростовых признаков в онтогенезе. С другой стороны, в условиях полевого опыта само генетически обусловленное различие наступления фенофаз может усилить несходство реакций. Действительно, при переходе к новой фазе развития происходит существенная перестройка генетических программ онтогенеза, меняется экспрессия, вплоть до изменения спектра активных генов (Батыгин, 1986). В период, когда два генотипа проходят разные фазы развития, общие колебания условий полевого опыта накладываются на разные генетические программы, что вызывает дополнительное несходство реакций, в том числе по форме кривых роста количественных признаков.

Анализ 13 дат наступления фенофаз для каждого из 22 сортов показывает, что их различия в один год для некоторых пар (i, j) сортов достигает 5-7 дней. Несходство формы реакции по названным причинам, связанным с различием этих дат, может вызвать корреляционные связи между FS_{ij} – суммами квадратов этих расхождений в днях по 13 датам и SS_{ij} – суммами квадратов отклонений от регрессионных моделей (2.4), (2.5) для тех же пар по 2-3-летним данным о росте трех признаков. Значения таких корреляций для 2-х составляющих параметра SS_{ij} , оцененных по данным о 231 пар 22 сортов, представлены в табл. 2.6.

Таблица 2.6. Коэффициенты корреляции: $r_{нг} - FS_{ij}$ с $SS_{нг}$ (показателями негладкости); $r_p - FS_{ij}$ с SS_p (расхождения)

Высота растений		Длина колоса		Площадь флагового листа	
$r_{нг}$	r_p	$r_{нг}$	r_p	$r_{нг}$	r_p
0,19*	0,18*	0,60**	0,87**	(0,03)	(0,08)
(0,46**)		(0,77**)			

Примечание. Без скобок – оценки, полученные на основе квадратичных регрессий (2.4), (2.5), в скобках – линейных, т.е. при $\gamma = \gamma_k = 0$.

* – значимо больше нуля при $P < 0,05$; ** при $P < 0,01$.

Степени скоррелированности несходства дат фенофаз и несходства формы реакции в онтогенезе существенно отличаются по трем признакам. Если по высоте растений корреляция слабая, по

площади флагового листа отсутствует, то по длине колоса она тесная и высоко достоверная. Возможно, это объясняется тем, что данные по фенофазам были неполны. Из 13 дат, отмеченных за 6 лет для каждого сорта, 10 зафиксированы до окончания роста длины колоса (обычно до конца июня), причем 6 из них относятся именно к наступлению колошения (конец мая – начало июня в зависимости от условий года). Это период наиболее интенсивного нарастания длины колоса, тогда как рост высоты растений и площади флагового листа в основном заканчивается раньше. Лишь 15-20% прироста высоты растения за 3 года отмечалось во второй половине июня, для флагового листа рост в основном оканчивается еще раньше. До конца мая – начало июня, когда заканчивается основной период нарастания двух последних признаков, за 6 лет было зафиксировано лишь 3 ряда дат наступления фенофаз. Следовательно, их различия по сортам слабее повлияли на величины FS_{ij} . Отметим, что расхождения дат по одной фенофазе для пар сортов слабо или нестабильно коррелируют с расхождениями по другим фенофазам: коэффициенты корреляции таких расхождений меняются по годам и парам фенофаз от 0,93 до -0,30.

Заметно также, что $SS_{\text{нр}}$ в рамках линейных регрессий сильнее, чем в рамках квадратичных, коррелирует со сдвигом фенофаз (FS) у пар сортов, т.е. наблюдалась ситуация, аналогичная представленной выше в табл. 2.4.

Следует также отметить, что стандартная методика определения наступления каждой фенофазы, основанная на глазомерной полевой оценке, не очень точна, вносит существенные случайные ошибки, снижающие оценки корреляций. При повышении точности оценки, например, при учете дат наступления фенофаз по каждому повторению опыта и полной фиксации всех фенофаз за 6 лет корреляции в табл. 2.6 были бы, по-видимому, несколько выше.

Несмотря на неполноту и недостаточную точность оценки дат фенофаз сортов, отмечены значимо меньшие нуля коэффициенты корреляции FS_{ij} и R_{ij} – показателями родства по 105 парам 15 сортов ($r=-0,36\pm 0,09^{**}$) и с I_{ij} – показателями Нея ($r=-0,24\pm 0,09^{**}$). Последнее вызвано плейотропными эффектами и(или) сцеплением генов, определяющими расхождение фенофаз, с генами, определяющими состав запасных белков.

Заключение

Характеристики отклонений от регрессионных моделей, описывающих относительную изменчивость ростового признака у пар генотипов в серии совместных полевых испытаний, могут служить

параметрами для количественной оценки сходства и различия формы реакции генотипов на изменение условий среды в онтогенезе и в экологических градиентах. Степень «негладкости» кривых относительной изменчивости признака у разных пар генотипов в онтогенезе эффективнее сравнивать в рамках линейных регрессий, а степень расхождения этих кривых по условиям совместного выращивания в серии опытов – в рамках квадратичных регрессий.

Гипотеза о сходстве относительной изменчивости признака в экологических градиентах и онтогенезе для серии полевых опытов подтверждена на данных о высоте растений сортов озимой пшеницы. По данным о трех ростовых признаках установлено, что предложенные параметры несходства формы реакции сортов не коррелируют с несходством их нормы реакции.

Между косвенными оценками несходства аллельного состава сортов и параметрами несходства их формы реакции по ростовым признакам в онтогенезе установлена слабая корреляционная связь. Между параметрами несходства формы реакции для разных ростовых признаков связь более тесная. В условиях полевых опытов несходство сортов по форме реакции ростового признака в онтогенезе существенно коррелирует с их несходством по срокам наступления тех фаз, в продолжении которых признак интенсивно нарастает.

Положительную корреляцию несходства формы реакции и аллельного состава генотипов можно объяснить следующим образом. Каждое существенное изменение условий выращивания в полевом эксперименте вызывает отклонения скорости роста многих количественных признаков, причем различия таких реакций двух сравниваемых генотипов в онтогенезе существенно зависят от большого числа полиморфных локусов. Позже происходит частичная компенсация отклонений физиологических процессов, в т.ч., ростовых, но также по-разному в зависимости от аллельного состава генотипов. Включение в метрику несходства формы реакции повторных оценок ростового признака по каждому году совместного испытания сортов отдельно увеличивает информацию о несходстве их аллельного состава по все большему числу локусов. Это связано с частичным переопределением генетической формулы количественного признака. Поэтому предложенные параметры несходства относительной изменчивости ростовых признаков могут служить косвенной оценкой несходства аллельного состава совместно испытываемых генотипов.

Глава 3. Наследование несходства в диаллельных скрещиваниях

Разработана аддитивно-доминантная модель и параметры для биометрико-генетического анализа парного несходства генотипов. Предполагается, что анализируемые величины метрики несходства определяются по количественным признакам чистых линий и гибридов F₁ диаллельной схемы скрещивания. Установлены связи параметров несходства, предложены алгоритмы их оценок и использования, в частности, для выявления максимально гетерозиготных гибридов. Приведены примеры разработанного метода анализа парного несходства для сортов и гибридов F₁ диаллельной схемы скрещивания пшеницы. В качестве метрик использовано несходство формы реакции генотипов по ростовому признаку – высоте растений, а также модуль разности одного количественного признака генотипов. Установлен характер наследования несходства, выявлены сорта, наиболее несходные по аллельному составу.

В последние десятилетия в селекционно-генетических исследованиях все шире используются генетические, биохимические и молекулярные маркеры (Глазко, Глазко, 2006; Lee, 1995; Reif et al., 2005; Couviour et al., 2011). Однако на практике далеко не всегда достаточно сведений о маркерах для получения, например, надежных оценок генетического расстояния – общего различия аллельного состава сравниваемых сортов. Поэтому в генетике количественных признаков и селекции растений по-прежнему применяют различные метрики парного несходства генотипов, косвенно оценивающие их генетическое расстояние. Среди них метрики, основанные на измерениях количественных признаков двух сравниваемых генотипов:

- простейшая метрика несходства – модуль разности величины одного признака 2-х генотипов, коэффициент корреляции средовой изменчивости их признака, а также показатели несходства формы реакции генотипов с учетом средовой и онтогенетической изменчивости признака (глава 2);

- на основе комплекса признаков: евклидово расстояние между генотипами, более общая метрика расстояния Махаланобиса и др. (Животовский, 1991).

Подобные метрики применяют, в частности, для оценки генетической дивергенции при подборе пар сортов для скрещивания (Смиряев и др., 1992). Косвенно оценивая аллельное несходство локусов количественных признаков, метрика имеет количественное выражение, зависящее от вкладов этих локусов и от влияния

модификационной изменчивости признаков.

Возникает вопрос – как наследуется это несходство в потомстве после скрещивания? Ответ позволит выявлять закономерности наследования несходства родительских генотипов, точнее прогнозировать и оптимизировать общее генетическое разнообразие потомства.

Разработан биометрико-генетический анализ наследования величин метрик несходства (Смиряев, 2008). Для этого использованы методы аналитического и имитационного моделирования.

3.1. Модель и параметры наследования парного несходства

Итак, генетическое расстояние, т.е. несходство аллельного состава двух генотипов, косвенно оценивается значением какой-либо метрики их парного несходства. Значение метрики несходства определяется вкладами некоторого количества полиморфных генов или блоков. Это те же гены, которые определяют различия величины признаков, используемых для построения метрики несходства генотипов, но вклады этих генов в значение метрики, естественно, другие. Далее предполагается, что значения метрики получены при обобщении измерений количественных признаков в диаллельном скрещивании, т.е. у p чистых линий и $p(p-1)/2$ гибридов F_1 .

В отличие от величины признака каждого генотипа значение метрики относится к их паре, что качественно усложняет анализ наследования несходства. Например, для пары родительских линий и их гибрида F_1 следует учитывать три значения метрики: несходство двух родителей, гибрида и первого родителя, гибрида и второго родителя (рис.3.1).

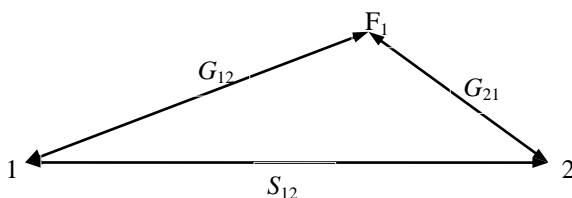


Рисунок 3.1. Схема генетических расстояний трех генотипов, оцениваемых метрикой парного несходства: S_{12} – расстояние между родителями 1 и 2, G_{12} – между гибридом F_1 и 1-м его родителем, G_{21} – между гибридом и 2-м его родителем.

Предполагается, что для метрики парного несходства выполняется правило треугольника метрического пространства, т.е.

$S_{12} \leq G_{12} + G_{21}$. Кроме того должны выполняться следующие условия, аналогичные принятым в обычном диаллельном анализе по Хейману для количественного признака (Науман, 1958):

1. Гомозиготность родительских форм;
2. Отсутствие множественного аллелизма (только два аллеля в каждом локусе, влияющем на изменчивость значений метрики);
3. Отсутствие неаллельных т.е. эпистатических взаимодействий: величина метрики несходства двух генотипов это просто сумма вкладов их полиморфных локусов;
4. Гены распределены независимо у исходных родительских форм;
5. Диплоидное расщепление (нормальное прохождение мейоза);
6. Отсутствие существенных различий по величине используемой метрики у реципрокных гибридов.

При моделировании для простоты предполагалось отсутствие случайных ошибок исходных данных. В рамках всех этих предположений в табл. 3.1 приведен модельный пример для четырех чистых линий при двух полиморфных локусах. Объяснение параметров σ_S^2 , V_i , W_i , W'_i , V'_i – в следующем разделе.

В скобках указаны значения метрики парного несходства генотипов при количественных вкладах локусов, выбранных в модельном примере. По локусу B : $A_1=2$; $\overline{A_1}=3$; $\underline{A_1}=1$, по локусу C : $A_2=3$; $\overline{A_2}=3$; $\underline{A_2}=0,5$. В частности, $S_{23}=A_1+A_2=5$, где A_1 , A_2 - количественные вклады локусов B , C в несходство 2-й и 3-й родительских линий, которые различаются по аллелям обоих локусов.

Например, несходство линии 2 ($BBcc$) и гибрида F_1 ($BbCc$), полученного от скрещивания родительских линий 2×3 : $G_{23} = \underline{A_1} + \overline{A_2} = 4$, а несходство этого гибрида и линии 3 ($bbCC$): $G_{32} = \overline{A_1} + \underline{A_2} = 3,5$. Здесь $\underline{A_1}$ - количественный вклад первого локуса, вызванный несходством Bb гибрида с BB родительской формы, в значение G_{23} , $\overline{A_2}$ - вклад второго локуса (несходство Cc и cc). Аналогично для G_{32} . Матрица табл. 3.1а симметрична относительно нулевой диагонали ($S_{ij} = S_{ji}$). Матрица табл. 3.1б в общем случае несимметрична ($G_{ij} \neq G_{ji}$).

Общее число значений метрики, используемых в анализе, из двух таблиц типа табл. 3.1 равно $3p(p-1)/2$, где p - количество родительских форм. При необходимости в анализе несходства генотипов диаллельной схемы можно вовлечь значения парного несходства между гибридами F_1 . Например, оценка генетического расстояния между гибридами, полученными от скрещивания 1×3 ($bbCc$) и 2×4 ($BbCc$) равна A_1 , между 1×3 и 3×4 ($BbCC$): $\overline{A_1} + \underline{A_2}$, между

1×4 (*BbCc*) и 2×3 (*BbCc*) расстояние равно 0 и т.д.

Таблица 3.1. Структура и количественные значения метрики несходства четырех родителей и их гибридов F₁ в модельном примере диаллельной схеме (пояснения в тексте):

a) S_{ij} – структура и значение метрики несходства i -го и j -го родителей ($i, j=1, \dots, 4$).

		<i>bbcc</i>	<i>BBcc</i>	<i>bbCC</i>	<i>BBCC</i>
		1	2	3	4
$i \setminus j$					
<i>bbcc</i>	1	(0)	A_1 (2)	A_2 (3)	A_1+A_2 (5)
<i>BBcc</i>	2	A_1 (2)	(0)	A_1+A_2 (5)	A_2 (3)
<i>bbCC</i>	3	A_2 (3)	A_1+A_2 (5)	(0)	A_1 (2)
<i>BBCC</i>	4	A_1+A_2 (5)	A_2 (3)	A_1 (2)	(0)
σ_S^2		4,33	4,33	4,33	4,33

б) G_{ij} – несходство гибрида F₁ с i -м родителем, G_{ji} – гибрида с j -м родителем ($i \leq j$).

		<i>bbcc</i>	<i>BBcc</i>	<i>bbCC</i>	<i>BBCC</i>	V_i	W_i
		1	2	3	4		
$i \setminus j$							
<i>bbcc</i>	1	(0)	$\overline{A_1}$ (3)	$\overline{A_2}$ (3)	$\overline{A_1+A_2}$ (6)	6,00	5,00
<i>BBcc</i>	2	$\overline{A_1}$ (1)	(0)	$\overline{A_1+A_2}$ (4)	$\overline{A_2}$ (3)	3,33	3,67
<i>bbCC</i>	3	$\overline{A_2}$ (0,5)	$\overline{A_1+A_2}$ (3,5)	(0)	$\overline{A_1}$ (3)	3,08	2,50
<i>BBCC</i>	4	$\overline{A_1+A_2}$ (1,5)	$\overline{A_2}$ (0,5)	$\overline{A_1}$ (1)	(0)	0,42	1,17
V'_i		0,42	3,08	3,33	6,00		
W'_i		1,17	2,50	3,67	5,00		

Общее число расстояний для разных пар гибридов F_1 в диаллельной схеме при p родителях равно $M(M-1)/2$, где $M=p(p-1)/2$. Аналогично можно прогнозировать несходство между генотипами популяции, ожидающейся в поздних поколениях – в потомстве любого гибрида F_1 .

Таким образом, определена биометрико-генетическая структура метрики парного несходства между различными генотипами – как гомозиготными, так и гетерозиготными. Модель может быть распространена на любое число полиморфных локусов.

Конкретные соотношения вкладов A_k , $\overline{A_k}$, $\underline{A_k}$ определяют характер наследования несходства по отдельным локусам и парам генотипов. В частности:

- полное доминирование одного из двух аллелей по k -му локусу (например, B над b): $\underline{A_k}=0$, $\overline{A_k}=A_k$, (рис. 3.2а). При полном доминировании родителя i над j по всем их полиморфным локусам: $G_{ij}=0$, $G_{ji}=S_{ij}$. То есть родитель i содержит все чисто доминантные аллели, и значение метрики его несходства с гибридом F_1 равно 0. Тогда j -й родитель рецессивен по отношению к i -му, т.к. содержит все чисто рецессивные аллели ($\overline{A_k} = A_k$ для любого k);

- неполное доминирование (например, B над b) по k -му локусу: $\overline{A_k} > \underline{A_k} > 0$ (рис. 3.2б);

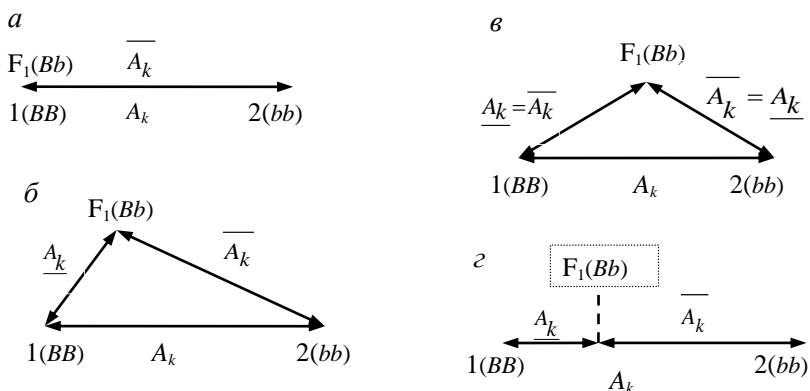


Рисунок 3.2. Различные схемы наследования несходства для одного (k -го) полиморфного локуса с двумя аллелями: BB – у 1-й гомозиготной родительской линии, bb – у 2-й

- аддитивная схема наследования парного несходства проявляется при $\overline{A_k} = \underline{A_k} \geq 0,5A_k$ (рис. 3.2в). Если по всем полиморфным локусам наблюдается такая схема, то $G_{ij}=G_{ji}$ и матрица типа табл. 3.1б - симметрична.

Графический анализ парного несходства

Кроме шести предположений метода Хеймана использованы некоторые параметры и подходы, аналогичные графическому анализу признака по Хейману (Науман, 1958), но отнесенные к метрике парного несходства.

Вводятся, в частности, параметры V_i и V'_i – дисперсии i -х строк и столбцов таблицы несходства гибридов F_1 с родительскими линиями (при четырех линиях это строки и столбцы табл. 3.1б), а также W_i и W'_i – ковариации i -х строк и столбцов таблицы несходства гибридов с i -ми строками табл. 3.1а несходства линий. Можно показать, что при выполнении шести предположений метода Хеймана между этими параметрами существуют две тесные регрессионные зависимости с одинаковыми коэффициентами:

$$W = a + b \times V \quad (3.1),$$

$$W' = a + b \times V' \quad (3.2).$$

Здесь a, b – МНК-оценки регрессий, полученные по значениям W_i, V_i и W'_i, V'_i ; $i=1, \dots, p$, p - число родительских линий.

Математический анализ показал, что все p точек с координатами $(W_i; V_i)$, как и в графическом методе Хеймана, лежат ниже или на самой ограничивающей параболе

$$W = \sqrt{\sigma_S^2 \times V} \quad (3.3),$$

где σ_S^2 – дисперсия любой строки или столбца таблицы несходства линий: при выполнении предположений метода дисперсии σ_S^2 одинаковы (последняя строка табл. 3.1а). Для точек $(W'_i; V'_i)$ ограничение (3.3) также справедливо.

Равенство σ_S^2 по i может служить тестом выполнения предположений метода, но только для значений S_{ij} , а равенство $\text{cov}(G_{ij}; G_{ji})$ по i – тестом для G_{ij} .

Пример для значений $W_i, V_i, W'_i, V'_i, \sigma_S^2$ из табл. 3.1 приведен на рис. 3.3. На графике нанесены как точки с координатами $(W_i; V_i)$, помеченные номерами родителей, так и $(W'_i; V'_i)$ – номерами родителей со штрихом (в скобках).

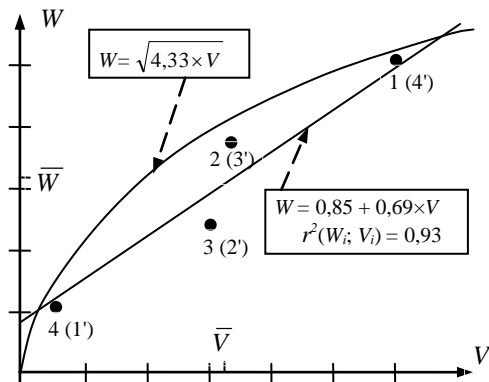


Рисунок 3.3. Модельный пример графического анализа значений метрики несходства для четырех родительских линий (1...4) и их диаллельных гибридов F_1 при двух полиморфных локусах. Вклады $A_1, \bar{A}_1, \underline{A}_1, A_2, \bar{A}_2, \underline{A}_2$ – из табл. 3.1а,б.

Числа W_i+V_i характеризуют доли рецессивных и доминантных генов метрики несходства i -той родительской линии. В частности, при полном доминировании всех аллелей i -ой линии (содержит все доминантные аллели, причем $A_k=0$) точка $(W_i; V_i)$ лежит на нижнем пересечении регрессии (3.1) и параболы (3.3). Точка $(W_j; V_j)$ j -той линии, содержащей все рецессивные аллели метрики несходства, лежит на верхнем их пересечении, а величина W_j+V_j , естественно, имеет максимальное значение среди всех родительских линий.

При неполном доминировании по локусам ($\bar{A}_k > \underline{A}_k > 0$) для родительских линий, содержащих как рецессивные, так и доминантные гены, их точки $(W_i; V_i)$ лежат тем выше вдоль регрессии (3.1), чем сильнее у них суммарный вклад рецессивных генов несходства (\bar{A}_k) по сравнению с доминантными (\underline{A}_k). Следовательно, значения контрастов $|W_i+V_i-W_j-V_j|$ положительно коррелируют с различием долей рецессивных и доминантных генов, определяющих несходство родителей i и j .

Для точек $(W'_i; V'_i)$ тех же родителей ситуация с пересечениями и

расположением их на регрессии (3.2) противоположная, в частности, для полностью рецессивной линии $W_i=V_i=0$.

В ситуации, когда по всем полиморфным локусам $\overline{A_k} = \underline{A_k}$ (аддитивная схема), регрессионные связи (3.1), (3.2) отсутствуют, т.к. все p точек ($W_i; V_i$) сливаются в одну.

Если хотя бы по некоторым локусам $\overline{A_k} \neq \underline{A_k}$, то корреляция между показателями несходства линий (S_{ij}) и контрастами $|W_i+V_i-W_j-V_j|$ позволяет судить о характере статистической связи вкладов A_k и контрастов $(\overline{A_k} - \underline{A_k})$ по локусам. Имитационное моделирование показало, что если эта корреляция в пределах 0,3-0,5, то связь слабая, если ниже 0,3, то она значимо отрицательная, если выше 0,5, то значимо положительная. На значения корреляции влияют также частоты доминантных и рецессивных аллелей по локусам в наборе родительских линий.

Можно показать, что при выполнении шести предположений метода Хеймана справедливы следующие равенства:

$$(V_i+V'_i)/2=\overline{V}=\overline{V'} \quad (3.4), \quad (W_i+W'_i)/2=\overline{W}=\overline{W'} \quad (3.5),$$

где \overline{V} , \overline{W} , $\overline{V'}$, $\overline{W'}$ - средние по всему набору родительских линий значения V_i, W_i, V'_i, W'_i . Именно по этой причине каждая i -я пара точек ($W_i; V_i$) и ($W'_i; V'_i$) на рис. 3.3 расположена симметрично относительно единой точки с координатами $(\overline{W}; \overline{V})$.

Совпадение регрессий (3.1) и (3.2), а также равенство по i значений $V_i+V'_i+W_i+W'_i$ (следствие равенств (3.4) и (3.5)) можно использовать в качестве тестов выполнения шести предположений, в частности, аддитивно-доминантного характера наследственной изменчивости метрики парного несходства. При анализе реальных опытных данных, естественно, необходимо учитывать влияние случайных средовых ошибок на оценки всех параметров.

Из равенств (3.4), (3.5) следует также, что, имея данные реального эксперимента и проверив статистическую справедливость предположений метода, далее для большей точности всего анализа можно пользоваться вместо двух пар параметров V_i, W_i и V'_i, W'_i , одной парой усредненных оценок:

$$\hat{W}_i=(\overline{W} + \overline{W'})/2+(W_i - W'_i)/2; \quad \hat{V}_i=(\overline{V} + \overline{V'})/2+(V_i - V'_i)/2.$$

При отсутствии случайных экспериментальных ошибок значения

\hat{W}_i и \hat{V}_i совпадают с V_i, W_i . Поэтому свойства пар оценок \hat{W}_i и \hat{V}_i те же, что у пар параметров W_i, V_i . В частности, между значениями \hat{W}_i и \hat{V}_i существует регрессионная связь (3.1).

Дополнительные генетико-статистические параметры и связи

Определенные выводы о наследовании несходства можно сделать, вводя специальные параметры.

Параметр $He_k = (\bar{A}_k + \underline{A}_k) / A_k$, где k - номер полиморфного локуса метрики несходства в наборе родительских линий, позволяет сопоставить вклады гомо- и гетерозиготных локусов в общее несходство. Этот параметр имеет ограничение $He_k \geq 1$. При $He_k > 1$ можно говорить о гетерозисе вклада по локусу k . Следует отметить, что в таких терминах гетерозис возможен и без доминирования, когда $\bar{A}_k = \underline{A}_k > 0,5A_k$ (рис. 3.2в). Если сумма вкладов $\bar{A}_k + \underline{A}_k$ минимальна (равна A_k - рис. 3.2з), то $He_k = 1$.

Параметр $Dm_k = (\bar{A}_k - \underline{A}_k)^2 / (\bar{A}_k^2 + \underline{A}_k^2)$ характеризует степень доминирования по метрике несходства в локусе (k). Этот параметр может меняться от 1 ($\underline{A}_k = 0$ - полное доминирование одного из аллелей), до 0 ($\bar{A}_k = \underline{A}_k$ - отсутствие доминирования, т.е. аддитивная схема наследования по локусу k).

Установлены связи между He_k, Dm_k и параметрами из предыдущего раздела. Если He_k равны по всем полиморфным локусам ($He_k = He$), то регрессионный коэффициент $b = 1/He$, а точки ($\hat{W}_i; \hat{V}_i$) лежат на прямой $\hat{W} = a + b \times \hat{V}$, при этом $\text{cor}(\hat{W}_i; \hat{V}_i) = 1$. В частности, если по всем локусам $Dm_k = 1$ ($He_k = He = 1$), то $b = 1$. Тогда для родительской линии i , содержащей все доминантной линии, выполнено $\hat{W}_i = \hat{V}_i = 0$.

Если He_k не равны по k , то b несколько отличается от $1/\overline{He}$, где \overline{He} - усредненные по локусам (k) значения He_k . Причина - возможная корреляционная связь по k параметра He_k с $(\bar{A}_k + \underline{A}_k)$.

При отсутствии корреляционных связей по k параметра He_k с $(\bar{A}_k + \underline{A}_k)$ и с $(\bar{A}_k - \underline{A}_k)$ среднее значение параметра доминирования Dm_k по локусам (k) оценивается отношением $1 - a / (b \hat{V})$. Здесь \hat{V} - средний по

линиям (i) параметр \hat{V}_i . В частности, при полном доминировании по всем полиморфным локусам $1-a/(b\hat{V})=1$, следовательно, коэффициент регрессий $a=0$.

Как отмечалось выше, при аддитивной схеме наследования метрики несходства ($\overline{A}_k = \underline{A}_k$ по всем полиморфным локусам) p точек $(\hat{W}_i; \hat{V}_i)$ сливаются в одну. Если, кроме того, значения He_k равны по локусам, то эта точка лежит на ограничивающей параболе (3.3), если не равны, то – ниже ее.

Оценка гетерозиготности гибридов в диаллельном скрещивании

Для решения различных теоретических и прикладных задач желательно сравнить степень гетерозиготности разных гибридов F_1 диаллельной схемы, т.е. доли их гетерозиготных локусов среди всех полиморфных локусов набора родительских линий. В частности, выбрать пару родителей, дающих максимально гетерозиготный гибрид F_1 , что обеспечит в популяции потомства максимальное генетическое разнообразие.

Для этих целей пригодны сами значения метрик парных несходств родительских линий. Однако ранги значений S_{ij} не обязательно совпадают с рангами гетерозиготности гибридов (Смиряев и др., 1999). В рамках представленной модели на значение метрики несходства родителей существенно влияют, например, неравные вклады A_k разных локусов. Непосредственно по метрике невозможно также установить есть ли среди гибридов F_1 полностью гетерозиготный. Показатели несходства гибридов с родителями (G_{ij}, G_{ji}) еще менее подходят для оценки гетерозиготности, т.к. их значения зависят от уровня доминирования и гетерозиса по метрике несходства. Единственным исключением является скрещивание полностью доминантного (по всем полиморфным локусам) с полностью рецессивным родителем

Более универсальным и точным параметром в подобных ситуациях оказывается корреляция значений метрики несходства для пары родительских линий (i, j) т.е. корреляция по родителям (m) двух рядов: S_{im} и S_{mj} , где $m=1, \dots, p$. Обозначим эту корреляцию через rS_{ij} . В табл. 3.2 приведен пример значений rS_{ij} – корреляции величин метрики из i -х строк и j -х столбцов табл. 3.1а.

Можно доказать математически, что в рамках предложенной модели несходства генотипов параметр rS_{ij} обладает рядом полезных

свойств.

1) Если для двух пар линий (i, j) и (l, m) выполнено равенство $rS_{ij}=rS_{lm}>-1$, то часть аллелей у них совпадает (остальные альтернативны). Обратное утверждение также справедливо: например, пары 1, 2 ($bbcc, BBcc$) и 3, 4 ($bbCC, BBCC$) имеют часть общие аллели bb и BB (CC и cc альтернативны). Тогда в предположениях табл. 3.1а: $rS_{12}=rS_{34}=0,38$. Следовательно, в результате двух скрещиваний – 1×2 и 3×4 получим гибриды F_1 с одинаковой, но не 100%-ной гетерозиготностью.

2) Если для двух линии (i, j) корреляции значений метрики несходства каждой из них с любой (l -ой) линией равны по модулю, но противоположны по знаку ($rS_{il}=-rS_{jl}$), то линии (i, j) имеют альтернативные аллели по всем полиморфным локусам набора родительских линий. Обратное также справедливо. Пример в табл. 3.2 – корреляции альтернативных по составу линий 2 ($BBcc$) и 3, ($bbCC$) с линией 1 ($bbcc$): $rS_{12}=0,38, rS_{13}=-0,38$.

3) Для любых двух родительских линий (i, j) с альтернативным аллельным составом по всем полиморфным локусам набора справедливо равенство $rS_{ij}=-1$, и наоборот. Пример в табл 3.2: $rS_{14}=-1, rS_{23}=-1$.

Таблица 3.2. Значения параметра rS_{ij} по модельным данным табл. 3.1а.

		<i>bbcc</i>	<i>BBcc</i>	<i>bbCC</i>	<i>BBCC</i>
		1	2	3	4
<i>i</i> \ <i>j</i>					
<i>bbcc</i>	1	1	0,38	-0,38	-1
<i>BBcc</i>	2	0,38	1	-1	-0,38
<i>bbCC</i>	3	-0,38	-1	1	0,38
<i>BBCC</i>	4	-1	-0,38	0,38	1

Первое свойство выполняется в точности при соблюдении всех шести предположений метода Хеймана; 2-е и 3-е – даже если не выполняется 4-е предположение о независимом распределении генов у родительских линий диаллельной схемы скрещиваний.

Все свойства rS_{ij} остаются в силе и для параметра rG_{ij} - корреляции по m двух рядов: $G_{im}+G_{mi}, G_{jm}+G_{mj}$, где $m=1, \dots, p$. Таблица этих корреляций, полученная по числовым данным табл. 3.1б, аналогична табл. 3. 2, но всюду вместо числа 0,38 появиться число -0,13,

а вместо $-0,38$ появится $+0,13$. Следовательно, можно использовать параметр rG_{ij} также, как rS_{ij} или пользоваться единым параметром $r_{ij}=(rG_{ij}+rS_{ij})/2$. В особой ситуации, когда He_k равны по всем полиморфным локусам ($He_k=He>1$) значения rG_{ij} и rS_{ij} совпадают.

Итак, если среди родительских форм есть пара линий (i, j), для которой $rS_{ij}=-1$ или $rS_{il}=-rS_{jl}$, где l – любая 3-я линия набора, то гибрид $F_1 i \times j$ обеспечит максимальное генетическое разнообразие потомства. Скрещивание такого гибрида с остальными линиями набора родителей или с их гибридами не увеличит генетическое разнообразие популяции потомства. Следует подчеркнуть, что максимальное генетическое разнообразие потомства по генам метрики несходства означает такое же разнообразие по генам всего комплекса количественных признаков, используемых для построения этой метрики.

В экспериментальных исследованиях необходимо учитывать ограничения представленных результатов моделирования. Так, в наборе родителей может не оказаться ни одной пары линий с аллелями, контрастными по всем полиморфным локусам. В таком случае, во-первых, можно использовать тенденцию: чем ближе значение $r_{ij}=(rG_{ij}+rS_{ij})/2$ к -1 , тем более несходны родители i, j . В дополнение к свойству 3 параметров rS_{ij} и rG_{ij} следует проверять их свойство 2.

Во-вторых, если не удастся выявить ни одной пары с $rG_{ij}=-1$, можно попытаться выбрать из всего набора 3-4 родительские линии, которые после скрещивания образуют популяцию, полиморфную по всем генам, определяющим метрику несходства в исходном наборе. Моделирование диаллельных скрещиваний наборов родителей с тремя полиморфными локусами (B, C, E) позволило установить следующее свойство.

4) Если для трех родительских линий сумма парных корреляции r_{ij} близка к -1 , то по каждому локусу, полиморфному в диаллельной схеме, среди трех есть две линии с альтернативными аллелями. Обратное также справедливо.

Например, для трех линий: 1) $BBccee$, 2) $bbCCee$, 3) $BBCCEE$ любая из трех парных корреляций больше -1 . Но для суммы этих трех корреляций выполнено: $r_{12}+r_{13}+r_{23} \approx -1$. То есть для получения популяции, полиморфной по всем трем генам, определяющим несходство родителей исходного набора, достаточно скрестить три названные линии.

Свойство 4, в отличие от свойства 3 справедливо при выполнении всех шести предположений метода. Возможно, что при большем числе полиморфных локусов и родительских линий, чем использованное при моделировании, свойство 4 не выполняется. Сумма их r_{ij} может оказаться меньше -1 .

Дополнительную информацию о характере распределения аллелей у пар максимально несходных родителей можно получить при сопоставлении их параметров rS_{ij} или rG_{ij} с контрастами $|W_i+V_i-W_j-V_j|$ (или, что тоже самое, с $|\hat{W}_i+\hat{V}_i-\hat{W}_j-\hat{V}_j|$). Например, у пар 1×4 и 2×3 (табл. 3.2) $rS_{14}=rG_{23}=-1$, а значения контрастов сильно различаются (судя по рис. 3.3 у первой пары $|W_i+V_i-W_j-V_j|$ значительно выше). Следовательно, доминантные (*B*, *C*) и рецессивные (*b*, *c*) гены несходства у родителей 1 и 4 распределены весьма неравномерно, а у 2 и 3, напротив, представлены почти в равных долях. Этот вывод может быть полезен для прогноза генетического разнообразия в поколениях самоопыления гибридов 1×4 и 2×3, особенно при сильном сцеплении генов, определяющих генетическое несходство.

3.2. Биометрико-генетический анализ несходства генотипов в диаллельных скрещиваниях пшеницы

Далее приведены два примера применения разработанной модели, характеризующей изменчивость значений метрик несходства генотипов, и соответствующего биометрико-генетического метода анализа. В первом примере (с яровой пшеницей) в качестве метрики использовано несходство по форме реакции генотипов в онтогенезе (см. глава 2). Во втором примере (с озимой пшеницей) – модуль разности одного признака растений.

Материалы и параметры: несходство по форме реакции

Значения метрики были получены на основе многократных измерений ростового количественного признака – высоты растений в диаллельном скрещивании сортов яровой пшеницы. Значения этих измерений служили исходными данными для всех вычислений.

Использованы четырехлетние данные – одновременные измерения высоты растений 8 сортов и гомозиготных форм яровой пшеницы различного происхождения из коллекций системы ВИР: 1) к-58152 (Бразилия), 2) Сибирская 65 (Россия), 3) Сибирская 3 (Россия), 4) РГ81220 (Индия), 5) Planet (Германия), 6) St. Mercheisto (Швеция), 7) SV Sonett (Швеция, 8) Opal (Швеция), а также 28 их гибридов F_1 . Опыт проводился на селекционной станции РГАУ-МСХА (Смиряев, Пыльнев, 2008). Каждый гибрид представлен растениями, полученными как от прямого, так и от обратного скрещивания 2 сортов, что позволяет не учитывать реципрокный эффект. Схема опыта в каждый из 4 лет – три рандомизированных блока. Любой родительский сорт и гибрид F_1 в блоке представлен тремя рядками по 12-20 растений. Каждое измерение

в блоке – результат усреднения высоты 12-20 растений среднего рядка. По годам число одновременных измерений высоты каждого сорта и гибрида F_1 в онтогенезе составляло от 5 до 9, общее число – 29 измерений за 4 года.

При реализации предложенного биометрико-генетического метода следует учесть влияние случайных ошибок опытных данных на все генетико-статистические параметры и связи модели (Смиряев, 1997). В частности, оценивали дисперсию средовой ошибки каждого измерения признака и автокорреляции этих ошибок между измерениями одного года (разделы 2.1 и 2.2). Это необходимо для корректного построения квадратичных регрессий для пар генотипов и последующего вычисления метрики их несходства по форме реакции в онтогенезе.

В качестве основной метрики несходства формы реакции использован средний квадрат отклонений от квадратичной регрессии. Таким образом, по всем генотипам анализируется S_{ij} – средний квадрат отклонений 29 экспериментальных измерений высоты от квадратичной регрессионной зависимости i -го и j -го сортов с учетом дисперсий и автокорреляции средовых ошибок измерений. Аналогично G_{ij} – несходство по форме реакции их гибрида F_1 с i -м сортом (G_{ji} – с j -м). В главе 2 подобные параметры несходства генотипов обозначались как ms_0 . По данным диаллельного скрещивания яровой пшеницы в рамках нового биометрико-генетического метода проведен анализ наследования этой метрики парного несходства.

Результаты анализа по несходству формы реакции

Реализованы тесты для проверки выполнения предположений метода (Смиряев, 2008), а именно дисперсионный анализ в повторениях величин двух параметров: σ_S^2 и $(V_i + W_i + V'_i + W'_i)/4$. Для получения данных в трех повторениях за все 4 года использован следующий прием.

Проводится случайный выбор серии из четырех отдельных повторений – измерений высоты растений в онтогенезе за четыре года. Например, за 1-й год взяты данные измерений только во 2-м блоке опыта, за 2-й год только в 3-м, за 3-й только в 1-м, за 4-й – во 2-м. Случайный выбор таких серий повторен трижды так, чтобы номера блоков за любой год в сериях не повторялись. Таким образом, однократно использованы все данные эксперимента, причем формально получена схема с тремя рандомизированными блоками – в каждом блоке 29 измерений высоты. Вычислены значения метрики (S_{ij} , G_{ij} и G_{ji}) в каждом повторении отдельно. На их основе получены параметры σ_S^2

и $(V_i+W_i+V'_i+W'_i)/4$ (табл. 3.3).

Результаты дисперсионного анализа каждого из двух параметров, представленные в табл. 3.4, показывают их значимые различия по сортам.

Таблица 3.3. Значения двух параметров 8 сортов диаллельной схемы в трех рандомизированных блоках (I, II, III) для проверки выполнения предположений метода. Номера сортов по списку из предыдущего раздела

сорт	1	2	3	4	5	6	7	8	
σ_S^2	I	0,409	0,314	0,316	0,360	0,469	0,251	0,280	0,325
	II	0,314	0,412	0,332	0,450	0,370	0,334	0,242	0,272
	III	0,383	0,397	0,357	0,603	0,410	0,373	0,229	0,477
$(V_i+W_i+V'_i+W'_i)/4$	I	0,348	0,312	0,285	0,335	0,342	0,237	0,267	0,269
	II	0,260	0,340	0,284	0,352	0,292	0,253	0,270	0,269
	III	0,352	0,314	0,317	0,441	0,328	0,304	0,223	0,342

Таблица 3.4. Дисперсионный анализ данных табл. 3.3

		<i>ms</i>	<i>df</i>	F_p	$F_{t, 5\%}$
σ_S^2	сорта	$1,29 \times 10^{-2}$	7	3,15	2,76
	повторения	$1,06 \times 10^{-2}$	2	2,60	3,74
	ошибки	$0,41 \times 10^{-2}$	14		
$(V_i+W_i+V'_i+W'_i)/4$	сорта	$4,42 \times 10^{-3}$	7	3,96	2,76
	повторения	$3,07 \times 10^{-3}$	2	2,74	3,74
	ошибки	$1,12 \times 10^{-3}$	14		

Парные сравнения сортов позволили установить, что значимость различий параметров, представленных в табл. 3.4, вызывают 4-й и 7-й сорта. Данные по остальным шести сортам и их гибридам F_1 удовлетворяют требованиям метода, поэтому только они использованы в дальнейшем анализе наследования парного несходства генотипов (табл. 3.5). По данным табл. 3.5 определили параметры σ_S^2 , W_i , V_i , W'_i и V'_i (последняя строка табл. 3.5а, последняя пара столбцов и

предпоследняя пара строк табл. 3.5б)

Таблица 3.5. Значения метрики несходства 6-и сортов и их гибридов F_1 диаллельной схемы. Окончательная нумерация сортов (i, j): 1) к58152, 2) Сибирская 65, 3) Сибирская 3, 4) Planet, st. 5) Mercheisto, 6) Opal.

а) S_{ij} – несходство i -го и j -го сортов ($i, j=1, \dots, 6$).

$i \backslash j$	1	2	3	4	5	6
1	0	1,52	1,29	1,68	1,45	1,53
2	1,52	0	1,10	1,56	1,12	1,07
3	1,29	1,10	0	1,39	0,98	1,00
4	1,68	1,56	1,39	0	0,76	0,98
5	1,45	1,12	0,98	0,76	0	0,63
6	1,53	1,07	1,00	0,98	0,63	0
σ_S^2	0,388	0,318	0,248	0,394	0,244	0,264

б) G_{ij} – несходство гибрида F_1 с i -м сортом, G_{ji} – гибрида с j -м сортом ($i \leq j$). Пояснения остальных параметров в тексте.

$i \backslash j$	1	2	3	4	5	6	V_i	W_i
1	0	1,37	1,17	1,66	1,48	1,25	0,351	0,362
2	0,97	0	0,97	1,13	1,22	1,11	0,204	0,226
3	1,04	1,05	0	1,36	0,95	0,91	0,214	0,227
4	0,94	1,04	0,94	0	0,76	0,91	0,148	0,222
5	0,90	0,88	0,94	0,80	0	0,60	0,128	0,162
6	1,05	0,91	0,60	0,69	0,71	0	0,131	0,173
V'_i	0,164	0,214	0,175	0,341	0,260	0,200		
W'_i	0,241	0,249	0,194	0,349	0,251	0,226		
\hat{V}_i	0,305	0,206	0,230	0,115	0,145	0,176		
\hat{W}_i	0,301	0,229	0,257	0,176	0,196	0,213		

Дополнительно сравнили линейные регрессионные зависимости W_i на V_i и W'_i на V'_i : при выполнении предположений биометрико-генетического метода регрессии не должны значимо различаться. Для проверки, кроме МНК-оценок коэффициентов a , b регрессий и ms_0 – средних квадратов отклонений от них (1 и 2 строки табл. 3.6), оценили ms_0 при использовании в регрессии W_i на V_i коэффициентов a , b

регрессии W'_i на V'_i и наоборот (3 и 4 строки табл. 3.6).

Таблица 3.6. Результаты сравнение регрессий W_i на V_i и W'_i на V'_i .

Вариант вычисления параметров	a	b	ms_0
1 W_i на V_i , оценки МНК	0,07	0,82	$3,9 \cdot 10^{-4}$
2 W'_i на V'_i , оценки МНК	0,09	0,72	$6,5 \cdot 10^{-4}$
3 W_i на V_i , но a, b из варианта 2	0,09	0,72	$4,9 \cdot 10^{-4}$
4 W'_i на V'_i , но a, b из варианта 1	0,07	0,82	$7,2 \cdot 10^{-4}$

Сравнение по критерию Фишера значений ms_0 для вариантов вычислений 3-го с 1-м ($F_p=1,26$) и 4-го со 2-м ($F_p=1,11$) показывает отсутствие значимых различий между регрессиями W_i на V_i и W'_i на V'_i ($F_t, 5\%=6,16$).

Все проведенные тесты позволяют объединить две пары параметров W_i, V_i и W'_i, V'_i в одну по формулам: $\hat{W}_i = (\bar{W} + \bar{W}')/2 + (W_i - W'_i)/2$; $\hat{V}_i = (\bar{V} + \bar{V}')/2 + (V_i - V'_i)/2$. Параметры \hat{V}_i и \hat{W}_i приведены в двух последних строках табл. 3.5б.

Для проведения графического анализа наследования метрики несходства, кроме регрессии значений \hat{W}_i на \hat{V}_i , следует построить ограничивающую параболу $\hat{W} = \sqrt{\sigma_S^2} \times \hat{V}$. Поскольку значения σ_S^2 , хотя и незначимо, но различаются по сортам (последняя строка табл. 3.5а), в качестве единой оценки σ_S^2 использовано их среднее значение, равное 0,309. Регрессия \hat{W}_i на \hat{V}_i и ограничивающая параболу $\hat{W} = \sqrt{0,309} \times \hat{V}$ представлены на рис. 3.4.

Используя разработанный биометрико-генетический метод (раздел 3.1) можно сделать некоторые выводы о характере наследования выбранной метрики парного несходства. Коэффициент регрессии $b=0,68 < 1$, следовательно He – среднее по полиморфным локусам (k) значение параметра He_k больше единицы – в схеме наследования метрики несходства есть гетерозис. Коэффициент детерминации для

регрессии (r^2) – близок к 1, следовательно, значение параметра He_k слабо различается по полиморфным локусам (k): $He_k \approx \overline{He} \approx 1/b = 1,47$. Учитывая структуру показателя $He_k = (\overline{A_k} + \underline{A_k})/A_k$, можно предположить, что значения выражений в числителе и знаменателе тесно коррелируют по локусам (k). Как следствие, должны коррелировать значения метрик S_{ij} и $G_{ij} + G_{ji}$ по всем сочетаниям i, j . Оценка $r(S_{ij}; G_{ij} + G_{ji}) = 0,92^{**}$ подтверждает это предположение.

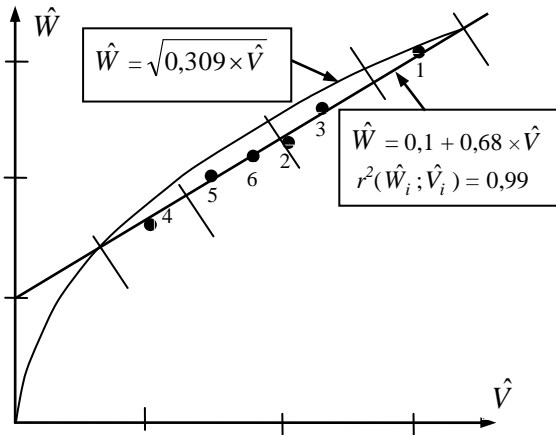


Рисунок 3.4. График для анализа значений метрики несходства родительских сортов и гибридов F_1 яровой пшеницы. Номера сортов (1,...,6) из табл. 3.5.

Точки (\hat{W}_i, \hat{V}_i) «растянуты» вдоль регрессии (рис. 3.4), следовательно, хотя бы по некоторым полиморфным локусам присутствует эффект доминирования $\overline{A_k} \neq \underline{A_k}$. Средняя по локусам степень доминирования $\overline{Dm} = 1 - a/(b\hat{V}) = 0,21$ – доминирование неполное. Учитывая выражение для $Dm_k = (\overline{A_k} - \underline{A_k})^2 / (\overline{A_k}^2 + \underline{A_k}^2)$, несложно оценить среднее по локусам (k) отношение $\underline{A_k} / \overline{A_k} \approx 0,46$.

Коэффициент корреляции между значениями метрики несходства сортов (S_{ij}) и несходством долей их рецессивных генов ($|\hat{W}_i + \hat{V}_i - \hat{W}_j - \hat{V}_j|$) равен $0,81^{**}$. Это свидетельствует о высоко значимой статистической связи вкладов A_k и контрастов $(\overline{A_k} - \underline{A_k})$ по локусам.

Учитывая отмеченную выше тесную связь $(\overline{A_k} + \underline{A_k})$ с A_k , можно предположить наличие тесной статистической связи $(\overline{A_k} + \underline{A_k})$ с $(\overline{A_k} - \underline{A_k})$. Это, в свою очередь, означает, что отношение $\underline{A_k} / \overline{A_k}$ мало меняется по локусам (k).

Сорт 1 несет более 75%, но менее 100% рецессивных аллелей в локусах, определяющих несходство шести сортов, сорта 2-й и 3-й – приблизительно 50-75%. 5-й и 6-й сорта – 50-75% доминантных генов, 4-й более 75%, но менее 100%.

Дополнительную информацию о наследовании несходства можно извлечь из некоторых корреляционных связей. Выявлены достоверные взаимосвязи парных (i, j) несходств шести сортов (S_{ij}) со следующими параметрами, используемыми в анализе:

- $(G_{ij}+G_{ji})/S_{ij}$ - уровнем гетерозиса их гибридов F_1 ($r=-0,80^{**}$),
- $|G_{ij}-G_{ji}|$ - уровнем доминирования ($r=-0,60^*$),
- $W_i+V_i+W_j+V_j$ суммами долей рецессивных аллелей этих пар сортов ($r=0,55^*$),
- модулями разностей тех же долей ($r=0,81^{**}$).

Таблица 3.7. Значения корреляций $r_{ij}=(rS_{ij}+rG_{ij})/2$, вычисленные по данным табл. 3.5. Нумерация сортов (i, j) из табл. 3.5.

$i \backslash j$	1	2	3	4	5	6
1	1	-0,32	-0,17	-0,59	-0,56	-0,55
2	-0,32	1	0	-0,34	-0,17	-0,07
3	-0,17	0	1	-0,35	-0,01	0,17
4	-0,59	-0,34	-0,35	1	0,48	0,31
5	-0,56	-0,17	-0,01	0,48	1	0,57
6	-0,55	-0,07	0,17	0,31	0,57	1

В соответствии с предложенным биометрико-генетическим методом для оценки степени гетерозиготности гибрида F_1 двух сортов (i, j) использованы полусуммы двух корреляций $r_{ij}=(rS_{ij}+rG_{ij})/2$, представленные в табл. 3.7. У пар сортов несходных по аллельному составу всех полиморфных локусов этот коэффициент должен равняться -1.

Однако на все генетико-статистические параметры и связи влияет модификационная изменчивость и случайные ошибки опытных данных, Их влияние не учитывалось при аналитическом и имитационном моделировании в процессе разработки биометрико-генетического метода. При анализе реальных экспериментальных данных оценки основных параметров должны быть скорректированы.

В частности, модификационная изменчивость не позволит проявиться генотипической корреляции, равной -1, для пары сортов с несходным аллельным составом по всем полиморфным локусам набора сортов. То есть, даже если истинное генотипическое значение корреляции равно -1, экспериментальная фенотипическая оценка $r_{ij}=(rS_{ij}+rG_{ij})/2$ будет не меньше $-H^2$, где H^2 – полусумма коэффициентов наследуемости параметров S_{ij} и $G_{ij}+G_{ji}$. Это следует из известной формулы для r_p – фенотипической корреляции признаков x и y (Рокицкий, 1974):

$$r_p = H_x H_y r_g + e_x e_y r_e \quad (3.6),$$

где r_g r_e – генотипическая и средовая корреляции признаков, $H = \sqrt{H^2}$, т.е. корень из коэффициента наследуемости в широком смысле; $e = \sqrt{1 - H^2}$.

Действительно, если оценивать влияние модификационной изменчивости отдельно на корреляцию rS_{ij} , то x и y – это значения метрики несходства (S_{ij}) в i -й строке и j -м столбце таблицы несходства сортов (табл. 3.5а). Средовые ошибки у этих значений одинаково распределены и не коррелируют, поэтому $H_x=H_y=H_S$ и $r_e=0$. Из формулы (3.6), следует, что $r_p \geq -H_S^2$. Аналогично, для оценок влияния модификационной изменчивости на корреляции rG_{ij} следует использовать H_G^2 значений $G_{ij}+G_{ji}$ из таблицы несходства гибридов (табл. 3.5б); в частности, для коэффициента фенотипической корреляции справедливо $r_p \geq -H_G^2$. Следовательно, минимально возможное значение фенотипической корреляции $r_{ij}=(rS_{ij}+rG_{ij})/2$ равно $-(H_S^2 + H_G^2)/2$.

Для получения значений S_{ij} и $G_{ij}+G_{ji}$ в трех повторениях был использован прием, описанный в начале раздела. По результатам дисперсионного анализа этих значений были получены оценки коэффициентов наследования: $H_S^2=0,83$, $H_G^2=0,74$. Следовательно, даже при различии аллельного состава двух сортов (i, j) по всем полиморфным локусам фенотипическое значение их параметра $r_{ij}=(rS_{ij}+rG_{ij})/2$ не может быть меньше $-(H_S^2 + H_G^2)/2=-0,79$.

Далее следует учесть, что рассуждения, связанные с формулой (3.6), справедливы, если выборка для оценки всех входящих параметров велика. В реальной ситуации, фенотипическая корреляция r_p (r_{ij}) оценивается по выборке объемом не больше числа родителей, так что оценки r_{ij} из матрицы типа табл. 3.7 получают дополнительную случайную ошибку. В частности, r_{ij} пары сортов, альтернативных по аллельному составу, может отклоняться от величины $-(H_S^2 + H_G^2)/2$. С помощью приема, описанного в начале раздела, была получена оценка коэффициента наследуемости параметра r_{ij} : $H_r^2 = 0,86$. Это значение мало отличается от H_S^2 и H_G^2 : из-за малого числа родителей точность рекомендаций по подбору пар на основе r_{ij} такая же, как по S_{ij} или $G_{ij} + G_{ji}$. Но, как отмечалось в разделе 3.1, параметр r_{ij} дает информацию о близости несходства двух сортов к альтернативному аллельному составу.

По данным табл. 3.7 минимальные фенотипические значения r_{14} , r_{15} и r_{16} больше $-0,79$ и мало различаются между собой. По-видимому, ни одна пара при скрещивании этих трех сортов (1×4 , 1×5 , 1×6) не даст максимально гетерозиготный гибрид F_1 . Дополнительно отметим, что у этих трех пар, судя по рис. 3.4, максимально велики контрасты $|\hat{W}_i + \hat{V}_i - \hat{W}_j - \hat{V}_j|$. Это означает, что доминантные и рецессивные гены, определяющие несходство сортов, у них распределены неравномерно.

Для приближения к полной аккумуляции генетического разнообразия всего набора из 6 сортов можно скрестить гибрид F_1 одной из трех названных пар родителей с еще одним сортом набора, максимально несходным с обоими родителями. В соответствие со свойством 4 парных корреляций r_{ij} (раздел (3.1)) подбор третьего сорта снова проводится по минимуму r_{ij} с первыми двумя. Так, по данным табл. 3.7, приемлемый вариант – скрещивание гибрида F_1 родителей 1 и 4 с сортом номер 2, который, судя по r_{ij} максимально (по сравнению с 3-м, 5-м и 6-м) несходен с 1-м и 4-м сортами.

Однако напомним, что свойство 4 было установлено при моделировании лишь для трех полиморфных локусов. Возможно, что при большем числе полиморфных локусов и родительских линий свойство 4 не выполняется: сумма их r_{ij} может оказаться меньше -1 . По этой причине, а также из-за наложения названных случайных ошибок оценок параметров r_{ij} и неполного соответствия данных другим модельным предположениям сумма $r_{14} + r_{12} + r_{24}$ оказалась меньше -1 .

Пример анализа несходства по модулю разности признака

Основой анализа послужили результаты диаллельного скрещивание 6 сортов озимой пшеницы: 1) Мироновская юбилейная, 2) Мироновская 808, 3) ППГ 186, 4) Карстен VIII, 5) Гейнес и 6) Кавказ. Использованные однолетние данные получены на селекционной станции РГАУ-МСХА (Федин и др., 1980; Смиряев, 2009). Схема опыта – два рандомизированных блока. Измеряли высоту растений на момент уборки. В качестве метрики парного несходства двух генотипов в каждом повторении диаллельной схемы использован модуль разности средней высоты (см) их растений.

Для проверки выполнения предположений метода реализован тест – дисперсионный анализ в двух повторениях только величин параметра $(V_i+W_i+V'_i+W'_i)/4$. Это связано с особым свойством выбранной простейшей метрики несходства – модуля разности единственного анализируемого признака. Можно показать, что в любой матрице S_{ij} значений модуля разности для пар родительских генотипов (i, j) , во-первых, всегда есть пара с максимальными и равными между собой значениями параметра σ^2_S , отличными от значений этого параметра у остальных генотипов. Это пара с предельно отличающимися значениями исходного признака. Однако такое отличие σ^2_S от остальных не означает, что схема наследования метрики не соответствует аддитивно-доминантной. Во-вторых, независимо от случайных ошибок опыта корреляция rS_{ij} для этой пары генотипов в точности равна -1, что, конечно, не означает несходства их аллелей по всем полиморфным локусам, влияющим на модуль разности признака.

Дисперсионный анализ показал отсутствие значимых различий параметра $(V_i+W_i+V'_i+W'_i)/4$ по родительским сортам (табл. 3.8).

Таблица 3.8. Дисперсионный анализ параметра $(V_i+W_i+V'_i+W'_i)/4$

	<i>ms</i>	<i>df</i>	F_p	$F_{t,05}$
сорта	366,70	5	2,57	5,06
повторения	1127,34	1	7,89	6,61
ошибки	142,8	5		

Дополнительно провели второй тест – сравнили линейные регрессионные зависимости W_i на V_i и W'_i на V'_i (табл. 3.9). Сравнение по критерию Фишера значений ms_0 для вариантов вычислений 3-го с 1-м ($F_p=2,24$) и 4-го со 2-м ($F_p=1,50$) показывает отсутствие значимых различий между регрессиями W_i на V_i и W'_i на V'_i ($F_{t,5\%}=6,16$).

Таблица 3.9. Результаты сравнение регрессий W_i на V_i и W'_i на V'_i

Вариант вычисления параметров		a	b	ms_0
1	W_i на V_i , оценки МНК	9,76	1,16	141,56
2	W'_i на V'_i , оценки МНК	-6,37	1,25	405,74
3	W_i на V_i , но a, b из варианта 2	-6,37	1,25	318,00
4	W'_i на V'_i , но a, b из варианта 1	9,76	1,16	607,52

Таблица 3.10. Значения метрики несходства (см) 6-и сортов озимой пшеницы и их гибридов F_1 . Номера сортов: 1) Мироновская юбилейная, 2) Мироновская 808, 3) ППГ 186, 4) Карстен VIII, 5) Гейнес, 6) Кавказ.

а) S_{ij} – несходство i -го и j -го сортов ($i, j=1, \dots, 6$).

$i \backslash j$	1	2	3	4	5	6
1	0	23,2	27,9	1,3	8,7	1,7
2	23,2	0	4,7	21,9	31,9	24,9
3	27,9	4,7	0	26,6	36,6	29,6
4	1,30	21,9	26,6	0	10,0	3,0
5	8,70	31,9	36,6	10,0	0	7,0
6	1,70	24,9	29,6	3,0	7,0	0
σ_S^2	147,9	156,7	220,5	128,1	220,5	165,3

б) несходство: G_{ij} – гибрида F_1 с i -м сортом, G_{ji} – гибрида с j -м сортом ($i \leq j$).

$i \backslash j$	1	2	3	4	5	6	V_i	W_i
1	0	14,5	16,1	4,3	0,2	4,7	49.2	76.8
2	8,7	0	9,1	5,2	4,1	4,0	11.4	7.2
3	11,8	4,1	0	3,0	5,8	12,1	23.6	44,9
4	3,0	16,7	23,6	0	6,4	1,5	90.2	106,0
5	8,9	27,8	30,8	16,4	0	16,0	132.4	159,8
6	6,4	20,9	17,5	1,5	9,0	0	71.5	101,3
V'_i	18,6	106,7	116,4	34,3	12,8	39,6		
W'_i	43,4	126,4	147,1	9,9	14,2	31,9		
\hat{V}_i	74,2	11,2	12,5	86,8	118,7	74,8		
\hat{W}_i	89,1	12,8	21,3	120,4	145,2	107,1		

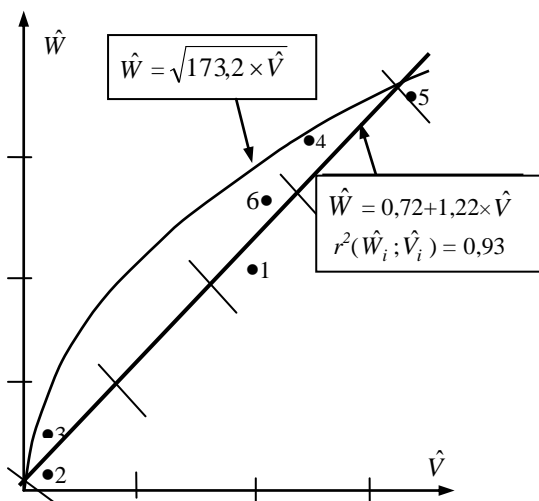


Рисунок 3.5. График для анализа значений метрики несходства 6-и родительских сортов и гибридов F₁ озимой пшеницы. Номера сортов из табл. 3.10а.

Исходные данные по высоте растений удовлетворяют требованиям метода, что позволяет оценить метрику несходства для средних по повторениям (табл. 3.10) и объединить две пары параметров W_i, V_i и W'_i, V'_i в одну \hat{W}_i, \hat{V}_i (две последние строки табл. 3.10б).

Регрессия \hat{W}_i на \hat{V}_i и ограничивающая парабола представлены на рис. 3.5.

Точки $(\hat{V}_i; \hat{W}_i)$ «растянуты» вдоль линии регрессии, следовательно, хотя бы по некоторым полиморфным локусам присутствуют эффекты доминирования. Сорта 2, 3 несут более 75%, но менее 100% доминантных аллелей в полиморфных локусах, сорта 1 и 6 – приблизительно 50-75% рецессивных генов. 4-й сорт – более 75%, 5-й почти 100%. Коэффициент регрессии $b > 1$. Следовательно, как показывает теоретический анализ (раздел 3.1), во-первых, в схеме наследования метрики несходства гетерозис изменчив по гибридам, во-вторых, эффекты доминирования существенно меняются по локусам.

Учитывая отмеченные выше особые свойства матрицы S_{ij} для выбранной метрики несходства (модуль разности единственного анализируемого признака), оценку гетерозиготности гибридов F₁ вели

только на основе значений параметра rG_{ij} , полученных из матрицы G_{ij} (табл. 3.11).

Дисперсионный анализ значений G_{ij} в двух повторениях дает оценку $H_G^2=0,88$. Можно сделать вывод, что минимум два гибрида F_1 с отрицательными параметрами rG_{ij} , близкими по модулю к этому значению, гетерозиготны практически по всем полиморфным локусам, а именно, 1×3 и 3×5 . Для последней пары контраст $|\hat{W}_i + \hat{V}_i - \hat{W}_j - \hat{V}_j|$ значительно больше, чем у первой. Следовательно, у 3-го и 5-го сортов доминантные и рецессивные гены распределены не равномерно.

Таблица 3.11. Значения корреляций rG_{ij} , вычисленные по данным табл. 3.10б.

$i \backslash j$	1	2	3	4	5	6
1	1	-0,70	-0,85	0,70	0,78	0,65
2	-0,70	1	0,75	-0,37	-0,76	-0,40
3	-0,85	0,75	1	-0,53	-0,84	-0,54
4	0,70	-0,37	-0,53	1	0,14	0,98
5	0,78	-0,76	-0,84	0,14	1	0,10
6	0,65	-0,40	-0,54	0,98	0,10	1

Заключение

Предложенный подход дает возможность проводить биометрико-генетический анализ любых метрик парной несходства генотипов, оцененных по значениям количественных признаков сортов и гибридов F_1 диаллельных скрещиваний чистых линий. Методика позволяет прогнозировать особенности изменчивости, ожидаемой в последующих поколениях. В частности, проводить подбор пар родительских сортов, обеспечивающих максимальное генетическое разнообразие потомства. Следует, однако, тщательно подбирать структуру метрики несходства. Так первая, более сложная метрика (несходство формы реакции) нацелена на оценку общего генетического разнообразия потомства. Вторая метрика – модуль различия одного признака позволяет прогнозировать разнообразие потомства именно по этому признаку.

Подход может быть объединен с другими методами. Так, используя первую метрику, при достаточно большом числе родителей в диаллельной схеме для подбора небольших их групп, генетически несходных в совокупности, можно также использовать кластерный

анализ. А именно, сначала объединять в кластеры генетически сходные генотипы – с максимальными положительными параметрами rS_{ij} , rG_{ij} . Затем скрещивать сорта из разных кластеров. Допустима также предварительная обработка комплекса признаков методом главных компонент, после чего к значениям отдельных компонент с учетом их смысловой интерпретации естественно применить вторую метрику.

Глава 4. Прогноз эффективности отбора в популяциях потомства по оценкам дивергенции родителей

Проведен анализ надежности нескольких косвенных количественных оценок генетической дивергенции родительских пар для прогноза эффективности отбора в популяциях их потомства. На основе результатов полевых экспериментов с мягкой яровой пшеницей проверяется упрощенное представление о положительных корреляционных связях генетического несходства родителей с доступной для отбора изменчивостью количественных признаков в потомстве. С помощью иллюстративных числовых и графических моделей проведен биометрико-генетический анализ изменчивости количественных признаков, связанных отношением «исходные–расчетные». Этот методический прием позволяет выявить сложные ситуации варьирования наборов полиморфных локусов, определяющих изменчивость признаков, и схем их наследования. Результаты анализа могут быть полезны в генетике количественных признаков и при планировании селекционного процесса.

До сих пор рассматривались, в основном, возможности прогноза степени гетерозиготности гибридов F_1 . Упрощенно предполагается: чем более несходны генетически родительские формы, тем шире полиморфизм популяции потомства по генам, определяющим изменчивость количественных признаков, в том числе хозяйственно ценных. То есть в популяции потомства за счет рекомбинации возникнет больше трансгрессивных вариантов, увеличатся среднеквадратические отклонения признаков, вследствие чего эффективность отбора новых удачных генотипов в потомстве должна повыситься.

Однако надежность подобных методов подбора родительских пар не всегда высока (Смирязев и др, 1992). Желательно в экспериментах сопоставить надежность подобных прогнозов, полученных на основе различных косвенных оценок генетической дивергенции родительских пар. Возможно, при анализе реально доступной для отбора изменчивости количественных признаков в поздних поколениях удастся понять причины снижения надежности прогнозов.

4.1. Проблема количественного сравнения родительских пар по генетическому несходству

Прежде всего, следует уточнить смысл выражения «степень

несходства» генотипов пары родителей в связи с уровнем изменчивости признаков, ожидаемой в расщепляющейся популяции их потомства.

Предположим, что в идеале полностью известен нуклеотидный состав каждой родительской формы – чистой линии одного вида. Достаточно ли опираться на общее генетическое несходство двух родителей, точнее на их так называемые генетические расстояния, определяемое с помощью методов биоинформатики (Смиряев, Панкина, 2008). Можно ли утверждать, например, для самоопылителей, что одна пара чистых линий «более несходна» по сравнению со второй, если у первой пары больше несходных нуклеотидов? Ожидается ли большее доступное отбору разнообразие по количественным признакам в позднем поколении – расщепленной популяции потомства от скрещивания первой пары?

По-видимому, не обязательно. Во-первых, потому, что большая часть генома у многих видов состоит из не транскрибируемой ДНК – по Cavalier-Smith (1985) от 30% до 95%. Возможно, несходство родителей в основном касается именно «молчащей» части генома. Во-вторых, даже если известно о большем несходстве нуклеотидных цепочек экспрессирующихся генов, нет гарантии, что большее генетическое расстояние родителей вызовет дополнительную изменчивость конкретных признаков в популяции потомства. Действительно, это несходство у двух пар родителей может касаться разных участков одних или разных генов. Кроме того, генетическая детерминация изменчивости количественных признаков в любой популяции потомства, как известно, существенно меняется по условиям выращивания (Драгавцев, 2003).

Далее следует учесть, что часто при косвенной оценке дивергенции учитываются различия родительских форм лишь по некоторым «удобным для изучения» маркерным локусам или нуклеотидным последовательностям, составляющим незначительную часть генома. Остается надеяться на положительную корреляционную связь подобных косвенных оценок общего несходства родителей (например, по маркерам на QTL или микросателлитным сайтам) с несходством «важных» генов, определяющих изменчивость хозяйственно ценных признаков.

Другой косвенный подход к подбору пар по степени дивергенции основан, как отмечено выше (глава 3), на несходстве фенотипов родителей. Точнее предполагается положительная корреляционная связь какой-либо метрики фенотипического несходства родителей с их несходством по «важным» генам ценных признаков.

Метрики для экспериментальной оценки дивергенции

Для оценки дивергенции 6 сортообразцов мягкой яровой пшеницы (табл. 3.5) использованы 2 метрики несходства, оцененные по микросателлитным локусам (см. ниже), а также 3 метрики, основанные на фенотипических проявлениях количественных признаков у родителей. Это метрики Махаланобиса и евклидова расстояния, а также метрика несходства формы реакции генотипов (главы 2 и 3).

Микросателлитный анализ 6 родительских сортообразцов проводил ст. научный сотрудник центра молекулярной биотехнологии РГАУ-МСХА к.б.н. М.Г. Дивашук. Были учтены 34 SSR-маркера, протестированных в предыдущих работах, по методикам описанным ранее (Крупин и др., 2011). Полученные длины микросателлитных сайтов использовали для косвенной оценки генетической дивергенции 15 пар 6 родителей. Основное предположение: чем более несходны длины сайтов с одинаковой локализацией у двух родителей, тем выше степень их общей генетической дивергенции. Несходство длин суммировалось по 34 сайтам.

Принимая во внимание возможные различия частоты мутаций разных микросателлитных сайтов, в частности, зависимость скорости изменения числа нуклеотидов от длины сайта (Watkins, 2007), была проверена надежность двух метрик генетической дивергенции 6 родителей. Первая – манхеттенское расстояние пары родителей имеет следующую структуру:

$$S_M = \sum_{n=1}^{\max} n \times m_n,$$

где m_n – число микросателлитных сайтов с одинаковой локализацией у двух родительских генотипов, но длиной, отличающейся на n нуклеотидов ($n=1..max$). Эта метрика не учитывает названные различия сайтов по скорости мутирования, что может вызвать преувеличенное влияние на S_M тех сайтов, которые в процессе эволюции и селекции меняют длину быстрее остальных.

Вторая метрика нивелирует эти различия скорости мутаций и оценивает лишь общее число микросателлитных сайтов с одинаковой локализацией у двух родителей, но разной длиной:

$$S_{01} = \sum_{n=1}^{\max} m_n.$$

Метрика S_{01} также имеет ограничение. В процессе создания

любого сорта с использованием рекомбинационной селекции в его геноме закрепляются участки хромосом от разных сортообразцов того же вида, входящих в родословную сорта. Эти сортообразцы в разной степени несходны между собой. В результате при оценке генетической дивергенции любой пары сортов с использованием микросателлитных сайтов их разная длина косвенно отражает несходство аллелей участков хромосом, дивергентных в разной степени. Метрика S_{01} не учитывает этот фактор т.к. игнорирует количественные различия длин микросателлитных сайтов с одинаковой локализацией.

Далее значения метрики Махаланобиса обозначены M , евклидово расстояния – E , несходства по форме реакции (главы 2, 3) – S , по сателлитам – S_M и S_{01} . Возможность использования пяти названных метрик парного несходства родителей для прогноза разнообразия, доступного для отбора по количественным признакам в популяциях потомства, далее проверена на основе экспериментальных данных.

Экспериментальный материал

В качестве материала использовали 6 родительских сортообразцов яровой мягкой пшеницы (названия в табл. 3.5) и 15 гибридных популяций, полученных при самоопылении гибридов F_1 их диаллельной схемы скрещиваний: F_3 (2009 г.), F_4 (2010 г.) и F_5 (2011 г.). Отбор в этих популяциях не проводился.

Ежегодно на полях лаборатории селекции и семеноводства полевых культур РГАУ-МСХА имени К.А. Тимирязева вручную проводили посев 6 сортообразцов и 15 гибридных популяций в 3 рандомизированных блоках. Каждая делянка состояла из пяти учетных и двух защитных рядков по 20 зерен в каждом. Размер делянки: $1\text{м} \times 1,26\text{м}$. Посев проводили поперек полосы шириной 1м с междурядьями 18 см , расстояния между соседними делянками 36 см , межполосная дорожка – 50 см . Уборка производилась вручную в фазе полной спелости, защитные и учетные рядки убирались отдельно.

Общее количество растений в учетных рядках из-за негативного влияния погодных условий разных лет колебалось в пределах от 70 до 100. На каждой делянке со всех пяти учетных рядков было использовано для измерений по 30 случайно выбранных растений. После уборки на них измеряли 7 признаков – элементов структуры урожая. Это высота растения (признак номер 1), длина колосового стержня главного колоса (2), число продуктивных колосков главного колоса (3), число зерен главного колоса (4), масса зерна главного колоса (5), число зерен боковых колосьев (6), масса зерна боковых

колосьев (7). Кроме того, анализировали три расчетных признака: число зерен с растения (8), масса зерна с растения (9), масса 1000 зерен (10).

Для посева семян любой гибридной популяции следующего поколения случайно выбирали одинаковое число семян каждого из 30 растений того же повторения предыдущего поколения. Благодаря этому влияние дрейфа генов было минимальным.

С целью экспериментальной оценки изменчивости 10 признаков в каждой из 15 популяций-семей F_3 , F_4 , F_5 отдельно вычисляли показатель \sqrt{D} – среднеквадратическое отклонение для каждого признака. Этот показатель оценивает реальное, доступное для отбора разнообразие признака в популяции. Кроме того, оценивали μ – среднее значения признака в каждой популяции.

Метрику Махаланобиса M парного несходства 6 родительских сортообразцов (всего 15 значений) оценивали по величинам всех их 10 признаков, каждый из которых был предварительно усреднен по данным за 4 года испытания. Для расчета 15 значений метрики M по этим усредненным данным использовали программу из пакета AGROS. Поскольку евклидово расстояние в отличие от метрики Махаланобиса не учитывает корреляционных связей между признаками, 15 значений парного евклидова расстояния 6 сортообразцов рассчитывали по той же программе, но все ковариации признаков задавали равными нулю.

15 значений метрики S – парного несходства формы реакции 6 родителей были оценены ранее в четырехлетнем эксперименте (S_{ij} в табл. 3.5)

Для оценки надежности прогнозов изменчивости признаков в потомстве по пяти метрикам-прогнозам (M , E , S , S_M и S_{01}) вычисляли r – коэффициенты корреляции между значениями этих метрик для 15-ти пар родителей и 15-ю экспериментальными значениями \sqrt{D} по каждому из 10 признаков в их потомстве (F_3 , F_4 или F_5).

Результаты и обсуждение: изменчивость корреляций

15 значений метрики Махаланобиса M и 15 значений евклидова расстояния E , полученные по одним четырехлетним данным о 6 родительских сортообразцах, имели между собой значимую, но невысокую корреляцию $r(E, M)=0,70$. Это можно объяснить тем, что в евклидовом расстоянии ковариации между 10 признаками считаются равными нулю. Отметим, что оценки этих ковариаций в метрике Махаланобиса, по-видимому, несут существенные статистические ошибки: слишком мала выборка родительских форм (6), по которой оценивались ковариации. Всего оценок ковариаций $10 \cdot 9/2=45$. Ясно,

что случайная ошибка 15 значений M , в расчете каждого из которых участвуют все 45 параметров одновременно, достаточно велика.

Метрики E и M слабо коррелировали с остальными тремя метриками несходства: $r(S, M)=0,39$, $r(S_M, M)=0,18$, $r(S_{01}, M)=0,52$, $r(S, E)=0,34$, $r(S_M, E)=0,20$, $r(S_{01}, E)=0,50$. Оценки корреляций остальных 3 метрик между собой в среднем выше: $r(S_{01}, S_M)=0,82$, $r(S, S_M)=0,48$, $r(S, S_{01})=0,46$, хотя две последние корреляции немного ниже критического уровня, равного 0,51 при $df=13$.

Далее вычисляли коэффициенты корреляции ($r \cdot 100$) метрик-прогнозов (M , E , S , S_M и S_{01}), косвенно оценивающих парное генетическое несходство 6 сортообразцов, с \sqrt{D} – наблюдаемой изменчивостью 10 признаков в их 15 гибридных популяциях: F_3 , F_4 , F_5 (табл. 4.1). В соответствии с обсуждаемым упрощенным представлением ожидали получить положительные корреляции. Однако результаты расчетов оказались неожиданными.

Например, для метрики S в F_3 единственный значимый коэффициент корреляции r проявился с изменчивостью продуктивности гибридных растений (9-й признак), причем r был отрицателен (-0,64*). В следующем поколении F_4 он также значимо отрицателен (-0,63*), а корреляции r с несколькими другими признаками несколько увеличились и также стали значимо отрицательными. То есть для них коэффициенты r имеют знак, противоположный по сравнению с ожидаемым по прогнозу S . Подобные отрицательные корреляции получены для остальных четырех метрик, хотя не все r значимы. Исключением стал признак 10 – масса 1000 зерен: в F_4 для него корреляция с S значимо положительна. Положительны корреляции для 10-го признака и с остальными четырьмя метриками в F_4 , хотя по трем из них корреляции незначимы по уровню.

Отмечена также особая тенденция изменения значимости корреляций показателя \sqrt{D} у признаков 6-9 с разными метриками в поколениях. Например, в F_3 и F_5 все эти корреляции с S незначимы (хотя и сходны с F_4). В тех же двух поколениях F_3 и F_5 корреляции \sqrt{D} тех же признаков с метриками M и S_{01} значимы. И, наоборот, в F_4 для S эти корреляции значимы, а для M и S_{01} – незначимы.

Следует изучить причины изменчивости корреляций r , проявившейся в расчетах с реальными экспериментальными данными. При этом можно выделить, как минимум, три типа изменчивости:

1) изменчивость по 3 годам корреляции r одной метрики при ее сопоставлении с показателем \sqrt{D} одного признака;

Таблица 4.1. Коэффициенты корреляции ($r \cdot 100$). Номера признаков – из раздела «Экспериментальный материал». Смысл параметра $\max = H_y \cdot 100$ пояснен в следующем разделе

Номер признака		1	2	3	4	5	6	7	8	9	10
F ₃ (2009)	<i>M</i>	12	-25	25	-46	-31	-60*	-62*	-61*	-65*	-19
	<i>E</i>	14	26	45	-29	-14	-79*	-66*	-79*	-69*	8
	<i>S</i>	-2	-32	-1	-4	-17	-23	-28	-4	-64*	-25
	<i>S_M</i>	65*	26	-2	08	-02	-40	-45	-36	-46	18
	<i>S₀₁</i>	56*	31	31	-23	-13	-67*	-73*	-67*	-74*	11
	max	94	79	82	65	40	70	70	75	62	86
F ₄ (2010)	<i>M</i>	27	5	26	3	4	-4	-26	-38	-24	12
	<i>E</i>	10	17	29	-2	0	-31	0	-28	-1	31
	<i>S</i>	-30	9	-1	-4	-52*	-53*	-52*	-57*	-63*	52*
	<i>S_M</i>	46	84*	44	20	7	-14	0	-16	-5	54*
	<i>S₀₁</i>	47	71*	51*	-1	5	-32	-11	-31	-13	43
	max	95	77	86	77	76	93	85	92	85	88
F ₅ (2011)	<i>M</i>	20	13	10	-9	-17	-76*	-64*	-76*	-63*	10
	<i>E</i>	28	43	7	20	9	-62*	-44	-61*	-41	24
	<i>S</i>	-15	19	19	21	-5	-41	-49	-34	-43	42
	<i>S_M</i>	40	47	-18	0	-30	-41	-42	-39	-42	55*
	<i>S₀₁</i>	65*	57*	1	1	-26	-72*	-70*	-73*	-69*	56*
	max	73	74	82	40	46	78	63	76	66	72

* - значимо на уровне $P < 0,05$

2) изменчивость корреляции r по 10 признакам (вплоть до изменения знака показателя r) для одной метрики в пределах одного года;

3) изменчивость корреляции r значений \sqrt{D} одного признака для данных одного года с оценками генетической дивергенции 6 родительских сортообразцов на основе 5 разных метрик.

Причины изменчивости первого типа

Одной из основных причин вариабельности коэффициента корреляции по трем годам является типичное для растений сильное взаимодействие генотип-среда по каждому признаку (см. раздел 1.1). Из примера на рис. 4.1 следует, что по годам меняются не только значения, но даже ранги показателя \sqrt{D} .

Напомним, что в отличие от показателя \sqrt{D} , значения метрики S для пар родительских форм были оценены ранее на основе 4-летних данных (табл. 3.5 *a*) и считаются постоянными как в поколениях, так и по признакам. То же относится к метрикам E и M . Метрики S_M и S_{01} , естественно, вообще не изменяются по годам. Поэтому наблюдаемая изменчивость рангов показателя \sqrt{D} 15 гибридных популяций в поколениях затрудняет экспериментальную оценку по r эффективности подбора пар с помощью любой метрики, отражающей генетическое несходство родительских форм.

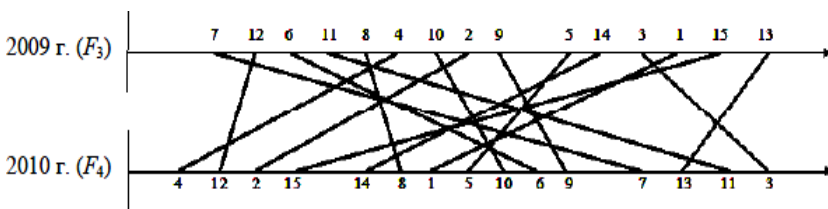


Рисунок 4.1. Пример изменения показателя \sqrt{D} для 15-и гибридных популяций по признаку «масса зерна с растения» за 2 года. Приведено к одному масштабу (Нгуен, Смирязев и др., 2010).

Второй причиной отсутствия значимых коэффициентов корреляции какого-либо признака, особенно в поколении F_3 (см. табл. 4.1), может выступать разрушение сформированных удачных генетических блоков несходных родительских генотипов вследствие их гибридизации, сопровождающейся рекомбинацией. Это может приводить к различию приспособленности и естественному отбору

гибридных генотипов в семьях – ранних поколениях самоопыления (Жученко, 1980; Шмальгаузен, 1968).

Третья возможная причина – снижение уровня остаточной гетерозиготности в 15 популяциях - семьях потомства, особенно при переходе от F_3 к F_4 .

Четвертая причина – возможное, причем разное по признакам, сцепление генов, определяющих их изменчивость. Из популяционной генетики известно (Bos, Caligari, 2008), что у самоопылителей сцепление может существенно влиять на частоты гомозиготных генотипов в поздних поколениях и, следовательно, на параметры \sqrt{D} .

В качестве еще одной причины варибельности r первого типа следует назвать изменчивость коэффициента наследуемости в широком смысле для показателя \sqrt{D} любого признака по годам. Это связано с модификационной изменчивостью признака и с разной степенью пестроты почвенного плодородия экспериментального поля по годам.

Действительно корреляции r_p , по сути, являются фенотипическими (обозначаются r_p), так как для их оценки использованы значения \sqrt{D} , усредненные по трем повторениям опыта. В главе 3 уже приводилась формула (3.6) для r_p двух показателей – x (в данном случае – значения какой-либо метрики парного несходства родителей) и y (в данном случае – показателей \sqrt{D} в популяциях потомства этих пар родителей):

$$r_p = H_x H_y r_g + e_x e_y r_e,$$

где r_g r_e – генотипическая и средовая корреляции двух показателей, $H = \sqrt{H^2}$ – корень из коэффициента наследуемости в широком смысле; $e = \sqrt{1 - H^2}$.

Если упрощенно считать, что значения любой метрики определены без модификационных отклонений, то $H_x=1$, $e_x=0$; тогда $r_p=H_y r_g$. Следовательно, модуль корреляции r_p для любого признака, даже если модуль истинной генотипической корреляции (r_g) близок к единице, не может быть больше значения H_y . Для любого года и признака величину H_y можно оценить из обычного дисперсионного анализа показателя \sqrt{D} в повторениях опыта, причем величина H_y не постоянна по годам. Значения $\max H_y \cdot 100$ приведены в табл.4.1.

Следует при этом отметить, что оценка H_y для каждого года и признака, а также сами оценки корреляций $r \cdot 100$, несут статистические ошибки из-за ограниченного объема выборки показателя \sqrt{D} (15

гибридов в трех повторениях) и неполного соответствия экспериментальных данных предположениям дисперсионного анализа (Смиряев и др., 2012). Поэтому модули некоторых значений $r \cdot 100$ в таблице оказались немного больше соответствующих оценок $\max = H_y \cdot 100$.

Причины изменчивости второго типа

Названные причины вариабельности коэффициентов корреляции r первого типа известны из литературных данных. Однако они не дают объяснения разнонаправленности корреляций отдельных метрик с показателем \sqrt{D} некоторых признаков в каком-либо конкретном поколении. Например, почему метрика S в F_4 достоверно положительно коррелирует с изменчивостью последнего признака, но достоверно отрицательно – с несколькими предыдущими (табл. 4.1). Это нельзя считать чистой случайностью, так как аналогичная ситуация повторяется для метрики S_{01} в F_5 .

Рассмотрим влияние на коэффициенты r эффектов генов, вызывающих скоррелированную изменчивость для пар самих количественных признаков. Эти эффекты генов отражены в коэффициентах их корреляций. Можно выделить два основных уровня парных корреляционных связей изучаемых исходных признаков: внутрисемейный (в пределах каждой семьи F_3 , F_4 или F_5) и межсемейный (рис. 4.2).

Первые – внутрисемейные корреляции признаков объясняются сцеплением генов, определяющих изменчивость признаков, и/или их плейотропными эффектами, а также, возможно, конкуренцией растений на делянке каждой популяции потомства. Для любой пары признаков в каждой из 15 семей оценивали коэффициенты внутрисемейных корреляции по измерениям 30 растений в трех повторениях, т.е. по 90 растениям.

Вторые (межсемейные) коэффициенты корреляции оценивают на основе обобщенных показателей, полученных внутри каждой из 15 семей-популяций в трех повторениях: μ – среднего значения признака в семье или \sqrt{D} – среднеквадратического отклонения признака в семье (рис. 4.2). Плейотропные эффекты и сцепление в этом случае влияют на межсемейные корреляции μ или \sqrt{D} опосредовано.

Имея данные в повторениях, каждую парную межсемейную фенотипическую ковариацию как между 15 значениями μ , так и \sqrt{D} , можно подразделить на генотипическую и средовую. После чего оценить коэффициенты корреляции: cor_p , cor_G , cor_e .

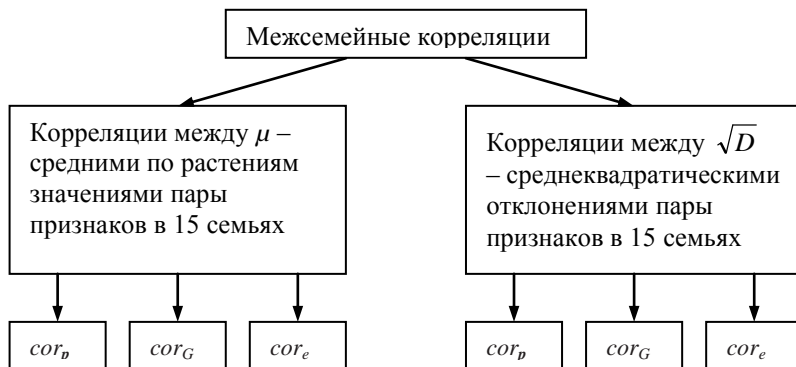


Рисунок 4.2. Структурная схема парных межсемейных корреляций признаков. cor_p , cor_G , cor_e – коэффициенты фенотипических, генотипических, средовых корреляций

Анализ межсемейных корреляций, например, в F_4 выявил плеяду четырех тесно коррелирующих по \sqrt{D} признаков с номерами 6, 7, 8, 9. Их оценки cor_p – парных фенотипических корреляций менялись в пределах от 0,85 до 0,98 (Нгуен, Смирязев и др., 2010). Причем cor_G – коэффициенты генотипических корреляций были даже несколько выше. То же относилось к корреляциям этих 4 признаков по μ – средним значениям в 15 семьях (от 0,73 до 0,98). Фактически названная группа из 4 признаков выступает как один признак, т.е. их изменчивость в основном определяется одними генами. Ситуация в F_5 аналогична, но корреляции несколько ниже.

Поэтому, в частности, каждый из 4 названных признаков в F_4 показал сходную по величине отрицательную корреляцию \sqrt{D} с S – метрикой парного несходства родителей по форме реакции (табл. 4.1). Корреляция S с \sqrt{D} в F_4 для 10-го признака (масса 1000 зерен), напротив, значимо положительная. Причем 10-й признак не коррелирует ни с одним из остальных девяти. Напомним, что 10-й признак является расчетным по отношению к паре других признаков. Его значение x_{10} для каждого растения внутри семьи равно:

$$x_{10} = x_9/x_8 \quad (4.1).$$

Внутрисемейные корреляции значений 9-го признака (x_9 – масса зерна с растения) с 8-м (x_8 – число зерен с растения) в пределах

каждой из 15 семей также весьма высоки: от 0,83 до 0,96. Следовательно, в любой семье можно достаточно надежно выразить признак x_9 через признак x_8 :

$$x_9 = a + b \cdot x_8 \quad (4.2),$$

где a, b – коэффициенты линейной регрессии.

Подставляя уравнение (4.2) в (4.1) и деля числитель в правой части на знаменатель, получаем:

$$x_{10} = a/x_8 + b \quad (4.3).$$

Из справочника Ллойда и Ледермана (1989) с учетом уравнения (4.3) получаем приблизительную оценку связи показателя \sqrt{D} для 8-го и 10-го признаков в пределах каждой семьи:

$$\sqrt{D}_{10} \approx |a| \cdot \sqrt{D}_8 / \mu_8^2 \quad (4.4),$$

где μ_8 – оценка математического ожидания 8-го признака в семье.

Учитывая соотношение (4.4) для 15 семей в предположении слабых колебаний коэффициента a по семьям (что справедливо при выявленной тесной межсемейной корреляции между \sqrt{D}_8 и \sqrt{D}_9), можно сделать следующий вывод. Если при изменении \sqrt{D}_8 по семьям знаменатель (μ_8^2) в уравнении (4.4) меняется скоррелированно с \sqrt{D}_8 , но растет быстрее, чем \sqrt{D}_8 , то уменьшение \sqrt{D}_8 по семьям сопровождается увеличением \sqrt{D}_{10} в них. Этих условий достаточно для возникновения положительной корреляции 15 значений S с \sqrt{D}_{10} при отрицательной корреляции S с \sqrt{D}_8 . Если же знаменатель в (4.4) растет приблизительно синхронно с числителем, то, несмотря на корреляцию \sqrt{D}_8 с S , значимой корреляции \sqrt{D}_{10} с S не будет. То есть связь \sqrt{D}_{10} в семьях с изменением генетической дивергенцией 15-и пар родительских сортообразцов не проявится ($r \approx 0$). Этот вывод справедлив не только для метрики S , но и для любой другой метрики генетической дивергенции родителей.

Отметим, что из уравнения (4.1) следует: изменчивость 10-го признака определяется теми же генами, что 8-го и 9-го. Напомним, что два последних признака относятся к одной плеяде: они тесно коррелируют между собой как в каждой семье, так и по 15 семьям. То есть доступная для отбора изменчивость (\sqrt{D}) всех трех признаков (8,

9, 10), хотя и определяется практически одними генами, но их проявление различается по признакам. То же относится к показателю μ_8 – значению признака, среднему по растениям семьи. Например, в пределах каждой семьи μ_8 вычисляется по тем же экспериментальным данным (измерениям x_8), что и $\sqrt{D_8}$ – среднеквадратическое отклонение признака. Поэтому изменчивость обоих показателей зависит от одних генов, но проявляющихся по-разному в μ_8 и $\sqrt{D_8}$.

Возможно, изменение знаков у коэффициентов корреляции для эффектов гетерозиса трех аналогичных взаимосвязанных признаков в первой строке табл. 1.8 объясняется подобной причиной, хотя структура алгебраической связи этих эффектов более сложная.

С помощью олигогенных биометрико-генетических моделей (табл. 1.6) несложно показать, что в подобных ситуациях схемы наследования трех признаков (8, 9, 10) внутри семей различаются. Например, если в какой-либо семье схема наследования для 8-го и 9-го признаков аддитивно-доминантная, то в схеме 10-го признака почти наверняка проявится эпистаз. Аналогична ситуация для межсемейной изменчивости показателей μ и \sqrt{D} любого расчетного признака. Подобные различные изменения схем наследования признаков будут рассмотрены в разделе 4.2 на основе простых числовых примеров.

В более общем случае, без предположения о тесной корреляции по семья между $\sqrt{D_9}$, $\sqrt{D_8}$ и μ_8 , с использованием оценок из справочника (Ллойд, Ледерман, 1989) можно показать:

$$\sqrt{D_{10}} \approx |V_9 - V_8| \cdot \mu_9 / \mu_8,$$

где $|V_9 - V_8|$ – модуль разности коэффициентов вариации 9-го и 8-го признаков.

При определенных вариантах совместного варьирования показателей 9-го и 8-го признаков положительная корреляционная связь $\sqrt{D_{10}}$ в семьях потомства с любой метрикой генетической дивергенции родителей приведет к отрицательной связи последнего с показателями $\sqrt{D_8}$, $\sqrt{D_9}$ или к отсутствию корреляционной связи с ними. Все зависит от различий экспрессии генов, определяющих изменчивость показателей μ , \sqrt{D} , V 8-го и 9-го признаков.

Существуют и другие формы связи между исходными и расчетными признаками. Например, x_0 – общую массу растения можно вычислить как сумму x_1 – массы полезной части, например колоса, и x_2 – остальной:

$$x_0 = x_1 + x_2.$$

Если внутрисемейные коэффициенты корреляции между x_1 и x_2 близки к -1 (например, из-за внутрисемейного полиморфизма по аттракции пластических веществ из соломы в колос (Смирязев и др., 1992)), то в любой семье показатель $\sqrt{D_0}$ будет связан с $\sqrt{D_1}$ и $\sqrt{D_2}$ приблизительно следующим образом:

$$\sqrt{D_0} \approx |\sqrt{D_1} - \sqrt{D_2}|.$$

Кроме того, положительная межсемейная регрессионная связь $\sqrt{D_1}$ и $\sqrt{D_2}$ при их одновременном увеличении может привести к снижению модуля разности, что будет означать снижение $\sqrt{D_0}$ в семьях. Коэффициенты корреляции любой метрики генетической дивергенции родителей с показателями $\sqrt{D_1}$ и $\sqrt{D_2}$ в семьях потомства снова будут иметь противоположный знак по сравнению с $\sqrt{D_0}$.

Из этих примеров следует, что биометрико-генетический анализ позволяет сформулировать различные условия, при которых полезная для отбора изменчивость расчетного количественного признака у потомков не будет коррелировать с конкретной метрикой генетического несходства родителей или будет иметь обратный знак корреляции по сравнению с исходными признаками. В определенных ситуациях именно эти причины, а не недостатки конкретных косвенных оценок дивергенции родителей, объясняют разнонаправленность корреляций. Это проявляется, например, в значимых корреляциях метрики S с изменчивостью (\sqrt{D}) в семьях для одних признаков при отсутствии корреляций той же метрики S с \sqrt{D} для других признаков. Аналогичное соотношение корреляций r метрики-прогноза S_{01} в F_5 с реальным показателем \sqrt{D} 10-го признака и с \sqrt{D} двух предыдущих признаков (табл. 4.1), по-видимому, объясняется теми же причинами.

Следует заключить, что увеличение генетической дивергенции родителей, т.е. общего несходства их аллельного состава, не гарантирует увеличения полезного для отбора разнообразия в потомстве по всем изучаемым количественным признакам. Для части признаков возможно отсутствие реакции или даже закономерное снижение ответной изменчивости в потомстве.

Причины изменчивости третьего типа

Следует подчеркнуть, что сравнивалась эффективность метрик, основанных на качественно отличных данных о 6 родительских сортообразцах. Так, S_M , S_{01} – отражают несходство длин их микросателлитных сайтов, M , E – несходство фенотипических значений 10 признаков сортообразцов на момент уборки, а также данных, полученных в онтогенезе по одному признаку – высоте растений (метрика S). Несмотря на эти качественные отличия, ни разу не проявилась ситуация, когда среди пяти метрик для одного признака в одном поколении были как значимо положительные, так и значимо отрицательные значения r . Но величины корреляций в табл. 4 менялись существенно. Максимальная вариабельность наблюдалась для 5-го признака в F_4 : коэффициент корреляции r для пяти метрик менялся от 0,07 до -0,52*.

Наблюдаемые различия корреляций r разных метрик с показателем \sqrt{D} одного признака в один год естественны и вызваны причинами, аналогичными рассмотренными в предыдущих двух разделах. Прежде всего, такая вариабельность r третьего типа указывают на ассоциацию пяти метрик с разными полиморфными генами, как сильно, так и слабо влияющими на изменчивость конкретного признака. Причем некоторые из этих ассоциаций наверняка носят нелинейный характер.

Кроме того, лишь упрощенно можно считать, что величины исследуемых метрик несходства родителей определены без ошибки. На самом деле они имеют различные и трудно оцениваемые ошибки.

Так, существенную ошибку выборочности вызывает ограниченное число проанализированных локусов при оценке метрик, вычисленных по молекулярным маркерам (Watkins, 2007). Чем меньше таких маркеров, тем в среднем менее равномерно в метрике отражено несходство разных участков генома родительских сортообразцов.

Отметим, что метрика S_M во всей табл. 4.1 имеет лишь 4 из 30 корреляций, значимо отличающихся от 0, тогда как для S_{01} 14 из 30 корреляций значимы. По-видимому, ошибка метрики S_M при оценке генетического несходства 6 сортообразцов больше, чем у S_{01} , что обусловлено их особенностями. Например, метрика S_M , как отмечалось выше, не учитывает различия микросателлитных сайтов по скорости мутирования.

Метрики M и E , основанные на несходстве количественных признаков родительских сортов, не учитывают известного явления конвергентного влияния научной селекции (Смиярев и др., 1992). А именно, сходство требований ко многим признакам сортов одной

культуры (модель идеального сорта) часто приводит к их фенотипическому сходству, в частности, по признакам их габитуса. Но это далеко не всегда означает сходства сортов по генотипам. Например, короткостебельность сортов может быть обусловлена разными сочетаниями аллелей в локусах с разной локализацией. Метрики M и E этого не учитывают: в них близость родительских форм по величинам количественных признаков подразумевает сходство этих форм по аллельному составу.

Возможно, и по этим причинам возникает вариабельность корреляций r 3-го типа, а многие корреляции в табл. 4.1 незначимо отличаются от 0.

По-видимому, бессмысленно искать «самую лучшую» метрику, гарантирующую надежный прогноз разнообразия потомства по всем хозяйственно ценным количественным признакам в любых условиях выращивания. При выборе метрик для практического использования необходимо учитывать названные ограничения и недостатки.

4.2. Моделирование наследственной изменчивости расчетных признаков

Основным средством изучения биологических объектов является анализ признаков, т.е. различительных знаков этих объектов. Человек выделяет те признаки и параметры, которые он считает естественными, существенными, важными (например, для таксономии) и т.п. Часто предпочтение отдается наглядным признакам или более удобным для измерения. Можно выделить признаки – индикаторы состояния живых объектов и систем (в медицине, экологии), а также связанные с потребностями человека (в сельском хозяйстве).

Подобные «естественные» признаки лишь опосредованно, косвенно и фрагментарно отражают комплексные процессы, протекающие в клетке, организме, популяции и в более сложных системах. Вряд ли именно эта мозаика признаков важна для самих биологических систем – продуктов эволюционного процесса, с которыми имеют дело генетики и селекционеры. Эволюционный процесс опирается на свои критерии существования, функционирования, выживания, превосходства, недостаточно понятные человеку.

Например, очевидно, что естественный отбор зависит от фенотипа, который человек определяет как совокупность признаков. Но каких именно, в каких комплексах и соотношениях? Длительные попытки ученых выразить через признаки и параметры основной принцип эволюции, который в упрощенной формулировке звучит как

«выживает лучшее», не дали четких результатов. Не удивительно, что искусственный отбор по отдельным признакам и комплексам часто вступает в противоречие с естественным отбором. В генетике и селекции проблема осмысления понятия «ген для фенотипа» остается дискуссионной (Kaplan J.M., Pigliucci, 2001).

Как уже отмечалось, выше у хозяйственно ценных признаков растений характер генотипической изменчивости в популяции весьма лабилен при изменении внешних условий, в онтогенезе и в зависимости от генотипов родителей. Меняется схема наследования количественных признаков, экспрессия отдельных генов, вплоть до смены спектров полиморфных локусов. Генетическая основа изменчивости признаков и их комплексов ускользает от исследователя.

Современный подход, основанный на представлении о генных сетях, формирующих признаки, снимает, казалось бы, подобные проблемы, позволяет опереться на фундаментальные знания биологии, в т.ч. молекулярной (Смирязев, Панкина, 2008а). Однако, пока это скорее теоретические построения, моделирующие генетическую детерминацию лишь отдельных признаков и функций живого организма. Генные сети формирования количественных признаков, когда они будут описаны, окажутся чрезвычайно сложными для практического использования. В селекционно-генетических исследованиях по-прежнему приходится использовать традиционный анализ не детерминации, а лишь изменчивости количественных признаков в популяции.

В предыдущем разделе при анализе экспериментальных данных были выявлены частные ситуации, когда корреляция метрик несходства родительских пар с изменчивостью исходных и расчетных признаков в популяции потомства менялась не только по величине, но и по знаку. Далее с помощью простых числовых и графических моделей изучен характер перестройки генетической изменчивости количественных признаков, связанных между собой отношением «исходные–расчетные». Этот методический прием позволяет глубже понять проблемы традиционного анализа биологических объектов на основе их фенотипических проявлений, а также уточнить некоторые рекомендации при планировании селекционно-генетических исследований.

Структура и генетический полиморфизм исходных и расчетных признаков

Под расчетным признаком далее понимается величина, полученная с помощью математического (арифметического,

алгебраического и т.п.) преобразования значений одного или нескольких исходных признаков, т.е. признаков непосредственно измеренных на биологическом объекте. Простейший пример из раздела 4.1: в селекции зерновых культур при оценке элементов структуры урожая определяют массу зерна растения (M_3) и число зерен ($Ч_3$), делят первый исходный признак на второй и получают расчетный – массу одного зерна (M_1) или массу 1000 зерен. В селекционно-генетических исследованиях используют также различные индексы (уборочный индекс, селекционные индексы и т.п.), которые фактически являются расчетными признаками.

Методическое достоинство анализа расчетных признаков – возможность прогноза их генотипической изменчивости на основе изменчивости исходных признаков, хотя, как будет продемонстрировано ниже, не во всех случаях.

По числу исходных и расчетных признаков можно выделить несколько отличающихся вариантов:

а) Один исходный и один расчетный признак. Например, в генетических исследованиях иногда используют преобразованные (расчетные) значения измеренного признака, например, анализируют их логарифмы. В отдельных случаях этот прием позволяет упростить схему наследственной изменчивости признака или адаптировать признак к требованиям биометрико-генетического анализа (Мазер, Джинкс, 1985).

б) Несколько исходных признаков, один расчетный. Пример: длина (d) и ширина (b) листа – два исходных признака. Расчетный – площадь листа (S), оцениваемая в ботанике и селекции по известной формуле $S = d \times b \times k$, где $k = 2/3$ – для многих видов растений. В разделе 1.1 эта формула была использована при изучении роста площади флагового листа в онтогенезе у ячменя.

в) Один или несколько исходных и несколько расчетных признаков. Например, исходный – урожайность сорта, полученная за ряд лет испытания набора сортов. Два расчетных – среднегодовалая урожайность сорта (оцененная по среднему арифметическому исходной выборки) и степень нестабильности урожайности (по дисперсии той же выборки). Учитывая явление переопределения генетической формулы исходного количественного признака в наборе сортов, однолетние урожайности можно рассматривать как несколько исходных признаков.

Следует отметить относительный характер понятий «исходный» и «расчетный» признаки. Можно, например, прямо отсчитать 100 зерен на растении и взвесить их – признак M_1 становится исходным. Тогда, перемножив M_1 на второй исходный – $Ч_3$, получаем оценку признака

Мз, который ранее считали исходным; теперь он становится расчетным.

Обычно генотипическая изменчивость нескольких количественных признаков в популяции определяется как разными, так и, частично, одинаковыми полиморфными локусами. С помощью простых моделей на числовых примерах проиллюстрируем изменения спектра полиморфных локусов у признаков, связанных отношением «исходные-расчетные». Для простоты ограничимся следующими модельными предположениями.

– Рассматриваются популяции диплоидной самоопыляющейся культуры, полученные в поколениях без отбора от скрещивания двух чистых линий. Например, после скрещивания типа $A_2A_2B_1B_1 \times A_1A_1B_2B_2$ в позднем поколении без отбора и сцепления у самоопылителей остаются четыре гомозиготных генотипа: $A_1A_1B_1B_1$, $A_1A_1B_2B_2$, $A_2A_2B_1B_1$, $A_2A_2B_2B_2$ с частотами по 0,25.

– Генотипическая изменчивость каждого количественного признака в популяции определяется малым числом олигогенов с двумя аллелями. Подобное предположение нередко используется в генетических исследованиях в связи с понятием «гены количественного признака» (QTL) (Liu et al., 2003; Long et al., 2006; Wu, Lin, 2006).

– Средовые модификации признака в каждом условиях выращивания малы по сравнению с генотипической изменчивостью.

Укрупнено, возможны три варианта связи генетического полиморфизма исходных и расчетных признаков:

1) Если изменчивость первого исходного признака зависит от одного набора локусов, а второго от совершенно другого, то изменчивость расчетного – от объединенного набора этих локусов.

2) Если изменчивость каждого из двух или нескольких исходных признаков определяется одним набором полиморфных локусов, то изменчивость расчетного признака, обычно, определяется тем же набором. Но в особых случаях этот набор может сократиться. Например, возможны ситуации, когда изменчивость двух исходных признаков зависит от полиморфного локуса, а расчетного – нет.

Пусть в популяции потомства величины каждого из двух исходных признаков (I, II) у четырех гомозигот, названных выше, можно выразить через произведение двух чисел. Первое число – влияние локуса A, второе – локуса B, причем первое одинаково для обоих признаков, второе – различается. Расчетный признак $III = I/II$. В табл. 4.2 показано, что на изменчивость расчетного признака III из двух локусов влияет только один – B. Причина в том, что каждый исходный признак примера, по существу, является расчетным – произведением

двух «внутренних» признаков, изменчивость каждого из которых зависит только от одного локуса – A или B . Причем первый «внутренний» признак-сомножитель, зависящий от A , одинаков для измеряемых признаков I и II, и при делении $\text{III}=\text{I}/\text{II}$ одинаковые числителя первого «внутреннего» признака-сомножителя на расчетный признак III сокращаются. В результате на изменчивость признака III локус A уже не влияет.

Таблица 4.2. Изменчивость двух исходных (I, II) и одного расчетного ($\text{III}=\text{I}/\text{II}$) признака. Числа – влияние каждого из двух локусов, указанных в скобках. Значения моделируемого признака – после знака «=».

I	B_1B_1	B_2B_2	II	B_1B_1	B_2B_2
A_1A_1	$3(A)2(B)=6$	$3(A)4(B)=12$	A_1A_1	$3(A)3(B)=9$	$3(A)1(B)=3$
A_2A_2	$5(A)2(B)=10$	$5(A)4(B)=20$	A_2A_2	$5(A)3(B)=15$	$5(A)1(B)=5$

$\text{III}=\text{I}/\text{II}$	B_1B_1	B_2B_2
A_1A_1	$(3 \times 2)/(3 \times 3)=2/3$	$(3 \times 4)/(3 \times 1)=4$
A_2A_2	$(5 \times 2)/(5 \times 3)=2/3$	$(5 \times 4)/(5 \times 1)=4$

3) Если изменчивость двух исходных признаков определяется как общими, так и разными полиморфными локусами, то изменчивость расчетного зависит от объединения последних. Но, что касается общих локусов, то возможны неожиданные ситуации, подобные рассмотренной в табл. 4.2.

Так в примере с расчетным признаком $M1=M3/Ч3$, если изменчивость каждого исходного зависит от своего полиморфного локуса (A для $M3$, B для $Ч3$), то изменчивость $M1$ определяется обоими локусами. Но, как отмечалось выше, понятия «исходные» и «расчетные» признаки относительны.

Если исходно измерить признаки $M1$ и $Ч3$, то расчетным станет признак $M3=M1 \times Ч3$. Генетический анализ, естественно, покажет, что на $M1$ влияют оба локуса, на $Ч3$ только их общий – B . Но из этого не следует, что на изменчивость расчетного $M3$ влияет набор из двух локусов. Локус B , как предполагалось изначально, на изменчивость $M3$ не влияет, и прямой генетический анализ это подтвердит. Формально влияние локуса B сократится при умножении $M1 \times Ч3$. Этот вывод можно подтвердить, анализируя числовой пример, представленный далее в табл. 4.5.

Аналогичные ситуации проявятся, когда один из исходных признаков по смыслу включает расчетный как часть. Например,

исходные – общая масса растения (M_0) и масса соломы (M_c), расчетный – масса колосьев $M_k = M_0 - M_c$. Если изменчивость M_k в популяции зависит, в основном, от одного набора локусов, M_c – от другого, то изменчивость M_0 , естественно, – от объединения этих наборов. Но, не зная этого, исходя только из связи $M_k = M_0 - M_c$, можно ошибочно предположить, что изменчивость M_k также зависит от всего набора полиморфных локусов признака M_0 . С другой стороны, если расчетный – $M_0 = M_k + M_c$, то подобной ошибки не возникает: изменчивость M_0 определяется объединением наборов полиморфных локусов M_k и M_c .

В сложных ситуациях, когда трудно понять биологическую взаимосвязь исходных и расчетных признаков, гарантированную оценку полиморфности может дать только генетический анализ каждого признака.

Перестройка схем наследования признаков

На протяжении многих десятилетий в ходе фенотипического анализа признаков у различных видов известные авторы: Н.В. Тимифеев-Ресовский (1925), М.М. Камшилов (1935), В.А. Драгавцев (2009), R. Fisher (1958) и др. неоднократно отмечали, что характер доминирования и плейотропии часто зависит не от свойства гена, а определяется всем процессом развития организма в конкретных внешних условиях. Некоторые выводы были также подтверждены с использованием генетико-математического моделирования.

Например, моделировались значения y – измеряемого количественного признака у двух чистых линий P_1 и P_2 и их гибрида F_1 (Смирязев, 1985). Родительские линии различались аллельным состоянием одного стабильно активного гена с двумя аллелями: $P_1(A_1A_1) \times P_2(A_2A_2) \rightarrow F_1(A_1A_2)$. Эффект этого гена имел аддитивную схему наследования и рассматривался как «внутренний» признак. Предполагалось, что в ходе онтогенеза – на пути от генотипа к величине растущего измеряемого признака (y) влияние «внутреннего» признака преобразуется одинаково по трем генотипам, но нелинейно (по аналогии с рис. 1.5). Стрелками на рис. 4.3 указаны направления роста измеряемого признака (y) относительно обобщенного индекса V . Последний отражает влияние внешних условий конкретных лет испытания в экологических градиентах и онтогенезе, как в регрессии на средние (глава 2).

При совместном анализе P_1 , P_2 , F_1 в разных участках диапазона величины V наблюдалось различные схемы наследования измеряемого признака y . Если в один год рост завершился в точке A , то значение

признака у гибрида F_1 находится точно посередине между родителями – схема наследования аддитивная. Если рост завершился в точке D_1 или D_2 , то схема наследования – полное доминирование; если – в диапазоне D_1 – D_2 , то сверхдоминирование. «Окончательная» схема наследования на момент уборки зависит от условий выращивания.

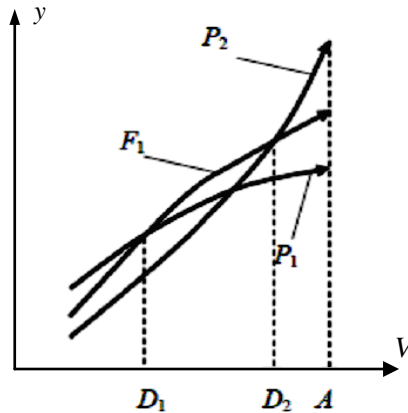


Рисунок 4.3. Результаты моделирования схемы наследования признака у в экологических градиентах и онтогенезе (Смиряев, 1985).

В такой ситуации нет переопределения спектра полиморфных локусов, а есть лишь изменение экспрессии единственного стабильно активного гена в разных внешних условиях. Причина изменения схем наследования измеряемого признака (y) – нелинейное преобразования в онтогенезе «внутреннего» признака с аддитивной схемой наследования на пути к измеряемому. В монографии А.А. Жученко «Экологическая генетика культурных растений» (1980) приведена таблица 24, содержащая десятки ссылок на работы, где в экспериментах с разными культурами для разных признаков наблюдалась вариабельность схемы наследования.

Итак, ранее был получен вывод, что часто доминирование, рецессивность, плейотропия – это свойства не гена, а изменчивого количественного признака. С помощью иллюстративных числовых и графических моделей признаков, связанных отношением «исходные-расчетные», далее будет показано, что подобная формулировка может быть распространена на эпистаз. Причем условия выращивания в моделях для простоты будут предполагаться постоянными.

Напомним, что под термином эпистаз в биометрической генетике понимаются любые межлокусные взаимодействия. Это означает, что численно эффекты влияния нескольких полиморфных

локусов на величину изучаемого количественного признака не просто суммируются: вклад одного гена зависит от вкладов других.

Напомним также структуру олигогенных биометрико-генетических моделей. Если изменчивость количественного признака в популяции определяется только одним локусом (A) с двумя аллелями (A_1, A_2), то генотипическая ценность трех генотипов в рамках F_∞ -метрики имеет следующую структуру. $A_1A_1: m+a; A_1A_2: m+d; A_2A_2: m-a$. Здесь m – среднее значение признака у двух гомозигот, $\pm a$ – эффект плюс и минус аллелей A_1 и A_2 , d – эффект доминирования. При двух полиморфных локусах, влияющих на изменчивость количественного признака, генотипическая ценность девяти генотипов приводилась в табл. 1.6. Эпистаза в схеме наследования признака нет, если $i=j=l=0$.

Наличие эпистаза в схеме наследования признаков затрудняет биометрико-генетический анализ, т.к. число параметров в схеме наследования резко возрастает. Это заметно даже при двух полиморфных локусах, где приходится оценивать четыре эффекта межлокусного взаимодействия (табл. 1.6). Обычно в биометрико-генетическом анализе эффективности отбора в расщепляющихся популяциях для простоты предполагается отсутствие эпистаза (Bos, Caligari, 2008). Однако более детальный анализ показывает, что наличие эпистаза в схеме наследования признака может значительно продлить эффективный отбор в поколениях (Rasmusson, Philips, 1997; Liu et al., 2003). Причина в том, что при эпистазе структура общей генотипической изменчивости признака кроме дисперсий аддитивных и доминантных эффектов включает дисперсии эффектов межлокусных взаимодействий. Таким образом, проблема эпистаза имеет как теоретическое, так и практическое значение.

Один исходный – один расчетный признак

В разделе 1.3 уже приводился пример возникновения «эпистаза-трансформации» по одному расчетному признаку y – высоте при отсутствии эпистаза для одного исходного «внутреннего» признака x – содержания ауксина (рис. 1.5). Этот тип взаимодействия генов ранее был рассмотрен в работах (Rasmusson, 1933; Forkmann, Seyffert, 1977). В частности, когда в поздних поколениях популяции самоопылителей остаются лишь четыре гомозиготных генотипа, в схеме наследования расчетного признака y возникает взаимодействие «гомозигота-гомозигота» ($i \neq 0$) при отсутствии эпистаза у исходного признака x .

Иногда удастся подобрать нелинейное преобразование величин признака x – изменить шкалу измерения таким образом, чтобы схема наследования расчетного признака y приблизилась к аддитивно-

доминантной (раздел 1.3). Различные нелинейные преобразования исходных экспериментальных данных, в частности $\ln(x+1)$, используются в биометрической генетике, например, для выравнивания средовых дисперсий (σ_e^2) перед применением дисперсионного анализа (Смиряев, Гохман, 1985). Фактически $\ln(x+1)$ – это новый расчетный признак. Однако если в схеме наследования исходного признака x не было эпистаза, то после преобразования, т.е. в новой шкале, эпистаз появится.

Еще одним примером ситуации «один исходный, один расчетный признак» являются аллометрические регрессионные зависимости двух количественных признаков, используемые при изучении онтогенеза животных. Часто в качестве исходного признака рассматривается длина отдельного органа или всего туловища, а расчетного – его масса (Зотин, 2009).

Несколько исходных – один расчетный признак

Для простоты в рамках двухлокусной модели рассмотрим поздние поколения самоопылителей, где остаются четыре гомозиготных генотипа. Структура генотипической ценности их признака представлена в четырех углах табл. 1.6. Присутствие эпистаза для гомозигот ограничивается эффектом $i \neq 0$.

Вернемся к примеру, где масса 1000 зерен (M_1 – расчетный признак) определяется как результат деления M_3 – массы зерна с растения на $Ч_3$ – число зерен на растении (два исходных признака). Пусть изменчивость обоих исходных признаков определяется двумя локусами (A, B), и в схеме наследования у четырех гомозигот для каждого из двух исходных признаков эпистаза, т.е. межлокусного взаимодействия, нет ($i=0$).

Зададим конкретные числа – эффекты двух локусов для исходных признаков по четырем гомозиготным генотипам из табл. 4.3. Для M_3 : $m_{M_3}=10, a_{1M_3}=1, a_{2M_3}=2, i_{M_3}=0$. Для $Ч_3$: $m_{Ч_3}=8, a_{1Ч_3}=-0,5, a_{2Ч_3}=1, i_{Ч_3}=0$. При подстановке этих параметров в табл. 1.6 получаем две табл. – 4.3а и 4.3б со значениями исходных признаков для четырех гомозиготных форм. Например, для гомозиготы $A_2A_2B_1B_1$ признака M_3 : $m_{M_3}-a_{1M_3}+a_{2M_3}=11$. Затем по формуле $M_1=M_3/Ч_3$ вычислим значения признака M_1 для тех же гомозиготных генотипов (табл. 4.3в).

Очевидно, что изменчивость расчетного признака M_1 определяется тем же набором локусов (A, B), что и у исходных. Приравнивая четыре значения из табл. 4.3в к структурам генотипической ценности четырех гомозигот из табл. 1.6 вычисляем четыре параметра для признака M_1 : $m_{M_1}=1,25, a_{1M_1}=0,20, a_{2M_1}=0,09,$

$i_{M1}=-0,02$. В схеме наследования признака $M1$ появился эпистаз (межлокусное взаимодействие) типа «гомозиготахгомозигота» при отсутствии эпистаза у двух исходных признаков.

Таблица 4.3. Генотипические ценности двух исходных признаков M_3 , $Ч_3$ (a , b), зависящих от одного набора из двух полиморфных локусов, и расчетного признака $M1=M_3/Ч_3$ (e).

a	B_1B_1	B_2B_2	b	B_1B_1	B_2B_2	e	B_1B_1	B_2B_2
A_1A_1	13	9	A_1A_1	8,5	6,5	A_1A_1	1,53	1,39
A_2A_2	11	7	A_2A_2	9,5	7,5	A_2A_2	1,16	0,93

Итак, для расчетного признака влияния тех же двух локусов на фенотип уже не просто суммируются – возникает зависимость вклада одного локуса от другого, чего не было для исходных признаков. Здесь отсутствие или наличие эпистаза – это не свойство генов, но результат конкретной вычислительной связи трех признаков, изменчивость которых в данном иллюстративном примере обусловлена одним и тем же набором локусов – A , B .

Можно подобрать числовые примеры, когда схема наследования у признака M_3 – без эпистаза ($i_{M_3}=0$), у признака $Ч_3$ – с эпистазом ($i_{Ч_3} \neq 0$), но в изменчивости расчетного признака $M1$ эпистаза не будет ($i_{M1}=0$). Так, фертильность корзинки подсолнечника (F) определяют как результат деления M_3 (масса семян) на d (диаметр корзинки): $F=M_3/d$. Схема наследования для первого исходного признака M_3 без эпистаза: $m=31,5$; $a_1=16$; $a_2=-2$; $i=0$ (табл. 4.4а). Для второго исходного признака d с эпистазом: $m=7,53$; $a_1=2,10$; $a_2=-1,47$; $i=0,10$ (табл. 4.4б).

Таблица 4.4. Генотипические ценности для двух исходных признаков M_3 , d (a , b), зависящих от одного набора из двух полиморфных локусов, и расчетного признака $F=M_3/d$ (e).

a	B_1B_1	B_2B_2	b	B_1B_1	B_2B_2	e	B_1B_1	B_2B_2
A_1A_1	45,5	49,5	A_1A_1	8,27	11	A_1A_1	5,5	4,5
A_2A_2	13,5	17,5	A_2A_2	3,85	7	A_2A_2	3,5	2,5

После подстановки значений из табл. 4.3 a и b в формулу $F=M_3/d$ получаем значения вычисляемого признака у четырех гомозигот (табл. 4.4в). Далее подставляем эти значения в табл. 1.6 и определяем схему наследования признака F : $m=4$; $a_1=1$; $a_2=0,5$; $i=0$. То есть при наличии эпистаза у одного из исходных признаков, в схеме наследования вычисляемого признака F эпистаза нет.

Аналогичные эффекты возможны при различных математических связях исходных и расчетных признаков, хотя и не всегда. Например, если расчетный признак равен сумме и/или разности любого числа исходных, в изменчивости которых эпистаза нет, то у расчетного признака эпистаз не проявится.

Также не обязательно предполагать, что изменчивость расчетных признаков определяется общим набором генов. Рассмотрим противоположную ситуацию, когда изменчивость первого расчетного признака (I) определяется локусом A , второго (II) – только локусом B , не сцепленным с A . В табл. 4.5 для признака I заданы значения параметров $m_I=6$, $a_I=2$ и представлена генотипическая ценность для двух гомозигот A_1A_1 и A_2A_2 . Для исходного признака II, аналогично, по локусу B : $m_{II}=4$, $a_{II}=1$. Для расчетного признака III, определяемого двумя локусами, по табл. 1.6 вычисляем: $m_{III}=1,60$, $a_{1\ III}=0,53$, $a_{2\ III}=-0,40$, $i_{III}=-0,13$. То есть в его схеме наследования появился эпистаз.

Таблица 4.5. Генотипические ценности двух исходных (I, II) признаков, зависящих от двух разных локусов, и одного расчетного признака (III=I/II).

I		II		III	B_1B_1	B_2B_2
A_1A_1	$m_I+a_I=6+2=8$	B_1B_1	$m_{II}+a_{II}=4+1=5$	A_1A_1	8/5	8/3
A_2A_2	$m_I-a_I=6-2=4$	B_2B_2	$m_{II}-a_{II}=4-1=3$	A_2A_2	4/5	4/3

Физиологический эффект насыщения, о котором говорилось выше в разделе 1.3 (рис. 1.5), не обязательное условие появления эпистаза у расчетного признака y . Например, пусть $y=z/x$. Схема наследования двух исходных признаков (z , x) – аддитивно-доминантная, причем они оба зависят от единого набора генов. Известно, что плейотропный эффект может вызывать жестко скоррелированную изменчивость признаков в популяции, например, $z=ax+b$, где a , b – константы (Смирязев и др., 1992). Тогда $y=z/x=a+b/x$. Пусть выполнены два условия: 1) $b<0$, 2) $a>|b/x|$ при любом значении x . Тогда несложно показать, что график зависимости y от x будет иметь форму, аналогичную рис. 1.5. Этого, как отмечалось выше, достаточно для возникновения эпистаза в схеме наследования признака y при отсутствии эпистаза у исходных признаков (z , x). Причина – структура арифметической связи этих трех признаков.

Несколько расчетных признаков

Рассмотрим числовой пример. Исходный – количественный признак (например, урожайность сорта), значения которого получены в отдельные годы испытания набора сортов. Два расчетных – среднемноголетнее значение признака у каждого сорта и его дисперсия, характеризующая нестабильность проявления признака у сорта по условиям выращивания.

Для простоты примем, что известны четыре основных варианта стрессовых факторов и вероятности наступления их, а также их сочетаний в любой год следующие:

$q_1=0,2$ – вероятность холодного года,

$q_2=0,4$ – вероятность эпифитотии конкретной болезни,

$q_3=0,3$ – вероятность совместного проявления двух первых стрессов,

$q_4=0,1$ – вероятность засухи, причем первые два стресса в такие годы не проявляются.

$$q_1+q_2+q_3+q_4=1.$$

В наборе из 4-х гомозиготных сортов аллели A_1 обеспечивают устойчивость к первому стрессору, B_1 – ко второму. Ген устойчивости к засухе нет, но она присутствует с вероятностью 0,1. Следовательно, сорт $A_1A_1B_1B_1$ – устойчив к двум первым стрессовым факторам; $A_1A_1B_2B_2$ – устойчив к первому, но неустойчив ко второму; $A_2A_2B_1B_1$ – наоборот; $A_2A_2B_2B_2$ – неустойчив к обоим.

Заданные значения параметров m и a , а также вычисленные по табл. 1.6 значения исходного признака для четырех вариантов условий выращивания четырех сортов представлены в табл. 4.6. В первых и вторых условиях выращивания проявляется изменчивость признака только по одному локусу (A или B), в третьих условиях – по двум локусам без эпистаза ($i=0$), в четвертых условиях выращивания генетического полиморфизма нет.

Для каждого (j -го) сорта из четырех значения двух расчетных признаков – среднемноголетнего значения (c/m) и дисперсии (дис) представлены в двух последних строках табл. 4.6. Они вычислены по обычным формулам математической статистики:

$$c/m = \sum_{k=1}^4 q_k x_{kj}; \quad \text{дис} = \sum_{k=1}^4 q_k \left(x_{kj} - \sum_{k=1}^4 q_k x_{kj} \right)^2,$$

где x_{kj} – значение признака j -й гомозиготы в k -х условиях выращивания из табл. 4.6.

Сопоставляя генотипические значения каждого из двух расчетных признаков со структурой четырех гомозигот из табл. 1.6, вычисляем генетические параметры двух расчетных признаков. Для с/м: $m_{с/м}=11,4$, $a_{1с/м}=1,9$, $a_{2с/м}=2,2$, $i_{с/м}=0$; для дис: $m_{дис}=15,2$, $a_{1дис}=2,9$, $a_{2дис}=-1,9$, $i_{дис}=-4,9$. Их анализ позволяет сделать следующие выводы.

Таблица 4.6. Генотипическая структура и значения исходного признака в четырех условиях выращивания, проявляющихся с вероятностями q_k ($k=1, \dots, 4$). с/м – среднемноголетнее значение признака, дис – дисперсия признака по четырем условиям выращивания.

q_k	$A_1A_1B_1B_1$	$A_1A_1B_2B_2$	$A_2A_2B_1B_1$	$A_2A_2B_2B_2$
0,2	$m_1+a_{11}==16+5=21$	$m_1+a_{11}==16+5=21$	$m_1-a_{11}==16-5=11$	$m_1-a_{11}==16-5=11$
0,4	$m_2+a_{22}==12+4=16$	$m_2-a_{22}==12-4=8$	$m_2+a_{22}==12+4=16$	$m_2-a_{22}==12-4=8$
0,3	$m_3+a_{13}+a_{23}==8+3+2=13$	$m_3+a_{13}-a_{23}==8+3-2=9$	$m_3-a_{13}+a_{23}==8-3+2=7$	$m_3-a_{13}-a_{23}==8-3-2=3$
0,1	$m_4=10$	$m_4=10$	$m_4=10$	$m_4=10$
с/м	15,5	11,1	11,7	7,3
дис	11,3	25,0	15,3	9,2

Во-первых, схема наследования среднемноголетнего значения признака без эпистаза – также как у исходного признака. Во-вторых, в схеме наследования дисперсии признака появился эпистаз. В-третьих, аллель B_1 , который является плюс-аллелем среднемноголетнего значения признака, т.к. увеличивает его значение, стал минус-аллелем для дисперсии, т.е. одновременно снижает степень нестабильности признака. Аллель же A_1 повышает как среднемноголетнее значение признака, так и его нестабильность. И, что важно, из этого примера следует: выводы о параметрах и схемах наследования расчетных признаков зависят не только от эффектов генов исходных признаков, но также от q_1, q_2, q_3, q_4 , т.е. от частот проявления тех или иных условий выращивания генотипов.

Проводя аналогию с формированием любого количественного признака в онтогенезе можно сказать, что генетическая детерминация его величины также зависит от внешних лимитов, меняющихся в онтогенезе. Эти лимиты влияют на эффекты генов, меняя, в частности, схему наследования признака.

Наследственная изменчивость популяционных параметров.

Популяционные параметры, а также параметры семьи по сути являются расчетными признаками. Они оцениваются по исходным количественным признакам совокупности генотипов, входящих в популяцию, семью или набор. Подобные параметры имеют самостоятельное значение в селекционно-генетических исследованиях.

Примеры:

- эффект гетерозиса (по признаку двух родителей и их гибрида F_1),
- значения комбинационных способностей – ОКС и СКС (по результатам диаллельного скрещивания нескольких родителей),
- величины т.н. главных компонент,
- значения селекционных индексов (на основе дисперсий и корреляций исходных признаков всего набора),
- сами дисперсии признаков, от которых зависит эффективность отбора в популяции и т.п.

Рассмотрим иллюстративный числовой пример расчета эффекта гетерозиса (H) для двух исходных (I, II) и одного расчетного признака (III), который равен произведению исходных (табл. 4.7). Предполагается, что обе родительские формы – чистые линии и схема наследования исходных признаков без эпистаза и доминирования. Тогда значения этих признаков у гибрида F_1 равны среднему двух родителей.

Таблица 4.7. Числовой пример оценки $H=F_1-(P_1+P_2)/2$ – эффектов гетерозиса для двух исходных и одного расчетного признака (III=I×II).

Поколение	I. Число плодов на растении (шт)	II. Средняя масса плода (г)	III. Продуктивность растения (г)
P_1	4	100	400
F_1	7	75	525
P_2	10	50	500
H	0	0	75

Обычно считается, что если у составляющих (исходных) признаков гетерозис отсутствует, то и по расчетному признаку гетерозиса быть не может. Однако в примере, несмотря на чисто аддитивную схему наследования каждого исходного признака, у расчетного признака – продуктивности гибрида F_1 в табл. 4.7 наблюдается значительный гетерозисный эффект. Ясно, что причина

не в особых «гетерозисных свойствах» генов, определяющих изменчивость исходных признаков, а в изменении схемы наследования расчетного признака, где наверняка появятся эффекты доминирования и эпистаза. Подобный результат в экспериментах был, в частности, получен при изучении гетерозиса у томата (Powers, 1944).

Также бессмысленно говорить об особых свойствах генов в ситуации, отраженной на рис. 4.3. Здесь наблюдается гетерозис в диапазоне внешних условий D_1 – D_2 . Причина – одинаковое по 3-м генотипам, но нелинейное преобразование в онтогенезе на пути к признаку единственного активного гена с аддитивной схемой наследования.

Отметим, что не любой расчетный признак допускает подобную ситуацию возникновения гетерозиса. Если расчетный признак – это сумма или разность исходных, не проявляющих гетерозиса, то у расчетного признака гетерозис также не проявится.

Еще один результат, важный для подбора пар в селекции можно получить при анализе изменения схем наследования расчетных популяционных параметров. Как отмечалось в начале главы, часто используется упрощенное предположение: чем более несходны генетически родительские формы, тем шире в позднем поколении популяции полиморфизм потомства по генам, определяющим изменчивость количественных признаков, в том числе хозяйственно ценных. Вследствие этого в потомстве должно увеличиться доступное отбору генотипическое разнообразие этих признаков.

Рассмотрим иллюстративный числовой пример, включающий две ситуации и показывающий, что подобное предположение не всегда справедливо. В обеих ситуациях из-за плейотропного эффекта два исходных признака – масса зерна ($Mз$) и число зерен ($Чз$) на растении связаны регрессионной зависимостью: $Mз = -5 + 0,5 \times Чз$. Расчетный признак масса 1000 зерен $M1$, который с учетом регрессии можно оценить прямо по $Чз$: $M1 = 1000 \times Mз / Чз = 1000(-5 / Чз + 0,5)$.

Напомним, что в позднем поколении популяции самоопылителей при одном полиморфном локусе остаются две гомозиготы, при двух – четыре. Тогда среднеквадратическое отклонение признака в популяции, определяющее доступную для отбора изменчивость для него, определяется по обычной формуле:

$\sqrt{D} = [\sum_i p_i (x_i - \mu)^2]^{0,5}$. Здесь p_k – доля k -й гомозиготы в популяции, x_k – значения ее признака, μ – среднее значение признака в популяции.

В первой ситуации популяция получена в позднем поколении от

скрещивания $A_1A_1 \times A_2A_2$, т.е. изменчивость признака Чз в этой популяции определяется одним полиморфным локусом A и характеризуется следующими двумя генетическими параметрами: $m_{чз}=50$, $a_{чз}=10$. С учетом регрессионной связи $M1=1000(-5/Чз+0,5)$ для признака $M1$ вычисляем его значения: у гомозиготы $A_1A_1=417$, у $A_2A_2=375$. Генетические параметры этого признака: $m_{M1}=396$, $a_{M1}=21$. Доли двух генотипов в популяции потомства $p_1=p_2=1/2$. Тогда по приведенной формуле получаем среднеквадратическое отклонение признака $M1$ в этой популяции, состоящей из двух гомозигот: $\sqrt{D}_{M1}=21$.

Во второй ситуации популяция получена в позднем поколении от скрещивания $A_1A_1B_1B_1 \times A_2A_2B_2B_2$, т.е. родительские формы генетически более несходны по сравнению с первой ситуацией. Изменчивость исходного признака Чз характеризуется следующими генетическими параметрами: $m_{чз}=70$ (из-за влияния второго полиморфного локуса значение $m_{чз}$ изменилось по сравнению с первой популяцией), $a_{1чз}$ по-прежнему равно 10, $a_{2чз}=8$, $i_{чз}=0$. Для четырех гомозигот значения исходного признака Чз и рассчитанного по регрессии признака $M1=1000(-5/Чз+0,5)$ представлены в табл. 4.8. В схеме наследования признака $M1$ появился эпистаз ($i_{M1}=-2,5$).

Таблица 4.8. Генотипическая структура (см. табл. 1.6) и значения признаков Чз, $M1$ для четырех гомозигот.

генотипы	признаки	
	Чз	$M1=1000 \times (-5/Чз+0,5)$
$A_1A_1B_1B_1$	$m_{чз}+a_{1чз}+a_{2чз}=70+10+8=88$	$m_{M1}+a_{1M1}+a_{2M1}+i_{M1}==426+10,8+8,8-2,5=443,2$
$A_1A_1B_2B_2$	$m_{чз}+a_{1чз}-a_{2чз}=70+10-8=72$	$m_{M1}+a_{1M1}-a_{2M1}-i_{M1}==426+10,8-8,8+2,5=430,5$
$A_2A_2B_1B_1$	$m_{чз}-a_{1чз}+a_{2чз}=70-10+8=68$	$m_{M1}-a_{1M1}+a_{2M1}-i_{M1}==426-10,8+8,8+2,5=426,5$
$A_2A_2B_2B_2$	$m_{чз}-a_{1чз}-a_{2чз}=70-10-8=52$	$m_{M1}-a_{1M1}-a_{2M1}+i_{M1}==426-10,8-8,8-2,5=403,8$

При отсутствии сцепления локусов доли четырех генотипов в популяции – позднем поколении от самоопылений без отбора: $p_1=p_2=p_3=p_4=1/4$. Снова по приведенной формуле получаем среднеквадратическое отклонение признака $M1$ в этой популяции,

состоящей из четырех гомозигот: $\sqrt{D}_{M1}=14$. Напомним, что в первой популяции, состоящей только из двух гомозигот $\sqrt{D}_{M1}=21$.

Итак, несмотря на большее генетическое несходство родительских форм во второй популяции по сравнению с первой, доступное для отбора разнообразие в позднем поколении по признаку M1 не увеличилось, а уменьшилось. Причина в том, что на величину параметра \sqrt{D}_{M1} влияет не только разнообразие исходного признака Чз, но также его среднее значение $m_{чз}$ и возникший эффект эпистаза. В общем смысле причина – изменение схемы наследования параметра \sqrt{D}_{M1} . Более подробный математический анализ возможных причин снижения параметра \sqrt{D} расчетных признаков, а также экспериментальное подтверждение возможности такого снижения были рассмотрены в разделе 4.1.

Заключение

Упрощенное представление о положительных корреляционных связях генетического несходства родителей с доступной для отбора изменчивостью количественных признаков в потомстве не всегда соответствует действительности. На основе результатов полевых экспериментов с мягкой яровой пшеницей, а также с помощью биометрико-генетического анализа выявлены возможные причины и ситуации вариабельности и разнонаправленности этих корреляционных связей. Увеличение генетической дивергенции родителей, т.е. общего несходства их аллельного состава, не гарантирует увеличения полезного для отбора разнообразия в потомстве по всем изучаемым количественным признакам. Для части признаков возможно отсутствие реакции или даже закономерное снижение ответной изменчивости в потомстве.

Простые биометрико-генетические модели позволили продемонстрировать ограниченность традиционных представлений о варьировании генетического полиморфизма количественных признаков и популяционных параметров, механизмах формирования и перестройки схем их наследования. В частности, набор полиморфных генов расчетного признака не всегда включает все полиморфные гены исходных признаков. Схемы наследования количественных признаков или популяционных параметров, даже если они полиморфных по единому набору генов, могут существенно меняться в зависимости от вычислительной связи этих признаков. Схемы наследования часто не

свойство генов, а самих признаков.

Последнее относится не только к доминированию, плейотропии, но и к эпистазу. В рассмотренных ситуациях межлокусное взаимодействие не соответствует традиционному пониманию причин эпистаза, которое связано с молекулярно-генетическими или биохимическими процессами. Мы не отрицаем «традиционных механизмов» эпистаза, но продемонстрированные межлокусные взаимодействия или их отсутствие – это лишь эффект конкретных вычислительных связей признаков.

Подобные выводы имеют значение не только для генетики количественных признаков. Они указывают на возможные ошибки при подборе родительских пар на основе упрощенных критериев: в частности, при прогнозе эффективности отбора в расщепляющихся популяциях – потомствах, предсказании схем наследования комплексного количественного признака по схемам его «составляющих», прогнозе гетерозиса, а также при выведении урожайных и стабильных сортов.

Практически для любого количественного признака можно подобрать «исходные» признаки таким образом, что он станет «расчетным». Следовательно, все результаты, представленные выше, относятся не к каким-то особым, но к любым количественным признакам.

При изучении генотипа и фенотипа человек пользуется лишь ограниченными наборами признаков, которые он считает естественными, важными, удобными и пр. Но следует осознать, что любой живой объект (организм, популяция, биоценоз и т.п.) порождает непрерывное поле бесконечного числа взаимосвязанных признаков. Действительно, по мере углубления знаний о живой природе увеличивается число признаков, доступных для оценки. Любое обобщение, преобразование признаков – это новый признак, несущий определенную информацию об биологических объектах. С другой стороны, любой известный признак есть результат обобщения других признаков – его составляющих. Одна из фундаментальных проблем теории эволюции и генетики – переход от мозаики удобных для человека признаков к поиску признаков, важных для природы. На этой основе, в частности, можно рассчитывать на смягчение противоречий между естественным отбором и целями селекционного процесса.

Библиографический список

- Батыгин Н.Ф. Онтогенез высших растений. М.: Агропромиздат. – 1986. 101 с.
- Глазко Т.Т., Глазко В.С. Молекулярно-генетические подходы в селекции зерновых // Известия ТСХА. – 2006. – вып. 4. – С. 100-107
- Горшкова В.А. Генетическое изучение сортообразцов ярового ячменя в системе диаллельных скрещиваний // Селекция и семеноводство зерновых культур НИ-ИСХ Центр.-Чернозем. полосы. — Каменная Степь. – 1988. – С. 72— 81.
- Демиденко Е.З. Линейная и нелинейная регрессия. М.: Финансы и статистика. – 1981. – 302 с.
- Драгавцев В.А., Утемишева Н.В. К проблеме онтогенетической изменчивости генетических параметров в растительных популяциях // Генетика. – 1975. – Т. 11. – №11. – С. 128-140.
- Драгавцев В.А. Генетика количественных признаков растений в решении селекционных задач: Автореф. дис. ... д-ра биол. наук. – М. 1983—36 с.
- Драгавцев В.А., Аверьянова А.Ф. Механизмы взаимодействия генотип-среда и гомеостаз количественных признаков растений // Генетика. – 1983. – Т. 19. – №11. – С. 1806-1810.
- Драгавцев В.А., Аверьянова А.Ф. Переопределение генетических формул количественных признаков пшеницы в разных условиях среды. — Генетика. – 1983. – Т. 19. – № 11. – С. 1811-1817.
- Драгавцев В.А., Литун П.П, Шкель И.М. Модель эколого-генетического контроля количественных признаков растений. – Докл. АН СССР. – 1984. – т. 274. – № 3. – С. 720-723.
- Драгавцев В. А., Цильке Р. А. и др. Генетика признаков продуктивности яровой пшеницы в Западной Сибири. Новосибирск: Наука. – 1984. 230 с.
- Драгавцев В.А. К проблеме генетического анализа полигенных количественных признаков растений. ВИР. СПб. – 2003. 35 с.
- Драгавцев В.А. Инновационная наукоемкая технология генетического улучшения полигенных. – экономически важных свойств растений. Изд. Елтон-2. – 2009. – http://archive.nbuv.gov.ua/portal/chem_biol/nvlnau/Biol/2009_1/Dragavtse_v01-09.pdf
- Животовский Л.А. Популяционная биометрия. М.: Наука. – 1991. 270 с.
- Жученко А.А. Экологическая генетика культурных растений. Кишинев: Штиинца. – 1980. 587 с.

- Зотин А.А. Закономерности роста и энергетического обмена в онтогенезе моллюсков. Автореф. ... док. биол. наук. М.: 2009. – 30 с.
- Камшилов М.М. Является ли плейотропия свойством гена? // Биол. Журнал. – 1935. 4. – № 1. – С. 113-120.
- Корочкин Л. И. Биология индивидуального развития. М: МГУ. – 2002. 264 с.
- Крупин П.Ю., Дивашук М.Г, Фесенко И.А, Карлов Г.И. Адаптация микросателлитных SSR-маркеров пшеницы для анализа геномов пырея среднего. – пырея удлиненного и пшенично-пырейных гибридов // Известия ТСХА. 2011. Вып. 3. С. 49–57.
- Ллойд Э, Ледерман У. Справочник по прикладной статистике. Т. 1. М.: Финансы и статистика. – 1989. 450 с
- Лобашев М.Е. Генетика. Л.: Наука. – 1967. 750 с.
- Лопатина Л.М., Папазов Д.Ю., Драгавцев В.А. О генетических механизмах становления и изменчивости параметров гомеостаза сортов сельскохозяйственных растений // Генетика. – 1986. – Т. 22. – № 9. – С. 2295-2302.
- Мазер К, Джинкс Дж. Биометрическая генетика. М.: Мир. – 1985. – 463 с.
- Нгуен Т.Т, Смиряев А.В. Баженова С.С. Оценка эффективности методов подбора родительских пар на примере мягкой яровой пшеницы // Известия ТСХА. – 2010. – Вып. 6. – С. 145-152.
- Полевой В.В. Фитогормоны. Л: ЛГУ. – 1982. 230 с.
- Пыльнев В.В. Влияние длительной селекции пшеницы на полиморфизм гиадинов // Известия ТСХА. – 1994. – №4. – С. 57-68.
- Пыльнев В.В., Смиряев А. В., Пун Хуэй Жу. Диаллельный анализ ростовых признаков ярового ячменя в онтогенезе. – Известия ТСХА. – 1995. – Вып. 3. – С. 75-82.
- Рокицкий П.Ф. Введение в статистическую генетику. Минск: Вышэйшая школа. – 1974. 447 с.
- Савченко В. К. Генетический анализ и сетевых пробных скрещиваниях. Минск: Наука и техника. – 1984. – 220 с.
- Силис Д.Я., Шмакова Т.В. Влияние экологических факторов на комбинационную способность сортов озимой мягкой пшеницы // Вестник сельскохозяйственной науки. – 1986. – №12. – С. 57-62.
- Смиряев А.В. Экологическая и онтогенетическая изменчивость количественного признака в серии сравнительных полевых испытаний // Доклады ВАСХНИЛ. – 1985. – № 8. – С. 23-25.
- Смиряев А.В, Гохман М.В. Биометрические методы в селекции растений. – М.: Агропромиздат. – 1985. - 215 с.
- Смиряев А.В. Моделирование относительной изменчивости

признаков в селекционно-генетических исследованиях растений. // Автореф. дис.... док. биол. наук. Л.: – 1987. 37 с.

Смиряев А.В. Регрессионные модели гетерозиса озимой пшеницы // Доклады ВАСХНИЛ. – 1990. – № 4. – С. 9-12.

Смиряев А.В., Тарутина Л.А. Регрессионные модели изменчивости количественных признаков гибридов F1 кукурузы. – Цитология и генетика. – 1991. – №3. – С. 23-30.

Смиряев А.В, Мартынов С.П., Кильчевский А.В. Биометрия для генетики и селекции растений. Учебник. Издательство МСХА. – 1992. – 268 с.

Смиряев А.В. Анализ сходства генотипов по эколого-онтогенетической изменчивости ростового признака. // Генетика. – 1997. – т. 33. – № 1. – С. 61-677.

Смиряев А.В. Пыльнев В.В., Тао Юн-Шен. Сопоставление параметров относительной изменчивости ростовых признаков и генетического сходства сортов озимой пшеницы // Генетика. – 1997. – т. 33. – № 1. – С. 68-77.

Смиряев А.В., Мартынов С.П., Толстова О.В. Прогноз гетерозиса и сравнение гетерозиготности гибридов F1 самоопылителей с помощью евклидова расстояния // Известия ТСХА. – 1999. – вып. 3. – С. 51-57.

Смиряев А.В., Панкина Л.К. Основы биоинформатики. – М.: ФГОУ ВПО РГАУ-МСХА им. К.А. Тимирязева. – 2008. – 102 с.

Смиряев А.В., Панкина Л.К. Моделирование генных сетей. – М.: ФГОУ ВПО РГАУ-МСХА им. К.А. Тимирязева. – 2008а. – 51 с.

Смиряев А.В. Биометрико-генетический анализ несходства генотипов. Модель и параметры // Генетика. 2008. – т. 44. – № 2. – С. 269-275.

Смиряев А.В., Пыльнев В.В. Биометрико-генетический анализ несходства генотипов в диаллельном скрещивании яровой пшеницы // Генетика. 2008. – Т. 44. – № 2. – С. 276-282.

Смиряев А.В. Наследование несходства. Биометрико-генетический анализ // сб. Молекулярная и прикладная генетика. Минск: ИЦиГ. – 2009. – С. 56-61.

Смиряев А.В. Тест на аллелизм по генам количественного признака // Генетика. 2011. т. 47. – № 4. С.

Смиряев А.В. Генетическая дивергенция родителей и изменчивость количественных признаков потомства. Причины несоответствия // Материалы международного симпозиума «Интерактивная ампелография и селекция». Краснодар КубГау. – 2011. <http://www.vitis.ru/pdf/is30.pdf>

Смирязев А.В., Хуцацария Т.И, Баженова С.С, Нгуен Т.Т. Совершенствование биометрического метода оптимизации выборки растений при экспериментальном сравнении гибридных популяций // Известия ТСХА. 2012. – Вып. 2. С. 71–86.

Смирязев А.В., Дивашук М.Г, Хуцацария Т.И, Баженова С.С, Нгуен Т.Т. Прогноз эффективности отбора в популяциях потомства по косвенным оценкам генетической дивергенции родителей // Известия ТСХА. – 2013. – Вып. 1. – С. 57-70.

Смирязев А.В. Моделирование наследственной изменчивости расчетных признаков // Известия ТСХА. – 2013. – Вып. 5. – С. 77-91.

Тимофеев-Ресовский Н.В. О фенотипическом проявлении генотипа. 1. Геновариация *radius incompletus* у *Dr. Funebri* // Журн. экспер. биол. – 1925. – Сер. А. – Т. 1. – с. 93-142.

Турбин Н.В, Хотылева Л.В., Тарутина Л.А. Диаллельный анализ в селекции растений. Минск: Наука и техника. – 1974. 215 с.

Федин М.А., Силис Д.Я, Смирязев. Статистические методы генетического анализа. М.: Колос. – 1980. 279 с.

Федулова И.М., Анисков Н.И. Показатели уровня стабильности основных элементов структуры урожайности ярового ячменя // Семеноводство и селекция с.-х. культур в Зап. Сибири. Новосибирск. – 1990. – С. 107-115.

Хассан М.А.М. Эффективность способов отбора пшеницы с высокой экологической адаптивностью в ранних поколениях гибридов: Автореф. дис. ... канд. биол. наук. – Л. 1989. – 23 с.

Шаптуренко М.Н., Тарутина Л.А, Мишин Л.А, Кильчевский А.В, Хотылева Л.А. Дивергенция ДНК как критерий отбора исходного материала перца сладкого (*Capsicum annuum* L.) Генетика. – 2014. Т. 50. – № 2. С.138-146

Шевелуха В.С. Рост растений и его регуляция в онтогенезе. М.: Колос. – 1992. 593 с.

Шмальгаузен И.И. Факторы эволюции. Теория стабилизирующего отбора. М.: Наука. – 1968. 451 с.

Шмидт В.С. Математические методы в ботанике. Л.: ЛГУ. – 1984. – 218 с.

Ahmad R., Quiros C. Inter-and intra-cluster heterosis in spring type oilseed rape (*Brassica napus* L.) hybrids and prediction of heterosis using SRAP molecular markers // SABRAO Journal of Breeding and Genetics. – 2011. vol. 43 N 1 p. 27-43

Barbosa-Neto N.J.F., Sorrels M. E., Cisar G. Prediction of heterosis in wheat using coefficient of parentage and RFLP-based estimates of genetic relationship // Genome. – 1996. vol.39. p.1142–1149.

Benin G., Matei G, Costa de Oliveira A. et al. Relationships between four measures of genetic distance and breeding behavior in spring wheat // Genetics and molecular research. – 2012. vol. 11 N 3. p. 2390-2400

Bos I., Caligari P. Selection Methods in Plant Breeding. Springer. – 2008. 460 p.

Cavalier-Smith T. (ed). The Evolution of Genome Size. Wiley. – New York. 1985. 523 p

Cho Y.I, Park C.W, Kwon S.W. Key DNA markers for predicting heterosis in F1 hybrids of *japonica* rice // Breeding Science. – 2004. vol. 54. N 3. 389-397

Couviour Le F., Faure S., Dubreuil P., Praud S. Analysis of genetic structure in a panel of elite wheat varieties and relevance for association mapping // Theor. Appl. Genet. – 2011. vol. 123. P.715–727.

Diers B.W., McVetty P.B.E., Osborn T.C. Relationship between heterosis and genetic distance based on restriction fragment length polymorphism markers in oilseed rape (*Brassica napus* L.) // Crop Sci. – 1996. 36: 79-83

Dubreuil P., Dufour P., Krejci E. Organization of RFLP diversity among inbred lines of maize representing the most significant heterotic groups // Crop Sci. – 1996. N 36. p. 790-799

Eberhart S.A., Russell W.A. Stability parameter for comparing varieties // Crop Sci. 1966. vol. 6. N 1. P. 36-40.

Esposito M., Bermedjo C. et al. Prediction of heterotic crosses for yield in *Pisum sativum* L. // Scientia Horticulturæ. – 2014. vol. 177. N 2. p. 53-62

Fisher R.A. The genetical theory of natural selection: Dover Press. N.U. – 1958. 215 p.

Forkmann G., Seyffert W. Simulation of quantitative characters by genes with biochemically definable action. VI. Modification of a simple model // Genetics. – 1977. vol. 85. N. 3. P. 557-572.

Han Y, Bonos S.A., Clarke B.B, Meyer W.A. Inheritance of resistance to gray leaf spot disease in perennial ryegrass // Crop Science Society of America. – 2006. vol. 46. N.3. P. 1143-1148.

Hashemi A.S., Nematzaden., G.A, Babaelan J.N.A, Ghasemi C.O. Study of gene effects for quantitative in rapeseed via diallel analysis // Journal of agricultural sciences and natural resources. – 2008. V. 15. N. 4. P. 81-96.

Hayman B.I. The theory and analysis of variance of diallel crosses // Genetics. – 1958. vol. 43. P. 63-85.

Jones R.V. Analysis of variance of the half diallel table // Heredity. – 1965. vol. 20. №. 1. p. 117-121.

Kang M. S., Board J. E., Aminha N., Zhang Y., Moreno O.J., Balzarini M. G. Diallel analysis of ear moisture loss rate. – SPAD chlorophyll. and harvest index in maize with the leafy (Lfy) gene // *Journal of New Seeds.* – 2005. vol. 7. Issue 2. P. 1-16.

Kaplan J.M., Pigliucci M. Genes 'for' phenotypes: A Modern History View // *Biology and Philosophy.* – 2001. – vol. 16 p. 189–213.

Kazunori T., Keiji N., Hiroyuki T., Kazuyuki O., Tomohiko Y. Relationship between the coefficient of parentage and sugar yield in sugar beet F1 hybrid // *Genetics.* – 2006. vol. 8. N.4. P.151-159

Kovallski C. J. Longitudinal analysis // *Growth.* – 1974. vol. 38. N 1. P. 131-169.

Lee M. DNA markers and plant breeding programs // *Advances in Agronomy.* – 1995. vol. 55. P. 165-177.

Liu P., Jun Zhu P.L., Yan Lu H.L. A method for marker-assisted selection based on QTLs with epistatic effects // *Genetica.* – 2003. vol. 119. N 1. p. 75-86

Long F., Chen Y.Q., Cheverud C.J., Wu R. Genetic mapping of allometric scaling laws // *Genet. Res, Camb.* – 2006. vol. 87. p. 207–216

Nei M. Genetic distance between populations // *Amer. Natur.* – 1972. vol. 106. N 949. P. 283-292.

Powers L. An expansion of Jones's theory for the explanation of heterosis // *Am. Naturalist.* – 1944. vol. 78. N 776. p. 275-280

Quero-Garcia J., Letourmy P, Ivancic A, Feldmann P, Courtois B, Noyer J.L, Lebot V. Hybrid performance in taro (*Colocasia esculenta*) in relation to genetic dissimilarity of parents // *TAG.* 2009. vol. 119. – N.2. P. 213-221.

Rao C.R. Some problems involving linear hypothesis in multivariate analysis // *Biometrika.* – 1959. vol. 46. N 1 P. 49-58.

Rasmusson D.C., Philips R.L. Plant breeding progress and genetic diversity from de novo variation and elevated epistasis // *Crop Science.* – 1997. N 2. p. 303-310

Rasmusson J. A contribution to the theory of quantitative character inheritance // *Hereditas.* – 1933. vol. 18. P. 245-261.

Reif J.C., Melchinger A.E, Frisch M. Genetic and mathematical properties of similarity and dissimilarity coefficients applied in plant breeding and seed bank management // *Crop Sci.* – 2005. vol.45. P. 1-7.

Sarawat P., Stodart F. L, Marshall D. R. Derivation of superior F5 lines from heterotic hybrids in pea // *Euphytica.* – 1994. vol. 73 N 3. p. 265-272

Watkins J.C. Microsatellite Evolution: Markov Transition Functions for a Suite of Models // *Theoretical population biology.* – 2007. vol. 71. P.

147–159.

Welch B. L. The generalization of Student's problem when several different variances are involved // *Biometrika* . – 1974 vol. 34. № 1. P. 28-35.

Wu R., Lin M. Functional mapping — how to map and study the genetic architecture of dynamic complex traits // *Nature Reviews. Genetics*. – 2006. N 7. p. 229-337.

Yang W., Yu X., Yang W., Liu B. Parental epigenetic difference in DNA methylation-level may play contrasting roles for different agronomic traits related to yield heterosis in maize // *African Journal of Biotechnology*. – 2011. vol. 10. N 46. p. 9253-9263

Научное издание

Смирязев Анатолий Владимирович

**ГЕНЕТИЧЕСКАЯ ДИВЕРГЕНЦИЯ
РОДИТЕЛЬСКИХ ФОРМ И
НАСЛЕДСТВЕННАЯ ИЗМЕНЧИВОСТЬ
ПОТОМСТВА.
Биометрико-генетический анализ**

Монография

Издается в авторской редакции
Корректурa автора