

**RUSSIAN ACADEMY OF SCIENCES
MINISTRY OF SCIENCE AND HIGH EDUCATION OF RUSSIA
AGROPHYSICAL RESEARCH INSTITUTE
FEDERAL STATE BUDGET SCIENCES INSTITUTION
«FEDERAL WILLIAMS RESEARCH CENTER OF FORAGE PRODUCTION
AND AGROECOLOGY»**

Yu.V. Chesnokov, N.V. Kocherina, V.M. Kosolapov

**MOLECULAR MARKERS
IN POPULATION GENETICS
AND BREEDING
OF CULTIVATED PLANTS**

Moscow, 2019

УДК 575.1:631.523:575.116.4

ББК 28.04:28.54:41.3

Ч51

Chesnokov Yu.V., Kocherina N.V., Kosolapov V.M. Molecular markers in population genetics and breeding of cultivated plants : monograph. — Moscow : Ugreshskaya Printing House LLC, 2019. — 200 p.

Information on the use of molecular and traditional genetic markers in population genetics and breeding of agricultural plants is presented. An assessment of the level of diversity in populations is given using different types of markers, both at the level of alleles and at the level of individual nucleotides, including the measurement of differentiation within different marker systems. The relationship between molecular divergence, phenotypic divergence and co-origin described. The factors affecting the reliable determination of quantitative trait loci highlighted. Special attention is paid to the mapping of Mendelian genes. Special attention paid to the physiological and biochemical characteristics and their role in the manifestation and identification of economically valuable properties of cultivated plant species. Information is provided on the main hardware and software commonly used to conduct statistical calculations and establish a disequilibrium of marker/trait linking in populations. The purpose of the main computer programs, the most widely used for genetic mapping and establishing a marker / trait association, is briefly discussed.

The book intended for students and graduate students of biological and agronomic specialties of universities, curators of collections, breeders, as well as specialists in the field of molecular biology, mathematical statistics, genetics and plant physiology.

Il. – 3, tabl. – 10, literature – 423

Reviewers: Academician of RAS V.P. Yakushev
 Academician of RAS V.I. Dolzhenko
 Academician of RAS N.I. Kashevarov

ISBN 978-5-91850-036-1

DOI 10.33814/monography_1614

РОССИЙСКАЯ АКАДЕМИЯ НАУК
МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ НАУЧНОЕ УЧРЕЖДЕНИЕ
«АГРОФИЗИЧЕСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ»
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ НАУЧНОЕ УЧРЕЖДЕНИЕ
«ФЕДЕРАЛЬНЫЙ НАУЧНЫЙ ЦЕНТР КОРМОПРОИЗВОДСТВА
И АГРОЭКОЛОГИИ ИМЕНИ В. Р. ВИЛЬЯМСА»

Ю. В. Чесноков, Н. В. Кочерина, В. М. Косолапов

МОЛЕКУЛЯРНЫЕ МАРКЕРЫ
В ПОПУЛЯЦИОННОЙ ГЕНЕТИКЕ
И СЕЛЕКЦИИ КУЛЬТУРНЫХ
РАСТЕНИЙ

Москва, 2019

УДК 575.1:631.523:575.116.4

ББК 28.04:28.54:41.3

Ч51

Чесноков Ю. В., Кочерина Н. В., Косолапов В. М. Молекулярные маркеры в популяционной генетике и селекции культурных растений : монография. — Москва : ООО «Угрешская Типография», 2019. — 200 с.

Представлена информация об использовании молекулярных и традиционных генетических маркеров в популяционной генетике и селекции сельскохозяйственных растений. Дается оценка уровня разнообразия в популяциях с помощью различных типов маркеров, как на уровне аллелей, так и на уровне отдельных нуклеотидов, включая измерение дифференциации внутри различных маркерных систем. Описаны взаимоотношения между молекулярной дивергенцией, фенотипической дивергенцией и происхождением. Освещаются факторы, влияющие на достоверное определение локусов количественных признаков. Отдельное внимание уделено картированию качественных менделевских генов. Особое внимание уделяется физиологическим и биохимическим признакам и их роли в проявлении и идентификации хозяйственно ценных свойств возделываемых видов растений. Приводятся информация об основных технических средствах и программном обеспечении, повсеместно используемых для проведения статистических расчетов и установления неравновесного сцепления маркер/признак в популяциях. Кратко обсуждены назначения основных компьютерных программ, наиболее широко используемых для проведения генетического картирования и установления ассоциаций маркер/признак.

Книга предназначена для студентов и аспирантов биологических и агрономических специальностей вузов, кураторов коллекций, селекционеров, а также специалистов в области молекулярной биологии, математической статистики, генетики и физиологии растений.

Ил. — 3, табл. — 10, библиограф. — 423

Рецензенты: академик РАН В. П. Якушев
академик РАН В. И. Долженко
академик РАН Н. И. Кашеваров

ISBN 978-5-91850-036-1

DOI 10.33814/monography_1614

- © Федеральное государственное бюджетное научное учреждение «Агрофизический научно-исследовательский институт», 2018
- © Федеральное государственное бюджетное научное учреждение «Федеральный научный центр кормопроизводства и агроэкологии имени В. Р. Вильямса», 2018

ВВЕДЕНИЕ

В начале 70-х годов прошлого столетия наметился существенный прогресс в исследованиях, проводившихся в области популяционной генетики. Такое развитие стало возможным благодаря использованию физико-химического метода белкового электрофореза для анализа различного рода ферментов. Хотя к этому периоду времени теоретический аспект данного методического подхода был уже довольно хорошо развит, особенно школой S.Wright [405, 406], анализ ферментных систем изучаемых организмов, в отсутствие различных и многочисленных молекулярно-генетических маркеров, мог быть подтвержден только лишь прямыми экспериментальными данными. В доказательство этому, в последующие тридцать лет появились сотни работ по изучению аллоферментного и изоферментного полиморфизма у различных организмов. Так, например, в конце 1980-х годов было опубликовано несколько основных библиографических обзоров, посвященных рассмотрению полиморфизма аллоферментов как в царстве растений, так и в царстве животных [164, 278].

Подробные исследования на уровне аллоферментов проводились в различных дисциплинах (физиология, биохимия, генетика, селекция) для решения различных по своим целям задач (изучение структуры популяций и полиплоидии, при проведении гибридизации и анализе гибридов, в систематике и др.) [221, 270]. Аллоферментный анализ относительно прост и легок в исполнении. Как правило, для проведения аллоферментного анализа готовится гомогенат из какой-либо ткани, и полученный экстракт разделяется в полиакриламидном или крахмальном геле. При этом белки, находящиеся в этом экстракте, последовательно разделяются по зарядам и размерам. После электрофореза гель окрашивается в соответствии с ферментативной активностью выявляемого фермента. Для этого добавляют ферментный субстрат и краситель, которые, при вступлении в определенную биохимическую реакцию, окрашивают белковые полосы ферментов в том положении на геле, которое соответствует миграционному положению ферментного белка. В зависимости от числа локусов, их состояния (гомо- или гетерозиготности), и конфигурации ферментной молекулы, проявляется от одной до нескольких полос на геле. Позиции этих полос могут быть полиморф-

ными, а поэтому информативными при определении генных локусов и групп сцепления.

Для анализа и выявления генетического полиморфизма в различного рода популяциях могут использоваться и изоферменты. Так, например, в ряде работ на томате, изоферменты были использованы не только при анализе генетического разнообразия [38, 375], но и при локализации экономически важных генов [366], при мониторинге чистоты гибридных семян [371], при выявлении интрогрессии генов и хромосом диких видов [312], при проведении пыльцевой селекции [415], а также при скрининге гаплоидных генотипов, регенерированных в культуре клеток и тканей *in vitro* [414]. Однако существует несколько ограничений, не позволяющих широко использовать алло- и изоферменты в качестве молекулярных маркеров [297]. Для того чтобы молекулярный маркер стал таковым, особенно при составлении молекулярно-генетических карт, строго обязательно выполнение, как минимум, следующих условий: 1) для наиболее полного маркирования каждой хромосомы (группы сцепления), в геноме должно быть достаточное количество маркерных локусов, которые равномерно располагались бы друг от друга на расстоянии не более чем 15-20 сМ; 2) маркерный локус должен быть достаточно полиморфным с таким расчетом, что любое, проводимое скрещивание, выявляло бы расщепление по интересующему исследователя маркерному локусу [75]. Это как раз и есть те два основных, предъявляемых молекулярным маркерам, требования, которым изоферменты не могут отвечать. Более полный список необходимых условий приведен в [14]. Кроме того, при использовании изоферментов в качестве маркеров, чисто техническим ограничением является возможность (точнее невозможность) детекции и установления ферментной активности, вследствие отсутствия соответствующих красителей, используемых для выявления ферментативной активности. Так, например, в различных организмах идентифицировано более 3500 ферментов, а известно всего лишь около 100 красителей, используемых при анализе ферментативной активности, и менее половины из них пригодны для анализа изоферментов растений [387].

Следует подчеркнуть, что иногда термины «изоферменты» и «аллоферменты» взаимозаменяют, что является некорректным. Изофер-

менты — это ферменты, которые распознают и расщепляют одни и те же субстраты, но не обязательно являются продуктами одного и того же гена. Изоферменты могут быть активны в различные стадии жизни организма или в различных клеточных тканях и/или органеллах. Аллоферменты — это изоферменты, которые являются продуктами ортологичных генов, и которые, вследствие наличия аллельных отличий между кодирующими их генами, различаются в своем аминокислотном составе по одной или нескольким аминокислотам.

Однако как бы ни были хороши белковые маркеры, они имеют целый ряд ограничений. Прежде всего, следует отметить то, что белки не являются первичными носителями генетической информации, а являются продуктом транскрипции и трансляции, кодирующих их генов. Помимо этого, полиморфизм белков, выявляемый одномерным электрофорезом, может быть подвержен как качественному, так и количественному изменению из-за воздействия на растения экологического стресса, обусловленного, например, недостатком элементов питания [79, 336-338, 351, 362]. Необходимо учитывать и возможность механических повреждений (порывов) и/или несоблюдение стандартизации условий экстракции белков и полипептидов при их выделении и очистке, а также при проведении электрофореза, что может приводить к появлению неспецифических электрофоретических спектров [77, 87, 265, 338, 339]. Вырожденность генетического кода также является ограничительным фактором, поскольку значительно сужает спектр, а значит и возможности, полиморфизма, выявляемого на уровне аминокислотных последовательностей по сравнению с полиморфизмом, который можно выявить на уровне первичных нуклеотидных последовательностей. Кроме того, как известно, не весь геном, а всего лишь его треть или даже меньше кодирует и экспрессирует какие-либо белки. Все это приводит к тому, что геномные изменения в некодирующей его части или регуляторной частях генов (а это более 70% генома), выпадает из поля зрения исследователей. С появлением методологий позволяющих оперировать с ДНК (основным носителем генетической информации на молекулярном уровне) внимание исследователей в значительной степени переключилось на нуклеиновые кислоты, как источник информативного полиморфизма. Это связано с тем, что каждая отдельно взятая нук-

леотидная последовательность уникальна по своей структуре. Последовательность ДНК может быть использована для любых исследований генетического разнообразия и взаимодействия между организмами. В последние десятилетия был разработан широкий спектр идентификации полиморфизма нуклеотидных последовательностей. Результатом таких работ явилась разработка молекулярных методов исследования организмов с использованием ДНК-маркеров. Совокупность этих методов, получившая название ДНК-фингерпринтинг, довольно широко используется в современных исследованиях для решения различных задач в различных биологических дисциплинах, в том числе при изучении различного рода генетических и селекционных популяций.

В целом появление молекулярных маркеров на уровне ДНК значительно ускорило последующее развитие и популяционной генетики [300]. Следует заметить, что если аллоферменты и изоферменты, также как другие типы белков, принадлежат к относительно однородному типу маркеров — биохимических, то молекулярные маркеры и методы исследований на уровне ДНК позволяют выявлять значительно большее количество гетерогенности генома, определяемое его полиморфизмом и формой эволюции различных геномных районов. Кроме того, наличие полиморфизма может быть выявлено и в геномах разнообразных клеточных компартаментов (хлоропласты, митохондрии), имеющих свою собственную наследственность и эволюцию. Благодаря результатам, полученным с помощью молекулярных ДНК маркеров, такая традиционная область популяционной генетики, как анализ генетической структуры природных и селекционных популяций был значительно консолидирован и укреплен. Пригодность маркеров нуклеиновых кислот и в особенности доступность к базам данных ДНК, не только открыли новые области их применения, но и сделали возможным уточнение большого количества полученных ранее данных по истории происхождения и эволюции разнообразных популяций, в том числе образцов Генных Банков генетических ресурсов растений (ГРР). Наглядным примером одного из разделов биологической науки, которые, несомненно, выиграли от применения ДНК-маркерных технологий, и которые активно развиваются, в том числе исследователями ГРР, является филогенетика (см., например, [1, 2, 4, 9]). Кроме того, дополнительное сравнение ге-

неалогической информации с географическим распределением видов и популяций, то есть с их филогеографией, позволило с большей отдачей реализовать на практике пути интеграции популяционной генетики в построении естественной системы организмов и установлении путей их эволюции.

В настоящей книге рассмотрены возможности широкого применения молекулярных маркеров в популяционной генетике и селекции сельскохозяйственных растений, а также их диких родичей. В первой главе, не создавая классификационного перечня методов, основанных на различных маркерах или маркерных системах, и не детализируя их, как это было сделано ранее [14, 17], с помощью сравнительного подхода кратко обрисованы особенности и свойства их применения, относящиеся к трем основным направлениям популяционной генетики дикорастущих и возделываемых видов растений, а именно: изучение биоразнообразия, как природного, так и сохраняемого в Генных Банках ГРР, дифференциации и дрейфа генов.

Во второй главе описывается как можно с помощью молекулярных маркеров получить более полную картину генетического разнообразия, существующего внутри отдельно взятых видов или их структурных комплексов, популяций, и прежде всего селекционных. Взаимоотношения между молекулярной дивергенцией, фенотипической дивергенцией и происхождением; взаимосвязь между молекулярной дивергенцией и гетерозисом; маркерная помощь при беккроссировании полигенного признака и использование молекулярных маркеров для защиты новых сортов, а также ряд других тем также рассматривается в этой главе. Следует отметить, что такого рода информация особенно важна для системного управления генетическими ресурсами растений, также как и для рационального использования ГРР в селекционных программах, на что обращается особое внимание во второй главе. Разделы второй главы, посвященные вкладу изучения разнообразия в практическую селекцию, а также применению методов классификации в гетерозисные группы и управлению генетическими ресурсами растений, написаны при участии к.с.-х.н. А.М.Артемяевой, за что авторы выражают ей особую благодарность. Завершается вторая глава описанием использования молекулярных маркеров генов при интрогрессии желаемых аллелей ге-

ном и/или QTL (Quantitative Trait Loci) в реципиентный геном селекционно значимых форм и отдельно взятых сортов.

Третья глава посвящена вопросам проявления, идентификации и локализации на хромосомах локусов количественных признаков, проявляющих свое действие в популяциях. Показаны принципы установления различия эффектов действия «диких» аллелей, ответственных за изменчивость количественных признаков в популяциях, включая определение QTL, учитывающих молекулярно-генетические маркеры обособленно и посредством анализа вариантов популяции F_2 . Рассмотрены достоинства и недостатки популяций, наиболее часто используемых для определения QTL. Проведено сравнение популяций F_2 , беккроссов, удвоенных гаплоидов и рекомбинантных инбредных линий. Освещается генетическая и молекулярная основа вариации количественных признаков в различного рода популяциях, эффекты доминирования, трансгрессия, эпистаз, эффекты действия окружающей среды и некоторые характерные особенности QTL.

В четвертой главе, являющейся логическим продолжением третьей, рассмотрены случаи контроля признака одним основным или, так называемым, менделевским геном. В этой главе описаны два основных пути, позволяющие провести в ограниченном районе генома поиск маркера, генетически сцепленного с интересующего исследователя геном. Эти два подхода предполагают использование для поиска и картирования основных или менделевских генов почти изогенные линии и совокупный сегрегационный анализ популяций, что позволяет не только решить проблему картирования, но и высвободить значительное количество времени, а также существенно сократить расходы финансово-экономических средств, требующихся на установление взаимосвязи «маркер-ген».

В пятой главе описаны конкретные примеры практического картирования и идентификации локусов количественных признаков, обуславливающих проявление физиологических и биохимических признаков в популяциях, в том числе в условиях абиотического стресса. Показано, что при рассмотрении влияния различных видов стрессоров и разнообразных стрессовых условий, включая общие наблюдения по целому ряду признаков, становится ясно, что QTL распределены по геному не

случайным образом, а довольно часто образуют кластеры. Такое кластерирование проявляется не только для коррелирующих свойств на уровне организации, которая может быть тривиальной, но и на уровне свойств и признаков различных организационных уровней. Обращается внимание на то, что в условиях стресса могут проявляться новые кластеры QTL, обозначая новые гены или группы генов, которые становятся более важными в объяснении изменчивости признаков в популяциях при их выращивании в определенных эколого-географических условиях. Это означает, что при отборе на устойчивость к стрессу, наиболее благоприятный аллель не будет найден в том же локусе, как и при условиях отсутствия стресса. В материалах пятой главы подчеркивается, что если ген является «физиологически» ключевым, то его изменение в селекционных поколениях популяций может оказаться противоположным желаемому направлению отбору, причем таким образом, что никакие QTL невозможно будет определить. Подчеркивается, что физиологи, генетики и селекционеры уже сегодня могут начать изучение охарактеризованные сегрегирующие маркерные популяции и маркер-специфичные генетически изогенные линии вместо простых разновидностей и гетерогенных по своему составу образцов растений.

В заключительной шестой главе представлена информация об основных технических средствах и программном обеспечении, повсеместно используемых для проведения статистических расчетов и установления неравновесного сцепления маркер/признак, а также выявления ассоциаций между целевыми признаками/генами и молекулярными маркерами которые основываются на идентификации неравновесного сцепления. Кратко обсуждены предназначения основных компьютерных программ, наиболее широко используемых для проведения генетического картирования и установления ассоциаций маркер/признак. Приводятся интернет-ссылки, позволяющие заинтересованным пользователям воспользоваться необходимым для них программным обеспечением.

Авторы будут благодарны за любые замечания или пожелания, высказанные читателями, и постараются их учесть в последующих изданиях данной книги.

Ю.В. Чесноков, Н.В. Кочерина, В.М. Косолапов

Глава I. НЕКОТОРЫЕ АСПЕКТЫ ИСПОЛЬЗОВАНИЯ МОЛЕКУЛЯРНЫХ И ТРАДИЦИОННЫХ МАРКЕРОВ В ПОПУЛЯЦИОННОЙ ГЕНЕТИКЕ

Как известно, научное развитие в популяционной генетике в основном происходит за счет реализации двухступенчатого алгоритма достижения поставленной цели. При этом исследователь либо (1) ограничивается описательным анализом генетического разнообразия внутри какого-то числа *организменных генетических единиц* (например, виды, популяции), либо, (2) исходя из описательного анализа разнообразия, продолжает дальше прояснять и устанавливать эволюционные взаимосвязи (размер популяций, уровень миграции, поток генов и т.п.). Сдвиг от ставших традиционными (фенотипических, биохимических) маркеров в сторону молекулярных, позволяет значительно расширить научно-практический опыт исследователя и упростить решение этих двух задач. Достигается это благодаря следующим преимуществам: существенному увеличению числа аллелей, выявляемых на локус; увеличению числа анализируемых или выявляемых локусов; реализации доступа к полиморфизму цитоплазматических геномов, а также возможности секвенирования и установления распределения аллелей [14, 17].

1.1. Мультиаллелизм

Иногда ферментные локусы могут иметь много аллелей. Однако такие аллели, как правило, являются аллелями с низко сбалансированной частотой их проявления [164]. Напомним, что, хотя некоторые молекулярно-генетические подходы, где используется ПЦР амплификация с использованием случайных праймеров (например, RAPD, AFLP), выявляют только два аллеля (присутствие или отсутствие фрагмента), другие методы (RFLP, SSCP, DGGE, STS, SSR) позволяют выявлять достаточно высоко мультиаллельные локусы (см. [14, 15, 17]). В естественных популяциях для одного микросателлитного локуса обычно можно выявить от 10 до 30 аллелей [26, 111, 325]. Благодаря этому мультиаллелизм микросателлитов оказался преимущественно востребованным в исследованиях посвященных изучению условий репродуктивного процесса и дрейфа генов [55]. В то же время аллоферменты вполне подхо-

дят для установления различий у самоопыляющихся и перекрестно опыляющихся видов растений [54], или для приблизительного описания ареала распространения пыльцы [21, 119]. Но если требуется реконструировать происхождение или родословную интересующих образцов или особей с помощью сравнительно небольшого числа локусов, то такую возможность предоставляют микросателлиты. Например, в исследованиях на различных видах дуба было показано [111, 355], что достаточно от четырех до шести локусов для идентификации родительских форм, даже если исходная родительская популяция по своим размерам была очень мала. Подобного рода данные могут быть использованы как для установления физического расстояния, на котором исследуемые формы находятся друг от друга, так и для выявления максимально возможной дистанции распространения пыльцы и семян.

1.2. Увеличение числа локусов

Как указывалось выше, многочисленные исследования, проведенные в 70-80-х годах прошлого столетия как на животных, так и на растениях [164, 278] позволили установить, что с помощью изоферментных маркеров полиморфизм может быть проанализирован только на ограниченном числе локусов (то же самое можно отнести и к другим типам белков). В растениях число локусов хромосом или генов, выявляемых с помощью белковых генетических маркеров, редко превышает 30 [374]. Причиной тому, помимо генетических, могут служить и чисто технические ограничения. Как известно, из-за присутствия многочисленных вторичных метаболитов, выделение ферментов из растений является достаточно деликатным делом и, кроме того, проведение почти всех специфично окрашивающих реакций лимитировано отсутствием соответствующих красителей, необходимых для выявления ферментативной активности [387]. Не следует забывать и о влиянии физиологических факторов, поскольку активность некоторых ферментных локусов является тканеспецифичной. Теоретические исследования [277] позволили установить, что для выявления генетического разнообразия необходимость увеличения числа анализируемых в геноме локусов более важна, чем увеличение числа анализируемых особей. Молекулярные маркеры, в отличие от алло- и изоферментов, предоставляют возможность прове-

дения более систематического исследования генома. Для изучения разнообразия, основанного на биологических особенностях естественных или селекционно-значимых генетических ресурсов растений, необходимо, как можно более точно, установить генетический полиморфизм внутри исследуемой популяции. Наиболее эффективно это можно сделать при помощи маркеров, позволяющих выявлять и определять гетерогенность генома. Более того, если при этом у исследователя имеется возможность использовать значительное число маркеров, ассоциированных с исследуемыми признаками, и к тому же нанесенных на генетические карты, то это гарантирует лучшее покрытие генома изучаемого генотипа и, как следствие, позволяет более точно определять и использовать генетическое разнообразие изучаемых образцов [11, 13, 14].

1.3. Анализ полиморфизма цитоплазматической ДНК

Цитоплазматические геномы имеют особенности, которые делают их эффективным инструментом для проведения исследований в популяционной генетике. Наследование цитоплазматических геномов обычно монородительское. Митохондрии передаются преимущественно по материнской линии. Для хлоропластов ситуация варьирует. У покрытосеменных они у 70% родов наследуются по материнской линии, а у остальных 30% родов наследование происходит через цитоплазму обоих родителей [167]. У голосеменных и особенно у Pinaceae и Taxodiaceae форма наследования варьирует внутри семейств, но хлоропласты преимущественно передаются по отцовской линии [173]. Несмотря на то, что цитологические исследования уже подтвердили различия в способе передачи цитоплазматического генома по отцовской и материнской линиям [64, 78], экспериментальные данные о характере наследования хлоропластов были получены в основном на основе сегрегационного анализа маркеров хлДНК (хлоропластной ДНК), а для мтДНК (митохондриальной ДНК) — в специальных контролируемых скрещиваниях [173, 316].

Изучение полиморфизма цитоплазматической ДНК создало новое направление исследований в популяционной генетике [251]. Дифференциация популяций на ядерном и цитоплазматическом уровнях, прежде всего, может быть применена для установления степени распростране-

ния семян и пыльцы. Кроме того, монородительское наследование цитоплазматических органелл упрощает реконструкцию родословной (происхождения). Например, отцовское наследование хлоропластов и материнское наследование митохондрий позволяют проследить пространственное распространение семян и пыльцы у представителей рода *Pinaceae* исключительно посредством изучения полиморфизма ДНК органелл [110]. И, наконец, поскольку виды растений посредством распространения семян могут колонизировать новые для их произрастания территории, то выявление пространственного распределения полиморфизма в геномах органелл позволяет установить пути колонизации того или иного вида или популяции [293]. К тому же анализ полиморфизма хлоропластной ДНК значительно облегчается наличием «универсальных» маркеров, подходящих для исследования хлоропластного генома у разных, как дикорастущих, так и культурных, видов растений.

1.4. Распределение аллелей

Как указывалось выше, при анализе полиморфизма алло- и изоферментов, обычно, изучают различные аллели, взаимоотношения между которыми не всегда поддаются какой-либо регламентации или классификации. Так, сделанные в начале 1970-х годов [281] попытки интерпретировать электрофоретические варианты изоферментных профилей как аллели, полученные один от другого вследствие изменения электрофоретического заряда (модель поэтапного мутирования белков), оказались недейственными и неэффективными. Особенно когда модель сталкивалась с экспериментальными данными, для которых различия в заряде не было единственной причиной, ответственной за электрофоретическую подвижность [138]. В случае молекулярных маркеров взаимоотношение между аллелями может быть установлено как функция формы мутации. Для районов tandemных повторов (мини- и микросателлиты, рибосомальная ДНК) аллели могут отличаться один от другого за счет добавления или изъятия одной или нескольких единиц повторов [340]. Благодаря этому, они могут быть систематизированы в соответствии с их размерами (модель поэтапного мутирования нуклеиновых кислот). Тем не менее, даже принимая во внимание особую форму эволюции геномных повторов, необходимо отметить, что два аллеля одной и

той же длины не обязательно идентичны по происхождению. Для других районов генома, каждая мутация может (в разумных пределах) быть принята во внимание как уникальная (модель неограниченного числа ДНК-аллелей). Поэтому два аллеля с идентичной последовательностью, с высокой долей вероятности, будут идентичны и по происхождению. В этом случае филогенетическое древо аллелей может быть построено исходя из их нуклеотидных последовательностей. Если только не было рекомбинации (например, в хлДНК), можно построить филогенетическое древо исходя из молекулярно-генетических карт сцепления или рестрикционных профилей различных аллелей.

Информацию о способах и путях пространственного и эволюционного распространения аллелей можно использовать для упрощения установления генетического разнообразия ГРР, поскольку доступность информации о последовательностях ДНК различных аллелей позволяет «закрепить» генетику популяций на «ретроспективном» пути исследований посредством изучения генеалогии аллелей (теория сращивания, [181]). На основе информации о последовательностях ДНК различных аллелей, можно получить множество разнообразных данных. Так, например, может быть установлено время объединения (сращивания) аллельных множеств, а также число поколений, разделяющих происходящих от одного общего предка аллелей, включая позиционное распределение различных узловых точек на их генеалогическом древе. В целом при анализе полиморфизма видов живых организмов подход сращивания предлагает основанный на теории вероятностей метод, позволяющий делать выводы о демографических и генетических параметрах исходя из генеалогии генов [343].

1.5. Анализ молекулярного разнообразия

Обилие различных вариантов доступных молекулярных маркеров привело к разработке значительного числа методов анализа генетического разнообразия, существующего в естественных популяциях дикорастущих и селекционных популяциях возделываемых видов растений. Эти методы разделены на две большие группы. Первая ограничена так называемой *фенотипической* интерпретацией данных. В этом случае различия между генетическими единицами определяются индексами

(коэффициентами) диссимилиарности [230] и мультивариантными статистическими методами, которые позволяют описать общую организацию разнообразия. Такой подход используют, если доступны маркеры, полученные в результате проведения генетического фингерпринтинга. Природа маркеров имеет значение только при качественном сравнении различных исследований: если не установлено, что локусы должны определять характерные черты (придавать индивидуальный характер), то границы применимости используемых маркеров определяются случаем исследования. Во второй группе, данные анализируют в *генетических терминах*, используя при этом кодоминантные (например, RFLP, STS, SSR, SNP) или доминантные маркеры (RAPD, AFLP), а выявленные локусы рассматривают *индивидуально*. В целом, молекулярно-генетический анализ богаче фенотипического и лучше подходит для построения логических умозаключений по эволюции и структуре популяций, в том числе и селекционных. Исходя из анализа нуклеотидного полиморфизма, разнообразие, отмечаемое на молекулярном уровне, может быть выражено количественно, в частности, посредством вычисления среднего числа замен на нуклеотидный сайт. Это число может быть установлено исходя либо из анализа первичной структуры нуклеотидной последовательности, либо из данных, полученных с помощью молекулярных маркеров.

1.5.1. Установление молекулярной дивергенции на основе первичной структуры нуклеотидной последовательности

Рассмотрим два аллеля известной нуклеотидной последовательности, которые проявляют различия по сайтам 5 и 14:

сайт		5	14
аллель 1		АССТ G СТАТСТТА C GACGGTCGCGATGATA	
аллель 2		АССТ C СТАТСТТА G GACGGTCGCGATGATA	

Предположим, что оба аллеля эволюционировали отдельно друг от друга за период времени T (T также может быть число поколений). Каждая последовательность подвергалась мутациям, которые явились результатом замены одних нуклеотидов другими. Пусть K будет среднее число замен на нуклеотидный сайт, продуцируемых последователь-

ностью за время T , n — размер последовательности, а n_d — число нуклеотидных сайтов, показывающих (устанавливающих) различия между двумя последовательностями после прохождения времени T ($n_d = 2$ в случае аллелей 1 и 2, сравниваемых в качестве примера выше). Если частота нуклеотидной замены низка, то число замен, наблюдаемых на период времени T , будет ближе к реально продуцируемому, т.е. $K = n_d/n$ ($2/30$ в нашем примере). С другой стороны, если уровень замен высок, некоторые замены в тех же нуклеотидных сайтах могут произойти повторно. Например, даже если аллели 1 и 2 содержат одно и то же основание А в сайте 8, возможно, что за период времени T произошло несколько замен, но конечный результат оказался таким же, как будто никаких замен не происходило. Вследствие этого, среднее число замен, наблюдавшихся после прохождения времени T , будет меньше реально продуцируемых. В свое время различными группами ученых были предложены принимавшие во внимание рекуррентные замены методы установления K [191, 199, 364]. Эти методы дают сравнимые результаты при $K < 0,5$ и поэтому наиболее часто используют самый простой из них [191]:

$$K = \left(- \frac{3}{4} \right) \ln \left(1 - \frac{4 n_d}{3 n} \right) \quad (1)$$

При более низких значениях K ($< 0,10$) может быть использован прямой метод ($K = n_d/n$) [274]. K еще называют *нуклеотидной дивергенцией* между двумя последовательностями.

Установление значения K имеет большое эволюционное значение. Если λ есть уровень нуклеотидной замены на единицу времени для одного сайта (равная частоте мутации m , рассматриваемой в рамках нейтральной теории), то мы будем иметь:

$$K = 2\lambda T \quad (2)$$

Коэффициент 2 получен на основании того факта, что замены возникали в двух (одной или другой) последовательностях. Из данного линейного уравнения, устанавливающего отношения между числом аккумулярованных замен и временем, берет свое начало концепция «молекулярных часов». Кроме того, его используют для установления времени дивергенции между последовательностями (табл. 1).

Таблица 1

**Сравнение нуклеотидного полиморфизма гена,
кодирующего большую субъединицу рибулозо-1,5-дифосфаткарбоксилазы
(*rbcL*) у различных видов растений (по [48])**

Виды	Длина последовательности	Число замен	K1	K2	K3
<i>Zea mays</i> (O)	1422	232	0,163	0,184	0,186
<i>Triticum aestivum</i> (O)	1425	227	0,159	0,179	0,181
<i>Serenoa repens</i> (M)	985	122	0,124	0,135	0,137
<i>Magnolia macrophylla</i> (M)	1425	165	0,116	0,126	0,128
<i>Spinacia oleracea</i> (O)	1425	214	0,150	0,168	0,171
<i>Quercus rubra</i> (M)	1425	173	0,121	0,132	0,134
<i>Itea virginica</i> (M)	1407	166	0,118	0,128	0,129
<i>Petunia hybrida</i> (O)	1425	209	0,147	0,163	0,166

O — однолетнее растение; M — многолетнее растение; K1 — число замен на нуклеотидную последовательность ($K = n_d/n$, см. текст); K2 — число замен на сайт с коррекцией по [191]; K3 — число замен на сайт с коррекцией по [199]. Коррекции, которые принимают в расчет рекуррентные замены в одном и том же нуклеотидном сайте, подтверждаются фактом, что значения K1 относительно высоки (см. текст). Следует заметить, что оба метода коррекции дают очень схожие результаты. Нуклеотидный полиморфизм варьирует в соответствии с однолетним или многолетним периодом вегетации изученных видов. Как следует из приведенных данных, молекулярная эволюция гена, кодирующего *rbcL*, происходила намного быстрее у однолетних растений, чем у многолетних.

1.5.2. Установление нуклеотидной дивергенции методами маркерных технологий

Методы маркерных технологий, разработанные на основе использования рестриктных ферментов или ПЦР амплификаций, объединяет тот факт, что они выявляют полиморфизм сайта узнавания последовательности в зависимости от используемого подхода. Это — сайт рестрикции в случае RFLP, сайт гибридизации праймеров для RAPD или два сайта в случае AFLP. Посредством набора сайтов узнавания, молекулярные маркеры позволяют измерять или устанавливать нуклеотидное разнообразие по шкале генома. Метод установления базируется на предположении, что любое изменение последовательности, например, замена, в сайте узнавания (или между двумя сайтами узнавания) выражается в присутствии или отсутствии полосы на электрофоретическом профиле. Средняя пропорция полос общих для двух последовательно-

стей G , в принципе, является индикатором или показателем среднего числа нуклеотидных замен на сайт узнавания. Подробности методических подходов, использованных для установления нуклеотидного разнообразия по шкале генома, описаны ранее для RFLP [156, 275, 276], для RAPD [83] и для AFLP [184]. Здесь мы лишь попытаемся кратко суммировать принципы этих подходов. Детали математических выкладок и расчетов можно найти в оригинальных статьях [83, 156, 184, 275, 276].

Для примера рассмотрим сайт узнавания длиной r (где r число оснований). Если время дивергенции T между двумя последовательностями A и B не слишком велико, то число нуклеотидных замен в данном сайте соответствует распределению Пуассона с математическим ожиданием λT . Вероятность P того, что сайт узнавания остается неизменным в течение периода T , равна:

$$P = e^{-r\lambda T} \quad (3)$$

Используя уравнение $K = 2 \lambda T$ можно получить для K :

$$K = \left(-\frac{2}{r}\right) \ln P \quad (4)$$

Более того, показано [275], что простое отношение может быть установлено между P и G :

$$P \approx G(3 - 2P_0)^{1/4} \quad (5)$$

Как указывают те же авторы [275], P может быть найдено с помощью итеративного метода, принимая ($P_0 = G^{1/4}$) в качестве исходного значения. Значение G может быть установлено исходя непосредственно из электрофоретического профиля:

$$G = 2M_{AB}/(M_A + M_B), \quad (6)$$

где M_{AB} обозначает число полос общих для двух последовательностей A и B , а M_A и M_B обозначают число полос у последовательностей A и B , соответственно.

Предположим, что при прохождении эволюционного процесса полоса сохранилась, сайты узнавания должны при этом оставаться не-

изменными и за то же время не должен образоваться полиморфизм, вызванный инсерцией-делецией (инделом). Так, в случае RFLP анализа, сравнение прямого ($K = 2\lambda T$) и не прямого методов ($P = G(3 - 2P_0)^{1/4}$) на модельных последовательностях, показало, что второй применим для значений K менее 0,05 и что использование рестрикционных ферментов, распознающих в качестве сайта рестрикции четыре пары оснований (п.о.), оказывается более предпочтительным по сравнению с ферментами, распознающими в качестве сайта рестрикции шесть п.о. [156, 193].

Применение данного метода в отношении продуктов амплификации, особенно RAPD, представляется более сложным, даже если результаты моделирования являются обнадеживающими [83]. Действительно, гомология амплифицированных фрагментов не всегда постоянна и несомненна, а нуклеотидная замена (особенно на 5'-конце сайта) не обязательно препятствует амплификации. Поэтому в таких случаях установление нуклеотидной дивергенции может оказаться недооцененным. Указанный риск, несомненно, меньше в случае AFLP, поскольку дополнительные нуклеотиды находятся на 3'-конце, и благодаря этому правила несоответствия не позволяют провести амплификацию. Теоретические расчеты показывают, что методы по выявлению разнообразия на нуклеотидном уровне во всем геноме, основанные на электрофоретических профилях AFLP, являются надежными, если K меньше чем 0,10 [184]. Однако при этом количество вариантов установления разнообразия увеличивается с увеличением K .

Следует отметить, что почти все эти подходы разработаны для случаев гаплоидного генома. Для того чтобы распространить их на диплоидные геномы, как предложено ранее [83], должны быть известны аллельные частоты, что наиболее трудно в случае доминантных маркеров (этот случай будет рассмотрен ниже).

Основная привлекательность непрямого метода заключается в том, что он предлагает возможность сравнения генетических параметров, установленных с помощью различных маркеров на уровне всего генома. Как будет показано в следующих разделах, параметры разнообразия можно определить на нуклеотидном уровне, параллельно то же самое может быть сделано на уровне аллелей. «Снижение» этих пара-

метров до нуклеотидного уровня делает возможными сравнения между различными маркерами.

1.6. Полиморфизм внутри популяции

Установление дивергенции между двумя аллелями, упомянутыми в предыдущем параграфе, в данном разделе будет экстраполировано на набор аллелей, присутствующих в популяции. В случае генетических ресурсов растений это могут быть естественные популяции (дикие виды или примитивные формы), сохраняемые в Генных Банках ГРР. Внутри-популяционный полиморфизм *de facto* является балансом различных эволюционных сил: мутаций как источника полиморфизма, миграций, приводящих к его распространению, и генетического дрейфа, имеющего тенденцию к сокращению возникающего полиморфизма. Как можно видеть из материала, представленного в предыдущей главе (см. также [14, 16]), различные методические подходы не равным образом эффективны в установлении полиморфизма. Поэтому уровень наблюдаемого полиморфизма будет зависеть как от уровня мутаций и рекомбинаций [7] в рассматриваемом районе, так и от использованного маркерного подхода. Более подробно характерные параметры внутривнутрипопуляционного полиморфизма рассмотрены рядом авторов ранее [53, 206, 400] и ими же сделаны статистические выкладки, позволяющие оценить возникающий внутри популяции полиморфизм. Рассмотрим некоторые из этих параметров.

1.6.1. Выражение внутривнутрипопуляционного полиморфизма

Существующий полиморфизм, как правило, вызван присутствием множественных вариантов внутри какой-либо генетической единицы. Если генетической единицей являются генотип, популяция или вид, то полиморфными вариантами, исходя из методических подходов, используемых для их определения, могут быть гаплотипы или нуклеотиды. Таким образом, понятие полиморфизма соотносится с маркером и с генетической единицей. Используемые в экологии [296], а также в популяционной генетике [53, 274], показатели полиморфизма, называемые в научной литературе индексами или коэффициентами полиморфизма, условно могут быть разделены на две большие категории: те, которые

принимают во внимание или в расчет число вариантов внутри проанализированных генетических единиц и те, которые основываются на относительной частоте вариантов.

1.6.1.1. Показатели полиморфизма, принимающие во внимание (в расчет) число вариантов

Число вариантов очень чувствительно к размеру, т.е. к количественной величине образца ГРП [159], и не может быть использовано в тех случаях, где эта величина варьирует значительно, даже если использованные статистические подходы корректируют показатели полиморфизма (индексы, коэффициенты) до постоянной величины [183, 120]. Проблемы становятся более острыми, если использованные при этом маркеры характеризуются проявлением одного, но часто встречающегося варианта или многочисленными, но редкими вариантами (случай изоферментов). Более того, число (маркерных?) вариантов может быть неинформативным, особенно, если оно граничит с верхним пределом (2 в случае RAPD или AFLP).

1.6.1.2. Показатели полиморфизма, принимающие во внимание (в расчет) частоты вариантов

Наиболее часто используемый индекс полиморфизма получен на основе показателя, выведенного при проведении исследований в экологии [341]. Он был перенесен в популяционную генетику [243] и введен в употребление под названием показатель *генного разнообразия* (h) [272, 274]. Генное (оно же генетическое в узком смысле этого слова) разнообразие определяют как вероятность того, что два варианта, отобранные внутри генетической единицы случайным образом, будут отличаться друг от друга. Этот показатель очень популярен в популяционной генетике, особенно при работе с ГРП, поскольку он эквивалентен уровням гетерозиготности в популяции при панмиктическом равновесии и потому его значение существенно зависит от количества наиболее часто встречающихся аллелей, и тем самым ограничивает число образцов, необходимых для его установления или оценки.

Были предложены и другие показатели: например, эффективное число аллелей [88], информационная функция Shannon-Weaver [220] и

пропорция гетерозиготных локусов. Однако все они выводились, тем или иным способом, либо из числа вариантов, либо из меры генетического (генного?, что в данном контексте равноценно генетическому) разнообразия. Более того, информация, которую они предоставляли, как правило, ограничена описанием и сравнением полученных значений в различных генетических единицах. Исходя из этих показателей, можно сделать лишь ограниченное число выводов или умозаключений относительно эволюции изучаемых популяций. С другой стороны, были не забыты и также широко изучались флуктуации в числе вариантов и в генном разнообразии, которые возникали в течение эволюции популяций [274]. Для популяций эволюционировавших согласно модели Wright-Fisher (постоянный размер популяции, отсутствие миграций между ними, тип мутаций соответствует наличию неограниченного числа аллелей с не упорядоченными аллельными вариантами), уравнения значений числа аллелей (A) и генетического разнообразия (h) могут быть записаны следующим образом [168, 274]:

$$A = \sum_{k=1}^x \frac{4N\mu}{4N\mu + k - 1}, \quad (7)$$

$$h = 4N\mu / (4N\mu + 1), \quad (8)$$

где N — это эффективный размер каждой из популяций, μ — доля (процент, пропорция) мутаций, а x — размер образца.

Таким образом, в отсутствии миграций уровень генного разнообразия зависит от дрейфа и частоты мутаций. Поскольку дрейф действует в одном и том же направлении для всех маркеров, то различия, наблюдаемые между маркерами, тем существеннее, чем выше частота мутаций (табл. 2 и 3). Наконец, проведенная для этих двух показателей многоплановая компиляция данных аллоферментного анализа, позволила в свое время предложить их в качестве стандартов при измерении генетического разнообразия [164].

Таблица 2

Измерение уровня внутрипопуляционного разнообразия посредством использования различных типов маркеров

Виды	Изоферменты			RFLP			RAPD			Литературный источник
	<i>nl</i>	<i>A</i>	<i>h</i>	<i>nl</i>	<i>A</i>	<i>h</i>	<i>nl</i>	<i>A</i>	<i>h</i>	
<i>Brassica campestris</i>	3	2,7	0,518	9	8,5	0,684				[253]
<i>Brassica oleracea</i>	4		0,318				64		0,488	[213]
<i>Hordeum vulgare</i>	7	4,86	0,439	13	4,46	0,469				[421]
<i>Pinus sylvestris</i>	20		0,269				22		0,327	[361]
<i>Picea mariana</i>	13	2,2	0,300				28	1,9	0,321	[185]
<i>Populus gremididata</i>	14	1,4	0,080	37	1,8	0,13	56	1,9	0,310	[224]
<i>Populus tremuloides</i>	13	2,8	0,250	41	2,7	0,250	61	2	0,300	[224]
<i>Quercus petraea</i>	8	3,3	00,381				31	1,9	0,233	[215]
<i>Zea mays</i>	15	1,69	0,179	35	3,46	0,472				[113]

nl — число проанализированных локусов; *A* — среднее число аллелей на локус; *h* — генетическое разнообразие внутри популяции. Эти сравнительные исследования были проведены на одних и тех же образцах (популяции и особи) и одними и теми же методами расчета (уравнение 9). Различия, наблюдаемые для *A* и *h*, не могут считаться характерным свойством различных образцов. С другой стороны, внутригенная выборка (число локусов) высоко разбалансирована между маркерами, особенно из-за недостаточного числа изоферментных локусов. В целом, как правило, генетическое разнообразие (*h*) высоко, если его устанавливают с помощью молекулярных маркеров, а не изоферментов. Частота появления мутаций в прошлом, несомненно, выше, чем в ближайшем времени. Это следует из того факта, что некодирующие районы подверглись нескольким селективным ограничениям, в отличие от кодирующих районов. И если маркеры находятся в генах, то вырожденность генетического кода, и тот факт, что не все аминокислотные замены выявляемы, объясняет эти различия. Для изоляции аллелей это выглядит по-другому (сравнение RFLP-изоферменты), даже несмотря на то, что в целом число аллелей кажется выше для RFLP (тем не менее, сравнение затруднено, поскольку число выявленных аллелей в RFLP напрямую связано с числом использованных рестриктных ферментов).

Корреляционный коэффициент разнообразия, оцененного с помощью двух различных типов маркеров на одних и тех же популяциях

Виды	N_p	N	1 тип маркера	n_1	2 тип маркера	n_2	Коэффициент корреляции	Вероятность	Литературный источник
<i>Zea mays</i>	10	30	изоферменты	20	RFLP	35	0,466	0,125	[113]
<i>Brassica campestris</i>	20	?	изоферменты	5	RFLP	4	0,544	0,025	[253]
<i>Hordeum vulgare</i>	19	?	изоферменты	7	RFLP	18	0,562	0,050	[421]
<i>Brassica oleraceae</i>	5	20-100	изоферменты	4	SSR	7	0,900	0,089	[311]
<i>Beta vulgaris</i>	6	5	RFLP	4	SSR	1	0,928	0,044	[97]
<i>Picea mariana</i>	12	11-30	изоферменты	14	RAPD	22	0,400	0,464	[185]
<i>Brassica oleraceae</i>	18	38-60	изоферменты	4	RAPD	64	-0,100	0,764	[213]
<i>Hordeum vulgare</i>	12	20	изоферменты	6	RAPD	77	0,685	0,024	[283]
<i>Pinus pinaster</i>	23	30	SSR	3	AFLP	122	-0,271	0,341	[240]
<i>Quercus petraea</i>	7	50-200	SSR	6	AFLP	155	-0,649	0,103	[241]
<i>Quercus robur</i>	7	50-200	SSR	6	AFLP	155	0,180	0,691	[242]

N_p — число популяций; N — средний размер образца популяций; n_1 и n_2 — число изученных локусов для маркеров 1 и 2. Для доминантных маркеров (AFLP, RAPD) данные были проанализированы либо методом G1, либо G2. G1 соответствует вычислению разнообразия, установленного посредством вычисления аллельных частот; G2 — тот же анализ, но исключающий локусы для которых частота гомозигот для нуль-аллелей меньше чем $3/N$ (рекомендации [232]). F — индекс фиксации используемый для межаллельных частот для доминантных маркеров. Показатели генетического разнообразия очень слабо коррелируют, когда их устанавливают с помощью различных маркерных подходов: всего лишь в 4 из 36 случаев изучения были найдены достоверные корреляции (ошибка I типа: 0,05, т.е. статистическая значимость корреляций находится на уровне 0,05 или 5%). Низкие коэффициенты корреляций, скорее всего, являются результатом плохого подбора проб или образцов, т.е. проб или образцов, не отвечающих необходимым критериям проводимого анализа, особенно по числу локусов. Другие случаи могут относиться к низкой дифференциации между популяциями по степени их биологического разнообразия [241].

1.7. Оценка уровня разнообразия с помощью различных типов маркеров

1.7.1. На уровне аллелей

Наиболее общее математическое выражение генетического разнообразия h дикорастущих и возделываемых видов растений дано в 1973 году [272]. Если объектами изучения являются гаплотипы или аллели данного локуса с A количеством вариантов, то уравнение можно записать так:

$$h = 1 - \sum_{i=1}^A p_i^2, \quad (9)$$

или так:

$$h = \sum_{i=1}^A \sum_{\substack{j=1 \\ j \neq i}}^A p_i p_j,$$

где p_i (или p_j) частота аллеля i (или j).

В этом уравнении предполагается, что аллели расположены в произвольном порядке. Если же они могут быть упорядочены на основе какой-либо «внешней», например, фенологической информации, а их взаиморасположение выражено как функция дистанции Δ_{ij} между аллелями i и j , то уравнение разнообразия примет вид:

$$h = \sum_{i=1}^A \sum_{\substack{j=1 \\ j \neq i}}^A p_i p_j \Delta_{ij} \quad (10)$$

Важно заметить, что это с количественной точки зрения больше не является вероятностью, по которой два аллеля выбранные случайным образом из популяции, будут отличаться, но есть средняя дистанция (или, как ее еще называют, генетическое расстояние) между аллелями, присутствующими в популяции.

Для большинства маркеров практически невозможно получить информацию, которая смогла бы облегчить упорядочение аллелей. В этом случае Δ_{ij} равно 1 для всех пар рассматриваемых аллелей и уравнение (10) становится эквивалентным уравнению (9). С другой стороны, для маркеров, основанных на использовании последовательностей тандемных повторов (например, микросателлитов, рДНК), преобразование одного аллеля в другой часто происходит посредством добавления или удаления единицы повтора. Следовательно, число единиц повторов, разделяющих два подобного рода маркера, может влиять на точность

установления генетической дистанции между ними. Наконец, если доступны последовательности ДНК аллелей, то Δ_{ij} может быть получено исходя из их нуклеотидной дивергенции K (уравнения (1) и (4)). В компетенции исследователя сделать выбор, каким образом устанавливать дистанцию между аллелями при измерении разнообразия. Такое решение, как правило, принимается в зависимости от целей проводимых изысканий. В случае если необходимо сравнить различные исследования, и если принимать в расчет Δ_{ij} , то тогда должна быть задана или определена избранная для измерения консенсусная единица.

1.7.2. Особый случай доминантных маркеров

Для того чтобы измерить аллельное разнообразие для доминантных маркеров (RAPD, AFLP), должны быть известны частоты аллелей — 1 (присутствие фрагмента) и 0 (отсутствие фрагмента). При этом, если популяция (этой популяцией может быть образец ГРП) подчиняется уравнению Харди-Вайнберга, то частота q нуль-аллеля (отсутствие фрагмента) может быть определена исходя из частоты Q особей, в которых электрофоретические профили не демонстрируют или не проявляют этот фрагмент: $q = \sqrt{Q}$. В случае же, если популяция не подчиняется уравнению Харди-Вайнберга, частоту q [80, 215] устанавливают так:

$$q = \left(\frac{-1}{2(1-F)} \right) \left(F - \sqrt{F^2 + 4(1-F)Q} \right), \quad (11)$$

где F — индекс фиксации, который определяет расхождение (отклонение) при равновесии между теоретической гетерозиготностью, H_e и наблюдаемой гетерозиготностью, H_o ($F = 1 - H_o/H_e$) [168]. Значение F может быть также получено и из известных кодоминантных маркеров, выявленных у тех же самых изучаемых видов или образцов. Очевидно, что уравнение (11) можно использовать только, если индекс фиксации F является тем же самым, который вычислялся для кодоминантных маркеров, использованных при установлении F , и тем же для которого определялось q . Строго говоря, уравнение (11) может быть применено только для нейтральных маркеров и при сравнимых уровнях мутирования. Если нет информации по F , доступной для изучаемой популяции, или если есть сомнения в достоверности полученных от различных мар-

керов значений, то F всегда возможно установить исходя из значений Q , определенного для двух следующих друг за другом поколений одной и той же популяции [232].

Если установление частоты нуль-аллеля производят исходя из частот гомозигот, проявляющих нуль-аллели, то в этом случае может обнаруживаться существенная погрешность измерения, но только если число гомозиготных генотипов при этом будет меньше трех [232]. Авторы рекомендуют обратить внимание на то, что локусы, для которых частота гомозигот по нуль-аллелям меньше чем $3/N_0$ (где N_0 размер образца), не должны сохраняться или поддерживаться в образце популяции, а при сравнении, выполненном с помощью кодоминантных маркеров, сделали выводы о том, что размеры исследуемого образца должны увеличиваться от 2 до 10 раз. Эти рекомендации также применимы и к оценке разнообразия образцов ГРР. Так использование доминантных маркеров для определения генетического разнообразия требует наличия большего количества локусов в образце, чем в случае кодоминантных маркеров, и элиминации «наихудших» полиморфных локусов. Увеличение числа локусов, которые облегчают использование мультилокусных маркеров, таких как RAPD или AFLP, может компенсировать потерю точности по каждому локусу. Это было установлено в специальных исследованиях [241]. Так, например, показано, что, несмотря на плохую точность на уровне одного локуса, AFLP предпочтительнее для оценки генетического разнообразия на уровне целого генома, потому что случайное распределение учитывает огромную гетерогенность (неоднородность), существующую в геноме. Компиляция данных сравнительных исследований по установлению генетического разнообразия ГРР, проведенных с разнообразными типами маркеров показывает, что различные оценки выявляемого разнообразия достоверно коррелировали менее чем в 12% случаев (табл. 3). Модельные исследования [241] выдвигают в качестве возможной гипотезы предположение, что причиной отсутствия корреляций может быть: во-первых, небольшое количество использованных для исследования локусов, во-вторых, низкая дифференциация популяций по их уровню разнообразия и, в-третьих, неуравновешенность ситуации при проведении анализа.

1.7.3. Оценка на нуклеотидном уровне

Правила оценки разнообразия на нуклеотидном уровне разработаны на основе принципа, который был предложен и используется для уровня аллелей. Возьмем пример аллелей с известными последовательностями и постоянной (константной) длиной n и рассмотрим каждый нуклеотидный сайт в качестве локуса с четырьмя «аллелями» (по числу оснований: А, Т, G, С).

В сайте k разнообразие нуклеотидов может быть оценено следующим образом [276, 363]:

$$h_k = 1 - \sum_{i=1}^4 x_{ik}^2, \quad (12)$$

где x_{ik} частота нуклеотида i (i варьирует от 1 до 4 для оснований А, Т, G и С, соответственно) в сайте k . Среднее разнообразие нуклеотидов h на протяжении n сайтов, составляющих последовательность, будет равно:

$$h = \frac{1}{n} \sum_k^n h_k \quad (13)$$

Это выражение эквивалентно тому, в которое введены переменные частот различных гаплотипов p_i и p_j :

$$h = \sum_i \sum_{j \neq i}^n p_i p_j \pi_{ij}, \quad (14)$$

где π_{ij} дивергенция нуклеотидов между гаплотипами i и j (K из уравнения (1)). Авторы данного уравнения [276] показывают, что π_{ij} , среднее число нуклеотидных замен, может быть оценено посредством $\pi_{ij} = n_d/n$, если π_{ij} меньше чем 0,02 и посредством аппроксимации Jukes и Cantor [191] (уравнение (1)), когда π_{ij} больше, чем 0,02.

Сравнение уравнений (12) и (14) с уравнением (10) показывает, что *разнообразие на нуклеотидном уровне эквивалентно разнообразию на уровне аллелей, если аллели распределены как функция их нуклеотидной дивергенции.*

1.8. Установление различий между популяциями ГРР

Существующее разнообразие анализируют, как правило, на нескольких, а не на одной популяции. В связи с этим может быть выделено два иерархических уровня: разнообразие между популяциями ГРР,

обычно называемое *дифференциация*, и внутрипопуляционное разнообразие, которое мы рассмотрели в предыдущем разделе. Безусловно, обобщение возможно и на других иерархических уровнях. Степень дифференциации между популяциями очень сильно зависит от времени расхождения или разделения популяций ГРР и поддерживаемого между ними обмена генами. Таким образом, информация о степени дифференциации популяций говорит нам об истории изучаемых популяций.

1.8.1. *Общее выражение дифференциации*

Дифференциация может быть выражена двумя способами, которые в итоге дают один и тот же результат [274]. Первый состоит из рассмотрения дистанций d_{xy} между всеми парами популяций x и y :

$$d_{xy} = h_{xy} - \frac{h_x + h_y}{2}, \quad (15)$$

где h_x (или h_y) — разнообразие в популяции x (или y соответственно) (уравнения (9), (10), (13) или (14) согласно рассматриваемому уровню), а h_{xy} — внутрипопуляционное разнообразие (если аллели i и j в упомянутых выше уравнениях получены от двух различных популяций x и y). Из этого, (все)общая дифференциация (d_{st}) вычисляется как среднее всех значений d_{xy} по всем парам популяций (включая реципрокные пары и даже те, которые ассоциированы со своей собственной популяцией):

$$d_{st} = \frac{1}{s^2} \sum_x^s \sum_y^s d_{xy}, \quad (16)$$

где s — число рассмотренных популяций.

Второй подход состоит из подразделения генетического разнообразия на два иерархических уровня (отдельные популяции и совокупность популяций). Разнообразие совокупности нескольких популяций ГРР рассматривают как разнообразие единой популяции (называют общим разнообразием, h_t) и вычисляют, принимая частоты в уравнениях (9), (10), (13) или (14) за средние частоты всех популяций, составляющих данную совокупность. Среднее разнообразие отдельных популяций вычисляют как среднее внутрипопуляционных разнообразий h_s . Можно показать [274], что дифференциация между популяциями (уравнение (16)) выражается также как различие разнообразий, установленных на двух иерархических уровнях.

$$d_{st} = h_t - h_s \quad (17)$$

Затем дифференциацию G_{st} можно выразить относительно к общему разнообразию:

$$G_{st} = \frac{d_{st}}{h_t} \quad (18)$$

В упрощенной модели эволюции популяций (без отбора), согласно которой все популяции одного и того же размера обмениваются равным числом генов (островная модель Wright, [406]), ожидаемое значение вычисляют с помощью следующего уравнения, устанавливающего отношения между миграцией и дрейфом [274]:

$$G_{st} = \frac{1}{1 + \left(4N \left(\frac{s}{s-1}\right) (\mu + m)\right)} \quad (19)$$

где m — уровень миграции, s — количество субпопуляций эффективного размера N , μ — частота (уровень) мутаций. Если уровень мутаций много меньше, чем уровень миграции и если число собранных популяций велик, то это отношение принимает более упрощенную форму [168]:

$$G_{st} = \frac{1}{1 + 4Nm} \quad (20)$$

Параметр дифференциации G_{st} был определен в рамках модели с фиксированными эффектами, т.е. учитывалась только дифференциация внутри собранных популяций. Модель со случайными эффектами введена [400, 402] в виде параметра θ , эквивалентного F_{st} Wright [244]. Параметр θ , выраженный на уровне аллелей (θ_w на уровне локуса), является коэффициентом корреляций между генами различных особей одной популяции. Две модели отличаются друг от друга по лежащим в их основе гипотезам, определению параметров дифференциации и методам оценки. При этом одни исследователи [400, 402] разработали гипотезу, согласно которой все популяции получены из одной популяции-предшественника и что эти популяции эволюционировали без миграций, мутаций или отбора, другой автор [274], в противоположность этому, не строил никаких гипотез по истории популяций. Однако, в конечном счете, ряд авторов, как в случае гаплоидных данных [299], так и в случае диплоидных данных [298], распространили разложение по Nei [274] (уравнение (17)) на модель случайных эффектов, без особо силь-

ных ограничений генетической гипотезы модели Weir и Cockerham [400, 402]. Вследствие этого последняя модель примиряет оба подхода. Сравнения, сделанные на основе экспериментальных данных, показывают, что оценка двух параметров (G_{st} и F_{st}) сходится, если число популяций высоко [67, 299]. Как следует из литературных источников, параметры дифференциации на уровне аллелей и на уровне нуклеотидов инспирируют (наполняют своим содержанием) то одну, то другую, то третью, т.е. одну из трех упомянутых выше моделей, которые далее будут обозначены как NE для Nei, WE для Weir и Cockerham и PO для Pons и Petit.

1.8.2. Измерение дифференциации внутри маркерных систем

Уравнения (15) и (18) устанавливают общие рамки для выражения дифференциации, которая может быть применена ко всем маркерам как на уровне аллелей, так и на уровне нуклеотидов.

1.8.2.1. Уровень аллелей

Значения G_{st} , полученные вычислением h_t и d_{st} при использовании уравнений (9), (16) или (17), имеют различные меры в соответствии с избранной (для счисления) моделью. Так же, как и в случае определения разнообразия, дистанция между аллелями может быть введена (уравнение (10)) аналогичным образом, как это было сделано в рамках модели WE [125] и модели PO [299]. Случай микросателлитов является особым. Так, например, если в качестве рабочей принята модель последовательных мутаций, то дифференциация может быть вычислена принятием различий в числе единиц повторов в качестве дистанции между аллелями. Методы оценки, представленные в литературе и примененные к особому случаю микросателлитов, были выведены на основе модели NE (параметр R_{st} , [344]) или модели WE (параметр Φ_{st} , [264]).

1.8.2.2. Уровень нуклеотидов

Дифференциация на уровне нуклеотидов была получена заменой членов уравнения выраженных в h в уравнении (15) на математические элементы, приведенные в уравнениях (12) и (14). Такое же математическое выражение, соответствующее модели NE, было введено и Nei

[273], но в качестве элемента γ_{st} . Дифференциация на нуклеотидном уровне была оценена также и с помощью анализа вариантов (параметр β аналог параметру θ модели WE) [401]. Другие доступные методы инспирированы больше моделью PO. Исключая сравнение популяций самих с собой в уравнении (14), ряд авторов [182, 231] приблизились к модели случайных эффектов (N_{st} для параметра, введенного ранее, F_{st} для параметра введенного позднее). Эти два параметра отличаются друг от друга только в тех случаях, когда они принимают, либо не принимают в расчет поправку Jukes и Cantor (уравнение (1)). В этой связи была предложена оценочная функция [175] без погрешности измерения по Nei [273], которая делает γ_{st} сравнимым с F_{st} по Wright [404].

Точно так же, как и в случае разнообразия, *дифференциация на нуклеотидном уровне эквивалентна дифференциации на уровне аллелей, если аллели ранжированы (организованы и систематизированы) как функция их нуклеотидной дивергенции.*

Опубликованные экспериментальные результаты относятся главным образом к данным первичных нуклеотидных последовательностей [63, 249]. Тем не менее, метод оценки нуклеотидной дивергенции на основе анализа рестриционных фрагментов или случайных амплификаций (уравнение (4)) может быть использован для установления нуклеотидной дифференциации, но только в том случае, если избранная в качестве рабочей гипотеза имеет к этому отношение.

Теоретически дифференциация относительно независима от типа использованного маркера, если маркеры нейтральны. Ее значение в уравнении не находится под влиянием уровня мутаций, если этот уровень ничтожен по сравнению с уровнем миграции (уравнение (20)). Следовательно, дифференциация зависит преимущественно от размера исследуемой популяции и уровня миграции, которые являются постоянными факторами, действующими в одном направлении и с одной амплитудой на различные маркеры. Экспериментальные данные, позволяющие сравнивать значения дифференциации, рассчитанной на основе использования различных маркеров, как правило, подтверждают эти наблюдения. Однако для высокополиморфных маркеров (микросателлиты) можно ожидать более низкую дифференциацию, чем для других маркеров, поскольку для них уровень мутаций, как правило, очень вы-

сок и, безусловно, не может постоянно сбрасываться со счетов. Теоретические же исследования, проведенные с помощью моделирования, показывают, что G_{st} не зависит от числа аллелей анализируемого маркера, тем самым, подтвердив «универсальную» природу этого параметра [254].

1.9. Поток генов

Анализ потока генов с помощью таких генетических маркеров, как изоферменты, всегда оставался ограничен непрямыми оценками числа мигрантов. При этом выводы делались на основе дифференциации, выявленной у соседствующих друг с другом природных популяций. Такие оценки всегда были очень сырыми, грубыми и приближительными, поскольку они строго зависят от нереального эволюционного сценария (преимущественно основанного на модели островов). Кроме того, порой требовался достаточно длительный период времени для их накопления на протяжении нескольких следующих одно за другим поколений. Однако, благодаря разработке методов отдельной оценки дифференциации ядерной и цитоплазматической ДНК, стало возможным подразделение общего потока генов на индивидуальные уровни миграции пыльцы и семян, что существенно улучшило не прямые методы анализа тока генов. В дополнение к этому, появление высоко полиморфных кодоминантных маркеров позволило производить реконструкцию родословной на основе изучения двух последовательных поколений и, как результат, проследить передвижение пыльцы и/или семян в пространстве.

1.9.1. *Непрямое измерение потока генов*

1.9.1.1. Сравнительные уровни миграции семян и пыльцы

Монородительское наследование цитоплазматических генов приводит к сокращению эффективного размера популяций ГРР, как минимум, наполовину по сравнению с ядерными генами [41, 42]. Уменьшение, обусловленное слабым потоком генов между популяциями, выражается в большей чувствительности к генетическому дрейфу и, как следствие, индуцирует более высокую дифференциацию между популяциями по цитоплазматическим генам. Такие умозаключительные про-

гнозы были подтверждены практическими аналитическими разработками и выводами [43, 291]. Свои теоретические выкладки сравнительной эволюции дифференциации у строго аллогамного растения авторы вполне убедительно обосновали с помощью хлоропластных и ядерных маркеров. Из полученных результатов анализа можно сделать три важных вывода [291]:

- если передача хлоропластов идет преимущественно по материнской линии, то дифференциация ядерного маркера явно меньше, чем дифференциация хлоропластного маркера ($\alpha > 0,5$, где α — уровень передачи женским родителем);
- если передача хлоропластов происходит с равной вероятностью обоими родителями ($\alpha = 0,5$), то дифференциация хлоропластного маркера будет того же порядка увеличения, что и для ядерного маркера;
- дифференциация хлоропластного маркера может внезапно увеличиваться, причем настолько быстро, насколько минимальная передача хлоропластов осуществляется другим родителем ($\alpha \approx 0,95$ при противопоставлении $\alpha = 1$).

Указанные теоретические выкладки позволяют найти отношение уровней миграции семенами и пыльцой со значениями G_{st} (островная модель цитоплазматических и ядерных генов) маркера с материнским типом наследования и с маркером, проявляющим двуродительский тип наследования. Для этого было предложено следующее уравнение [121, 122]:

$$\frac{m_p}{m_g} = \frac{2(1/G_{st_m} - 1) - (1/G_{st_n} - 1)}{(1 - 1/G_{st_m})}, \quad (21)$$

где m_p и m_g соответственно обозначают уровни миграции посредством пыльцы и посредством семян, а G_{st_m} и G_{st_n} — значения дифференциации для маркеров с материнским и двуродительским типами наследования, соответственно. Таблица 4 иллюстрирует контраст между дифференциацией на ядерном уровне и дифференциацией генома, наследуемого по материнской линии, так же как асимметрию токов генов по этим двум векторам.

Таблица 4

Дифференциации по ядерным (Gst_n), хлоропластным (Gst_m) генам и отношение уровней миграции посредством пыльцы (m_p) и семян (m_g) у различных видов растений

Виды	Gst_n	Литературный источник	Gst_m	Литературный источник	m_p/m_g
<i>Argania spinosa</i>	0,25	[120]	0,60	[120]	2,5
<i>Eucalyptus nitens</i>	0,30	[266]	0,62	[63]	1,8
<i>Hordeum spontaneum</i>	0,48	[52]	0,75	[271]	1,3
<i>Quercus petraea</i>	0,024	[416]	0,90	[291, 292]	500
<i>Quercus robur</i>	0,032	[416]	0,92	[291, 292]	286
<i>Silene alba</i>	0,16	[252]	0,63	[252]	7

В соответствии с теоретическими выкладками, дифференциация по ядерным генам, передаваемым обоими родителями, всегда меньше, чем дифференциация по цитоплазматическим генам, передаваемым семенами. Эта разница объясняет различия в уровнях миграции по пыльцевому и семенному векторным направлениям. Асимметрия между этими уровнями, как представляется, еще раз подчеркивает разницу в биологии видов. Виды с тяжелыми семенами (например, дубы) распространяют свои семена на короткие расстояния (распространение преимущественно за счет сил гравитации, хотя незначительное число семян может быть перенесено на длительное расстояние птицами, такими, например, как сойки). Вследствие этого, у таких видов и существует большая асимметрия.

1.9.1.2. Происхождение колонизирующих генотипов

Анализ дифференциации в подмножествах популяций, использовавшихся при своем распространении, как это было показано выше, различные стратегии и динамики колонизации, может помочь в установлении происхождения колонизирующих генотипов. Эволюция дифференциации в теоретическом варианте была описана для групп популяций, обменивающихся друг с другом равным числом мигрантов (островная модель Wright, [406]). В этой модели при равновесии между миграцией и дрейфом, дифференциация того же порядка увеличения, что и для подмножества рассматриваемых популяций. Это явление еще более резко выражено на границе ареала распространения, т.е. там, где виды больше всего подвержены колонизации и поэтому такая ситуация должна приводить к более высоким уровням дифференциации ГРР, чем это может наблюдаться в центральной части ареала распространения. Это наблюдение было приведено [252] при аргументации аналитического выражения дифференциации между недавно образовавшимися популяциями (F_{st0} в выражении McCauley, [252]):

$$F_{st0} = \frac{1}{2k} + \Phi \left(1 - \frac{1}{2k} \right) F_{stc}, \quad (22)$$

где k обозначает число колонизаторов, формирующих новую популяцию, Φ — вероятность того, что два гена отобраны случайным образом из основателей популяции, произошедших из одной и той же популяции-источника, и F_{stc} — дифференциация между популяциями, установившимися ранее (исходные популяции). Такими исходными популяциями могут быть популяции диких родичей культурных растений, если речь идет об образцах возделываемых видов, сохраняемых в Генных Банках. С другой стороны, если Φ стремится к нулю, то дифференциация между вновь созданными популяциями образуется главным образом за счет числа основателей. При этом колонизация стремится к разрушению существовавшей раньше структуры, а основатели представляют собой особей, происходящих из разных популяций. В противоположность этому, если Φ стремится к 1, то основатели происходят из ограниченного числа популяций, а колонизация стремится к увеличению дифференциации. Таким образом, величина Φ позволяет получить важ-

ную информацию по динамике колонизации, а значит и о возможных уровнях сходства и дифференциации образцов Генных Банков ГРР.

Как показано ранее [252], k можно установить посредством демографических данных, F_{st_o} и F_{st_c} высчитывают исходя из показателей образца вновь образованных и показателей образца установившихся популяций, а Φ может быть выведено из уравнения (22). В примере, приведенном в работе [252], значения Φ были 0,73 для аллоферментов и 0,89 для ДНК хлоропластов, указывая на то, что колонизаторы распространились преимущественно из одной популяции-источника. Более того, незначительное различие между двумя значениями Φ наводит на мысль, что некоторые колонизаторы происходят от одних родителей, присутствовавших в популяции-источнике.

1.9.2. Прямое измерение потока генов

В противоположность к непрямому измерению потока генов, устанавливаемому на основе длительной дифференциации среди популяций ГРР, измерение дифференциаций, реализовавшихся за более короткий срок, не зависит от эволюционных моделей и базируется на изучении миграций, произошедших за последние два поколения. Благодаря стабильности методов молекулярного ДНК-анализа (например, микросателлитов), каждому генотипу данной природной популяции ГРР можно подобрать уникальный генетический отпечаток. Затем необходимо сравнить отпечатки у различных особей и, если отпечатки совместимы друг с другом, устанавливают генетические взаимоотношения. Вслед за этим обычно проводят анализ родословной. Делают это посредством установления взаимоотношений в назначенных взаимосвязях типа «родители-потомки» между членами подросткового (ювенильного) и взрослого континентов в исследуемой природной популяции ГРР. Такой популяцией может быть, например, популяция дикорастущих родичей культурных растений. Если маркеры высоко вариабельны, то задание может быть выполнено только после исключения всех несочетаемых или несовместимых пар генотипов. В большинстве случаев такая процедура приводит преимущественно к назначению в качестве родителя взрослого генотипа, который проявляет мультилокусное построение, сочетаемое с таковым у ювенильного растения. Если же не

могут быть применены простые методы исключения, например, когда сочетаемыми являются более чем один, т.е. несколько взрослых генотипов, то в таких случаях используют методы максимального правдоподобия и наиболее вероятных родителей отбирают в соответствии с оценкой и расчетом правдоподобия [255]. Для этого заранее необходимо в общих чертах наметить ареал (какую-то площадь или участок), в котором все взрослые растения генотипированы и с которыми совместимы генотипы ювенильной когорты. Если изучаемый ареал находится в непрерывной (сплошной) популяции, то может случиться так, что сочетания совсем не будет найдено и реальный родитель окажется за пределами ареала изучения. В дополнение к этому, если маркерная система не полиморфна в достаточной степени, вполне вероятно, что совместимый родительский генотип, идентифицированный внутри ареала изучения, окажется неправильно установленным родителем и что реальный родитель находится за пределами избранного для изучения ареала (скрытый ток генов). Доступные сейчас компьютерные программы, позволяют вычислять отношения правдоподобия для различных возможных родителей и устанавливать скрытый поток генов, а также недопущение вероятности неверного отбора [148, 244]. Это следует учитывать при привлечении в коллекции ГРР новых образцов культурных и особенно дикорастущих видов растений во избежание повторного сбора уже имеющихся в коллекции генотипов.

У растительных организмов анализ родословной может сталкиваться с ограничениями практического свойства. Так, в случае условий, когда известна численность всех взрослых генотипов растений, было установлено, что анализ родословной приводит к идентификации менее чем 50% мужских родителей в ареале изучения площадью от 5 до 10 га. Лабораторные условия, как правило, не позволяют производить оценку и апробацию более чем нескольких сотен взрослых генотипов и поэтому уменьшают размер изучаемого ареала. Следовательно, для видов, у которых наблюдаются большие дистанции распространения пыльцы, этот метод игнорирует более чем половину таких событий популяции [112, 355]. Во-вторых, если производят анализ родословной, например, устанавливают отцовство и материнство, то данный метод позволяет идентифицировать родительские пары, но не отдельное установление от-

цовского и материнского родителя. В этом случае могут оказаться полезны маркеры цитоплазматических геномов, но только если они полиморфны внутри ареала. Альтернативный метод заключается в отдельном изучении ДНК оболочки семян и зародыша, что тоже позволяет устанавливать родительский генотип, поскольку, как известно, что генотип оболочки идентичен диплоидному генотипу материнского растения [153].

В качестве анализа родословной зачастую требуется экстенсивное генотипирование. В последнее время стали разрабатывать альтернативные методы, которые позволят обойти генотипирование всех предполагаемых родителей и тем самым значительно сократить объемы работ. Так, например, предлагается сравнивать аллельные частоты в пулах пыльцы, опыляющих различные женские растения, собранные в популяции. Как установлено, в этом случае, мужская гаметная гетерогенность среди женских родителей, выраженная как коэффициент генетической дифференциации Φ_{ft} , должна быть обратно пропорциональна среднему распространению пыльцы и физическому расстоянию, разделяющему матерей [28, 348]. Если расстояние между женскими растениями достаточно велико (больше чем в 5 раз среднего расстояния распространения пыльцы), то Φ_{ft} становится независимым от дистанции, разделяющей материнские растения, и зависит только от расстояния распространения пыльцы. Данный подход не требует высокополиморфных маркеров, но если мужской или женский мультилокусные массивы (совокупности или множества) не могут быть однозначно, т.е. недвусмысленно, разделены на уровне потомков, то метод должен быть отклонен.

Заключение

Использование различных молекулярных маркеров для описания генетического полиморфизма, существующего в естественных популяциях ГРР, чаще всего ограничено лишь сравнительным анализом видов, что, как правило, не позволяет делать какие-либо обобщающие заключения по структуре изучаемого генетического разнообразия. В то же время, ранее с помощью ряда семейств маркеров были проведены мета-анализы ([164], для изоферментов; [280], для RAPD). Интересно в этих

случаях то, что если сравнение проводили с эволюцией развития признаков изучаемых видов, то указанные маркеры выявляли схожие тенденции вариаций между видами по разнообразию и дифференцировке. Однако тенденции проявления вариаций внутри видов могут различаться, если исследования разнообразия проводить с помощью различных маркеров, но на одном и том же наборе популяций. Как показано в таблице 3 уровни разнообразия ГРР, выявленные с помощью разных маркерных систем, в редких случаях существенно коррелируют. Низкие корреляции, наблюдаемые между различными типами маркеров, несомненно, подчеркивают недостаток в апробировании отобранных локусов и очень высокую гетерогенность полиморфизма внутри генома. Поэтому контроль (мониторинг) за генетическим разнообразием ГРР следует вести по возможно большему числу локусов. Согласно теоретическим выкладкам, если для анализа использованы различные типы маркеров, то в противоположность разнообразию, дифференциация среди популяций ГРР реже выражает конгруэнтные результаты. Применение сравнительного маркерного анализа полиморфизма генома ядра и цитоплазматических компартментов дало толчок более глубокому изучению тока генов у растений. Кроме того, теоретические прогнозы эволюции цитоплазматических геномов подтверждены экспериментальными результатами: ядерная дифференциация всегда меньше, чем цитоплазматическая дифференциация (табл. 4). Быстрое развитие ДНК- и, прежде всего, микросателлитных маркеров (для ядерных и даже хлоропластных геномов) сделало возможным выявление эффективных физических расстояний миграции пыльцы и семян, а также установление успеха членов популяции при репродукции.

Глава II. НЕКОТОРЫЕ АСПЕКТЫ ПРИМЕНЕНИЯ МОЛЕКУЛЯРНЫХ МАРКЕРОВ В СЕЛЕКЦИИ

Одной из важнейших областей практического применения молекулярных маркеров является селекция растений. Можно выделить два основных направления их использования в этой области растениеводства. Во-первых, с помощью маркеров можно получить более полную картину генетического разнообразия, существующего внутри отдельно взятых видов или их структурных комплексов, популяций (см. Главу I). Эта информация особенно важна для системного управления генетическими ресурсами растений, так же как и для рационального использования ГРР в селекционных программах [10, 11]. Наиболее перспективные направления применения подходящих для этих целей молекулярных маркеров приведены в первой части настоящей главы. Во-вторых, маркеры дают возможность построения насыщенных генетических карт, которые, в свою очередь, позволяют идентифицировать и локализовать хромосомные локусы, имеющие количественный или качественный эффект действия, а также устанавливать степень проявления этого эффекта [13, 15] (см. Главы III и IV). Такая информация сама по себе представляет прямой и непосредственный интерес для селекционеров. Самым простым примером тому может служить использование маркеров генов при интрогрессии желаемых аллелей в реципиентный геном или геномы. Рассмотрению этих вопросов посвящена вторая часть настоящей главы.

2.1. Вклад изучения разнообразия в практическую селекцию

Некоторые из методов, описанные в предыдущих наших работах [14, 17], могут быть использованы как для анализа генотипов, составляющих генетическое разнообразие, или компонентов коллекций ГРР (в последующем именуемые — образцы, популяции или инбредные (селекционные) линии), так и для их непосредственного применения в селекционных программах. Для начала рассмотрим взаимоотношения различного рода полиморфизмов, выявляемых с помощью маркеров, фенотипические различия и, так называемый, коэффициент сопроисхождения, т.е. коэффициент вероятности общего происхождения. Исходя

из характера этих взаимоотношений, как будет показано ниже, можно предложить ряд подходов их практического применения. Вслед за этим вопрос об использовании маркеров будет рассматриваться в русле решения двух задач, которые так или иначе связаны с созданием гибридов и гибридных сортов — это (1) классификация материала в предбридинговые гетерозисные группы и (2) предсказание гибридной (силы) величины. И, наконец, кратко будет рассмотрено использование молекулярно-генетических методических подходов для защиты сортов растений.

2.1.1. Основные методы анализа данных

Среди методов, используемых для изучения генетического разнообразия, определение генетической дистанции всегда являлось предметом особого интереса и внимания селекционеров растений. В этой связи, более четверти века тому назад, было показано, что если данные интерпретировать исходя из аллельных вариаций в данном локусе, то по аллельным частотам можно установить различные индексы дистанций [101]. В таком случае дистанция D между двумя наиболее часто используемыми в том или ином исследовании популяциями x и y будет равна [322]:

$$D_{xy} = \frac{1}{2L} \sum_{l=1}^L \sum_{a_l=1}^{n_l} (f_x^{a_l} - f_y^{a_l})^2, \quad (23)$$

где L — число локусов, использованных для вычисления дистанции, n_l — общее число аллелей в локусе l , а $f_x^{a_l}$ (или $f_y^{a_l}$) — частота аллеля a_l в локусе l в популяции x (или y соответственно). Это выражение эквивалентно уравнению (15) в предыдущей главе. В случае чистых линий D_{xy} просто равно пропорции локусов, для которых линии зафиксировали различные аллели. Когда информация не может быть проанализирована в пределах аллельных вариаций (например, когда используется кодировка «наличие-отсутствие» полос), то в таких случаях обычно используется дистанция по Nei и Li [275]. С другой стороны, если в значениях аллельных вариаций можно анализировать только какую-то часть информации, то в этом случае может оказаться полезным определение взвешенных средних дистанций по Rodgers и Nei и Li [105]. Точность установления дистанции увеличивается с числом принятых во внимание

локусов (табл. 5). Следует заметить, что доверительные интервалы не симметричны (даже если установленная дистанция равна 50%). Например, для 10 маркеров реальная дистанция (для всего генома) равная 30,8% может привести к нулевому значению установленной дистанции с вероятностью 2,5% $[(1 - 0,308)^{10}]$, причем таким образом, что доверительный интервал для нулевой дистанции будет находиться в интервале от 0% до 30,8%. Для данного числа маркеров равномерное их расположение на генетической карте, по сравнению со случайным распределением, в целом также повышает точность установления дистанций [106, 395]. С другой стороны, для увеличения точности установления генетических дистанций в расчет может быть принята дистанция на карте между генетическими маркерами. Для этого, как правило, используют соответствующие статистические подходы [106].

Таблица 5

Точность установления генетических дистанций между двумя линиями как функция числа (N) использованных для счисления дистанции локусов (по [29])

N	D				
	0%	10%	20%	30%	50%
10	0-30,8	0,2-44,5	2,5-55,6	6,6-65,2	18,8-81,2
20	0-16,8	1,2-31,7	5,7-43,6	11,8-54,3	27,2-72-8
30	0-11,5	2,1-26,5	7,7-38,5	14,7-49,4	31,3-68,7
50	0-7,1	3,3-21,8	10,0-33,7	17,8-44,5	35,6-64,4
100	0-3,6	4,9-17,6	12,7-29,2	21,2-40,0	39,9-60,1
200	0-1,8	6,2-15,0	14,7-26,2	23,7-36,9	42,9-57,1

Генетические дистанции могут служить основой для классификации интересующего исследователя, например, куратора коллекций или селекционера, материала. Таким материалом могут быть линии, образцы, популяции и т.п. На сегодняшний день используют различные алгоритмы определения генетических дистанций. Наиболее часто применяют такой алгоритм счисления среднего сцепления (обычно называемый UPGMA — невзвешенный среднеарифметический парно-групповой метод), в котором дистанция между двумя группами определяется как средняя дистанция между элементами групп, а также метод Ward [397], который минимизирует вариацию внутри формируемых групп. Эти ана-

лизы могут быть дополнены другими наглядными подходами установления степени генетической изменчивости, такими как анализ главных компонент или факториальный анализ соответствий.

2.1.2. Взаимоотношения между молекулярной дивергенцией, фенотипической дивергенцией и происхождением

Достоверная корреляция между дистанцией и происхождением обычно отмечается, когда в расчет берется достаточно большое число генетических маркеров (см. обзор [256]). Низкие или короткие дистанции могут наблюдаться только для высоких уровней происхождения (см., например, [346]). Тем не менее, для схожих степеней происхождения дивергенция между двумя особями, как правило, не одна и та же из-за генетического дрейфа и эффекта отбора. Например, в случае получения рекомбинантных инбредных линий посредством последовательного самоопыления потомства, полученного от скрещивания линий А и Б, некоторые линии получаются ближе к А чем другие, даже если все линии имеют один и тот же коэффициент происхождения с А. Это может быть проиллюстрировано результатами работы [37], в которой показано, что у кукурузы вклад данного родителя в полученные от него инбредные линии, подобным упомянутому выше способом, может варьировать между 80% и 20%. Если говорить об информации о происхождении, то при ее получении генетические маркеры имеют явное преимущество, поскольку принимают во внимание эффекты дрейфа и селекционного отбора и, вследствие этого, обеспечивают более аккуратную оценку генетической схожести.

Во многих исследованиях делались попытки установить точные взаимоотношения между индексами дистанций, выявленными с помощью генетических маркеров. Первоначально в таких работах использовали изоферментные маркеры [89]. Впоследствии стали применять молекулярные маркеры [29, 60, 61, 105, 269]. В целом, если результаты анализов изобразить графически, то на графике они, как правило, показывают треугольновидную форму распределения отношения взаимоотношений между двумя типами дистанций. Это означает, что если две особи близки на уровне молекулярных маркеров, то почти всегда они имеют схожий фенотип. С другой стороны, большие дистанции на

уровне маркеров не позволяют делать какие-либо умозаключения ни относительно сходства, ни относительно фенотипического расхождения исследуемых особей. Такие взаимоотношения могут быть объяснены полигенной основой признаков обычно принимаемых в расчет при оценке и установлении фенотипической дистанции. Если предположить наличие четырех биаллельных локусов с равными эффектами и, если благоприятные и неблагоприятные аллели обозначить + и -, соответственно, то линии, несущие комбинации аллелей ++ -- и -- ++ будут иметь один и тот же фенотип, даже если они отличаются по каждому локусу (детально данный анализ приведен в [60]).

Такие особенности и свойства взаимосвязи между индексами дистанций, установленными на фенотипическом уровне и индексами дистанций, определенными с помощью молекулярных маркеров, имеют два существенных практических результата или следствия. Во-первых, агрономические признаки особи могут быть предсказаны, если они на уровне генетических маркеров подобны таковым особи с известным фенотипом. Это особенно важно для морфологических признаков, которые требуют для своего определения длительного периода времени или их определение является очень дорогостоящим. Подобного рода прогнозирование может быть сделано эмпирически или формализовано посредством подходящих для этих целей статистических методов (см. раздел, посвященный прогнозированию гетерозиса). Второе следствие заключается в возможности распознавать, являются ли две особи или образца, близкие фенотипически, схожими на уровне генетических маркеров. Если предположить, что маркерные локусы дают представительный пример (образчик) всего генома, то особи с высоким уровнем симилярности на уровне маркеров, скорее всего, несут схожие аллели основных QTL или генов, в то время как особи с высокой дивергенцией одних и тех же молекулярных маркеров вероятнее всего содержат различные аллели. Это особенно важно для осуществления длительного сохранения и практического использования генетических ресурсов. Например, образцы, которые схожи как на фенотипическом уровне, так и на уровне маркеров могут рассматриваться как излишние (или резервные) и, вследствие этого, некоторые из них могут быть либо удалены из коллекции или объединены для упрощения и удешевления их сохранения. С другой

стороны, образцы, которые схожи на фенотипическом уровне, но отличаются на уровне маркеров, могут быть определены как оригинальные.

2.2. Применение методов классификации: гетерозисные группы и управление генетическими ресурсами

Первые экспериментальные исследования по созданию гибридных сортов кукурузы выявили существование корреляционных взаимоотношений между проявлением свойств гибридов и происхождением их родительских линий (исходных популяций или ареала распространения культуры). Во-первых, было обнаружено, что скрещивания между линиями, которые являются генетически родственными (по родословной) обычно приводят к получению менее продуктивных гибридов, чем те, которые получены от неродственных родительских линий [169, 207]. Более того, отдельные пары из некоторых популяций проявляют комплементарность при их использовании в гибридной селекции. Обычно, в этом случае, к получению хороших гибридов приводят линии, полученные из одной популяции, если их скрещивают с линиями не из той же, а из другой популяции [207]. Как правило, именно на основе таких эмпирических результатов селекционеры и определяют необходимые им пребридинговые гетерозисные группы, включающие в себя линии с сопоставимым эффектом скрещивания и часто с общим происхождением (например, традиционные сорта-популяции со свободным типом опыления). Подобным образом были созданы группы Lancaster и Reid Yellow Dent, широко используемые в Corn Belt (кукурузный пояс) в США.

Некоторые авторы изучали репрезентативные наборы линий пребридинговых гетерозисных групп посредством генетических маркеров, при этом статистический анализ полученных результатов проводился с помощью методов, описанных ранее: иерархическая классификация на основе установления молекулярных дистанций или факториальный анализ. Во всех исследованиях, проведенных на кукурузе, классификация групп, определенная на основе молекулярной информации, согласовывалась с традиционной классификацией пребридинговых гетерозисных групп [44, 114, 115, 152, 226, 259, 260]. Такие же результаты были получены и на подсолнечнике [147, 376]. А работа, выполненная на кукурузе

[114], проиллюстрировала дивергенцию между американскими и европейскими кремнистыми линиями этой культуры. Выявленные между группами различия отражали их расхождения на уровне аллельных частот по целой серии локусов. Результаты [114] приводят к двум основным заключениям важным для практической селекции. Во-первых, если имеются гетерозисные группы, хорошо идентифицированные на основе большого числа эмпирических исследований, то генетические маркеры могут быть использованы для установления филогенетического происхождения генетического материала ГРР (линии, образца, популяции) неясного генезиса и его таксономического, генетического и иного рода месторасположения по отношению к этим группам. Как правило, порядка 60 мультиаллельных локусов вполне достаточно для получения ответа на подобного рода вопрос, который периодически возникает у любого исследователя, имеющего дело с исходным селекционным материалом и не только [114]. Такое подразделение пребридинговых групп значительно облегчает выбор тестеров, которые должны быть скрещены с изучаемым материалом с целью установления его комбинационной способности. С другой стороны, если внутри интересующих исследователя разновидностей нет идентифицированных гетерозисных групп, так же как и в случае с вновь отобранными для работы генотипами или генотипами у которых гибрид был получен недавно, анализ с помощью генетических маркеров может помочь установить структуру изучаемой популяции в историко-селекционном аспекте ее формирования (см., например, результаты [30] по луку). В качестве второго шага, в этом случае, необходимо оценить какие пары групп являются наиболее подходящими для создания гибридов с высокими агрономическими или хозяйственно ценными качествами.

Даже если задача и не заключается в определении пребридинговых гетерозисных групп анализ структуры генетического разнообразия ГРР и определение совокупностей внутри видов или популяций может оказаться очень полезным для рационального управления и использования генетических ресурсов в селекции [103, 197, 258]. Так, на основе выработанной классификации и ее практического применения, посредством перекрестного опыления отдельных особей одной группы, может быть составлен так называемый «общий пул». Такие пулы представля-

ют интерес для реализации «пребридинговых» программ по увеличению частоты и расширению спектра генотипической изменчивости, что крайне важно при создании сортов [139]. Классификации также могут быть использованы как один из основополагающих принципов отбора при создании стержневых коллекций (core collections), являющихся репрезентативной выборкой генетического разнообразия всей коллекции [51]. Такую коллекцию можно создать, например, посредством отбора определенного числа образцов из каждой группы, образовавшихся в процессе классификации. Для создания стержневых коллекций были предложены и другие подходы, например, оптимизация отбора образцов на основе повторяющегося или, так называемого, итеративного процесса отбора [157, 389]. Как известно, такие стержневые коллекции полезны при управлении генетическими ресурсами растений. В частности, они упрощают обмены между программами и странами, и могут быть использованы для вычленения отдельных особей из больших по объему коллекций, в которых собраны образцы со специфичными признаками (например, с устойчивостью к заболеваниям).

2.2.1. Взаимосвязь между молекулярной дивергенцией и гетерозисом

С помощью теоретических расчетов можно показать [126], что степень гетерозиса гибрида, определяемая как различия между характеристиками данного гибрида и средними показателями его родителей, соотносится с дивергенцией родителей по QTL, проявляющих эффекты доминирования. В рамках диаллельной модели показатель гетерозиса H_{xy} , наблюдаемый при скрещивании двух популяций x и y может быть выражен следующим образом:

$$H_{xy} = \sum_{l=1}^L \sum_{a_l=1}^2 \frac{1}{2} d_l (f_x^{a_l} - f_y^{a_l})^2, \quad (24)$$

где d_l — эффект доминирования QTL l (другие обозначения точно такие же, как и в уравнении (23)). В гибридах F_1 гетерозис связан с гетерозиготностью QTL, которые проявляют эффект доминирования. Если локусы имеют равные эффекты доминирования, то величина гетерозиса будет прямо пропорциональна средней величине гетерозиготности. На основе этой теоретической выкладки, а также упомянутых выше экспериментальных результатов по оценке эффекта инбридинга, многие генети-

ки пытались предсказать гетерозис исходя из индексов различий между родителями. Результаты, полученные на основе морфологических критериев, в целом оказались разочаровывающими (см. [218]). Это частично может быть объяснено причинами, обозначенными ранее: линии, с одними и теми же значениями количественного признака, могут содержать различные аллели для каждого конкретного локуса [70]. В этом случае гетерозис, предсказанный для гибрида между двумя линиями данного типа, будет нулевым, в то время как реально наблюдаемый — максимальным.

Мы видим, что генетические маркеры так же, как и количественные признаки, предоставляют информацию о дивергенции генотипов по индивидуальным локусам. Поэтому среднее значение гетерозиготности гибрида может быть установлено исходя из анализа его родительских линий. На основе этого предположения было проведено несколько исследований, в которых авторы пытались предсказать гетерозис посредством соизмерения индексов дистанций, установленных с помощью генетических маркеров [217]. Анализ результатов, полученных на кукурузе и других видах растений [100], приводит к нескольким общим заключениям:

- в целом, прогнозы, сделанные с помощью многочисленных молекулярных, например, RFLP, маркеров кажутся более эффективными, чем предсказания, сделанные на основе ограниченного числа ферментных (белковых) маркеров (см. в особенности исследования [346]), что может быть объяснено лучшим покрытием генома и более высоким числом аллелей молекулярных (RFLP) маркеров на локус;
- когда наблюдаются существенные корреляции коэффициентов, индексы дивергенций могут проявлять бóльшую взаимосвязь со специфичной комбинационной способностью [352], также называемой специфичным гетерозисом [145], чем с непосредственными характеристиками гибридов.

Такие выводы можно объяснить с помощью моделей количественной генетики. Прежде всего, следует отметить, что оценка гибридов частично зависит от QTL, имеющих однозначно аддитивный эффект, и для которых гетерозиготность не предоставляет какое-либо преимущество, в то время как специфичный гетерозис не зависит от этих локусов.

Это заключение представляет интерес только в тех пределах, в которых специфичная комбинативная способность существует как компонент гибридного значения, что является наиболее трудным для предсказания на основе «классических» агрономических композиционных построений. Общая комбинационная способность данной инбредной линии может быть более аккуратно установлена на основе средних показателей характеристик потомков гибрида, полученных с помощью подходящих для этих целей тестеров. При этом результаты исследований, как правило, напрямую зависят от изучаемого материала. Однако следует отметить, что максимальная корреляция значений, прежде всего, была получена в исследованиях, включавших родственные гибриды (скрещивания между родственными линиями). Так, например, исследования [345] включали гибриды, имеющие высокий коэффициент родственности (около 0,9) и, вследствие этого, вели к высокой корреляции. С другой стороны исследования [152] не включали родственные гибриды и, как результат, не приводили к выявлению существенных достоверных взаимоотношений. Корреляция уменьшалась, когда родственные гибриды убирали из набора исследуемых гибридов [22, 61, 70]. Эти результаты подтверждают общую тенденцию, установленную на уровне изоферментов у кукурузы [133] и впоследствии на горчице [334]. И, наконец, в исследованиях, выполненных на гетерозисных «межгрупповых» гибридах (см. выше), не выявлено никакого взаимоотношения между характеристиками гибрида и дивергенцией по ним родителей [44, 257]. Аналогичная информация была получена и на межсубвидовом уровне у риса [326, 407].

Как видно из представленных выше работ, полученные в них данные довольно многоплановые. И, тем не менее, несмотря на их многоплановость, такие разнообразные результаты можно проанализировать теоретически. Для этого, прежде всего, следует предположить, что все используемые генетические маркеры являются фенотипически нейтральными. Кроме того, для того, чтобы значение было предсказуемым, между маркерными аллелями и аллелями QTL должно существовать неравновесное сцепление (linkage disequilibrium, LD) [76]. Неравновесие особенно высоко внутри вновь образованных популяций, полученных из ограниченного числа родителей. Это может быть проиллюстрировано

результатами работы [360], в которой получили картирующую популяцию F_3 при проведении скрещивания линий кукурузы В73 и Мо17. Каждое изученное растение F_3 было скрещено с каждой из родительских линий, и затем была оценена средняя продуктивность соответствующего гибридного потомства. Установленная продуктивность высоко коррелировала (коэффициент корреляции был равен 0,68) со средней гетерозиготностью в маркерных локусах. Как известно, внутри генетической группы неравновесное сцепление может существовать, если группа создана из ограниченного числа родителей недавно или подверглась достоверному влиянию процесса генетического дрейфа. У аллогамных видов, предполагающих панмиктический тип репродукции, неравновесное сцепление будет ослабевать с каждым поколением, приводя к уменьшению корреляций во взаимоотношении между генетической дистанцией и гетерозисом [76].

В свете вышеизложенного особенно интересен случай гибридов между линиями, происходящими от различных генетических групп, поскольку, во-первых, для отношения существующего между дивергенцией и характеристиками этих гибридов должно быть не только неравновесное сцепление между маркерами и QTL внутри каждой рассматриваемой группой, но и ассоциации, выявляемые между маркерными аллелями и аллелями QTL, должны быть одинаковыми для различных гибридных групп [73]. Вторым условием, в большинстве случаев не всегда выполнимым, является наличие линий представителей различных групп, которые, как правило, получены от различных популяций и которые эволюционировали независимо друг от друга на протяжении очень длительного периода времени. И, наконец, когда внутри- и межгрупповые гибриды рассматривают одновременно, то различия между генетическими группами могут объяснить существование неравновесного сцепления (специфичные аллели групп сцеплены между собой), даже если неравновесие равно нулю внутри каждой из групп. Таким образом, существующая возможность различения и, соответственно, разделения межгрупповых и внутригрупповых гибридов, позволяет увеличить предсказуемость получения гибридов с желаемыми характеристиками, как это проиллюстрировано результатами работы, выполненной на бобах [279].

На степень проявления взаимосвязи между гетерозисом и молекулярной дивергенцией, особенно эпистазом и относительно важными эффектами доминирования у QTL, могут действовать и другие генетические факторы. Однако результаты совмещения экспериментальных и теоретических данных, как оказалось, могут совпадать с общепринятыми представлениями о взаимоотношениях между гетерозисом и молекулярной дивергенцией, а также с условиями, в которых ожидаемое предсказание окажется эффективным [256].

Наконец, важность гибридов между неродственными инбредными линиями приводит к поиску других подходов предсказания гетерозиса. Эти подходы основаны на идеях, развитых ранее и заключающихся в том, что если два гибрида $A \times B$ и $A' \times B'$ имеют родителей близкородственных на уровне генетических маркеров (A' близок к A и B' близок к B), то их гетерозис или специфичные комбинативные способности будут схожими. Для формализации этого подхода был разработан метод [36], основанный на BLUP (best linear unbiased prediction, [172]). Дистанции между линиями применяли здесь для установления коэффициентов сопроисхождения [35], которые затем использовали для установления ковариаций между особями, чьи характеристики известны, с одной стороны, и особями, чьи характеристики необходимо предсказать, с другой стороны. Затем, на основе полученных ковариаций и известных характеристик, посредством использования соответствующей формулы в рамках рассматриваемого подхода, были сделаны необходимые математические расчеты и на их основе ожидаемые прогнозы. Экспериментальные результаты демонстрируют, что этот подход может быть эффективным в ситуациях, где метод установления генетических дистанций невозможно применить [71]. Следует заметить, что использование BLUP позволяет увеличить точность в определении оценки общей комбинационной способности инбредных родительских линий. Это также может быть использовано на любом этапе селекционной программы, в частности для того, чтобы лучше оценить индивидуальные качества и характеристики особей.

2.3. Использование маркеров для защиты новых сортов

Молекулярные ДНК-маркеры могут быть использованы для оцен-

ки сходства двух особей по многим локусам, и преимущество таких маркеров заключается в том, что они явно превосходят в этом фенотипические маркерные признаки. Это открывает отличные перспективы их использования для правовой защиты вновь создаваемых сортов [349]. Возможность такого применения молекулярных маркеров вызвала неподдельный интерес селекционеров, несмотря на то, что использование молекулярных ДНК-маркеров для этих целей все еще пока не принято и не одобрено к применению UPOV (Union pour la Protection des Obtentions Vegetales) — организацией, занимающейся персонализацией и защитой новых сортов. Кратко основную идею использования молекулярных маркеров в идентификации сортов можно выразить так: если после анализа многих полиморфных локусов новый сорт Б идентичен существующему сорту А, можно утверждать, что сорт Б есть копия сорта А. Следует заметить, что утверждение неидентичности сортов А и Б на основе разницы по одному маркеру не обеспечивает надлежащей защиты, так как генотип особи по данному локусу может быть легко изменен, например, простым обратным скрещиванием. Это было также декларировано в конвенции UPOV от 1991 года, в которой говорилось, что понятие «существенных различий» должно быть обязательным образом введено (www.upov.int/en/publications/conventions/ndex.html). Понятие «существенных различий» применимо и в ситуации, когда линия Б, даже несмотря на то, что она отличается от линии А, по своим экономическим характеристикам и использованному при ее получении селекционному процессу подобна линии А. Это значит, что должны быть установлены соответствующая граница или некое пороговое значение, определяющее различия между сортами. Например, как минимум 25% локусов должны иметь фиксированные различающиеся аллели, как это было предложено в [347]. Точность в установлении таких различий — это ключевой момент при проведении подобного рода анализов и экспертиз.

2.4. Перспективы изучения генетического разнообразия

Упомянутые выше методические подходы в большинстве своем уже повсеместно используются в различных селекционных программах за рубежом. В России интерес к ним также увеличивается (см. [17]).

Среди других потенциальных применений молекулярных маркеров, в частности, для изучения генетического разнообразия образцов дикорастущих и культивируемых видов растений, следует отметить подбор и идентификацию родительских форм, на основе которых получают новые селекционные популяции, что до сих пор остается ключевым моментом для большинства селекционеров. Ценность популяции для проведения селекционно значимого отбора зависит не только от среднего значения ее характеристик, но также и от степени их варьирования [139]. Среднее популяции в целом может быть предсказано исходя из среднего характеристик родителей. Молекулярные маркеры позволяют прогнозировать некоторую степень варьирования, что, как правило, бывает особенно полезным. Во-первых, изучение свойств взаимоотношения между фенотипом и молекулярными дистанциями приводит к заключению, что если предстоит получить селекционную популяцию из двух линий, которые близки на фенотипическом уровне, то варьирование полученной популяции будет выше для линий, которые различаются по маркерным локусам, нежели в случае линий, которые схожи по ним. Кроме того, в рамках определенной гипотезы, можно также показать, что варьирование количественного признака внутри популяции напрямую коррелирует или зависит от разнообразия этой популяции по QTL, ответственным за ее вариацию. Поэтому разнообразие, установленное на уровне маркерных локусов, может являться предсказанием степени разнообразия популяции. Некоторые экспериментальные результаты поддерживают тот факт, что популяции, полученные от родителей с высокой маркерной симилярностью, действительно проявляют низкое генетическое варьирование [236] (для сои). Однако другие исследования показывают низкие корреляционные коэффициенты между генетической вариацией популяций и молекулярной дистанцией между родительской парой ([259], для кукурузы; [200], для сои; [59], для яровой пшеницы), и, вследствие этого, необходимы дополнительные исследования для установления возможностей и ограничений данного подхода.

2.5. Маркерная помощь селекции (отбору)

Любой вклад молекулярных маркеров в селекцию основан на перспективности использования маркеров, сцепленных с генами, отвечаю-

щими за изменчивость отбираемых признаков [358]. Таким образом, термин «маркерная помощь селекции (отбору)» относится к любой форме селекции (отбора), использующей маркеры для:

- управления рекомбинацией, с целью накопления желаемых генов или сегментов хромосом в одном генотипе и
- установления, как минимум частично, генотипического значения (ценности, полезности) особи посредством идентификации генотипа в различных маркерных локусах.

Указанные стратегические подходы могут быть использованы как для получения или создания новых и/или специфичных генотипов, так и для улучшения существующих популяций посредством рекуррентной селекции.

2.5.1. Создание генотипов

2.5.1.1. Маркерная помощь при беккроссировании генотипов с моногенным признаком

Некоторые моногенные признаки, такие, например, как устойчивость к заболеваниям и насекомым, трудно или дорого оценивать в процессе проведения циклов беккроссирования. Более того, в случае переноса рецессивных генов для выявления рецессивных гомозигот необходимо «вставить» одно поколение самоопыления после скрещивания с рекуррентным родителем, что по сравнению с вариантом переноса доминантного гена, удлиняет процесс переноса. Молекулярные маркеры предлагают решение этих проблем. Кодоминантные маркеры, позволяющие идентифицировать гетерозиготные генотипы, значительно облегчают процесс. Однако, как мы указывали ранее [13], использование практически изогенных линий и BSA (bulk segregant analysis) является приемом, в котором выбирают маркеры, тесно сцепленные с интрогрессивным геном. Эти маркеры обычно доминантны (RAPD, AFLP и др.), но они также могут быть использованы в маркерной помощи беккроссированию. Такое становится возможным в случаях сцепления доминантного аллеля маркерного локуса с желаемым аллелем, который необходимо интрогрессировать. Это означает, что ситуация становится похожей на ситуацию с кодоминантными маркерами и «вставка» само-

опыленного поколения уже не требуется. Другое достоинство применения маркеров заключается в том, что особи могут быть отобраны намного проще, чем в случае традиционной беккроссной селекции и к тому же без нанесения какого-либо вреда окружающей среде. Таким образом, использование молекулярных маркеров экономит время и, что совершенно очевидно, позволяет существенно сократить расход денежных средств.

Более того, как известно, основным недостатком классического беккроссирования является тот факт, что не только ген родителя-донора, но также и фрагмент хромосомы, окружающий этот ген, интродуцируется в растение-реципиент. Поскольку мейоз рекомбинирует геном довольно грубо, то длина донорного фрагмента, несущего селективируемый ген уменьшается очень медленно и только через поколения. Так, например, после пяти беккроссов остается примерно около 30 сМ от фрагмента хромосомы в 100 сМ [166, 353]. Это было проверено экспериментально на различных линиях томата, в которых ген устойчивости к вирусу табачной мозаики *Tm-2* был интрогрессирован посредством беккроссов [413]. Полученные авторами результаты свидетельствуют о том, что практическая изогенность интрогрессированных линий относительна. Зная, что в геноме в среднем наблюдается несколько десятков генов на сМ (и даже если все они не являются полиморфными), становится ясным, что интрогрессированные линии могут существенно отличаться от рекуррентной линии по различным признакам. Молекулярные маркеры предлагают частичное решение этой проблемы за счет использования методологии *полиморфизма длин рестриктивных интродуцированных сегментов*. Кроме того, маркеры предоставляют возможность *ускорить возврат к изогенности* за пределами интродуцированного сегмента. В идеале для этого должно быть два типа маркеров:

- те, которые фланкируют находящийся под вопросом локус, и, тем самым определяют справа и слева «окно», причем таким образом, чтобы можно было определить ближайший (насколько возможно) кроссовер гена. Практически невероятно, что в одном поколении будет найден целевой двойной рекомбинант. Вследствие этого, выявление кроссоверов будет происходить последовательно. Сначала в ВС1 (первое беккроссное поколение) необходимо найти рекомбинацию с

одной стороны гена и затем в ВС2 с другой. Вероятность этих событий может быть определена как функция размера популяции и позиции маркеров, находящихся по соседству с геном, что позволяет вычислить размер популяции, с которой необходимо производить манипуляции [177].

- те, которые позволяют ускорить возврат к рекуррентному родителю и, вследствие этого, они должны быть распределены по всему геному. Как правило, для достижения этих целей порядка двух маркеров на плечо хромосомы вполне достаточно. В основном такие маркеры используют после ВС2, преимущественно в ВС3. Однако они могут быть использованы и на каждом из предыдущих шагов, если в распоряжении исследователя имеются несколько растений с благоприятным рекомбинантным событием.

За три поколения с помощью молекулярных маркеров можно получить такую изогенность, которая будет лучше, чем изогенность, полученная традиционными методами селекции. Теоретические исследования показывают, что такая стратегия, по сравнению с традиционными технологиями, сокращает время, как минимум, на два поколения [178]. Это очень хорошо проиллюстрировано экспериментами на кукурузе [310]. Так, ген *Vt*, который кодирует инсектицидный белок, был интродуцирован трансгенозом в одну из линий кукурузы. Генетический анализ показал, что *Vt* располагался в единичном локусе на хромосоме 1. Необходимо было беккроссированием перенести этот ген в другую хозяйственно ценную линию. Посредством молекулярных маркеров возврат к геному рекуррентного родителя был осуществлен уже после трех беккроссов. Это, как было доказано, оказалось эквивалентно тому, что можно было бы получить за шесть поколений, используя методы традиционной селекции. После третьего беккросса в отобранных особях оставался гетерозиготным только фланкирующий ген район (между 2% и 24% общей длины хромосомы). После четвертого беккросса расстояние сократилось максимум до 8,4%, что соответствует дистанции 14,5 сМ.

Тот же процесс можно применить и в случае нескольких генов, но число поколений и/или размер популяции при этом должны быть увеличены.

2.5.1.2. Маркерная помощь при беккроссировании полигенного признака

2.5.1.2.1. Перенос хромосомного сегмента, несущего QTL

Описанный выше процесс, или его варианты, порой применяют для переноса, представляющего интерес идентифицированного в процессе выявления и локализации QTL хромосомного сегмента, который называют QТА (от англ. Quantitative Trait Allele). В этом случае, донорным генотипом является один из родителей, участвовавших в скрещивании, применявшегося для определения QTL, а реципиентный родитель в качестве второго генотипа. Помимо существенной трудности фенотипической оценки, существует еще одна трудность, которая заключается в том, что в каждом беккроссном поколении QТА будет передаваться с вероятностью менее чем $\frac{1}{2}$ нежели какой-либо ген. Действительно, как это было показано нами ранее [13], доверительный интервал позиции QTL, т.е. хромосомный сегмент, где QTL может находиться, обычно довольно велик. Это значит, что в течение мейоза существует вероятность рекомбинации внутри сегмента. Для сохранения QТА необходимо увеличить на $\frac{1}{2}$ вероятность с таким расчетом, чтобы не было рекомбинации внутри сегмента. Для сегмента в 20 сМ, маркированного по краям и в середине, эта вероятность передачи составит около 0,4 вместо 0,5, т.е. $(1-0,1)^2$, с приближением, что 10 сМ соответствует $p = 0,1$. Кроме того, показано, что использование трех маркеров на QТА позволяет осуществлять контроль за переносом [177].

Известен пример переноса QТА, при использовании для этого маркеров, у томата [369]. Авторы нашли QTL для серии признаков в ВС2 популяции, полученной в результате скрещивания сорта культурного вида *Lycopersicon esculentum* (рекуррентный родитель) с диким видом *L. pimpinellifolium* (донорный родитель). Фенотипические значения каждой ВС2 особи были измерены как в поколении ВС3, так и в популяции, полученной в результате скрещивания данной особи с тестером. Найденные три QTL (один — оказывающий влияние на массу плодов и два других — на их размер) были независимо интрогрессированы для получения практически изогенных линий. Сравнение этих линий после скрещивания с тестером показало, что в каждом случае интрогрессия существенно модифицировала значение признака в направлении ожи-

даемого эффекта данного QTL, измеренного в BC2. Так, QTA *L. pimpinellifolium* расположенный на хромосоме 8, делал плоды круглее, чем у культивируемого вида (–38% по шкале длины). Другой QTA, определяющий размер плодов, расположенный на хромосоме 1, удлинял плоды (+8%). Необходимо заметить, что в двух последних случаях QTA, с которыми проводились манипуляции, давали эффект противоположный тому, что ожидался у родительского фенотипа: плоды *L. pimpinellifolium* круглые и очень маленькие, а у *L. esculentum* — большие и удлиненные. Благодаря этому было возможным отличить дикую форму, которая исходно была не похожа на что-либо другое, от QTA улучшающего некоторые «культивируемые» признаки культурного сорта. Этот подход, сделавший использование молекулярных маркеров намного более простым, чем в предыдущем варианте, несомненно, является многообещающим методом для маркер вспомогательной селекции, включая практическое использование генетических ресурсов растений.

2.5.1.2.2. Перенос нескольких хромосомных сегментов

Принцип беккроссирования с несколькими генами может быть применен и для случая переноса нескольких желаемых QTA. Если принять во внимание вероятность отсутствия или нереальности рекомбинации внутри сегмента, то в этом случае проведенные вычисления показывают, что для популяции размером примерно в 300 особей не представляется возможным осуществлять манипуляции одновременно более чем с четырьмя QTA [177]. Очевидно, что проведение подобного рода действий гораздо легче для популяций с большим числом особей и с маркерами, которые очень тесно или даже непосредственно сцеплены с генами.

Если обрисовать схематично, то беккроссирование может быть выполнено, если придерживаться некоторых правил, согласно которым необходимо:

- отдавать приоритет желаемым QTA донорного родителя. Эти QTA, которые будут всегда поддерживаться в гетерозиготном состоянии благодаря и посредством получения поколений беккроссов, не должны быть потеряны.
- отбирать особи, которые несут желаемые QTA, для их изогенности с

рекуррентным родителем за пределами манипулируемых фрагментов.
– заканчивать процесс беккроссирования одним или двумя самоопылениями для закрепления интрогрессированных QTA.

Эксперимент по изучению признака урожая зерна у кукурузы ясно показал роль и значение генетического построения вспомогательного маркерного генотипа для повышения комбинационной способности родителей гибрида [177, 179]. Так гибрид В73 × Мо17 был улучшен интрогрессией желаемого QTA двух других линий, Тх303 и Oh43. Поиск QTL показал наличие шести хромосомных районов, для которых Тх303 может улучшить В73 и шесть других, с помощью которых Oh43 может улучшить Мо17. Три беккросса, следующие за двумя самоопылениями, были выполнены в каждом из этих случаев. Отбор по маркерам (RFLP) авторы начали проводить со второго беккросса: особи отбирали на основе числа перенесенных сегментов и процента генома рекуррентного родителя. Результаты оказались обнадеживающими. Скрещивания между «улучшенной В73» и Мо17, а также между В73 и «улучшенной Мо17», дали гибриды, которые в целом имели лучшие характеристики, чем контрольный В73 × Мо17. Частота улучшения составила 8-15%. Анализ улучшенных линий показал, что максимальное число перенесенных сегментов не превышало трех и что большинство содержало только один сегмент, полученный от донорного родителя: сегмент 6L для В73 и сегмент 4S для Мо17. Эти два QTL, по-видимому, имели наиболее существенный эффект действия на генотип улучшаемых линий. Ни одна из линий не содержала все шесть ожидаемых фрагментов, поскольку вероятность получить их была очень мала. Ясно, что такое улучшение невозможно осуществить в столь короткие сроки при использовании традиционных методов селекции.

2.5.1.3. Маркерная помощь при отборе по потомству

В поколении F₂, вместо проведения беккроссирования, среди особей, полученных самоопылением (F₃, F₄ и т.д.), можно попытаться отобрать те, которые получили лучшие QTA.

Для того чтобы отобрать линии на основе значения их характеристик *per se*, обычно необходимо произвести отбор линий, которые могут быть или были получены от одной особи. Самоопыление гетерозигот-

ной особи не позволяет напрямую получить величину значения характеристик этих линий. Такая величина, как известно, указывает пределы отбора по потомству, как впрочем, и преимущества гаплодиплоидизации. С другой стороны, явное превосходство маркеров заключается в том, что они позволяют нам предсказывать это значение для маркированных локусов. Пусть a_l будет равно половинному значению различия между двумя гомозиготами по QTL l . Для данного признака среднее значение линий, что может быть определено исходя из генотипа i , может быть записано как:

$$\hat{G}^i = \sum_l \hat{g}_l^i, \quad (25)$$

где $\hat{g}_l^i = a_l \Psi_l^i$, $\Psi_l^i = -1$ или $+1$ в соответствии с тем, является ли генотип в избранном маркере гомозиготным по аллелю от неблагоприятного родителя или от благожелательного, соответственно. Генотипическое значение семьи, полученной в результате самоопыления ее исходного пра-родителя, также можно прогнозировать, принимая во внимания эффекты доминирования. Примером действенности данного подхода может служить работа, выполненная на томатах [285]. Авторы сравнили эффективность отбора, проведенного на основе только лишь маркеров с отбором по фенотипу, при этом отбор выполнялся на особях F_2 , а ответ изучался в семьях F_3 . Для pH плодов (признак, который наследуется средне) отбор по маркерам имел эффективность, сравнимую с отбором по фенотипу. В противоположность этому, для массы плодов, более наследуемый признак, отбор по фенотипу был более эффективным, в то время как для концентрации растворимых сахаров, плохо наследуемый признак, отбор на основе маркеров был наиболее эффективным (табл. б). Таким образом, как следует из результатов данной работы [285], отбор по фенотипу более эффективен для признаков с высокой наследуемостью, поскольку он базируется на источниках изменчивости всех QTL одновременно, в то время как отбор по маркерам задействует только те QTL, которые сцеплены с маркерами. Следовательно, в случае низкой наследуемости признака генетические значения лучше устанавливать посредством молекулярных маркеров, нежели по фенотипу, если размер популяции при этом, по определяемым QTL, достаточно велик.

Преимущество отбора только по маркерам для улучшения признака «урожай зерна» также было проверено и у кукурузы [356]. Отбор F₆ семей (потомство совокупно самоопыленных растений F₂) на основе только информации полученной с помощью маркеров, сцепленных с определенными в F₂ QTL, позволило получить в среднем такой же результат, как и в случае отбора по фенотипу. В то же время, менее чем 50% изменчивости в данном случае можно было объяснить, используя для этих целей только лишь маркеры.

Таблица 6

Относительная эффективность отбора по родословной в F₂ на основе значений, определенных с помощью маркеров, в сравнении с отбором по фенотипу в скрещивании *Lycopersicon esculentum* x *L. chesmanii* (по [285])

	Масса плодов	pH плодов	Концентрация растворимых веществ
Наследуемость	0,45	0,25	0,15
Число установленных QTL	7	5	4
R ² (%) – процент фенотипической изменчивости QTL	72	34	44
Относительная эффективность маркеров	0,66	1,07	2,20

С другой стороны, если целью является отбор линий по их комбинационной способности с тестером, то в этом случае необходимо установить эффекты QTL по данной способности. Таким образом, значение линий, каждая из которых получена от одной отдельной особи, будут спрогнозированы посредством суммирования эффектов каждого QTL, в соответствии с имеющимися в наличии аллелями. В то же время обрабатывать такой материал надо так, как если бы работа велась с аддитивным признаком [139].

2.5.2. Маркерная помощь рекуррентной селекции (рекуррентному отбору)

Если число QTA, которые должны быть объединены в одном геноме, велико, то вероятность получить необходимый генотип за один цикл маркерной помощи создания такого генотипа быстро уменьшается. Поэтому, как представляется, изначально лучше получить популяцию, в

которой частота желаемого генотипа будет приумножена за счет увеличения частот каждого из необходимых QTA. Сделать это можно посредством проведения нескольких циклов отбора следующих за рекомбинацией. Такая маркерная помощь рекуррентному отбору, базирующаяся только на использовании маркеров, может быть осуществлена с или без учета эффектов QTA. Однако существует проблема, так называемой, немаркерной генетической изменчивости, которая для комплексных агрономических признаков, остается в целом достаточно высокой (примерно 50% и выше). Эта изменчивость может быть использована только в рамках «традиционных» методов рекуррентного отбора, и только если наследуемость для этого достаточна. Таким образом, для максимального увеличения эффективности генетического процесса будет достаточно скомбинировать оба типа отбора (на уровне маркеров и по фенотипу) по рекуррентной схеме. При этом первая стадия этой схемы должна быть применена к популяции F_2 , где QTL и были определены.

2.5.2.1. Рекуррентный отбор, основанный только на маркерах

Возможен прямой отбор маркеров, дающих одинаковое значение для каждого QTA. Цель такого метода — увеличить частоты нужных маркеров, т.е. маркеров, которые сцеплены с желаемым QTA. Теоретические исследования [142, 179, 229] показали, что частота отбираемых аллелей увеличивается быстро. Например, для 10-ти независимых локусов частота желаемых аллелей повышается с 0,50 до 0,90 всего за три поколения отбора. Благодаря этому частота получения необходимого генотипа возрастает с 0,001 до 0,35. Вслед за этим, вполне доступно получить наилучший маркерный генотип для улучшенной популяции. Эффективность такого подхода зависит в основном от того, насколько тесно сцеплены маркер и QTL. Если они сцеплены не тесно, то рекомбинация между маркерами и QTL быстро уменьшит часть изменчивости, объясняемой маркерами. В таком случае использование фланкирующих маркеров позволит увеличить эффективность отбора. В то же время, принимая во внимание уровень фиксации, а также риск рекомбинации между маркерами и ассоциированными с ними QTL, лучше ограничить применение этого метода лишь некоторыми подходящими для этих целей циклами отбора.

2.5.2.2. Рекуррентный отбор по аддитивному значению, прогнозируемому с помощью маркеров

Для увеличения эффективности схемы рекуррентного отбора важно спрогнозировать аддитивное значение особей — кандидатов на отбор, причем только то значение, которое передается из одного поколения в другое. Это значение прогнозируют по-разному, сообразно тому улучшался ли материал с целью получения гибридного или линейного сорта.

На подготовительном этапе получения гибридных сортов или линий, материал можно улучшить, используя его комбинационную способность со специфичным тестером. Как описано выше, такую комбинационную способность можно спрогнозировать посредством простого суммирования эффектов QTL. Если целью является улучшение внутрипопуляционной комбинационной способности, то тогда необходимо иметь возможность устанавливать аддитивные эффекты. Без маркеров аддитивное значение особи, как правило, определяют посредством изучения потомства, полученного по нескольким сопряженным друг с другом схемам. При наличии маркеров, если известны частота маркера и ассоциированный с ним эффект, появляется непосредственная возможность установить это значение. В поколении F_2 при частоте каждого маркерного аллеля равной 0,5 аддитивный эффект прямо пропорционален половине различия между двумя гомозиготами. Однако, в зависимости от того как долго проводился отбор, аддитивный эффект аллеля меняется, поскольку он зависит от частоты встречаемости аллеля. Поэтому можно получить средний эффект действия данного аллеля в комбинации с другими аллелями популяции [139]. Возможность установления аддитивного значения QTL, по-видимому, является основным вкладом молекулярных маркеров в рекуррентный отбор. Для каждого индивидуального кандидата на отбор можно посчитать значение, проявляемого им аддитивного эффекта ассоциированное с картированным QTL. Общее аддитивное значение получают простым сложением величин, спрогнозированных для каждого QTL. Если целью является получение линейных сортов, то, в этом случае, необходимо отбирать по величине значений характеристик линий, которые могли бы быть получены из

данного генотипа [139]. Как и при отборе по родословной, преимущества маркеров заключаются в том, что они позволяют установить нужное значение посредством вычисления $\frac{1}{2}$ величины различий между двумя гомозиготами и суммированием предсказанных для каждого QTL значений.

Точность прогнозируемых аддитивных значений зависит от генетических значений R^2 установленных QTL (% генетической изменчивости, объясняемой с помощью маркеров), наследуемости признака и размера изучаемой популяции. Последний должен быть тем больше, чем ниже наследуемость. В случае популяции достаточно большого размера, если процент генетической изменчивости, объясняемой с помощью маркеров больше, чем наследуемость в узком смысле этого слова, отбор на основе прогнозируемых значений, при их равной интенсивности, будет более эффективным, чем отбор только по фенотипу [210]. Поскольку отбор, основанный только на маркерах, не требует полевой оценки и поскольку идентификация генотипов по маркерам производится очень быстро, продолжительность селекционного цикла может быть уменьшена, как минимум, на одно поколение. Такой анализ можно проводить как в условиях недостаточности временного интервала для осуществления циклов или манипуляций традиционной селекции, так и в условиях искусственного выращивания материала или в период межсезонья. Вследствие этого, маркерная селекция, на единицу времени, может быть более выгодной, чем классический отбор по фенотипу, даже когда маркерный отбор менее эффективен на цикл. Порой через несколько циклов селекционного отбора возникает проблема рекомбинации между установленными QTL и их маркерами. Из-за этого относительная эффективность данного метода будет падать с каждым циклом, поэтому строгое сцепление между маркерами и QTL всегда предпочтительней. В идеале, конечно, было бы неплохо иметь доступные маркеры в самих QTL. При отсутствии такой ситуации, как в случае предыдущей схемы отбора только по маркерам, лучше сократить схему реализации отбора всего до нескольких циклов. По-видимому, для короткого периода времени такой подход будет более эффективным, нежели упомянутая выше схема, в которой каждому QTL придается одинаковая значимость. Обратное будет верным для длительного временного периода.

Таблица 7

Результаты экспериментов (по [357]) по изучению признака урожай зерна, сравнение цикла отбора только по маркерам (после определения QTL) и цикла классического фенотипического отбора (массовый отбор)

Тип отбора	Ответ (реакция)	
	Урожай зерна	Изменчивость в частоте желаемых QTA
Положительный отбор по: только маркеры	151,2	+0,19
	151,7	+0,06
Негативный отбор по: только маркеры	107,7	-0,19
	122,4	-0,06

Результаты экспериментов [357] на популяции F_2 кукурузы ясно показывают вклад маркеров в селекционный процесс — цикл отбора только на основе маркеров является столь же продуктивным, как и селекция по фенотипу, а для некоторых признаков, даже более эффективным (табл. 7). Эксперименты также показали, что частота желаемых маркерных аллелей увеличивается много больше с отбором на основе маркеров, нежели с отбором по фенотипу, в то время как конечный результат получается один и тот же. Отбор по фенотипу, как правило, использует другие источники изменчивости, в частности, немаркированные QTL. Вследствие этого, становится очевидным, что комбинированный отбор, основанный как на фенотипе, так и на маркерах, является более эффективным. К такому же заключению можно прийти, проанализировав результаты экспериментов, выполненных и на томатах (табл. 6) [285].

2.5.3. Комбинированный отбор, основанный на фенотипе и маркерах

2.5.3.1. Принцип и ожидаемая эффективность

В некоторых экспериментах, преимущественно на популяциях F_2 , маркерный отбор на единицу времени оказался вполне эффективным. Кроме того, он не требовал полевой перепроверки. Однако через несколько поколений его эффективность может упасть очень быстро из-за фиксации желаемых QTA и рекомбинации между QTL и маркерами. В этом случае, существует вероятность переоценки взаимоотношения ме-

жду маркерами и улучшенным признаком и, как результат, исследователь может ошибочно продолжить отбор только по маркерам. В то же время, существует много QTL с очень слабыми для их выявления эффектами, изменчивость которых не может быть использована при отборе. При одновременном отборе как по фенотипу, так и по маркерам задействуются все источники изменчивости и, вследствие этого, окончательный результат будет намного лучше.

При таком подходе маркеры необходимо рассматривать как ассоциированные признаки: их генетическая корреляция с основным признаком прямо пропорциональна квадратному корню процента изменчивости, объясняемой с помощью маркеров. Из общих соображений по много признаковому отбору можно предсказать, что наиболее эффективный путь использования двух источников информации — это комбинирование их в индексе I фенотипического значения P и аддитивного значения, предсказанного по маркерам A [210]:

$$I = b_1P + b_2A, \quad (26)$$

где b_1 и b_2 коэффициенты частичной регрессии истинного генетического значения G у P и у A , соответственно, которые могут быть получены на основе регрессионной теории с помощью модели случайных эффектов. Можно показать, что:

$$b_1 = \frac{h^2(1-r_G^2)}{1-r_p^2} \quad (27) \quad \text{и} \quad b_2 = \frac{1-h^2}{1-r_p^2}, \quad (28)$$

где h^2 наследуемость в узком смысле этого слова, r_p^2 (или r_G^2) часть фенотипической (или генотипической) изменчивости, объясняемой маркерами. Как указывают авторы [210]:

$$n_p^2 = h^2r_G^2, \quad (29)$$

это уравнение показывает, что при фиксированном r_G^2 вес (мера), приданный фенотипическому значению, увеличивается с наследуемостью признака, в то время как вес (мера), данный маркерам ослабевает.

Если наследуемость высока, то вклад молекулярных маркеров мал. Преимущество комбинированного отбора на основе фенотипа и

маркеров может проявиться только для достаточно низкой наследуемости. Однако если наследуемость получается слишком низкой, относительно того какой она должна быть для данного числа генотипированных растений, то число ошибок в определении QTL (фальшивых позитивов, низкая точность при определении эффектов QTL) будет возрастать. В таком случае маркеры ничего не вкладывают в отбор и даже могут приводить к ошибкам. Для того чтобы сохранить преимущества комбинированного отбора, размер популяции должен быть больше, особенно, если наследуемость низка. Как правило, только у выборки от 100 до 300 растений имеется наследуемость, которая максимизирует относительный интерес комбинирования отбора по фенотипу и по маркерам и он находится в пределах 0,15-0,20 [141, 267].

Очевидно, что эффективность молекулярных маркеров тем выше, чем больше процент генетической изменчивости, объясняемой с помощью маркеров. Обычно это результат Nh^2r_G , который необходимо принимать во внимание, при этом N — это размер популяции [143, 210]. Таким образом, область действенности маркерной помощи комбинированному отбору соответствует очень низкой наследуемости (между 0,10 и 0,30), очень высокой плотности маркеров и достаточно большому размеру популяции [141].

Проблема, которая появляется с этим типом отбора — затраты в финансовом плане, поскольку переоценки взаимоотношения маркер-признак требуются при проведении каждого селекционного цикла. Если совместить маркерную селекцию и отбор по фенотипу, сравнив в их оптимуме по одним и тем же вложениям, то комбинированная маркерная помощь отбору, как представляется, имеет небольшие преимущества [268]. Эффективное решение, включая материальные вложения в генотипирование и проведение эксперимента, может заключать в себе перемежающееся комбинирование «фенотип + маркерная селекция» с отбором только по маркерам [142, 180]. Причем отбор только по маркерам можно сделать для двух или трех циклов с целью получения так называемых, «форсированных» поколений, поскольку такой подход не требует полевой оценки. Следовательно, прогресс на единицу времени здесь будет выше, чем в случае совмещенного отбора.

2.5.3.2. Результаты моделирования

Ожидаемая эффективность комбинированного отбора была проверена моделированием [422], но на большом размере популяции, который дает преимущества даже при низкой наследуемости. Моделирование [149, 150, 180] привело к следующим наблюдениям:

- прогресс достигает своего плато намного быстрее, если неравновесие маркер-QTL не переоценивается в каждом поколении: изменчивость, объясняемая с помощью маркеров, быстро исчезает и наилучшие результаты можно получить посредством использования другого подмножества маркеров, установленных в новом исследовании взаимодействия маркер-признак;
- индекс (показатель) эффектов плотности и числа маркеров, кажется относительно низким по сравнению с индексом (показателем), полученным для размера популяции;
- метод комбинированного отбора лучше всего подходит для случаев с низкой наследуемостью;
- отбор не должен быть слишком жестким или строгим, особенно когда QTL находятся в фазе отталкивания;
- увеличение числа маркеров на хромосому не обязательно приводит к более достоверному результату отбора.

2.5.3.3. Сравнение на основе создания генотипов

Методы маркерной помощи отбору не сразу приводят к получению особых, т.е. обладающих исключительными свойствами, генотипов из популяций ГРР. Однако если отбор в популяции, проведенный на основе только маркеров, сравнить с маркерной помощью беккроссам, то становится очевидно, что цель в обоих случаях одна: собрать в одном генотипе максимальное число желаемых аллелей. Построение генотипа беккроссированием, управляемое отбором, является результатом предпочтительных рекомбинаций. На уровне популяций процесс более вероятностный (пробабалистичный) благодаря увеличению частот желаемых QTA. В обоих случаях результатом является полученная в конце процесса селекционная популяция, из которой должны быть отобраны лучшие генотипы, предназначенные облегчить беккроссирование.

Схожим образом рекуррентный фенотип + маркерный отбор увеличивают частоту ассоциаций интересующих исследователя генов посредством увеличения частот маркированных и немаркированных желаемых QTA. Здесь также необходимо отобрать лучшие генотипы, которые более эффективно могут быть получены посредством гаплоидизации, безусловно, если этот подход может быть реализован на рассматриваемых видах. И, конечно, необходимо учитывать маркерную помощь отбору по родословной. Преимущества комбинированного отбора «фенотип + маркер» заключается в том, что он использует все источники изменчивости. Скрещивания между комплементарными растениями могут быть рассмотрены для увеличения вероятности трансгрессии с целью сохранения изменчивости немаркированных QTL. Однако при наличии достаточного большого числа маркированных QTL управление системой становится более трудным и более затратным в финансовом плане, нежели случайное перекрестное скрещивание отселектированных растений.

2.5.3.4. Выбор между использованием маркеров и увеличением числа репликаций (репродукции, повторности, воспроизводства)

Для комбинированной маркерной помощи отбору важным преимуществом является то, что маркеры позволяют устанавливать генотипическое значение, которое равно степени увеличения наследования. Действительно, генетическое значение является более точно предсказуемым за счет использования маркеров, ассоциированных с QTL, поскольку каждая особь, несущая маркер QTA является своеобразной репликой нужного QTA. Чем больше особей содержат данный QTA, тем больше точность отбора. Таким образом, оценка по маркерам может быть заменена настоящими репликами особей, с хорошо известными средними значениями для увеличения наследуемости. Оба способа увеличения наследуемости, как это было показано ранее [140, 141], должны быть в арсенале любого селекционера. Селекционер должен иметь возможность выбора между тем, чтобы делать большее число репродукций без использования маркеров, с одной стороны, и использованием маркеров, что позволяет существенно уменьшить число репродукций, с

другой стороны. Таким образом, проблема из биологической по существу становится чисто экономической.

Если снова вернуться к маркерной помощи отбору, то, как представляется, применение молекулярных маркеров может сократить время селекционного процесса в межсезонный период, и, прежде всего, за счет отсутствия необходимости проведения полевой оценки. Более того, маркеры просто незаменимы для управления процессом рекомбинации для того, чтобы аккумулировать желаемые аллели в одном генотипе так быстро насколько это возможно. Но для того, чтобы использовать достоинства и преимущества молекулярных маркеров полностью необходимо разработать новые схемы рекуррентной селекции или создания рекуррентных генотипов.

Заключение

Генетические маркеры могут быть использованы на различных стадиях реализации селекционных программ: оптимизация сохранения генетических ресурсов, подбор родительских пар, идентификация желаемых аллелей и получение генотипов, аккумулирующих в себе такие аллели. В целом ряде лабораторий биотехнологических и селекционных фирм, таких как TraitGenetics, Enza Zaden, Icon Genetics и др., для некоторых из этих направлений исследований молекулярные маркеры применяют уже повседневно. Значительное число частных и государственных программ уже сейчас на постоянной основе могли бы использовать анализ генетической изменчивости с целью организации и классификации селекционного или иного материала в различного рода группы (или для подтверждения статуса существующих групп), а также маркерную помощь селекции при проведении различного рода беккроссов. Другие области применения молекулярных маркеров на сегодня по большей части остаются прерогативой научно-исследовательской деятельности в биологической и сельскохозяйственной науках. Это в основном случаи селекционного отбора генотипов по количественным признакам, для которых основополагающие принципы уже разработаны, но для которых пока еще накоплено не так много экспериментальных данных. Поэтому данная область генетики и селекции растений пока остается лишь своеобразным полигоном для разработки новых и оптимизации существ-

вующих методических подходов и приемов. Последующее развитие этих методов будет зависеть от полученных результатов, так же как и от изменения цены информационных данных (генотипирования особи в расчете на один локус). Тем не менее, молекулярные маркеры уже необратимо дополнили собой ряд имеющихся, разработанных к сегодняшнему дню и доступных для селекционеров методов по управлению генетической изменчивостью, включая ее практическое использование [11, 14, 17, 94, 412].

Глава III. КАРТИРОВАНИЕ И ХАРАКТЕРНЫЕ ОСОБЕННОСТИ ХРОМОСОМНЫХ ЛОКУСОВ, ОПРЕДЕЛЯЮЩИХ ПРОЯВЛЕНИЕ КОЛИЧЕСТВЕННЫХ ПРИЗНАКОВ В ПОПУЛЯЦИИ

Формальная популяционная генетика имеет дело с дискретными, четко различимыми, чаще всего противительными вариациями (или проявлениями) признака, такими как: круглые/морщинистые семена у гороха, белая/красная окраска цветов у петунии, нормальный/карликовый рост растения и др. Существует значительное число признаков, проявление которых контролируется одним геном, и для которых «дикий» фенотип ясно отличается от «мутантного» фенотипа. Молекулярный анализ показывает, что, в своей массе, мутантные фенотипы являются следствием крупных изменений в структуре гена (например, ген *r* у гороха, определяющий морщинистую форму семени, инактивирован инсерцией; [40]). В природных популяциях мутации обычно редки, поскольку они, как правило, подвергаются действию естественного отбора, являющегося результатом их явно выраженного фенотипического эффекта.

В то же время, множество признаков в популяциях проявляют *непрерывно варьирующую* изменчивость. По таким, так называемым, количественным признакам, особи не могут быть *разделены на классы*, как в случае дискретной, четко различимой изменчивости, но могут быть *измерены*. Такое продолжительное распределение может быть объяснено без нарушения законов менделевской генетики, посредством простого предположения, что *несколько* сегрегирующих локусов влияют на изменение признака, проявление которого может зависеть от взаимодействия «генотип – окружающая среда». С начала 1980-х годов такие локусы стали называть локусами количественных признаков или QTL. Тем не менее, этот термин не может прояснить, чем данные, так называемые, локусы количественных признаков отличаются по своей природе от тех, что ответственны за дискретную изменчивость. Было выдвинуто предположение о том, что существует возможность с большой точностью *устанавливать различия эффектов действия* «диких» (или активных) аллелей, ответственных за изменчивость количествен-

ных признаков популяции. Причем даже с большей точностью, чем в случае дискретного противопоставления «мутант – дикий тип». Исследования QTL признака высоты растений у кукурузы, лучший пример подтверждения правильности данной гипотезы, поскольку многие из них картированы в тех же районах, что и гены карликовости [31, 171]. Например, у потомства кукурузы QTL высоты растения были найдены на хромосоме 9 в районе гена карликовости *d3*. Возможно это один локус с различными аллельными эффектами, отвечающими как за проявление QTL, так и за проявление мутантного фенотипа в случае аллеля с потерей исходной функции [380]. Идентификация некоторых генов с количественными эффектами у растений поддерживает эту концепцию [62, 109, 411].

До последнего времени генетика количественных признаков или количественная генетика базируясь на статистике, не могла обеспечить общее описание генетических свойств этих признаков, за исключением определения более или менее верного числа генов, относительного значения аддитивных вариантов, доминирования и эпистаза вариантов в целом и расчета результата отбора на основе сходства между родственными генотипами. Даже если предположить, что этот раздел биологической науки достиг большого успеха в генетическом улучшении растений и животных, традиционная генетика количественных признаков до сих пор не может оценить *индивидуальные* эффекты QTL, их межлокусное (доминирование) и внутрилокусное (эпистаз) взаимодействие, а также их позицию в геноме, также как она не может дать никаких данных по их молекулярной идентификации. И это несмотря на то, что принцип картирования QTL известен с давних пор. В 1923 году впервые была установлена связь между окраской (моногенный признак) и массой (полигенный признак) семян гороха и высказано предположение, что ген, контролирующей окраску, генетически сцеплен с одним или несколькими факторами, контролирующими признак массы семян [330]. В последующие десятилетия последовательно были опубликованы другие примеры того же типа, но они никогда не приводили к систематизированному методу идентификации QTL. Как известно, маркеры с визуальным эффектом (фенотипические) присутствуют в сравнительно небольшом количестве в данном генетическом окружении (геноме) и мо-

гут противоречить или даже мешать изучению количественного признака (например, фенотип карликовости очень трудно использовать в качестве маркера для определения QTL признака роста). Несколько лучше дело обстоит с изоферментными маркерами, которые являют собой образец примерной иллюстрации предложенных в 1961 году принципов анализа QTL [378], но лишь молекулярные ДНК-маркеры, охватывающие весь геном, оказались способны привести к появлению нового подхода в идентификации и локализации локусов QTL в количественной генетике. Достаточно подробно различные типы маркеров, их достоинства и недостатки, рассматривались ранее [10, 11, 14] и в данном случае обсуждаться не будут.

3.1. Принципы картирования QTL

Картирование QTL основываются на систематизированном поиске неравновесного сцепления (LD) между маркерными локусами и QTL. Для этого необходимо следующее:

- иметь сегрегирующую популяцию (в которой существуют взаимоотношения между неравновесным сцеплением и генетической дистанцией). Для генетического картирования наиболее эффективными являются такие популяции, которые получены в результате скрещивания двух гомозиготных линий;
- для каждой особи популяции посредством серии маркерных локусов распределенных по геному определить ее генотип (отобранные маркеры должны позволить построить насыщенную генетическую карту);
- измерить значения количественного признака для каждой особи популяции;
- использовать биометрические методы для нахождения маркеров, чье генотипическое проявление коррелировало бы с признаком(ами), и установить генетические параметры идентифицированных QTL.

Существует три пути определения значения признака для каждой особи:

- замерить собственные значения особи (в этом случае признак должен иметь высокую наследуемость) или среднее у генетически идентичных особей (если последнее возможно то, лучше, если измерения бу-

дуг проводиться у удвоенных гаплоидов или рекомбинантных инбредных линий);

- вычислить среднее значение семьи, полученной от самоопыленной особи. Увеличение точности, таким образом, позволит ограничить число особей которые должны быть генотипированы. Однако в случае популяции F_2 установленная степень доминирования будет уменьшена на половину [359];
- установить среднее потомства, полученного от скрещивания изучаемой особи и гомозиготы «тестерного» генотипа. Данный случай интересен для селекции растений, в частности, когда необходимо получить гибридные сорта. В этом случае доминирование не может быть оценено, поскольку нет гетерозигот Q_1Q_2 , но есть два генотипа Q_1T и Q_2T (1:1), где T — аллель тестера.

Биометрические подходы нахождения коррелирующих с признаком маркеров могут быть разделены на две категории: (1) те, которые базируются на помаркерном (т.е. маркер за маркером) анализе и (2) те, которые берут в расчет одновременно два или более маркеров.

3.1.1. Определение QTL, учитывающих маркеры обособленно

Этот подход, впервые предложенный в 1976 году [350] и активно используемый с начала 1980-х годов [372], основывается на анализе варiances или на тесте Стьюдента (так называемый t -тест). Принцип состоит в простом нахождении отношений между генотипом *каждого* маркерного локуса и значением рассматриваемого количественного признака [107, 117, 196, 360].

3.1.2. Нахождение QTL посредством анализа варiances популяции F_2

Случай потомства F_2 наиболее подходящий для иллюстрации поиска QTL на основе принципа «маркер за маркером», поскольку это единственный случай, при котором один локус может быть представлен тремя возможными отличающимися друг от друга генотипами. Другие наиболее часто используемые популяции рассмотрены в разделе *Сравнение F_2 , беккроссов, удвоенных гаплоидов и рекомбинантных инбредных линий*.

В случае кодоминантного маркерного локуса M , особи популяции F_2 распределяются на три генотипических класса M_1M_1 , M_1M_2 и M_2M_2 . Для каждого класса изучаемого количественного признака вычисляется среднее и проводится анализ дисперс с целью определения достоверных различий между найденными средними. Если это проделать для каждого отдельно взятого маркера популяции (при этом становится возможным специфичное подразделение популяции), то можно построить карту генетических районов, где есть ассоциация между маркером и количественным признаком. Данный эффект в упрощенном виде проиллюстрирован на рисунке 1. Если маркер генетически сцеплен с полиморфным QTL, то подразделение потомства по маркерному локусу соответствует подразделению того же потомства по QTL. Наличие достоверных различий между средними означает, что эффекты двух аллелей QTL достоверно различаются по определяемому результату (проявлению). В дальнейшем, для «упрощенной» интерпретации, будем исходить из того, что только *один* QTL находится поблизости от маркера. Однако на практике обычно случается так, что два близлежащих QTL «мимикрируют» действие одного прилежащего к ним QTL и только исследования по точному картированию может различить их [124, 284].

Если уровень рекомбинации r между маркером и QTL равен нулю, то существуют два параметра, которые определяют действие QTL и которые легко установить — это аддитивный эффект a и степень доминирования d . В этом случае аддитивный эффект a равен $(m_{22} - m_{11})/2$, где m_{22} и m_{11} средние значения гомозигот генотипов M_2M_2 и M_1M_1 (поскольку $r = 0$, то генотипы в QTL будут, соответственно, Q_2Q_2 и Q_1Q_1). Аддитивный эффект также «разжижает» линии проходящие сквозь средние значения генотипов M_2M_2 и M_1M_1 . Это происходит, если генотипы выстраиваются согласно числу доз какого-нибудь одного из аллелей. Понятно, что a может быть отрицательным. Знак a показывает, который из родителей несет «предпочтительный» аллель, т.е. аллель, дающий наибольшее значение признака. Согласно нашим обозначениям, положительное значение a означает, что желаемый аллель получен от родителя 2, а отрицательное — от родителя 1. И наконец, удобно рассматривать эффект двойного замещения одного аллеля на другой в виде следующего выражения: $2a = m_{22} - m_{11}$.

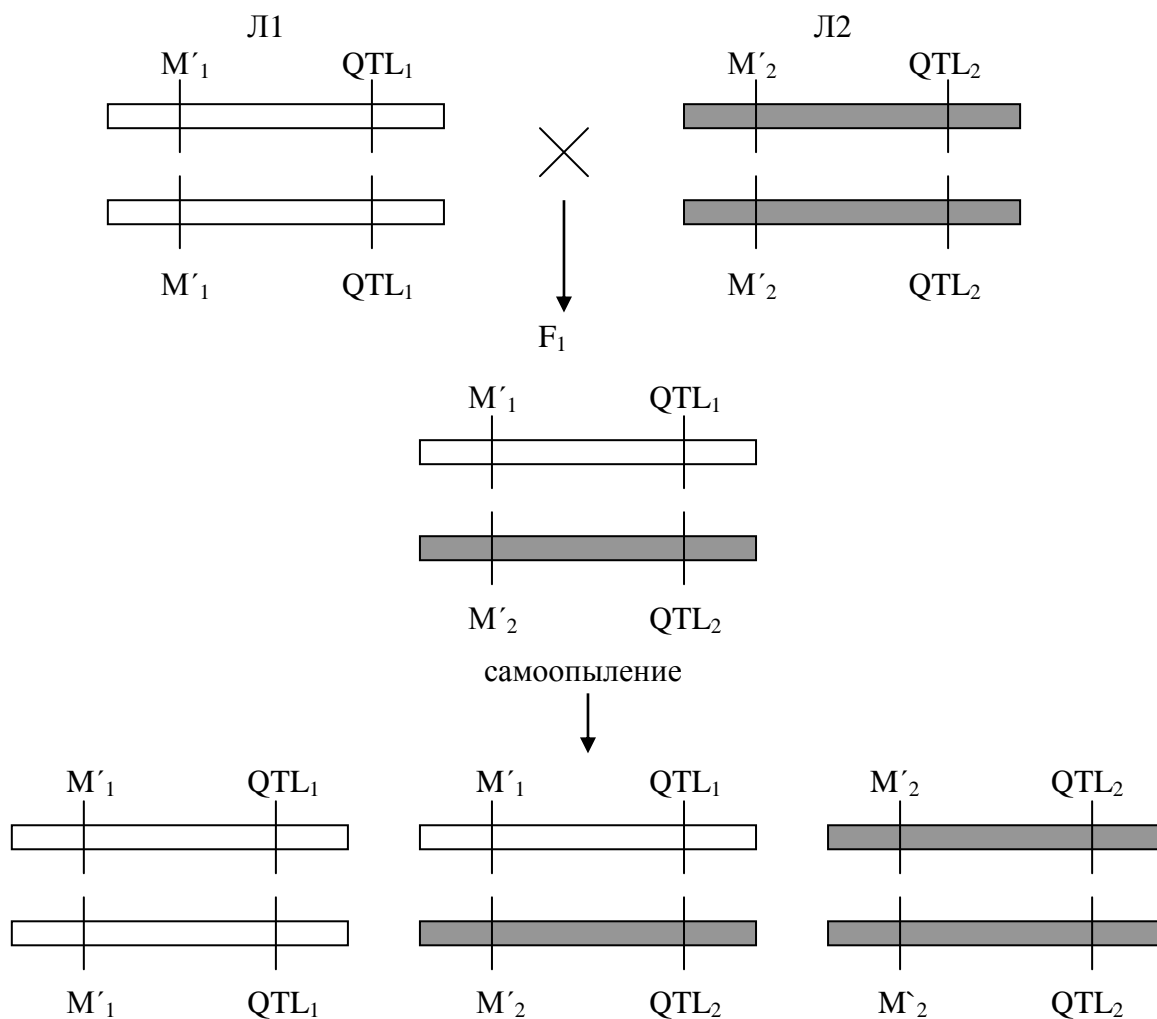


Рисунок 1. Интерпретация различий между значениями модельного одиночного QTL прилежащего к маркеру. На схеме представлены только случаи, в которых не было рекомбинации между маркером и обозначенными QTL. QTL₁ и QTL₂ — два различных аллеля QTL.

Степень доминирования — это различие между значениями гетерозигот M_1M_2 и полусуммой гомозигот: $d = [m_{12} - (m_{11} + m_{22})]/2$. Однако более содержательно использовать отношение d/a , которое позволяет классифицировать доминирование. Так, если $d/a = 0$, то это случай аддитивности. Если $0 < |d/a| < 1$, то проявляется частичное доминирование, а в случае $d/a = 1$ — полное доминирование. И, наконец, неравенство $|d/a| > 1$ означает сверхдоминирование.

В общем случае, QTL и маркеры связаны не полностью. Последствия рекомбинации в данном случае таковы, что особи гомозиготные по маркеру не все гомозиготны по сцепленному с ним QTL. Этот феномен является следствием того, что в исследуемом локусе присутствует

другой аллель в пропорции, зависящей от уровня рекомбинации между маркером и QTL. Другими словами, установленный аддитивный эффект — это аддитивный эффект в маркерном локусе a' . Отношение между a и a' зависит только от уровня рекомбинации между маркером и QTL: $a' = a(1 - 2r)$. Например, QTL имеющий аддитивный эффект $a = 7,5$ г для признака вес фруктов, приводит к кажущемуся аддитивному эффекту $a = 5$ г, если он локализован на расстоянии 20,3 сМ от маркера (согласно дистанции по Haldane [13]).

3.1.3. Измерение эффектов QTL

Есть два общеприменимых способа определить «эффект» QTL. Прежде всего это, как указано выше, относится к различиям эффектов между аллелями, которые могут быть определены как аддитивный эффект (a или a') или как эффект двойной замены ($2a$ или $2a'$). Например, группой немецких ученых установлено [107], что в скрещивании между сортами гороха Erygel и 661 маркер, расположенный на группе сцепления b , имел место кажущийся эффект соответствующий 14 см высоты растения, что отличалось от значений размеров особей M_1M_1 и M_2M_2 по данному маркеру. Схожим образом, на хромосоме 4 найден маркер с эффектом двойного замещения равный 3,4 (измерения проводили по шкале от 0 до 5) по устойчивости к *Ascochyta blight* [107]. Аддитивный эффект или эффект двойного замещения иногда выражают в единицах стандартного отклонения: a/σ_T или $2a/\sigma_T$, где σ — фенотипическое стандартное отклонение популяции.

С другой стороны, каждый сегрегирующий QTL привносит свою часть в проявление общей фенотипической изменчивости. Часть фенотипической вариабельности, определяемой маркером (например, QTL, сцепленным с ним), обозначают как R^2 и устанавливают вычислением соотношения суммы квадратов к общей сумме квадратов, $R^2 = SS_M / (SS_M + SS_E)$, где SS_M — сумма квадратов, а SS_E — ошибка суммы квадратов. Например, $R^2 = 0,15$ означает, что 15% фенотипической изменчивости является следствием расщепления QTL. Если маркер не строго сцеплен с QTL, уменьшение различий между средними уменьшает сумму квадратов и таким образом уменьшает значение R^2 . Сумма значений R^2 различных независимых QTL каждого отдельного признака

не может превышать 1 (100% наблюдаемой изменчивости), но на практике достижение этого значения практически невозможно. Это объясняется несколькими причинами. Во-первых, QTL со слабыми эффектами, потенциально многочисленными, не могут быть определены из-за ограниченной силы тестов. Во-вторых, причиной части общей изменчивости являются эффекты воздействия окружающей среды (наследование признака никогда не равно 1). И, наконец, в-третьих, часть изменчивости может быть результатом эпистаза между QTL (см. раздел *Эпистаз* ниже).

Эти два способа оценки эффекта действия QTL, по понятным причинам, не зависят от того, каким методом анализа вариантов их вычисляли в F_2 , и они остаются адекватными в независимости от того, какой метод для их вычисления был использован.

3.1.4. Ограничения определения QTL по одиночным маркерам

Основное достоинство определения QTL по одиночным маркерам — это простота. Кроме того, для статистического анализа больших серий может быть использовано любое компьютерное программное обеспечение. Такого рода анализы достаточно устойчивы к отклонениям от норм и насыщенная генетическая карта в данном случае не нужна. Более того, рассматриваемые модели могут быть легко изменены, например, введением переменной в качестве кофактора или принятием во внимание комплексного расчета. Однако если плотность маркеров низка, то в оценке действительной позиции и эффектов QTL этот метод не подходит. В этой связи уже сейчас разработаны другие более эффективные методы, которые позволяют картировать QTL, находящиеся в интервале между двумя маркерами и точно оценивать их эффекты в данной позиции.

3.2. Определение QTL по двум или более маркерам

3.2.1. Методы интервального картирования

Методы интервального картирования основаны на гипотезе, что в интервале между сцепленными друг с другом маркерами G и D , с уровнем рекомбинации r , существует один QTL. Генотипические значения каждого из возможных генотипов в двух локусах G и D (например, в F_2

их девять) могут быть выражены как функция значений a и d данного QTL и уровнями рекомбинации между QTL и фланкирующих его маркеров. При решении такой системы уравнений устанавливают уровни рекомбинации между QTL и его маркерами, а также значения a и d [202]. Другой популярный метод выявления наличия QTL в интервале между двумя маркерами основывается на определении, так называемого, LOD (logarithm of odds) score — оценке десятичного логарифма шансов. Для каждой позиции хромосомы (например, с шагом 2 сМ) вычисляется десятичный логарифм отношения правдоподобия:

$$\text{LOD} = \log_{10} \frac{V(a_1, d_1)}{V(a_0, d_0)}, \quad (30)$$

где $V(a_1, d_1)$ значение функции правдоподобия в случае присутствия QTL, где оценки параметров будут равны a_1 и d_1 , а $V(a_0, d_0)$ значение функции правдоподобия в случае отсутствия QTL (нуль-гипотеза), что бывает, когда $a_0 = 0$ и $d_0 = 0$ [211]. Например, $\text{LOD} = 2$ означает, что гипотеза, альтернативная нулевой, является в $10^2 = 100$ раз более вероятной, $\text{LOD} = 3$ — в $10^3 = 1000$ раз, и так далее. В большинстве исследований пороговым значением принято считать $\text{LOD} = 3$. Таким образом, кривая LOD может быть прослежена как функция позиции на группе сцепления. Максимум кривой, если она соответствует определенному пороговому пределу, указывает на наиболее вероятное местонахождение QTL. Доверительный интервал позиции QTL обычно определяют как фрагмент хромосомы соответствующий понижению максимального значения LOD на 1 единицу, что означает уменьшение отношения правдоподобия в десятки раз. Подробное описание определения интервала доверительности можно найти у [235]. Для наиболее частого количества образцов (обычно от 100 до 200 особей), и для QTL, имеющих средние эффекты, доверительные интервалы редко меньше 20 сМ.

Существует также подход, равный по своей силе с упомянутым выше, суть которого состоит в использовании тестов и установлении модели линейной регрессии [161] вместо комплексных вычислений максимума правдоподобия. Кроме того, этот метод имеет преимущества, поскольку он менее зависим от типа экспериментальных популяций и более гибок в использовании.

Сравнение эффективности выявления QTL методом индивидуальных маркеров и методом фланкирующих маркеров указывает, что для интервалов менее 20 сМ последний подход в расчетах лишь на 5% лучше, чем предыдущий. Вследствие этого, метод интервального картирования очень быстро набирает свою популярность [315].

Основное ограничение интервального картирования заключается в том, что такой подход не позволяет выявить точное определение близлежащих QTL на той же хромосоме. Таким образом, в случае, когда два QTL локализованы близко друг от друга и действуют в одном направлении проявления признака (фаза спаривания гамет), существует риск выявления локуса-«призрака» между двумя реально существующими QTL [247]. С другой стороны, если два QTL достаточно далеки друг от друга (более чем 50 сМ), то легче установить эффект и позицию второго фиксированием позиции первого [211]. Однако эта процедура эффективна только лишь в том случае, если оба QTL не локализованы на смежных интервалах, и если, как минимум один, имеет относительно сильный эффект. Альтернативный метод заключается в поиске QTL в двух размерностях посредством рассмотрения пар QTL с различными комбинациями маркерных интервалов и с помощью использования множественной линейной регрессии [161, 247]. Справедливости ради, необходимо отметить, что в случае, когда на хромосоме присутствует одиночный QTL, последний подход менее эффективен и более неповоротлив в использовании, чем классический метод интервального картирования.

3.2.2. Методы множественных маркеров

В конце прошлого столетия был предложен иной подход, основанный на регрессии аддитивных эффектов маркеров, a'_i , по количеству $(1 - 2r_i)$, где r_i уровень рекомбинации между маркером i и гипотетическим QTL [195]. При локализации QTL с линейным отношением $a'_i = a(1 - 2r_i)$, регрессия которого определяется a , существует возможность установления аддитивного эффекта QTL. Для данной группы сцепления такая регрессия действенна на уровне каждых 2 сМ. Например, r_i соответствует частотам рекомбинации между гипотетической позицией QTL и каждым из маркеров i . Наиболее вероятная позиция QTL

та, при которой остаток регрессионного уравнения наименьший. В качестве достоинства данного метода можно назвать простоту, но он и имеет и некоторые ограничения. В частности, данный подход не может быть использован для доминантных маркеров популяций F_2 .

Примерно в это же время был предложен способ, основанный на комбинации интервального картирования и регрессии по маркерам. Такой подход должен был помочь более точно определять множественные QTL [186]. На практике, в качестве кофакторов в данном случае используют привлечение других присутствующих в геноме, представленных своими маркерами QTL, что уменьшает часть остаточной изменчивости индуцируемой их сегрегацией. Выбор маркеров, которые должны быть включены в рассматриваемую гипотетическую схему, является важным этапом. Так, например, для каждой хромосомы наиболее достоверные маркеры регрессионной модели было предложено отбирать последовательно друг за другом [187]. Можно также провести предварительный поиск QTL простым интервальным картированием и затем выбрать маркеры, наиболее близко прилежащие к выявленным QTL [419]. Новые и усовершенствованные методы идентификации и локализации QTL, использующие кофакторы, появляются быстро и, благодаря этому, у исследователей появляется новый инструмент, позволяющий повышать точность оценки эффектов QTL и локализации их позиции.

Кроме того, был предложен подход линейной модели, в котором посредством нескольких сцепленных QTL с помощью кодоминантных маркеров могут быть определены хромосомные эффекты [321]. Этот метод представляет интерес в случае, если ассоциированные QTL имеют очень мягкие индивидуальные эффекты или противоположные знаки, но он не подходит для определения отдельных, неассоциированных QTL.

3.3. Факторы, влияющие на определение QTL

Хотя принцип определения QTL сравнительно прост, существует несколько факторов, которые влияют на результаты исследований, и которые должны быть учтены для оптимизации постановки экспериментов.

3.3.1. *Определение силы и изменчивости внутри классов*

В статистике ошибка I типа (α) есть вероятность наблюдения эффекта фактора, когда его нет, т.е. наблюдать эффект случайным образом. Применительно к определению QTL, это соответствует вероятности определения «фальшивого» QTL. Ошибка II типа (β) соответствует вероятности не наблюдать эффект фактора, когда этот эффект есть. В случае определения QTL данный вариант соответствует вероятности того, что тест не выявит QTL, когда этот локус есть. Сила теста, $1 - \beta$, есть вероятность принятия верного решения. В нашем случае — это определение реального QTL. Таким образом, при проведении исследований по идентификации QTL необходимо находить компромиссное решение для ограничения числа фальшивых позитивов без особо сильного уменьшения силы теста. Вдобавок к этому, выбор α и β всегда будет зависеть от целей эксперимента. Так, например, если требуется провести предварительное установление «вероятных» позиций QTL может быть допущена высокая ошибка I типа α (около 0,05 или 0,1), а если планируется проведение клонирования QTL, то тогда должно быть взято значительно более низкое значение α . Для данного количества образцов сила определения QTL зависит, с одной стороны, от аддитивного эффекта QTL (очень маленькие различия эффектов между аллелями не могут быть найдены достоверно) и, с другой стороны, от изменчивости внутри генотипических классов или внутриклассовая вариабельность. Это изменение зависит от нескольких факторов:

- Во-первых, от эффектов воздействия окружающей среды (контроль изменений, возникающих под воздействием окружающей среды, увеличивает силу теста).
- Во-вторых, от других сегрегирующих QTL в геноме. Действительно, особи генотипического класса данного QTL не фиксированы по другим распределенным по остатку генома QTL и, таким образом, могут представлять различные фенотипические значения (проявления признака), в соответствии с их генотипом в других QTL. Эпистатические эффекты могут лишь ухудшать или отягчать этот феномен. Для того чтобы улучшить силу детекции локусов QTL с мягким эффектом проявления, находящихся в генетическом окружении, где есть QTL с сильным эффектом, может быть использована двустадийная модель

анализа вариантов, в которой одним из факторов должен быть маркер QTL с сильным эффектом. При обобщении этого принципа были предложены описанные выше методы множественных маркеров [187, 419]. Кроме того, экспериментальное решение данной проблемы кроется в работе с популяциями, в которых QTL с сильным эффектом фиксированы посредством самоопыления (см. подсекцию *Популяции в которых генетическое окружение частично фиксировано*).

- В-третьих, от дистанции между маркером и QTL. Как можно видеть, рекомбинация приводит к присутствию обоих аллелей QTL в каждом классе генотипов у маркера. Это увеличивает внутриклассовое расхождение, которое связано с уменьшением различий между средними, что, в свою очередь, редуцирует силу теста при увеличении r . Если при проведении анализов выявлен слабый эффект действия QTL, то нельзя сделать однозначное заключение (как минимум с помощью методов индивидуальных маркеров), что это: QTL с сильным эффектом, но находящийся поодаль от маркера, или QTL со слабым эффектом, но тесно сцепленный с маркером. Тем не менее, тот факт, что рекомбинация влияет на формы внутриклассового распределения, может помочь в различении данных двух QTL [27, 195, 403].

3.3.2. Размер популяции и плотность маркеров

В начале 1990-х годов было показано [93], что сила определения QTL и точность расположения (локализация) QTL в большей степени зависит от размера популяции, чем от плотности маркеров. Это значит, что, если средняя плотность маркеров достигла 20-25 сМ, любые дополнительные значения должны быть получены за счет анализа большего числа особей, но не за счет увеличения количества маркеров. Результаты исследования влияния размера изучаемой популяции на силу выявления QTL, как функции их эффекта [74], показывают, что QTL с большим эффектом будет обнаружен с достаточно высокой вероятностью и практически независимо от численности популяции, но для определения QTL со средним эффектом ($R^2 = 5\%$) необходимо использовать большее число особей.

Следует также заметить, что для данного общего числа особей лучше увеличить число генотипов популяции, чем увеличить число ре-

пликаций на генотип. Одна репликация на большом числе генотипов позволяет достичь максимума силы. Дополнительные данные по изучению силы теста, как функции размера и типа изучаемых популяций, были опубликованы ранее [201].

3.3.3. Выбор статистических рисков

Во многих экспериментальных исследованиях ошибка I типа, выбранная для анализа вариантов по индивидуальным маркерам, равна $\alpha = 0,05$ или $\alpha = 0,01$. Поскольку, как правило, проводится большое число анализов, такие высокие значения ошибки I типа неизбежно приводят к выявлению фальшивых QTL (100 независимых тестов с $\alpha = 0,05$ продуцируют в среднем пять фальшивых позитивов). Даже если маркеры не независимы, должны выбираться низкие уровни α . Выбор должен опираться на число маркеров, при этом необходимо помнить, что подобного рода выбор увеличивает ошибку II типа (т.е. уменьшает силу теста). Схожим образом, при интервальном картировании, *всеобщая* ошибка I типа (при $\alpha = 0,05$ для целого генома) придает довольно высокое значение LOD на интервал, размер которого, в свою очередь, зависит от плотности маркеров и генетической длины генома. Так, например, для генома в 1000 сМ с маркерами через каждые 15 сМ была предложена пороговая величина LOD равная 2,4 [211]. Вслед за этим появился метод [314], с помощью которого можно было производить аналитическое вычисление приближенной пороговой величины LOD для каждой группы сцепления. Пороговая величина вычислялась как функция длины и маркерной плотности данной группы сцепления. Одновременно были предложены тесты преобразования [81], основанные на случайном переборе (resampling) данных для оценки эмпирической и устойчивой пороговой величины. И, наконец, необходимо заметить, что уменьшение числа анализируемых особей уменьшает силу проводимых тестов, т.е. настоящие QTL будут «утрачены», особенно те, которые производят слабый эффект, а выявленные QTL не будут достоверными. Риск, что установленный QTL является фальшивым позитивом, равен выбранной исследователем ошибке I типа α , и не зависит от размера популяции.

3.3.4. Отклонение от нормальности

Метод интервального картирования, основанный на LOD score, применяемый для признаков с непрерывной изменчивостью, дает нормальное распределение математических остатков. Для признаков, измеряемых по шкале (часто градуированной от 0 до 5), например, устойчивости к болезням и/или признаков с высоко ассиметричным распределением, при использовании данного метода существует риск получить отклоняющиеся от нормы результаты. В таком случае можно применять непараметрические тесты, которые не требуют выдвижения предположения о распределении данных [208] или специфичные методы для признаков в классах [160]. Тем не менее, несмотря на высказанные опасения, показано [313], что если изучаемая популяция достаточно велика, то тест интервального картирования, основанный на регрессии, настолько же эффективен, как и непараметрические тесты, и поэтому он достоверен в случае большого диапазона распределений.

3.4. Достоинства и недостатки популяций, наиболее часто используемых для определения QTL

3.4.1. Сравнение популяций F_2 , беккроссов, удвоенных гаплоидов и рекомбинантных инбредных линий

При использовании популяций беккроссных, удвоенных гаплоидов или рекомбинантных инбредных линий, в отличие от популяций F_2 , как правило, сравнивают только значения двух средних по данному маркеру. Такие расчеты могут быть сделаны посредством критерия Стьюдента (t -тест) или более сложным методом. Но не все эффекты QTL возможно оценить с помощью этих популяций. Для удвоенных гаплоидов и рекомбинантных инбредных линий, может быть установлена a , но не доминирование. Для беккроссов, где гомозиготный генотип сравнивается с гетерозиготным, выявленный эффект ($m_{12} - m_{11}$) или ($m_{12} - m_{22}$), будет равен $\pm a$ или ($\pm a'$) только в рамках предположения об отсутствии доминирования. В противном случае, найденное значение, поскольку оно равно $\pm (a + d)$ или $\pm (a - d)$, будет зависеть от двух параметров, причем без какой-либо возможности оценить их отдельно друг от друга (за исключением варианта, где используют реципрокные

беккроссы). При полном доминировании, где $|d| = |a|$, QTL может быть определен, только если доминантный аллель не является фиксированным аллелем рекуррентного родителя. Для рекомбинантных инбредных линий рекомбинация между маркером и QTL будет уменьшать установленный аддитивный эффект быстрее, чем в случае других типов популяций. Так, для *высоко* рекомбинантных инбредных линий, т.е. линий при создании которых, для увеличения уровня рекомбинации, между потомством F_2 и поколениями самоопыления случайным образом «вставлены» несколько поколений, можно показать, что $a' = a(1 - 2R)$, где $R = 2r/(1 + 2r)$ [13]. Таким образом, для фиксированных r и a , отношение аддитивных эффектов установленных у маркера между рекомбинантными инбредными линиями и популяциями F_2 будет $1/(1 + 2r)$. Например, если $r = 0,1$ то аддитивный эффект у маркера для рекомбинантных инбредных линий будет оставлять 0,8 уровня других популяций.

Если примерно половина особей популяции F_2 являются гетерозиготами по данному локусу, то в этом случае они не могут быть информативными для установления a . Таким образом, при равном размере популяции, сила тестов для двойных гаплоидов или рекомбинантных инбредных линий должна быть однозначно превосходящей. Однако этот прирост частично (но только частично) компенсируется тем фактом, что внутриклассовая вариация выше для двух последних типов популяций из-за отсутствия гетерозигот среди других QTL. Для рекомбинантных инбредных линий аддитивный эффект с определенным r редуцируется быстрее, чем уменьшение силы тестов. В то же время, по несколько особей каждого генотипа все же могут быть проанализированы, как, например, это делается в случае удвоенных гаплоидов.

Если есть доминирование в маркерном локусе, смешение особей гомозиготных по одному из аллелей с гетерозиготами будет уменьшать силу теста для выявленных QTL. Кроме того, такое смешение будет мешать установлению генетических параметров, включая определение доминирования у данного QTL. И, наконец, если фенотипическое значение особей F_2 оценено, как это обычно делается, по среднему их потомков F_3 , то найденная степень доминирования будет равна половине

реальной степени доминирования, поскольку потомство гетерозиготного генотипа F_2 , как известно, включает в себе только 50% гетерозигот.

3.4.2. Потомство, полученное от гетерозиготных родителей

Если гомозиготные родительские линии недоступны (у строго аллогамных видов и видов с длинным периодом генерации, как, например, у древесных), то определение QTL затруднено, поскольку родители, в данном случае, могут различаться более чем по двум аллелям и поэтому фаза (спаривания или расхождения гамет) для генетического сцепления маркер-QTL может отличаться от одной семьи к другой. И, тем не менее, различающееся друг от друга потомство все же может быть использовано. Так, например, если известны генотипы маркерных локусов прародителей («бабушек» и «дедушек»), то это может облегчить детекцию QTL посредством использования фаз ассоциации между смежными маркерами, которые, в свою очередь, тоже должны быть идентифицированы. Такой тип популяции, содержащей три различных поколения (прародители, родители и потомство), был привлечен для локализации QTL плотности древесины у тополя [49]. У эвкалипта, напротив, для изучения возможности вегетативного размножения использовали информацию о двух поколениях семей потомков одних и тех же родителей [158]. Более подробно о решении проблем определения QTL в популяциях F_1 , полученных от нефиксированных родителей, описано в посвященных этому вопросу исследованиях [219].

3.4.3. Популяции с частично фиксированным генетическим окружением

В данном генетическом окружении степень силы определения отдельных (единичных или специфичных) QTL всегда ниже по отношению к силе многочисленных и/или тех QTL, которые имеют более сильный эффект. Популяции, используемые чаще всего, как правило, содержат особи, геномы которых (статистически) включают в себя фрагменты обоих родительских геномов в равных частях. Такая ситуация предпочтительна для определения QTL со слабым эффектом, поскольку она, как минимум, максимизирует число QTL в расщеплении, а также возможные эпистатические взаимодействия, и, вследствие этого, приводит

к большой внутриклассовой вариации. По этой причине и во избежание возможных проблем при проведении идентификации QTL у томата были созданы картирующие интрогрессивные линии [123]. Для этого авторы последовательными беккроссами и отбором с помощью RFLP маркеров получили примерно 50 линий культивируемого вида *Lycopersicon esculentum* геном, каждый из которых содержал хромосомный сегмент (33 сМ в среднем) дикого вида *L. pennellii*. Таким образом, полученная популяция представляла собой своеобразную разновидность «геномной библиотеки» *L. pennellii*, находящейся в геноме *L. esculentum*. Для поиска и идентификации QTL полученные линии сравнили по интродуцированным фрагментам с родительской культивируемой формой. Благодаря фиксации остатка генома, сравнение удалось провести даже более эффективно, чем в случае классического скрещивания. Так, например, в исследованной популяции удалось определить 23 QTL признака содержания сухого вещества плодов и 19 массы плодов [124], в то время как для популяции рекомбинантных инбредных линий, полученных в аналогичном скрещивании [154], для тех же признаков выявлены лишь 7 и 13 QTL, соответственно.

Полученный материал, в свою очередь, послужил отправной точкой для экспериментов по точному картированию. С этой целью провели скрещивание между интродуцированной и обычной культивируемой линиями и получили популяцию F₂. В полученной популяции F₂ особи, рекомбинантные по интродуцированному фрагменту, были самоопылены и в результате выделены фиксированные линии для различных субгрупп исходного интродуцированного фрагмента, которые и использовали для точного картирования. Например, на длинном плече хромосомы 2, где был найден QTL признака массы плода, на фрагменте в 55 сМ авторы выявили три QTL, в районах 3,2 сМ, 3,7 сМ и 14,1 сМ, соответственно [124].

Другой способ получить частично фиксированное генетическое окружение заключается в поиске QTL в так называемых «усовершенствованных» или «продвинутых» беккроссах (BC₂, BC₃, BC₄) [373]. Такая стратегия с успехом использована для скрининга позитивных аллелей из диких родственными видами, что позволило сэкономить на получении нескольких поколений, улучшающих исследуемые линии, поскольку

ку определение QTL и отбор генотипов делался одновременно в одном и том же поколении (см. также секцию *Трансгрессия*).

3.4.4. Селективное генотипирование

Большая сила теста, а также значительная экономия временных и финансовых затрат, могут быть достигнуты за счет молекулярного генотипирования исключительно тех особей, которые проявляют крайние значения изучаемого признака (так называемое селективное генотипирование) [211]. Такие особи действительно имеют высокий шанс аккумуляции (накопления) либо множества благоприятных, либо множества неблагоприятных аллелей для различных QTL изучаемого признака, причем таким образом, что для каждого маркера сцепленного с QTL, одновременно будут существовать значительные различия между средними и меньшие вариации для внутриклассовой генетической изменчивости. Поскольку фенотипические данные обычно не так трудно получить, как данные по молекулярному картированию, то при фенотипировании можно работать с большими популяциями и, таким образом, иметь высококонтрастные крайние классы. Но этот подход бесполезен, если необходимо измерить несколько признаков, поскольку вряд ли генетически одинаковые особи окажутся в крайних классах (за исключением случая, когда существуют сильные корреляции между изучаемыми признаками). С другой стороны, аддитивные эффекты могут быть переоценены, поскольку в данном случае генотипы различных QTL не будут более независимыми: особи с высоким значением фенотипического признака зафиксируют благоприятные аллели в большинстве из своих QTL, а часть особей с меньшим фенотипическим значением признака зафиксируют неблагоприятные аллели [91].

В подобного рода случаях для признаков с высокой наследственностью и для QTL с большими эффектами может быть использован совокупный сегрегационный анализ (bulked segregant analysis — BSA), который, в свое время, был предложен для картирования основных (менделевских) генов (см. Главу IV) [68, 396]. Найденные маркеры, коррелирующие с изучаемым признаком, также могут соответствовать одному или нескольким несцепленным с ними QTL. Поэтому подобного рода QTL необходимо проверять картированием [379]. Так, например, с

помощью метода BSA было найдено [365] несколько RAPD маркеров, очень тесно сцепленных с QTL, определяющих признак содержания олеиновой кислоты у рапса. Другие авторы [238, 239], анализируя крайние фенотипические классы популяции F_2 сои, содержащей 280 рекомбинантных инбредных линий, полученных от 69 особей, применили этот же подход для проверки надежности QTL компонентов урожая. Для этих целей они гибридизовали RFLP пробы, которые позволяли выявлять QTL в F_2 , со смесью ДНК особей крайних фенотипов и, таким образом, смогли подтвердить первичное сцепление. Вслед за этим другими авторами [92] показано, что данный подход можно сделать более эффективным, если аллельные частоты в каждой группе выразить количественно.

3.5. Генетическая и молекулярная основа вариации количественных признаков

3.5.1. QTL (почти) всегда определяемы

С того времени, когда стало возможным проведение картирования QTL, несколько независимых исследовательских групп показали, что даже с популяциями умеренных размеров (иногда чуть больше 50 особей), некоторые QTL практически всегда определимы почти для всех типов признаков [194, 367]. Это значит, что во многих случаях, большая часть изменчивости может быть объяснена сегрегацией нескольких QTL с основными (менделевскими) эффектами. Случаи, когда находят одиночные QTL, объясняющие 20% изменчивости или даже больше, нередки. А поскольку R^2 в целом учитывает и вариацию окружающей среды, то генетическая изменчивость, приписываемая QTL, обычно больше. Данные полученные на кукурузе, злаках, капусте, томате и других растениях, у которых было картировано множество QTL, указывают на то, что эффекты QTL, определяемые их R^2 , графически распределяются подобно с L-образной кривой, посредством нескольких QTL, имеющих сильный или очень сильный эффект, и большего количества QTL, имеющих средний или слабый эффект [194]. Позднее было показано [46, 47], что биохимические, генетические и статистические факторы вполне могут объяснить такое распределение.

Для популяций с наиболее частым размером (от 60 до 400 особей), и в зависимости от изучаемых признаков, обычно определяют от одного до десяти QTL [194, 367]. Эти числа дают минимально возможную оценку количеству QTL в исследуемых сегрегирующих популяциях по нескольким причинам:

- некоторые QTL имеют эффект действия ниже определяемой пороговой величины. Когда изучают большие популяции, QTL со слабым эффектом можно определить вполне эффективно. Так, например, для популяции F_2 размером в 1700 особей было определено несколько QTL, объясняющих менее чем 1% общей фенотипической изменчивости [117].
- благодаря наличию нескольких сцепленных QTL может наблюдаться эффект хромосомного сегмента.
- если два QTL со сравнимыми эффектами тесно сцеплены, то QTL будет выявлен только при наличии фазы спаривания гамет, т.е. если положительные аллели в двух локусах получены от одного и того же родителя.

Более того, как это обычно бывает в формальной генетике, мономорфные локусы в исследуемой популяции не могут быть обнаружены. Поэтому серии QTL, ответственные за изменчивость признака, как правило, не соответствуют серии генов, которые должны быть функциональны при проявлении признака, имеющего явно выраженный характер. Если признак рассматривает, например, концентрацию конечного продукта линейного метаболического пути, то *все* гены пути должны быть функциональны для данного продукта, который в свою очередь должен присутствовать, но только *полиморфная часть этой группы* в изучаемой популяции может влиять на изменчивость этого признака. Более того, порой бывает, что серии QTL для данного признака в двух различных популяциях только частично перекрываются [31, 370]. Наиболее подходящее объяснение данного явления, в добавление к возможным эффектам эпистаза (см. секцию *Эпистаз* ниже), по-видимому, заключается в том, что это не одни и те же локусы, которые в обеих популяциях являются полиморфными.

3.5.2. Эффекты доминирования

Даже если тесты достоверности проводить достаточно редко, то, как представляется, и в этом случае эффекты доминирования должны обнаруживаться довольно часто. Ряд авторов считают, что в некоторых исследованиях, значение $|d/a|$ однозначно падает ниже 1, если существует сверхдоминирование для определяемого QTL [360]. Однако в другой интерпретации предполагается, что исследователи имеют дело не с одним, а с двумя сцепленными QTL в фазе расхождения гамет, с эффектом доминирования для каждого из них. Таким образом, двойные гетерозиготы $+/-+$ будут иметь более высокое фенотипическое значение, чем каждая из двойных гомозигот $+/--$ и $-/++$. Например, благодаря экспериментам по тонкому картированию с использованием практически изогенных линий, происхождение установленного различного сверхдоминирования в районе 37 сМ, определяющего содержание растворимых веществ в плодах томата, может быть интерпретировано как результат сцепления между двумя доминантными QTL в фазе расхождения гамет [124]. Если тонкое картирование не позволяет сделать выбор между двумя выдвигаемыми предположениями, то тогда для установления истины может понадобиться проведение позиционного клонирования локуса или локусов QTL.

3.5.3. Трансгрессия

Явление *трансгрессии* наблюдается тогда, когда особи потомства превышают изменчивость родительских форм по данному признаку. Этот феномен может быть генетически проанализирован посредством картирования QTL. Трансгрессию легко объяснить на основе простой модели с двумя независимыми локусами. Если родители имеют, соответственно, генотипы $+ -/+ -$ и $- +/+ +$, то их потомство будет содержать особи $++/++$ и $--/--$, генотипические значения которых будут превосходить значения родительских форм. Обобщая и распространяя эту модель на несколько локусов можно ожидать, что чем больше родители различаются по рассматриваемым признакам, тем меньше трансгрессивных особей среди потомков, и тем больше родители будут генетически гомогенными в различных определяемых QTL (т.е. можно ожидать, что если «благоприятные» аллели зафиксированы для большинства

QTL, то другие аллели будут «неблагоприятными»). В исследованиях многих признаков на томате удалось подтвердить этот прогноз [99]. В противоположность этому трансгрессии и QTL можно обнаружить также в популяциях, чьи родители не проявляют никаких различий по изучаемым признакам, просто потому, что данное значение признака можно получить у различных генотипов для QTL-локусов признака (это можно ожидать исключительно с неродственными особями).

Интересно заметить, что даже когда в качестве родителей для получения картирующей популяции, выбирают высоко контрастные особи, нередко находятся фракции QTL, показывающие эффект противоположный тому, который можно ожидать исходя из значений признака, наблюдаемого у родительских форм. При установлении QTL контролирующего устойчивость к *Ascochyta blight* у гороха [107], авторы обнаружили, что устойчивый родительский сорт имел фиксированный аллель чувствительности для одного из трех выявленных QTL (того, который проявлял самый слабый эффект). Исходя из этого, устойчивость данной линии может быть улучшена интрогрессией устойчивого аллеля, найденного у чувствительного родителя. Этот тип интрогрессии был эффективно применен на томате при межвидовых скрещиваниях между культурным видом *Lycopersicon esculentum* и родственными ему дикими видами *L. hirsutum* и *L. pimpinellifolium*. Анализ улучшающих беккроссов (BC3) позволил картировать QTL агрономически важные признаки [33]. Последующее создание практически изогенных линий, несущих интрогрессии для желаемых диких QTL-аллелей, позволило выявить наблюдаемое фенотипическое улучшение для большинства исследованных признаков [34].

3.5.4. Эпистаз

Попытки поиска эпистаза между QTL предпринимались неоднократно. Посредством n -кратного дисперсионного анализа существующих взаимодействий, в принципе, можно проследить взаимодействия любого порядка между локусами QTL. Однако на практике этот подход ограничен взаимодействиями между парами QTL, поскольку увеличение числа локусов быстро редуцирует число особей на генотипический класс. При взаимодействиях второго порядка (между тремя QTL) в F_2

образуется 27 генотипических классов, что приводит к очень низкому числу (несколько штук) особей на класс и очень низкой силе теста. Более того, даже принимая в расчет взаимодействия первого порядка, двумерный анализ вариантов необходимо проводить многократно. Так для N локусов, необходимо провести $N(N - 1)/2$ тестов, что в свою очередь поднимает проблему выбора ошибки I типа. Несмотря на эти трудности, эпистатические взаимодействия QTL были установлены в различных исследованиях и на различных культурах. В некоторых случаях достоверные взаимодействия между QTL наблюдались, например, для устойчивости к стрессу, вызванному низкими дозами фосфора [317], и архитектурному построению у кукурузы [108], солеустойчивости у томата [50], устойчивости к холоду у ячменя [282], компонентам устойчивости к *Phytophthora capsici* у перца [216] и даже молекулярным биохимическим признакам, таким как концентрация белка [90].

Тем не менее, целый ряд авторов подчеркивали, что эпистатическое взаимодействие QTL происходит нечасто [286, 332, 360, 367 и др.]. Эти наблюдения, со всей очевидностью, противоречат как многочисленным результатам классической количественной генетики, которая часто обнаруживает варианты эпистаза, так и теории метаболических процессов, в которой эпистаз определяется как внутреннее присущее качество или свойство [192]. Как было установлено, основная причина такого противоречия, несомненно, является статистической: используемые популяции часто имеют ограниченный размер и, вследствие этого, анализ вариантов, применяемый для определения взаимодействий, имеет ограниченную силу. Действенным способом увеличения силы анализа вариантов является элиминация «фонового шума» или «шума генетического окружения», возникающего из-за других QTL. При этом в качестве родителей изучаемых популяций используют практически изогенные линии с интересующими исследователя QTL. Например, израильские исследователи, используя полудиаллельную схему скрещивания, скрестили 10 интрогрессивных линий томата, содержащих в геноме культурного томата *Lycopersicon esculentum* (хозяйское генетическое окружение) специфический фрагмент генома *L. pennellii* [124]. Благодаря этому авторы смогли показать и доказать многочисленные достоверные эпистатические взаимодействия QTL. Схожим образом, проведена про-

верка эпистатического взаимодействия между QTL вовлеченных в эволюционное происхождение кукурузы из ее предшественника (теосинте). При этом два аллельных QTL теосинте, по которым в предварительных экспериментах установлены значительные морфологические различия между теосинте и кукурузой, были интрогрессированы в генетическое окружение (геном) кукурузы. Растения каждого из четырех двулокусных гомозиготных классов двух QTL авторы сравнили между собой. В результате, как на морфологическом уровне, так и на уровне мРНК одного из генов (*tb1*), вовлеченных в эволюцию кукурузы, наблюдались эффекты сильного взаимодействия между QTL [228].

С другой стороны, отсутствие эпистаза между QTL может наблюдаться не потому, что QTL не проявляют эпистатического взаимодействия с другими отдельно взятыми QTL, а потому, что их эффект, по видимому, независим от хозяйского генетического окружения. Как показано ранее [108], эффекты двух QTL одомашнивания кукурузы являются более слабыми, когда они сегрегируют в генетическом окружении «теосинте», нежели в F₂ генетического окружения «кукуруза × теосинте». Схожим образом при изучении трех популяций рекомбинантных инбредных линий кукурузы, полученных в результате трех возможных скрещиваний между тремя линиями, показано [72] достоверное влияние хозяйского генетического окружения на экспрессию нескольких QTL признаков времени цветения и урожая.

3.5.5. Генетический анализ корреляций между признаками

Картирование QTL дает уникальный подход для изучения генетических основ корреляций между признаками. Когда признаки коррелируют, то можно ожидать, что как минимум некоторые из их QTL окажутся общими (или, как минимум, генетически сцепленными). Было отмечено, что это действительно и может быть с пользой применено на практике. Например, в уже упоминавшейся популяции F₂ гороха, для самых сильных QTL высоко коррелирующих друг с другом признаков высоты растения, времени цветения и числа узелков, был найден один общий район [107]. Такое же явление наблюдалось у пшеницы [19] и ячменя для признаков высоты растения и времени цветения [39]. Для последнего вида дополнительно установлены два специфических QTL,

определяющих высоту растения и отдельно семь QTL признака времени цветения. У кукурузы в условиях стресса недостатка питания по фосфору, где уровни роста проростков и корней молодых побегов коррелируют друг с другом, поиск QTL показал, что четыре из шести идентифицированных хромосомных районов были вовлечены в контроль обоих признаков, каждый из оставшихся двух контролировали только по одному из признаков [317]. Возможность различия общих QTL от специфичных может быть полезно при селективном отборе или скрининге селекционных образцов. Если два желаемых признака отрицательно коррелируют (обратно пропорциональны), идентификация специфичных QTL позволяет оценить пределы, до которых каждый из признаков может быть улучшен без негативного последствия на другой признак. В этой связи был предложен [204, 205] статистический тест, в котором, используя информацию коррелирующих признаков, можно локализовать QTL, контролирующие одновременно несколько признаков. Авторы показали, что такой подход эффективнее увеличивает силу определения QTL, чем при использовании методологии поиска «признак за признаком».

В случае явной ко-локализации QTL, контролирующей различные признаки, нет прямого способа, позволяющего сделать выбор между присутствием одиночного QTL с плейотропным эффектом или двумя сцепленными QTL. В последнем варианте ключ к решению может дать тонкое картирование. Тем не менее, даже в этом варианте нет соразмерности частей доказательства: невозможность нахождения двух QTL не доказывает, что существует только один QTL-локус, но может указывать, что поблизости есть сцепление. Дополнительно, помимо применения методов молекулярной биологии, можно полагаться и на не прямые аргументы:

- если известно, что мутация одновременно влияет на оба признака, то это значит, что возможны плейотропные связи. Если в дополнение к этому мутация и QTL ко-локализованы, то мутантный локус представляет собой подходящий QTL-кандидат.
- если оба признака одновременно имеют долю не в одном, а в двух или более очевидно общих QTL, то тем самым вероятность плейотропии подкрепляется. Действительно, если совпадение позиции не

невероятно из-за размера доверительных интервалов позиций QTL, то множественные совпадения весьма вероятны и должны быть вследствие случайности событий.

- могут выдвигаться и анализироваться предположения физиологического и/или морфологического порядка. Если, например, найденный у потомства кукурузы общий QTL признака содержания абсцизиевой кислоты и признака устьичной проводимости подчиняется воздействию водного стресса, то возможны причинные связи или взаимозависимость [303](см. Главу V).
- если есть плейотропия, то наследование ассоциации маркер-признак на протяжении нескольких поколений будет стабильным, до тех пор, пока в случае сцепления из-за рекомбинации оно не спадет. В то же время, в случае тесного сцепления наблюдаемый спад может быть очень медленным.
- и, наконец, в случае плейотропии корреляция должна наблюдаться даже в отсутствии неравновесного сцепления (LD), т.е. у неродственных особей.

Если необходимо анализировать генетическую основу сложных (комплексных) признаков, фокусируясь при этом на их *составных частях*, которые, как предполагается, имеют более высокую наследственность, то картирование QTL становится особенно привлекательным. Например, признак урожая зерна у кукурузы может быть проанализирован посредством изучения серии характеристик початка [359], признак высоты растения — посредством исследования числа узелков, длины междоузлий и высоты стебля у початка [116] и т.д. Другие авторы [392] изучали признак урожая зерна у кукурузы и шесть из определяющих его компонентов (число початков на растение, средняя длина и диаметр початка, среднее число рядов, и т.д.). Из 14 идентифицированных районов, несущих QTL, девять влияли более чем на один компонент, а два контролировали четыре и пять компонентов, соответственно (с различными корреляционными значениями между признаками). Можно также попытаться оценить роль физиологических компонентов в экспрессии признаков развития или морфологии. Так, например, был исследован случай явной ко-локализации QTL [65], отвечающих за активность ферментов метаболизма углерода, и QTL роста у кукурузы. Схожим образом,

рассмотрение компонентов сложных признаков при исследовании устойчивости к заболеванию позволяет пополнить наши знания о механизмах реализации этого процесса. Например, у кукурузы инфицированной грибом *Exserohilum turcium*, вместо проведения общей фенотипической оценки признаков заболевания, отдельно изучали размер и число поражений [134]. Авторы показали, что эти два компонента находятся под независимыми генетическими контролями, поскольку их QTL (два и три, соответственно) не совпадают. Так же было показано [216], что устойчивость к *Phytophthora capsici* у перца, исследованная посредством изучения шести отдельных компонентов данного признака, соответствующих различным реакциям органов и стадиям инфекции, контролировалась несколькими QTL, некоторые из которых были специфичны к одному компоненту, другие оказались общими для нескольких.

3.5.6. Эффекты действия окружающей среды

Окружающая среда является тем фактором, который может иметь и имеет существенное влияние на эффекты проявления QTL. Неизбежный результат данного факта заключается в том, что взаимодействие генотип-среда может частично модифицировать проявление генотипов таким образом, что амплитуда эффектов QTL будет варьировать от одних условий внешней среды к другим, и даже если QTL определен в одних условиях внешней среды, то в других условиях тот же QTL может быть и не найден, и наоборот. Такое явление довольно часто встречается, даже если условия внешней среды влияют на проявление QTL самым разнообразным образом согласно признакам, ранжированию выбранных условий внешней среды и размеру популяции [19, 56, 170, 203, 285, 331, 360]. Важная концепция, на которой базируются эти исследования, заключается в том, что некоторые QTL можно выявить во всех или почти во всех исследуемых условиях внешней среды и поэтому они «универсальны», в то время как другие находят лишь в определенных условиях внешней среды и, вследствие этого, они — «специфичны». Такое подразделение необходимо брать в расчет при реализации так называемой маркерной вспомогательной селекции, причем делать это необходимо согласно типу и ранжированию условий внешней среды для которых создаются сорта.

Если давать определения в методологических терминах, то, как правило, в большинстве исследований просто ищут QTL для каждого из условий внешней среды и затем сравнивают выявленные QTL. Другие исследования используют модели анализа вариантов или множественной регрессии включая, в добавление к эффектам QTL и окружающей среды, взаимодействия «генотип – окружающая среда» [32, 188, 170, 417]. Для изучения таких случаев и картирования QTL были разработаны соответствующие компьютерные программы [204, 418]. При разработке и апробации более сложного метода [323], рассмотрены данные [170] по урожайности зерен у ячменя в 16-ти различающихся условиях окружающей среды. Авторы впервые провели принципиальный компонентный анализ взаимодействия «матрица генотипа × окружающая среда» и нашли QTL для основных компонентов этого взаимодействия. Благодаря такому подходу установлены четыре хромосомных сегмента, которые отвечали за отличия в проявлении признаков в различных условиях внешней среды.

3.5.7. Сравнительное картирование QTL

Используя данные сравнительного картирования, ряд исследователей попытались установить могут ли QTL данных признаков быть найдены в гомологичных позициях геномов родственных видов. Наиболее наглядный пример обнаружен в роде *Vigna* (семейство *Fabaceae*), где QTL признака массы семян был картирован в геноме видов *V. unguiculata* и *V. radiata* [128]. На основе результатов проведенного анализа с использованием молекулярных маркеров районы хромосом, несущие QTL с наибольшим эффектом для каждого вида (R^2 равно 36,5% и 32,5%, соответственно), проявили гомологию. Дополнительным аргументом в пользу идентичности установленных QTL стало эпистатическое взаимодействие. Идентичные QTL, с которыми выявленные основные QTL имели эпистатическое взаимодействие, также были расположены в гомологичных районах. Вдобавок к этому, применение молекулярных проб общих для *Fabaceae* позволило установить, что QTL признака массы семян у гороха [379] и у сои [250] соответствуют одному из двух ранее идентифицированных QTL *Vigna*. Компиляция QTL признаков одомашнивания у риса, кукурузы, сорго и сахарного тростника так-

же выявила совпадение локализации QTL более частое, чем можно было бы ожидать просто по случаю, особенно для признаков размера зерна, размалывания и времени цветения при коротком дне [287]. И, наконец, аналогичные выводы сделаны в отношении и трех видов рода *Lycopersicon*, включая культурный томат [154, 285].

Полученные результаты интересны не только как дополнительные примеры эволюционной идиоадаптации синтении. Примечательным фактом является то, что виды, которые дивергировали много лет тому назад (семейство *Poaceae*, например, насчитывает порядка 60 миллионов лет) и, жившие в очень разнообразных экологических условиях, до сих пор выявляют гомологичные полиморфные локусы, объясняющие изменчивость некоторых признаков. И в то же время, все это напоминает нам, что сравнение двух потомств на внутривидовом молекулярно-генетическом уровне может привести к получению частично различающихся серий QTL.

3.5.8. Некоторые характерные особенности QTL

Характеризовать QTL задача не из легких. Из-за слабых фенотипических эффектов и шума генетического окружения, производимого другими QTL, а также из-за влияния условий окружающей среды, идентификация локусов QTL приобретает комплексный и более многогранный характер. У растений на молекулярном уровне со всей определенностью были охарактеризованы всего несколько QTL. С другой стороны, совершенно очевидно, что познание природы QTL несет огромный фундаментальный и прикладной аспект. В этой связи следует помнить, что современная интерпретация словосочетания «идентификация QTL» обычно подразумевает необходимость использования одного из двух методических подходов для придания реального смыслового значения данному термину. Во-первых, если ничего не известно о физиологических и молекулярных основах признака, необходимо использовать генетические и молекулярно-биологические подходы для того, чтобы клонировать изучаемый QTL («позиционное клонирование»). Во-вторых, если установлено, что в проявление признака вовлечены гены с известной функцией, более правильным будет искать те из них, чей полимор-

физм составляет определенную часть изменчивости изучаемого признака (подход «генов-кандидатов»). Рассмотрим эти подходы.

3.5.8.1. Клонирование QTL

Для идентификации и выделения QTL методом позиционного клонирования могут быть использованы генетические подходы, описанные нами ранее [13]. Однако отсутствие дискретности, присущей менделеевским признакам, или, другими словами, наличие продолжительности вариации у количественных признаков делает задачу более сложной. Прежде всего, по статистическим причинам. Это значит, что научный план проведения экспериментов должен быть по возможности всеобъемлющ. В случае использования практически изогенных линий или метода BSA для скрининга генотипов необходимо воспользоваться маркерами, фланкирующими QTL, поскольку проведение фенотипического скрининга с целью клонирования QTL при этом обычно невозможно. По этой же причине для установленных районов QTL требуется картирование с высоким разрешением. Позиционное клонирование реально осуществимо только лишь тогда, когда QTL локализован в интервале меньше, чем одна сантиморганида и пока только у видов с маленьким размером генома. В идеале необходимо достичь дистанции между маркерами и QTL равной примерно размеру YAC или BAC клона или даже меньше. Для идентификации «неясных» QTL может понадобиться проведение трансформации с целью установления истинной функциональной активности клонированных QTL. С другой стороны, вероятнее всего, другие доказательства могут оказаться более эффективными, например те, что упомянуты в Главе IV при описании методологии клонирования основных (менделевских) генов: картирование с высоким разрешением единичных генов-кандидатов с целью нахождения ко-сегрегирующих с признаком клонов, биоинформативный анализ первичных аминокислотных и нуклеотидных последовательностей или изучение соответствующей экспрессии (см. ниже). В любом случае идентификация QTL посредством клонирования, по-видимому, должна быть возможна или, как минимум, легче для QTL с большими эффектами и/или для признаков, у которых известны мутантные формы (напри-

мер, клонированные гены карликовости очень удобны для характеристики QTL высоты растения).

Несмотря на перечисленные трудности, возникающие в реализации данного подхода, в ряде случаев удалось добиться успешного позиционного клонирования QTL. Например, у томата клонирован QTL *fw2.2* в значительной мере ответственный за существенное увеличение размера плода, произошедшее в процессе одомашнивания вида. Картирование с высоким разрешением позволило идентифицировать YAC-клон с этим QTL. При переносе методом трансгеноза космиды, выделенной из *fw2.2* района мелкоплодных диких видов и содержащей один из пяти генов, найденных в YAC (*ORFX*), в сорта с большими плодами, произошло уменьшение размера плодов до предсказанного и ожидаемого уровня [132]. В потомстве томата же, состоящем из 7000 растений F₂, с высокой точностью был картирован локус QTL, увеличивающий содержание сахара в плодах. В данном случае, при использовании практически изогенных линий для идентификации упомянутого QTL, картирование высокого разрешения позволило установить предельные значения идентифицируемого QTL в районе размером 484 п.о., включающем экзон и интрон гена инвертазы — фермента, который модулирует силу налива плодов томата [135]. У риса посредством картирования высокого разрешения был идентифицирован ген основного количественного признака, *Hdl*, контролирующий ответ на изменение фотопериода. Для этого авторы использовали 1505 сегрегантов, которые позволили им определить геномный район размером всего лишь в 12 т.п.о. Этот район стал своеобразным геном-кандидатом для *Hdl*. Последующий анализ выявил, что *Hdl* соответствует гену гомологичному *CONSTANS* у *Arabidopsis* [411]. И, наконец, с помощью схожих подходов у *Arabidopsis* клонирован QTL времени цветения [118].

3.5.8.2. Гены-кандидаты

Подход «генов-кандидатов» фактически есть ничто иное как два разных подхода, которые в чем-то похожи друг на друга (табл. 8).

Подходы идентификации генов-кандидатов

«Функциональные» гены-кандидаты	«Позиционные» гены кандидаты
Гены с известной функцией ↓	Генетическое картирование QTL ↓
Полиморфизм? ↓	Гены/мутации в районе? ↓
Представляет ли этот полиморфизм хотя бы часть изменчивости признака?	Соответствует ли хотя бы один/одна из них QTL?

3.5.8.3. «Функциональные» гены-кандидаты

Для некоторых количественных признаков физиология служит отличительной чертой исследуемых функций. Если доступны соответствующие гены, то являются ли они генами-кандидатами локусов QTL изучаемого признака или нет, зависит от их воздействия на эффекты полиморфизма (любого), проявляющегося в изменчивости признака [295].

У растений первым опубликованным примером гена с известной функцией, который соответствовал QTL, по-видимому, был ген *Srunken-2*, кодирующий одну из субъединиц АДФ-глюкозопирофосфорилазы (АДФ-ГПФ) — ключевого фермента в синтезе сахаров зерна. Американские ученые в популяции примерно из 100 семей F₃ кукурузы нашли QTL с очень сильным эффектом, определяющим содержание сахаров, и локализовали его точно в районе расположения гена АДФ-ГПФ [155]. Другим примером служит сахарозосинтетаза, ответственная за гидролиз сахаров на глюкозу и фруктозу. Французские исследователи в популяции рекомбинантных инбредных линий кукурузы обнаружили QTL, определяющие активность этого фермента, один из которых они сумели локализовать в девятой хромосоме точно в месторасположении структурного гена *Sh1* [65]. Более того, в том же районе авторы локализовали QTL признака содержания гексозы и один QTL раннего прорастания. Таким образом, полиморфизм *Sh1* может быть причиной вариации ак-

тивности фермента, который в свою очередь, объясняет присутствие QTL продуктов реакции фермента и влияет на признак роста растения.

Другим примером идентификации QTL посредством использования подхода функционального гена-кандидата является изучение признака содержания майзина (англ. *maysin*) в пестиках початка кукурузы. Майзин — это флавоноид, который ингибирует рост и развитие гусеницы американской хлопковой совки (*Helicoverpa zea* Boddie). Пути его биосинтеза хорошо известны, а гены, кодирующие основные ферменты, и большинство регуляторных факторов этого процесса клонированы. Используя в качестве маркеров геномные и кДНК пробы этих генов, группа ученых [62] провела в популяции F₂ поиск локусов QTL, определяющих содержание майзина. В результате ими выявлен ($R^2 = 58\%$) активатор транскрипции *p1*, играющий одну из основных ролей в биосинтезе майзина. Найденный QTL действовал аддитивным образом, причем с большими различиями эффектов между аллелями: средние классы Q_1Q_1 , Q_1Q_2 и Q_2Q_2 оказались, соответственно, равны 0,023%, 0,280% и 0,630% свежего веса пестиков початка. Отличительная, по сравнению со структурными генами, кодирующими ферменты, роль данного генного регулятора, может быть, простой случайностью, но это согласуется с теорией метаболического контроля [192]. Кроме того, авторы показали, что вариации в эффективности структурных генов ферментов в комплексных путях метаболизма в целом имеют умеренные последствия в отличие от вариаций регуляторных факторов. Приведенный пример также иллюстрирует важную точку зрения, заключающуюся в том, что проблема характеристики QTL не обязательно является проблемой идентификации и клонирования *новых* генов. Даже если известны все гены и их функции в модельных объектах, то вопрос об их влиянии на количественную вариацию признака все равно останется нерешенным.

3.5.8.4. «Позиционные» гены-кандидаты

Для многих признаков нет физиолого-биохимических данных служащих ориентиром при выборе генов с известной функцией. В этом случае стратегия заключается в картировании QTL настолько точно, насколько это возможно, и в поиске района QTL для генов или мутаций, которые могут оказаться хорошими кандидатами.

Особый интерес вызывают мутации, придающие ярко выраженное значение изучаемому признаку. Например, QTL-локус хромосомы 1 кукурузы, влияющий на длину стеблей у потомства, полученного при скрещивании кукурузы и теосинте, совпадает с мутацией *Teosinte branched 1 (Tb1)*, придающей кукурузе фенотип частично схожий с фенотипом теосинте, т.е. обуславливает развитие длинных латеральных стеблей. Последовательное клонирование *Tb1* с помощью генетических и молекулярных методических подходов (см. ниже) подтвердило, что локус *Tb1* соответствует QTL, для которого ранее были найдены три типа аллелей: полностью активный аллель кукурузы; аллель, названный *tb1 + teosinte*, проявляющий примерно половину активности *Tb1* теосинте и инактивированный аллель *tb1* мутантов кукурузы.

Модельные виды с малыми полностью секвенированными геномами особенно полезны в контексте подхода генов-кандидатов. Вполне иллюстративен в этом отношении пример гена *CO (CONSTANS)* резушки Тале (*Arabidopsis thaliana*). Мутант (*co*) задерживает время цветения при длинном фотопериоде, но не при коротком, и, тем самым, уменьшает реакцию на яровизацию. Зонд гена *CO* и еще 10 дополнительных проб молекулярных маркеров, сцепленных с геном *CO* в соответствующей ему группе сцепления примерно в 7,5 сМ, были картированы в популяции беккроссов *Brassica nigra* посредством RFLP. Из-за произошедших в процессе эволюции генома *B. nigra* дупликаций, группы сцепления были найдены в трех различных районах, два из которых несли QTL времени цветения, совпадающие с гомологами *CO*. Таким образом, полученные результаты обозначили гомолог *CO* арабидопсиса в качестве подходящего гена-кандидата времени цветения у *B. nigra* [209].

3.5.9. Проверка генов-кандидатов

Независимо от того, использовалась ли методология позиционного клонирования или применялась стратегия генов-кандидатов, роль выделившегося гена в изменчивости количественного признака может быть установлена посредством различных не исключающих друг друга подходов.

3.5.9.1. Секвенирование аллелей генов-кандидатов

Присутствие полиморфизма среди аллелей гена-кандидата должно быть установлено до начала изучения его функций. При этом исследованию подлежит не только кодирующий район гена, но также и его регуляторные части. Как показано для *Tb1* кукурузы (см. выше), полиморфизм может находиться и в промоторном районе гена. Секвенирование аллельных серий гена-кандидата, найденных в популяции с минимальным неравновесным сцеплением (LD), является другим способом, с помощью которого можно подтвердить роль кандидата, например, посредством поиска полиморфизма на уровне единичного полиморфизма (англ. single nucleotide polymorphism, SNP), коррелирующего с признаком. Достоверная корреляция не предполагает случайных взаимоотношений, поскольку LD на короткой дистанции может объяснить наблюдаемую корреляцию. Как минимум, она указывает, что наблюдаемый полиморфизм признака генетически коррелирует с SNP.

3.5.9.2. Определение экспрессии QTL или активности генов-кандидатов

Необходимым (но не достаточным) условием отбора кандидата является то, что уровень его экспрессии (мРНК или концентрация белка) и/или его активность (для ферментов, транспортеров, транскрипционных факторов и т.п.) должны быть вариабельными в изучаемой популяции. В дополнение к этому экспрессия или активность QTL/гена должны быть со-локализованы с QTL изучаемого фенотипического признака. Разработка высокопроизводительных методов анализа экспрессии транскриптов и белков делает этот подход более легким.

3.5.9.3. Трансформация

Трансформационный генно-инженерный перенос или трансформация фрагментов ДНК, несущих предполагаемые QTL, в модельные (и не только) объекты, способна помочь принять определенное решение относительно исследуемых QTL. Однако трансформация может выдвинуть специфичные проблемы, связанные с QTL, поскольку в целом задача сводится не к комплементированию неактивного гена активным, а к сравнению эффектов двух активных аллелей гена или QTL, что может оказаться статистически трудным. Кроме того, гомологичная

трансформация пока еще не возможна у растений, и поэтому месторасположение и/или эффекты дозы могут привести к неоднозначным результатам [12].

3.5.9.4. Количественная комплементация

Если выявляемый QTL имеет большой эффект и известен мутант, который определяет крайнее значение признака, то для изучения кандидата может быть использована генетическая комплементация. Для этого проводят скрещивания между особью гетерозиготной по QTL и особью гомозиготной по мутации или дикой/мутантной гетерозиготой. При доминировании «максимальных» аллелей, расщепление в потомстве будет наблюдаться только если QTL и локус, несущий мутацию, полностью совпадут (соответственно, 1:1 и 3:1). В противном случае будет наблюдаться комплементация, рецессивные аллели двух локусов будут замаскированы (скрыты), а популяция не проявит расщепление. Такой вариант генетического взаимодействия продемонстрирован для локуса *Tb1* [108, 109]. Мутация, придававшая кукурузе фенотип, в некоторой степени, схожий с фенотипом теосинте, оказалась одним из QTL ответственных за различия между кукурузой и теосинте.

Таблица 9

Количественная комплементация*

Кандидат, соответствует QTL			Кандидат, не соответствует QTL		
	ПИЛ _С			ПИЛ _С	
	$\frac{+}{+}$	$\frac{m}{m}$		$\frac{Q_{X+}}{Q_{X+}}$	$\frac{Q_{Xm}}{Q_{Xm}}$
ПИЛ _Q			ПИЛ _Q		
$\frac{Q_H}{Q_H}$	$\frac{+}{Q_H}$	$\frac{m}{Q_H}$	$\frac{Q_{H+}}{Q_{H+}}$	$\frac{Q_{X+}}{Q_{H+}}$	$\frac{Q_{Xm}}{Q_{H+}}$
$\frac{Q_L}{Q_L}$	$\frac{+}{Q_L}$	$\frac{m}{Q_L}$	$\frac{Q_{L+}}{Q_{L+}}$	$\frac{Q_{X+}}{Q_{L+}}$	$\frac{Q_{Xm}}{Q_{L+}}$

*ПИЛ_Q и ПИЛ_С соответствуют практически изогенным линиям для QTL и локуса-кандидата. Q_X — аллель QTL линий +/+ и m/m, в случае когда ген-кандидат не соответствует QTL.

Однако при реализации этого теста однозначный результат может быть получен лишь в редких случаях, поскольку QTL, как правило,

имеют слабые эффекты и проявляют доминирование. В таких случаях следует применить общий подход по установлению количественной комплементации, даже если сам тест и его интерпретация окажутся более сложными [127]. Рассмотрим несколько подробнее. Пусть Q_H и Q_L (H и L — высокий и низкий или максимальный и минимальный, соответственно) будут два аллеля QTL обнаруженных в потомстве. Посредством проведения беккроссных скрещиваний с помощью маркеров, получают пару практически изогенных линий (ПИЛ, англ. near isogenic lines, NIL) Q_H/Q_H и Q_L/Q_L . Одновременно необходимо получить пару ПИЛ дикого типа и мутанта по гену-кандидата, с генотипами $+/+$ и m/m , соответственно. Скрещивания $Q_H/Q_H \times +/+$, $Q_L/Q_L \times +/+$, $Q_H/Q_H \times m/m$ и $Q_L/Q_L \times m/m$ дадут представленные в таблице 9 генотипы, соответственно тому, соответствует ли кандидат району QTL (один локус) или нет (два локуса). Если кандидат соответствует QTL, то тогда различия между генотипическими значениями m/Q_H и m/Q_L , как ожидается, должны быть выше чем различия между $+/Q_H$ и $+/Q_L$ (если только аллель + полностью рецессивен по отношению к аллелям Q_H и Q_L ; если аллель + строго доминантен над Q_H и Q_L , то тогда не будет различий между $+/Q_H$ и $+/Q_L$). Если кандидат не соответствует QTL, то в этом случае не ожидается, что различия между Q_X+/Q_H+ и Q_X+/Q_L+ будут достоверно отличаться от тех, что будут получены между Q_Xm/Q_H+ и Q_Xm/Q_L+ , каким бы то ни было доминирование между аллелем Q_X и аллелями Q_H и Q_L (различия равны нулю, если Q_X строго доминирует над Q_H и Q_L). Таким образом, следует помнить, что этот тест тоже не является однозначным, поскольку выявленные достоверные различия могут быть результатом эпистатического взаимодействия между геном-кандидатом и одним или несколькими другими генами. В этом случае кандидат будет соответствовать QTL, оказывающему влияние на признак, но не изучаемому QTL локусу.

Заключение

Первые эксперименты по локализации QTL у растений берут свое начало с середины 1980-х годов. Однако уже на сегодня становится очевидным, что идентификация QTL позволит расширить наше понимание молекулярной и физиологической основы изменчивости комплексных

признаков. В селекции растений это делает проведение мониторинга самих QTL даже в большей степени привлекательным, чем поиск и идентификация их маркеров.

Накопленные на нескольких видах растений данные выявили явные тенденции в архитектонике геномов, в особенности, в частоте случаев явной ко-локализации между локусами количественных признаков и генами с количественными эффектами (например, генами, контролирующими устойчивость к патогенам, карликовости, чувствительности к фотопериоду, архитектурному построению и др.). Даже если огромная задача по идентификации и локализации QTL останется и должна будет случаем за случаем подтверждать, что QTL и гены с количественными эффектами принадлежат одним и тем же локусам, полученные результаты, также как и первые эксперименты по идентификации и характеристике QTL у растений, животных и человека, ясно подтверждают непрерывный характер спектра между дискретной и непрерывной изменчивостью. В этом контексте становится понятным, что картирование конкретных генов более предпочтительно, чем картирование анонимных маркеров, поскольку они являются источниками генов-кандидатов. Программы систематического секвенирования EST (англ. expressed sequence tags; — последовательности с невыявленной функцией, экспрессированные в исследуемом объекте целевым образом) у растений помогут ускорить получение таких карт. В особенности, выделение кДНК, специфичных к определенной стадии развития, органа или полученных под воздействием какого-либо стресса и так далее, а также выбор согласно изучаемому признаку, может установить предварительный тест. Необходимо снова хочется подчеркнуть, что проведенное полное секвенирование геномов *Arabidopsis thaliana*, риса (*Oryza sativa*) и других возделываемых растений представляет огромный и поистине неоценимый источник информации, который можно и должно использовать при решении задач, стоящих перед генетиками, селекционерами и кураторами коллекций генетических ресурсов растений, сохраняемых в Генных Банках различных стран мира.

Глава IV. КАРТИРОВАНИЕ ОСНОВНЫХ (МЕНДЕЛЕВСКИХ) ГЕНОВ

В случае если признак контролируется одним геном, сцепленные с ним маркеры, могут быть найдены посредством создания генетических карт сцепления [13, 15]. Однако неразумно картировать целый геном (если при этом необходимо найти маркеры в очень ограниченном районе) для поиска маркеров в очень ограниченном районе. Кроме того, вероятность найти маркеры, тесно сцепленные с искомым геном, используя подобного рода методологию, ничтожно мала. Существует два пути, позволяющие решить эту проблему и высвободить при этом значительное количество времени и денег — это использование практически изогенных линий (NIL) и совокупного сегрегационного анализа (bulked segregant analysis, BSA). Рассмотрим эти подходы.

4.1. Подходы к картированию основных (менделевских) генов

4.1.1. Практически изогенные линии

Изогенные линии можно определить как линии, которые имеют один и тот же генотип (аналогичные генотипы). Линии, практически изогенные (NIL) по данному локусу, — это линии, имеющие один и тот же генотип, за исключением локуса, который содержит фиксированные, но различные аллели. На практике такой случай встречается редко, поскольку в данном локусе генома должны произойти мутации и эти мутации должны быть зафиксированы. Для получения генетически схожего материала проводят следующие один за другим беккроссы между «донорной» линией, несущей желаемый аллель, и реципиентной (или рекуррентной) линией, содержащей другой аллель, после чего материал фиксируют посредством самоопыления. Для хромосомы, несущей интрогрессированный аллель, этот процесс приводит к умеренной изогенности: в каждом беккроссном поколении, отбор особей, содержащих желаемый аллель, приводит к переносу части хромосомы с данным геном («hitch-hiking» эффект) (рис. 2).

С помощью такого материала маркеры, сцепленные с геном, могут быть найдены сравнением донорной линии D , рекуррентной линии R и линии R' , получившей ген (R и R' практически изогенны). Для данного маркера возможны три ситуации:

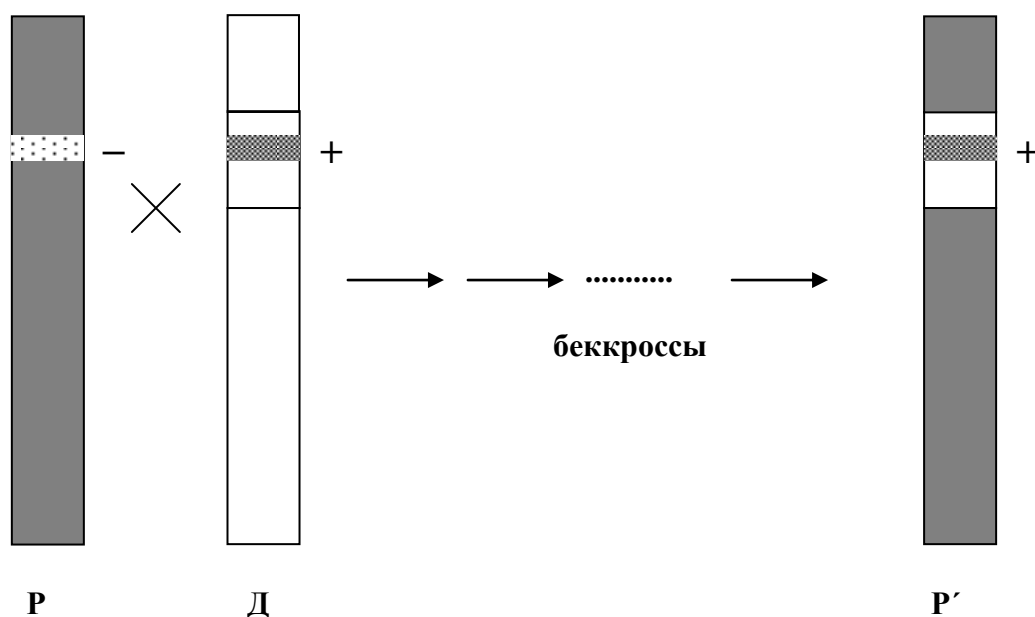


Рисунок 2. Практически изогенные линии (NIL). В течение процесса последовательных беккроссов, направленных на замену нежелательного аллеля (-) «реципиентной» линии (P) желаемым аллелем (+) «донорной» линии (D), часть генома родителя D переносится с интрогрессируемым аллелем.

- D, P и P' идентичны: локус является мономорфным и неинформативным (рис. 3а);
- D отличается от P, но P и P' идентичны: аллель D неинтрогрессирован в P' (рис. 3б). Следовательно, маркер не сцеплен с интрогрессированным геном.
- P и P' различаются и P' идентичен D (рис. 3в): аллель D интрогрессирован в P'. Следовательно, маркер сцеплен с интрогрессированным геном.

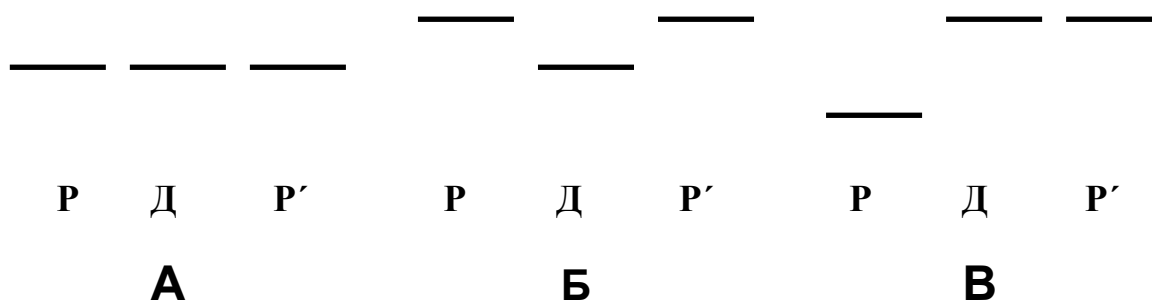


Рисунок 3. Возможные профили электрофореграмм при проведении сравнительного электрофоретического анализа практически изогенных линий (NIL). P — реципиентная линия; D — донорная линия; P' — линия, получившая часть генома родителя D. А — неполиморфный маркер; Б — полиморфизм за пределами интрогрессируемого района; В — полиморфизм между P и P' как результат интрогрессии исследуемого района.

Поскольку в процессе беккроссирования интрогрессии подвергается совершенно незначительная часть генома, то, как правило, чаще всего встречаются первые две ситуации. Поэтому для поиска маркеров изучаемого гена достаточно сравнить Р и Р', но при этом Д должен быть использован в качестве контроля. Такая постановка проверочного эксперимента позволит удостовериться, что аллель Р' действительно получен от Д, а не является технической (ложные полосы на геле) или генетической ошибкой (генетическое загрязнение перекрестным опылением).

Для сравнительного анализа подойдет любой тип маркера, но лучше использовать геномный ДНК-фингерпринтинг, поскольку он может предоставить широкий набор маркеров (электрофоретических полос) в одном эксперименте, например, AFLP, ISSR и др. Тот факт, что эти маркеры, как правило, доминантны, не является проблемой, поскольку в данном случае для любой фазы сцепления или расхождения гамет, различия могут быть найдены. Если необходимо, то из доминантных маркеров можно получить кодоминантные маркеры (например, SCAR). Чтобы удостовериться, что найденные подобным образом маркеры сцеплены с изучаемым геном, их необходимо картировать.

4.1.2. Совокупный сегрегационный анализ

В первой половине 1990-х годов с помощью практически изогенных линий были картированы несколько основных (менделевских) генов, прежде всего определяющих устойчивость к заболеваниям, [104, 246, 261, 262, 290, 328, 388]. Создание такого материала — процедура довольно длительная по времени, поскольку требуется получить несколько поколений, и иногда выявляются так называемые «фальшивые позитивы» (сегмент родителя Д интрогрессируется, но не несет желаемый ген). В этом случае используют методологический подход, получивший название, совокупный сегрегационный анализ (BSA) [24, 263]. Этот подход заключается в сравнении совокупности ДНК отдельных особей, имеющих один и тот же генотип. Например, необходимо определить различия между совокупностью особей *RR* (устойчивая гомозигота) и совокупностью особей *rr* (неустойчивая гомозигота) для популяции F_2 , рекомбинантных инбредных линий или удвоенных гаплоидов,

расщепляющихся по гену устойчивости. Только сцепленные с геном маркеры покажут различия, либо изменением месторасположения электрофоретической полосы на геле (кододоминантные маркеры), либо самим фактом ее присутствия/отсутствия (доминантные маркеры; снова подчеркнем, что методы ДНК-фингерпринтинга наиболее подходящи для BSA анализа). Как указывалось выше, эти сцепления должны быть проверены картированием идентифицированных проявляющих различия маркеров.

Минимальный размер совокупности определяют установлением максимальной вероятности возможности различить несцепленный полиморфный маркер от сцепленного с геном. Эта вероятность зависит как от типа популяции, так и от наследования маркера. Например, для доминантного маркерного расщепления в популяции F_2 , вероятность нахождения независимого сцепленного с геном маркера, при размере совокупности n , равна $2(1/4)^n[1 - (1/4)^n]$. Если в совокупности будет всего лишь 4 особи, то тогда доля фальшивых позитивов составит 1%, а в случае 10 особей: 2×10^{-6} . Именно поэтому, как правило, для анализа используют от 10 до 15 особей. С другой стороны, BSA не обязательно дает ответ по принципу «все или ничего». Иногда, вследствие того, что в совокупности могут присутствовать один или несколько отдельных индивидуальных рекомбинантных генотипов, а сцепление маркера и изучаемого гена есть не у всех отобранных для анализа особей, между совокупностями могут наблюдаться различия в интенсивности полос.

И, тем не менее, несмотря на эти ограничительные особенности, целый ряд исследователей с успехом использовали BSA для картирования основных (менделевских) генов у растений. Например, ген, определяющий пол у фисташкового дерева [176], а также ген – восстановитель фертильности у рапса [95]. Еще больше примеров можно найти для генов устойчивости к болезням, например, у салата [198, 234, 263], бобов [162] или сосны [98]. Ряд других примеров приведен у Michelmore [261].

Этот метод можно использовать также и для заполнения пробелов в картах. Для этого совокупности должны быть составлены на основе генотипов маркеров, фланкирующих интервал, который необходимо заполнить, при этом во внимание принимают только особи нереконбинантные по данным локусами (например, особи G_1D_1/G_1D_1 сравнивают с

особями G_2D_2/G_2D_2). Схожим образом маркеры концов хромосом могут быть найдены посредством BSA анализа периферических маркеров групп сцепления. Таким образом, шаг за шагом устанавливают точку генома, в которой нет «удлиняющих» карту маркеров [318].

В принципе BSA и NIL взаимосвязаны друг с другом. Совокупности выглядят как практически изогенные линии (NIL) с гетерозиготным геномом, потому что в случае локусов, не связанных с желаемым геном, они содержат две родительские аллели в статистически равных пропорциях. Тем не менее, при доминировании в изучаемом локусе и маркере, маркеры, которые находятся в фазе расхождения (англ. repulsion) с геном, не могут быть определены в совокупности беккроссов и F_2 , даже, если они полностью генетически сцеплены, поскольку электрофоретическая полоса появится в двух совокупностях (табл. 10). Именно поэтому, в среднем, вдвое больше маркеров необходимо протестировать в случае использования двойных гаплоидов или рекомбинантных инбредных линий. Существует только один способ обойти эту проблему — сравнить совокупность rr проявляющую полосу на геле с родителем RR , который не проявляет ее.

Таблица 10

Совокупный сегрегационный анализ и нахождение маркеров сцепленных основными (менделевскими) генами⁽¹⁾

Типы популяций	Фаза		Информативные маркеры
	сцепление	расхождение	
F_1 (генотипы)	$R1/r0$	$R0/r1$	
Удвоенные гаплоиды	[R1] и [r0]	[R0] и [r1]	все
Беккроссы ⁽²⁾	[R1] и [r0]	[R1] и [r1]	только если есть сцепление
F_2	[R1] и [r0]	[R1] ⁽³⁾ и [r1]	только если есть сцепление
Рекомбинантные инбредные линии	[R1] и [r0]	[R0] и [r1]	все

(1) — сравнение потомства, когда доминирование отмечалось для изучаемого локуса u у маркера. R и r обозначают, соответственно, доминантный и рецессивный аллели, а 1 и 0 — присутствие или отсутствие полосы полностью сцепленного с геном маркера. Фенотипы двух групп, полученные в каждой ситуации, обозначены в скобках.

(2) — беккроссирование проведено на родительскую гомозиготу с рецессивным аллелем изучаемого гена (в противном случае сегрегация не будет выявлена).

(3) — этот фенотипический класс гетерогенен по особям [R1] и [R0] в пропорции 2:1.

4.2. Использование маркеров для клонирования основных (менделевских) генов

4.2.1. Принцип позиционного клонирования

Ген с неидентифицированными мРНК или белком может быть выделен с помощью транспозонинсерционного мутагенеза или позиционного клонирования. Преимуществом транспозонинсерционного мутагенеза является способность мобильных элементов к инактивации генов, в которые они встраиваются. Если мутант найден по его фенотипическому проявлению, то ген может быть выделен с помощью мобильного элемента, явившегося причиной мутации. Как правило, такое возможно только для видов, у которых хозяйский генетический материал может вызывать транспозиции мобильных элементов, как, например, у кукурузы.

Позиционное клонирование не столь специфично, поскольку для выделения гена использует тесно сцепленные с геном маркеры. В исходном методе или «прогулке по хромосоме» [324] желаемый ген получают постепенным приближением к нему с помощью скринирования перекрывающихся друг с другом геномных клонов (космиды, YAC, BAC) [58, 398]. Помимо трудностей тестирования каждого нового клона на наличие в нем нужного гена этот подход у высших растений может занимать значительное количество времени из-за размеров их геномов. Даже если маркер находится на расстоянии 1 сМ от искомого гена, физически это может составлять несколько сотен или тысяч т.п.о. (тысяч пар оснований), которые почти невозможно полностью проскринировать. «Прыжки по хромосоме» [324] в определенном смысле являются усовершенствованием этого подхода, но методология поиска маркеров, основанная на использовании NIL или BSA, совместно с точными картирующими методами, представляет собой более сильное орудие для поиска и клонирования генов. Идея заключается в том, что на практически изогенных линиях можно проскринировать несколько тысяч маркеров и найти среди них те, которые находятся на генетической дистанции, соответствующей физическому размеру YAC, BAC или космидному клону. При скринировании библиотек с большими фрагментами ДНК с помощью молекулярных маркеров, как минимум для видов с не-

большими геномами, можно совершенно точно отобрать клоны, содержащие интересующие исследователя гены. Такой подход порой называют «посадка на хромосому» (chromosome landing) [368].

4.2.2. Картирование высокого разрешения

Если полиморфные маркеры уже отобраны, то порядок их картирования должен определяться поиском такого маркера, который находится ближе всего к изучаемому гену. Обычно для этого применяют картирование высокого разрешения, требующее анализа большого числа особей несколькими десятками маркеров (если, например, геном размером 2000 сМ скринирован 8000 маркерами, то около желаемого гена в интервале 10 сМ можно ожидать наличие порядка 40 маркеров). Чтобы получить разрешение меньше чем десять сантиморганов, что является минимумом для нахождения клона, содержащего искомым ген, понадобится более 3000 беккроссных особей (при условии, что с вероятностью 0,95 необходимо выявить, по крайней мере, одну рекомбинацию). Одним из способов уменьшить объем работы — анализировать особи популяции только по двум маркерам, о которых известно, что они фланкируют ген и лишь потом, для последующего анализа с помощью других маркеров, привлекать только те особи, о которых известно, что они рекомбинировали. Другая стратегия основывается на анализе совокупности особей [82]. Она заключается в разделении особей популяции согласно генотипам изучаемого локуса, с последующим анализом ДНК отделённых друг от друга совокупностей генотипов. Уровень рекомбинации между маркером и желаемым геном может быть вычислен исходя из пропорции совокупностей с профилями, которые выявляют хотя бы один кроссовер и в которых его нет. Как только каждая совокупность будет проанализирована серией маркеров, можно провести трехточковые тесты и идентифицировать наиболее близко расположенные к гену маркеры.

4.2.3. Поиск необходимого гена в геномных клонках

Для скринирования библиотек с большими фрагментами геномной ДНК (YAC, BAC или космиды) используют маркеры, наиболее тесно сцепленные с геном. Отобранные положительные клоны среди дру-

гих генов, как правило, содержат и необходимый ген. Одним из способов выявить подходящих кандидатов является биоинформативный анализ нуклеотидной последовательности выявленного клона. Клон также может быть использован для скрининга библиотек кДНК, что приведет к получению дополнительных генов-кандидатов. Для установления, какой из выявленных кандидатов является необходимым, самым прямым и действенным методом является трансформация. Проведение генетической трансформации позволяет получить практически однозначный ответ о функциональном назначении и роли выявленного гена-кандидата в проявлении изучаемого признака или о сцеплении с ним, поэтому ее следует проводить всякий раз, когда это возможно. Другой подход — это картирование с высоким разрешением отобранного клона с целью идентификации в нем гена, который ко-сегрегирует с признаком. И наконец, для некоторых признаков, действенным способом является изучение экспрессии гена-кандидата. Данный подход базируется на том факте, что экспрессия вовлеченного в ответ на стресс гена должна индуцироваться или контролироваться этим стрессом.

Кроме того, для выделения генов у различных видов растений, как правило, с небольшими размерами геномов, с успехом использовалось позиционное клонирование. Так, например, были клонированы: ген *ABI3* ответственный за уменьшение воздействие абсцизиевой кислоты в *Arabidopsis* [151], ген, кодирующий омега-3-десатуразу арабидопсиса [25], ген *Pto* устойчивости к *Pseudomonas syringae* томата [245], ген *Axr1* нечувствительности *Arabidopsis* к ауксину [222] и некоторые другие. Информацию о генах устойчивости к заболеваниям и о симбиотических генах корней растений можно прочесть в ряде обзоров [163, 189, 262, 354].

Заключение

Несмотря на то, что большинство хозяйственно ценных признаков являются количественными, в геноме живых организмов имеются признаки, которые контролируются единичными или основными (менделевскими) генами. Выявление и картирование таких генов, наряду с картированием QTL, также остается одной из основных задач генетиков и селекционеров. Использование практически изогенных линий и сово-

купного сегрегационного анализа позволяет решать эту задачу наиболее эффективно. Накопленные к сегодняшнему дню данные выявили ясные тенденции в архитектонике геномов, в особенности в частоте случаев явной ко-локализации между локусами хромосом, определяющих проявление количественных признаков, и основными (менделевскими) генами, проявляющими в своем большинстве моногенное наследование и такие же эффекты действия (например, гены, контролирующие устойчивость к патогенам, карликовости, чувствительности к фотопериоду, архитектурному построению и др.). Даже если огромная задача по идентификации и локализации QTL останется и должна будет раз за разом, эксперимент за экспериментом подтверждать, что QTL и гены с количественными эффектами принадлежат одним и тем же локусам, полученные результаты, также как и первые эксперименты по идентификации и характеристике QTL у растений, животных и человека, ясно указывают на непрерывный характер спектра между дискретной и непрерывной изменчивостью. В этом контексте, становится понятным, что картирование конкретных основных (менделевских) генов более предпочтительно, чем картирование анонимных маркеров, поскольку они являются источниками генов-кандидатов. Программы систематического секвенирования EST (англ. expressed sequence tags; — последовательности с неустановленной функцией, экспрессированные в исследуемом объекте целевым образом), позиционное клонирование, скринирование YAC и BAC геномных библиотек, а также картирование высокого разрешения у растений помогут ускорить получение таких карт. В особенности, выделение кДНК специфичных к определенной стадии развития, органа или полученных под воздействием какого-либо стресса.

С появлением все большего и большего количества насыщенных генетическими маркерами молекулярных карт, полученных для наиболее распространенных видов растений, а также с доступностью таких карт для все более широкого круга исследователей и сфер их деятельности, физиологи и генетики, а в перспективе и селекционеры, будут нуждаться только в скрининге родителей известных картированных популяций для установления вариаций в экспрессии представляющих интерес признаков, что позволит методом экстраполяции оценить соответствующую картированную популяцию по изучаемому признаку. Разуме-

ется, что использование QTL анализа для тестирования генов-кандидатов будет требовать наличия картированных генов, в том числе основных (менделевских), в уже изученной популяции. Однако техники и методологии картирования с каждым десятилетием становятся все легче и быстрее, и нет принципиальных ограничений (пожалуй, только за исключением финансовых возможностей) для обеспечения исследователей современным лабораторным оборудованием, необходимым для проведения молекулярно-биологических работ и исследований в этом направлении.

Глава V. АБИОТИЧЕСКИЙ СТРЕСС И ЛОКУСЫ КОЛИЧЕСТВЕННЫХ БИОХИМИЧЕСКИХ И ФИЗИОЛОГИЧЕСКИХ ПРИЗНАКОВ У РАСТЕНИЙ

Несмотря на то, что генетики и физиологи, изучающие количественные признаки растений, имеют дело с изменчивостью, подходы которые они применяют для изучения природы исследуемой ими изменчивости, у них различаются. Генетики в основном заинтересованы в объяснении природы и использовании генетической основы изменчивости, а физиологи обычно стремятся минимизировать генетическую компоненту, для того чтобы сконцентрироваться на изменчивости, возникающей под воздействием окружающей среды. Наследуемость, являющаяся частью общей фенотипической изменчивости, обусловленной генетическими различиями, используется генетиками для определения того, может ли признак использоваться в генетико-селекционных исследованиях (т.е. является достаточно изменчивым или полиморфным), в частности, для картирования соответствующих генов или локусов хромосом. При этом если исследователи имеют дело со сложными или количественными признаками, то они учитывают тот факт, что несколько локусов могут определять один признак. Появление молекулярных маркеров позволило начать картирование локусов количественных признаков или QTL (quantitative trait loci) [286]. Это делает возможным физическое разделение генетических компонентов (или их составляющих) количественных признаков на небольшое число хромосомных локусов и оценить вклад каждого из таких локусов в определение генетической изменчивости у изучаемого организма. Ранние попытки локализовать QTL основных агрономических свойств выявили, что позиции QTL не оставались стабильными от одного эксперимента или установленного локуса к другому, что указывало на значительный эффект воздействия окружающей среды на проявление признака. Чтобы установить это в ряде экспериментов под воздействием различных условий окружающей среды была протестирована стабильность QTL. У томата, например, 350 потомков поколения F₂ выращивались в трех различных экологических точках (две в США и одна в Израиле). В результате проведенных экспериментов было выявлено 29 QTL, причем 15 из них были специфичны

лишь для какой-то одной точки и только 5 оказались общими для всех трех точек тестирования [285]. В то же время более слабый эффект воздействия окружающей среды был обнаружен у кукурузы, которая выращивалась в шести различных экологических точках [360]. В этом случае пропорция идентифицированных общих QTL в проведенных экспериментах была выше.

Поиск QTL, проявление которых связано с реакцией на окружающую среду, может показаться несколько странным, поскольку генетически локализовать взаимодействие «генотип-среда» невозможно, особенно в случае количественных признаков, когда признак формируется под воздействием различных лимитирующих факторов среды и проявление одного и того же признака в разных условиях среды определяют разные гены или генные хромосомные локусы. В то же время, так называемая изменчивость окружающей среды — это переменный фактор общей изменчивости организма, который остается после исключения генетической изменчивости из общей составляющей того или иного признака. Это значит, что изменчивость окружающей среды включает в себя не только соответствующую климатическую изменчивость, являющуюся по своей сути различного рода стрессорами, но и значительную часть экспериментальной ошибки, с которой приходится сталкиваться при проведении и воспроизведении экспериментов. При установлении QTL, имеющих отношение к стрессу, возникающему под воздействием окружающей среды, необходима тщательная постановка эксперимента с соответствующими повторами. К сожалению, в этом ключе проведено сравнительно небольшое количество экспериментов.

Таким образом, дополнительная идентификация и проведение анализа выявленных локусов QTL, позволили бы установить присутствие или отсутствие ко-локализации QTL-признаков, определяющих некоторые физиологические процессы (например, налив зерна, ответ на тот или иной абиотический или биотический стресс и др.), а также возможна ли ко-локализация QTL и генов-кандидатов, связанных с тем или иным QTL. Ко-локализация означает ко-сегрегацию, которая может быть использована селекционерами и биотехнологами растений при получении новых растительных форм и реализации селекционных программ. Ко-локализация QTL для некоторых признаков может опреде-

лять ко-регуляцию экспрессии. Более того, если QTL ко-локализованы с генами с известной функцией, то эти гены представляют интерес как гены-кандидаты и могут быть использованы для объяснения фенотипических корреляций между признаками [302, 303].

На сегодня этот подход проиллюстрирован полевыми экспериментами по анализу ряда агрономических признаков и экспериментами, проведенными в теплице и направленными на более точный анализ эффектов различных стрессов преимущественно на ранних стадиях развития растений. Так, например, с помощью QTL проанализировано несколько различных количественных признаков и ответов на изучаемый стресс [302, 303], таких как урожайность, рост корней и проростков, фотосинтетический обмен газов, содержание абсцизовой кислоты (АК) в ксилеме, содержание углеводов и активность расщепляющих их ферментов (инвертазы, сахарозосинтазы и АДФ-глюкозофосфорилазы).

5.1. Некоторые элементы методологии

Принцип и методология определения QTL достаточно подробно была описана ранее [13, 14, 303] (см. также Главы III и IV). Коротко, основным необходимым условием для определения QTL необходимы, во-первых, линии расщепляющихся популяций, полученные от двух гомозиготных родителей. При этом чтобы получить достаточную точность эксперимента, минимальное число линий, которое необходимо проанализировать, должно быть равно 100-150. Сегрегационными линиями могут быть расщепляющиеся популяции F_2 , удвоенные гаплоиды или рекомбинантные инбредные линии (recombinant inbred lines, RIL). Рекомбинантные инбредные линии обычно получают как минимум четырехкратным самоопылением поколения F_2 , произошедшего из одного семени. Основное достоинство такой процедуры заключается в том, что полученные RIL являются практически гомозиготными, т.е. генетически однородными. Вторым требованием является картирование популяций с достаточной плотностью насыщения генетическими маркерами таким образом, чтобы число групп сцепления на генетической карте было равно числу хромосом. При соблюдении этих условий для каждой сегрегационной линии становится возможным распознавать генотип (либо первого, либо второго родителя) по каждому отдельно взятому маркеру.

Принцип определения QTL заключается в измерении вариантов признаков в процессе изучения во всех сегрегационных популяциях и последующей проверке одного маркированного локуса за другим с целью установления зависимости среднего значения признака от генотипа в данном локусе. Если результат статистически достоверен, то это значит, что QTL, определяющий изменчивость данного признака, расположен поблизости от данного маркера. Этот точечный пошаговый метод, основанный на ANOVA (ANalysis Of VAriance — статистический тест, который количественно устанавливает различия между значениями двух или более групп), на сегодня постепенно замещается интервальным картированием, которое позволяет идентифицировать локализацию между маркерами (как, например, MAPMARKER-QTL). Более современный метод, названный методом комбинированного интервального картирования (англ. composite interval mapping, CIM), совмещает в себе интервальное картирование и введение кофакторов, которые с одной стороны увеличивают чувствительность определения, но с другой стороны также увеличивают и риск выявления фальшивых позитивных QTL [335]. Наиболее важным параметром, вычисляемом в дополнение к установлению локализации QTL, является величина R^2 , представляющая процент фенотипической вариации объясняемой QTL.

5.2. QTL и абиотический стресс

5.2.1. Высокая температура (термоустойчивость)

Хотя большинство данных по картированию QTL физиологических параметров, так или иначе, связаны с определением водного статуса растений, одни из первых исследований направленных на идентификацию QTL признаков имеющих отношение к климатическим адаптациям были эксперименты по выявлению QTL термоустойчивости пыльцы [137]. При этом для этих целей авторы использовали всего 45 популяций RIL, с помощью которых им удалось установить пять и шесть QTL прорастания пыльцы и роста пыльцевой трубки соответственно.

5.2.2. Состояние засухи на вегетативных стадиях развития

5.2.2.1. Корневая система, морфология и признаки проростков у риса

Совершенно очевидно, что морфология корней является важным фактором при определении устойчивости растений к засухе. В одном из первых экспериментов по QTL анализу, проведенном в 1995 году, удалось картировать пять связанных с этим признаком QTL [69]. Эксперименты проводились на рисе (203 RIL в тройной повторности), который выращивался как в теплице, так и в полевых условиях. Двенадцать из сорока хромосомных районов, несущие QTL ассоциированные с полевой устойчивостью к засухе, также содержали QTL морфологии корней. Большинство идентифицированных QTL определяли толщину корней и соотношения: корень/проросток, сухой вес корней/побег. Толщина и сухой вес корней на побег являлись признаками, которые, как было установлено, в последнюю очередь подвержены влиянию окружающей среды в условиях теплицы. Корреляции экспериментов, проведенных в теплице с устойчивостью к засухе, в полевых условиях оказались достоверными, но не очень предсказуемыми. Эти первые результаты позволили авторам идентифицировать аллели в маркерных локусах, ассоциированных с фенотипом корней, и получить засухоустойчивые сорта. С целью изучения приспособления растений к осмотическому стрессу и выявления признаков устойчивости к дегидратации с использованием тех же рисовых популяций, но на гораздо меньшем количестве образцов (52 RIL) были проведены дальнейшие более полные исследования. Авторы предположили, что один основной локус, ассоциированный с осмотолерантностью должен быть гомеологичен к схожему локусу того же признака у пшеницы. В популяциях, осмотолерантность и признак устойчивости к дегидратации негативно коррелировали с признаком морфологии корней, ассоциированного с устойчивостью к дегидратации. К счастью высокая приспособленность к осмотическому стрессу и признак устойчивости к дегидратации оказались генетически связанными с одним и тем же родительским аллелем. Это вселяет надежду на разрыв нежелательного сцепления между осмотолерантностью и некоторыми признаками морфологии корней [223].

Последующие исследования также базировались на изучении морфологии корней, но без учета влияния водного стресса. В этом случае использовали 105 линий риса, полученных из популяций удвоенных гаплоидов [409]. Авторы установили, что основное количество QTL сконцентрировано в определенных компактных районах хромосом 1, 2, 3, 6, 7, 8 и 9. Индивидуальные QTL отвечали за варьирование признаков от 4 до 22%. Взаимодействие между QTL и маркерами, расположенными на разных хромосомах, показали антагонистические эффекты. Сравнение с другими популяциями показало, что от одного до трех QTL на признак были общими.

Другие авторы исследовали QTL признаков, имеющих отношение к засухоустойчивости у проростков риса [84]. В двух различных экологических точках на 85-105 линиях удвоенных гаплоидов под воздействием различных степеней засухи они измеряли признаки скручивания и высыхания листьев, относительное содержание воды и относительный уровень роста. Как показали проведенные эксперименты, некоторые QTL были общими для всех исследованных признаков. Среди одиннадцати возможных QTL скручивания листьев три (на хромосомах 1, 5 и 9) оказались общими во всех трех проведенных испытаниях, а четыре дополнительных QTL (на хромосомах 3, 4 и 9) — лишь в двух. Один QTL признака высыхания листьев на хромосоме 4 и один относительного содержания воды на хромосоме 1 совпали по своему проявлению в двух различных эколого-географических экспериментах, в то время как не было идентифицировано ни одного QTL относительного уровня роста, проявление которых было бы обусловлено воздействием стресса, вызванного недостаточностью обеспечения влагой. Некоторые из QTL скручивания и высыхания листьев, а также относительного содержания воды картированы в тех же локусах на хромосомах, что и QTL контролирующей морфологию корней, и которые были идентифицированы в предыдущих исследованиях, включавших те же самые расщепляющиеся маркированные популяции. Локализация некоторых QTL, установленная в этом исследовании, оказалась точно такой же, как и локализация аналогичных QTL, определенная другими авторами с помощью других популяций.

5.2.2.2. Осмотолерантность у риса

Как это уже отмечалось ранее [223, 420] адаптация к осмотическому стрессу или осмотическая приспособленность (ОП) может играть роль в засухоустойчивости, обеспечивая возможность поддержания клеточного тургора при снижении водного потенциала. Эта проблема была рассмотрена на примере ячменя с использованием QTL-методологии [377]. Набор рекомбинантных инбредных линий ячменя, полученных в результате скрещивания сортов Tadmor (засухоустойчивый родитель) и Er/Arp (чувствительный родитель), выращивали в климатической камере в условиях различной влажности грунта (14% и 100% полевой емкости почв по этому признаку, соответственно). Карта сцепления была построена с использованием 167 RIL генотипированных посредством 78 RFLP, 32 RAPD и трех морфологических маркеров. Благодаря сильному эффекту воздействия условий окружающей среды на изучаемые свойства (изменчивость варьировала от 0,04 до 0,44) интервал-картирование и одномаркерный ANOVA-анализ позволили определить три QTL признака относительного содержания воды (ОСВ), четыре QTL общего осмотического потенциала, два QTL осмотического потенциала при полном тургоре и один QTL ОП при 14% влажности почвы. В поливных условиях выявлены только два QTL: один для ОСВ и один для осмотического потенциала при полном тургоре. Два хромосомных района оказались ответственны за некоторые изменения свойств, сопутствующие ОП, и поэтому могут считаться как районы, контролируемые ОП. Они располагались на хромосомах 1 (7H) и 6 (6H), в то время как другие районы были специфичны только для какого-то одного свойства. Никаких основных QTL найдено не было. В то же время, как представляется, геномные районы на хромосоме 1 (7H) у ячменя, вовлеченные в регуляцию сопутствующих ОП свойств, должны быть консервативны в отношении изменчивости ОП у злаков.

5.2.2.3. Листья и абсцизовая кислота (АК) у кукурузы

Большинство исследований по изучению эффектов засухи у растений проводились на вегетативных органах, прежде всего на листьях. Потеря воды, как правило, приводит к двум типам реакций в зависимости от возраста растительных тканей: 1) быстрой остановке роста клеток

в молодых растущих органах, 2) уменьшению уровня фотосинтеза и изменению метаболизма сахарозы в зрелых листьях. АК, выделяемая кончиками корней и циркулирующая в ксилеме, играет важную триггерную роль в подобного рода ответных реакциях. Сталкиваясь с многочисленными взаимозависимыми взаимодействиями очень трудно установить причинные связи, которые объяснили бы наблюдаемый ответ целого растения, вызываемый несколькими предполагаемыми ключевыми элементами. Прояснение этого вопроса особенно важно для возделываемых растений, поскольку идентификация таких основных элементов или генов позволит получить способ воздействия на организм с целью повышения адаптации или усиления устойчивости растительного организма к стрессу. С этой точки зрения, использование методологии QTL наиболее подходящий подход, поскольку она позволяет устанавливать корреляции между различными признаками. Кроме того, ко-локализация QTL может предоставить информацию по сцеплению между признаками на разных уровнях интеграции (от активности ферментов до обмена газов или роста листьев).

Первая попытка выявить QTL, определяющие количественные уровни АК по отношению к степени проявления водного стресса, была предпринята на расщепляющихся маркированных популяциях F₂ пшеницы и кукурузы [307, 309]. Несмотря на то, что плотность маркеров была довольно низка (32 у кукурузы) авторам удалось продемонстрировать принципиальную возможность применения данного метода. Более тщательный анализ проведен уже на 81 популяции кукурузы F₂ и 84 маркерах. В дополнение к АК проблема водного стресса была перепроверена и в отношении других свойств, таких как устьичная проводимость, водный потенциал, тургор, число корней, сила натяжения корней, флуоресценция хлорофилла и некоторых параметров пыльников [214]. За счет увеличения числа маркеров месторасположение большинства QTL АК подтверждено на хромосомах 1, 3, 6, 7 и 8, но на хромосоме 2 выявлены новые QTL АК. Когда авторы приняли во внимание все признаки, ими был сделан вывод о том, что QTL различных признаков имели тенденцию к ко-локации с одновременным образованием кластеров. Так, например, в середине короткого плеча хромосомы 1 возле маркера umc11 картированы сразу четыре QTL: числа корней, си-

лы натяжения корней, АК ксилемы и устьичной проводимости. Схожим образом QTL как минимум четырех признаков ко-локализованы на хромосомах 3 и 7.

Для того чтобы рассортировать бессистемные взаимоотношения был предложен усовершенствованный метод [214], основанный на сравнении двух возможных регрессионных линий, полученных после схематизирования аллельных эффектов для двух признаков по двенадцати маркерам объединяющих максимальные дистанции. В этом случае, если регрессия признак1/признак2 была более достоверна, чем регрессия признак2/признак1, то предполагалось, что признак2 контролируется признаком1. Например, $r = -0,067^*$ для регрессии проводимость/АК и $r = -0,43^{NS}$ для реципрокной регрессии может означать, что количество АК более вероятно контролируется проводимостью, чем наоборот. Хотя этот результат хорошо согласуется с проведенными ранее физиологическими исследованиями, сам метод не вполне четкий, поскольку с теоретической точки зрения он не берет в расчет возможные различия в изменчивости каждого признака.

Другой путь установления корреляций между признаками был предложен при изучении межвидовых сравнений [308]. Принцип базируется на использовании Закона гомологических рядов в наследственной изменчивости Н.И.Вавилова [3] и том предположении, что сцепленные гены у одного вида схожим образом связаны друг с другом и у другого родственного вида. Для того чтобы установить, что ассоциация между геном и QTL или между генами не является артефактом, проводят сравнение картированных позиций гомологичных генов, локализованных в точке изучения. Авторы перепроверили предыдущие данные, которые показали, что у стоящих особняком листьев риса имеет место негативная корреляция между накоплением АК и размером листа (меньшие по размеру листья продуцировали больше АК). Из установления факта ко-локализации различных QTL и выявления того, что только один QTL АК и один QTL размера листа согласовывались с противоположным эффектом действия аллелей, авторы заключили, что причиной ассоциации между количеством АК и размером листа в большей степени являлось генетическое сцепление, чем прямой эффект размера листа на аккумуляцию АК или *vice versa*. Особый интерес представлял

Vrn1 — основной ген ответственный за яровизацию у пшеницы. Этот ген оказался ассоциирован с основными эффектами QTL, определяющих размер листьев, число побегов и накопление АК. Однако локализация гомологичных генов у *Oryza sativa* не выявила какого-либо совпадения месторасположения генов АК и QTL морфологии листьев у риса.

В ряде исследований рассмотрены взаимосвязи проявления активности QTL АК по отношению к уровню водного стресса в полевых условиях. Так, например, в течение двух лет проводились замеры концентрации L-АК в листьях на двух стадиях развития (элонгация стебля и закладка пыльников) [383]. Хотя ни число линий (80), ни число картированных маркеров (106) не было очень высоко, авторы определили шестнадцать различных QTL, и только четыре из них согласовались между собой в пределах образцов. В то же время, эти четыре QTL учитывали 66% фенотипической изменчивости. Аллели, которые увеличивали концентрацию L-АК, были привнесены родителем с высоким содержанием АК в тканях (Os420). Два наиболее важных QTL были картированы на хромосоме 2 возле *csu133* и *csu109a*. Эффекты, ассоциированные с QTL, расположенным возле *csu133*, более резко проявляли свое действие в тканях вблизи пыльников. Доверительные интервалы для четырех первичных QTL L-АК не перекрывали интервалов картированных позиций аналогичных QTL у мутантов с ухудшенным биосинтезом АК. Проведены исследования также и по измерению устьичной проводимости, температуре листьев, относительному содержанию воды в листьях и учету массы зерен [327]. Анализ эффектов каждого района QTL по исследуемым признакам показал, что концентрация L-АК представляет собой индикатор уровня стресса засухи, испытанной растением во время сбора (отбора) образцов, поскольку увеличение содержания L-АК в тканях в основном ассоциировано с уменьшением устьичной проводимости, падением урожая зерна и с увеличением температуры листьев. Противоположные результаты по этим признакам наблюдались для районов QTL на хромосоме 7 возле RFLP-локуса *asg8*, что предполагает, в данном случае, наличие плейотропного эффекта.

5.2.2.4. Клеточно-мембранная стабильность у риса

Клеточно-мембранная стабильность (КМС), которая, как предпо-

лагают, должна быть весьма важна при предотвращении тока ионов в условиях стресса, впервые была исследована с использованием метода QTL-картирования в 2000 году [382]. Авторы избрали этот признак для анализа засухоустойчивости у риса. Для этого в условиях теплицы были изучены 104 линии удвоенных гаплоидов *Oryza sativa* L., полученных из генотипов F₁ — потомков скрещивания CT9993-5-10-1-M × IR62266-42-6-2. 50-дневные растения подвергались пропорционально увеличивающемуся стрессу посредством приостановки подачи воды или недостаточного полива. После этого на образцах листьев, собранных как в контроле (нормальный полив), так и у подвергнутых стрессу растений (60-65% полива от контроля), проведен стандартный тест на КМС. И хотя существенных различий между двумя родителями по засухоустойчивости обнаружено не было, также как и среди 104 линий в условиях стресса, КМС у исследованных линий оказалась различной. Это указывает на то, что изменчивость по КМС генотипична по своей природе. Наследуемость изучаемого признака в проведенных экспериментах составила 34%. Карта сцепления маркированной популяции включала 145 RFLP, 153 AFLP и 17 микросателлитных маркеров. Все они были использованы для QTL анализа. Сравнительное интервальное картирование выявило девять предполагаемых QTL КМС расположенных на хромосомах 1, 3, 7, 8, 9, 11 и 12. Количество фенотипической изменчивости, определяемое индивидуальными QTL, ранжировало от 13,4% до 42,1%. Исследователи обнаружили четыре достоверных ($p < 0,05$) пары дигенных взаимодействий между выявленными определяющими КМС QTL. Эти результаты продемонстрировали возможность использования КМС в качестве селективного признака в селекции по увеличению засухоустойчивости у риса.

5.3. Условия засухи на стадии созревания и QTL компонентов урожая

5.3.1. Кукуруза

Урожай кукурузы сильно падает, когда водный стресс приходится на фазу цветения растений. В соответствии с этим основная задача для многих селекционеров — это улучшение или повышение урожая в ус-

ловиях засухи. Однако это трудная задача, поскольку признак засухоустойчивости обычно полигенен и коррелирует с другими признаками частично. Для того чтобы попытаться найти решение этой сложной проблемы был использован метод идентификации QTL, с конечной целью разработать стратегии маркер вспомогательной селекции по данному признаку у кукурузы [319]. В течение двух лет растения кукурузы выращивали при трех разных поливных условиях: хороший полив, а также промежуточный и строгий водный стресс. При этом авторы измеряли следующие агрономические признаки: время созревания пыльников (ВСП), время цветения мужского и женского соцветий, урожай зерна (УЗ), число початков, число зерен на початок и масса 100 зерен. В качестве родителей для получения 234 самоопыленных F_2 растений в экспериментах использованы две тропические инбредные линии, проявляющие контрастный ответ на засуху. Основные измерения проводили на растениях F_3 . Все 234 линии F_2 были генотипированы по 142 локусам.

Первый результат воздействия водного стресса, обнаруженный у маркированных кукурузных линий, заключался в установлении увеличения генетической изменчивости ВСП и широкого разброса наследуемости признака времени цветения каждого из соцветий. Многочисленные QTL, определенные для этих признаков, составляли 47-48% отмечаемой фенотипической изменчивости. В условиях недостатка полива четыре QTL оказались общими для времени цветения мужского и женского соцветий, один — для ВСП и для времени цветения мужского соцветия, а четыре — для ВСП и для времени цветения женского соцветия. Число ко-локализаций QTL было связано с уровнем корреляций между признаками. Так, например, признак ВСП проявил трансгрессивный характер. Установленные значения ВСП у исследованных линий кукурузы были либо выше, либо ниже значений этого же признака у родительских форм. При этом, как оказалось, засухочувствительный родитель привносил аллели, обнаруженные в двух QTL, редуцирующих ВСП (4 дня). Аллели, привнесенные устойчивым родителем в других четырех QTL, отвечали за 7-дневное уменьшение ВСП. Эти четыре QTL, составляющие примерно 9% карты сцепления, оставались стабильны годами и не зависели от уровня стресса. Авторы [319] аргументировано показали,

что маркерная селекция, основанная на определении QTL ВСП, является сильным инструментом для улучшения засухоустойчивости у тропических кукурузных инбредных линий.

Во второй части своих исследований те же авторы [320] анализировали компоненты урожайности: урожай зерна, число початков, число зерен на початок и массу 100 зерен. Засуха выражалась в 60% увеличении УЗ в условиях жесткого стресса. Семьи, которые давали лучший урожай в условиях полива, как было установлено, должны быть пропорционально более подвержены стрессу. Их урожай уменьшался под воздействием жесткого стресса и обратно пропорционально проявлялся в условиях засухи. Это значит, что селекция по улучшению урожайности в поливных условиях не очень эффективна, в отличие от селекции по этому же признаку в условиях засухи. В то же время, в работе не были установлены QTL компонентов урожая для более чем 13% фенотипической изменчивости. Другое важное наблюдение заключалось в том, что позиции QTL оказались недостаточно консервативны по отношению к водному режиму. Это было квалифицировано авторами как «несовместимость» [320], но, по-видимому, это явление в большей степени является индикатором того, что стресс-специфичные гены взаимодействуют друг с другом (см. ниже).

Использование комплексного интервального картирование (Composite Interval Mapping, CIM) позволило оценить взаимодействие «QTL-окружающая среда» ($Q \times E$) и таким образом идентифицировать «стабильные» в условиях засухи QTL. Два таких QTL УЗ, обнаруженные на хромосомах 1 и 10, совпадали с двумя QTL, определяющими число зерен. Более того, четыре геномных района были идентифицированы как ответственные за проявления признаков УЗ и ВСП. Для некоторых из них, аллельные вклады короткого ВСП на хромосоме 10 соответствовали редуцированию урожая. Из полученных результатов авторы заключили [320], что для улучшения урожая в условиях засухи с помощью маркерной селекции необходимо комбинировать «лучшие» QTL, вовлеченные в становление компонентов урожая, с «лучшими» QTL вовлеченными в ВСП. При этом отбираемые QTL должны быть стабильными во всех предполагаемых условиях окружающей среды и

представлять настолько большой насколько возможно процент фенотипической изменчивости.

Схожий анализ был проведен и для генотипов кукурузы адаптированных к температурным перепадам воздуха [136, 329]. Измерения компонентов урожая включали: длину початка, вес початка, вес зерен, число зерен, вес 50 зерен и признаки, имеющие отношение к развитию и цветению растений (ВСП, время цветения, высота растения). 142 рекомбинантные инбредные популяции 142 семей, полученные в результате скрещивания В73 × Н99 и генотипированные по 173 локусам (RFLP, микросателлитам и AFLP), были оценены в поливных и водно-стрессовых условиях. Индекс засухоустойчивости авторы определяли как отношение между средними значениями признаков в двух различных экологических условиях окружающей среды. Для признаков компонентов урожая достоверные позитивные корреляции обнаружены для двух водных режимов. При этом более чем 50% QTL оказались общими, а аллельное увеличение признака привнесено в основном от линии В73. В то же время, несколько QTL идентифицированы как общие для двух или более признаков. Однако при определении индекса устойчивости, большинство QTL были специфичны для одного компонента и различались для тех, что контролировали основные признаки. В дополнение можно сказать, что основная часть аллельного увеличения устойчивости предоставлена линией Н99.

При анализе признаков развития некоторые исследователи [329] не только картировали соответствующие QTL, но и сравнили между собой районы хромосом, в которых располагались локусы засухоустойчивости связанные с признаками развития растений и временем цветения. Анализ сцепления выявил, что в условиях контроля и стресса для признаков времени цветения мужского соцветия и высоты растения, основные идентифицированные QTL совпадали для обоих признаков. В противоположность этому признак времени цветения женского соцветия и ВСП, регулировались различными QTL как в контроле, так и в условиях стресса. Все QTL подтверждающие устойчивость к засухе идентифицированы в различных хромосомных локусах, как это было установлено в сравнении с картированными позициями факторов контролирующих признак *per se*. Данное наблюдение позволило предположить [329], что

устойчивость растений в ее различных компонентах, не может быть отнесена к присутствию предпочтительных аллельных комбинаций, контролирующих признак, а базируется на физиологических свойствах не напрямую ассоциированных с контрольным элементом того или иного свойства. Эта концепция заслуживает особого внимания, поскольку она содержит в себе элементы определения стратегии дальнейшего развития маркерной селекции. Возможное логическое допущение, объясняющее наблюдаемое несовпадение между так, называемыми факторами, позволяющими проведение селекции по засухоустойчивости и QTL, может заключаться в том, что эти факторы контролируются полигенно, и что в зависимости от условий среды происходит переопределение контролирующих локусов, как это наблюдалось для инвертаз кукурузы [288].

5.3.2. Сорго

Сорго является одним из наиболее засухоустойчивых растений и может быть предложено в качестве модельного объекта по оценке механизмов устойчивости. В ряде исследований одновременно проанализировано 98 рекомбинантных инбредных линий сорго, полученных от скрещивания двух контрастных линий, T × 7078 (устойчивая до начала цветения и чувствительная после цветения) и B35 (чувствительная до начала цветения и устойчивая после цветения) [384]. Вся популяция была оценена в условиях засухи и контроля (отсутствие засухи) по признакам, ассоциированным как с засухоустойчивостью в период после завершения цветения, так и с компонентами развития зерна. QTL были картированы в 13 геномных районах, ассоциированных с одним или более признаками засухоустойчивости в период после цветения. Идентифицированы два QTL с большими эффектами по признакам урожайности и проявления «зеленой окраски» в период засухи после времени цветения. Эти локусы были также ассоциированы с урожайностью в условиях полной ирригации подтверждая, что идентифицированные QTL устойчивости имеют плеiotропный эффект на урожайность в условиях отсутствия засухи. Анализ признаков указывает на наличие многих локусов, которые ассоциированы как со скоростью, так и с продолжительностью развития зерен. Высокая скорость и короткий период развития зерен были в целом ассоциированы с большим размером зерна, но толь-

ко два из этих локусов были ассоциированы с различиями в стабильности их проявления в условиях засухи. Эти результаты значительно отличаются от результатов, полученных для кукурузы [319, 320], у которой QTL компонентов урожая были расположены на тех же самых позициях в условиях наличия засухи и в контроле (отсутствие засухи).

Важность свойства «сохранение зеленого цвета» (СЗЦ) как признака устойчивости в период прохождения цветения была подтверждена для зерен сорго [85]. QTL контролирующие старение и созревание были измерены на рекомбинантных инбредных линиях, помеченных 142 RFLP маркерами, для того, чтобы изучить их возможную взаимосвязь в условиях засухи в период прохождения цветения. RIL и их родительские формы были оценены в четырех различных условиях произрастания. Простое интервальное картирование выявило семь QTL признака СЗЦ и два QTL созревания. Три основных QTL СЗЦ делали вклад в 42% фенотипической изменчивости ($LOD = 9,0$) и четыре минорных QTL многозначительно делали вклад в дополнительные 25% фенотипической изменчивости в номинальную характеристику свойства СЗЦ. Два основных QTL делали вклад в 40% и 17% (соответственно, $LOD = 10,0$ и $LOD = 4,9$) фенотипической изменчивости соответственно. Сложное интервальное картирование подтвердило изложенные выше результаты посредством дополнительного анализа взаимодействия «QTL \times окружающая среда». Хотя номинальная характеристика свойства СЗЦ имеет взаимосвязь с созреванием (коэффициент корреляции $r = 0,22$, $p < 0,05$), шесть из семи QTL признака СЗЦ были независимы от QTL, влияющих на созревание. Схожим образом один QTL созревания был независим от QTL свойства СЗЦ. Один QTL признака СЗЦ, однако, был картирован по соседству от QTL созревания и все маркеры по соседству от независимого QTL созревания обладали высокосignимой корреляционной связью ($p < 0,1$) с номинальной характеристикой свойства СЗЦ. Молекулярно-генетический анализ QTL влияющих на свойство СЗЦ и созревание совместно с ассоциацией между этими двумя взаимнообразно зависимыми признаками обеспечивает основу для дальнейших исследований молекулярно-физиологических механизмов и для улучшения засухоустойчивости растений посредством многоступенчатых благоприятствующих QTL.

Идентификация генетических факторов, вовлеченных в фенотипическое проявление признака СЗЦ по отношению к ответу растений на засуху у сорго, была перепроверена [408] с целью уменьшения чувствительности растений в период стресса после цветения. QTL контролирующие проявление признака СЗЦ и содержание хлорофилла были определены при помощи RFLP карты полученной от популяции рекомбинантной инбредной линии. Четыре QTL признака СЗЦ были идентифицированы и локализованы на трех группах сцепления (2 на группе А и по одному на группах Q и J). Три QTL по содержанию хлорофилла объясняли 25-30% фенотипической изменчивости, в то время как два QTL признака СЗЦ на группе А объясняли 13-20% и 20-30% соответственно, во всех испытаниях, проведенных в различных экологических точках в течение двух лет. Три хлорофильных QTL совпадающих с тремя QTL районами, контролирующими проявление признака СЗЦ (группы А и Q) давали 46% фенотипической вариабельности. QTL признака СЗЦ на группе А были также расположены на районах, содержащих гены ключевых фотосинтетических ферментов, белков теплового шока и АК-зависимых генов. Такое пространственное расположение показало, что группа сцепления А важна для засухо- и жароустойчивости, а также получения урожая у сорго. Авторы исследования [408] полагают, что высокоразрешающее картирование и клонирование совместимых QTL, определяющих проявление свойства СЗЦ может помочь в получении засухоустойчивых гибридов.

5.3.3. Засоление

Растения возделываемых форм томата очень чувствительны к засолению как на стадии прорастания, так и на стадии вегетации. Однако его дикие родственные виды проявляют значительную устойчивость к засолению. Для изучения проблемы устойчивости томата к засолению QTL солеустойчивости были интродуцированы в инбредные линии, полученные от межвидового скрещивания солечувствительного *Lycopersicum esculentum* и солеустойчивого *L. pimpinifolium* [130, 131]. Авторы проанализировали 190 семей полученных из популяций семей возвратного скрещивания, насыщенных 151 RFLP маркером. Солевой стресс достигался посредством применения раствора содержащего

150 mM NaCl + 15 mM CaCl₂ (-0,85 МПа). Измерялись процент прорастания семян и признаки вегетативного роста. Результаты выявили наличие небольших, но важных корреляций ($r = -0,22$, $p < 0,05$) между уровнем прорастания семян и процентом выживания растений в условиях солевого стресса. Выявлено семь и пять QTL определяющих солеустойчивость в процессе прорастания семян и вегетативного роста, соответственно. Наряду с тем, что в большинстве случаев расположение QTL прорастания отличалось от локализации QTL вегетативного роста, наблюдались и некоторые совпадения месторасположения QTL, что согласовалось с небольшой фенотипической корреляцией, наблюдаемой между обоими признаками. В терминах практической селекции это означает, что возможно проведение одновременного улучшения признака устойчивости к засолению как на стадии прорастания семян, так и на стадии вегетативного роста растений. Всего было идентифицировано семь хромосомных сайтов, имевших существенные эффекты на проявление признака солеустойчивости. При этом образец *L. pimpinifolium* имел предпочтительные QTL аллели в шести сайтах. Процент фенотипической вариации, определяемый индивидуальными QTL, ранжировал от 6,5 до 15,6%, а совокупное действие всех существенных для проявления данного признака QTL составляло до 44,5% от общего фенотипического варьирования. В дополнение можно сказать, что всего идентифицировано 12 парных эпистатических взаимодействий, включая четыре между QTL-сцепленными и QTL-несцепленными районами, а также восемь между QTL-несцепленными локусами. Трансгрессивные фенотипы наблюдались в направлении солечувствительности. Авторы заключили [131], что высокое соответствие между фенотипами противоположных семейств и их QTL генотипами указывает на то, что солеустойчивость томата может быть улучшена с помощью маркерной селекции посредством межвидовой генетической изменчивости.

Citrus — это другой солечувствительный вид растений. В качестве первого шага по идентификации QTL, имеющих отношение к солевому стрессу, был рассмотрен эффект 40 mM NaCl на родительские разновидности *Poncirus trifoliata*, *Citrus grandis*, а также на их потомство обратного скрещивания и F₁ [381]. Анализ концентрации ионов Na⁺ и Cl⁻ в различных тканях помог идентифицировать 38 признаков как для усло-

вий засоленности, так и условий отсутствия засоления. Всего на основе достоверного LOD score ($> 3,0$) картировано 73, так называемых, потенциальных QTL. Определено семнадцать районов, 8 из которых с большим эффектом QTL. Многие QTL накопления ионов Na^+ и соотношения Cl^-/Na^+ образовали кластеры на отдельных группах сцепления.

Ряд проведенных исследований по изучению солеустойчивости у растений направлены на изучение более общих вопросов применимости, как в отношении любой неспецифичной популяции, так и в отношении найденного маркера, который должен быть ассоциирован с QTL в исследуемой популяции [129]. В отсутствие адекватных генов-кандидатов, определяющих солеустойчивость, генетики и селекционеры обычно используют подход соотнесения «локуса количественного признака/маркер вспомогательной селекции» (QTL/MAS). Так, например, в картированной обычным образом популяции риса удалось обнаружить вероятные маркеры ионного транспорта и селективности, идентифицированные при анализе AFLP. Однако девять из этих маркеров показали неспецифичную ассоциацию со схожими признаками в близкородственных популяциях рекомбинантных инбредных линий или при отборе сортов. Эти результаты свидетельствуют против любых ожиданий общей применимости маркеров по физиологическим признакам. Авторы заключили [129], что необходима конкретная информация о генах, вовлеченных в эти процессы. Путь по идентификации и подтверждению действия этих генов поможет установить условия, при которых QTL дифференцированно проявляют себя в результате экспериментального воздействия.

5.4. Идентификация генов-кандидатов

Как это ранее отмечалось разными авторами, использование QTL в качестве селективного признака может быть рискованным, особенно, если маркер не очень тесно связан с геном, отвечающим за изменение признака. Действительно, увеличение дистанции между маркером и соответствующим локусом увеличивает шанс прерывание связки из-за рекомбинации между ними. Таким образом, лучшая ситуация заключается в том, когда маркером является сам ген. Процесс выявления таких генов, объясняющих природу изменчивости признака, получил название

идентификация генов-кандидатов. Такой подход обычно состоит из двух фаз. Поиск гена-кандидата посредством оценки генов картированных около QTL локуса с целью нахождения гена, который сильнее всего генетически связан с признаком. Вторая фаза выявления гена-кандидата заключается в установлении того факта, что его молекулярный полиморфизм имеет эффект воздействия на генетическую изменчивость признака. Описанная выше процедура и есть процедура известная как позиционная идентификация гена-кандидата. Другой не исключающий позиционной идентификации метод (подход), названный «функциональная идентификация гена-кандидата», основывается на *a priori* выборе гена(ов), функционально связанного с признаком. Этот подход, который может пренебрегать QTL шагом, широко применяется в исследованиях генома человека. Корреляция между степенью проявления признака и аллельным полиморфизмом гена-кандидата обеспечивает строгий аргумент в оценке степени проявления гена-кандидата [102, 302]. У растений «функциональный подход» может быть проиллюстрирован примером определения роли полиморфизма *pr1* — гена, кодирующего транскрипционный фактор, который в свою очередь определяет содержание майзина в шелковидных нитеобразных пестиках початков неспелой кукурузы [62].

«Позиционный подход» предпочтителен в случае использования практически изогенных линий (NIL), которые отличаются друг от друга лишь отдельными интересующими исследователя зонами хромосом. Для получения посредством рекомбинации таких линий с редуцированным размером зон необходимо проведение возвратного скрещивания, что, в конечном счете, позволяет проводить прямую идентификацию гена. В ряде удачно поставленных экспериментов [135] было продемонстрировано, что желаемый аллельный эффект в QTL, отвечающем за содержание глюкозы и фруктозы у томата, определяется единичной нуклеотидной заменой в 484 пар нуклеотидном (п.о.) районе, охватывающем экзон и интрон инвертазы специфичной для плодов. Более точные исследования выявили, что этот район являлся горячей точкой рекомбинации. Важной особенностью этой работы явилось то, что QTL были определены не в результате использования RIL. Для идентификации QTL авторы, посредством системной интрогрессии фрагментов

хромосом дикого вида (*Lycopersicon pinnellii*) в культивируемый вид, а также рядом повторных возвратных скрещиваний и отбором линий, несущих различные хромосомные фрагменты, получили замещенные линии. Каждая линия являлась практически изогенной по отношению к культивируемой родительской форме за исключением интрогрессированного фрагмента. Позиция маркерного района была картирована, а локализация QTL легко определена идентификацией этих линий, представляющих собой эффект проявления генотипа по признаку.

5.5. QTL биохимических признаков и гены-кандидаты

5.5.1. QTL активности ферментов

Обычно гены, контролирующие агрономические признаки, трудно идентифицировать, в то время как гены-кандидаты, вовлеченные в биохимические количественные признаки, определяющие активность ферментов, концентрации различных субстратов или продуктов биохимических реакций *a priori* легче определять с помощью QTL. Например, когда QTL ферментной активности установлены, можно ответить на вопрос какой из этих QTL ко-локализован со структурным геном данного фермента? Этот подход впервые был применен в исследовании, в котором по биохимическим признакам анализировали сравнительно небольшую популяцию (всего 65 RIL) [65]. Авторы измеряли активности ключевых ферментов углеводородного метаболизма, а также концентрации их субстратов и продуктов реакций (сахарозы, глюкозы, фруктозы и крахмала) в листьях кукурузы верхних и нижних ярусов. В упомянутых выше RIL для каждого идентифицированного фермента, например, АДФ-глюкозопирофосфорилазы, сахарозосинтазы (нижние листья), сахарозофосфатсинтазы и двух инвертаз (нижние листья) были получены соответствующие кДНК и картированы соответствующие локусы QTL. Измерения делались на стадии 3-4 листьев в нескольких повторностях по каждому признаку. Бесспорная ко-локализация была установлена между количеством субстрата и продуктов реакций, а также QTL активности одного из следующих соответствующих структурных генов: сахарозофосфатсинтазной активностью и *Sps*-локусом на восьмой

хромосоме, а также сахарозосинтазной активностью и *Sh1*-локусом на хромосоме 9.

Последующий анализ родительских форм и изученных RIL кукурузы показал, что при потере воды одной из ранних реакций растений на данный стресс является повышение активности вакуолярной инвертазы, которая синхронизирована с увеличением содержания глюкозы и фруктозы, в то время как фотосинтетический уровень уменьшается немного позже [289].

Этот ответ на воздействие стресса оказался тесно связан с уровнем мРНК одного из генов инвертазы (*Ivr2*) кодирующего вакуолярную изоформу. Поскольку активность вакуолярной инвертазы изначально была различна у обоих родителей, то это открывало возможность поиска QTL имеющих отношение к водному стрессу, обусловленному данным биохимическим признаком. Кроме того, на тех же самых растениях авторы замерили различные фотосинтетические признаки такие как: водный статус листьев, а также рост корней и листьев. Для получения результата почти 120 RIL выращивались до стадии четвертого нормального листа, после чего полив на 9 дней приостанавливался и производились измерения. Эксперимент повторялся четыре раза. Во всех случаях вычислялись средние значения, что минимизировало влияние эффекта выращивания в теплице на определяемые показатели.

Несмотря на то, что QTL изучаемых признаков выявлены на всех 10 хромосомах, распределение QTL по хромосомам у кукурузы оказалось неслучайным. Как показал анализ полученных результатов, основные кластеры QTL сосредоточены на хромосомах 1, 4, 5, 6, 7 и 9. Кластер соотносился к данному хромосомному району, если более 3 QTL находились в пределах 10 сМ. Следуя этому определению, удалось установить более чем 12 таких кластеров, обычно содержащих QTL различных фенотипических классов (углеводородного метаболизма, фотосинтеза, водного стресса, роста и др.). Например, два кластера QTL содержащих 6 и 12 QTL из четырех классов, присутствовали на хромосомах 1 и 7, а шесть кластеров, содержащих QTL трех классов — локализованы на хромосомах 1, 4, 7 и 9. Из всех выявленных кластеров пять, расположенные на хромосомах 1, 4, 7 и 9, по-видимому, определяли реакцию растений на физиологический стресс. В противоположность это-

му, QTL многих морфологических признаков (длина, ширина, площадь и число листьев, высота растений, относительное содержание воды и транспирация) были картированы в тех же самых позициях в различных условиях произрастания. Такие QTL могут быть названы универсальными и можно высказать предположение, что это основные локусы физиологического гомеостаза. Возможное объяснение подобного рода кластерирования заключается в том, что эти же гены или близлежащие локусы контролируют схожие признаки. Это предсказуемо для признаков, принадлежащих той же фенотипической группе, например, длина, ширина и площадь листа, но менее очевидно, когда признаки принадлежат разным группам. В последнем случае можно предположить, что некоторые основные гены контролируют признаки на различных организационных уровнях. Другое общее наблюдение относится к кластерной композиции определенных групп признаков. Например, основная пропорция признаков роста оказалась ассоциирована с углеводородным метаболизмом в большей степени, чем с признаками фотосинтеза [289]. Это согласуется с тем фактом, что отбор по уровню фотосинтеза или Рубиско в целом не дает значительного улучшения урожайности [86].

Реакция растений на водный стресс может быть проанализирована сравнением распределения QTL схожих признаков, выявленных в как условиях наличия, так и отсутствия стресса. В результате проведенных экспериментов установлено, что количество QTL, которое может быть идентифицировано у растений в условиях стресса гораздо больше, чем в контроле. Так, например, 52% всех выявленных QTL были специфичны только для условий водного стресса и лишь 16% оказались общими как для контроля, так и для условий стресса. Причем основные стресс специфичные кластеры оказались локализованы лишь на хромосомах 1, 4, 5, 7 и 9. Некоторые отвечали только за физиологические признаки, другие в основном за морфологические либо как за те, так и за другие вместе взятые. Такое стресс-зависимое распределение означает, что в условиях недостатка воды различные наборы генов являются наиболее важными элементами адаптации и позволяют растениям приспособлять их метаболизм и рост к новым условиям окружающей среды. Кластерование QTL из различных классов, наблюдавшееся снова, по-видимому, означает то, что гены, контролирующие эти QTL, с малой вероятностью

кодируют одну и/или простую функцию, поскольку, как правило, несколько функций вовлечены в процесс адаптации к стрессу, вызванному недостатком воды. В этом случае чрезвычайно важна роль регуляторных генов. Однако из сравнения параметров поглощения двуокси углерода и содержания АК в условиях стресса вытекает явное несогласие с таким кластированием генов. Оба признака должны коррелировать, поскольку увеличение ксилемной АК предполагает закрывание устьиц. В действительности, ко-локализация QTL очень редка, и отмечалось лишь некоторыми авторами [214]. Возможное объяснение полученных результатов экспериментов, скорее всего, кроется в процедуре проведения анализа образцов [289]. Каждое растение анализировали спустя 9 дней после прекращения полива. В течение этого времени эффективная концентрация АК необходимая для закрывания устьиц, скорее всего, уже была достигнута.

Используя преимущества RFLP карты, основанной на маркерах хорошо согласующихся с известными функциями, особенно с метаболизмом углеводов [66], проведен сравнительный анализ локализаций QTL активности ферментов и соответствующих структурных генов, который выявил интересные ко-локализации, позволяющие применять установленные гены в качестве генов-кандидатов. Можно привести три примера таких генов-кандидатов. У кукурузы QTL активности АДФ-глюкозопирофосфорилазы листьев картирован в той же позиции на хромосоме 1, что и локус специфичного для листьев гена малой субъединицы этого же фермента [301]. На хромосоме 8 отмечена ко-локализация [65] активности SPS (sucrose-phosphate-synthase) и *Sps*-гена. Третьим примером может быть ко-локализация отмеченная для инвертазы и водного стресса. Как указывалось ранее, одна из первых реакций водного стресса в зрелых листьях — это существенное увеличение количества инвертазы в вакуолях и одновременно содержания гексозы. Данная реакция генотип-зависима и, вследствие этого, широко варьирует в RIL (трангрессионный эффект). Как результат, у кукурузы несколько QTL идентифицированы как в контроле (3 QTL), так и в условиях водного стресса (7 QTL) на хромосомах 1, 2 и с 5 по 10. Наиболее интересная локализация выявлена на хромосоме 5. В этом случае один QTL контроля и один водного стресса объясняли 17% и 5% изменчиво-

сти, соответственно, причем оба QTL, как оказалось, расположены очень близко к гену *Ivr2*, который кодирует вакуолярную инвертазу. Дальнейший анализ экспрессии этого гена в родительских линиях продемонстрировал, что *Ivr2* только один из шести известных генов инвертазы либо клеточной стенки, либо их вакуолярных форм, который, чтобы иметь такой же профиль индукции как и у вакуолярной инвертазы, должен быть индуцируем водным стрессом в листьях, представлять такую же генотип-зависимую экспрессию и ферментную активность, а также присутствовать в клетке в соответствующих удельных пропорциях. Все эти доводы поддерживают *Ivr2* в качестве гена-кандидата для инвертазных QTL кукурузы [288]. Другие QTL инвертазной активности были найдены недалеко от углеводородных QTL и некоторые из них образовывали совместные «стрессовые кластеры».

5.5.2. QTL анализа белков 2D-электрофорезом

Двумерный (2D)-электрофорез, также может быть использован для анализа различных количественных признаков у растений. Так, например, после проведения 2D-электрофореза, с помощью специальных компьютерных программ становится возможным количественно идентифицировать интенсивность отдельных белков (белковых электрофоретических зон, БЭЗ). Поскольку БЭЗ обычно являются генетически варьируемыми, несколько сотен зон потенциально представляют собой количественные признаки, проявляющиеся на биохимическом уровне, и могут быть использованы в этом качестве. Более того, существенный прогресс в секвенировании выделенных из геля полипептидов уже сейчас позволяет проводить идентификацию все большего и большего числа неизвестных ранее белков. Так, стало возможным одновременно определять QTL количества белка в БЭЗ (PQLs, Protein Quantity Loci) и, в случае, если белок в БЭЗ — фермент, QTL его ферментной активности, либо QTL соответствующего фенотипического признака. Ранее в результате проведенного биохимического анализа группой ученых [102] было установлено, что 78 белков из 413 исследованных имели достоверную количественную изменчивость (увеличение или ее уменьшение), а 38 из них проявляли различный экспрессионный уровень в обоих родительских генотипах. Количественный уровень одиннадцати белков

увеличивался в растениях в условиях стресса, а 8 белков обнаруживались только в подвергнутых стрессу растениях, при этом отбор велся на основе результатов микросеквенирования аминокислотных последовательностей, находящихся внутри аминокислотной цепи. Сравнительный анализ аминокислотных последовательностей установил, что 16 из секвенированных белков симилярны идентифицированным ранее белкам. В дополнение к уже известным, вовлеченным в реакцию водного стресса белкам (например, RAB17, который подвержен влиянию АК), были идентифицированы несколько ферментов (например, элолаза и триозофосфатизомераза) вовлеченных в основной метаболизм растительных организмов, включая гликолиз и цикл Кребса. Также были установлены и некоторые другие ферменты, такие, например, как кафеат-О-метилтрансфераза, индуцирование которой связано с процессом одревеснения растений. Впервые методология PQL проиллюстрирована на примере фосфоглицератмутаза и некоторых других белков кукурузы, сверхэкспрессия которых наблюдалась в условиях засухи [102].

Заключение

При рассмотрении влияния различных видов стрессоров и различных стрессовых условий, включая общие наблюдения по целому ряду признаков, становится ясно, что QTL распределены по геному не случайным образом, а довольно часто образуют кластеры. Такое кластерирование проявляется не только для коррелирующих свойств на уровне организации, которая может быть тривиальной, но и на уровне свойств и признаков различных организационных уровней. Это наблюдение подтверждено рядом авторов [5-7, 14, 19, 165, 394]. Теоретически такая точка зрения была сформулирована в конце 70-х годов XX века [96]. Эта теория [96] вводила так называемое понятие «linkat» как «набор локусов, которые агрегируют в одном и том же хромосомном секторе в течение дифференцировки видов. Эти linkats (=кластеры) проявляют сильный эпистатический эффект и в целом проявляют или представляют ко-адаптированные функции». Каждый linkat должен быть защищен против собственного разобщения низким уровнем рекомбинации, а его содержание должно быть наполнено генными дупликациями таким образом, чтобы все это поддерживало функциональные свойства

linkat в ряду поколений. Наши знания по организации геномов, накопленные к сегодняшнему моменту, хорошо согласуются с высказанными ранее [96] теоретическими выкладками.

В условиях стресса проявляются новые кластеры QTL, указывая на то, что новые гены или группы генов становятся более важными в объяснении изменчивости признаков. Это наблюдалось, например, у подвергнутой водному стрессу кукурузы, для некоторых районов хромосом 1 и 5. Это означает, что при отборе на устойчивость к стрессу, наиболее благоприятный аллель не будет найден в том же локусе, как и при условиях отсутствия стресса. Хотя накопленных до сегодняшнего дня данных явно не достаточно, это, вне всякого сомнения, заслуживает внимания для проверки в отношении некоторых видов растений, подвергаемых различным стрессам (водному стрессу, засолению и т.п.), особенно если некоторые QTL стресса окажутся общими. Например, некоторые биохимические процессы, такие как реакции антиокисления, являются частью условий возникновения целого ряда стрессов. В любом случае, очень важным элементом в достижении прогресса является идентификация генов, соответствующих определенным QTL. Методология генов-кандидатов дает средство для достижения поставленной цели. Для стресса этот подход в основном ограничивается структурными генами метаболизма углеводов. Хотя ко-локализация для одного признака может наблюдаться для трех ферментов, наиболее близких к QTL, которые расположены повсеместно. Эти локусы, как представляется, вовлечены в стресс-зависимую регуляцию, и поэтому очень важно идентифицировать их.

Глава VI. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ ГЕНЕТИЧЕСКОГО КАРТИРОВАНИЯ И АНАЛИЗА АССОЦИАЦИЙ МАРКЕР/ПРИЗНАК

Для того чтобы получить необходимую информацию о месторасположении QTL в геноме проводят их молекулярно-генетическое картирование, в результате которого определяют не только их локализацию, но и то, от какого родителя получен тот или иной аллель QTL, процент фенотипической изменчивости, определяемый каждым из выявленных QTL, и, конечно же, молекулярный маркер или маркеры с которыми сцеплены идентифицированный локус или локусы на группах сцепления. Одним из важнейших требований для картирования QTL является возможность производить обмеры и статистический анализ вариантов исследуемых признаков для установления достоверных корреляций между QTL и тем или иным маркером. Подробности применения таких подходов были описаны нами ранее [8, 13].

Еще одним из необходимых требований, на основе которого строятся современные генетические карты — это наличие ассоциаций маркер-признак и неравновесного сцепления (LD) между ними [14, 15, 367]. LD связано с исторически или эволюционно неравным и разнородным уровнем рекомбинации специфичных аллелей в различных локусах, определяющих проявление изучаемой генетической изменчивости в популяции. Неравновесное сцепление может быть установлено статистически и его выявление уже широко используется для картирования и даже клонирования генов, определяющих сложные признаки у человека, животных и растений [14, 16, 18]. С изменением направления вектора усилий по идентификации и локализации генов/локусов хромосом от дуродительского скрещивания к естественным популяциям линий или образцов коллекций ГРР, и от традиционного QTL картирования к LD ассоциативному анализу, исследователи получили новый мощный инструмент по картированию интересующих их генов. Ассоциативное или LD картирование позволяет не только проводить анализ, классифицировать и группировать естественные и селекционные популяции (дикие виды, примитивные формы, сорта, линии), но и более эффективно ис-

пользовать их в практической селекции с целью улучшения возделываемых культур.

Построение генетических карт, с использованием для этих целей молекулярных маркеров, и применение созданных карт для анализа ассоциаций маркер/признак, являются двумя предварительными условиями, требуемых для реализации маркер вспомогательной селекции (marker assisted selection, MAS). Отличительной чертой проведения генетического картирования и установления ассоциаций маркер/признак, является оперирование большими объемами данных, полученных в результате экспериментальной оценки различных классов генетических маркеров и различного рода картирующих популяций. Для оперативного решения этой проблемы на сегодняшний день существует большое число статистических методологий и подходов, а также информационно-программных средств, подходящих для анализа различных типов популяций и маркеров.

В настоящей главе приведено краткое описание предназначений основных компьютерных программ, наиболее широко используемых для проведения генетического картирования и установления ассоциаций маркер/признак. Без точной, статистически выверенной информации о такого рода ассоциациях невозможно проведение маркер-вспомогательной селекции, что в свою очередь, как ожидается, будет способствовать ускорению селекционного процесса, включая создание новых сортов с желаемыми характеристиками, в том числе для их применения в системе точного земледелия. В то же время мы надеемся, что в ближайшее время будут разработаны и использованы в практических целях новейшие программные обеспечения, которые позволят более эффективно на современном информационном уровне находить и оценивать ассоциации маркер/признак в новых типах популяций.

6.1. Построение генетической карты

Генетическая карта может быть построена посредством использования сегрегирующих популяций различных типов для видов с различным уровнем ploидности. Первым и наиболее часто используемым программным обеспечением для построения карт является MAPMAKER/EXP, которое было разработано в 1987 году [212]. Почти

все молекулярные карты, основанные на первом поколении молекулярных маркеров (RFLP), были построены с помощью этого программного продукта. Альтернативой является MAP MANAGER CLASSIC — графическая, интерактивная программа для картирования менделевских локусов. Она проводит все необходимые расчеты и построения на основе информации, полученной при межвидовых скрещиваниях с доминантными маркерами, беккроссных (backcross, BC) или рекомбинантных инбредных линий (RIL) у экспериментальных растений или животных ([237]; <http://www.mapmanager.org/mapmgr.html>).

Иногда для построения карт могут быть нужны некоторые специальные статистические модификации, при этом используют маркеры со строгим расхождением с сегрегацией в популяциях. MAPDISTO (<http://mapdisto.free.fr/>) — программа для картирования генетических маркеров в случае сегрегационного расхождения при использовании таких расщепляющихся экспериментальных популяций как BC, удвоенные гаплоиды (double haploid, DH) и RIL. Она может: (1) рассчитывать и рисовать генетические карты посредством графического интерфейса и (2) облегчать анализ маркерных данных, экспонируя сегрегационные расхождения благодаря различиям в жизнеспособности гамет или зигот.

Карты или данные от множественных популяций, полученных в результате различных скрещиваний, могут быть объединены в одну или консенсусную карту посредством совместного картирования (joint mapping). JOINMAP представляет собой программное обеспечение для построения карт генетического сцепления для нескольких типов картирующих популяций: BC₁, F₂, RIL, и DH, регенерированные из F₁ и F₂ поколений, а также семейств полных сибсов, полученных в результате скрещиваний неблизкородственных родителей (<http://www.kyazma.nl/index.php/mc.JoinMap/>). Она может объединять (соединять — join) данные, полученные из нескольких источников в интегрированную карту и реализовывать некоторые другие функции, включая определение групп сцепления, автоматическое определение фазы для семейств полных сибсов, полученных в результате скрещиваний неблизкородственных родителей, некоторые построения диагностических графиков и карт [390].

Программным обеспечением со сравнительными функциями является СМАР, которое было разработано, как основанное на использовании интернет-технологий, средство, позволяющее пользователям видеть сравнения генетических и физических карт. Пакет программы включает в себя инструменты для исправления данных карт (<http://www.gmod.org/wiki/СMap>; [399]).

6.2. Картирование QTL, основанное на сцеплении

В экспериментах по картированию было продемонстрировано сцепление/ассоциации между целевыми признаками/генами и молекулярными маркерами, которое основывается на генетическом сцеплении и неравновесном сцеплении (LD). Средства технической поддержки, необходимые для установления ассоциаций генотип-фенотип, обычно включают в себя: (1) статистические методы и программные средства для установления, оценки и сравнения ассоциаций генотип-фенотип посредством картирования сцепления, LD или ассоциативного картирования и картирования *in silico*, использующие для этого единообразные популяции, множественные популяции или все генетические ресурсы со всей доступной информацией обо всех разнообразных испытаниях по годам, временам года и местам их проведения; (2) статистические методы и программные средства для идентификации эффектов генетического фона или генетической среды окружения, аллелях локусов количественных признаков (QTL) во множественных локусах и множественных аллелях в локусе; (3) программные средства, облегчающие процесс оценки от сцепленных маркеров до функциональных маркеров и генов-кандидатов; и (4) программные средства, облегчающие управление генетическими популяциями, картами и сопутствующими маркерными и фенотипическими данными.

Существует много коммерческих и находящихся в свободном доступе пакетов программного обеспечения, предназначенных для установления ассоциаций между маркерными генотипами и фенотипами признаков. Наиболее распространенными и часто используемыми являются QTL Cartographer, MAPQTL, PLABQTL и QGENE. Все они оперируют только с двуродительскими популяциями, в то время как MCQTL [190] также проводит QTL картирование в многоаллельных ситуациях,

включая двуродительские сегрегирующие популяции, а также наборы (серии, ряды) двуродительских или двуаллельных популяций. Наиболее часто используемым программным обеспечением в 1980-1990-х годах было MAPMAKER/QTL, являющее сестринской программой MAPMAKER/EXP, разработанной [212] (http://www-genome.wi.mit.edu/genome_software). Это ПО основывается на установлении максимального правдоподобия сцепления между маркером и фенотипом, используя для этого интервальное картирование, которое имеет дело с несложным QTL и несколькими стандартными популяциями. Другое раннее программное обеспечение MAPL (MAPping and QTL analysis; <http://lbn.ab.a.u-tokyo.ac.jp/software.html>; [385]) позволяет пользователям получать результаты по уровню расщепления, тесту сцепления, значению рекомбинации, группировать маркеры, располагать маркеры в определенном порядке посредством метрического многомерного масштабирования, рисовать карту и графически изображать генотип и карту QTL посредством интервального картирования и дисперсионного анализа (analysis of variance, ANOVA).

Широко используемым для картирования QTL является программное обеспечение QTL Cartographer (<http://statgen.ncsu.edu/qtlcart/cartographer.html>), которое обеспечивает выполнение нескольких статистических методов, используя множество маркеров одновременно, включая сложное интервальное картирование (CIM) и множественное композитное интервальное картирование (англ. multiple composite interval mapping, MCIM). Взаимодействие между идентифицированными QTL также может быть установлено. PLABQTL использует составное интервальное картирование со многими функциями схожими с QTL Cartographer. QTL может быть локализовано и охарактеризовано в популяциях, полученных самоопылением потомков двуродительского скрещивания, или получением линий удвоенных гаплоидов (DH). Простое и составное интервальное картирование производят, используя модель множественной регрессии. В качестве дополнительной функции ко многим другим пакетам ПО, может быть использован анализ взаимодействия QTL-окружающая среда [386].

Программа QGENE (<http://www.qgene.org/>) предназначена для проведения сравнительного анализа наборов данных по картированию

QTL посредством вычислений, наиболее часто используемых пользователем. Кроме того, она написана вместе с дополнительным программным модулем построения (структурирования) программной системы, что позволяет ее расширять. QGENE начиналась в 1991 году как программа для картирования и моделирования популяций, в которую потом была добавлена возможность QTL анализа. Сравнительно недавно QGENE была переписана на языке Java, что позволяет ее использование на любом компьютере с любой оперативной системой. Она предлагает наиболее общепринятые методы QTL картирования и позволяет их параллельное (одновременное) сопоставление. Ее интерфейс по желанию может быть описан на языке пользователя, для этого необходимо самостоятельно написать трансляционный файл к программе. QGENE может быть использована для анализа признаков, QTL и пермутаций (математических перестановок), а также моделирования популяций и признаков.

Некоторые пакеты ПО могут быть использованы для построения карт сцепления у отдаленных скрещиваемых видов растений. Например, программа ONEMAP, созданная на языке программирования R, предоставляет такие возможности. Для этого используют семейства полных сибсов, полученных в результате скрещиваний неблизкородственных родителей (<http://cran.r-project.org/web/packages/OneMap/index.html>; [144]). Другим ПО является MAPQTL — программное обеспечение, предназначенное для расчета позиций QTL (<https://www.kyazma.nl/index.php/MapQTL/>; [391]), которое может быть использовано для нескольких типов картирующих популяций, включая BC₁, F₂, RIL, DH и семейства полных сибсов, полученных в результате скрещиваний неблизкородственных родителей. Она может быть применена для картирования QTL посредством интервального, композитного интервального и непараметрического картирования, с функцией для автоматического кофакторного отбора и проведения теста пермутации.

Несколько компьютерных программ для картирования рассматривают эпистаз (взаимодействия неаллельных генов) при картировании QTL. EPISTACY — это SAS программа, созданная для тестирования всех возможных двулокусных комбинаций для эпистатических взаимодействий или эффектов количественных признаков. Программа является

настоящим SAS программным шаблоном, который пользователи должны настроить для того, чтобы она подходила к их данным. В простейших случаях пользователи могут изменить только названия файлов, содержащих их данные. Однако алгоритмы программы основаны на методе наименьших квадратов и нет возможности проводить интервальное картирование [174].

Байесовскому картированию QTL в последние годы было уделено много внимания и, благодаря этому, на его основе были разработаны специальные компьютерные программы. Например, программное обеспечение BQTL (байесовское картирование локусов количественных признаков — Bayesian Quantitative Trait Locus mapping) было разработано для картирования генетических признаков в популяциях, полученных от скрещивания линий и в RIL (<http://hacuna.ucsd.edu/bqtl>; [45]). Оно выполняет: (1) установление максимального правдоподобия в мультигенных моделях; (2) байесовское оценивание мультигенных моделей посредством аппроксимаций Лапласа; и (3) интервальное и композитное интервальное картирование генетических локусов. Программа BLADE (байесовское картирование неравновесного сцепления — Bayesian LinkAge DisEquilibrium mapping) была создана для байесовского анализа гаплотипов с целью неравновесного LD (linkage disequilibrium) картирования (<http://www.people.fas.harvard.edu/~junliu/TechRept/03.html>; [225, 227]). MULTIMAPPER — программное обеспечение для байесовского картирования QTL при анализе BC, DH и данных поколения F₂, полученного в результате скрещивания инбредных линий [248]. Программа MULTIMAPPER/OUTBRED дополнила этот список популяциями, полученными от беспородных линий (<http://www.rni.helsinki.fi/~mjs/>).

Несколько пакетов ПО по картированию было разработано для картирования QTL, необходимых в особых специфичных случаях. Программа MCQTL разработана для одновременного картирования QTL во множественных скрещиваниях и популяциях (<http://www.genoplante.com>; [190]). Она позволяет анализировать обычные популяции, полученные в результате скрещивания инбредных линий, и может соединять (связывать) семейства при допущении, что месторасположение QTLу всех у них одно и то же. Более того, возможно также диаллельное моделиро-

вание эффектов QTL, если используют множественные родственные семейства. Программа MAPPOP была предложена для селективного картирования и картирования при выбранном интервале значений случайной величины (bin mapping), посредством выбора хороших образцов из картирующих популяций, и для локализации новых маркеров на предсуществующих картах [393]. В дополнение была разработана программа QTLNetwork, для картирования и визуализации генетической архитектуры, подчеркивающей сложные признаки, в экспериментальных популяциях, полученных в результате скрещивания двух инбредных линий (<http://ibi.zju.edu.cn/software/qtlnetwork>; [410]).

Как только базирующиеся на Интернет–технологиях программные средства стали все более и более востребованы, использующее Интернет–технологии QTL аналитическое программное обеспечение стало доступным. В качестве примера можно привести программу WEBQTL, созданная как интерактивная Интернет страница, полезная для изучения и анализа генетической модуляции тысяч фенотипов, собранных за более чем 30-летний период сотнями исследователями, используя референсные панели рекомбинантных инбредных штаммов мышей (<http://www.genenetwork.org/home.html>). WEBQTL включает плотные, проверенные на ошибки генетические карты, так же как и наборы данных экстенсивной экспрессии генов (фирмы Affymetrix), полученные в результате анализов более чем 35 штаммов мышей. Используя интернет-технологии удобный для пользователя пакет QTL EXPRESS (<http://qtl.car.ed.ac.uk>; [333]), предназначенный для картирования QTL в беспородных популяциях, был разработан для скрещенных линий, семейств полусибс, а также ядерных семейств и сибсовых пар. Этот пакет предлагает две опции для тестов значимости QTL: тесты пермутации для определения эмпирических уровней значимости и метод ступенчатого перехода от малых значений к большим или наоборот (bootstrapping) для установления эмпирических доверительных интервалов локализации QTL. Фиксированные эффекты/ковариаты (независимые переменные) могут быть подогнаны, а модели могут включать одиночные или множественные QTL.

6.3. Картирование посредством установления неравновесного сцепления

В последнее десятилетие ассоциативное или LD картирование стало очень популярным. Оно использует неструктурированные популяции, содержащие неродственных особей, образцы коллекций генетических ресурсов или случайным образом отобранные сорта [14, 16]. До ассоциативного картирования генотипированные единицы подвергаются статистическому анализу для удаления наиболее существенного фактора — популяционной структуры, которая может быть причиной ложноположительных ассоциаций, возникающих в большей степени из-за зависящих от обстоятельств, побочных корреляций, чем за счет настоящего сцепления. Для этих целей может быть использована программа STRUCTURE (<http://pritchardlab.stanford.edu/structure.html>; [304]). Для ассоциативного картирования были разработаны некоторые пакеты ПО, в которые функциональные возможности анализа структуры популяций уже включены. STRAT, как программа, дополняющая STRUCTURE, использует структурированный ассоциативный метод для ассоциативного картирования разрешающий достоверные (действенные) методы изучения «случай-контроль» даже при наличии структуры у популяции (<https://web.stanford.edu/group/pritchardlab/software/STRAT.html>; [305]).

6.3.1. Картирование *QTL*, основанное на *LD*

TASSEL — это универсальная программа для анализа признаков посредством установления ассоциаций, эволюционных отношений и сцепления, которая выполняет различные генетические анализы, включая ассоциативное картирование, установление разнообразия и расчет неравновесного сцепления (<http://sourceforge.net/projects/tassel/>; [423]). Анализ LD между генотипами и фенотипами может быть выполнен либо посредством общей линейной модели, либо смешанной линейной модели. Общая линейная модель позволяет пользователям анализировать сложные (комплексные) полевые исследования, взаимодействия с окружающей средой и эпистаз. Смешанная модель разработана для обработки полигенных эффектов на множественных уровнях связанностей, включая информацию родословной. Эти анализы позволяют проводить установление и исследование LD у многих видов растений и животных.

Другие пакеты программного обеспечения включают в себя программу мультиаллельного межаллельного неравновесного анализа (Multiallelic Interallelic Disequilibrium Analysis Software, MIDAS), которая была разработана для анализа и визуализации межаллельного неравновесия между мультиаллельными маркерами (<http://www.oege.org/software/midas/index.shtml>, [146]) и PEDGENIE (<http://www-genepi.med.utah.edu/Genie/PedGenieDetail.html> [23]), которая создана как универсальное средство для анализа ассоциаций и перенос (распределение) неравновесия между генетическими маркерами и признаками в семьях произвольного размера и структуры. В эту программу может быть введена любая родословная любого размера от отдельных независимых особей до больших генеалогий. Анализ независимых особей и семей может проводиться совместно.

GENERECON (<http://www.daimi.au.dk/~mailund/GeneRecon/>) — другая программа, которая применяет для ассоциативного картирования теорию сращивания (слияния). Она основывается на методе Монте-Карло по байесовской схеме цепи Маркова для тонкого ассоциативного картирования, используя при этом высокоплотные маркерные карты у животных. GENERECON предлагает недвусмысленные модели генеалогии образца в случае родственных хромосом по локусу, обуславливающему заболевание. Исследуемый случай и данные контроля представляются в виде информации о генотипе или гаплотипе, при этом программа производит оценку ряда параметров, наиболее важный из которых установление позиции локуса заболевания на хромосоме [233].

6.3.2. Полногеномное ассоциативное картирование

Исследования по полногеномным ассоциациям (genome-wide association, GWA) на сегодня довольно часто применяют для нахождения связи между генетической изменчивостью и общими заболеваниями у человека, а также агрономически важных признаков у растений. В идеале хорошо оснащенные GWA исследования включают в себя измерения сотни тысяч полиморфизмов, выявляемых на уровне единичного нуклеотида (single nucleotide polymorphism; SNP) у тысяч особей. Чистый объем данных, генерируемый этими экспериментами, порождает очень высокие аналитические запросы. В ней есть ряд важных этапов,

которые необходимо пройти в процессе анализа этих данных, многие из которых представляют собой критические моменты, ограничивающие производительность системы. Данные должны быть загружены и анализированы для выполнения изначального контроля качества до начала процедуры тестирования неравновесного сцепления. Оценка результатов может включать последующий статистический анализ, например, тест пермутаций, или последующий контроль качества ассоциированных маркеров, например, проверка недоработанных интенсивностей генотипирования. Наконец, существенные ассоциации должны быть расположены в соответствии с приоритетом, используя для этого функциональные и биологические методы интерпретации, просмотр и редактирование имеющихся в распоряжении биологических примечаний и пояснений, информационных путей и модели неравновесного сцепления [294]. GOLDSURFER2 (GS2) универсальная программа для анализа и визуализации GWA исследований была предложена ([294]; <http://www.well.ox.ac.uk/gs2>). GS2 — это интерактивное, удобное для пользователя графическое приложение, которое может быть использовано на любых этапах GWA проекта от контроля качества исходных данных и анализа биологической оценки до проверки достоверности полученных результатов. Программа составлена на языке Java и может быть использована на всех платформах. С помощью GS2 очень большие наборы данных (например, 500000 маркеров и 5000 образцов) могут быть качественно оценены, быстро проанализированы и интегрированы в информацию о геномных последовательностях. При этом могут быть отображены и функционально оценены кандидаты SNP.

Другие программные средства, предназначенные для исследований GWA, включают: GENOMIZER — не зависящая от типа операционной системы Java программа, определенная для анализа GWA экспериментов (<http://www.ikmb.uni-kiel.de/genomizer>), PLINK — набор инструментальных средств программирования, разработанный для полногеномного анализа неравновесного сцепления (<http://zzz.bwh.harvard.edu/plink/>; [306]), MAPBUILDER — созданная для широкого хромосомного ассоциативного картирования (<http://bios.ugr.es/BMapBuilder>; [20]) и калькулятор расчета степени ассоциаций с двустадийной композицией (Calculator for Association with Two Stage design, CATS), который рас-

считывает степень и другие необходимые величины для двустадийных GWA исследований (<http://csg.sph.umich.edu/abecasis/CaTS/>; [342]).

Результаты масштабных GWA исследований все чаще и чаще размещаются в общедоступных базах данных. Но доступное программное обеспечение для анализа и интерпретации наборов GWA данных часто может быть затруднительным для использования [57]. В настоящее время существует острая потребность в разработке усовершенствованных программных обеспечений для обработки наборов GWA данных, результаты которых было бы достаточно легко анализировать и исследователям проще обмениваться ими друг с другом, а также необходимо повысить качество таких программных продуктов, которые имели бы дело с увеличивающимся расчетными запросами, предъявляемыми этими наборами данных.

Заключение

Таким образом, исходя из выше изложенного, можно заключить, что на сегодняшний день существует большое количество разнообразных и разноплановых программных обеспечений, позволяющих проводить не только статистическую обработку данных, строить генетически карты, но и довольно эффективно устанавливать, а также оценивать ассоциации маркер/признак. И хотя картирование с помощью молекулярных маркеров требует специализированной технической оснащенности (например, программного обеспечения, лабораторного оборудования), а также опыта, знаний и компетентности в понимании разрешающей способности того или иного метода, включая интерпретацию полученных результатов, совершенно очевидно, что в перспективе исследователи для того, чтобы прокартировать интересующий их ген или локусы хромосом, не будут нуждаться в том, чтобы начинать с нуля широкомасштабный физиолого-генетический эксперимент по идентификации и анализу QTL или установлению ассоциаций маркер/признак, а смогут оперировать существующими, постоянно пополняемыми базами данных и современным программным обеспечением, позволяющими проводить статистически достоверный и оперативный анализ внутригеномных взаимодействий.

СПИСОК ЛИТЕРАТУРЫ

1. **Артемьева А.М., Чесноков Ю.В., Клоке Э.** Морфолого-географический и молекулярно-генетический анализ коллекции белокочанной капусты ВИР. *Доклады Российской Академии Сельскохозяйственных Наук*. 2008. 5: 14-18.
2. **Артемьева А.М., Чесноков Ю.В., Клоке Э.** Генетическое разнообразие и внутривидовые филогенетические взаимоотношения культур вида *Brassica rapa* L. по результатам анализа микросателлитов. *Информационный Вестник ВОГиС*. 2008. 12(4): 608-619.
3. **Вавилов Н.И.** Закон гомологических рядов в наследственной изменчивости. М.-Л. 1935.
4. **Вишнякова М.А., Бурляева М.О., Алпатьева Н.В., Чесноков Ю.В.** RAPD-анализ видового полиморфизма рода Чина (*Lathyrus* L.) сем. Fabaceae Lindl. *Информационный Вестник ВОГиС*. 2008. 12: 595-607.
5. **Драгавцев В.А.** К проблеме генетического анализа полигенных количественных признаков растений. СПб. ВИР. 2003. 35 с.
6. **Жученко А.А.** Экологическая генетика культурных растений (адаптация, рекомбиногенез, агробиоценоз). Кишинев, Штиинца. 1980. 588 с.
7. **Жученко А.А., Король А.Б.** Рекомбинация в эволюции и селекции. М.: Наука. 1985. 400 с.
8. **Кочерина Н.В., Артемьева А.М., Чесноков Ю.В.** Использование лод-оценки в картировании локусов количественных признаков у растений. *Доклады Российской Академии Сельскохозяйственных Наук*. 2011. 3: 14-17.
9. **Рыжова Н.Н., Бурляева М.О., Кочиева Е.З., Вишнякова М.А.** Использование ITS последовательностей для оценки таксономических отношений у представителей трибы Viciae (Adans.) Bronn сем. Fabaceae Lindl. *Экологическая генетика*. 2007. 3: 5-12.
10. **Чесноков Ю.В.** ДНК фингерпринтинг и анализ генетического разнообразия у растений. *Сельскохозяйственная биология*. 2005. 1: 20-40.
11. **Чесноков Ю.В.** Генетические ресурсы растений и современные методы ДНК-типирования. СПб, ВИР. 2007а. 80 с.
12. **Чесноков Ю.В.** Генно-инженерные манипуляции у растений и их естественная основа. СПб.: ВИР. 2007б. 80 с.
13. **Чесноков Ю.В.** Картирование локусов количественных признаков у растений. ВИР, СПб. 2009. 100 с.
14. **Чесноков Ю.В.** Молекулярно-генетические маркеры и их использование в предселекционных исследованиях. СПб: АФИ. 2013. 116 с.

15. **Чесноков Ю.В.** Разновидности сцепления генетических маркеров с целевым геном и локусами хромосом. *Агрофизика*. 2018. 2: 40-45.
16. **Чесноков Ю.В., Артемьева А.М.** Ассоциативное картирование у растений. *Сельскохозяйственная биология*. 2011. 5: 3-16.
17. **Чесноков Ю.В., Косолапов В.М.** Генетические ресурсы растений и ускорение селекционного процесса. Москва: ООО «Угрешская типография», 2016. 172 с.
18. **Чесноков Ю.В., Почепня Н.В., Бёрнер А., Ловассер У., Гончарова Э.А., Драгавцев В.А.** Эколого-генетическая организация количественных признаков растений и картирование локусов, определяющих агрономически важные признаки у мягкой пшеницы. *Доклады Академии Наук*. 2008. 418(5): 693-696.
19. **Чесноков Ю.В., Почепня Н.В., Козленко Л.В., Ситников М.Н., Митрофанова О.П., Сюков В.В., Кочетков Д.В., Ловассер У., Бёрнер А.** Картирование QTL, определяющих проявление агрономически и хозяйственно ценных признаков у яровой мягкой пшеницы (*Triticum aestivum* L.) в различных экологических регионах России. *Вавиловский журнал генетики и селекции*. 2012. 16(4/2): 970-986.
20. **Abad-Grau M.M., Montes R., Sebastiani P.** Building chromosome-wide LD maps. *Bioinformatics*. 2006. 22: 1933–1934.
21. **Adams W.T.** Gene dispersal within forest tree populations. *New Forests*. 1992. 6: 217-240.
22. **Ajmone-Marsan P., Castiglioni P., Fusari F., Kuiper M., Motto M.** Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. *Theor. Appl. Genet.* 1998. 96: 219-227.
23. **Allen-Brady K., Wong J., Camp N.J.** PedGenie: an analysis approach for genetic association testing in extended pedigrees and genealogies of arbitrary size. *BMC Bioinformatics*. 2006. 7: 209.
24. **Arnheim N., Strange C., Erlich H.** Use of pooled DNA samples to detect linkage disequilibrium of polymorphic restriction fragments and human disease, Studies of the HLA class II loci. *Proc. Natl. Acad. Sci. USA*. 1985. 92: 6970-6974.
25. **Arondel V., Lemieux B., Hwang I., Gibson S., Goodmann H.M., Somerville C.R.** Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. *Science*. 1992. 258: 1353-1355.
26. **Ashley M.V., Dow B.D.** The use of microsatellite analysis in population biology: background, methods and potential application. In: *Molecular Ecology and Evolution*. Schierwater B., Streit B., Wagner G.P., DeSalle R. (eds.). Birkhauser, Verlag. 1994. pp.185-201.

27. **Asins M.J., Carbonell E.A.** Detection of linkage between restriction fragment length polymorphism markers and quantitative traits. *Theor. Appl. Genet.* 1988. 76: 623-626.
28. **Austerlitz F., Smouse P.E.** Two-generation analysis of pollen flow across a landscape. II. Relation between Φ_{ft} pollen dispersal and interfemale distance. *Genetics.* 2001. 157. 851-857.
29. **Bar-Hen A., Charcosset A.** Relationship between molecular and morphological distances in a maize inbred lines collection: Application for breeders' right protection. In: *Biometrics in Plant Breeding: Application of Molecular Markers. Proc. 9th Meeting of the Eucarpia Section Biometrics in Plant Breeding.* Van Oijen J.W., Jansen J. (eds.). 1994. pp.57-66.
30. **Bark O.H., Heavey M.J.** Similarities and relationships among populations of the bulb onion estimated by RFLPs. *Theor. Appl. Genet.* 1995. 90: 407-414.
31. **Beavis W.D., Grant D., Albertsen M., Fincher R.** Quantitative trait loci for plant height in for maize populations and their associations with qualitative genetic loci. *Theor. Appl. Genet.* 1991. 83: 141-145.
32. **Beavis W.D., Keim P.** Identification of quantitative trait loci that are affected by environment. In: *Genotype by Environment Interaction.* Kang M.S., Gauch H.G. (eds.). CRC Press Inc. 1996. pp.123-149.
33. **Bernacchi D., Beck-Bunn T., Emmatty D., Eshed Y., Inai S., Lorez J., Petiard V., Sayama H., Uhlig J., Zamir D., Tanksley S.D.** Advanced backcrossQTL analysis in tomato. I. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL-alleles derived from *Lycopersicon hirsutum* and *L.pimpinellifolium*. *Theor. Appl. Genet.* 1998. 97: 170-180.
34. **Bernacchi D., Beck-Bunn T., Eshed Y., Lorez J., Petiard V., Uhlig J., Zamir D., Tanksley S.D.** Advanced backcross QTL analysis in tomato. II. Identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum*. *Theor. Appl. Genet.* 1998. 97: 381-397.
35. **Bernardo R.** Estimation of coefficient of coancestry using molecular markers in maize. *Theor. Appl. Genet.* 1993. 85: 1055-1062.
36. **Bernardo R.** Prediction of maize single cross performance using RFLPs and information from related hybrids. *Crop Sci.* 1994. 34: 20-25.
37. **Bernardo R., Kahler A.I.** North American study on essential derivation in maize: inbreds developed without and with selection from F₂ populations. *Theor. Appl. Genet.* 2001. 102: 986-992.
38. **Bernatzky R., Tanksley S.D.** Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. *Genetics.* 1986. 112: 887-898.
39. **Bezant J., Laurie D., Pratchett N., Chojecki J., Kearsey M.** Marker regression mapping of QTL controlling flowering time and plant height

- in spring barley (*Hordeum vulgare* L.) cross. *Heredity*. 1996. 77: 64-73.
40. **Bhattacharyya M.K., Smith A.M., Ellis T.H.N., Hedley C., Martin M.** The wrinkled seed character of pea described by Mendel is caused by a transposon-like insertion in a gene encoding starch-branching enzyme. *Cell*. 1990. 60: 115-122.
 41. **Birky C., Jr.** Evolution and variation in plant chloroplast and mitochondrial genomes. In: *Plant Evolutionary Biology*. *Gottlieb L.D., Jain S.K. (eds.)*. Sinauer Associates. 1988. pp.21-46.
 42. **Birky C., Jr.** Evolution and population genetics of organelle genes: mechanisms and models. In: *Evolution at the Molecular Level*. *Selander R.K., Clark A.G., Whittam T.S. (eds.)*. Sinauer Associates. 1991. pp.112-134.
 43. **Birky C.W., Fuerst P., Maruyama T.** Organelle gene diversity under migration, mutation and drift: equilibrium expectations, approach to equilibrium, effect of heteroplasmic cells, and comparison to nuclear genes. *Genetics*. 1989. 121: 613-627.
 44. **Boppenmeier J., Melchinger A.E., Brunklaus-Jung E., Geiger H.H., Herrmann R.G.** Genetic diversity for RFLPs in European maize inbreds: I. Relation to performance of Flint x Dent crosses for forage traits. *Crop Sci*. 1992. 32: 895-902.
 45. **Borevitz J.O., Maloof J.N., Lutes J., Dabi T., Redfern J.L., Trainor G.T., Werner J.D., Asami T., Berry C.C., Weigel D., Chory J.** Quantitative trait loci controlling light and hormone response in two accessions of *Arabidopsis thaliana*. *Genetics*. 2002. 160: 683-696.
 46. **Bost B., De Vienne D., Hospital F., Moreau L., Dillmann C.** Genetic and nongenetic bases for the L-shaped distribution of quantitative trait loci effects. *Genetics*. 2001. 157: 1773-1787.
 47. **Bost B., Dillmann C., De Vienne D.** Fluxes and metabolic pools as model traits for quantitative genetics. I. The L-shaped distribution of gene effects. *Genetics*. 1999. 153: 2001-2012.
 48. **Bousquet J., Strauss S.H., Doerksen A.H., Price R.A.** Extensive variation in evolution rate of *rbcL* gene sequences among seed plants. *Proc. Natl. Acad. Sci. USA*. 1992. 89: 7844-7848.
 49. **Bradshaw H.D., Stettler R.F.** Molecular genetics of growth and development in *Populus*. *Proc. Natl. Acad. Sci. USA*. 1995. 139: 963-973.
 50. **Breto M.P., Asins M.J., Carbonell E.A.** Salt tolerance in *Lycopersicon* species. III. Detection of QTLs by means of molecular markers. *Theor. Appl. Genet*. 1994. 88: 395-401.
 51. **Brown A.H.D.** Core collections: a practical approach to genetic resources management. *Genome*. 1989. 31: 318-324.
 52. **Brown A.N.D., Munday J.** Population genetic structure and optimal sampling of land races of barley from Iran. *Genetica*. 1982. 58: 85-96.
 53. **Brown A.N.D., Weir B.S.** Measuring genetic variability in plant popu-

- lations. In: *Isozymes in Plant Genetics and Breeding. Part A.* Tanksley S.D., Orton T.J. (eds.). Elsevier, Amsterdam. 1983. pp.219-239
54. **Brown A.H.D., Burdon J.J., Jarosz A.M.** Isozyme analysis of plant mating system. In: *Isozyme in Plant Biology.* Soltis D.E., Soltis P.S. (eds.). 1989. pp.73-105.
 55. **Bruford M.W., Wayne R.K.** Microsatellites and their application to population genetics studies. *Curr. Op. Dev.* 1993. 3: 939-943.
 56. **Bubeck D.M., Goodman M.M., Bevis W.D., Grant D.** Quantitative trait loci controlling resistance to gray leaf spot in maize. *Crop Sci.* 1993. 33: 838-847.
 57. **Buckingham S.D.** Scientific software: seeing the SNPs between us. *Nature Methods.* 2008. 5: 903-908.
 58. **Burke D.T., Carle G.F., Olson M.V.** Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science.* 1987. 236: 806-812.
 59. **Burkhamer R.L., Lanning S.P., Martens R.J., Martin J.M., Talbert L.E.** Prediction progeny variance from parental divergence in hard red spring wheat. *Crop Sci.* 1998. 38: 243-248.
 60. **Burstin J., Charcosset A.** Relationship between phenotypic and marker distances: theoretical and experimental investigations. *Heredity.* 1997. 118: 519-526.
 61. **Burstin J., Charcosset A., Barriere Y., Herbert Y., De Vienne D., Damerval C.** Molecular markers and protein quantities as genetic descriptors in maize. II. Prediction of performance of hybrids for forage traits. *Plant Breeding.* 1995. 114: 427-433.
 62. **Byrne P.F., McMullen M.D., Snook M.E., Musket T.A., Theuri J.M., Widstrom N.W., Wiseman B.R., Coe E.H.** Quantitative trait loci and metabolic pathways: genetic control of the concentration of maysin, a corn earworm resistance factor, in maize silks. *Proc. Natl. Acad. Sci. USA.* 1996. 93: 8820-8825.
 63. **Byrne M., Moran G.F.** Population divergence in the chloroplast genome of *Eucaliptus nitens*. *Heredity.* 1994. 73: 18-28.
 64. **Camefort H.** Fecondation et proembryogenese chez les Abietacees (notion de neocytoplasme). *Rev. Cytol. et Biol. Veg.* 1969. 32: 253-271.
 65. **Causse M.A., Rocher J.P., Henry A.M., Charcosset A., Prioul J.L., De Vienne D.** Genetic dissection of the relationship between carbon metabolism and early growth in maize, with emphasis on key-enzyme loci. *Mol. Breed.* 1995. 1: 259-272.
 66. **Causse M., Santoni Damerval C., Maurice A., Charcosset A., Deatrick J. De Vienne D.** A composite map of expressed sequences in maize. *Genome.* 1996. 39: 418-432.
 67. **Chakraborty R., Leimar O.** genetic variation within a subdivided population. In: *Population Genetics and Fishery management.* Ryman N.,

- Utter F. (eds.). University of Washington Press, Seattle and London. 1987. pp.89-120.
68. **Chalmers K.J., Barua U.M., Hackett C.A., Thomas W.T.B., Waugh R., Pawell W.** Identification of RAPD markers linked to genetic factors controlling the milling energy requirement of barley. *Theor. Appl. Genet.* 1993. 87: 314-320.
 69. **Champoux M.C., Wang G., Sarkarung S., Mackill D.J., Otoole J.C., Huang N., McCouch S.R.** Locating genes associated with root morphology and drought avoidance in rice via linkage to molecular markers. *Theor. Appl. Genet.* 1995. 90: 969-981.
 70. **Charcosset A.** Prediction of heterosis. In: *Reproductive Biology and Plant Breeding. Proc. XIIIth Eucarpia Meeting. Dattee Y., Dumas C., Gallais A. (eds.)*. Springer-Verlag. 1992. pp.365-369
 71. **Charcosset A., Bonnisseau B., Touchebeuf O., Burstin J., Dubreuil P., Barriere Y., Gallais A., Denis J.B.** Prediction of maize hybrid silage performance using marker data: comparison of several models for sprcific combining ability. *Crop Sci.* 1998. 38: 38-43.
 72. **Charcosset A., Dubreuil P., Essiox L., Gallais A.** Utilisation des techniques de marquage génétique pour l'analyse de la variabilité génétique et la prediction du phénomène d'hétérosis. *Le Selectionneur Français.* 1994. 44: 3-13.
 73. **Charcosset A., Essioux L.** The effect of population structure on the relationship between heterosis and heterozygosity at marker loci. *Theor. Appl. Genet.* 1994. 89: 336-343.
 74. **Charcosset A., Gallais A.** Estimation of the contribution of quantitative trait loci (QTL) to the variance of a quantitative trait by means of genetic markers. *Theor. Appl. Genet.* 1996. 93: 1193-1201.
 75. **Charcosset A., Gallais A.** Application of markers in selection. In: *Molecular Markers in Plant Genetics and Biotechnology. De Vienne D. (ed.)*. Science Publishers, Inc. Enfield, NH. 2003. pp.153-176.
 76. **Charcosset A., Lefort-Buson M., Gallias A.** Relationships between heterosis and heterozygosity at marker loci: a theoretical computation. *Theor. Appl. Genet.* 1991. 81: 571-575.
 77. **Chesnokov Yu.V., Shutov A.D.** 11S seed storage globulins: are they reliable as molecular markers? In: *Recent Res. Devel. Genet. Breeding.* 2004. V.1. pp.181-194.
 78. **Chesnoy L.** L'origine des organites du cytoplasme embryonnaire chez les gymnosperms. *Bull. Soc. Bot. Fr. Actual. Bot.* 1987. 134: 51-56.
 79. **Chiaese P., Ohkama-Ohtsu N., Molvig L., Godfree R., Dove H., Hocart C., Fujiwara T., Higgins T.J., Tabe L.M.** Sulphur and nitrogen nutrition influence the response of chickpea seeds to an added, transgenic sink for organic sulphur. *J. Exp. Bot.* 2004. 55: 1889-1901.

80. **Chong D.K.X., Yang R.C., Yeh F.C.** Nucleotide divergence between populations of trembling aspen (*Populus tremuloides*) estimated with RAPDs. *Curr. Genet.* 1994. 26: 374-376.
81. **Churchill G.A., Doerge R.W.** Empirical threshold values for quantitative trait mapping. *Genetics.* 1994. 138: 963-971.
82. **Churchill G.A., Giovannonni J.J., Tanksley S.D.** Pooled-sampling makes high-resolution mapping practical with DNA markers. *Proc. Natl. Acad. Sci. USA.* 1993. 90: 16-20.
83. **Clark A.G., Lanigan C.M.S.** Prospects for estimating nucleotide divergence with RAPDs. *Mol. Biol. Evol.* 1993. 10: 1096-1111.
84. **Courtois B., McLaren G., Sinha P.K., Prasad K., Yadav R., Shen L.** Mapping QTLs associated with drought avoidance in upland rice. *Mol. Breed.* 2000. 6: 55-66.
85. **Crasta O.R., Xu W.W., Rosenow D.T., Mullet J., Nguen H.T.** Mapping of post-flowering drought resistance trait in grain sorghum: association between QTLs influencing premature senescence and maturity. *Mol. Gen. Genet.* 1999. 262: 579-588.
86. **Crosbie T.M., Pearce R.B., Mock J.J.** Selection for high CO₂ exchange rate among inbred lines of maize. *Crop Sci.* 1981. 21: 629-631.
87. **Crouch M., Sussex I.M.** Development and storage-protein synthesis in *Brassica napus* L. embryos in vivo and in vitro. *Planta.* 1981. 153: 64-74.
88. **Crow J.F., Kimura M.** The number of alleles that can be maintained in a finite population. *Genetics.* 1964. 49: 725-738.
89. **Damerval C., De Vienne D.** Divergence morphologique et divergence moleculaire: 1. Apport des margueurs proteiques. In: *Les Distances Genetiques.* Lefort-Buson M., De Vienne D. (eds.). INRA, Paris. 1985.
90. **Damerval C., Maurice A., Josse J.M., De Vienne D.** Quantitative trait loci underlying gene product variation – a novel perspective for analyzing regulation of genome expression. *Genetics.* 1994. 137: 289-301.
91. **Darvasi A., Soller M.** Selective genotyping for determination of linkage between a molecular marker and a quantitative trait locus. *Genetics.* 1992. 85: 353-359.
92. **Darvasi A., Soller M.** Selective DNA pooling for a determination of linkage between a molecular marker and quantitative trait locus. *Genetics.* 1994. 138: 1365-1373.
93. **Darvasi A., Weinreb A., Minke V., Weller J.I., Soller M.** Detection marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics.* 1993. 134: 943-951.
94. **Dekkers J.C., Hospital F.** The use of molecular genetics in the improvement of agricultural populations. *Nat. Rev. Genet.* 2002. 3: 22-32
95. **Delourme R., Bouchereau A., Hubert N., Renard M., Landry B.S.** Identification of RAPD markers linked to a fertility restorer gene for the

- Ogura radich cytoplasmic male sterility of rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* 1994. 88: 741-748.
96. **Demarly Y.** The concept of linkat. In: Proceedings of the Conference on Broadening Genetic Bases of Crops. Zeven A.C., van Harten A.M. (eds.). PUDOC. Wageningen. pp. 257-265.
 97. **Desplanque B., Boudry P., Broomberg K., Saumitou-Laprade P., Cuguen J., Van Dijk H.** Genetic diversity and gene flow between wild, cultivated and weedy forms of *Beta vulgaris* L. (Chenopodiaceae), assessed by RFLP and microsatellite markers. *Theor. Appl. Genet.* 1999. 98: 1194-1201.
 98. **Devey M.E., Delfino-Mix A., Kinloch B.B., Neale D.B.** Random polymorphic DNA markers tightly linked to a gene for resistance to white pine blister rust in sugar pine. *Proc. Natl. Acad. Sci. USA.* 1995. 92: 2066-2070.
 99. **De Vicente M.C., Tanksley S.D.** QTL analysis of transgressive segregation in an interspecific tomato cross. *Genetics.* 1993. 134: 585-596.
 100. **De Vienne D.** Marquerurs moleculares en genetique et biotechnologies vegetales. INRA, Paris. 2002.
 101. **De Vienne D., Damerval C.** Mesures de la divergence genetique: 3. Distances calculees a partier de marqueurs moleculares. In: *Les Distances Genetiques.* Lefort M., De Vienne D. (eds.). INRA, Paris. 1985.
 102. **De Vienne D., Leonardi A., Damerval C., Zivy M.** Genetics of proteome variation for QTL characterization: application of drought-stress responses in maize. *J. Exp. Bot.* 1999. 50: 303-309.
 103. **Deu M., Gonzalez de Leon D., Glaszmann J.C., Degremont I., Chantereau J., Lanaud C., Hamon P.** RFLP diversity in cultivated sorghum in relation to racial differentiation. *Theor Appl. Genet.* 1994. 88: 834-844.
 104. **Diers B.W., Mansur L., Imsande J., Shoemaker R.C.** Mapping *Phytophthora* resistance loci in soybean with restriction fragment length polymorphism markers. *Crop Sci.* 1992. 32: 377-383.
 105. **Dillmann C., Bar-Hen A., Guerin D., Charcosset A., Murigneux A.** Comparison of RFLP and morphological distances between maize *Zea mays* L. inbred lines. Consequences for germplasm protection purposes. *Theor. Appl. Genet.* 1997. 95: 92-102.
 106. **Dillmann C., Charcosset A., Goffnet B., Smith J.S.C., Dattee Y.** Best linear unbiased estimators of the molecular genetic distance between inbred lines. In: *Advances in Biometrical Genetics.* Krajewski P., Kaczmarek Z. (eds.). Poznan. 1997. p.105-110.
 107. **Dirlewanger E., Isaak P.G., Ranade S., Belajouza M., Cousin R., De Vienne D.** Restriction fragment length polymorphism analysis of lo-

- ci associated with disease resistance genes and development traits in *Pisum sativum* L. *Theor. Appl. Genet.* 1994. 88: 17-27.
108. **Doebley J., Stec A., Gustus C.** *Toesinte branched 1* and origin of maize: evidence for epistasis and the evolution of dominance. *Genetics.* 1995. 144: 333-346.
 109. **Doebley J., Stec A., Hubbard L.** The evolution of apical dominance in maize. *Nature.* 1997. 386: 485-488.
 110. **Dong J., Wagner D.B.** Paternally inherited chloroplast polymorphism in *Pinus*: estimation of diversity and population subdivision, and tests of disequilibrium with a maternally inherited mitochondrial polymorphism. *Genetics.* 1994. 136: 1187-1194.
 111. **Dow B.D., Ashley M.V.** Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus makrocarpia*. *Mol. Ecol.* 1996. 5: 615-627.
 112. **Dow B.D., Ashley M.V.** High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. *J. Heredity.* 1998. 89: 62-70.
 113. **Dubreuil P., Charcosset A.** Genetic diversity within and among maize populations: a comparison between isozymes and nuclear RFLP loci. *Theor. Appl. Genet.* 1998. 96: 577-587.
 114. **Dubreuil P., Dufour P., Krejci E., Causse M., de Vienne D., Gallais A., Charcosset A.** Organization of RFLP diversity among inbred lines of maize representing the most significant heterotic groups. *Crop Sci.* 1996. 36: 790-799.
 115. **Dudley J.W., Saghal Maroof M.A., Rufener G.K.** Molecular markers and grouping of parents in maize breeding programs. *Crop Sci.* 1991. 31: 718-723.
 116. **Edwards M.D., Helentjaris T., Wright S., Stuber C.W.** Molecular-marker facilitated investigations of QTL in maize. IV. Analysis based on genome saturation and restriction fragment length polymorphism markers. *Theor. Appl. Genet.* 1992. 83: 765-774.
 117. **Edwards M.D., Stuber C.W., Wendel J.F.** Molecular-marker facilitated investigation of QTL in maize. I. Number, genetic distribution and types of gene action. *Genetics.* 1987. 116: 113-125.
 118. **El-Din El Assal S., Alonso-Blanco C., Peters A.J., Raz V., Koorneef M.** A QTL for flowering time in *Arabidopsis* reveals a novel allele of CRY2. *Nat. Genet.* 2001. 29: 435-440.
 119. **Ellstrand N.C.** Gene flow among seed plant populations. *New Forests.* 1992. 6: 241-256.
 120. **El Mousadik A., Petit R.J.** High levels of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* (L.) Skeels) endemic to Morocco. *Theor. Appl. Genet.* 1996. 92: 832-839.

121. **El Mousadik A., Petit R.J.** Chloroplast DNA phylogeography of the argan tree of Morocco. *Mol. Evol.* 1996. 5: 547-555.
122. **Ennos R.A.** Estimating the relative rates of pollen and seed migration among plant populations. *Heredity.* 1994. 72: 250-259.
123. **Eshed Y., Abu-Abied M., Saranga Y., Zamir D.** *Lycopersicon esculentum* lines containing small overlapping introgressions from *L.pennellii*. *Theor. Appl. Genet.* 1992. 83: 1027-1034.
124. **Eshed Y., Zamir D.** An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield associated QTLs. *Genetics.* 1995. 141: 1147-1162.
125. **Excofier L., Smouse P.E., Quattro J.M.** Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics.* 1992. 131: 479-491.
126. **Falconer D.S.** Introduction to Quantitative Genetics. 2nd ed. Longman, London. 1961. 340 p.
127. **Falconer D.S., MacKay T.F.C.** Quantitative trait loci. In: *Introduction to Quantitative Genetics.* 4th ed., Longman Group Ltd. Harlow. England. 1996. pp.12-17
128. **Fatokun C.A., Mennancio-Hautea D.I., Danesh D., Young N.D.** Evidence for orthologous seed weight genes in cowpea and mung bean based on RFLP mapping. *Genetics.* 1992. 132: 841-846.
129. **Flowers T.J., Koyama M.L., Flowers S.A., Sudhakar C., Singh K.P., Yeo A.R.** QTL: their place in engineering tolerance of rice to salinity. *J. Exp. Bot.* 2000. 51: 99-106.
130. **Foolad M.R.** Comparison of salt tolerance during seed germination and vegetative growth in tomato by QTL mapping. *Genome.* 1999. 42: 727-734.
131. **Foolad M.R., Chen F.Q., Lin G.Y.** RFLP mapping of QTLs conferring salt tolerance during germination in an interspecific cross of tomato. *Theor. Appl. Genet.* 1998. 97: 1133-1144.
132. **Frary A., Nesbitt T.C., Grandillo S., Knaap E., Cong B., Liu J., Meller J., Elber R., Alpert K.B., Tanksley S.D.** *fw 2.2*: a quantitative trait locus key to the evolution of tomato fruit size. *Science.* 2000. 289: 85-88.
133. **Frei O.M., Stuber C.W., Goodman M.M.** Use of allozymes as genetic markers for predicting performance in maize single cross hybrids. *Crop Sci.* 1986. 26: 37-42.
134. **Freyark P.J., Lee M., Woodman W.L., Martinson C.A.** Quantitative and qualitative trait loci affecting host-plant response to *Exserohilum turcicum* in maize (*Zea mays* L.). *Theor. Appl. Genet.* 1993. 87: 537-544.

135. **Fridman E., Pleban T., Zamir D.** A recombinant hotspot delimits a wild QTL for tomato sugar content to 484 bp within an invertase gene. *Proc. Natl. Acad. Sci. USA*. 2000. 97: 4718-4723.
136. **Frova C., Krajewski P., di Fronzo N., Villa M., Sari-Gorla M.** Genetic analysis of drought tolerance in maize by molecular markers. I. Yield components. *Theor. Appl. Genet.* 1999. 99: 280-288.
137. **Frova C., Sari-Gorla M.** Quantitative trait loci (QTLs) for pollen thermotolerance detected in maize. *Mol Gen Genet.* 1994. 245: 424-430.
138. **Fuerst P.A., Ferrell R.E.** The stepwise mutation model: an experimental evaluation utilizing hemoglobin variants. *Genetics*. 1980. 94: 185-201.
139. **Gallais A.** *Theorie de la Selection en Amelioration des Plantes*. Masson ed. 1990.
140. **Gallais A.** La selection assistee par marqueurs. *Le Seletionneur Francais*. 1995. 43: 43-62.
141. **Gallais A., Charcosset A.** Efficiency of marker-assisted selection. In: *Biometrics in Plant Breeding: Applications of Molecular Markers. Proc. 9th Meeting of the Eucarpia Section Biometrics in Plant Breeding*. Van Oijen J.W., Jansen J. (eds.). 1994. p. 91-99.
142. **Gallais A., Dillmann C., Hospital F.** An analytical approach of marker-assisted selection with selection on markers only. In: *Proc. 10th Meeting of the Eucarpia Section Biometrics in Plant Breeding*. Krajewski P., Kaczmarek Z. (eds.). IGR, Poznan, 1997. p. 111-116.
143. **Gallais A., Rives M.** Detection, number and effects of QTLs for a complex character. *Agronomie*. 1993. 13: 723-738.
144. **Garcia A.A., Kido E.A., Meza A.N., Souza H.M., Pinto L.R., Pastina M.M., Leite C.S., Silva J.A., Ulian E.C., Figueira A., Souza A.P.** Development of an integrated genetic map of a sugarcane (*Saccharum* spp.) commercial cross, based on a maximum-likelihood approach for estimation of linkage and linkage phases. *Theor. Appl. Genet.* 2006. 112: 298-314.
145. **Gardner C.O., Eberhart S.A.** Analysis and interpretation of the variety cross diallele and related populations. *Biometrics*. 1966. 22: 439-452.
146. **Gaunt T.R., Rodriguez S., Zapata C., Day I.N.M.** MIDAS: software for analysis and visualization of interallelic disequilibrium between multiallelic markers. *BMC Bioinformatics*. 2006. 7: 227.
147. **Gentzbittel I., Zhang Y.X., Vear F., Griveau B., Nicolas P.** RFLP studies of genetic relationships among inbred lines of the cultivated sunflower, *Heliantus annuus* L.: evidence for distinct restorer and maintainer germplasm pools. *Theor. Appl. Genet.* 1994. 89: 419-425.
148. **Gerber S., Mariette S., Struff R., Bodenes C., Kremer A.** Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. *Mol. Ecol.* 2000. 17: 1037-1048.

149. **Gimelfarb A., Lande R.** Simulation of marker-assisted selection in hybrid populations. *Genet. Res. Camb.* 1994. 63: 39-47.
150. **Gimelfarb A., Lande R.** Marker-assisted selection and marker-QTL associations in hybrid populations. *Theor. Appl. Genet.* 1995. 91: 522-528.
151. **Giraudat J., Hauge B.H., Valon C., Smalle J., Parcy F., Goodman H.M.** Isolation of the *Arabidopsis ABI3* gene by positional cloning. *The Plant Cell.* 1992. 4: 1251-1261.
152. **Godshalk E.B., Lee M., Lamkey K.R.** Relationship of restriction fragment length polymorphisms to single-cross hybrid performance of maize. *Theor. Appl. Genet.* 1990. 80: 273-280.
153. **Godoy J.A., Jordano P.** Seed dispersal by animals: exact identification of source trees with endocarp DNA microsatellites. *Mol. Ecol.* 2001. 10: 2275-2283.
154. **Goldman I.L., Paran I., Zamir D.** Quantitative trait locus analysis of a recombinationinbred line population derived from a *Lycopersicon esculentum* × *Lycopersicon cheesmanii* cross. *Theor. Appl. Genet.* 1995. 90: 925-932.
155. **Goldman I.L., Rocherford T.R., Dudley J.W.** Quantitative trait loci influencing protein and starch concentration in the Illinois Long Term Selection maize strains. *Theor. Appl. Genet.* 1993. 87: 217-224.
156. **Gonzales-Candelas F., Elena S.F., Moya A.** Approximate variance of nucleotide divergence between two sequences estimated from restriction fragment data. *Genetics.* 1995. 140: 1443-1446.
157. **Gouesnard B., Bataillon T.M., Decoux G., Rozale C., Schoen D.J., David J.L.** MSTRAT: An algorithm for building germplasm core collection by maximizing allelic or phenotypic richness. *J. Heredity.* 2001. 92: 93-94.
158. **Grattapaglia D., Bertolucci F.L., Sederoff R.R.** Genetic mapping of QTLs controlling vegetative propagation in *Eucalyptus grandis* and *E.urophylla* using a pseudo-testcross mapping strategy and RAPD markers. *Theor. Appl. Genet.* 1995. 90: 933-947.
159. **Gregorius H.R.** The probability for losing an allele when diploid genotypes are sampled. *Biometrics.* 1980. 36: 643-652.
160. **Hackett C.A., Weller J.I.** Genetic mapping of quantitative trait loci for traits with ordinal distributions. *Biometrics.* 1995. 51: 1252-1263.
161. **Haley S.D., Knott S.A.** A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity.* 1992. 69: 315-324.
162. **Haley S.D., Mikals P.N., Stavely J.R., Byrum J., Kelly J.D.** Identification of RAPD markers linked to a major rust resistance gene block in common bean. *Theor. Appl. Genet.* 1993. 86: 505-512.

163. **Hammond-Kosack K.E., Jones J.D.G.** Plant disease resistance genes. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 1997. 48: 575-607.
164. **Hamrick J.L., Godt M.J.W.** Allozyme diversity in plant species. In: *Plant Population Genetics, Breeding and Genetic Resources*. Brown A.H.D., Glegg M.T., Kahler A.L., Weir B.S. (eds.). Sinauer Associates. 1989. p.43-64.
165. **Hancock J.F.** Plant Evolution and the Origin of Crop Species. 3rd ed. CABI, MPG Books Group, London, UK. 2012. 246 p.
166. **Hanson W.D.** Early generation analysis of length of heterozygous chromosome segments around a locus held heterozygous with back-crossing or selfing. *Genetics*. 1959. 44: 833-837.
167. **Harris S.A., Ingram R.** Chloroplast DNA and biosystematics: the effect of intraspecific diversity and plasmid transmission. *Taxon*. 1991. 40: 393-412.
168. **Hartl D.L., Clark A.D.** Principle of Population Genetics. Sinauer Associates. 1989. 682 p.
169. **Hayes H.K., Johnson I.J.** The breeding of improved selfed lines of corn. *J. Am. Soc. Agron.* 1939. 31: 710-724.
170. **Hayes P.M., Lui B.H., Knapp S.J., Chen F., Jones B., Blake T., Franckowiak J., Rasmusson D., Sorrells M., Ullrichs E., Wesenberg D., Kleinhofs A.** Quantitative trait locus effects and environmental interaction in a simple of North American barley germplasm. *Theor. Appl. Genet.* 1993. 87: 392-401.
171. **Helentjaris T.** A linkage map for maize based on RFLPs. *Trends Genet.* 1987. 3: 217-221.
172. **Henderson C.R.** Best linear unbiased estimation and prediction under a selection model. *Biometrics*. 1975. 31: 423-447.
173. **Hipkins V.D., Krutovskii K.V., Strauss S.H.** Organelle genomes in conifers. Structure, evolution and diversity. *Forest Genet.* 1994. 1(4): 179-189.
174. **Holland J.B.** EPISTACY: a SAS program for detecting two-locus epistasis interactions using genetic marker information. *J. Hered.* 1998. 89: 374-375.
175. **Holsinger K.E., Mason-Gamer R.J.** Hierarchical analysis of nucleotide diversity in geographically structured populations. *Genetics*. 1996. 142: 629-639.
176. **Hormaza J.I., Dollo L., Polito V.S.** Identification of RAPD marker linked to sex determination in *Pistacia vera* using bulked segregant analysis. *Theor. Appl. Genet.* 1994. 89: 9-13.
177. **Hospital F., Charcosset A.** Marker-assisted introgression of quantitative trait loci. *Genetics*. 1997. 147: 1469-1485.
178. **Hospital F., Chavalet C., Mulsant P.** Using markers in gene introgression breeding programs. *Genetics*. 1992. 132: 1199-1210.

179. **Hospital F., Goldringer I., Openshaw S.** Efficient marker based recurrent selection for multiple quantitative trait loci. *Genet. Res. Camb.* 2000. 75: 357-368.
180. **Hospital F., Moreau L., Lacoudre F., Charcosset A., Gallais A.** More on the efficiency of marker-assisted selection. *Theor. Appl. Genet.* 1997. 95: 1181-1189.
181. **Hudson S.S.** Gene genealogies and coalescent process. *Oxf. Surv. Ecol. Biol.* 1990. 7: 1-44.
182. **Hudson R.R., Slatkin M., Maddison W.P.** Estimation of levels of gene flow from DNA sequence data. *Genetics.* 1992. 132: 583-589.
183. **Hurlbert S.H.** The nonconcept of species diversity: a critique and alternative parameters. *Ecology.* 1971. 52: 577-586.
184. **Innan H., Terauchi R., Kahl G., Tajima F.** A method for estimating nucleotide diversity from AFLP data. *Genetics.* 1999. 151: 1157-1164.
185. **Isabel N., Beaulieu J., Bousquest J.** Complete congruence between gene diversity estimates derived from genotypic data at enzymes and random amplified polymorphic DNA loci in black spruce. *Proc. Natl. Acad. Sci. USA.* 1995. 92: 6369-6373.
186. **Jansen R.C.** Interval mapping of multiple quantitative trait loci. *Genetics.* 1993. 135: 205-211.
187. **Jansen R.C., Stam P.** High resolution of quantitative traits into multiple loci via interval mapping. *Genetics.* 1994. 136: 1447-1455.
188. **Jansen R.C., Van Ooijen J.W., Stam P., Lister C., Dean C.** Genotype-by-environment interaction in genetic mapping of multiple quantitative trait loci. *Theor. Appl. Genet.* 1995. 91: 33-37.
189. **Jones J.D.G.** Putting knowledge of plant disease resistance genes to work. *Curr. Opin. Plant Biol.* 2001. 4: 281-287.
190. **Jourjon M.F., Jasson S., Marcel J., Ngom B., Mangin B.** MCQTL: multi-allelic QTL mapping in multi-cross design. *Bioinformatics.* 2005. 21: 128-130.
191. **Jukes T.H., Cantor C.R.** Evolution of protein molecules. In: *Mammalian Protein Metabolism.* Munro H.N. (ed.). Academic Press. 1969. p.21-132.
192. **Kacser H., Burns J.A.** The molecular basis of dominance. *Genetics.* 1981. 97: 639-666.
193. **Kaplan N.** Statistical analysis of restriction map data and nucleotide sequence data. In: *Statistical Analysis of DNA Sequence Data.* Weir B.S. (ed.). Mark Dekker. 1983. p. 75-106.
194. **Kearsey M.J., Farquhar A.G.I.** QTL analysis in plants; where are we now? *Heredity.* 1998. 80: 137-142.
195. **Kearsey M.J., Hyne V.** QTL analysis: a simple "marker-regression" approach. *Theor. Appl. Genet.* 1994. 89: 698-702.

196. **Keim P., Diers B.W., Olson T.C., Shoemaker R.C.** RFLP mapping in soybean: association between marker loci and variation in quantitative traits. *Genetics*. 1990. 126: 735-742.
197. **Keim P., Shoemaker R.C., Palmer R.G.** Restriction fragment length polymorphism diversity in soybean. *Theor. Appl. Genet.* 1989. 77: 786-792.
198. **Kesseli R., Witsenboer H., Stanghellini M., Vandermark G., Michelmore R.** Recessive resistance to *Plasmopara lactucae-redicis* maps by bulked segregant analysis to cluster of dominant disease resistance genes in lettuce. *Mol. Plant-Microbe Interact.* 1993. 6: 722-728.
199. **Kimura M.** A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 1980. 16: 111-120.
200. **Kisha T.J., Sneller C.H., Diers B.W.** Relationship between genetic distance among parents and genetic variance in populations of soybean. *Crop Sci.* 1997. 37: 1317-1325.
201. **Knapp S.J., Bridges W.S.** Using molecular markers to estimate quantitative trait locus parameters: power and genetic variances for unreplicated and replicated progeny. *Genetics*. 1990. 126: 769-777.
202. **Knapp S.J., Bridges W.C., Birkes D.** Mapping QTL using molecular linkage maps. *Theor. Appl. Genet.* 1990. 79: 583-592.
203. **Koester R.P., Sisco P.H., Stuber C.W.** Identification of quantitative trait loci controlling days to flowering and plant height in two near isogenic lines of maize. *Crop Sci.* 1993. 33: 1209-1216.
204. **Korol A., Ronin Y., Itzovich A., Peng J., Nevo E.** Enhanced efficiency of QTL mapping analysis based on multivariate complexes of quantitative traits. *Genetics*. 2001. 157: 1789-1803.
205. **Korol A., Ronin Y., Kirzhner V.** Interval mapping of quantitative trait loci employing correlated trait complex. *Genetics*. 1995. 140: 1137-1147.
206. **Kremer A., Petit R.J., Pons O.** Measures of polymorphism within and between populations. In: *Molecular Tools for Screening Biodiversity in Plants and Animals*. Karp A., Ingram D.S., Isaac P.G. (eds.). Chapman and Hall, London. 1997.
207. **Krug C.A., Viegas G.P., Paolieri L.** Híbridos comerciais de Milho. *Bragantia*. 1943. 3: 367-552.
208. **Kruglyak L., Lander E.S.** A nonparametric approach for mapping quantitative trait loci. *Genetics*. 1995. 139: 1421-1428.
209. **Lagercrantz U., Putterill J., Coupland G., Lydiate D.** Comparative mapping in *Arabidopsis* and *Brassica*, fine scale genome colinearity and congruence of genes controlling flowering time. *Plant J.* 1996. 9: 13-20.
210. **Lande R., Thompson B.** Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*. 1990. 124: 743-756.

211. **Lander E.S., Botstein D.** Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*. 1989. 121: 185-199.
212. **Lander E.S., Green P., Abrahamson J., Barlow A., Daly M.J., Lincoln S.E., Newburg L.** MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics*. 1987. 1: 174-181.
213. **Lanner-Herrera C., Gustafsson M., Falt A.S., Bryngelsson T.** Diversity in natural populations of wild *Brassica oleracea* as estimated by isozyme and RAPD analysis. *Genet. Resour. Crop Evol.* 1996. 43: 13-23.
214. **Lebreton C., Lazicjancic V., Steed A., Pekic S., Quarrie S.A.** Identification of QTL for drought responses in maize and their use in testing causal relationships between traits. *J. Exp. Bot.* 1995. 46: 853-865.
215. **Le Corre V., Dumolin-Lapegue D., Kremer A.** Genetic variability at enzyme and RAPD loci in sessile oak *Quercus petraea* (Matt.) Liebl.: the role of history and geography. *Mol. Ecology*. 1997. 6 : 519-529.
216. **Lefebvre V., Palloix A.** Both epistatic and additive effects of QTLs are involved in polygenic induced resistance to disease: a case study, the interaction pepper-*Phytophthora capsici* Leonian. *Theor. Appl. Genet.* 1996. 93: 503-511.
217. **Lefort-Buson M.** Distances génétiques et hétérosis: 1. Mise en évidence d'une relation entre heterosis and divergence genetique. In: *Les Distances Genétiques. Lefort-Buson M., De Vienne D. (eds.)*. INRA, Paris. 1985. pp.111-118
218. **Lefort-Buson M.** Distances génétiques et hétérosis: 4. Utilisation des criteres biometriques. In: *Les Distances Genétiques. Lefort-Buson M., De Vienne D. (eds.)*. INRA, Paris. 1985. pp.143-158
219. **Leonards-Schippers C., Gieffers W., Schafer-Pregl R., Ritter E., Knapp S.J., Salamini F., Gebhardt C.** Quantitative resistance to *Phytophthora infestans* in potato: a case study for QTL mapping in allogamous species. *Genetics*. 1994. 137: 67-77.
220. **Lewontin R.C.** The apportionment of human diversity. *Evol. Biol.* 1973. 6: 381-398.
221. **Lewontin R.C., Hubby J.L.** A molecular approach to study of genetic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics*. 1966. 54: 595-609.
222. **Leyser H.M.O., Lincoln C.A., Timpte C., Lammer D., Turner J., Estelle M.** *Arabidopsis* auxin-resistance gene *AXRI* encodes a protein related to ubiquitin-activating enzyme E1. *Nature*. 1993. 364: 161-164.
223. **Lilley J.M., Ludlow M.M., McCouch S.R., Otoole J.C.** Location QTL for osmotic adjustment and dehydration tolerance in rice. *J. Exp. Bot.* 1996. 47: 1427-1436.

224. **Liu Z., Furnier G.R.** Comparison of allozyme, RFLP and RAPD markers for revealing genetic variation within and between trembling aspen and bigtooth aspen. *Theor. Appl. Genet.* 1993. 87: 97-105.
225. **Liu J.S., Sabatti C., Teng J., Keats B.J.B., Risch K.** Bayesian analysis of haplotypes for linkage disequilibrium mapping. *Genome Research.* 2001. 11: 1716–1724.
226. **Livini C., Ajmone-Marsan P., Melchinger A.E., Messmer M.M., Motto M.** Genetic diversity of maize inbred lines within and among heterotic groups revealed by RFLPs. *Theor. Appl. Genet.* 1992. 84: 17-25.
227. **Lu X., Niu T., Liu, J.S.** Haplotype information and linkage disequilibrium mapping for single nucleotide polymorphisms. *Genome Research.* 2003. 13: 2112–2117.
228. **Lukens L., Doebley J.** Epistatic and environmental interactions for quantitative trait loci involved in maize evolution. *Genet. Research.* 1999. 74: 291-302
229. **Luo Z.W., Thompson R., Wooliams J.A.** A population genetics model of marker-assisted selection. *Genetics.* 1997. 146: 1173-1183.
230. **Lynch M.** The similarity index and DNA fingerprinting. *Mol. Biol. Evol.* 1990. 7: 478-484.
231. **Lynch M., Crease T.J.** The analysis of population survey data on DNA sequence variation. *Mol. Biol. Evol.* 1990. 7: 377-394.
232. **Lynch M., Milligan B.G.** Analysis of population genetic structure with RAPD markers. *Mol. Ecol.* 1994. 3: 91-99.
233. **Mailund T., Schierup M.H., Pedersen C.N.S., Madsen J.N., Hein J., Schausser L.** GeneRecon – a coalescent based tool for fine-scale association mapping. *Bioinformatics.* 2006. 22: 2317-2318.
234. **Maisonneuve B., Bellec Y., Anderson P., Michelmore R.W.** Rapid mapping of two genes for resistance to downy mildew from *Lactuca serriola* to existing cluster of resistance genes. *Theor. Appl. Genet.* 1994. 89: 96-104.
235. **Mangin B., Goffinet B., Rebai A.** Constructing confidence intervals for QTL location. *Genetics.* 1994. 138: 301-308.
236. **Manjarrez-Sandoval P., Carter T.E.J., Webb D.M., Burton J.W.** RFLP genetic similarity and coefficient of percentage as genetic variance predictors for soybean yield. *Crop Sci.* 1997. 37: 698-703.
237. **Manly K.F.** A Macintosh program for storage and analysis of experimental genetic mapping data. *Mammalian Genome.* 1993. 4: 303–313.
238. **Mansur L.M., Lark K.G., Kross H., Oliviera A.** Interval mapping of quantitative trait loci for reproductive, morphological and seed traits of soybean (*Glycine max* L.). *Theor. Appl. Genet.* 1993. 86: 907-913.
239. **Mansur L.M., Orf J., Lark K.G.** Determining the linkage of quantitative trait loci to RFLP markers using extreme phenotypes of recombi-

- nant inbreds of soybean (*Glycine max* L. Merr.). *Theor. Appl. Genet.* 1993. 86: 914-918.
240. **Mariette S., Chagne D., Lezier C., Pastuszka P., Raffin A., Plomion C., Kremer A.** Genetic diversity within and among *Pinus pinaster* populations: comparison between AFLP and microsatellite markers. *Heredity*. 2001. 86: 469-479.
241. **Mariette S., Le Corre V., Austerlitz F., Kremer A.** Sampling within the genome for measuring within-population diversity: trade-offs between markers. *Mol. Ecol.* 2002. 34: 75-79.
242. **Mariette S., Cottrell J., Csaikl U., Goikoechea P., König A., Lowe A.J., van Dam B.C., Barrenche T., Bodenes C., Streiff R., Burg K., Groppe K., Munro R.C., Tabbener H., Kremer A.** Comparison of levels of genetic diversity detected with AFLP and microsatellite marker within and among mixed *Q. petraea* (Matt.) Liebl. and *Q. robur* L. stands. *Silvae Genetica*. 2002. 51: 72-79.
243. **Marshall D.R., Jain S.K.** Genetic polymorphism in natural populations of *Avena fatua* and *A. barbata*. *Nature*. 1969. 221: 276-278.
244. **Marshall T.C., Slate J., Kruuk L.E.B., Pemberton J.M.** Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* 1998. 7: 639-655.
245. **Martin G.B., Brommonschenkel S.H., Chunwongse J., Frary A., Ganai M.W., Spivey R., Wu T., Earle E.D., Tanksley S.D.** Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science*. 1993. 262: 1432-1436.
246. **Martin G.B., Williams J.G.K., Tanksley S.D.** Rapid identification of markers linked to a *Pseudomonas* resistance gene in tomato by using random primers and near-isogenic lines. *Proc. Natl. Acad. Sci. USA*. 1991. 88: 2336-2340.
247. **Martinez O., Curnow R.N.** Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. *Theor. Appl. Genet.* 1992. 85: 480-488.
248. **Martinez V., Thorgaard G., Robison B., Sillanpaa M.J.** An application of Bayesian QTL mapping to early development in double haploid lines of rainbow trout including environmental effects. *Genet. Research*. 2005. 86: 209-221.
249. **Mason-Gamer R.J., Holsinger K.E., Jansen R.K.** Chloroplast DNA haplotype variation within and among populations of *Coreopsis grandiflora* (Asteraceae). *Mol. Biol. Evol.* 1995. 12: 371-381.
250. **Maughan P.J., Maroof M.A.S., Buss G.R.** Molecular-marker analysis of seed-weight: genomic locations, gene action, and evidence for orthologous evolution among three legume species. *Theor. Appl. Genet.* 1996. 93: 574-579.

251. **McCauley D.E.** The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends Ecol. Genet.* 1995. 10: 198-202.
252. **McCauley D.E., Raveill J., Antonovics J.** Local founding events as determinants of genetic structure in plant metapopulation. *Heredity.* 1995. 75: 630-636.
253. **McGrath J.M., Quiros C.F.** Genetic diversity at isozyme and RFLP loci in *Brassica campestris* as related to crop type and geographical origin. *Theor. Appl. Genet.* 1992. 83: 783-790.
254. **McDonald J.H.** Detecting natural selection by comparing geographic variation in protein and DNA polymorphisms. In: *Non-neutral Evolution. Theories and Molecular Data.* Golding B. (ed.). Chapman and Hall. 1994. pp.88-101.
255. **Meagher T.R., Thompson E.** The relationship between single parent and parent pair genetic likelihoods in genealogy reconstruction. *Theor. Pop. Biol.* 1986. 29: 87-106.
256. **Melchinger A.E.** Genetic diversity and heterosis. In: *The Genetics and Exploitation of Heterosis in Crops.* Coors J.G., Pandey S. (eds.). *International Symposium on the Genetics and Exploitation of Heterosis in Crops.* CIMMYT, Mexico. 17-22.08.1997. 1999. chap. 10. pp.99-118.
257. **Melchinger A.E., Boppenmeier J., Dhillon B.S., Pollmer W.G., Herrmann R.G.** Genetic diversity for RFLPs in European maize inbreds: II. Relation to performance of hybrids within versus between heterotic groups for forage traits. *Theor. Appl. Genet.* 1992. 84: 672-681.
258. **Melchinger A.E., Graner A., Singh M., Messmer M.M.** Relationships among European barley germplasm. I. Genetic diversity among winter and spring cultivars revealed by RFLP. *Crop Sci.* 1994. 34: 1191-1199.
259. **Melchinger A.E., Messmer M.M., Lee M., Woodman W.L., Lamkey K.R.** Diversity and relationships among U.S. maize inbreds revealed by restriction fragment length polymorphisms. *Crop Sci.* 1991. 31: 669-678.
260. **Messmer M.M., Melchinger A.E., Boppenmeier J., Herrmann R.G., Brunklaus-Jung E.** RFLP analysis of early-maturing European maize germplasm. I. Genetic diversity among flint and dent inbreds. *Theor. Appl. Genet.* 1992. 83: 1003-1012.
261. **Michelmore R.** Molecular approaches to manipulation of disease resistance genes. *Ann. Rev. Phytopatol.* 1995. 15: 393-427.
262. **Michelmore R.** Isolation of disease resistance genes from crop plants. *Curr. Opin. Biotechnol.* 1995. 6: 145-152.
263. **Michelmore R.W., Paran I., Kesseli R.V.** Identification of markers linked to disease resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions using segregating populations. *Proc. Natl. Acad. Sci. USA.* 1991. 88: 9828-9832.

264. **Michalakis Y., Excoffier L.** A genetic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics*. 1996. 142: 1061-1064.
265. **Mimouni B., Robin J.-M., Azanza J.-L.** Comparative studies of 11S globulin constituents of *Brassica napus* L. and of its related species *Brassica campestris* L. and *Brassica oleracea* L. *Plant Sci*. 1990. 67: 183-194.
266. **Moran G.F.** Patterns of genetic diversity in Australian tree species. *New Forests*. 1992. 6: 49-66.
267. **Moreau L., Charcosset A., Hospital F., Gallais A.** Marker-assisted selection efficiency in populations of finite size. *Genetics*. 1998. 148: 1353-1355.
268. **Moreau L., Lemarie S., Charcosset A., Gallais A.** Economic efficiency of marker assisted selection. *Crop Sci*. 2000. 40: 329-337.
269. **Moser H., Lee M.** RFLP variation and genealogical distance, multivariate distance, heterosis, and genetic variance in oats. *Theor. Appl. Genet*. 1994. 87: 947-956.
270. **Murphy R.W., Sites J.W. Jr., Buth D.G., Haufler C.H.** Proteins I: isozyme electrophoresis. In: *Molecular Systematics*. Hillis D.M., Moritz C. (eds.). Sinauer Associates, Sunderland, MA. 1990. pp.45-126.
271. **Neale D.B., Saghai Maroof M.A., Allard R.W. Zhang Q., Jorgensen R.A.** Chloroplast DNA diversity in populations of wild and cultivated barley. *Genetics*. 1988. 120: 1105-1110.
272. **Nei M.** Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA*. 1973. 70: 3321-3323.
273. **Nei M.** Evolution of human races at the gene level. In: *Human Genetics. Part A. The unfolding genome*. Bonne-Tamir B., Cohen T., Goodman P.M. (eds.). Alan R. Liss, New York. 1982. pp.167-181.
274. **Nei M.** *Molecular Evolutionary Genetics*. Columbia University Press. 1987. 512 p.
275. **Nei M., Li W.H.** Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA*. 1979. 76: 5269-5273.
276. **Nei M., Miller J.C.** A simple method for estimation average number of nucleotide substitution within and between populations from restriction data. *Genetics*. 1990. 125: 873-879.
277. **Nei M., Roychoudhury A.K.** Sampling variances of heterozygosity and genetic distance. *Genetics*. 1974. 76: 379-380.
278. **Nevo E., Beiles A., Ben-Shlomo R.** The evolutionary significance of genetic diversity. Ecological, demographic and life history correlates. In: *Evolutionary Dynamics of Genetic Diversity*. Mani G.S. (ed.). Springer-Verlag (Lecture Notes in Biomathematics, no.53). 1984. pp.13-213.

279. **Nienhuis J., Sills G.** The potential of hybrid varieties in self-pollinating vegetables. In: *Reproductive Biology and Plant Breeding. Proc. XIIIth Eucarpia Meeting. Dattee Y., Dumas C., Gallias A. (eds.)*. Springer-Verlag. 1992.
280. **Nybom H., Bartish I.V.** Effects of life history traits and sampling strategies on genetic diversity estimates obtained with RAPD markers in plants. *Perspectives in Plant Ecol. Evol. Systematics*. 2000. 3/2: 93-114.
281. **Ohta T., Kimura M.** A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in finite population. *Genet. Res. Camb.* 1973. 22: 201-204.
282. **Pan A., Hayes P.M., Chen T.H.H., Blake T., Wright S., Karsai I., Bedo Z.** Genetic analysis of the components of winter hardiness in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 1994. 89: 900-910.
283. **Papa R., Attene G., Barcaccia G., Ohgata A., Konishi T.** Genetic diversity in landrace populations of *Hordeum vulgare* L. From Sardinia, Italy, as revealed by RAPDs, isozymes and morphophenological traits. *Plant Breeding*. 1998. 117: 523-530.
284. **Paterson A.H., De Verna J.W., Lanini B., Tanksley S.D.** Fine mapping of quantitative trait loci overlapping recombination chromosomes, in an interspecific cross of tomato. *Genetics*. 1990. 124: 735-742.
285. **Paterson A.H., Damon S., Hewitt J.D., Zamir D., Rabinowitch H.D., Lincoln S.E., Lander E.S., Tanksley S.D.** Mendelian factors underlying quantitative traits in tomato: comparison across scpecies, generations ans environments. *Genetics*. 1991. 127: 181-187.
286. **Paterson A.H., Lander E.S., Hewitt J.D., Peterson S., Loncoln S.E., Tanksley S.D.** Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature*. 1988. 335: 721-726.
287. **Paterson A.H., Lin Y.R., Li Z., Schertz K.F., Doebley J.F., Pinson S.R., Liu S.C., Stansel J.W., Irvine J.E.** Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. *Science*. 1995. 269: 1714-1718.
288. **Pelleschi S., Guy S., Kim J.Y., Pointe C., Mahe A., Barthes L., Leonardi A., Prioul J.L.** Ivr2, a candidate gene for QTL of vacuolar invertase activity in maize leaves. *Plant Mol. Biol.* 1999. 39: 373-380.
289. **Pelleschi S., Rocher J.P., Prioul J.L.** Effect of water restriction on carbohydrate metabolism and photosynthesis in mature maize leaves. *Plant Cell Environ.* 1997. 20: 493-503.
290. **Penner G.A., Chong J., Levesque-Lemay G., Molnar S.J., Fedak G.** Identeficationof a RAPD marker linked to the oat stem rust gene *PG3*. *Theor. Appl. Genet.* 1993. 85: 702-705.
291. **Petit R.J., Kremer A., Wagner D.B.** Finite island model for organelle and nuclear genes in plants. *Heredity*. 1993. 71: 630-641.

292. **Petit R.J., Kremer A., Wagner D.B.** Geographic structure of chloroplast DNA polymorphism in European oaks. *Theor. Appl. Genet.* 1993. 78: 122-128.
293. **Petit R.J., Brewer S., Bondacs S., Burg K., Cheddadi R., Coart E., Cottrell J., Csaikl U.M., van Dam B.C., Deans J.D., Fineschi S., Finkeldey R., Glaz I., Goicoechea P.G., Jensen J.S., König A.O., Lowe A.J., Madsen S.F., Matyas G., Munro R.C., Popescu F., Slade D., Tabbener H., de Vries S.M.G., Ziegenhagen B., de Beaulieu J.L., Kremer A.** Identification of refugia and postglacial colonization routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecol. Mgmt.* 2002. 35: 456-462.
294. **Pettersson F., Morris A.P., Barnes M.R., Cardon L.R.** Goldsurfer2 (Gs2): a comprehensive tool for the analysis and visualization of genome wide association studies. *BMC Bioinformatics.* 2008. 9: 138.
295. **Pflieger S., Lefebvre V., Causse M.** The candidate gene approach in plant genetics: a review. *Mol. Breeding.* 2001. 7: 275-291.
296. **Pielou E.C.** Ecological Diversity. Wiley, New York. 1975. 165 p.
297. **Pogson G.H., Mesa K.A., Boutilier R.G.** Genetic population structure and gene flow in the Atlantic cod *Cadus morhua*: a comparison of allozyme and nuclear RFLP loci. *Genetics.* 1995. 139: 375-385.
298. **Pons O., Chaouche K.** Estimation, variance and optimal sampling strategy of gene diversity. II. Diploid locus. *Theor. Appl. Genet.* 1996. 91: 122-130.
299. **Pons O., Petit R.J.** Estimation, variance and optimal sampling strategy. I. Haploid locus. *Theor. Appl. Genet.* 1995. 90: 462-470.
300. **Powell J.R.** Molecular techniques in population genetics: a brief history. In: *Molecular Ecology and Evolution: Approaches and Applications.* Schierwater B., Streit B., Wagner G.P., Desalle R. (eds.). Birkhauser Verlag. 1994. pp.131-157.
301. **Prioul J.L., Jeannette E., Reyss A., Gregory N., Giroux M., Hannah L.C., Causse M.** Expression of ADP-glucose pyrophosphorylase in maize (*Zea mays* L.) grain and source leaf during grain filling. *Plant Physiol.* 1994. 104: 179-187.
302. **Prioul J.L., Pelleschi S., Sene M., Thevenot C., Causse M., de Vienne D., Leonardi A.** From QTLs for enzyme activity to candidate genes in maize. *J. Exp. Bot.* 1999. 50: 1281-1288.
303. **Prioul J.L., Quarrie S., Causse M., De Vienne D.** Dissecting complex physiological functions into elementary components through the use of molecular quantitative genetics. *J. Exp. Bot.* 1997. 48: 1151-1163.
304. **Pritchard J.K., Stephens M., Donnelly P.** Inference of population structure using multilocus genotype data. *Genetics.* 2000. 155: 945-959.

305. **Pritchard J.K., Stephens M., Rosenberg N.A., Donnelly P.** Association mapping in structured populations. *Am. J. Hum. Genet.* 2000. 67: 170–181.
306. **Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J., de Bakker P.I.W., Daly M.J., Sham P.C.** PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am. J. Hum. Genet.* 2007. 81: 559–575.
307. **Quarrie S.A., Guili M., Calestani C., Steed A., Marmirolli N.** Location of a gene regulating drought-induced abscisic acid production on the long arm of chromosome 5 of wheat. *Theor. Appl. Genet.* 1994. 89: 794-800.
308. **Quarrie S.A., Laurie D.A., Zhu J.H., Lebreton C., Semikhodskii A., Steed A., Witsenboer H., Calestani C.** QTL analysis to study the association between leaf size and abscisic acid accumulation in droughted rice leaves and comparisons across cereals. *Plant Mol. Biol.* 1997. 35: 155-165.
309. **Quarrie S., Steed A., Lebreton C., Guili M., Calestani C., Marmirolli N.** Location of a gene regulating drought-induced abscisic acid production in wheat and maize and associated physiological trait. *Russian J. Plant Physiol.* 1994. 41: 565-571.
310. **Ragot M., Biasioli M., Delbut M.F., Dell'orco A., Malgarini L., Thevenin P., Vernoy J., Vivant J., Zimmermann R., Gay G.** Marker-assisted backcrossing: a practical example. In: *Techniques et utilisations des marqueurs moleculaires*. Ed. INRA, Paris. 1995. pp.45-46.
311. **Raybould A.F., Mogg R.J., Clarke R.T., Gliddon C.J., Gray A.J.** Variation and population structure at microsatellite and isozyme loci in wild cabbage (*Brassica oleraceae* L.) in Dorset (UK). *Genet. Resour. Crop Evol.* 1999. 46: 351-360.
312. **Rick C.M., DeVerna J.W., Chetelat R.T., Stevens M.A.** Meiosis in sesquidiploid hybrids of *Lycopersicon esculentum* and *Solanum lycopersicoides*. *Proc. Natl. Acad. Sci. USA.* 1986. 83: 3580-3583.
313. **Rebai A.** Comparison of methods for regression interval mapping in QTL analysis with non-normal traits. *Genet. Res. Camb.* 1997. 69: 69-74.
314. **Rebai A., Goffinet B., Mangin B.** Approximate thresholds of interval mapping tests for QTL detection. *Genetics.* 1994. 138: 235-240.
315. **Rebai A., Goffinet B., Mangin B.** Comparing power of different methods for QTL detection. *Biometrics.* 1995. 51: 87-99.
316. **Reboud X., Zeyl C.** Organelle inheritance in plants. *Heredity.* 1994. 72: 132-140.
317. **Reiter R.S., Coors J.G., Sussman M.R., Gabelman W.H.** Genetic analysis of tolerance to low phosphorus stress in maize using RFLPs. *Theor. Appl. Genet.* 1991. 82: 561-568.

318. **Reiter R.S., Williams J., Feldman K., Rafalski J.A., Tingey S.V., Scolnik P.A.** Global and local genome mapping in *Arabidopsis thaliana* recombinant inbred lines and random amplified polymorphic DNAs. *Proc. Natl. Acad. Sci. USA*. 1992. 89: 1477-1481.
319. **Ribaut J.M., Hoisington D.A., Deutsch J.A., Jiang C., Gonzalez-delon D.** Identification of quantitative trait loci under drought conditions in tropical maize. 1. Flowering parameters and the anthesis-silking interval. *Theor. Appl. Genet.* 1996. 94: 887-896.
320. **Ribaut J.M., Jiang C., Gonzalesdelon D., Edmeades G.O., Hoisington D.A.** Identification of quantitative trait loci under drought conditions in tropical maize. 2. Yield components and marker-assisted selection strategies. *Theor. Appl. Genet.* 1997. 94: 887-896.
321. **Rodolphe F., Lefort M.** A multi-marker model for detecting chromosomal segments displaying QTL activity. *Genetics*. 1993. 134: 1277-1288.
322. **Rogers J.S.** Measures of genetic similarity and genetic distance. In: *Studies in Genetics. VII. Wheeler M.R. (ed.)*. University of Texas Publ. 7213. 1972. pp.145-153.
323. **Romagosa I., UllrichS.E., Han F., Hayes P.M.** Use of the additive main effects and multiplicative interaction model in QTL mapping for adaptation in barley. *Theor. Appl. Genet.* 1996. 93: 30-37.
324. **Rommens J.M., Iannuzzi M.C., Kerm B.S., Drum M.L., Melmer G., Dean M., Rozmahel R., Cole J.L., Kennedy D., Hidaka N., Zsiga M., Buchwald M., Riordan J.R. Tsui L.C., Collins F.S.** Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science*. 1989. 245: 1059-1065.
325. **Rongwen J., Akkaya M.S., Bhagwat A.A., Lavi U., Cregan P.B.** The use of microsatellite DNA markers for soybean genotype identification. *Theor. Appl. Genet.* 1995. 90: 43-48.
326. **Saghai-Marooof M.A., Yang G.P., et al.** Correlation between molecular marker distance and hybrid performance in U.S. southern long grain rice. *Crop. Sci.* 1997. 37: 145-150.
327. **Sanguineti M.C., Tuberosa R., Landi P., Salvi S., Maccaferri M., Casarini E., Conti S.** QTL analysis of drought related traits and grain yield in relation to genetic variation for leaf abscisic acid concentration in field-grown maize. *J. Exp. Bot.* 1999. 50: 1289-1297.
328. **Sarfatti M., Katan J., Fluhr R., Zamir D.** An eflp marker in tomato linked to the *Fusarium oxysporum* resistance gene *I2*. *Theor. Appl. Genet.* 1989. 78: 755-759.
329. **Sari-Gorla M., Krajewski P., di Fronzo N., Villa M., Frova C.** Genetic analysis of drought tolerance in maize by molecular markers. II. Plant height and flowering. *Theor. Appl. Genet.* 1999. 99: 289-295.

330. **Sax K.** The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics*. 1923. 8: 552-560.
331. **Schon C.C., Lee M., Melchinger A.E., Guthrie W.D., Woodman W.L.** Mapping and characterization of quantitative trait loci affecting resistance against second-generation European corn borer in maize with the aid of RFLPs. *Heredity*. 1993. 70: 648-659.
332. **Schon C.C., Melchinger A., Boppenheimer J., Brunklaus-Jung E., Herrmann R.G., Seitzer J.F.** RFLP mapping in maize: quantitative trait loci affecting testcross performance of elite European flint lines. *Crop Sci*. 1994. 34: 378-388.
333. **Seaton G., Haley C.S., Knott S.A., Kearsley M., Visscher P.M.** QTL Express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics*. 2002. 18: 339-340.
334. **Sekhon M.S., Gupta V.P.** Genetic distance and heterosis in Indian mustard: developmental isozymes as indicators of genetic relationships. *Theor. Appl. Genet*. 1995. 91: 1148-1152.
335. **Sene M., Causse M., Damerval C., Thevenot C., Prioul J.L.** Quantitative trait loci affecting amylose, amylopectin and starch content in maize recombinant inbred lines. *Plant Physiol. Biochem*. 2000. 38: 459-472.
336. **Shewry P.R., Casey R.** (eds.) Seed Proteins. The Netherlands: Kluwer Academic Publishers. 1999. 883 p.
337. **Shewry P., Casey R.** Seed proteins. In: *Seed Proteins*. Shewry P., Casey R. (eds.). Kluwer Acad. Publishers. Dordrecht, London. 1999. pp.1-10.
338. **Shewry P., Franklin J., Parmar S., Smith S.J., Mifflin B.J.** The effects of sulphur starvation on amino acid and protein compositions of barley grain. *J. Cereal Sci*. 1983. 1: 21-31.
339. **Shwenke K.D., Raab B., Plietz P., Damaschun G.** The structure of the 12S globulin from rapeseed (*Brassica napus* L.). *Die Nahrung*. 1983. 27: 165-175.
340. **Shriver M.D., Li J., Charraborty R., Boerwinkle E.** VNTR allele frequency distributions under the stepwise mutation model: a computer simulation. *Genetics*. 1993. 134: 983-993.
341. **Simpson E.H.** Measurement of diversity. *Nature*. 1949. 163: 688.
342. **Skol A.D., Scott L.J., Abecasis G.R., Boehnke M.** Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat. Genet*. 2006. 38: 209-213.
343. **Slatkin M.** Inbreeding coefficients and coalescence times. *Genet. Res. Camb*. 1991. 58: 167-175.
344. **Slatkin M.** A measure of population subdivision based on microsatellite allele frequencies. *Genetics*. 1995. 139: 457-462.
345. **Smith J.S.C., Smith O.S.** The use of morphological, biochemical and genetic characteristics to measure distance and to test for minimum dis-

- tance between inbred lines of maize (*Zea mays* L.). UPOV Workshop. Versailles. 1989.
346. **Smith O.S., Smith J.S.C., Bowen S.L., Tenborg R.A., Wall S.J.** Similarities among a group of elite maize inbreds as measured by pedigree, F1 grain yield, grain yield, heterosis and RFLPs. *Theor. Appl. Genet.* 1990. 80: 833-840.
 347. **Smith J.S.C., Smith O.S., Bowen S.L., Tenborg R.A., Wall S.J.** The description and assessment of distances between inbred lines of maize. III. A revised scheme for the testing of distinctiveness between inbred lines utilizing DNA RFLPs. *Maydica.* 1991. 36: 213-226.
 348. **Smouse P.E., Dyer R.J., Westfall R.D., Sork V.L.** Two-generation analysis of pollen flow across a landscape. I. Male gamete heterogeneity among females. *Evolution.* 2001. 55: 260-271.
 349. **Soller M., Beckmann J.S.** Genetic polymorphism in varietal identification and genetic improvement. *Theor. Appl. Genet.* 1983. 67: 837-843.
 350. **Soller M., Brody T., Genizi A.** On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in cross between inbred lines. *Theor. Appl. Genet.* 1976. 47: 35-39.
 351. **Spencer D., Rerie W.G., Randall P.J., Higgins T.J.V.** The regulation of pea seed storage protein genes by sulphur stress. *Austr. J. Plant Physiol.* 1990. 17: 355-363.
 352. **Sprague G.F., Tatum L.A.** General vs. specific combining ability in single crosses of corn. *J. Am. Soc. Agron.* 1942. 34: 923-932.
 353. **Stam P., Zeven A.C.** The theoretical proportion of the donor genome in near-isogenic lines of self-fertilizers bred by backcrossing. *Euphytica.* 1981. 30: 227-238.
 354. **Stougaard J.** Genetics and genomics of root symbiosis. *Curr. Opin. Plant Biol.* 2001. 4: 328-335.
 355. **Streiff R., Ducouso A., Lexer C., Streinkellner H., Gloessl J., Kremer A.** Pollen dispersal inferred from paternity analysis in mixed oak stand of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. *Mol. Ecol.* 1999. 8: 831-841.
 356. **Stromberg L.D., Dudley J.W., Rufener G.K.** Comparing conventional early generation selection with marker assisted selection in maize. *Crop Sci.* 1994. 34: 1221-1225.
 357. **Stuber C.W.** Marker-based selection for quantitative traits. *Vortr. Pflanzenzuchtg.* 1989. 16: 31-49.
 358. **Stuber C.W.** Mapping and manipulating quantitative traits in maize. *Trends Genet.* 1995. 11: 477-481.
 359. **Stuber C.W., Edwards M.D., Wendel J.F.** Molecular-marker-facilitated investigations of QTL in maize. II. Factors influencing yield and its component traits. *Crop Sci.* 1987. 27: 639-648.

360. **Stuber C.W., Lincoln S.E., Wolf D.W., Helebtjaris T., Lander E.S.** Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics*. 1992. 132: 823-839.
361. **Szmidt A.E., Wang X.R., Lu M.Z.** Empirical assessment of allozyme and RAPD variation in *Pinus sylvestris* (L.) using haploid tissue analysis. *Heredity*. 1996. 76: 412-420.
362. **Tabe L., Hagan N., Higgins T.J.V.** Plastisity of seed protein composition in response to nitrogen and sulphur availability. *Curr. Opin. Plant Biol.* 2002. 5: 212-217.
363. **Tajima F.** Measurement of DNA polymorphism. In: *Mechanisms of Molecular Evolution*. Takahata N., Clark A.G. (eds.). Sinauer Associates. 1993. pp.37-59.
364. **Tajima F., Nei M.** Estimation of evolutionary distance between nucleotide sequences. *Mol. Biol. Evol.* 1984. 1: 269-285.
365. **Tanhuanpaa P.K., Vilkki J.P., Vilkki H.J.** Mapping of a QTL for oleic acid concentration in spring turnip rape (*Brassica rapa* spp. *Olifera*). *Theor. Appl. Genet.* 1996. 92: 952-956.
366. **Tanksley S.D.** Molecular markers in plant breeding. *Plant Mol. Biol. Rept.* 1983. 1: 3-8.
367. **Tanksley S.D.** Mapping polygenes. *Ann. Rev. Genet.* 1993. 27: 205-233.
368. **Tanksley S.D., Ganai M.W., Martin G.B.** Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes. *Trends Genet.* 1995. 11: 63-68.
369. **Tanksley S.D., Grandillo S., Fulton T.M., Zamir D., Eshed Y., Petiard V., Lopez J., Beckbunn T.** Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. *Theor. Appl. Genet.* 1996. 92: 213-224.
370. **Tanksley S.D., Hewitt J.** Use of molecular markers in breeding for soluble solids content in tomato: a re-examination. *Theor. Appl. Genet.* 1988. 75: 811-823.
371. **Tanksley S.D., Jones R.A.** Application of alcohol dehydrogenase allozymes in testing the purity of F₁ hybrids of tomato. *Hort. Science*. 1979. 16:179-181.
372. **Tanksley S.D., Medina-Filho H., Rick C.M.** Use of naturally-occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. *Heredity*. 1982. 49: 11-25.
373. **Tanksley S.D., Nelson N.** Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.* 1996. 92: 191-203.

374. **Tanksley S.D., Orton T.J.** Isozymes in plant genetics and breeding. Elsevier. 1983. Part A. 516 pp.
375. **Tanksley S.D., Rick C.M.** Isozymic gene linkage map of the tomato: application in genetics and breeding. *Theor. Appl. Genet.* 1980. 57:161-170.
376. **Tersac M., Blanchard P., Brunel D., Vincourt P.** relations between heterosis and enzymatic polymorphism in populations of cultivated sunflowers (*Heliantus annuus* L.). *Theor. Appl. Genet.* 1994. 88: 49-55.
377. **Teulat B., This D., Khairallah M., Borries C., Ragot C., Sourdille P., Leroy P., Monneveux P., Charrier A.** Several QTLs involved in osmotic adjustment trait variation in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 1998. 96: 688-698.
378. **Thoday J.M.** Location of polygenes. *Nature.* 1961. 191: 368-370.
379. **Timmerman-Vaughan G.M., McCallum J.A., Frew T.J., Weeden N.F., Russell A.C.** Linkage mapping of quantitative trait loci controlling seed weight in pea (*Pisum sativum* L.). *Theor. Appl. Genet.* 1996. 93: 431-439.
380. **Touzet P., Winkler R.G., Helentjaris T.** Combined genetic and physiological analysis of a locus contributing to quantitative variation. *Theor. Appl. Genet.* 1995. 91: 200-205.
381. **Tozlu I., Guy C.L., Moore G.A.** QTL analysis of Na⁺ and Cl⁻ accumulation related traits in an intergeneric BC₁ progeny of *Citrus* and *Poncirus* under saline and nonsaline environments. *Genome.* 1999. 42: 692-705.
382. **Tripathy J.N., Zhang J., Robin S., Nguyen T.T., Nguyen H.T.** QTLs for cell-membrane stability mapped in rice (*Oryza sativa* L.) under drought stress. *Theor. Appl. Genet.* 2000. 100: 1197-1202.
383. **Tuberosa R., Sanguineti M.C., Landi P., Salvi S., Caarini E., Conti S.** RFLP mapping of quantitative trait loci controlling abscisic acid concentration in leaves of drought-stressed maize (*Zea mays* L.). *Theor. Appl. Genet.* 1998. 97: 744-755.
384. **Tuinstra M.R., Grote E.M., Goldsdrough P.B., Ejeta G.** Genetic analysis of post-flowering drought tolerance and components of grain development in *Sorghum bicolor* (L.) Moench. *Mol. Breed.* 1997. 3: 439-448.
385. **Ukai Y., Osawa R., Saito A., Hayashi T.** MAPL: a package of computer programs for construction of DNA polymorphism linkage maps and analysis of QTL. *Breed. Sci.* 1995. 45: 139-142 (in Japanese).
386. **Utz H.F., Melchinger A.E.** PLABQTL: a program for composite interval mapping of QTL. *Journal of Agricultural Genomics.* – 1996. – <http://www.cabi-publishing.org/jag/papers96/paper196/indexp196.html> (доступно с 30 июня 2007 г.).

387. **Vallejos C.E.** Enzyme activity staining. In: *Isozymes in Plant Genetics and Breeding*. Tanksley S.D., Orton T.J. (eds.). Elsevier, Amsterdam. 1983. pp.37-49.
388. **Van der Beek J.G., Verkerk R., Zabel P., Lindhout P.** Mapping strategy for resistance genes in tomato based on RFLPs between cultivars: *Cf9* (resistance to *Clamidosporium fulvum*) on chromosome 1. *Theor. Appl. Genet.* 1992. 84: 106-112.
389. **Van Hintum T.J.L.** Comparison of marker systems and construction of a core collection in a pedigree of European spring barley. *Theor. Appl. Genet.* 1994. 89: 991-997.
390. **Van Ooijen A.J.** JoinMap 4.0: Software for the Calculation of Genetic Linkage Maps in Experimental Populations. Plant Research International B.V & Kyazma B.V. Wageningen, Netherlands. 2006. 45 p.
391. **Van Ooijen J.W.** MapQTL 6. Software for the mapping of quantitative trait loci in experimental populations of diploid species. Wageningen, Netherlands. 2009. 60 p.
392. **Veldboom L.R., Lee M.** Molecular marker facilitated studies of morphological traits in maize. II. Determination of QTLs for grain yield and yield components. *Theor. Appl. Genet.* 1994. 89: 451-458.
393. **Vision T.J., Brown D.G., Shmoys D.B., Durrett R.T., Tanksley S.D.** Selective mapping: a strategy for optimizing the construction of high-density linkage maps. *Genetics.* 2000. 155: 407-420.
394. **Walbot V., Cullis C.A.** Rapid genome change in higher plants. *Ann. Rev. Plant Physiol.* 1985. 36: 367-396.
395. **Wang J., Bernardo R.** Variance of marker estimates of parental contributions to F2 and BC1 derived inbreds. *Crop Sci.* 2000. 40: 659-665.
396. **Wang G.L., Paterson A.H.** Assessment of DNA pooling strategies for mapping of QTLs. *Theor. Appl. Genet.* 1994. 88: 355-361.
397. **Ward J.H.** Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assos.* 1963. 58: 236-244.
398. **Ward E.R., Jen C.C.** Isolation of single-copy-sequence clones from a yeast artificial chromosome library of randomly-sheared *Arabidopsis thaliana* DNA. *Plant Mol. Biol.* 1990. 14: 561-568.
399. **Ware D.H., Jaiswal P., Ni J., Yap I.V., Pan X., Clark K.Y., Teytelman L., Schmidt S.C., Zhao W., Chang K., Cartinhour S., Stein L.D., McCouch S.R.** Gramene, a tool for grass genomics. *Plant Physiol.* 2002. 130: 1606-1613.
400. **Weir B.S.** Genetic Data Analysis. Sinauer associates, New York. 1990. 377 p.
401. **Weir B.S., Basten J.C.** Sampling strategies for distances. *Biometrics.* 1990. 46: 551-582.
402. **Weir B.S., Cockerham C.C.** Estimating F-statistics for the analysis of population structure. *Evolution.* 1984. 38: 1358-1370.

403. **Weller J.I.** Mapping and analysis of quantitative trait loci in *Lycopersicon* (tomato) with the aid of genetic markers using approximate maximum likelihood methods. *Heredity*. 1987. 59: 413- 421.
404. **Wright S.** The genetical structure of populations. *Ann. Eugen.* 1951. 15: 323-354.
405. **Wright S.** Evolution and the Genetics of Populations. Vol.1. Genetics and Biometric Foundations. The University of Chicago Press. 1968. 480 p.
406. **Wright S.** Evolution and the Genetics of Populations. Vol.2. The Theory of Gene Frequencies. The University of Chicago Press. 1969. 520 p.
407. **Xiao J., Yuan L., McCouch S.R., Tanksley S.D.** Genetic diversity and its relationship to hybrid performance and heterosis in rice as revealed by PCR-based markers. *Theor. Appl. Genet.* 1996. 92: 647-643.
408. **Xu W.W., Subudhi P.K., Crasta O.R., Rosenow D.T., Mullet J.E., Nguyen H.T.** Molecular mapping of QTLs conferring stay-green in grain sorghum (*Sorghum bicolor* L. Moench). *Genome*. 2000. 43: 461-469.
409. **Yadav R., Courtois B., Huang N., McLaren G.** Mapping genes controlling root morphology and root distribution in a double-haploid population of rice. *Theor. Appl. Genet.* 1997. 94: 619-932.
410. **Yang J., Hu C., Hu H., Yu R., Xia Z., Ye X., Zhu J.** QTLNetwork: mapping and visualizing genetic architecture of complex traits in experimental populations. *Bioinformatics*. 2008. 24: 721–723.
411. **Yano M., Katayose Y., Ashikari M., Yamanouchi U., Monna L., Fuse T., Baba T., Yamamoto K., Umehara Y., Nagamura Y., Sasaki T.** *Hdl* a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. *Plant Cell*. 2000. 12: 2473-2484.
412. **Young N.D.** A cautiously optimistic vision for marker-assisted breeding. *Mol. Breed.* 2000. 5: 500-510.
413. **Young N.D., Tanksley S.D.** RFLP analysis of the size of the chromosomal segments retained around the *Tm-2* locus of tomato during back-cross breeding. *Theor. Appl. Genet.* 1989. 77: 353-359.
414. **Zamir D., Tanksley S.D., Jones R.A.** Genetic analysis of the origin of plants regeneration from anther tissues of *Lycopersicon esculentum* Mill. *Plant Sci. Letters*. 1981. 21:223-227.
415. **Zamir D., Tanksley S.D., Jones R.A.** Haploid selection for low temperature tolerance of tomato pollen. *Genetics*. 1982. 101:129-137.
416. **Zanetto A., Roussel G., Kremer A.** Geographic variability of inter-specific differentiation between *Quercus robur* L., *Quercus petraea* (Matt.) Liebl. *Forest Genet.* 1994. 1: 111-123.

417. **Zehr B.E., Dudley J.W., Chojeki J., Maroof M.A.S., Mowers R.P.** Use of RFLP markers to search for alleles in a maize population for improvement of an elite hybrid. *Theor. Appl. Genet.* 1992. 83: 903-911.
418. **Zeng Z.B.** Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proc. Natl. Acad. Sci. USA.* 1993. 90: 10972-10976.
419. **Zeng Z.B.** Precision mapping of quantitative trait loci. *Genetics.* 1994. 136: 1457-1468.
420. **Zhang J.X., Klueva N.Y., Wang Z., Wu R., Ho T.H.D., Nguyen H.T.** Genetic engineering for abiotic stress resistance in crop plants. *In Vitro Cell. Develop. Biol. Plant.* 2000. 36: 108-114.
421. **Zhang Q., Saghai-Maroof M.A., Kleinhofs A.** Comparative diversity analysis of RFLPs and isozymes within and among populations of *Hordeum vulgare* ssp. *spontaneum*. *Genetics.* 1993. 134: 909-915.
422. **Zhang W., Smith C.** Computer simulation of marker-assisted selection utilizing linkage disequilibrium. *Theor. Appl. Genet.* 1992. 83: 813-820.
423. **Zhang Z., Bradbury P.J., Kroon D.E., Casstevens T.M., Buckler E.S.** TASSEL 2.0: a software package for association and diversity analyses in plants and animals. Poster P956/CP012 presented at Plant and Animal Genomes XIV Conference, 14–18 January 2006. San Diego, California. 2006. http://www.intl-pag.org/14/abstracts/PAG14_C012.html

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	5
Глава I. НЕКОТОРЫЕ АСПЕКТЫ ИСПОЛЬЗОВАНИЯ МОЛЕКУЛЯРНЫХ И ТРАДИЦИОННЫХ МАРКЕРОВ В ПОПУЛЯЦИОННОЙ ГЕНЕТИКЕ	12
1.1. Мультиаллелизм	12
1.2. Увеличение числа локусов.....	13
1.3. Анализ полиморфизма цитоплазматической ДНК.....	14
1.4. Распределение аллелей.....	15
1.5. Анализ молекулярного разнообразия	16
1.5.1. Установление молекулярной дивергенции на основе первичной структуры нуклеотидной последовательности	17
1.5.2. Установление нуклеотидной дивергенции методами маркерных технологий.....	19
1.6. Полиморфизм внутри популяции	22
1.6.1. Выражение внутривидового полиморфизма.....	22
1.6.1.1. Показатели полиморфизма, принимающие во внимание (в расчет) число вариантов.....	23
1.6.1.2. Показатели полиморфизма, принимающие во внимание (в расчет) частоты вариантов.....	23
1.7. Оценка уровня разнообразия с помощью различных типов маркеров	27
1.7.1. На уровне аллелей	27
1.7.2. Особый случай доминантных маркеров.....	28
1.7.3. Оценка на нуклеотидном уровне	30
1.8. Установление различий между популяциями ГРР.....	30
1.8.1. Общее выражение дифференциации	31
1.8.2. Измерение дифференциации внутри маркерных систем	33
1.8.2.1. Уровень аллелей	33
1.8.2.2. Уровень нуклеотидов	33
1.9. Поток генов.....	35

1.9.1. Непрямое измерение потока генов	35
1.9.1.1. Сравнительные уровни миграции семян и пыльцы	35
1.9.1.2. Происхождение колонизирующих генотипов	38
1.9.2. Прямое измерение потока генов	39
Заключение	41
Глава II. НЕКОТОРЫЕ АСПЕКТЫ ПРИМЕНЕНИЯ	
МОЛЕКУЛЯРНЫХ МАРКЕРОВ В СЕЛЕКЦИИ	43
2.1. Вклад изучения разнообразия в практическую селекцию.....	43
2.1.1. Основные методы анализа данных	44
2.1.2. Взаимоотношения между молекулярной дивергенцией, фенотипической дивергенцией и сопроисхождением	46
2.2. Применение методов классификации: гетерозисные группы и управление генетическими ресурсами	48
2.2.1. Взаимосвязь между молекулярной дивергенцией и гетерозисом	50
2.3. Использование маркеров для защиты новых сортов	54
2.4. Перспективы изучения генетического разнообразия.....	55
2.5. Маркерная помощь селекции (отбору).....	56
2.5.1. Создание генотипов.....	57
2.5.1.1. Маркерная помощь при беккроссировании генотипов с моногенным признаком	57
2.5.1.2. Маркерная помощь при беккроссировании полигенного признака	60
2.5.1.2.1. Перенос хромосомного сегмента, несущего QTL.....	60
2.5.1.2.2. Перенос нескольких хромосомных сегментов	61
2.5.1.3. Маркерная помощь при отборе по потомству	62
2.5.2. Маркерная помощь рекуррентной селекции (рекуррентному отбору)	64
2.5.2.1. Рекуррентный отбор, основанный только на маркерах.....	65
2.5.2.2. Рекуррентный отбор по аддитивному значению, прогнози- руемому с помощью маркеров	66

2.5.3. Комбинированный отбор, основанный на фенотипе и маркерах	68
2.5.3.1. Принцип и ожидаемая эффективность.....	68
2.5.3.2. Результаты моделирования.....	71
2.5.3.3. Сравнение на основе создания генотипов	71
2.5.3.4. Выбор между использованием маркеров и увеличением числа репликаций (репродукции, повторности, воспроизводства).....	72
Заключение	73
Глава III. КАРТИРОВАНИЕ И ХАРАКТЕРНЫЕ ОСОБЕННОСТИ ХРОМОСОМНЫХ ЛОКУСОВ, ОПРЕДЕЛЯЮЩИХ ПРОЯВЛЕНИЕ КОЛИЧЕСТВЕННЫХ ПРИЗНАКОВ В ПОПУЛЯЦИИ.....	
3.1. Принципы картирования QTL.....	77
3.1.1. Определение QTL, учитывающих маркеры обособленно.....	78
3.1.2. Нахождение QTL посредством анализа вариантов популяции F_2	78
3.1.3. Измерение эффектов QTL.....	81
3.1.4. Ограничения определения QTL по одиночным маркерам	82
3.2. Определение QTL по двум или более маркерам	82
3.2.1. Методы интервального картирования	82
3.2.2. Методы множественных маркеров	84
3.3. Факторы, влияющие на определение QTL.....	85
3.3.1. Определение силы и изменчивости внутри классов.....	86
3.3.2. Размер популяции и плотность маркеров	87
3.3.3. Выбор статистических рисков.....	88
3.3.4. Отклонение от нормальности	89
3.4. Достоинства и недостатки популяций, наиболее часто используемых для определения QTL	89
3.4.1. Сравнение популяций F_2 , беккроссов, удвоенных гаплоидов и рекомбинантных инбредных линий	89
3.4.2. Потомство, полученное от гетерозиготных родителей	91

3.4.3. Популяции с частично фиксированным генетическим окружением.....	91
3.4.4. Селективное генотипирование	93
3.5. Генетическая и молекулярная основа вариации количественных признаков	94
3.5.1. QTL (почти) всегда определяемы	94
3.5.2. Эффекты доминирования.....	96
3.5.3. Трансгрессия	96
3.5.4. Эпистаз.....	97
3.5.5. Генетический анализ корреляций между признаками.....	99
3.5.6. Эффекты действия окружающей среды	102
3.5.7. Сравнительное картирование QTL	103
3.5.8. Некоторые характерные особенности QTL	104
3.5.8.1. Клонирование QTL.....	105
3.5.8.2. Гены-кандидаты.....	106
3.5.8.3. «Функциональные» гены-кандидаты	107
3.5.8.4. «Позиционные» гены-кандидаты.....	108
3.5.9. Проверка генов-кандидатов.....	109
3.5.9.1. Секвенирование аллелей генов-кандидатов	110
3.5.9.2. Определение экспрессии QTL или активности генов-кандидатов	110
3.5.9.3. Трансформация	110
3.5.9.4. Количественная комплементация	111
Заключение	112
Глава IV. КАРТИРОВАНИЕ ОСНОВНЫХ (МЕНДЕЛЕВСКИХ) ГЕНОВ	114
4.1. Подходы к картированию основных (менделевских) генов	114
4.1.1. Практически изогенные линии.....	114
4.1.2. Совокупный сегрегационный анализ	116
4.2. Использование маркеров для клонирования основных (менделевских) генов	119

4.2.1. Принцип позиционного клонирования.....	119
4.2.2. Картирование высокого разрешения	120
4.2.3. Поиск необходимого гена в геномных клонах	120
Заключение	121
Глава V. АБИОТИЧЕСКИЙ СТРЕСС И ЛОКУСЫ КОЛИЧЕСТВЕННЫХ БИОХИМИЧЕСКИХ И ФИЗИОЛОГИЧЕСКИХ ПРИЗНАКОВ У РАСТЕНИЙ	124
5.1. Некоторые элементы методологии	126
5.2. QTL и абиотический стресс	127
5.2.1. Высокая температура (термоустойчивость)	127
5.2.2. Состояние засухи на вегетативных стадиях развития	128
5.2.2.1. Корневая система, морфология и признаки проростков у риса.....	128
5.2.2.2. Осмотолерантность у риса.....	130
5.2.2.3. Листья и абсцизовая кислота (АК) у кукурузы	130
5.2.2.4. Клеточно-мембранная стабильность у риса	133
5.3. Условия засухи на стадии созревания и QTL компонентов урожая.....	134
5.3.1. Кукуруза.....	134
5.3.2. Сорго	138
5.3.3. Засоление	140
5.4. Идентификация генов-кандидатов.....	142
5.5. QTL биохимических признаков и гены-кандидаты	144
5.5.1. QTL активности ферментов.....	144
5.5.2. QTL анализа белков 2D-электрофорезом.....	148
Заключение	149
ГЛАВА VI. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ ГЕНЕТИЧЕСКОГО КАРТИРОВАНИЯ И АНАЛИЗА АССОЦИАЦИЙ МАРКЕР/ПРИЗНАК.....	151
6.1. Построение генетической карты	152
6.2. Картирование QTL, основанное на сцеплении	154

6.3. Картирование посредством установления неравновесного сцепления	159
6.3.1. Картирование QTL, основанное на LD	159
6.3.2. Полногеномное ассоциативное	160
Заключение	162
СПИСОК ЛИТЕРАТУРЫ	163

**Yuriy Valentinovich Chesnokov
Natalia Viktorovna Kocherina
Vladimir Mikhailovich Kosolapov**

**Юрий Валентинович Чесноков
Наталья Викторовна Кочерина
Владимир Михайлович Косолапов**

**Molecular markers in population genetics
and breeding of cultivated plants**

**Молекулярные маркеры в популяционной генетике
и селекции культурных растений**

Научное издание

Верстка, оригинал-макет Н. И. Георгиади

Подписано в печать 23.04.2019 г.
Бумага «Снегурочка». Формат 60×84 ¹/₁₆.
Гарнитура «Таймс». Печать ризографическая
Усл. печ. л. 11,1. Тираж 1000. Заказ 026

ООО «Угрешская Типография»
т. 700-12-29, 700-06-66
111621, Москва, ул. Оренбургская, 15