

Молекулярные маркеры – инструмент исследования генетического разнообразия

1 Введение

ДНК маркеры полезны в фундаментальных (напр., в филогенетическом анализе и в поиске полезных генов) и прикладных исследованиях (напр., в мар-

керной селекции, при установлении отцовства и контроле движения пищевых продуктов). В этой части рассматривается, главным образом, их примене-

Вставка 70 ДНК, РНК и белки

ДНК (дезоксирибонуклеиновая кислота) организована в пары хромосом, каждая из которых наследуется от одного из родителей. Каждый ген индивидуума, следовательно, представлен двумя копиями, называемыми аллелями, по одной на каждой хромосоме из пары. У млекопитающих гены расположены вдоль хромосом, разделены длинными, часто повторяющимися последовательностями ДНК. Гены состоят из кодирующих последовательностей (экзоны), разделенных интронами. Последние не несут информации, кодирующей белки, но иногда играют роль в регуляции генной экспрессии. Инструкция, кодируемая генами, реализуется в двух процессах. Первый – транскрипция (копирование) генетической информации в другой тип нуклеиновой кислоты, РНК (рибонуклеиновая кислота). Экзоны и интроны транскрибируются в молекулу первичной матричной РНК (мРНК). Затем эта молекула редактируется, процесс редактирования включает удаление интронов, объединение экзонов и добавление уникальных изменений к каждому из концов мРНК. Таким образом создается зрелая молекула мРНК, которая затем транспортируется к структурам, известным как рибосомы, расположенным в цитоплазме клетки. Рибосомы состоят из рибосомной РНК (рРНК) и белков, и обеспечивают второй процесс – трансляцию генетической информации,

ранее скопированной в мРНК, в полипептид (собственно белок или одну из цепей белкового комплекса). Молекула мРНК прочитывается, или транслируется по три нуклеотида (или кодона) за раз. Комплементарность между кодоном мРНК и антикодоном транспортной РНК (тРНК), которая приносит соответствующую аминокислоту к рибосоме, обеспечивает формирование нового полипептида, содержащего специфическую последовательность аминокислот.

Не все гены транслируются в белки; некоторые выполняют свои функции как молекулы РНК (такие как рРНК и тРНК, участвующие в трансляции). Недавно были открыты новые функции РНК в процессе редактирования мРНК и регуляции генной экспрессии (Storz и др., 2005; Aravin, Tuschl, 2005; Wienholds, Plasterk 2005). Несомненно, некодирующая РНК может быть ключевым участником в различных регуляторных процессах (Vertone и др., 2004; Сlor и др., 2006). Таким образом, для исследований генетических характеристик на клеточном, тканевом и организменном уровнях доступны три типа молекул: ДНК, содержащая закодированную инструкцию, РНК, переносящая эту инструкцию в клеточную «фабрику»; и белки, которые построены согласно этой инструкции и обеспечивают функционирование клеток и организмов.

РАЗДЕЛ 4

ние для характеристики разнообразия ГРЖ, и поиска функциональных вариантов важных генов. Важно отметить, что РНК и белки тоже содержат ключевую информацию, и поэтому заслуживают параллельных исследований; их роль в поисках функциональных вариантов также рассматривается ниже.

Разнообразие организмов обусловлено изменчивостью последовательностей ДНК и влиянием факторов среды. Генетическая изменчивость значительна, и каждый индивидуум данного вида, за исключением монозиготных близнецов, несет уникальные последовательности ДНК. ДНК-варианты являются следствием мутаций, происходящих вследствие замены одного нуклеотида (однонуклеотидный полиморфизм – single nucleotide polymorphisms, SNP), вставок или потерь фрагментов ДНК разной длины (от одного до нескольких тысяч нуклеотидов), или дубликаций и инверсий фрагментов ДНК. ДНК-варианты классифицируются как «нейтральные», если они не вызывают изменений в метаболических или фенотипических характеристиках, и, следовательно, не подвергаются ни позитивному, ни негативному, ни балансирующему отбору; все остальные варианты рассматриваются как «функциональные». Мутации в ключевых нуклеотидах кодирующей последовательности могут изменять аминокислотную последовательность белка и приводить к появлению новых функциональных вариантов. Такие варианты могут увеличивать или уменьшать метаболическую эффективность по сравнению с исходным «диким типом», могут полностью

утрачивать свою функцию, или даже добавлять новую функцию. Мутации в регуляторной области могут изменять уровни и характер экспрессии гена; например, переключение генов по схеме вкл./выкл. или ниже/выше при экспрессии белков в определенных тканях на различных стадиях развития или в разных физиологических состояниях.

Хотя доказано, что анализ одного типа биомолекул очень много дает для понимания биологических явлений, параллельные крупномасштабные исследования ДНК, РНК и белков открывают новые перспективы в интерпретации и моделировании сложности живых организмов.

Вставка 72 Современные направления развития молекулярной биологии

Текущие революционные достижения в молекулярно-биологических исследованиях в области разведения домашнего скота и сохранения его генетического разнообразия включают:

- 1) полное секвенирование геномов наиболее важных видов домашних животных;
- 2) разработку методов оценок полиморфизма по локусам, распространенным по всему геному (например, методы определения SNP);
- 3) разработку технологии микроматриц для измерения транскрипции генов в большом масштабе.

Информация, получаемая в результате секвенирования всего генома (завершено для кур и почти полностью – для свиней и крупного рогатого скота), объединяемая с SNP-технологиями, будет существенно ускорять поиск важных генов. Картирование главных генов количественных признаков (Quantitative trait loci - QTL) для идентификации участков хромосом, влияющих на признаки - мишени действия этих генов, и выявления присутствия генов-кандидатов, локализованных в этих районах, исследование характера их экспрессии (напр., с использованием микроматриц или протеомного анализа) и функции у разных видов, все это взятое вместе дает возможность идентифицировать ключевые гены и вскрывать комплексность физиологической регуляции признаков – мишеней исследований.

Дальнейшее обсуждение этих достижений см. ниже.

Вставка 71 Новые «–омик» научные дисциплины

Геномика картирует гены и генетические варианты у индивидуумов и групп. Это позволяет приблизиться к пониманию перевода генетической информации в метаболические функции и фенотипические характеристики. Такой подход вскрывает биологические процессы и их взаимодействия с факторами среды. Геномика включает в себя комбинацию высокоточных технологий, таких как протеомика и метаболомика, с методами биоинформатики, которые обеспечивают возможность обработки, анализа и интеграции большого количества данных.

Появились новые научные дисциплины с суффиксом «-омика». В этих областях современные достижения в приготовлении, идентификации и секвенировании ДНК, РНК и белков, и крупномасштабное накопление и анализ данных, привели к революции в наших знаниях. Возникла глобальная, интегрированная точка зрения на внутреннюю сеть биологических молекул, вовлеченных в комплексные биологические процессы. Структурная геномика, транскриптомика, протеомика, за которыми следует метаболомика, интерактомика и еще более высокие уровни сложности, формируют системную биологию (Hood и др., 2004; Vox 71).

Изучение комплексных биологических явлений - новое передовое направление исследований, которые требуют мощных молекулярных технологий, компьютеров с высокими скоростями и большой памятью, новых подходов к анализу данных и интегрированности междисциплинарных оценок (вставка 72).

2 Роль молекулярных технологий в описании генетического разнообразия

Информация по генетическому разнообразию существенна для оптимизации стратегий сохранения и использования ГРЖ. Если возможности сохранения ограничены, часто возникает необходимость определения приоритетов. Принято считать, что новые молекулярные инструменты позволяют идентифицировать гены, участвующие в формировании многих признаков, включая адаптивные признаки, и полиморфизм, приводящий к функциональным генетическим вариантам (QTN – нуклеотиды количественных признаков - Quantitative Trait Nucleotides). Однако у нас нет достаточных знаний для того, чтобы на основании функционального молекулярного разнообразия строить выбор в отношении приоритетности сохранения, до сих пор необходимы альтернативные оценки. Фенотипическое описание обеспечивает грубую оценку средних значений по функциональным вариантам генов, присутствующих у данных индивидуумов или популяций. Однако до сих пор большинство фенотипов большинства домашних видов остаются не оцененными.

Первая роль. В отсутствие надежных фенотипических и QTN данных, или в дополнение к существующим данным, наиболее быстрой и рентабельной мерой генетического разнообразия является оценка полиморфизма с использованием анонимных молекулярно-генетических маркеров. При условии, что уникальные популяции, имевшие специфическую эволюционную историю по нейтральным маркерам (например, в связи с изоляцией предков или независимой доместикацией), по-видимому, будут нести уникальные варианты функциональной изменчивости, анонимные маркеры, вероятно, могут дать непрямую информацию о функциональных генах важных признаков. Молекулярные методы оказываются полезными при исследованиях происхождения и доместикации домашних видов и их последующей миграции, а также для получения информации по эволюционным взаимосвязям (филогенетические деревья), и установления географических областей скрещиваний между популяциями, имеющими разное генетическое происхождение. В главе 3.1 представлена схема молекулярных методов оценки генетического разнообразия внутри и между породами.

Вторая роль. Эффективная численность популяции (N_e) является показателем, на основании которого оценивается эффективное число животных в популяции, участвующих в воспроизводстве и вносящих свои гены в следующее поколение. N_e тесно связана с уровнем инбридинга и генетическим дрейфом в популяции и, следовательно, является важным показателем для оценки степени опасности для популяции (см. части А и Е). Традиционные подходы к получению надежных оценок N_e для племенных популяций основываются на родословных или на переписи. В развивающихся странах часто невозможно получить необходимые данные об изменчивости репродуктивного успеха и интервала между поколениями в популяциях. Поэтому молекулярные подходы могут представлять многообещающую альтернативу (подробнее см. в главе 3.2).

Третья роль. Главным приоритетом в управлении ГРЖ является сохранение пород с уникальными признаками. Самое большое значение, особенно для развивающихся стран, имеют способность жить и давать продукцию в критических условиях и обладать устойчивостью к инфекционным заболеваниям.

РАЗДЕЛ 4

Адаптация и устойчивость к заболеваниям – комплексные неморфологические признаки, их нелегко измерить. Их можно исследовать в экспериментах, в которых животные подвергаются специфическим средовым воздействиям или инфицируются соответствующими агентами. Однако такие эксперименты трудоемки и дорогостоящи и поднимают вопросы о защите животных. Именно поэтому чрезвычайный интерес исследователей вызывает идентификация генов, контролирующих сложные признаки. Для выявления таких генов используются разные подходы. Созданные инструменты для выявления функциональной изменчивости описаны в главе 3.3.

3 Краткий обзор молекулярных методов

В этом разделе описаны наиболее важные современные молекулярные методы, разработанные для оценки генетического разнообразия и выявления функциональных вариантов. Во вставке 73 описывается, как ДНК и РНК экстрагируются из биологического материала и подготавливаются для анализа. Характеристики чаще всего используемых молекулярных маркеров приведены во вставке 74, и составление выборки (очень важный аспект молекулярных исследований) обсуждается во вставке 75.

Полиморфизм белков был первым поколением маркеров, используемых для генетических исследований домашних животных. Однако количество полиморфных локусов, доступных для анализа, и уровень их наблюдаемого полиморфизма часто низки, что сильно ограничивает их применение в исследованиях генетического разнообразия. Благодаря развитию новых технологий, полиморфизмы ДНК стали предпочтительными маркерами в исследованиях генетической изменчивости с использованием молекулярных методов (вставка 74).

3.1 Методы, использующие ДНК-маркеры для оценки генетического разнообразия

Маркеры ядерной ДНК

В настоящее время для определения полиморфизма ядерной ДНК в распоряжении имеется ряд

Вставка 73 Экстракция и наработка ДНК и РНК

Первый шаг в анализе ДНК, РНК и белка – их выделение из биологического образца и очистка. Для этого имеется целый ряд протоколов и готовых реакционных смесей (китов). Применяемая стратегия определяется источником материала и выделяемыми молекулами. Например, экстракция ДНК из цельной крови или лейкоцитов относительно легка, в то время как ее экстракция из обработанных пищевых продуктов много труднее. Экстракция РНК из поджелудочной железы очень трудна из-за высокой скорости процессов посмертной деградации в этом органе. Очистка ДНК, РНК и белков является ключевым фактором, от которого зависит надежность конечного результата.

После выделения ДНК (или РНК) из клеток, следующим шагом является получение тысяч или миллионов копий определенного гена или участка ДНК. Фрагмент ДНК может быть размножен в микроорганизмах, обычно *E. coli*, или *in vitro* с использованием полимеразной цепной реакции (ПЦР). Этим методом, принесшим Нобелевскую Премию его создателю, Кэри Мюллису, экспоненциально умножают (амплифицируют) любой сегмент ДНК с известной последовательностью. Ключевым компонентом в ПЦР является ДНК-полимераза, выделенная из *Thermus aquaticus*, микроорганизма, адаптированного к жизни и размножению при очень высоких температурах. Эта термостабильная *Taq*-полимераза (по названию *Thermus aquaticus*) обеспечивает репликацию цепей ДНК в циклическом режиме и приводит к геометрическому росту количества копий ДНК – мишени амплификации. Цикл ПЦР включает три этапа: i) денатурация ДНК при 90–95°C для разделения молекулы ДНК на две одноцепочечные последовательности, которые служат матрицами; ii) отжиг пар коротких одноцепочечных олигонуклеотидов (праймеров, или затравок), комплементарных к областям – мишеням, фланкирующим амплифицируемый фрагмент (фрагмент интереса), при 45–65°C; iii) удлинение вновь синтезируемых цепочек ДНК, начиная от праймеров, осуществляемое *Taq*-полимеразой при 72°C. Этот цикл повторяется, обычно 25–45 раз, позволяя накапливать достаточное количество ампликонов (фрагмент гена или ДНК, синтезируемый с использованием ПЦР) для анализа.

Вставка 74

Широко используемые ДНК-маркеры

Полиморфизм длин рестриционных фрагментов (ПДРФ/RFLP) определяется при использовании ферментов рестрикции (рестриктаз), которые разрезают ДНК только в точных местах «сайтов рестрикции» (например, рестриктаза EcoRI разрезает ДНК в сайте, определяемом палиндромной последовательностью GAATTC). В настоящее время наиболее часто ПДРФ/RFLP используют вслед за ПЦР (ПЦР–ПДРФ/PCR–RFLP), для того, чтобы выявить аллели, отличающиеся по нуклеотидным последовательностям в сайте рестрикции. Фрагмент гена амплифицируется с использованием ПЦР и затем обрабатывается специфическим ферментом рестрикции, который разрезает только одну аллельную форму. Переваренные таким образом ампликоны затем разделяются электрофорезом.

Микросателлиты, или SSR (Simple Sequence Repeats), или STR (Simple Tandem Repeats) состоят из участков ДНК длиной в 2 – 6 п.о. (пар оснований) – tandemно повторенных много раз (например, CACACACACACACA). Они распространены по всему эукариотическому геному. Микросателлиты имеют относительно малые размеры и могут, следовательно, легко амплифицироваться при использовании ПЦР на ДНК, экстрагируемой из различных источников, например, кровь, корни волос, кожа или даже фекалии. Полиморфизмы могут быть визуализированы на секвенирующем геле, и при наличии автоматического ДНК-секвенатора можно анализировать большое количество образцов (Goldstein, Schlötterer, 1999; Jame, Lagoda, 1996). Микросателлиты гипервариабельны; они часто имеют десятки аллелей по одному локусу, отличающихся один от другого по числу повторов. Их все еще предпочитают использовать для изучения разнообразия, а также для анализа отцовства и картирования локусов количественных признаков (QTL), хотя в ближайшем будущем от них могут отказаться в результате развития методов микроматриц (или чипов)

для анализа SNP. ФАО опубликовала разработанные ISAG–FAO Консультативной группой по генетическому разнообразию животных рекомендации по набору микросателлитных локусов для изучения изменчивости главных сельскохозяйственных видов, которые были (см. DAD-IS библиотеку <http://www.fao.org/dad-is/>).

Минисателлиты обладают теми же характеристиками, что и микросателлиты, но длина повторов составляет от десяти до нескольких сотен пар оснований (п.о.). Микро и минисателлиты известны также как VNTR- полиморфизмы (Варирующее количество tandemных повторов - Variable Number of Tandem Repeats).

Полиморфизм длин амплифицируемых фрагментов (Amplified fragment length polymorphisms - AFLP) является методом ДНК-финггерпринта (отпечатки пальцев), который выявляет фрагменты рестрикции ДНК способом их амплификации в ПЦР.

STS маркер (Меченный сайт последовательности - Sequence Tagged Site) является ДНК последовательностью, которая встречается в геноме только раз в известном месте. Они необязательно бывают полиморфными и используются для построения физических карт.

SNP являются вариантами по одному нуклеотиду, которые не меняют общую длину последовательности ДНК в этом регионе. SNP встречаются по всему геному. Они широко распространены и встречаются в геноме человека с частотой один SNP на каждую 1000 пар оснований (Sachinandam и др., 2001). Большинство SNP локализуется в некодирующих областях и не имеет прямого влияния на фенотип индивидуума. Однако некоторые введенные мутации в экспрессирующиеся последовательности или области, влияющие на экспрессию генов (промоторы, энхансеры), могут вызывать изменения в структуре белка или регуляции. Такие SNP предоставляют определенные возможности для выявления функциональных генетических вариантов.

разных маркеров. Для изучения разнообразия наиболее часто используются микросателлиты.

Микросателлиты

В настоящее время микросателлиты (вставка 74) являются наиболее популярными маркерами в ис-

следованиях генетических характеристик домашних животных (Sunnucks, 2001). Их высокая скорость мутирования и кодоминантный характер наследования позволяют оценивать внутри и межпородное генетическое разнообразие и генетическое смешивание пород, даже если они близко родственны.

РАЗДЕЛ 4

Вставка 75
Формирование выборки
генетического материала

Составление выборки является первым и наиболее важным этапом при любом изучении изменчивости. В идеале выборки должны быть репрезентативными и включать неродственных животных. Как правило, выборка объема 30 - 50 хорошо подобранных индивидуумов на породу считается достаточной для обеспечения предварительной оценки различий между породами и внутривидового разнообразия, если оценивается достаточное количество независимых маркеров (например, 20–30 микросателлитов; Nei, Roychoudhury, 1974; Nei, 1978). Однако фактический объем выборки может варьировать в разных случаях и даже может быть меньше в случае высоко инбредных локальных популяций. В случае широко распространенной популяции, подразделенной на различные экотипы, объем выборки должен быть больше.

Выбор неродственных образцов достаточно прост в случае однозначного определения породы, когда он основан на племенной книге или записях родословных. Напротив, в полудиких популяциях при отсутствии письменной регистрации такой выбор может оказаться много труднее. В таких случаях рекомендуется использование географического критерия, т.е. выбирать одно-несколько (неродственных) животных на стадо в разных стадах, распространенных в широкой географической области. Запись географических координат, фотодокументация мест сбора, животных и стад очень ценны – для проверки кроссбридинга в случае неожиданных выбросов, или для выявления интересной географической картины генетического разнообразия. Тщательно подобранный набор образцов – ценный ресурс, который может долгое время использоваться для получения значимых результатов даже при плохой технологии. Наоборот, смещенная выборка приведет к искаженным или трудно интерпретируемым результатам, даже если использованы самые передовые молекулярные методы.

В анализе микросателлитных данных определенные противоречия связаны с выбором модели их мутирования – неограниченная (неограниченное появление аллелей, случайно отличающихся по длине – количеству повторов) или пошаговая (последовательное изменение количества повторов) модель мутиро-

вания (Goldstein и др., 1995). Однако имитационные исследования показали, что неограниченная модель мутирования, в общем, соответствует оценкам внутривидовой изменчивости (Takezaki, Nei, 1996).

Среднее число аллелей в расчете на один локус в популяции, наблюдаемая и ожидаемая гетерозиготность (H_o и H_e) являются наиболее общими параметрами для оценки внутривидовой изменчивости. Наиболее простым параметром для оценки расхождения между популяциями является генетическая дифференциация, или индекс фиксации. Предложено много вариантов оценок (напр., F_{ST} и G_{ST}), причем наиболее широко используется F_{ST} (Weir, Basten, 1990), которая оценивает степень генетической дифференциации субпопуляций на основании расчета стандартизированных вариантов частот аллелей между популяциями. Для значений F_{ST} между парами популяций может быть оценена статистическая достоверность (Weir, Cockerham, 1984) нулевой гипотезы об отсутствии генетической дифференциации между популяциями и, следовательно, различий между генетическими структурами популяций (напр., Mbulu и др., 2003). Может быть выполнен иерархический анализ молекулярной дисперсии (пакет программ AMOVA) (Excoffier и др., 1992) для оценки распределения разнообразия внутри и между группами пород.

Микросателлитные данные также широко используются для оценки генетических взаимоотношений между популяциями и индивидуумами путем оценки генетических расстояний (например, Beja-Pereira и др., 2003; Ibeagha-Awemu и др., 2004; Joshi и др., 2004; Sodhi и др., 2005; Tapio и др., 2005). Наиболее широко используемая мера генетического расстояния – стандартное генетическое расстояние D_S (Nei, 1972). Однако для близко родственных популяций, в которых основным фактором генетической дифференциации является генетический дрейф, что часто происходит в случае пород домашнего скота, особенно в развивающихся странах, рекомендуется использование модифицированных расстояний Кавалли-Сфорца (DA) (Nei и др., 1983). Генетические взаимоотношения между породами могут быть выявлены через реконструкцию их филогении, причем наиболее часто используется метод ближайших соседей (neighbour-joining - N-J) (Saitou, Nei, 1987). Однако главный недостаток реконструкции филогенетическо-

го древа заключается в предположении, что эволюция его ветвей не может образовывать сеть, то есть, ветви могут расходиться, но не могут появляться за счет пересечения. Это предположение редко оказывается справедливым для домашнего скота, где новая порода часто возникает в результате скрещиваний между двумя или более предковыми породами. Таким образом, результаты визуализации эволюции пород, полученные путем реконструкции филогенетических деревьев, следует воспринимать с осторожностью.

Для анализа смешения микросателлитных данных из разных популяций предлагаются методы многомерного анализа или недавно появившиеся методы кластеризации на основе подходов Байеса (Pritchard и др., 2000). Вероятно, примером самого всестороннего исследования такого типа у домашнего скота является изучение крупного рогатого скота на всем африканском континенте (Hanotte и др., 2002), которое выявило генетические следы происхождения, вторичных передвижений и дифференциации пастбищного крупного рогатого скота Африки.

Молекулярно-генетические данные, связанные с другими источниками и дополненные, такими как археологические свидетельства и письменные записи, дают полезную информацию о происхождении, дальнейших перемещениях и развитии генетического разнообразия у домашних видов. Картирование происхождения современного генетического разнообразия потенциально позволяет делать выводы о том, где может быть найдена функциональная генетическая изменчивость вида, для которого существует только ограниченное количество данных о фенотипической изменчивости.

Объединенный анализ микросателлитных данных, полученных в независимых исследованиях, крайне желателен, но редко возможен. Прежде всего, потому, что большинство популяционно-генетических исследований с использованием ДНК маркеров ограничивается небольшим количеством пород, часто из одной и той же страны (Vaunang и др., 2004). Часто используются разные группы маркеров, рекомендованные ФАО, а не генотипирование стандартных наборов во всех проектах. Использование различных микросателлитных систем для генотипирования приводит к различиям в оценках числа аллелей одного и того же локуса в

разных исследованиях. Для того чтобы стимулировать использование одинаковых маркеров, ФАО в настоящее время предлагает обновленный, ранжированный список³ микросателлитных локусов для главных видов домашних животных. ФАО рекомендует использовать маркеры в порядке их ранжирования для того, чтобы максимизировать количество маркеров, совместно используемых в независимых исследованиях. Для некоторых видов доступна ДНК от стандартных животных. Например, алиquotная стандартная ДНК овец и коз, использованная в Эконоген-проекте Европейского Союза (ЕС), распространяется и в других крупномасштабных проектах Азии и Африки. Эти образцы могут быть затребованы через Интернет-сайт проекта Эконоген (Econogene Website - <http://www.econogene.eu>).

Имеется только несколько примеров крупномасштабных исследований генетического разнообразия домашних видов. Hillel и др. (2003) и SanCristobal и др. (2006а) исследовали, соответственно, разнообразие кур и свиней в Европе; Hanotte и др. (2002) получили данные по крупному рогатому скоту практически всего африканского континента; Tarjo и др. (2005) оценили разнообразие овец в странах Северной Европы; и Саïon и др. (2006) исследовали разнообразие овец в Европе и на Ближнем и Среднем Востоке. Однако для большинства видов такой всесторонний обзор все еще отсутствует. Продолжающаяся тесная координация между крупномасштабными проектами обещает дать общую оценку генетического разнообразия некоторых видов, таких как овцы и козы, уже в ближайшем будущем. Тем временем развиваются новые методы анализа, позволяющие выполнять мета-анализ наборов данных, которые включают только несколько пород и/или только некоторые общие маркеры. (Freeman и др., 2006). Такая глобальная перспектива оценки разнообразия домашнего скота будет чрезвычайно ценна для воссоздания картины происхождения и истории популяций domesticированных животных и, косвенно, популяций человека. Это также позволит высветить региональные и локальные «горячие точки» генетического разнообразия, на которые могут быть направлены усилия по сохранению.

³ Списки и правила можно найти в библиотеке DAD-IS по адресу <http://www.fao.org/dad-is>

РАЗДЕЛ 4

SNP

SNP (вставка 74) используется в изучении генетического разнообразия как альтернатива микросателлитам. Доступен ряд технологий по выявлению и типированию SNP-маркеров (см. обзор Syvänen, 2001). Будучи диаллельными маркерами, SNP имеют существенно меньшее информационное содержание, и для получения того же уровня информации, какой можно получить при использовании стандартной панели из 30 микросателлитных локусов, необходимо использовать большее их количество. Однако постоянно развивающиеся молекулярные технологии увеличивают автоматизацию и уменьшают стоимость типирования SNP. Похоже, что в ближайшем будущем это позволит выполнять параллельные анализы большого числа маркеров по низкой цене. С такой перспективой выполняются крупномасштабные проекты по ряду видов домашних животных для идентификации миллионов (напр., Wong и др., 2004) и подтверждения нескольких тысяч SNP, для выявления блоков гаплотипов в геноме. Так же как информация о последовательностях, SNP позволяют непосредственно сравнивать и объединять результаты анализа различных экспериментов.

В будущем SNP, по-видимому, будут привлекательными маркерами для изучения генетического разнообразия, поскольку их легко использовать в оценке и функциональной, и нейтральной изменчивости. Однако критической становится предварительная стадия выявления SNP или отбора SNP из базы данных. SNP могут быть выявлены с использованием различных экспериментальных протоколов, таких как секвенирование, одноцепочечный конформационный полиморфизм (SSCP) или денатурирующая высокоэффективная жидкостная хроматография (DHPLC), или *in silico*, путем выравнивания и сравнения множества последовательностей одной и той же области, представленных в публичных базах данных геномных и экспрессирующихся (EST) последовательностей. Если данные получены для случайно сформированных выборок, к ним невозможно применять стандартные оценки популяционно-генетических параметров. Распространенный пример – когда SNPs, исходно идентифицированные в маленькой выборке (панели) индивидуумов, далее типированы на большой выборке хромосом. Такой

протокол, предпочтительно отбирая SNP с промежуточными частотами, сместит оценки распределения аллельных частот по сравнению с ожидаемым для случайной выборки. SNP действительно считается перспективным методом для будущего применения в популяционно-генетическом анализе; однако должны быть разработаны статистические методы, учитывающие особенности каждого метода выявления (Nielsen, Signorovitch, 2003; Clark и др., 2005).

AFLP

AFLP являются доминантными диаллельными маркерами (Vos и др., 1995). Можно оценивать изменчивость одновременно по многим локусам и выявлять единичные нуклеотидные замены в неизвестных участках генома, в которых может присутствовать неизвестный функциональный ген, несущий данную мутацию. Однако неудобством метода является то, что они наследуются доминантно (невозможно отличить гомозиготу по доминантному аллелю от гетерозиготы); это уменьшает возможности их использования в исследованиях генетического разнообразия внутри породы и при инбридинге. Тем не менее, профили AFLP высоко информативны при оценке взаимосвязей между породами (Ajmone-Marsan и др., 2002; Negri и др., 2006; De Marchi и др., 2006; SanCristobal и др., 2006b) и близкими видами (Buntjer и др., 2002).

Митохондриальные ДНК-маркеры

Полиморфизм митохондриальной ДНК (мтДНК) широко используется в филогенетических исследованиях и при изучении генетического разнообразия. Гаплоидная мтДНК, которую несут митохондрии в цитоплазме клеток, имеет материнский характер наследования (индивидуумы наследуют мтДНК от своих матерей, но не от отцов) и высокую скорость мутирования; мтДНК не рекомбинирует. Эти характеристики позволяют биологам реконструировать эволюционные взаимосвязи между видами и внутри них на основании оценок распределения мутаций в мтДНК. МтДНК-маркеры могут также обеспечивать быстрый способ выявления гибридизации между видами или подвидами домашних животных (напр., Nijman и др., 2003).

Полиморфизм в последовательности гипервариабельного района D-петли, или контролирующей области мтДНК вносит очень большой вклад в иден-

тификацию диких предковых видов domesticiрованных видов животных, установление географического распределения генетического разнообразия и в понимание процессов domestикации животных (см. обзор Bruford и др., 2003). Например, происхождение современного европейского крупного рогатого скота из Среднего Востока (средиземноморский центр происхождения) было продемонстрировано недавно в работе Troy и др. (2001). У *Bos taurus* обнаружено четыре материнских линии, также показана утрата бычьими генетического разнообразия в процессе миграции человека из средиземноморского Плодородного полумесяца (Fertile Crescent) во времена неолита. Таким же способом были выявлены происхождение от множества матерей и наличие трех мтДНК-линий у коз (Luikart и др., 2001), с вероятными центрами их происхождения в Азии и Средиземноморье. Недавно была обнаружена третья линия мтДНК у аборигенных китайских овец (Guo и др., 2005), четвертая – у аборигенных китайских коз (Chen и др., 2005) и пятая – у китайского крупного рогатого скота (Lai и др., 2006). У азиатских кур найдено девять различных мтДНК ветвей (Liu и др., 2006), что позволяет предполагать их полифилетическое происхождение в Южной и Юго-Восточной Азии. Все эти данные свидетельствуют о том, что наши современные представления о domestикации домашних животных и их генетическом разнообразии далеки от завершения. Для дальнейшего обсуждения происхождения domesticiрованных видов домашних животных см. раздел 1, часть А.

3.2 Использование маркеров для оценки эффективной численности популяций

Hill (1981) предложил использовать гаметическое неравновесие по ДНК-полиморфизмам для оценки эффективной численности популяций (N_e). Оценка основывается на генотипировании сцепленных маркеров (микросателлитов или SNP). Ожидаемая корреляция между частотами аллелей сцепленных локусов является функцией N_e и частот рекомбинаций между ними. Следовательно, N_e может быть оценена по наблюдаемому неравновесию. Hauges и др. (2003) предложил аналогичный подход, основанный на гомозиготности сегментов хромосом, кроме того, этим методом потенциально возможно оценивать

N_e для более ранних поколений и, следовательно, можно судить, увеличивался или уменьшался раз-

Вставка 76 Картирование QTL

Если QTL для признака-мишени существуют, плюс- и минус- аллельные варианты неизвестного, отвечающего за признак гена (Q и q), будут сегрегировать совместно с аллелями ближайшего M1 маркера (M1 и m1), которые мы можем генотипировать в лаборатории. Предположим, что M1 ко-сегрегирует с Q и m1 с q, это означает, что M1 и Q расположены рядом в одной и той же хромосоме, а m1 и q в гомологичной хромосоме (M1Q и m1q).

Предположим также, что популяция F2, полученная от скрещивания гетерозиготных индивидов F1, генотипирована. В результате генотипирования потомки F2 группируются в соответствии с их маркерными генотипами (M1M1 и m1m1; M2M2 и m2m2; ... MmMn и mnmn), и далее сравниваются средние фенотипы этих групп.

Если отсутствуют QTL, сцепленные с данным маркером (например, с M2), тогда отсутствуют значимые различия между средними фенотипическими значениями исследуемого признака у потомков с генотипами M2M2 и m2m2. Противоположная ситуация будет наблюдаться тогда, когда у потомков, сгруппированных по генотипу по маркеру M1, окажется, что группа M1M1 большей частью несет вариант QQ по QTL, а группа m1m1 представлена главным образом qq. В этом случае наблюдаются значимые различия средних значений между группами потомков и, следовательно, определяется присутствие QTL. У видов, таких как куры и свиньи, у которых обычно линии и породы скрещиваются в коммерческих целях, такая процедура может быть выполнена в экспериментальных популяциях (F2, BC), тогда как у жвачных обычно используется анализ в двух (определение по дочерям - daughter design – DD) или трех (определение по внукам - grand-daughter design – GDD) поколениях потомков. В DD сегрегация гетерозиготных маркеров производителя (поколение I) прослеживается у дочерей (поколение II), фенотип которых оценивается. В GDD, сегрегация гетерозиготных у деда-производителя маркеров (поколение I) прослеживается у его сыновей-полусибсов (поколение II), а оценка связей маркеров с фенотипическими характеристиками выполняется у их дочерей-внучек (поколение III).

РАЗДЕЛ 4

мер существующей популяция в прошлом. Исследования показали, на примере ряда данных, что голштино-фризская порода крупного рогатого скота в прошлом подверглась существенному сокращению N_e , в то время как эффективная численность популяций человека возрастает, что согласуется с данными переписей и исследованиями родословных.

3.3 Молекулярные инструменты для выявления функциональной изменчивости

Подходы, основанные на картировании положения: картирование локусов количественных признаков (QTL)

Генетические маркеры ведут себя как менделирующие признаки; другими словами, они подчиняются законам сегрегации и независимого наследования, впервые описанным Менделем. Два гена, локализованные в одной хромосоме, физически сцеплены и имеют тенденцию наследоваться вместе. При прохождении мейоза рекомбинации между гомологичными хромосомами могут разрушать это сцепление. Частота рекомбинаций между двумя генами, локализованными в одной и той же хромосоме, зависит от расстояния между ними. Частота рекомбинаций между маркерами, следовательно, является показателем степени их сцепления: чем ниже частота рекомбинации, тем ближе маркеры. Создание генетических карт использует это свойство для установления наиболее вероятного порядка маркеров и расстояний между ними.

В общем, на практике картирование достигается после оценки совместной сегрегации аллелей полиморфных маркеров в структурированных экспериментальных популяциях (например, F2 или обратное скрещивание) или существующих популяциях в селекционных программах (семьи полных сибсов или полусибсов). Для большинства видов домашних животных имеются генетические карты с высокой плотностью распределения маркеров, от нескольких сотен до нескольких тысяч.

Для идентификации QTL для данного признака, семья с сегрегацией признака генотипируется по нескольким картированным молекулярным маркерам, относительно равномерно распределенным по

геному (вставка 76). Существует ряд статистических методов, позволяющих устанавливать присутствие существенных QTL в данном маркерном интервале, но все усложняется тем фактом, что семьи обладают высоким уровнем неравновесия по сцеплению, то есть, большими сегментами хромосом, которые передаются без рекомбинаций от родителей к потомкам.

В результате экспериментов по картированию QTL идентифицируются участки хромосом, часто распространяющиеся на половину хромосомы, по которым выявляются значимые эффекты на проявление исследуемого признака. В современных исследованиях активно используются методы картирования для выявления QTL, влияющих на признаки, связанные с адаптацией. Примерами таких признаков являются, повышенная резистентность к колонизации и экскреции сальмонеллы (Tilquin и др., 2005) и чувствительность к развитию синдрома легочной гипертензии (Rabie и др., 2005) у кур; и толерантность к трипаносоме у крупного рогатого скота (Hanotte и др., 2002).

За стадией картирования QTL обычно следует уточнение положения QTL на карте (QTL тонкое картирование). Для достижения этой цели анализируются дополнительные маркеры и все, представленные выше, дополнительные рекомбинационные события между ними в исследуемой области. Недавно был разработан и использован удачный подход для тонкого картирования области хромосомы BTA14, несущей существенный QTL для процента жира в молоке и других признаков (Farnir и др., 2002). В этом подходе использовалась история рекомбинаций в предыдущих поколениях для ограничения положения на карте небольшим участком в 3.8 cM (сантимограна); такой размер участка позволяет проводить позиционное клонирование гена (DGAT1) (Grisart и др., 2002).

Вслед за тонким картированием среди генов, локализованных в выделенном районе, могут быть выявлены гены, определяющие проявление признака. Гены-кандидаты могут быть найдены у того же вида (например, когда имеется достаточно полная карта экспрессирующихся последовательностей - EST карта, или когда геном полностью секвенирован) или в ортологических участках модельных организмов, для которых имеется полная геномная информация.

Иногда ключевая информация о функции гена приходит из неожиданных источников. Так было с

Вставка 77

Популяционная геномика

Недавно был предложен альтернативный подход для выявления геномных областей, несущих соответствующие гены. Он состоит в обнаружении «следов отбора» через подход «популяционной геномики» (Black и др., 2001; Luikart и др., 2003). Подход популяционной геномики для картирования QTL основан на трех главных принципах:

- 1) на нейтральные локусы по всему геному будут действовать сходным образом генетический дрейф, демография и эволюционная история популяций;
- 2) локусы, находящиеся под действием отбора, будут отличаться по своему поведению и, следовательно, характер их изменчивости будет резко отличаться: будут обнаруживаться утрата изменчивости (рост изменчивости, если локусы находятся под давлением балансирующего отбора), неравновесие по сцеплению, и увеличение/уменьшение показателей Gst/Fst ;
- 3) через эффекты «путешествия на попутках» отбор будет влиять на сцепленные маркеры, что позволит обнаруживать «следы отбора» (выбросы из общей картины), которые могут быть выявлены путем генотипирования большого количества маркеров вдоль хромосомы и идентификации кластеров, изменчивость которых «выпадает» из общей картины. Такой подход использует фенотипические данные на уровне породы (или субпопуляции внутри породы), а не на индивидуальном уровне, таким образом, хорошо дополняя классический подход по картированию QTL в пределах родословных.

Подход популяционной геномики позволяет также идентифицировать гены, подвергающиеся сильному давлению отбора и в конечном счете фиксирующиеся

в породе, и в частности, гены, участвующие в адаптации к экстремальным условиям окружающей среды, устойчивости к болезням и т.д. Большинство таких признаков, существенных для устойчивого разведения животных, трудно, или даже невозможно исследовать с использованием классических методов картирования QTL или связанных с ними подходов. Потенциал популяционной геномики недавно был исследован с теоретической точки зрения (Beaumont, Balding, 2004; Bamshad, Wooding, 2003) и в экспериментальных исследованиях с различными типами маркеров в природных популяциях (AFLP маркеры: Campbell, Bernatchez, 2004; микросателлиты: Kayser и др., 2003; SNP: Akey и др., 2002). В настоящее время этот подход используется в проекте Эконоген (<http://lasig.epfl.ch/projets/econogene>). В результате предварительного анализа получены данные о том, что для поведения трех SNP в генах MYH1 (миозин 1), MEG3 (каллипин) и CTSD (катепсин В) характерно значимое отклонение (Pariset и др., 2006).

В том же самом проекте был разработан новый подход на основании метода пространственного анализа (Spatial Analysis Method - SAM) для выявления следов естественного отбора в геномах доместивированных и диких видов животных (Joost, 2006). Предварительные результаты, полученные с использованием этого метода, соответствуют полученным путем теоретического моделирования в популяционной генетике, например, разработанного Beaumont и Balding (2004). SAM делает следующий шаг по сравнению с классическими подходами, так как он разработан для выявления параметров среды, ассоциированных с селектированными маркерами.

геном миостатина, функция которого была впервые установлена у мышей, а потом оказалось, что он локализуется у крупного рогатого скота в том районе хромосомы, где ранее был картирован ген синдрома двойной мускулатуры (McPherron, Lee, 1997).

Ясно, что идентификация генов, отвечающих за фенотипические характеристики (гены количественных признаков – QTG), и функциональных мутаций (QTN) комплексных признаков все еще остается

очень важной задачей, и необходимо разработать ряд подходов для того, чтобы уменьшить количество позиционированных на карте генов-кандидатов. В этом отношении определяющей является информация о функции гена. Однако, нам все еще не известна возможная функция(ии) большинства генов, идентифицированных в результате секвенирования геномов и кДНК (комплементарная ДНК). Именно поэтому исследование профилей генов экспрес-

РАЗДЕЛ 4

сии в сочетании с описанным выше позиционным подходом может обеспечить появление полезной информации для идентификации генов-кандидатов, контролирующих комплексные признаки. Такой комбинированный подход определяется как генетическая геномика (Haley, de Koning, 2006). Новые достижения в исследованиях профилей геной экспрессии описаны в следующем разделе.

В настоящее время исследуются альтернативные подходы для выявления адаптивных генов с использованием генетических маркеров (вставка 77). Сейчас они находятся на экспериментальной стадии, и только дальнейшие исследования позволят оценить их плодотворность.

Конечная цель картирования QTL состоит в идентификации QTG, и, в конечном счете, QTN. Хотя до настоящего времени у домашнего скота известно мало примеров, существуют мутации, которые могут оказать непосредственное воздействие на маркерную селекцию и на принятие решения о сохранении. По мере увеличения числа обнаруженных QTG и QTN в ближайшем будущем необходима разработка специальных моделей сохранения, учитывающих функциональные признаки и мутации.

Исследование характера геной экспрессии

В прошлом проявление специфических признаков, таких как адаптация и устойчивость, можно было оценить только на фенотипическом уровне. В настоящее время транскриптом (совокупность всех транскриптов в клетке или ткани) и протеом (совокупность всех белков) могут быть прямо исследованы с использованием высокоточных технологий, таких как дифференциальное проявление (differential display – DD) (Liang, Pardee, 1992), кДНК-AFLP (Bachem и др., 1996), серийный анализ геной экспрессии (SAGE) (Velculescu и др., 1995; 2000), масс-спектрометрия, белковые и ДНК-микроматрицы. Эти методы обусловили прорыв в анализе РНК и белков, позволяя параллельно анализировать фактически все экспрессирующиеся в данное время гены в ткани. Таким образом, эти методы вносят свой вклад в расшифровку геновых сетей, лежащих в основе большинства комплексных признаков.

–Омик технологии часто сравнивают с включением света перед фреской Микеланджело вместо исполь-

зования для ее освещения факела, который позволяет видеть только часть целого. Полный вид позволяет понять представленное и оценить его красоту. В реальности, возможности этих методик в настоящее время соответствуют трудностям и стоимости их применения и анализа получаемых данных. Очень трудным является выделение гомогенной клеточной популяции, что является важным исходным требованием для большинства исследований профилей геной экспрессии. Большое количество параллельных анализов приводит к снижению стоимости одного анализа, но к высокой суммарной стоимости эксперимента. На всех этапах экспериментальных исследований требуется дорогое оборудование и высокий технический профессионализм. К этому прибавляются общие трудности работы с РНК, по сравнению с ДНК. РНК очень чувствительна к разрушению, этому приходится уделять особенно много внимания при экстрагировании из тканей с высокой метаболической активностью. Консервация образцов и манипуляции с ними, несомненно, являются ключевыми условиями успеха экспериментов по анализу РНК. Применение нанотехнологий для анализа биологических молекул открывает многообещающие перспективы в решении этих проблем (Sauer и др., 2005).

Следующая проблема – обработка данных. Молекулярная база данных, такая, как профили геной экспрессии, могут создаваться в относительно короткое время. Однако стандартизация данных, полученных в разных лабораториях, требует согласованного анализа различных наборов биологических данных. Существенным для эффективного анализа молекулярных сетей являются именно соглашения по стандартизации, также как и по созданию взаимосвязанных баз данных.

Профили транскрипции

В этом разделе представлено короткое описание методов SAGE и микроматриц. Описания других методов можно найти в ряде современных обзоров (напр., Donson и др., 2002). SAGE создает полный профиль экспрессии ткани или клеточной линии. Метод включает создание полной библиотеки мРНК, позволяющей выполнять количественный анализ целых транскриптов, экспрессирующихся или инактивированных на определенных стадиях клеточной

активности. Метод основан на трех принципах: (i) необходимо иметь короткие мишени – последовательности (9-14 пар оснований), полученные из определенного района внутри каждого мРНК-транскрипта, содержащего достаточно информации для уникальной идентификации одного специфического транскрипта; (ii) последовательности-мишени должны быть сцеплены вместе для формирования длинных ДНК-молекул (конкатемеры), которые могут быть клонированы и секвенированы – секвенирование клонов конкатемеров приводит к быстрой идентификации множества индивидуальных мишеней; (iii) уровень экспрессии оценивается по количеству мишеней в сумме транскриптов.

Микроматрицы используются для сравнения в одном эксперименте уровней экспрессии мРНК нескольких тысяч генов в двух биологических системах, например, у животных в нормальной среде и у животных в экстремальной среде. Технология микроматриц обеспечивает также понимание временного и пространственного порядков экспрессии генов в ответ на действие широкого спектра факторов, которому подвергается организм.

Очень маленькие объемы раствора ДНК печатаются на подложках, сделанных из непористых материалов, например, стекло, формируя небольшие пятна (споты), диаметром от 100 до 150 мк. В настоящее время с использованием робототехники на предметном стекле для микроскопа можно распечатать около 50 000 комплементарных ДНК (кДНК). ДНК-микроматрицы содержат несколько тысяч известных и несколько тысяч неизвестных генов. На микроматрице могут быть распечатаны фрагменты кДНК или заранее подготовленные олигонуклеотиды. В последнем случае достигается высокий уровень специфичности и воспроизводимости, однако их разработка возможна только тогда, когда известна последовательность таких олигонуклеотидов. Использование микроматрицы основано на принципе «гибридизации», то есть экспонировании двух одноцепочечных последовательностей ДНК, или одной ДНК и одной РНК, с последующим измерением количества образовавшихся двуцепочечных молекул. Экспрессия мРНК может быть измерена качественно и количественно. Наличие мРНК указывает на активность гена в

ткани и обычно прямо связано с наработкой белка, транслируемого с этой мРНК.

Профили генной экспрессии вносят вклад в понимание биологических механизмов и, следовательно, облегчение идентификации генов-кандидатов. Например, пул генов, участвующих в проявлении трипанотолерантности у крупного рогатого скота, был охарактеризован с использованием метода SAGE (Berthier и др., 2003), и анализа кДНК-микроматриц (Hill и др., 2005). Параллельное исследование многих генов может позволить идентифицировать гены-«господа», отвечающие за фенотипические характеристики, неопределяемые в анализе дифференциальной экспрессии генов. Такие гены-«господа», например, могут быть представлены разными аллелями, которые все экспрессируются на одном и том же уровне, однако, с разной эффективностью обеспечивают экспрессию подчиненных генов. В этом случае ген-«господин» можно найти путем использования знаний о метаболических путях или через подход по проявлению QTL (expression QTL-eQTL) (Lan и др., 2006). При таком подходе в сергирующей популяции измеряется уровень экспрессии подчиненных генов. Количество транскрипта каждого гена обрабатывается как фенотипическая характеристика, и QTL, который влияет на генную экспрессию, можно обнаружить с использованием методологии, описанной выше. Следует отметить, что анализ данных для обнаружения QTL является все еще весьма трудным для исполнения. Это верно и для методики профилирования транскриптов из-за возникновения большого количества ложных сигналов.

Профилирование белков

Систематическое изучение структуры белков, посттрансляционных модификаций, белковых профилей, взаимодействий белок-белок, белок-нуклеиновая кислота, белок-малые молекулы, и пространственной и временной экспрессии белков в эукариотических клетках важно для понимания комплексных биологических явлений. Белки необходимы для структуры живых клеток и их функций.

Структура белка может быть выявлена дифракцией рентгеновских лучей или ядерно-магниторезонансной спектроскопией. Первое требует большого количества кристаллического белка, что является существенным ограничением. Для того, чтобы по-

РАЗДЕЛ 4

нять функцию белка и взаимодействия белок-белок на молекулярном уровне, было бы полезно определить структуру всех белков в клетке или организме. К настоящему времени, однако, этого еще не достигли. Интересно, что количество вариантов различных белков, появляющихся в процессе белкового синтеза (в частности, в результате альтернативного сплайсинга и/или пост-трансляционной модификации), существенно больше, чем количество генов в геноме.

Масс-спектрометрия (аналитический метод для определения молекулярных масс) в комбинации с хроматографией или техникой электрофоретического разделения является в настоящее время предпочтительным методом для идентификации эндогенных белков в клетках, характеристики пост-трансляционных модификаций и определения относительного содержания белков (Zhu и др., 2003). Двумерный гель-электрофорез уникален в отношении большого числа белков (>10 000), которые можно разделить и визуализировать в одном эксперименте. Белковые пятна вырезаются из геля, затем подвергаются протеолитическому перевариванию, и затем белки идентифицируются с использованием масс-спектрометрии (Aebersold, Mann, 2003). Однако стандартизация и автоматизация двумерного гель-электрофореза достигается с трудом, и использование получающегося белкового рисунка, отражающего карту протеомы, оказывается успешным только в некоторых случаях. Дополняющая методика, жидкостная хроматография, автоматизируется легко и может быть прямо объединена с масс-спектрометрией. Основанные на аффинности методы протеомики, которые основываются на микроматрицах, являются альтернативным подходом к белковому профилированию (Lueking и др., 2003), они могут использоваться и для выявления взаимодействий белок-белок. Такая информация существенна для алгоритмического моделирования биологических путей. Однако специфичность связывания остается проблемой в применении белковых микроматриц, поскольку невозможно точно предсказать перекрестную реактивность. Существуют альтернативные подходы для выявления взаимодействий белок-белок, такие, как система двух гибридов (Fields, Song, 1989). Однако ни один из современных методов не позволяет количественно определить связывающиеся белки, и остается неясным, в какой

Вставка 78

Базы данных биологических молекул

Существует ряд баз данных, в которых собирается информация по биологическим молекулам:

Базы данных по секвенированной ДНК:

- Европейская лаборатория молекулярной биологии (EMBL):
<http://www.ebi.ac.uk/embl/index.html>
- ГенБанк: <http://www.ncbi.nlm.nih.gov/>
- Банк ДНК-данных Японии (DDBJ):
<http://www.ddbj.nig.ac.jp>

Базы данных белков:

- Швейцарский-ПРОТ:
<http://www.expasy.ch/sprot/sprot-top.html>
- Информационные ресурсы белков (PIR):
<http://pir.georgetown.edu/pirwww/>
- Банк данных белков (PDB):
<http://www.rcsb.org/pdb/>

Сервис по идентификации генов, сайты**Био-портала:**

- Геномные странички: <http://www.hgmp.mrc.ac.uk/GenomeWeb/nuc-geneid.html>
- ВСУ поисковое устройство:
<http://searchlauncher.bcm.tmc.edu/>
- МОЛБИОЛ: <http://www.molbiol.net/>
- Инструменты биомолекулярных исследований Педро: http://www.biophys.uni-duesseldorf.de/BioNet/Pedro/research_tools.html
- ExPASy сервер Молекулярной биологии:
<http://www.expasy.ch/>

Базы данных, представляющие интерес в связи с domestцированными животными:

- <http://locus.jouy.inra.fr/cgi-bin/bovmap/intro.pl>
- <http://www.cgd.csiro.au/cgd.html>
- <http://www.ri.bbsrc.ac.uk/cgi-bin/arkdb/browsers/>
- <http://www.marc.usda.gov/genome/genome.html>
- <http://www.ncbi.nlm.nih.gov/genome/guide/pig/>
- <http://www.ensembl.org/index.html>
- <http://www.tigr.org/>
- <http://omia.angis.org.au/>
- <http://www.livestockgenomics.csiro.au/ibiss/>
- <http://www.thearkdb.org/>
- <http://www.hgsc.bcm.tmc.edu/projects/bovine/>

степени наблюдаемые взаимодействия представляют физиологические взаимодействия белок-белок.

Методы, основанные на матрицах, разрабатываются также для выявления ДНК-белок взаимодействий *in vitro* и *in vivo* (см. обзор Sauer и др., 2005), и идентификации неизвестных белков, связывающихся с последовательностями, участвующими в регуляции гена. ДНК микроматрицы эффективно применяются для выявления ДНК-связывающих комплексов в ядерных экстрактах, тогда как белковые микроматрицы, главным образом, используются для идентификации неизвестных ДНК-связывающих белков на уровне всего протеома. В будущем эти два метода будут способствовать выявлению деталей регуляторных сетей транскрипции.

Многие методы предсказания функции белка основываются на их гомологии с другими белками и их локализации внутри клетки. Предсказание функций белка достаточно сложны, и также требуют методов выявления взаимодействия белок-белок и выявления связывания белков с другими молекулами, поскольку именно в этих процессах связывания белки реализуют свою функцию.

4 Роль биоинформатики

Без возможностей анализа экспоненциально растущего количества биологических данных, развитие высокопроизводительных технологий было бы бесполезным. Эти технологии требуют методов хранения в электронных базах данных (вставка 78), связанных с разработкой специального программного обеспечения, позволяющего обновлять, запрашивать и искать данные. Информация должна быть легкодоступной и гибкой по отношению к запросам, чтобы можно было осуществлять поиск информации, которая может быть проанализирована для объяснения метаболических путей и роли участвующих в них белков и генов.

Для объединения информации из разных источников и получения новых знаний из существующих данных биоинформатика является определяющей областью исследований. Она также обладает мощным потенциалом для моделирования структур, функций и динамики молекулярных систем, и, следовательно, служит для формулирования гипотез и развития экспериментальной работы.

5 Заключение

Молекулярная характеристика может играть важную роль в раскрытии истории, оценке разнообразия, самобытности и популяционной структуры ГРЖ. Она также может помочь избежать избыточного инбридинга при генетическом управлении маленькими популяциями. Многие исследования описывают внутри- и межпопуляционное разнообразие – некоторые в весьма крупном масштабе. Однако эти исследования фрагментарны, их трудно сравнивать и обобщать. Более того, не проведены всесторонние международные обследования соответствующих видов. По этой причине стратегическое значение имеет разработка методов объединения существующих, частично перекрывающихся наборов данных, и обеспечение стандартизации образцов и маркеров для будущего использования в качестве стандартов для исследований во всех странах. Сеть лабораторий, собирающих образцы автохтонной зародышевой плазмы, которая станет доступной для научного сообщества на определенных условиях, будет способствовать выполнению глобального обследования генетического разнообразия.

Маркерные технологии эволюционируют и, похоже, что микросателлиты последовательно замещаются на SNP. Эти маркеры очень перспективны, поскольку число их в геноме велико, и они пригодны для автоматизации анализа и генотипирования. Однако эффективность SNP в изучении разнообразия у видов животных до сих пор остается недостаточно исследованной. К этому вопросу необходимо относиться достаточно критично, для того, чтобы избежать накопления искаженных данных.

Методы анализа данных также эволюционируют. Новые методы позволяют изучать разнообразие, не прибегая к предположениям *a priori* о структуре исследуемой популяции; использовать разнообразие для выявления адаптивных генов (например, используя популяционную геномику, см. вставка 77); обобщать информацию, полученную из различных источников, включая социально-экономические и экологические параметры, для расстановки приоритетов по сохранению (см. часть E). Принятие правильной стратегии

РАЗДЕЛ 4

формирования выборок и систематический сбор фенотипических и экологических данных, остаются ключевыми требованиями для использования полного потенциала новых технологий и подходов.

Кроме изучения нейтральной изменчивости ведется активный поиск генов, влияющих на ключевые признаки. В первую очередь изучаются такие признаки как устойчивость к заболеваниям, продуктивность, качество конечной продукции. В этих целях используется ряд стратегий и новых высокоэффективных —омик технологий. Идентификация QTN открывает новые возможности и ставит новые задачи в управлении ГРЖ. Информация об адаптивном разнообразии дополняет фенотипическое и нейтральное генетическое разнообразие, и может быть использована для управления ГРЖ и создания инструментов для решения вопросов по их сохранению. Идентификация в определенных популяциях уникальных аллелей или комбинаций аллелей по адаптивным признакам может усилить обоснование их сохранения и направленного использования. Селекция с помощью генов потенциально может уменьшить разрыв в эффектив-

ности отбора, обычно существующий между большими популяциями, разводящимися в индустриальных системах производства, и небольшими локальными популяциями, где не могут быть применены системы популяционной генетической оценки и схемы селекции. Селекция с помощью маркеров и генов, однако, не всегда может представлять наилучшее решение. Эти подходы необходимо оценивать и оптимизировать на основе последовательного анализа каждого случая, принимая во внимание краткосрочные и долгосрочные воздействия на популяционную структуру и степень инбридинга, стоимость и выгоды, выраженные в экологических и социально-экономических параметрах – в особенности по влиянию на экономическое положение людей.

Как в случае других успешных технологий, очень желательно, чтобы преимущества научных достижений в области молекулярного описания стали глобальными, внося свой вклад в улучшение понимания, использования и сохранения мировых ГРЖ для пользы настоящих и будущих поколений человека.

Вставка 79

Словарь: молекулярные маркеры

В связи с задачами этого подраздела используются следующие определения:

Ген-кандидат: любой ген, который реально может вызывать отличия в наблюдаемых характеристиках животных (например, в устойчивости к болезням, продукции молочных белков или росте). Такой ген может быть кандидатом, поскольку локализован в определенном хромосомном районе, который предположительно участвует в контроле признака, или считается, что его белковый продукт может прямо принимать участие в формировании признака (например, гены белков молока в продукции белков молока).

ДНК: генетическая информация в геноме кодируется дезоксирибонуклеиновой кислотой (ДНК), которая хранится в клеточном ядре. ДНК состоит из двух цепочек, организованных в двойную спираль,

каждая цепочка состоит из нуклеотидов, включающих сахар (дезоксирибозу), фосфат, и одно из четырех химических оснований: аденин (А), гуанин (G), цитозин (С) и тимин (Т). А на одной цепочке всегда соединяется с Т на другой цепочке двумя водородными связями, тогда как С всегда соединяется с G тремя водородными связями. Таким образом, две цепочки комплементарны друг другу.

Комплементарная ДНК (кДНК): последовательность ДНК, синтезируемая путем обратной транскрипции последовательности мРНК. Этот тип кДНК включает экзоны и нетранслируемые последовательности 5' и 3' концов гена, но никогда не включает ДНК-интроны.

- продолжение следует

Вставка 79 (продолжение)

Словарь: молекулярные маркеры

Генетический маркер: полиморфизм ДНК, который может быть легко установлен при молекулярном или фенотипическом анализе. Маркер может находиться в гене или в ДНК с неизвестной функцией. Поскольку участки ДНК, лежащие близко друг к другу, имеют тенденцию наследоваться вместе, маркеры часто используются как не прямой путь для отслеживания особенностей наследования гена, который до сих пор не идентифицирован или чья приблизительная локализация известна.

Гаплотип: сокращение словосочетания «гаплоидный генотип», генетическое строение индивидуальной хромосомы. В случае диплоидных организмов, гаплотип будет содержать одного члена пары аллелей по каждому сайту. Это может относиться к ряду маркеров (например, однонуклеотидный полиморфизм - SNP), для которых обнаружено, что они все статистически ассоциированы с одной хромосомой. Из этого следует, что идентификация нескольких аллелей в блоках гаплотипов может однозначно идентифицировать все другие полиморфные участки в этом районе. Такая информация очень полезна для изучения генетики не только комплексных признаков.

Сцепление: Ассоциация генов и/или маркеров, лежащих рядом друг с другом на хромосоме. Сцепленные гены и маркеры имеют тенденцию наследоваться вместе.

Неравновесие по сцеплению (LD): термин, используемый в исследованиях по популяционной генетике для неслучайных ассоциаций аллелей двух или более локусов, необязательно одной и той же хромосомы. Это не то же самое, что сцепление, которое описывается как связь двух или более локусов на хромосоме с ограниченной рекомбинацией между

ними. LD описывает ситуацию, в которой некоторая комбинация аллелей или генетических маркеров возникает с частотой большей или меньшей в популяции, чем можно было бы ожидать, исходя из случайного образования гаплотипов аллелей, основываясь на их частотах. Неравновесие по сцеплению вызывается приспособительным взаимодействием между генами или такими неадаптивными процессами, как популяционная структура, инбридинг, или стохастические эффекты. В популяционной генетике неравновесие по сцеплению характеризует распределение гаплотипов по двум или более локусам.

Технология микроматриц: новый способ изучения того, как большие количества генов взаимодействуют друг с другом и как клеточные регулирующие сети одновременно управляют обширными батареями генов. Метод использует робототехнику для точного распределения крошечных капелек, содержащих функциональную ДНК, на стеклянных подложках. Затем исследователи прикрепляют флуоресцентные метки к мРНК или кДНК из клеток, которые они изучают. Меченые пробы связываются с цепочками кДНК на подложках. Подложки размещают в сканирующий микроскоп, в котором измеряется яркость каждого флуоресцирующего пятна; яркость показывает, сколько присутствует специфической мРНК, что указывает на активность соответствующего гена.

Праймер: короткая (одноцепочечная) олигонуклеотидная последовательность, используемая в полимеразной цепной реакции (ПЦР)

РНК: рибонуклеиновая кислота является одноцепочечной нуклеиновой кислотой, содержащей четыре основания, три из них присутствуют и в ДНК (А, С и G). Т в РНК заменен на урацил (U).

РАЗДЕЛ 4

Источники

- Aebersold, R. & Mann, M.** 2003. Mass spectrometry-based proteomics. *Nature*, 422 (6928): 198–207. Review.
- Ajmone-Marsan, P., Negrini, R., Milanesi, E., Bozzi, R., Nijman, I.J., Buntjer, J.B., Valentini, A. & Lenstra, J.A.** 2002. Genetic distances within and across cattle breeds as indicated by biallelic AFLP markers. *Animal Genetics*, 33: 280–286.
- Akey, J.M., Zhang, G., Zhang, K., Jin, L. & Shriver, M.D.** 2002. Interrogating a high-density SNP map for signatures of natural selection. *Genome Research*, 12(12): 1805–14.
- Aravin, A. & Tuschl, T.** 2005. Identification and characterization of small RNAs involved in RNA silencing. *Febs Letters*, 579(26): 5830–40.
- Bachem, C.W.B., Van der Hoeven, R.S., De Bruijn, S.M., Vreugdenhil, D., Zabeau, M. & Visser, R.G.F.** 1996. Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analyses of gene expression during potato tuber development. *The Plant Journal*, 9: 745–753.
- Bamshad, M. & Wooding, S.P.** 2003. Signatures of natural selection in the human genome. *Nature Reviews Genetics*, 4(2): 99–111. Review.
- Baumung, R., Simianer, H. & Hoffmann, I.** 2004. Genetic diversity studies in farm animals – a survey, *Journal of Animal Breeding and Genetics*, 121: 361–373.
- Beaumont, M.A. & Balding, D.J.** 2004. Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, 13(4): 969–80.
- Beja-Pereira, A., Alexandrino, P., Bessa, I., Carretero, Y., Dunner, S., Ferrand, N., Jordana, J., Laloe, D., Moazami-Goudarzi, K., Sanchez, A. & Cañon, J.** 2003. Genetic characterization of southwestern European bovine breeds: a historical and biogeographical reassessment with a set of 16 microsatellites. *Journal of Heredity*, 94: 243–50.
- Berthier, D., Quere, R., Thevenon, S., Belemsaga, D., Piquemal, D., Marti, J. & Maillard, J.C.** 2003. Serial analysis of gene expression (SAGE) in bovine trypanotolerance: preliminary results. *Genetics Selection Evolution*, 35 (Suppl. 1): S35–47.
- Bertone, P., Stolc, V., Royce, T.E., Rozowsky, J.S., Urban, A.E., Zhu, X., Rinn, J.L., Tongprasit, W., Samanta, M., Weissman, S., Gerstein, M. & Snyder, M.** 2004. Global identification of human transcribed sequences with genome tiling arrays. *Science*, 306: 2242–2246.
- Black, W.C., Baer, C.F., Antolin, M.F. & DuTeau, N.M.** 2001. Population genomics: genome-wide sampling of insect populations. *Annual Review of Entomology*, 46: 441–469.
- Bruford, M.W., Bradley, D.G. & Luikart, G.** 2003. DNA markers reveal the complexity of livestock domestication. *Nature Reviews Genetics*, 4: 900–910.
- Buntjer, J.B., Otsen, M., Nijman, I.J., Kuiper, M.T. & Lenstra, J.A.** 2002. Phylogeny of bovine species based on AFLP fingerprinting. *Heredity*, 88: 46–51.
- Campbell, D. & Bernatchez, L.** 2004. Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Molecular Biology and Evolution*, 21(5): 945–56.
- Cañon, J., Garcia, D., Garcia-Atance, M.A., Obexer-Ruff, G., Lenstra, J.A., Ajmone-Marsan, P., Dunner, S. & The ECONOGENE Consortium.** 2006. Geographical partitioning of goat diversity in Europe and the Middle East. *Animal Genetics*, 37: 327–334.
- Chen, S.Y., Su, Y.H., Wu, S.F., Sha, T. & Zhang, Y.P.** 2005. Mitochondrial diversity and phylogeographic structure of Chinese domestic goats. *Molecular Phylogenetics and Evolution*, 37: 804–814.
- Clark, A.G., Hubisz, M.J., Bustamante, C.D., Williamson, S.H. & Nielsen, R.** 2005. Ascertainment bias in studies of human genome-wide polymorphism. *Genome Research*, 15: 1496–1502.

- Clop, A., Marcq, F., Takeda, H., Pirottin, D., Tordoix, X., Bibe, B., Bouix, J., Caiment, F., Elsen, J.M., Eychenne, F., Larzul, C., Laville, E., Meish, F., Milenkovic, D., Tobin, J., Charlier, C. & Georges, M.** 2006. A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nature Genetics*, 38: 813–818.
- De Marchi, M., Dalvit, C., Targhetta, C. & Cassandro, M.** 2006. Assessing genetic diversity in indigenous Veneto chicken breeds using AFLP markers. *Animal Genetics*, 37: 101–105.
- Donson, J., Fang, Y., Espiritu-Santo, G., Xing, W., Salazar, A., Miyamoto, S., Armendarez, V. & Volkmuth, W.** 2002. Comprehensive gene expression analysis by transcript profiling. *Plant Molecular Biology*, 48: 75–97.
- Excoffier, L., Smouse, P.E. & Quattro, J.M.** 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, 131: 479–491.
- Farnir, F., Grisart, B., Coppieters, W., Riquet, J., Berzi, P., Cambisano, N., Karim, L., Mni, M., Moio, S., Simon, P., Wagenaar, D., Vilkkki, J. & Georges, M.** 2002. Simultaneous mining of linkage and linkage disequilibrium to fine map quantitative trait loci in outbred half-sib pedigrees: revisiting the location of a quantitative trait locus with major effect on milk production on bovine chromosome 14. *Genetics*, 161: 275–287.
- Fields, S. & Song, O.** 1989. A novel genetic system to detect protein–protein interactions. *Nature*, 340: 245–246.
- Freeman, A.R., Bradley, D.G., Nagda, S., Gibson, J.P. & Hanotte, O.** 2006. Combination of multiple microsatellite data sets to investigate genetic diversity and admixture of domestic cattle. *Animal Genetics*, 37: 1–9.
- Goldstein, D.B., Linares, A.R., Cavalli-Sforza, L.L. & Feldman, M.W.** 1995. An evaluation of genetic distances for use with microsatellite loci. *Genetics*, 139: 463–471.
- Goldstein, D.B. & Schlötterer, C.** 1999. *Microsatellites: evolution and applications*. New York, USA. Oxford University Press.
- Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P., Spelman, R., Georges, M. & Snell, R.** 2002. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. *Genome Research*, 12: 222–231.
- Guo, J., Du, L.X., Ma, Y.H., Guan, W.J., Li, H.B., Zhao, Q.J., Li, X. & Rao, S.Q.** 2005. A novel maternal lineage revealed in sheep (*Ovis aries*). *Animal Genetics*, 36: 331–336.
- Haley, C. & de Koning, D.J.** 2006. Genetical genomics in livestock: potentials and pitfalls. *Animal Genetics*, 37(Suppl 1): 10–12.
- Hanotte, O., Bradley, D.G., Ochieng, J.W., Verjee, Y. & Hill, E.W.** 2002. African pastoralism: genetic imprints of origins and migrations. *Science*, 296: 336–339.
- Hayes, B.J., Visscher, P.M., McPartlan, H.C. & Goddard, M.E.** 2003. A novel multilocus measure of linkage disequilibrium to estimate past effective population size. *Genome Research*, 13: 635–643.
- Hill, E.W., O’Gorman, G.M., Agaba, M., Gibson, J.P., Hanotte, O., Kemp, S.J., Naessens, J., Coussens, P.M. & MacHugh, D.E.** 2005. Understanding bovine trypanosomiasis and trypanotolerance: the promise of functional genomics. *Veterinary Immunology and Immunopathology*, 105: 247–258.
- Hill, W.G.** 1981. Estimation of effective population size from data on linkage disequilibrium. *Genetics Research*, 38: 209–216.
- Hillel, J., Groenen, M.A., Tixier-Boichard, M., Korol, A.B., David, L., Kirzhner, V.M., Burke, T., Barre-Dirie, A., Crooijmans, R.P., Elo, K., Feldman, M.W., Freidlin, P.J., Maki-Tanila, A., Oortwijn, M., Thomson, P., Vignal, A., Wimmers, K. & Weigend, S.** 2003. Biodiversity of 52 chicken populations assessed by microsatellite typing of DNA pools. *Genetics Selection Evolution*, 35: 533–557.

РАЗДЕЛ 4

- Hood, L., Heath, J.R., Phelps, M.E. & Lin, B.** 2004. Systems biology and new technologies enable predictive and preventative medicine. *Science*, 306: 640–643.
- Ibeagha-Awemu, E.M., Jann, O.C., Weimann, C. & Erhardt, G.** 2004. Genetic diversity, introgression and relationships among West/Central African cattle breeds. *Genetics Selection Evolution*, 36: 673–690.
- Jarne, P. & Lagoda, P.J.L.** 1996. Microsatellites, from molecules to populations and back. *Tree*, 11: 424–429.
- Joshi, M.B., Rout, P.K., Mandal, A.K., Tyler-Smith, C., Singh, L. & Thangaraj, K.** 2004. Phylogeography and origin of Indian domestic goats. *Molecular Biology and Evolution*, 21: 454–462.
- Joost, S.** 2006. *The geographical dimension of genetic diversity*. A GIScience contribution for the conservation of animal genetic resources. École Polytechnique Fédérale de Lausanne, Switzerland. (PhD thesis)
- Kayser, M., Brauer, S. & Stoneking, M.** 2003. A genome scan to detect candidate regions influenced by local natural selection in human populations. *Molecular Biology and Evolution*, 20: 893–900.
- Lai, S.J., Liu, Y.P., Liu, Y.X., Li, X.W. & Yao, Y.G.** 2006. Genetic diversity and origin of Chinese cattle revealed by mtDNA D-loop sequence variation. *Molecular Phylogenetics and Evolution*, 38: 146–54.
- Lan, L., Chen, M., Flowers, J.B., Yandell, B.S., Stapleton, D.S., Mata, C.M., Ton-Keen Mui, E., Flowers, M.T., Schueler, K.L., Manly, K.F., Williams, R.W., Kendzioriski, C. & Attie, A.D.** 2006. Combined expression trait correlations and expression quantitative trait locus mapping. *PLoS Genetics*, 2: 51–61.
- Liang, P. & Pardee, A.B.** 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science*, 257: 967–997.
- Liu, Y.P., Wu, G.S., Yao, Y.G., Miao, Y.W., Luikart, G., Baig, M., Beja-Pereira, A., Ding, Z.L., Palanichamy, M.G. & Zhang, Y.P.** 2006. Multiple maternal origins of chickens: out of the Asian jungles. *Molecular Phylogenetics and Evolution*, 38: 12–19.
- Luiking, A., Possling, A., Huber, O., Beveridge, A., Horn, M., Eickhoff, H., Schuchardt, J., Lehrach, H. & Cahill, D.J.** 2003. A nonredundant human protein chip for antibody screening and serum profiling. *Molecular and Cellular Proteomics*, 2: 1342–1349.
- Luikart, G., England, P.R., Tallmon, D., Jordan, S. & Taberlet, P.** 2003. The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, 4: 981–994.
- Luikart, G., Gielly, L., Excoffier, L., Vigne, J.D., Bouvet, J. & Taberlet, P.** 2001. Multiple maternal origins and weak phylogeographic structure in domestic goats. *Proceedings of the National Academy of Science USA*, 98: 5927–5932.
- Mburu, D.N., Ochieng, J.W., Kuria, S.G., Jianlin, H. & Kaufmann, B.** 2003. Genetic diversity and relationships of indigenous Kenyan camel (*Camelus dromedarius*) populations: implications for their classification. *Animal Genetics*, 34(1): 26–32.
- McPherron, A.C. & Lee, S.J.** 1997. Double muscling in cattle due to mutations in the myostatin gene. *Proceedings of the National Academy of Science USA*, 94: 12457–12461.
- Negrini, R., Milanese, E., Bozzi, R., Pellecchia, M. & Ajmone-Marsan, P.** 2006. Tuscany autochthonous cattle breeds: an original genetic resource investigated by AFLP markers. *Journal of Animal Breeding and Genetics*, 123: 10–16.
- Nei, M.** 1972. Genetic distance between populations. *The American Naturalist*, 106: 283–292.
- Nei, M.** 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89: 583–590.
- Nei, M. & Roychoudhury, A.K.** 1974. Sampling variances of heterozygosity and genetic distance. *Genetics*, 76: 379–390.
- Nei, M., Tajima, F. & Tateno, Y.** 1983. Accuracy of estimated phylogenetic trees from molecular

- data. II. Gene frequency data. *Journal of Molecular Evolution*, 19: 153–170.
- Nielsen, R. & Signorovitch, J.** 2003. Correcting for ascertainment biases when analyzing SNP data: applications to the estimation of linkage disequilibrium. *Theoretical Population Biology*, 63: 245–55.
- Nijman, I.J., Otsen, M., Verkaar, E.L., de Ruijter, C. & Hanekamp, E.** 2003. Hybridization of banteng (*Bos javanicus*) and zebu (*Bos indicus*) revealed by mitochondrial DNA, satellite DNA, AFLP and microsatellites. *Heredity*, 90: 10–16.
- Pariset, L., Cappuccio, I., Joost, S., D'Andrea, M.S., Marletta, D., Ajmone Marsan, P., Valentini A. & ECONOGENE Consortium** 2006. Characterization of single nucleotide polymorphisms in sheep and their variation as an evidence of selection. *Animal Genetics*, 37: 290–292.
- Pritchard, J.K., Stephens, M. & Donnelly, P.** 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155: 945–959.
- Rabie, T.S., Crooijmans, R.P., Bovenhuis, H., Vereijken, A.L., Veenendaal, T., van der Poel, J.J., Van Arendonk, J.A., Pakdel, A. & Groenen, M.A.** 2005. Genetic mapping of quantitative trait loci affecting susceptibility in chicken to develop pulmonary hypertension syndrome. *Animal Genetics*, 36: 468–476.
- Saitou, N. & Nei, M.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4: 406–425.
- Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G., Sherry, S., Mullikin, J.C., Mortimore, B.J., Willey, D.L., Hunt, S.E., Cole, C.G., Coggill, P.C., Rice, C.M., Ning, Z., Rogers, J., Bentley, D.R., Kwok, P.Y., Mardis, E.R., Yeh, R.T., Schultz, B., Cook, L., Davenport, R., Dante, M., Fulton, L., Hillier, L., Waterston, R.H., McPherson, J.D., Gilman, B., Schaffner, S., Van Etten, W.J., Reich, D., Higgins, J., Daly, M.J., Blumenstiel, B., Baldwin, J., Stange-Thomann, N., Zody, M.C., Linton, L., Lander, E.S. & Altshuler, D.; International SNP Map Working Group.** 2001. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*, 409: 928–933.
- SanCristobal, M., Chevalet, C., Haley, C.S., Joosten, R., Rattink, A.P., Harlizius, B., Groenen, M.A., Amigues, Y., Boscher, M.Y., Russell, G., Law, A., Davoli, R., Russo, V., Desautels, C., Alderson, L., Fimland, E., Bagga, M., Delgado, J.V., Vega-Pla, J.L., Martinez, A.M., Ramos, M., Glodek, P., Meyer, J.N., Gandini, G.C., Matassino, D., Plastow, G.S., Siggens, K.W., Laval, G., Archibald, A.L., Milan, D., Hammond, K. & Cardellino, R.** 2006a. Genetic diversity within and between European pig breeds using microsatellite markers. *Animal Genetics*, 37: 189–198.
- SanCristobal, M., Chevalet, C., Peleman, J., Heuven, H., Brugmans, B., van Schriek, M., Joosten, R., Rattink, A.P., Harlizius, B., Groenen, M.A., Amigues, Y., Boscher, M.Y., Russell, G., Law, A., Davoli, R., Russo, V., Desautels, C., Alderson, L., Fimland, E., Bagga, M., Delgado, J.V., Vega-Pla, J.L., Martinez, A.M., Ramos, M., Glodek, P., Meyer, J.N., Gandini, G., Matassino, D., Siggens, K., Laval, G., Archibald, A., Milan, D., Hammond, K., Cardellino, R., Haley, C. & Plastow, G.** 2006b. Genetic diversity in European pigs utilizing amplified fragment length polymorphism markers. *Animal Genetics*, 37: 232–238.
- Sauer, S., Lange, B.M.H., Gobom, J., Nyarsik, L., Seitz, H. & Lehrach, H.** 2005. Miniaturization in functional genomics and proteomics. *Nature Reviews Genetics*, 6: 465–476.
- Sodhi, M., Mukesh, M., Mishra, B.P., Mitkari, K.R., Prakash, B. & Ahlawat, S.P.** 2005. Evaluation of genetic differentiation in *Bos indicus* cattle breeds from Marathwada region of India using microsatellite polymorphism. *Animal Biotechnology*, 16: 127–137.
- Storz, G., Altuvia, S. & Wassarman, K.M.** 2005. An abundance of RNA regulators. *Annual Review of Biochemistry*, 74: 199–217.
- Sunnucks, P.** 2001. Efficient genetic markers for population biology. *Tree*, 15: 199–203.
- Syvänen, A.C.** 2001. Accessing genetic variation genotyping single nucleotide polymorphisms. *Nature Reviews Genetics*, 2: 930–941.

РАЗДЕЛ 4

- Takezaki, N. & Nei, M.** 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*, 144: 389–399.
- Tapio, M., Tapio, I., Grislis, Z., Holm, L.E., Jeppsson, S., Kantanen, J., Miceikiene, I., Olsaker, I., Viinalass, H. & Eythorsdottir, E.** 2005. Native breeds demonstrate high contributions to the molecular variation in northern European sheep. *Molecular Ecology*, 14: 3951–3963.
- Tilquin, P., Barrow, P.A., Marly, J., Pitel, F., Plisson-Petit, F., Velge, P., Vignal, A., Baret, P.V., Bumstead, N. & Beaumont, C.** 2005. A genome scan for quantitative trait loci affecting the Salmonella carrier-state in the chicken. *Genetics Selection Evolution*, 37: 539–61.
- Troy, C.S., MacHugh, D., Bailey, J.F., Magee, D.A., Loftus, R.T., Cunningham, P., Chamberlain, A.T., Sykes, B.C. & Bradley D.G.** 2001. Genetic evidence for Near-Eastern origins of European cattle. *Nature*, 410: 1088–1091.
- Velculescu, V.E., Vogelstein, B. & Kinzler, K.W.** 2000. Analyzing uncharted transcriptomes with SAGE. *Trends in Genetics*, 16: 423–425.
- Velculescu, V.E., Zhang, L., Vogelstein, B. & Kinzler, K.W.** 1995. Serial analysis of gene expression. *Science*, 270: 484–487.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J. & Kuiper, M.** 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23: 4407–4444.
- Weir, B.S. & Basten, C.J.** 1990. Sampling strategies for distances between DNA sequences. *Biometrics*, 46: 551–582.
- Weir, B.S. & Cockerham, C.C.** 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, 38: 1358–1370.
- Wienholds, E. & Plasterk, R.H.** 2005. MicroRNA function in animal development. *FEBS Letters*, 579: 5911–5922.
- Wong, G.K., Liu, B., Wang, J., Zhang, Y., Yang, X., Zhang, Z., Meng, Q., Zhou, J., Li, D., Zhang, J., Ni, P., Li, S., Ran, L., Li, H., Zhang, J., Li, R., Li, S., Zheng, H., Lin, W., Li, G., Wang, X., Zhao, W., Li, J., Ye, C., Dai, M., Ruan, J., Zhou, Y., Li, Y., He, X., Zhang, Y., Wang, J., Huang, X., Tong, W., Chen, J., Ye, J., Chen, C., Wei, N., Li, G., Dong, L., Lan, F., Sun, Y., Zhang, Z., Yang, Z., Yu, Y., Huang, Y., He, D., Xi, Y., Wei, D., Qi, Q., Li, W., Shi, J., Wang, M., Xie, F., Wang, J., Zhang, X., Wang, P., Zhao, Y., Li, N., Yang, N., Dong, W., Hu, S., Zeng, C., Zheng, W., Hao, B., Hillier, L.W., Yang, S.P., Warren, W.C., Wilson, R.K., Brandstrom, M., Ellegren, H., Crooijmans, R.P., van der Poel, J.J., Bovenhuis, H., Groenen, M.A., Ovcharenko, I., Gordon, L., Stubbs, L., Lucas, S., Glavina, T., Aerts, A., Kaiser, P., Rothwell, L., Young, J.R., Rogers, S., Walker, B.A., van Hateren, A., Kaufman, J., Bumstead, N., Lamont, S.J., Zhou, H., Hocking, P.M., Morrice, D., de Koning, D.J., Law, A., Bartley, N., Burt, D.W., Hunt, H., Cheng, H.H., Gunnarsson, U., Wahlberg, P., Andersson, L., Kindlund, E., Tammi, M.T., Andersson, B., Webber, C., Ponting, C.P., Overton, I.M., Boardman, P.E., Tang, H., Hubbard, S.J., Wilson, S.A., Yu, J., Wang, J., Yang, H.; International Chicken Polymorphism Map Consortium.** 2004. A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature*, 432: 717–722.
- Zhu, H., Bilgin, M. & Snyder, M.** 2003. Proteomics. *Annual Review of Biochemistry*, 72: 783–812.