



ИРКУТСКОМУ ГОСУДАРСТВЕННОМУ
УНИВЕРСИТЕТУ
100 ЛЕТ

АКТУАЛЬНЫЕ ПРОБЛЕМЫ СОВРЕМЕННОЙ ГЕНЕТИКИ

ГЕНЕТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА БИОРАЗНООБРАЗИЯ

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Иркутский государственный университет»
Биолого-почвенный факультет
Лимнологический институт СО РАН
Министерство продовольствия и сельского хозяйства ЛНР
ГОУ ЛНР Луганский национальный аграрный университет

АКТУАЛЬНЫЕ ПРОБЛЕМЫ СОВРЕМЕННОЙ ГЕНЕТИКИ

ГЕНЕТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА БИОРАЗНООБРАЗИЯ

Учебное пособие

Составители:

Д. Ю. Щербаков, В. Е. Харченко



УДК 574/575(075.8)
ББК 28.04я73
А43

Печатается по решению учебно-методической комиссии
биолого-почвенного факультета ИГУ

Рецензенты:

канд. биол. наук *Ю. С. Букин*,
канд. биол. наук *Г. В. Юринова*

А43 **Актуальные** проблемы современной генетики: генетические методы анализа биоразнообразия : учеб. пособие / сост. Д. Ю. Щербаков, В. Е. Харченко. – Иркутск : Изд-во ИГУ, 2018. – 123 с.

ISBN 978-5-9624-1588-8

Учебное пособие содержит начальные сведения об актуальных проблемах современной генетики и генетических методах анализа биоразнообразия, а также задания для практических занятий, направленные на закрепление материала, и рекомендации, которые можно использовать в исследовательской работе.

Предназначено для студентов, аспирантов и специалистов в области биологии, экологии и сельского хозяйства, деятельность которых связана с вопросами анализа и регуляции биоразнообразия.

УДК 574/575(075.8)
ББК 28.04я73

ISBN 978-5-9624-1588-8

© Щербаков Д. Ю., Харченко В. Е., сост., 2018
© ФГБОУ ВО «ИГУ», 2018

Оглавление

| | |
|--|----|
| Введение | 5 |
| 1. Теоретические основы биоразнообразия | 6 |
| 1.1. Понятие вида и форм внутривидовой изменчивости | 6 |
| 1.1.1. Основные концепции вида | 6 |
| 1.1.2. Формы внутривидовой изменчивости | 9 |
| 1.1.3. Центральные и периферические популяции | 11 |
| 1.2. Механизмы, предотвращающие скрещивания разных видов.... | 12 |
| 1.3. Мутационная изменчивость как источник биоразнообразия ... | 13 |
| 1.4. Ненаследуемая (модификационная) изменчивость | 17 |
| 1.5. Распределение значений признака..... | 18 |
| 1.6. Действие отбора на распределение значений признаков..... | 21 |
| 1.7. Географическая изменчивость видов и их ареалы | 23 |
| 1.8. Экология видообразования..... | 28 |
| 1.8.1. <i>K/R</i> -стратегии адаптации | 28 |
| 1.9. Генетическая устойчивость в популяциях | 29 |
| 2. Основные молекулярно-генетические методы анализа биоразнообразия | 37 |
| 2.1. Полимеразная цепная реакция (PCR) | 38 |
| 2.2. Экстракция ДНК..... | 40 |
| 2.3. Штрихкод жизни | 42 |
| 2.4. Генетические маркеры, используемые для исследования биоразнообразия | 48 |
| 2.5. Метагеномный анализ..... | 51 |
| 2.5.1. Shotgun-метагеномный анализ..... | 51 |
| 2.6. Микросателлитные повторы ДНК и биоразнообразие | 53 |
| 2.7. Гипотеза молекулярных часов | 57 |
| 2.8. Анализ демографических процессов и миграции..... | 64 |

Практикум

| | |
|--|----|
| <i>Занятие 1. Анализ биоразнообразия на основании последовательностей ДНК</i> | 69 |
| <i>Занятие 2. Мутации как элементарный эволюционный материал, регулирующий биоразнообразие</i> | 73 |

| | |
|--|-----|
| <i>Занятие 3.</i> Анализ гаплотипического и нуклеотидного разнообразия..... | 77 |
| <i>Занятие 4.</i> Анализ результатов филогенетических построений и молекулярные часы | 82 |
| <i>Занятие 5.</i> Представления результатов анализа последовательностей ДНК | 88 |
| <i>Занятие 6.</i> Использование молекулярно-генетических методов для анализа географической изменчивости | 91 |
| <i>Занятие 7.</i> Анализ биоразнообразия при помощи построения эволюционных сценариев..... | 94 |
| <i>Занятие 8.</i> Метагеномный анализ..... | 97 |
| <i>Занятие 9.</i> Анализ биоразнообразия при помощи микросателлитных повторов ДНК | 104 |
| | |
| Заключение | 111 |
| Задания для контроля умений и навыков студентов | 112 |
| Словарь терминов | 113 |
| Рекомендуемая литература | 115 |
| Рекомендуемые веб-ресурсы | 116 |
| Приложения | 117 |
| Использованная литература | 122 |

ВВЕДЕНИЕ

Современные подходы к анализу проблем биоразнообразия и, соответственно, круг вопросов, которые стали доступны для изучения, резко изменились в последние десятилетия и в связи с кардинальными переменами в методическом потенциале, и вследствие углубления знаний в этой области. Так было, когда в популяционной генетике, а потом и в других биологических дисциплинах стали активно использоваться методы математической статистики, а потом и биохимические методы. Это привело к появлению и расцвету синтетической теории эволюции. Вторжение молекулярно-генетических подходов привело к дополнению наших взглядов на эволюционный процесс и позволило использовать нейтральные мутации для определения возраста событий и описания демографической истории популяций. В настоящее время геномные и постгеномные технологии создают базу для дальнейшего прогресса нашего понимания законов развития и функционирования живого мира.

Учебное пособие «Актуальные проблемы современной генетики: генетические методы анализа биоразнообразия» посвящено изучению изменчивости и разнообразия живых организмов, процессам микро- и макроэволюции живых организмов, а также образованию экотопов и развитию экосистем. Понимание причинно-следственных связей между возникновением и изменениями степени и характера внутривидовой изменчивости, видообразованием и экосистемами относится к фундаментальным знаниям, которые необходимы для приобретения профессиональных компетенций, необходимых студентам биологических и аграрных направлений и специальностей.

1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ БИОРАЗНООБРАЗИЯ

В описательной зоологии и ботанике простое накопление наблюдений может иметь ценность, но этого никогда не бывает в физике. Это обстоятельство составляет для учёного, искушённого в теоретической науке, одну из характерных черт описательной зоологии или ботаники; для него это просто «охота за мухами», т. е. вопрос коллекционирования, а не проницательности.

S. Toulmin [1953]

Вид, конечно, понятие не столько ботаническое или зоологическое, сколько генетическое.

П. М. Жуковский

1.1. ПОНЯТИЕ ВИДА И ФОРМ ВНУТРИВИДОВОЙ ИЗМЕНЧИВОСТИ

1.1.1. Основные концепции вида

При анализе биоразнообразия мы оперируем разными таксономическими единицами, характеризующими положение интересующего нас организма на эволюционном «дереве жизни»: царство, отдел, класс, отряд, семейство, тип, род, вид и др. Однако в природе реально существует только вид.

«Вид» является наиболее дискуссионным понятием в биологии, однако без него разговор о биоразнообразии будет беспредметным. На разных этапах развития естествознания представления о виде менялись.

По мнению Ж. Б. Ламарка (1751), как в ботанике, так и в зоологии вид должен состоять из совокупности сходных между собой особей, которые остаются такими же при размножении, при этом он акцентировал внимание на существовании «случайных различий», половых различий и разновидностей в пределах одного вида.

Согласно **типологической (монотипической) концепции**, вид представляет собой группу особей, морфологически сходных с эталоном-образцом (базиотипом, представителями типовой серии, происходящими из типового местообитания), который хранится в общедоступной коллекции, а идентификация вида проводится на основании диагностических признаков, чаще всего морфологических. Признак является **диагностическим**, если имеет специфическое состояние, свойственное только для особей данного вида, но отсутствующее у всех особей других видов.

Изменчивость, возникающая под влиянием условий среды или мутаций, может привести к преобразованию фенотипа. Поэтому морфологические критерии имеют условную ценность для диагностики таксона.

При анализе биоразнообразия применение этой концепции осложняется тем, что она исключает формы модификационной изменчивости и не позволяет дифференцировать виды-близнецы.

Монотипические виды, по мнению Н. И. Вавилова [1928], «существуют только до того времени, как они изучаются в гербарии, но при детальном ботанико-географическом изучении они распадаются на множество форм». *Например*: виды *Zea mays*, *Triticum durum*, *Pisum sativum* и другие имеют сотни тысяч форм внутривидовой изменчивости.

Биологическая концепция вида была сформулирована сторонниками синтетической теории эволюции. Она предполагает, что виды состоят из популяций, обладающих внутренней генетической интегрированностью, вследствие того, что все особи вида имеют общую генетическую программу, исторически сложившуюся в ходе эволюции [Тимофеев-Ресовский, 2009].

Виды представляют собой группы популяций, в состав которых входят особи, которые имеют потенциальную возможность скрещиваться и давать плодовитое потомство, но при этом они репродуктивно изолированы от других таких же групп [Майр, 1974]. Таким образом, морфологическое сходство особей в пределах вида как бы уходит на второй план.

При анализе биоразнообразия политипическая концепция вида позволяет лучше, чем типологическая, представить систему образования и преобразования видов. Так как вид, как и любой

объект, можно детализировать до бесконечности, обращая внимание на морфологические, физиологические, этологические и прочие подробности его существования, которые могут сделать задачу анализа биоразнообразия невыполнимой.

В ходе истории формирования вид гармонично приспосабливается к условиям своего существования, а наличие репродуктивных барьеров позволяет это сохранить. Межвидовая гибридизация происходит крайне редко, но в тех редких случаях, когда она имеет место, обычно возникает дисбаланс, негативно отражающийся на их жизнеспособности.

Каждый вид можно рассматривать как биологический эксперимент. Попадая в новые условия, он может оказаться как в эволюционном тупике, так и на старте новых широких возможностей адаптации.

Один и тот же генотип может порождать разные фенотипы в разных условиях среды. Экстремальные обстоятельства могут вызвать реализацию возможностей, которые не проявляются в условиях, близких к нормальным. Внутри- и межпопуляционная изменчивость не согласуется с типологической концепцией вида.

В центре ареала популяции, как правило, имеют более высокую плотность и большую индивидуальную изменчивость, чем на периферии, так как они населены обычно менее плотно, а индивидуальная изменчивость там меньше. Причина состоит в том, что лишь ограниченное число генотипов способно существовать вблизи границы ареала вида, на пределе возможностей существования, так как видовые и подвидовые признаки – адаптивны.

Обычно описание вида приводится упрощённо, так как учитывать все нюансы его модификационной изменчивости не имеет смысла. К. Линней (1753) в работе «Философия в ботанике» писал, что «нищета не знает законов», подразумевая, что исследователь, наблюдая объект только при каких-то определённых условиях, может выработать стандарт его восприятия и начать его излишне детализировать. Пока не будет принята концепция политипического вида, каждая отличающаяся популяция может быть возведена в ранг настоящего вида. Утверждение, что политипические виды отсутствуют, свидетельствует о недостаточности выборки из периферических или периферически изолированных по-

пуляций или о том, что авторы склонны называть видом отличный географический изолят. В таких случаях особенно полезным может оказаться использование молекулярных признаков – нуклеотидных последовательностей маркерных генов.

Виды политипичны и состоят из популяций, распределённых как в пространстве, так и во времени, им свойственны как сходство, так и изменчивость.

1.1.2. Формы внутривидовой изменчивости

Виды представляют собой реальные единицы эволюции, в которых временно воплощаются гармоничные, хорошо интегрированные генные комплексы. Видообразование представляет собой метод, с помощью которого эволюция движется вперёд посредством продуцирования новых генных комплексов, обеспечивающих адаптацию к новым экологическим условиям и само биоразнообразие.

Формы (морфобиологические группы) представляют собой группы особей одного вида, отличающихся друг от друга по своим морфологическим, анатомическим признакам, биологическим или физиологическим и биохимическим особенностям, продолжительности онтогенеза и времени смены фаз развития и пр.

Формы внутривидовой изменчивости могут быть **ареальными** и **безареальными**.

Ареальные формы внутривидовой изменчивости: подвид, популяция, экотип, клон.

Подвид (географическая раса, экотип) – совокупность фенотипически сходных популяций некоторого вида, населяющих часть ареала этого вида и существенно отличающихся от других популяций этого вида. *Например:* в пределах вида *Betula nana* L. выделяют подвиды *Betula nana* subsp. *exilis* (Sukaczew) Hultén и *Betula nana* subsp. *nana* (<http://www.efloras.org/> FNA Vol. 3 Betulaceae). *Betula nana* subsp. *exilis* распространён на северо-востоке Азии и северо-западе Северной Америки, на Аляске, Юконе до зал. Гудзон. *Betula nana* subsp. *nana* распространён на северо-востоке Северной Америки и Гренландии. То есть ареалы их распространения удалены, и между ними находится преграда – залив Гудзон, который препятствует их скрещиванию между собой.

Внутри подвида можно выделить климатические разновидности, и чем разнообразнее условия в ареале вида, тем больше можно ожидать у него климатипов и экотипов.

Клоном называется группа особей, происходящих от одной особи в результате бесполого размножения. *Например*, картофель – *Solanum tuberosum* L. можно размножать вегетативно – клубнями, при этом различиями, возникающими в результате соматического мутагенеза, можно пренебречь.

Популяция представляет собой совокупность организмов, свободно скрещивающихся между собой и образующих плодовитое потомство. При этом как диагностические, так и любые другие признаки (морфологические, экологические и пр.) могут быть полиморфными. *Например*, «пшеницы-двуручки» при осеннем посеве ведут себя как озимые, а при весеннем как яровые, т. е. наблюдается полиморфизм по продолжительности периода вегетации.

Разбиение вида на популяции бывает весьма условно. При исследовании природных популяций это разбиение зачастую операционально. Считается, что мы имеем дело с двумя популяциями, если удастся выделить две группы особей, у которых вероятность скрещивания внутри группы достоверно выше, чем если родители принадлежат к разным группам. Если к тому же географическая граница между группами достаточно резкая, то принято говорить о существовании у вида «популяционной структуры».

Безареальными называются следующие формы **внутривидовой изменчивости**: мутантные формы (лузусы, аберрации), гибридные формы. Такие разновидности имеют большое значение для селекционной практики. Резкую границу между лузусами и аберрациями иногда провести трудно. *Например*, ель европейская змеевидная – *Picea abies* var. *Virgata* формирует преимущественно ветви первого порядка, в результате образуется плакучая крона. У ясеня (*Fraxinus*) и акации белой (*Robinia pseudoacacia*) листья обычно сложные, но бывают и простые.

Искусственно созданные формы внутривидовой изменчивости

Сорт представляет собой популяцию организмов, искусственно созданных человеком и обладающих определёнными наследственными особенностями (морфологическими, физиоло-

гическими и пр.), проявляющимися при определённых условиях среды (культивирования, агротехники и пр.).

Экады – формы ненаследуемой (модификационной) изменчивости, приуроченные к определенным местообитаниям и носящие приспособительный характер. *Например*, мелколистность древесных растений связана с их солевыносливостью и заморозкоустойчивостью [Булыгин, 1991].

Со временем между формами внутривидовой изменчивости может развиваться устойчивая репродуктивная изоляция и они могут стать самостоятельными видами. Однако преждевременное возведение их в ранг вида может затруднить представления исследователей о путях видообразования.

Исчезновение форм внутривидовой изменчивости свидетельствует о том, что данный генотип оказался нежизнеспособным к адаптации в данных условиях среды, но для вида в целом это не критично.

При анализе биоразнообразия формы внутривидовой изменчивости иногда возводят в ранг вида или даже относят к другому роду, в результате вопросы состава экосистем и путей их эволюции становятся крайне запутанными.

1.1.3. Центральные и периферические популяции

Каждый вид представляет собой независимую генетическую систему, репродуктивно изолированную от других симпатрических видов и экологически совместимую с ними. Видообразование может происходить постепенно (географическое видообразование) или мгновенно (полиплоидия). Репродуктивная изоляция означает наличие механизмов, которые охраняют гармонично коадаптированный генофонд от разрушительного влияния генотипов из других генофондов. В ходе видообразования формы внутривидовой изменчивости могут приобретать репродуктивную изоляцию.

В центре ареала вида чаще всего наблюдается высокая степень изменчивости, так как условия среды оптимальны для вида и в них может выживать максимальное количество различных форм. Условия обитания периферических и изолированных попу-

ляций обычно близки к критическим условиям для данного вида, в них выживают только те генотипы, которые могут к ним адаптироваться. В таких популяциях численность вида может резко сокращаться в неблагоприятные годы и увеличиваться в благоприятные. В отличие от центральной части видового ареала для них характерны резкое колебание численности и высокая генетическая изменчивость [Майр, 1974].

Высокая степень специализации генотипа к условиям среды может оказаться критичной в случае резких изменений среды, поэтому, при анализе биоразнообразия, важно выявлять факторы, лимитирующие распространение вида.

1.2. МЕХАНИЗМЫ, ПРЕДОТВРАЩАЮЩИЕ СКРЕЩИВАНИЯ РАЗНЫХ ВИДОВ

Видообразование предполагает приобретение видом независимой генетической системы, репродуктивно изолированной от других симпатрических видов.

Классификация изолирующих механизмов [Майр, 1974]:

- докопулятивные (презиготические):
 - биотопическая, сезонная, географическая, экологическая изоляция;
 - этологическая (поведенческая) изоляция;
 - механическая изоляция;
- посткопулятивные (постзиготические):
 - гибель гамет;
 - гибель зигот;
 - стерильность или низкая жизнеспособность гибридов.

Механизмы формирования репродуктивной изоляции и ее поддержания являются предметом многочисленных исследований, составляя основу эволюционной биологии. Репродуктивные барьеры очень специфичны и могут иметь промежуточные состояния. *Например*, распространена ситуация, когда самки вида А уже изолированы от самцов вида Б, а самки вида Б все еще могут давать плодовые гибриды с самцами вида А. Именно так устроена изоляция во многих парах сестринских видов *Drosophila*.

Для многих организмов использование генетического критерия путем скрещивания требует много времени, причём больше, чем живёт сам исследователь. Поэтому всё активнее применяют методы молекулярно-генетического анализа, который может прояснить филогенетическую близость изучаемых объектов и иногда – восстановить последовательность событий, приведших к изоляции.

1.3. МУТАЦИОННАЯ ИЗМЕНЧИВОСТЬ КАК ИСТОЧНИК БИОРАЗНООБРАЗИЯ

Источником наследственной изменчивости являются **мутации**. Этот термин впервые ввел Хуго де Фриз (1886), на основании изучения растений ослинника (*Oenothera*) он сформулировал мутационную теорию, которую изложил в книге «Мутации и периоды мутаций при происхождении видов» (1901).

Мутация, чаще всего, возникает в результате изменения одного гена и состоит в замещении, дупликации или делеции одной или нескольких пар оснований. В результате мутации возникают новые аллели, которые наследуются так же, как и исходные. Генные мутации представляют собой изменения последовательностей нуклеотидов, которые могут приводить к изменению фенотипа. Причиной их возникновения могут быть ошибки копирования в ходе репликации или модификации ДНК в клетках зародышевой линии под действием мутагенов различной природы. Физические мутагены: радиация, высокая и низкая температура, механические воздействия, ультразвук. Химические мутагены: различные органические и неорганические соединения.

Мутация – это результат сложной цепи биохимических реакций, и их частоты неравномерно распределены по геному. Мутации могут проявляться в изменении морфологии, физиологии, биохимии, поведения и пр. Существует множество различных классификаций мутаций, в основу которых положены различные принципы.

Типы мутаций

1. Генные – точечные мутации (внутригенные), приводят к изменению нуклеотидов внутри гена:

- к потере нуклеотидов (делеции);
- удвоению нуклеотидов (дупликации);
- вставке нуклеотидов;
- изменению порядка чередования нуклеотидов.

2. Хромосомные:

- транслокации (обмен участками между двумя негомологичными хромосомами);
- инверсия – поворот участка хромосомы на 180°;
- дупликации, нехватки, инверсии, инсерции.

3. Геномные:

- кратное изменение числа хромосом (полиплоидия и гаплоидия);
- добавление или потеря отдельных хромосом (анеуплоидия).

Мутации могут быть доминантными и рецессивными.

По относительному влиянию на жизнеспособность и плодовитость организма мутации бывают полезные, нейтральные и вредные.

Мутации, возникающие в гаметах и клетках, из которых образуется зигота, – наследуются. Мутации, возникающие в соматических клетках, не наследуются, но могут распространяться в популяции путём вегетативного размножения (клонирования).

По появлению мутации могут быть прямые и обратные (очень редкое явление).

В природе мутации возникают случайно, но их шансы на сохранение зависят от того, полезны они или нет в данных условиях среды. Мутации случайны в том смысле, что вероятность их возникновения не обусловлена потребностью. Однако условия среды могут оказаться благоприятными для распространения мутации в популяции или наоборот. Все гены участвуют в межгенных взаимодействиях. Поэтому в зависимости от генотипа влияние мутантного аллеля на жизнеспособность может существенно изменяться.

Несмотря на то что на уровне морфологии и физиологии мутации были обнаружены у многих организмов, долгое время не удавалось проследить их на молекулярном уровне. Однако на протяжении последних двух десятилетий были найдены методы их идентификации.

Мутации обычно имеют множественные – плеiotропные эффекты, обусловленные межгенными взаимодействиями, поэтому их определение как положительных, отрицательных или нейтральных довольно условно, так как один и тот же эффект в разных условиях может быть положительным, а в других отрицательным. Учитывая, что эффектов от мутации может быть несколько, целесообразней проводить оценку каждого из них и применительно к конкретным условиям. Вся селекция построена по принципу «тришкиного кафтана», когда использование одного из ценных в хозяйственном отношении эффекта всегда влечет за собой какие-либо отрицательные последствия. *Например*, формы гороха, которые имеют усатые листья, обладают высокой засухоустойчивостью, но низкой урожайностью.

Летальные мутации, казалось бы, не могут считаться как положительные, но это – для конкретного организма, а в целом для вида они несут положительный эффект, так как позволяют элиминировать безнадёжные в плане адаптации генотипы. Кроме того, мутации, которые летальны в гомозиготном состоянии, могут повышать жизнеспособность гетерозигот – моногенный гетерозис. *Например*: ген, вызывающий серповидно-клеточную анемию у человека, в гетерозиготном состоянии предотвращает заболевание малярией и поэтому широко распространён в популяциях людей, живущих в зоне распространения малярии.

Фенотипы организмов, отличающихся всего по одному гену, могут не отличаться от исходных фенотипов вовсе, могут иметь слабо заметные отличия, но могут отличаться сильнее, чем представители разных семейств. Так у *Arabidopsis thaliana* (модельного объекта для генетических исследований) растения дикого типа *Landsberg 0* (рис. 1, *A*) существенно отличаются от растений мутантных линий. В частности, в отличие от дикого типа у мутантов *bp-1 – brevipedicellus* цветоножки сильно укорочены и поэтому их плоды повёрнуты вниз (рис. 1, *B*); у растений мутантной линии – *ap1-1 – apetala* – цветок лишён лепестков (рис. 1, *C*); у растений мутантной линии *dw* (№ 232 в каталоге NASC) – *dwarf* – карлик практически редуцирован стебель и цветки располагаются одиночно или образуют соцветие из 2–4 цветков, что существенно меньше, чем у дикого типа (рис. 1, *D*).



Рис. 1. Мутантные линии *Arabidopsis thaliana*

- A* – *Landsberg erecta* – дикий тип;
- B* – *bp-1* – *brevipedicellus* – короткая цветоножка;
- C* – *ap1-1* – *apetala* – цветок без лепестков;
- D* – *dw* (№ 232 в каталоге NASC) – *dwarf* – карлик

Процесс возникновения мутаций – мутагенез по происхождению может быть *естественный* (спонтанный) или *искусственный* (индуцированный).

1.4. НЕНАСЛЕДУЕМАЯ (МОДИФИКАЦИОННАЯ) ИЗМЕНЧИВОСТЬ

Изменения под влиянием условий среды (модификационная изменчивость) не наследуются, хотя могут менять облик животных и растений в значительной степени. *Например*, почвенно-климатические условия также могут в значительной степени влиять на проявление как качественных, так и количественных признаков. Соли кальция, которые содержатся в мергельных почвах, делают клеточную оболочку более плотной и тормозят её растяжение. В результате растения, растущие на мергельных почвах, обычно имеют более мелкие листья, меньшую высоту и меньшее количество цветков, чем те, которые растут на чернозёмах.

Типы модификационной изменчивости:

● *Онтогенетическая* изменчивость – индивидуальная изменчивость:

- возрастная;
- сезонная;
- травматическая.

● *Социальная*.

● *Экологическая* – групповая изменчивость:

- средовая;
- экотопическая;
- сезонная.

Модификационная (паратипическая) изменчивость не передаётся по наследству и носит приспособительный характер (табл. 1).

Таблица 1

| Модификационная изменчивость | Мутационная изменчивость |
|----------------------------------|-------------------------------------|
| Ненаследуемые | Наследуемые |
| Носят приспособительный характер | Не носят приспособительный характер |
| Предсказуемы | Непредсказуемые |
| Массовые | Единичные |
| Обратимы | Необратимы |
| Не затрагивают ДНК | Затрагивают ДНК |

Модификационная изменчивость определяется размахом **нормы реакции**.

Норма реакции – пределы, в которых может изменяться фенотипическое проявление признака, в зависимости от условий внешней среды.

Мутации изменяют норму реакции. Учитывая, что в популяции существует множество мутаций в рецессивном состоянии, то в случае их проявления следует ожидать изменения фенотипа и нормы реакции признака, но это не может быть поводом для выделения такой группы особей в новый вид. Чем лучше изучен вид, тем более вероятно, что в нём будут обнаружены экологический полиморфизм или непрерывная экологическая изменчивость [Майр, 1974].

1.5. РАСПРЕДЕЛЕНИЕ ЗНАЧЕНИЙ ПРИЗНАКА

При исследовании отбора, а также просто для иллюстрации различных его механизмов часто используются вариационные ряды, которые представляют собой гистограмму, полученную на основании анализа частоты встречаемости ранжированных признаков. Для её построения в пределах выборки образцов проводят измерения значений какого-либо признака. Общую совокупность значений признака (от наименьшего до наибольшего) разбивают на несколько равных интервалов. Число интервалов зависит от числа измерений, и многие программы оценивают это число автоматически (*например*, высота растений *ler Arabidopsis thaliana* при определённых условиях среды варьировала в пределах от 15 до 22 см). Подсчитывают число измерений, которые попадают в каждый из интервалов, в результате чего получается ряд чисел, равный числу интервалов. Полученные значения откладывают на графике в виде кривой (рис. 2). В результате сумма всех частот будет равна единице, или 100 %.

Ширину разброса значений можно считать приблизительной оценкой разнообразия организмов по исследуемому признаку. Большое значение имеет также форма распределения. Распределения, имеющие единственный максимум, называются одномодальными. Если одномодальные распределения асимметричны, то их называют смещёнными. Отдельный интерес часто представляют распределения, у которых два или более максимума. Такие

распределения называются, соответственно, бимодальными или полимодальными. Обычно они отражают изменчивость в пределах выборки, и чем полнее выборка в пределах нормы реакции, тем ближе она будет к нормальному распределению. Например, у мутантной линии *Arabidopsis thaliana* число цветков на побегах варьирует от 1 до 25 и имеет распределение, близкое к нормальному (рис. 2, А).

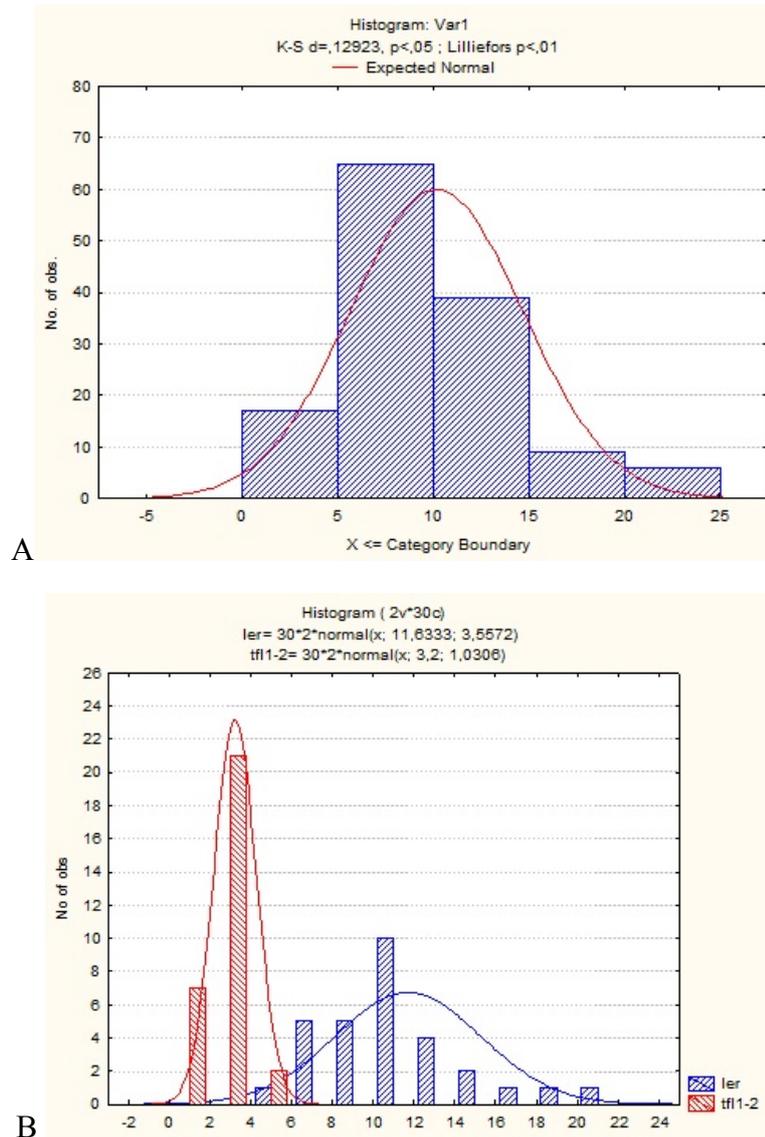


Рис. 2. Вариационные ряды изменчивости числа цветков у *Arabidopsis thaliana*.

А – у линии *ler* – данные взяты с учётом повторностей и с разных побегов;
 В – у линий *tfl1-2 ler* – данные одной выборки

У мутантной линии *tfl1-2* норма реакции уже, чем у *ler*, и число цветков у них изменяется в пределах от 1 до 6 (рис. 2, В). Причём если ограничиться небольшой выборкой, взятой из определённых условий и с определенной части растения, например только с главного побега, то их распределение признака не будет охватывать всю широту нормы реакции, а будет далеким от нормального. Поэтому нормальное распределение признака в пределах ареала вида может иметь решающее значение при анализе биоразнообразия, так как может свидетельствовать о фрагментарном анализе признака в пределах нормы реакции.

Разброс значений признака характеризует размах нормы реакции.

В пределах возможных значений признака можно выделить зоны оптимума и депрессии.

Зона оптимума – наиболее вероятное значение признака в данных условиях среды.

Зона депрессии – менее вероятное значение признака в данных условиях среды.

В целом общий разброс складывается из разброса, возникающего в результате генетической изменчивости в популяции и влияния условий среды. Это выражается формулой

$$\sigma^2_{\Sigma} = \sigma^2_e + \sigma^2_g,$$

где σ^2_{Σ} – общая дисперсия значений признака складывается из: σ^2_e – дисперсии, обусловленной влиянием среды, и σ^2_g – дисперсии, обусловленной генетическим разнообразием.

Из этой формулы следует, что у группы особей с одинаковыми генотипами все равно будет наблюдаться разброс значений признака, возникающий в результате небольших вариаций воздействия окружающей среды в процессе их индивидуального развития. Таким образом, формируется разнообразие фенотипов, и именно на них действуют различные формы отбора. Это действие приводит к тем или иным изменениям формы распределений значений признаков.

1.6. ДЕЙСТВИЕ ОТБОРА НА РАСПРЕДЕЛЕНИЕ ЗНАЧЕНИЙ ПРИЗНАКОВ

Естественным отбором, согласно взглядам Дарвина, называется небольшое, обусловленное влиянием окружающей среды изменение шансов оставить плодовитое потомство. Эта точка зрения принципиально отличается от экстремистского понимания естественного отбора как «выживания наиболее приспособленных», которое получило широкое распространение благодаря формулировке, предложенной одним из основоположников либерального мировоззрения, экономистом и философом Спенсером.

Несмотря на то что в «Происхождении видов» [Darwin, 1882] нет определения естественного отбора в современном понимании, там содержатся рассуждения: «Много мелких различий, которые появляются у потомства от одних и тех же родителей... могут быть названы индивидуальными различиями... Эти индивидуальные различия имеют для нас самое большое значение, так как они часто наследуются... таким образом они позволяют материалам для естественного отбора действовать и накапливаться». Четыре основных типа отбора проиллюстрированы на рис. 3.

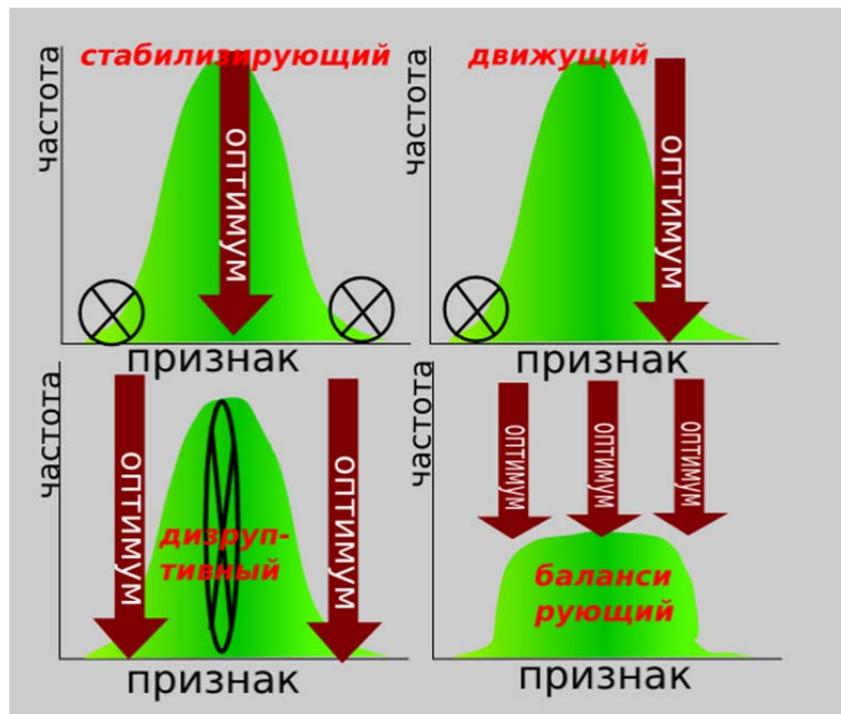


Рис. 3. Типы отбора

При **стабилизирующем отборе** сокращается доля крайних форм независимо от того, в какую сторону от оптимума произошло отклонение. При этой форме отбора разнообразие организмов уменьшается.

Движущий отбор происходит, если оптимум значения признака сдвигается в сторону одного из крайних значений (см. рис. 3), в результате чего максимум распределения сдвигается в соответствующую сторону. Поскольку любой отбор действует только на уже существующее разнообразие, максимум будет постепенно смещаться в сторону отбора. При этой форме отбора степень разнообразия постепенно уменьшается.

Дизруптивный отбор, в отличие от двух предыдущих механизмов, происходит, если возникают два оптимума, один вблизи максимального, другой – вблизи минимального значений признака (см. рис. 3). Со временем кривая распределения значений становится двухмодальной (т. е. с двумя пиками), впадина между максимумами становится все глубже. Предполагается, что именно разные вариации дизруптивного отбора участвуют в видообразовании. При этой форме отбора общее разнообразие увеличивается.

Балансирующий отбор поддерживает разнообразие как таковое. При этой форме отбора нет какого-то определенного оптимума (или оптимум очень широк), в результате поддерживается максимальное разнообразие значений признака (см. рис. 3).

Модификационная и мутационная изменчивость могут развиваться в популяции параллельно. *Например*, если в условиях с коротким периодом вегетации растения имеют меньшую высоту, чем обычно, то в случае появления среди них растения с карликовой мутацией его потомство может получить адаптивные преимущества, так как в таком случае развивается направленный отбор.

В природных популяциях распространение мутаций происходит под покровом модификаций [Шмальгаузен, 1957]. В селекционной практике отбор проводится искусственно и формы сохраняются целенаправленно, хотя в дикой природе они могли быть элиминированы отбором.

1.7. ГЕОГРАФИЧЕСКАЯ ИЗМЕНЧИВОСТЬ ВИДОВ И ИХ АРЕАЛЫ

Бесчисленные аспекты географической изменчивости вида и существование «полувинов» доказывают широкое распространение географического видообразования.

Географическая изоляция представляет собой чисто внешний фактор (действие которого полностью обратимо) и сама по себе не ведёт к образованию видов. Её роль заключается лишь в том, чтобы сделать возможной беспрепятственную генетическую перестройку популяций, которая служит предпосылкой для создания изолирующих механизмов.

Любой признак (морфологический, физиологический и пр.) может быть подвержен географической изменчивости в пределах ареала вида.

Ареал (от лат. *area* – площадь, пространство) – это территория (акватория), на которой распространён и проходит полный цикл своего развития данный таксон.

В случае протяжённого ареала размах этой изменчивости может быть весьма велик, поскольку в каждой части ареала вид должен отвечать требованиям местных условий.

Поскольку условия среды постепенно изменяются с изменением широты и долготы, то следует ожидать у разных видов в значительной мере параллельную изменчивость.

Правило, сформулированное в 1847 г. немецким биологом Карлом Бергманом, гласит, что среди сходных форм теплокровных животных наиболее крупными являются те, которые живут в условиях более холодного климата – в высоких широтах или в горах. *Например*: самые крупные виды медведей – полярные *Ursus maritima*, самые мелкие – индийские и африканские. Это правило не соблюдается для беспозвоночных и растений, где, наоборот, в экстремальных условиях созревание ускоряется, а размеры тела половозрелых особей – уменьшаются.

Фенотипический признак не означает, что он существенен для ареала в данном подвиде, а означает, что генотип, который возник в процессе эволюции в данной области, в результате отбора, его обуславливает. *Например*: рецессивный летальный аллель

гена серповидно-клеточной анемии приводит к устойчивости к малярии и широко распространен среди населения малярийных районов.

Физиологическая адаптация локальной популяции, имеющая адаптивное значение, при перенесении в другие условия может исчезнуть.

Параллельные изменения с изменениями широты и долготы широко распространены.

Климатические правила – это чисто эмпирические обобщения, описывающие параллелизм между морфологическими и физическими характеристиками среды.

Некоторые правила применимы при анализе видов изменчивости в пространстве и во времени.

1. Популяции внутри вида обычно отличаются генетически, а иногда и фенотипически.

2. Степень различий между разными популяциями вида варьирует от почти полного отсутствия до уровня, близкого к межвидовому.

3. Область, занимаемая внешне идентичными популяциями, может быть крайне мала или может включать весь ареал вида.

4. Различные признаки вида могут варьировать независимо.

5. Все признаки, используемые для различения видов, могут быть подвержены географической изменчивости в той или иной степени.

Признаки данной популяции имеют, обычно, генетическую основу и в большинстве случаев обнаруживают тенденцию сохранять относительное постоянство на протяжении ряда лет.

Экотипическая адаптация локальных популяций способствует повышению генетического разнообразия вида.

Географическая изменчивость проявляется на популяционном уровне, и это следует учитывать при определении ареала вида.

Границы ареала вида ограничены условной линией, за которой селективные факторы среды препятствуют успешному размножению. Отдельные особи могут появляться за этой линией, но при постоянных условиях они не могут там закрепиться. Даже если эта граница расширяется в период ослабления действия лимитирующих факторов, при их усилении она возвращается на преж-

нее место. В результате граница вида, хотя и динамична в целом, остаётся стабильной.

Сдвиги границы вида возможны, в частности, в связи с изменением климатических условий.

Границы распространения вида иногда совпадают с некоторыми климатическими факторами. Например: граница распространения *Rubus pepegrina* в Западной Европе совпадает с январской изотермой +4,5 °С [Тахтаджян, 1978].

Устойчивые долговременные изменения климата порождают изменения в распространении многих организмов. В этой связи необходимо отметить, что часть видов, которые внесены под охрану на границах своих ареалов, просто передвинули эти границы по причине временных климатических изменений или по причине завышения таксономического ранга локальных форм географической изменчивости.

Выдающийся ботаник В. Л. Комаров [1951] считал, что вид – это морфологическая система, помноженная на географическую определённость. Однако, учитывая, что границы ареала вида обусловлены влиянием лимитирующих факторов среды и изменяются в ходе истории его формирования, то их положение изменяется во времени и пространстве. *Например*, если температура является лимитирующим фактором для вида растений, то по мере потепления граница ареала будет продвигаться на север.

В зависимости от размеров и очертаний ареала различают виды: эндемики, космополиты.

Эндемики – это виды, которые встречаются только в одном месте. *Например*, *Pelargonium graveolens* L. Her. – эндемик Капских гор, кенгуру – эндемик Австралии.

К сожалению, в современной литературе термин «эндемизм» не всегда корректно используется. *Например*, иногда говорят о «сибирских» эндемиках, что предполагает весьма обширный ареал распространения с довольно большим разнообразием условий среды.

Космополиты – это виды (или надвидовые таксоны), которые имеют очень широкие ареалы, включающие несколько континентов. *Например*: *Rattus norvegicus* – серая крыса, *Drosophila melanogaster* – дрозофила, *Poa annua* – мятлик однолетний и др.

В целом соблюдается закономерность: чем шире видовой ареал, тем шире его спектр адаптаций.

Ареал сплошной, или непрерывный, если особи вида встречаются во всех подходящих для их жизни местах обитания. *Например*, серый воробей, домовая муха.

Однако ни один вид не встречается на площади своего ареала повсеместно, а заселяет только свойственные ему экологические ниши.

Ареал прерывистый (дизъюнктивный, разорванный), если он распадается на несколько разобобщенных территорий и предполагает географическую изоляцию между его участками. *Например*, у голубой сороки *Cyanopisca cyana* одна часть ареала находится в Европе, а другая – в Восточной Азии.

Участки, заселенные одной формой внутривидовой изменчивости, – **дизъюнкция**. Если на участке встречается одна форма, то дизъюнкция однородная, а если несколько, то разнородная (или гетерогенная).

Как для морских, так и для наземных живых организмов характерны **биполярные** разрывы ареалов, когда они встречаются одновременно в умеренно холодных регионах Северного и Южного полушарий, но отсутствуют в разделяющих их тропических зонах.

Территорию, на которой происходит становление вида, называют **первичным** ареалом, изменение исходной территории приводит к формированию **вторичного** ареала.

Иногда определить ареал достаточно проблематично из-за миграций (*например*, у лососёвых рыб или перелётных птиц).

Преобразования ареалов могут быть связаны с разными причинами:

- возникновением орографических или других географических преград;
- сукцессией сообществ, с которыми были связаны те или иные таксономические группы живых организмов, в результате изменений климата либо антропогенной деятельности;
- сокращением площади суши.

Дробление ареалов вследствие горообразования, антропогенного влияния и пр. может способствовать изоляции популяций, формированию репродуктивной изоляции и видообразова-

нию. Однако в результате инбридинга сокращается амплитуда изменчивости и жизнеспособность вида может снизиться, а численность может начать сокращаться вплоть до полного исчезновения популяции на изолированной территории.

В пределах ареала особи одного и того же вида могут отличаться по причине модификационной и внутривидовой изменчивости. В оптимальных для вида условиях могут выживать самые разные формы, поэтому в центре видообразования обычно наблюдается максимальное разнообразие, но по мере удаления из зоны оптимума будут оставаться только те формы, для которых данные условия благоприятны. Поэтому по мере удаления от центра происхождения вида внутривидовая изменчивость обычно сокращается. Кроме того, на краю ареала порой могут выживать только абберрантные формы, которые оказались бы неконкурентоспособны в центре видового ареала (рис. 4).

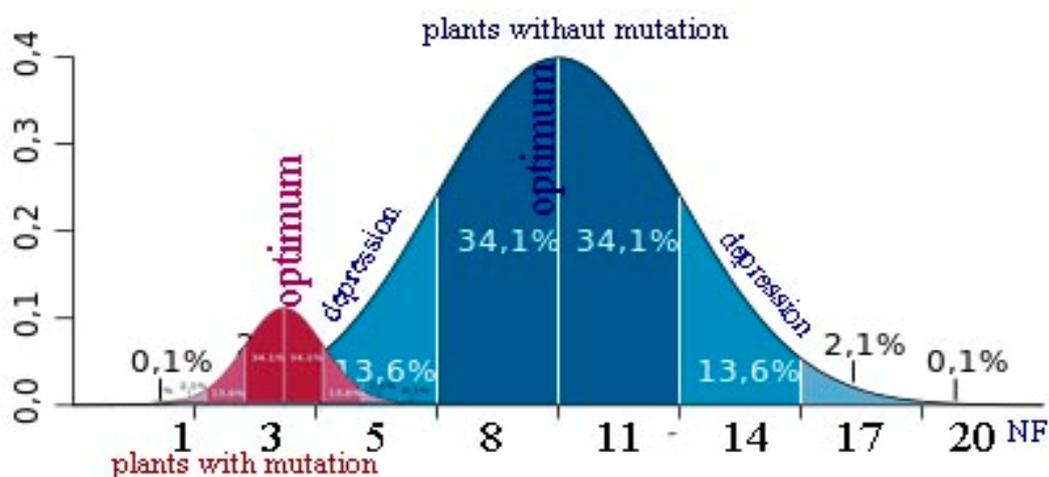


Рис. 4. Направленный отбор

Вид постоянно преобразуется, в связи с этим был сформулирован **Закон Красной королевы**: «чтобы оставаться на прежнем месте, нужно очень быстро бежать». Это значит, что условия обитания вида постоянно меняются и, чтобы он мог сохранить свой ареал, внутривидовая изменчивость должна происходить постоянно, но оставаться будут только те формы, для которых данные условия будут благоприятными.

1.8. ЭКОЛОГИЯ ВИДООБРАЗОВАНИЯ

Границы ареалов видов определяются лимитирующим действием климатических факторов. Они довольно лабильны и изменяются во времени. Климатические условия подвержены колебаниям, которые вызывают расширение и сужение границ ареалов. В периоды сокращения ареалов в пригодных для обитания местах часто сохраняются реликтовые популяции, которые могут достичь уровня вида.

Любая граница вида, определяемая климатическими факторами, находится в состоянии динамической стабильности. Существование многих изолятов можно объяснить только как результат происшедших в прошлом расширений ареалов в такие периоды, когда климат был для этого благоприятен.

Самое активное видообразование должно происходить в местах, которые особенно богаты географическими преградами.

Большое число подвидов свидетельствует о значительной локализации популяций, т. е. об уменьшении потока генов. Виды, состоящие из более крупных особей, имеют меньше форм внутривидовой изменчивости, но более широкие ареалы.

Не существует стандартной скорости видообразования.

1.8.1. К/R-стратегии адаптации

Чем лучше организм защищен от неблагоприятных условий среды и чем меньше его зависимость от среды, тем больше роль конкуренции между особями одного вида. В таких условиях с селективной точки зрения гораздо выгоднее производить потомков, хорошо подготовленных к самостоятельной жизни (К-стратегия), чем большое число потомков (R-стратегия).

Если имеет место значительная конкуренция за самок и период заботы о потомстве длителен, большое селективное преимущество будет давать продолжительность жизни. Это увеличивает внутривидовую конкуренцию и вызывает тенденцию к дальнейшему уменьшению числа потомков. (Многолетние деревья и множество семян, и недолговечные организмы со сменой поколений через 10 мин.)

Противоположная стратегия – R. Она состоит в продуцировании большого количества потомков, что приводит к уменьшению энергетических затрат в расчете на один организм. Процент выживания при этой стратегии крайне низок. Выгода состоит в том, что даже небольшое селективное преимущество, которое у K-стратегов было бы нивелировано дрейфом генов, подхватывается отбором и закрепляется в ряду поколений. То же относится и к признакам, понижающим жизнеспособность. Для R-стратегов характерна высокая эволюционная гибкость и способность к быстрой оккупации любых пригодных областей.

1.9. ГЕНЕТИЧЕСКАЯ ИЗМЕНЧИВОСТЬ В ПОПУЛЯЦИЯХ

В популяциях ограниченного размера в ряду поколений происходят случайные изменения частоты аллелей в тех случаях, когда их влияние на жизнеспособность организмов одинаково, такие аллели называются нейтральными. Эти колебания особенно значительны, если популяции состоят из небольшого (десятки) числа особей. В малочисленных популяциях это явление сильно уменьшает генетическое разнообразие (сравни рис. 5, A, C, где показано случайное изменение частот аллелей в популяции численностью в 20 и в 2000 особей). Такие случайные колебания частот аллелей называются **генетическим дрейфом**.

Крайним случаем дрейфа является ситуация, когда один из аллелей совсем исчезает из популяции не в результате действовавшего против него отбора, а в результате цепи случайностей.

Влияние дрейфа генов на генетическое разнообразие в популяции зависит не только от численности половозрелых особей в популяции, но и от целого ряда других факторов, к числу которых относятся соотношение полов, пloidность организмов, чередование многочисленных и малочисленных поколений и т. п.

Для того чтобы сравнивать популяционные явления у разных организмов, было введено понятие **эффективной численности популяции** N_e . В двух популяциях организмов с различными биологическими свойствами, но одинаковой эффективной численностью популяции влияние дрейфа генов на генетический по-

лиморфизм одинаково. Чем меньше эффективная численность популяции, тем меньше сказывается влияние аллеля на жизнеспособность. В малочисленных и многочисленных популяциях один и тот же аллель может вести себя в первом случае – как нейтральный, а во втором – как селективно значимый. В результате при низких значениях эффективного размера популяций аллели, определяющие признаки с меньшим адаптивным потенциалом, могут полностью вытеснить более ценные аллели в результате случайных (стохастических) процессов.

Влияние генетического дрейфа на частоты двух аллелей одного гена при отсутствии отбора может быть смоделировано (см. рис. 5).

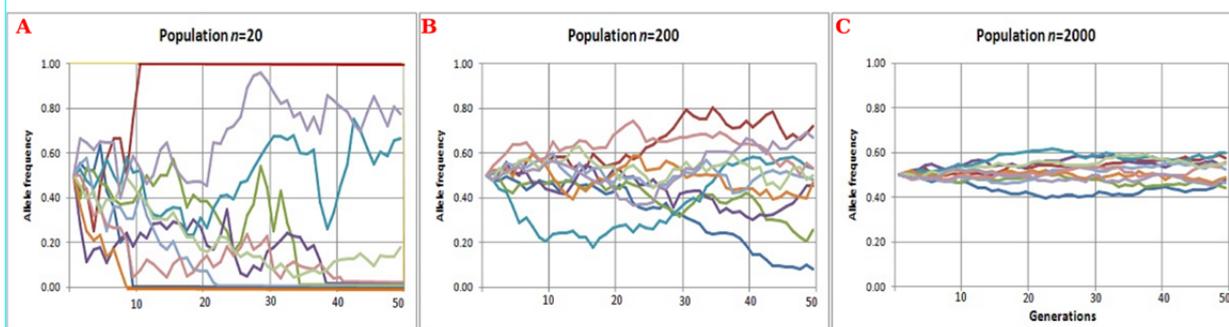


Рис. 5. Возможные сценарии судьбы мутантных аллелей в популяциях при отсутствии отбора. Постоянное число организмов в популяциях: *A* – 20; *B* – 200; *C* – 2000

В начале в каждой из популяций по 50 % аллелей находятся в одном состоянии, а 50 % – в альтернативном. Каждая кривая описывает изменение во времени доли первого аллеля в ряду поколений (время отложено по оси X и измеряется в поколениях). Условия компьютерного эксперимента таковы, что численность популяции остается постоянной, а аллели (за долями которых наблюдают по оси Y) одинаково сказываются на жизнеспособности и гомозигот и гетерозигот.

Использование при описании природных популяций эффективной численности вместо физической численности особей позволяет не только сравнивать между собой разные виды и популяции, но и прогнозировать их на основании теоретических расчётов.

На сегодняшний день предложено довольно много формул, позволяющих вычислить эффективную численность популяции, зная ее физическую численность, и учитывающих факторы, о которых было сказано выше. С другой стороны, зная эффективную численность, можно оценить физическую численность популяции. Это бывает полезно, когда измерить численность популяции оказывается сложной задачей, например для диких животных.

Судьбу мутантных аллелей в эволюции моделировали с самого начала развития теоретической популяционной генетики. Достаточно вспомнить, что среди условий, которые необходимы для выполнения закона Харди – Вайнберга, есть необходимость одинакового влияния аллелей на жизнеспособность. То есть – отсутствие отбора, или, другими словами, нейтральность породившей один из этих аллелей мутации. Результаты сравнения первых же последовательностей биополимеров (это были аминокислотные последовательности глобинов позвоночных) привели к выводу о линейности скорости накопления этих замен во времени. Оказалось, что этот эффект, получивший название «гипотезы молекулярных часов», можно полностью объяснить, если предположить, что большинство аллелей фиксируется в популяциях в результате дрейфа генов. Этот взгляд получил название «теории нейтральной эволюции», но в последнее время этот термин практически не используется, а взгляды «нейтралистов» стали практически составной частью синтетической теории эволюции.

Исчезновение аллелей в популяциях низкой численности приводит к уменьшению доли гетерозигот. Экспериментально определенная доля гетерозигот называется **наблюдаемой гетерозиготностью**, в отличие от **ожидаемой гетерозиготности**, которую можно вычислить с помощью правила Харди – Вайнберга. В малочисленной популяции велика вероятность инбридинга (близкородственного скрещивания).

Для количественной характеристики генетических процессов в популяциях используется несколько показателей, которые вычисляются из формулы Харди – Вайнберга.

Пусть p и q – частоты двух аллелей одного гена в популяции. Частотой в данном случае называют долю аллелей определенного вида от общего числа посчитанных. Для популяции, размер кото-

рой стремится к бесконечности, частота равна вероятности, что позволяет пользоваться математической теорией вероятности для вывода формул.

Пусть N диплоидов образуют $2N$ аллелей. Тогда q будет равно числу аллелей данного типа, поделенному на $2N$. Если в исследуемой популяции присутствуют только два аллеля некоего гена с частотами p и q , то p и q представляют собой вероятности встретить соответствующий аллель. По определению согласно теории вероятностей можно написать простое уравнение:

$$p + q = 1.$$

Вычислить частоты генотипов в следующем поколении можно, воспользовавшись правилом перемножения вероятностей: вероятности образования гомозигот будут p^2 и q^2 , а вероятность образования гетерозиготы – $2pq$. Следовательно, правило Харди – Вайнберга можно записать для частот генотипов так:

$$p^2 + 2pq + q^2 = 1.$$

Средний член в этом уравнении как раз и позволяет, зная частоты аллелей, подсчитать долю гетерозигот в равновесной, т. е. не подвергающейся никаким внешним воздействиям популяции. Важнейшее следствие этого правила состоит в том, что вероятности генотипов от поколения к поколению не изменяются и остаются теми же p и q .

Ниже рассматривается использование микросателлитных локусов для того, чтобы охарактеризовать процессы в популяциях или видах. Поскольку для этого типа маркеров характерен мультиаллелизм, то потребуется обобщение формулы Харди – Вайнберга для этого случая.

Пусть есть N аллелей некоего локуса, обозначенных как $x_1, \dots, x_i, \dots, x_N$.

Тогда для частот этих аллелей будет соблюдаться правило

$$x_1 + \dots + x_i + x_N = 1,$$

а генотипы в равновесной популяции будут подчиняться правилу

$$(x_1 + \dots + x_i + x_N)^2 = 1.$$

Отсюда можно вычислить ожидаемую гетерозиготность:

$$H_{\text{exp}} = 1 - (x_1 + \dots + x_i + x_N)^2,$$

что для двухаллельного частного случая принимает простую форму:

$$H_{\text{exp}} = 1 - (p + q)^2.$$

Наблюдаемую долю гетерозигот в популяции обозначают H_{obs} . Эту величину можно использовать, чтобы определить и подсчитать **коэффициент гетерозиготности**:

$$F_I = \frac{H_{\text{exp}} - H_{\text{obs}}}{H_{\text{exp}}}.$$

Если этот коэффициент больше нуля, то мы имеем дело с заметной долей близкородственных скрещиваний, скорее всего, в относительно малочисленной популяции. Если этот коэффициент становится равным единице, то это будет означать, скорее всего, самооплодотворяющуюся «популяцию».

Распространенной причиной инбридинга является самоопыление. Оно становится предпочтительным, если организмы или их опылители являются редкими. Это может повысить шансы на сохранение и распространение видов, на ограниченный срок в стабильных условиях среды. Однако, по мнению Stebbins [1957], самоопыляющиеся виды редко оказываются долгоживущими. Так как вследствие инбридинга накопление вредных мутаций за счет самоопыления может привести к увеличению скорости исчезновения вида [Medawar, 1952] и вероятность этого будет возрастать по мере изменения условий его произрастания. Однако «вредность» и «полезность» мутаций являются условными категориями. Отдельные изменения фенотипических признаков (цвета, формы, размера и пр.), возникшие под влиянием мутаций, могут не иметь значения в ходе отбора, но только весь спектр изменений, обусловленных мутацией, применительно к конкретным условиям среды может позволить понять, вредной или полезной она является.

Аутбридинг – скрещивание неродственных организмов.

Точность определения таксономического статуса группы особей имеет большое значение для селекционной практики, так как определяет успешность проводимых селекционных мероприятий.

Межвидовые скрещивания редко оказываются удачными, в то время как внутривидовые ведут к эффекту гетерозиса, который приводит к тому, что гибриды превосходят исходные родительские организмы по ряду признаков [Лобашёв, 1970].

При скрещивании чистых линий обычно наблюдается эффект гетерозиса, который имеет большую хозяйственную ценность, однако в следующих поколениях он затухает вследствие расщепления. Закрепление гетерозиса можно осуществлять путем перевода гибридного организма с нормального полового размножения на бесполое (вегетативное или апомиктичное). Эффекты гибридизации, как и эффекты мутаций, могут иметь как положительный, так и отрицательный или нейтральный эффект в зависимости от влияния окружающей среды. *Например*: мутация $l(2)gl$ – в гомозиготном состоянии вызывает злокачественную опухоль мозга у личинок III возраста *Drosophila melanogaster*, однако в гетерозиготном состоянии она повышает устойчивость к резким колебаниям температуры.

В популяциях организмы могут иметь значительные морфологические отличия, но при этом представлять собой разные стадии развития одного организма. *Например*, у кишечно-полостных есть две формы существования у аурелии: полипы – прикрепленная форма, которая размножается почкованием, и медузы – свободно плавающая форма, которая размножается половым путём.

Кроме того, существуют группы видов-близнецов, у которых при сходном фенотипе и даже совместном проживании наблюдается строгая репродуктивная изоляция, так как их сходство может быть адаптацией, которая развивалась на разных генотипических основах. *Например*, полёвка обыкновенная ($2n = 46$) и полёвка восточноевропейская ($2n = 54$).

В ходе онтогенеза организмы адаптируются к условиям окружающей среды и могут претерпевать модификационную изменчивость [Лобашёв, 1970] вследствие соматических мутаций и пр., однако они не наследуются. Для животных и растений существует определённая специфика развития, в частности условия среды могут прерывать онтогенез (морфогенез) на разных стадиях, это может способствовать их адаптации к условиям среды, но такая адаптация не наследуется. Наследоваться, однако, может

способность переживать нарушения такого рода. Однако если в популяции возникнет мутация, предполагающая формирование соответствующего фенотипа, то она будет иметь адаптивные преимущества и через несколько поколений может получить широкое распространение в данной популяции.

Появление в популяции аллеля, положительно влияющего на жизнеспособность, может приводить к тому, что в результате действия естественного отбора его частота будет возрастать, до тех пор, пока он не вытеснит менее эффективные аллели. В процессе увеличения численности такой аллель «захватит с собой попутчиков» – нейтральные аллели, которые находятся в соседних локусах. Частота нового аллеля, находящегося под действием положительного отбора, увеличивается, а соседние, сцепленные с ним аллели, наследуются попутно, увеличивая свою частоту и создавая локальный эффект бутылочного горлышка.

Например, эволюционная адаптация в ряду поколений благодаря появлению благоприятной мутации (обозначена красным) приводит к уменьшению генетического разнообразия в прилегающей области (рис. 6).

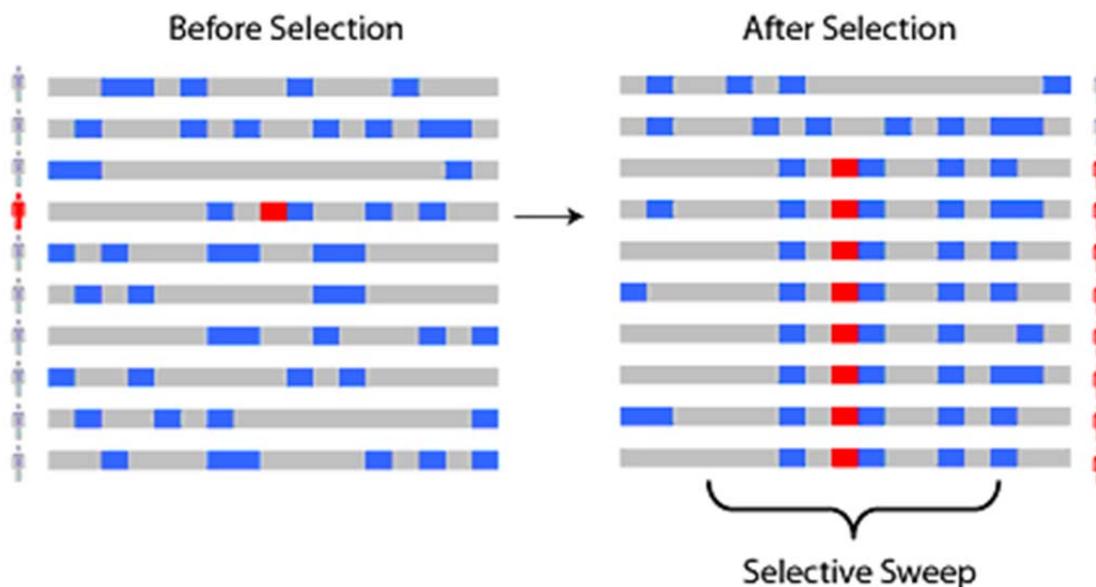


Рис. 6. Эффект уменьшения генетического разнообразия генов, соседних с геном, находящимся под действием отбора

Серым цветом обозначены предковые аллели, синим цветом – производные аллели (непредковые), красным цветом – показан аллель, который находится под действием положительного отбора [по Schaffner&Sabeti, 2008]

Из-за низкой частоты рекомбинации обеднение генетического разнообразия охватывает более широкий район. Таким образом, в результате движущего отбора происходит уменьшение генетического разнообразия. В результате появляется возможность эмпирически выделить район хромосомы, в котором недавно произошла положительная мутация, и, потом, использовать методы молекулярной генетики для ее поиска и расшифровки молекулярного механизма, вызвавшего это изменение.

2. ОСНОВНЫЕ МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА БИОРАЗНООБРАЗИЯ

Для анализа биоразнообразия применяется множество молекулярно-биологических методов, среди которых можно выделить три группы.

1. Идентификация организмов с помощью одного или нескольких стандартных фрагментов ДНК, уникальных для своих носителей, называется **штрихкодированием** (см. ниже).

2. На основании штрихкодов может быть идентифицирован видовой состав биотопа, в котором организмы находятся на такой стадии онтогенеза, которая не позволяет идентифицировать их визуально или представлена фрагментами организмов. Естественно, такой подход – **метагеномный анализ** – не учитывает очень многие параметры – стадию онтогенеза, размер, ярусность и пр. Однако, в ряде случаев, метагеномный подход оказывается единственным, который может дать хоть какое-то представление о видовом составе организмов.

3. **Микросателлитный анализ** основан на использовании очень быстро эволюционирующих молекул. Обычно его используют для идентификации личности, отцовства и пр. *Например*, в зоологии он был применён для выяснения отцов в рое пчёл, а у людей – получил широкое распространение для выяснения путей этногенеза.

Все перечисленные подходы имеют свои преимущества и недостатки, которые необходимо учитывать при планировании собственных экспериментов. Соответствующие вопросы мы подробно обсудим ниже. Но начать все же следует с того, что объединяет все перечисленные подходы: все они так или иначе используют полимеразную цепную реакцию (ПЦР), или амплификацию ДНК.

2.1. ПОЛИМЕРАЗНАЯ ЦЕПНАЯ РЕАКЦИЯ (PCR)

Изобретение полимеразной цепной реакции (ПЦР или PCR) в конце 80-х – начале 90-х гг. революционизировало всю молекулярную биологию.

Полимеразная цепная реакция основана на общей особенности ДНК-зависимых ДНК-полимераз, которая способна только продолжать уже имеющуюся цепочку нуклеотидов. Эта особенность синтеза ДНК позволяет практически безошибочно удваивать целый геном в процессе каждого жизненного цикла клетки, но ни одна из ДНК-полимераз не может начинать синтез ДНК *de novo*. Не обнаружено ни одного случая, когда этот фермент связался с каким-либо кусочком ДНК и двумя дезоксирибонуклеотид трифосфатами, из которых синтезируется ДНК, отщепил бы от одного из них пирофосфат и образовал бы первую фосфодиэфирную связь, при этом соблюдая принцип комплементарности.

ДНК-полимераза **способна только продолжать уже имеющуюся цепочку нуклеотидов.** Эта особенность синтеза ДНК позволяет практически безошибочно удваивать целый геном в процессе каждого жизненного цикла клетки. То есть в живой клетке удвоение ДНК начинается с появления «насечки» (по-англ. *nick* – зарубка) – гидролизованной фосфодиэфирной связи, причем гидролиз может происходить только одним из двух возможных способов. То есть в живой клетке удвоение ДНК начинается с появления насечки – гидролизованной фосфодиэфирной связи. Если эта насечка содержит свободную 3'-гидроксильную группу, то ДНК-полимераза может удлинять цепь нуклеотидов, присоединяя новые один за другим. Наличие или отсутствие фосфата на 5'-конце перед ДНК-полимеразой не имеет значения.

Участок ДНК, прилегающий к этому свободному гидроксилу (3'-концу), – называется праймером.

Последовательность событий PCR состоит из трех стадий:

- 1) нагрев (денатурация);
- 2) охлаждение (отжиг);
- 3) инкубация в подходящих для работы ДНК-полимеразы условиях (элонгация).

Эта последовательность событий превращается в полимерную цепную реакцию, если её производить с одним и тем же препаратом ДНК многократно. Необходимо отметить, что в идеале каждое повторение полного цикла должно приводить к удвоению участка генома, который находится между праймерами. В результате, если, например, перед началом эксперимента в пробирке было x молекул ДНК, содержащих нужный нам участок, и 3 стадии реакции провели n раз, то максимальное число молекул этого участка N можно оценить по формуле $N = 2^n x$. Количество ДНК вне этого участка будет также нарастать, однако согласно линейной зависимости $N_1 = knx$. Обычно используется не менее 20 циклов (повторов трёх стадий). В идеальных условиях это привело бы к увеличению числа копий интересующего нас участка ДНК примерно в миллион раз, тогда как остальная ДНК увеличила бы свою концентрацию раз в 10, такое увеличение концентрации называется степенью амплификации, обычно его удается достичь в результате 30–35 циклов. Это вызвано различными причинами, и не в последнюю очередь – истощением реакционной смеси. Экспоненциальное увеличение числа копий ДНК в процессе ПЦР роднит этот процесс с реакцией распада урана-235. Такая кинетика в физической химии носит название «цепной». Именно таково происхождение слова «цепная» в ПЦР.

В ПЦР требуется периодически нагревать реакционную смесь почти до температуры кипения, однако общеизвестно свойство ферментов терять свою активность при повышенной температуре из-за денатурации. В связи с этим можно было бы ожидать, что эта смесь тоже денатурируется, однако этого не происходит благодаря использованию термостабильной ДНК-полимеразы – выделенной из экстремофилов (бактерий и архей, обитающих в очень горячей воде). Первой такой бактерией стала *Thermus aquaticus*, для которой оптимальной является температура выше 70 °С. Затем стали использовать ферменты из микроорганизмов, населяющих кальдеры гейзеров и «черные курильщики» – вулканические выходы на глубине нескольких километров, где вода прогревается существенно выше 100 °С. В настоящее время используют термостабильные полимеразы, гены которых были введены в *Escherichia coli*.

Синтетический праймер – олигонуклеотид длиной 20–30 нуклеотидов – добавляют к раствору ДНК, после чего его нагревают до 94 °С и ДНК полностью денатурирует. Затем, в процессе охлаждения смеси олигонуклеотид, из-за своей высокой концентрации, быстро прилипает к комплементарным участкам. Таким образом, праймеры быстро будут находить комплементарные им участки ДНК и образовывать с ними двойные спирали – готовые старты для ДНК-полимеразы. В смеси, содержащей ДНК-полимеразу и нуклеотиды, а также прочие необходимые ингредиенты, начнется синтез ДНК, и этот синтез, в подавляющем большинстве случаев, начнется с искусственных праймеров. Если добавить в смесь два праймера, направленных навстречу друг другу и комплементарных участкам на разных цепях, то фрагмент ДНК, ограниченный этой парой праймеров, будет удваиваться, а вся остальная геномная ДНК – нет.

Фрагмент ДНК с праймерами на флангах называется ***ампликоном***.

В настоящее время существует бесчисленное число вариаций метода ПЦР, которые используются в самых разных целях. Те варианты, которые наиболее широко используются в молекулярной экологии, мы обсудим детально в соответствующих разделах.

Технологии ПЦР и её модификациям посвящен целый ряд специальных монографий и больших обзоров, в частности:

<http://dna-technology.ru/files/images/metodichki/OsnoviPCR.pdf>;

<http://evrogen.ru/kit-user-manuals/Evrogen-PCR-recommendation.pdf>.

2.2. ЭКСТРАКЦИЯ ДНК

Подавляющее большинство молекулярно-генетических методов анализа биоразнообразия начинаются с экстракции ДНК, и очень часто именно эта стадия оказывается критической для успеха всего проекта. Поэтому на ней следует остановиться детально.

Таксономическое разнообразие сообществ в природе очень велико. Образцами для анализа могут быть пробы со дна океана, вулкана, соскобы с поверхности космической станции, с участка леса и пр. Они могут включать не только живые организмы (бактерии, растения, животные, грибы и пр.), но и разные типы суб-

стратов и других примесей. Соответственно, для их анализа используют разные методики выделения ДНК.

В общем, требования к процедуре экстракции ДНК сводятся к тому, чтобы все ткани были максимально разрушены и при этом произошла очистка от веществ, ингибирующих ДНК-полимеразу (например, кислых полисахаридов, похожих на гепарин). Все эти методы сочетают механическое разрушение (например, замораживание в жидком азоте с последующим механическим измельчением в замороженном состоянии) с химическими воздействиями, которые обеспечивают переход ДНК в растворимое состояние. Функции применяемых компонентов растворителя представлены в табл. 2 на примере протокола Doyle and Dickson [1987].

Таблица 2

Протокол Doyle and Dickson [1987]

| Компоненты раствора | Функция |
|--|---|
| 2 % СТАВ (цетил триметил-аммоний бромид) | Катионный детергент (гидрофобная часть положительно заряжена): – растворяет мембраны; – при низкой ионной силе образует с нуклеиновыми кислотами и кислыми полисахаридами нерастворимые в воде соли |
| 4M NaCl | Обеспечивает высокую ионную силу, необходимую для удержания ДНК в растворе и для того, чтобы она диссоциировала с гистонами |
| 0,1M Na ₂ EDTA | Связывает ионы двухвалентных металлов |
| 0,1M 2-меркаптоэтанол | Восстанавливает дисульфидные мостики |
| 0,1M Трис-HCl pH 8,0–8,6 | Буфер, предотвращает избыточное закисление смеси при нагревании до 65 °C |
| Поливинил пирролидон | Способствует осаждению многих ингибиторов ДНК-полимеразы |

Дальнейшая очистка ДНК производится путем центрифугирования, экстракции хлороформом или фенол-хлороформом, в результате которого происходит отделение белковых примесей и повторное осаждение этиловым или изопропиловым спиртом. Другие протоколы для экстракции ДНК основаны на сочетании тех же самых функций, и различия между ними специфичны для объекта исследования.

2.3. ШТРИХКОД ЖИЗНИ

Одно из наиболее эффективных и эффективных приложений ПЦР при исследовании биоразнообразия – программа создания «Штрихкода жизни» (синоним – баркодинг). Эта программа получила свое название по аналогии со штрихкодом, который используется для идентификации товаров (рис. 7, А). Аналогия состоит в том, что использование стандартных генов в качестве маркеров видов может позволить однозначную идентификацию организма. Более того, последовательность ДНК можно представить в виде серии цветных полосок (рис. 7, В).



Рис. 7. Примеры штрихкодов

А – штрихкод товара; В – графическое представление штрихкода жизни
[www.barcoding.si.edu]

Последовательность действий при применении штрихкода жизни

1. Выделить тотальную ДНК из образца ткани исследуемого организма.
2. Экстракция ДНК.
3. ПЦР-амплификация одного или нескольких стандартных молекулярных маркеров.
4. Определение нуклеотидной последовательности ампликона (обычно выполняется специализированными фирмами).

5. Сравнение нуклеотидной последовательности с известными последовательностями из банка данных, например NCBI [<http://ncbi.nlm.nih.gov>].

Зная штрихкоды, можно решать самые разнообразные задачи. Например, выделяя ДНК из вечной мерзлоты, можно изучать вымершую флору и фауну. Анализ штрихкодов необходим для выяснения путей передачи заболеваний, контроля качества лекарственных растений и т. п.

Созданием и поддержкой проекта, посвященного созданию и совершенствованию штрихкода живущих на Земле организмов, занимается международная некоммерческая организация International Barcode of Life (iBOL) (Международный штрихкод жизни [<http://www.barcodeoflife.org/>]). Цель этого проекта – создать базу данных (для максимально возможного числа видов живых организмов. Каждому виду присваивается цифровой «штрихкод» (как товару в супермаркете). Он указывает на название вида, численность, местообитание и пр. Чтобы его узнать, биологи изучают строение определенного участка ДНК (для животных это фрагмент митохондриального гена, который кодирует одну субъединицу фермента цитохром-*b* оксидазы). Строение этого участка ДНК различается у разных видов птиц, насекомых, рыб, млекопитающих и других животных. Его достаточно просто определить в любом живом организме или в его остатках.

Метод баркодинга направлен на изучение глобального биоразнообразия. К 2015 г. ученые поставили задачу собрать в «библиотеке» штрихкоды для 500 тыс. видов: животных, растений и грибов. На сегодня в ней содержится 167 тыс. видов, одних только бабочек 60 тыс. видов. Библиотека птиц покрывает около 40 % всех видов.

Отдельного внимания заслуживает вопрос о маркерах, которые используются для штрихкодирования различных групп живых организмов.

Требования к маркерам для штрихкодирования:

- 1) возможность амплифицировать из ДНК;
- 2) нуклеотидные последовательности праймеров должны быть очень консервативны и иметь минимум изменений в ходе эволюции у многих видов;

3) последовательности, заключенные между этими праймерами, должны иметь достаточно высокую изменчивость на межвидовом уровне, обеспечивая тем самым надежную идентификацию, но уровень их внутривидовой изменчивости не должен быть слишком высок или быть близок к уровню межвидовой изменчивости.

У подавляющего большинства животных этим требованиям удовлетворяет 5'-концевой фрагмент расположенного в митохондриальном геноме гена, кодирующего первую субъединицу цитохром-*b* оксидазы (обозначается COI или *coxI*). Для менее точных таксономических исследований используются последовательности генов, кодирующих субъединицы рибосомных ДНК, а для того, чтобы разделить очень близкие виды – следует использовать последовательности транскрибируемых спейсеров – фрагментов ДНК, расположенных между генами рРНК, которые транскрибируются, но сразу же вырезаются в ядре, обеспечивая эквимоллярность синтеза рРНК. Обозначаются эти фрагменты ITS-1 и ITS-2. Для идентификации прокариот, архей и простейших, включая грибы, в основном также используются гены рибосомных РНК.

Для растений наиболее распространено использование последовательностей трех генов – *matK*, *rbcL*, and *rpoC1* (табл. 3). Они кодируют, соответственно, хлоропластную матуразу К, расположенную в интроне другого гена (*trnK*) (рис. 8), большую субъединицу хлоропластного гена рибулезобифосфаткарбоксилазы и β'-субъединицу РНК-полимеразы.

Таблица 3

Гены, наиболее распространённые для штрихкодирования

| Сокращённое название гена | Полное название гена | Локализация нуклеотидных кислот | Степень консервативности относительно других генов данного вида |
|---------------------------|---|---------------------------------|---|
| Животные | | | |
| <i>coxI</i> | I субъединица цитохром- <i>b</i> оксидазы | Митохондрии | Низкая |
| SSU rRNA (18S rRNA) | Ген малой субъединицы рибосомной РНК | Ядро | Высокая |
| LSU rRNA (28S rRNA) | Ген большой субъединицы рибосомной РНК | Ядро | Высокая |

| Сокращённое название гена | Полное название гена | Локализация нуклеотидных кислот | Степень консервативности относительно других генов данного вида |
|----------------------------|---|---------------------------------|---|
| Растения | | | |
| <i>matK</i> | maturase K | Хлоропласт | Высокая |
| <i>rbcL</i> | большая субъединица хлоропластного гена рибулезобифосфат-карбоксилазы | Хлоропласт | Средняя |
| <i>rpoC1</i> | β' -субъединица РНК-полимеразы | Хлоропласт | Высокая |
| <i>ITS2</i> и <i>ITS-1</i> | Транскрибируемые спейсеры рДНК | Хлоропласт | Низкая |
| SSU rRNA (18S rRNA) | Ген малой субъединицы рибосомной РНК | Ядро | Высокая |

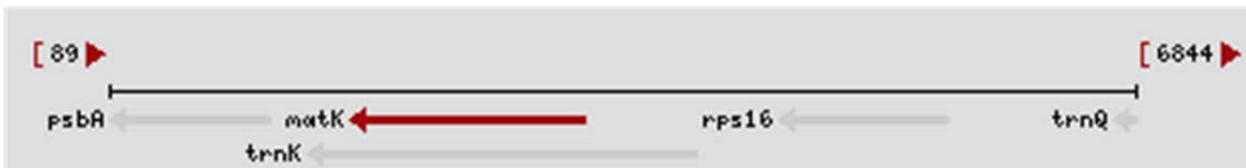


Рис. 8. Карта района хлоропластного генома *Zea mays*, прилегающая к гену *matK*

К сожалению, до сих пор нет единого мнения по поводу генов, пригодных для штрихкодирования растений. Причина состоит в очень неравномерных скоростях накопления нуклеотидных замен (см. ниже параграф, посвященный молекулярным часам). Проблема состоит в том, что помещенные в базу данных разными группами исследователей нуклеотидные последовательности, определенные для разных видов, оказываются идентичными на 100 %.

Неопределенности, возникающие при идентификации растений, можно проиллюстрировать примером, приведенным на рис. 9. На этом рисунке представлена эффективность видовой идентификации различных сосудистых растений Канады при использовании трех различных маркеров и трех различных методов сравнения нуклеотидных последовательностей с базой данных о штрихкодах (три панели, слева направо – методы BLAST, mothur1

и *mothur2*, детали методов изложены в соответствующих руководствах, и здесь мы их опускаем). В качестве 100 % использована видовая идентификация, сделанная профессиональными ботаниками, специализирующимися в систематике исследованных групп растений.

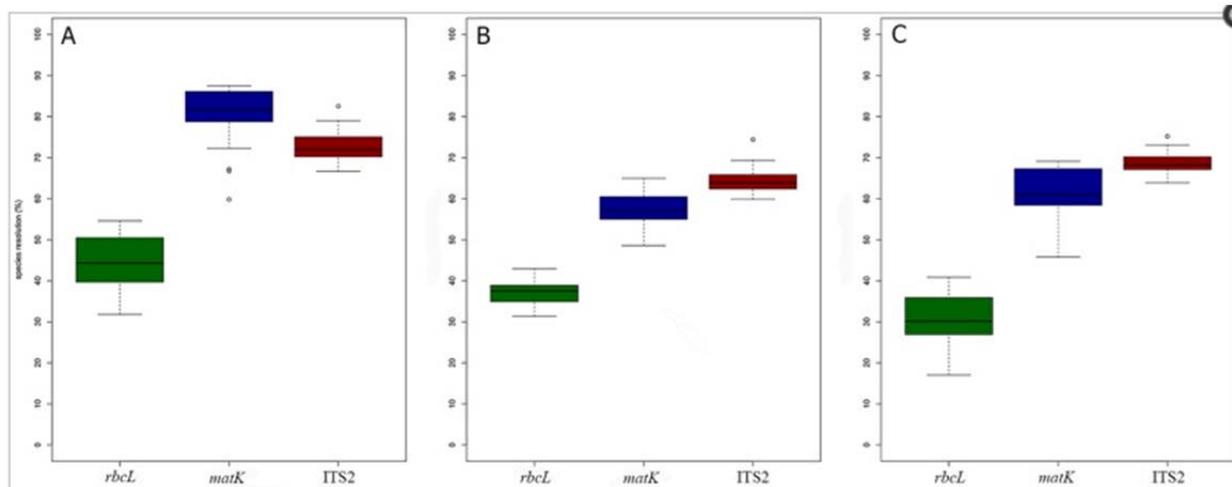


Рис. 9. Эффективность определения видов с помощью штрихкода [Testing the Efficacy ... , 2017]

Последний пример подчеркивает важнейшую роль, которую в штрихкодировании играет профессиональная идентификация видов. Ошибка определения организма, последовательность маркерных генов которого отсутствует в базе данных, будет затем многократно воспроизведена всеми пользователями этой базы и создаст риск неверных оценок экологической обстановки. Следует помнить и о том, что читателей таксономической специализированной литературы – единицы, а применение штрихкодов для решения самых различных задач в основном прикладного характера носит массовый характер и исчисляется тысячами.

Другая проблема состоит во вредных последствиях таксономических ревизий как таковых. Любая трансформация существующей системы должна быть серьезно обоснована, поскольку будет сопровождаться соответствующими изменениями в широко используемых базах данных, к тому же с большой задержкой, что сделает трудно сравнимыми результаты исследований биоразнообразия разными группами исследователей и в разное время. Особенно серьезной угрозой в этом контексте оказывается неоправданное дробление видов. Примеры его многочисленны, а ре-

зультат – появление большого количества синонимичных идентификаций при попытке сравнения штрихкода с базой. Другими словами, столбцы на рис. 9 в результате такой «деятельности» будут смещаться вниз.

С другой стороны, следует отметить несомненные достоинства метода штрихкодирования.

Для идентификации организма, принадлежащего к уже имеющемуся в базе виду, достаточно любого фрагмента организма. При этом не имеет значения, на какой стадии развития этот организм находится. Данное свойство штрихкодирования особенно важно при анализе природных образцов. «Морфологическое» определение с точностью до вида зачастую возможно только для половозрелых организмов, которые в пробах обычно составляют меньшинство. Остальные при традиционном исследовании определить не удастся. *Например*, широко известны мутации у *Drosophila melanogaster*, которые приводят к развитию вместо жужжалец полноценной второй пары крыльев. На основании одного этого признака можно было бы усомниться в принадлежности такого организма к Diptera. Штрихкод может снять все сомнения.

Метод объективен: несмотря на цепочку неизбежных преобразований систематики, результаты, подтвержденные штрихкодированием, сохраняют свою значимость и в будущем.

При массовом прикладном использовании штрихкодирование не требует непосредственного участия профессиональных таксономистов, специализирующихся на группах организмов, которых касается работа.

Метод штрихкодирования – на сегодняшний день является наиболее эффективным методом видовой идентификации, не зависящим ни от стадии развития организма, ни от того, какой фрагмент находится в распоряжении исследователя. Соответственно, он находит себе множество различных применений. *Например*, помогает выявлять «жульничество» производителей, в частности можно выявить подлог продукции из рыбы, растений и пр., который невозможно определить иначе.

Штрихкодирование позволяет получить информацию о вымерших видах, остатки которых законсервированы в вечной мерзлоте. Специальный проект BarFrost (баркодинг вечной мерзлоты) позволяет определить видовую принадлежность остатков возрастом от 10 тысяч до нескольких сотен тысяч лет.

С помощью штрихкодирования удастся исследовать видовой состав содержимого желудков и кишечника разных животных и определить их рацион. *Например:* в университете Аделаиды изучили рацион верблюдов, которые оказывают большое влияние на экологию. Верблюды поедают примерно 80 % всех растений, произрастающих на территории.

Другое важнейшее направление использования молекулярной идентификации – инвазивные виды. Их представители попадают в экосистему издалека и могут нанести большой вред местным видам и целым экосистемам. Среди таких видов – сорняки, насекомые – вредители сельского хозяйства и переносчики болезней, грибы, нематоды.

О качестве воды в водоемах можно судить по видовому разнообразию водных обитателей. Но на то, чтобы вручную разобрать пробы воды и донных осадков и под микроскопом найти и определить до вида все обитающие там организмы, могут уйти месяцы работы. Штрихкодирование позволяет сделать это легко и быстро.

2.4. ГЕНЕТИЧЕСКИЕ МАРКЕРЫ, ИСПОЛЬЗУЕМЫЕ ДЛЯ ИССЛЕДОВАНИЯ БИОРАЗНООБРАЗИЯ

Набор генетических маркеров, используемых в исследованиях биоразнообразия, весьма широк и зависит от конкретной задачи и возможностей.

Маркеры, используемые без определения нуклеотидных последовательностей.

RFLP (Restriction Fragment Length Polymorphism). Полиморфизм длин рестрикционных фрагментов. Определенный фрагмент разрезают эндонуклеазами рестрикции – ферментами, узнающими определенную последовательность ДНК, *например* GGATCC, и гидролизующими ее в этих местах. Любая нуклеотидная замена в таком локусе предотвращает гидролиз, и на электрофореграмме появится полоса, соответствующая сумме полос, между которыми расположен мутировавший сайт. Эти маркеры относительно недороги и широко использовались до распространения методов секвенирования ДНК (рис. 10).

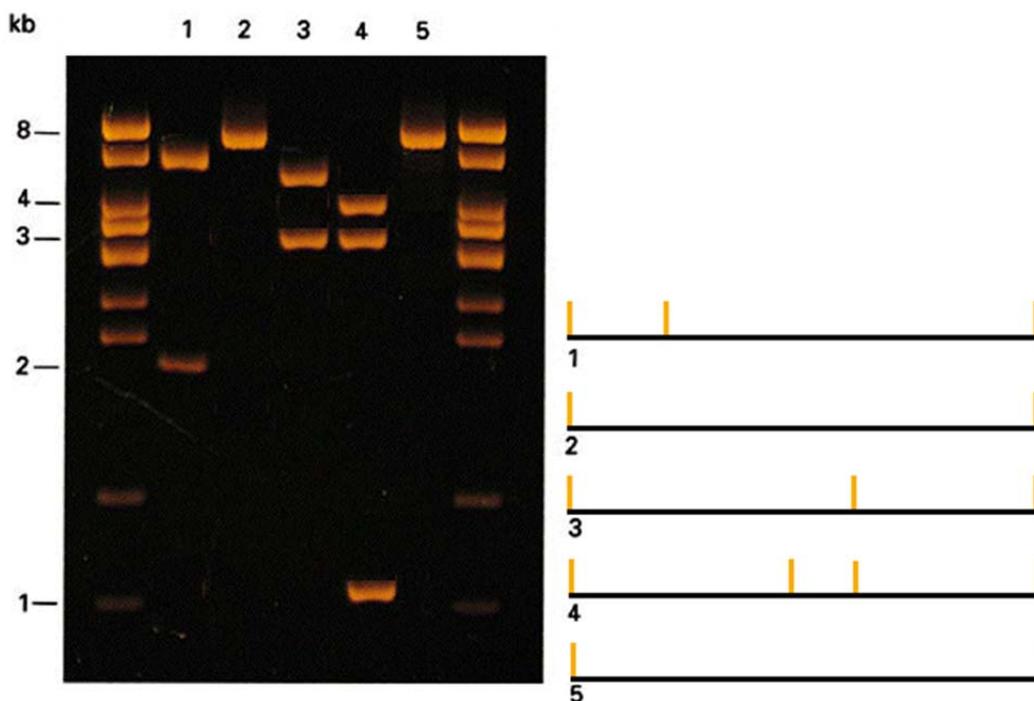


Рис. 10. Пример RFLP

Крайние дорожки – маркеры молекулярного веса. Слева – электрофореграмма, окрашенная флуоресцентным красителем, справа – карты расположения сайтов, по которым прошел гидролиз фрагмента ДНК в соответствующих дорожках

RAPD – Randomly Amplified Polymorphic DNA (случайно амплифицированная полиморфная ДНК). Для амплификации используется единственный праймер, что позволяет амплифицировать ДНК, находящиеся в составе инвертированных повторов. В качестве аллельного состояния такого признака приводят его длину в нуклеотидах. Полиморфизм возникает в результате инсерций или делеций в RAPD-фрагменте. В настоящее время эти маркеры используются очень редко из-за очень низкой воспроизводимости экспериментальных методов (рис. 11).

AFLP – Amplified Fragment Length Polymorphism (полиморфизм длин рестриктов амплифицированных фрагментов). Отличается от RFLP только тем, что для гидролиза используются предварительно амплифицированные фрагменты ДНК, что существенно ускоряет и удешевляет метод. Можно считать более современным вариантом RFLP.

Аллоферменты (или изоферменты). Аллельные формы полипептидов, обладающих определенной ферментативной функцией. В результате точечных мутаций, изменяющих заряд аминокислот, ферменты изменяют свой заряд и мигрируют с разной

скоростью в условиях неденатурирующего электрофореза. По окончании фракционирования гель окрашивают гистохимически для того, чтобы выявить фракции (рис. 12).

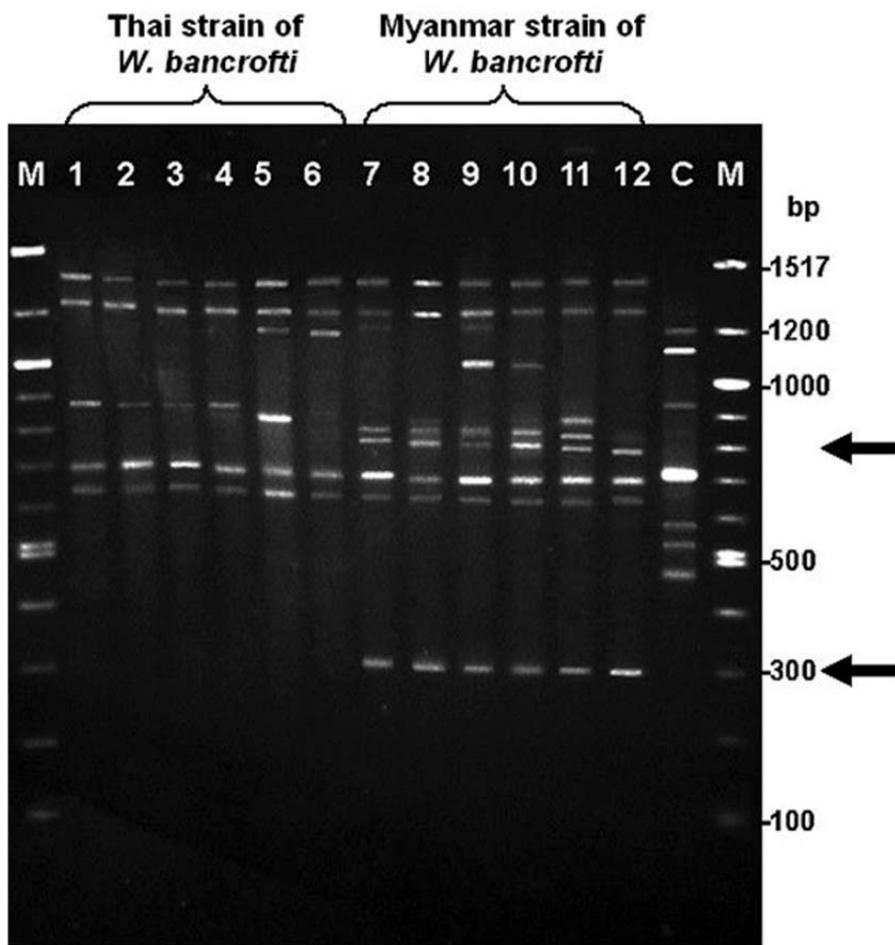


Рис. 11. Электрофореграмма RAPD-маркеров, позволяющих отличить два штамма патогенных нематод из Таиланда и Бирмы [Nuchprayoon, Junree, Poovorawan, 2007]

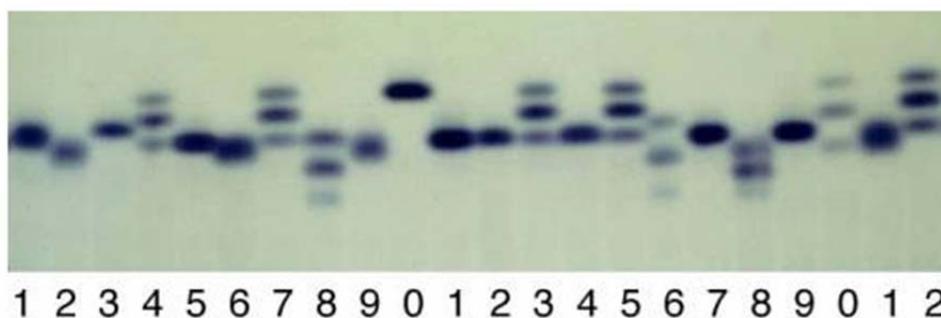


Рис. 12. Крахмальный электрофорез в неденатурирующих условиях с последующим гистохимическим окрашиванием 6-фосфоглюконат дегидрогеназы

2.5. МЕТАГЕНОМНЫЙ АНАЛИЗ

Метагеномный анализ препаратов ДНК, экстрагированной из природных образцов так, как если бы она была получена из одного организма. Это определение было предложено в 1998 г. в работе Handelsman [1998]. Современное определение этой области исследования несколько шире: «Метагеномика – это применение современных методов геномики без выделения и культивирования в лаборатории отдельных видов» [Chen, Pachter, 2005]. Парадоксальность этого метода состоит в том, что он сразу отменяет пропорциональную представленность организмов. Представим себе экосистему, состоящую из всего двух видов, А и В. Пусть в какой-то точке они присутствуют поровну. В результате дифференциального сродства праймеров к ДНК разных видов, это соотношение с большой вероятностью окажется искажено.

Поэтому в различных лабораториях разрабатывают многочисленные методические ухищрения, направленные на то, чтобы придать метагеномному анализу полуколичественный характер. Пока удалось достичь результата только по отношению к малой субъединице рибосомной РНК (SSU rRNA) и только для одноклеточных организмов. Этот недостаток искупается тем, что шансы быть хоть в какой-то мере представленными есть у всех организмов данной экосистемы. Если объектом анализа являются, *например*, почвенные микроорганизмы, то классическими методами (культивирование на средах) можно оценить не более одного-двух процентов от их таксономического разнообразия.

Метагеномика позволяет максимально выявить видовой состав сообщества, с учётом тех организмов, которые при прочих равных остаются незамеченными. *Например*, сине-зелёные водоросли или плесневые грибы редко учитываются при анализе в экосистеме леса.

2.5.1. SHOTGUN-МЕТАГЕНОМНЫЙ АНАЛИЗ

После экстракции и очистки ДНК исследователь может выбрать одну из двух основных стратегий. Простейшей с точки зрения эксперимента и наиболее сложной для дальнейшего анализа является стратегия *shotgun* («шотган»).

Термин *shotgun* часто встречается в современной молекулярной биологии и связанных с ней областях. По происхождению он связан с привычками североамериканских ковбоев во времена широкого распространения там револьверов – стрелкового оружия с вращающимся барабаном. Так называли манеру стрельбы, применявшуюся ковбоями в салунах: они держали спусковой крючок нажатым, а свободной рукой раскручивали барабан, пока не кончались патроны, не забывая при этом размахивать своим оружием. Такой способ исключает прицеливание, но не исключает случайного попадания. Это точно передаёт суть данного метода и поэтому широко используется для обозначения любых неприцельных методов. В применении к метагеномам *shotgun* означает, что всю ДНК фрагментируют примерно до размера около 500 пар нуклеотидов и затем так или иначе определяют нуклеотидные последовательности полученных фрагментов. По сумме этих последовательностей пытаются составить представление о разнообразии геномов метагенома. Это был самый первый вариант метагеномного анализа. Он появился задолго до создания технологий высокопроизводительного секвенирования.

В первоначальной форме *shotgun*-метод выглядел следующим образом: фрагментированную ДНК встраивали в плазмиды – кольцевые молекулы ДНК, способные к самостоятельной репликации в бактериальной клетке. Эти плазмиды вводили в состав бактерий в условиях, когда на одну клетку могла приходиться только одна молекула плазмиды. После этого клетки высевали на среду, позволяющую выживать только клеткам, содержащим плазмиды, и получали колонии, каждая из которых представляла клон – потомство одной молекулы из состава ДНК-образца (отсюда термин «клонирование»). Из отдельных клонов выделяли плазмидную ДНК и определяли нуклеотидные последовательности максимального количества встроек.

Главный недостаток этого подхода состоит в том, что если у исследователя получались, *например*, два фрагмента ДНК, нуклеотидные последовательности не перекрывались, то невозможно было выбрать одну из двух гипотез:

- 1) речь идёт о двух, отдельно расположенных фрагментах одного генома;
- 2) речь идёт о двух геномах разных организмов.

По сравнению с этим уникальные последовательности эволюционируют на 3–4 порядка медленнее даже у самых «быстрых» генов. Простой расчет показывает принципиальное значение быстрой эволюции. Предположим, мы следим за 20 различными независимыми микросателлитными локусами, скорость эволюции каждого из которых составляет 10^{-4} на поколение. Следовательно, вероятность того, что за одно поколение данный локус не мутирует, составит $1 \cdot 10^{-4}$. Поскольку все локусы мутируют независимо друг от друга, вероятность того, что ни один из десяти не мутирует, будет $(1 \cdot 10^{-4})^{10} = 0,99$. Для 20 это будет уже 0,98. Другими словами, вероятность обнаружить мутантов уже в первом поколении будет хоть и весьма мала, но заметна даже при небольшом объеме выборки. Естественно, в ряду поколений процент мутантов будет возрастать.

Вдобавок для каждого из микросателлитных локусов может быть большое количество аллелей с примерно одинаковой жизнеспособностью. Весьма распространены случаи, когда таких аллелей может быть десятков и более.

Предположим, мы анализируем один локус с 10 аллелями, каждый из которых присутствует в панмиктической популяции с одинаковой частотой (1/10). В соответствии с законом Харди – Вайнберга для n аллелей ожидаемая гетерозиготность $H_{exp} = 1 - 1/n = 0,9$, что легко позволяет отслеживать следы инбридинга, однако для наших целей важнее другая характеристика: вероятность каждого из генотипов будет $P = p(1 - p)p(1 - p)$, а так как в нашем примере $p = 0,1$, то $P = 0,0081$. Это вероятность для любого из генотипов по одному-единственному локусу найти такой же. Она очень мала. Если таких локусов 10 и в каждом по 10 аллелей, то вероятность случайно найти себе «близнеца» в идеальном случае составит $8,1 \cdot 10^{-13}$. Сравните с числом людей на Земле. Именно поэтому набор относительно небольшого числа микросателлитных локусов можно считать генетическим отпечатком пальцев, и именно этот метод стандартно используется для идентификации личности. Теми же свойствами анализа микросателлитов определяется и их использование в экологии. Они подходят для анализа миграций и относительно недавних демографических событий.

Экспериментальная сторона микросателлитного анализа

Как уже отмечалось выше, один и тот же элемент может входить в состав большого количества микросателлитов. Для анализа же требуется выделить определенный локус. Это делается с помощью ПЦР, который устроен так, что праймеры расположены по сторонам микросателлитного повтора. В результате амплификации различных аллелей одного микросателлитного локуса и последующего электрофоретического разделения амплифицированных фрагментов ДНК (ампликонов) получаются картины, подобные схематическому изображению на рис. 15.

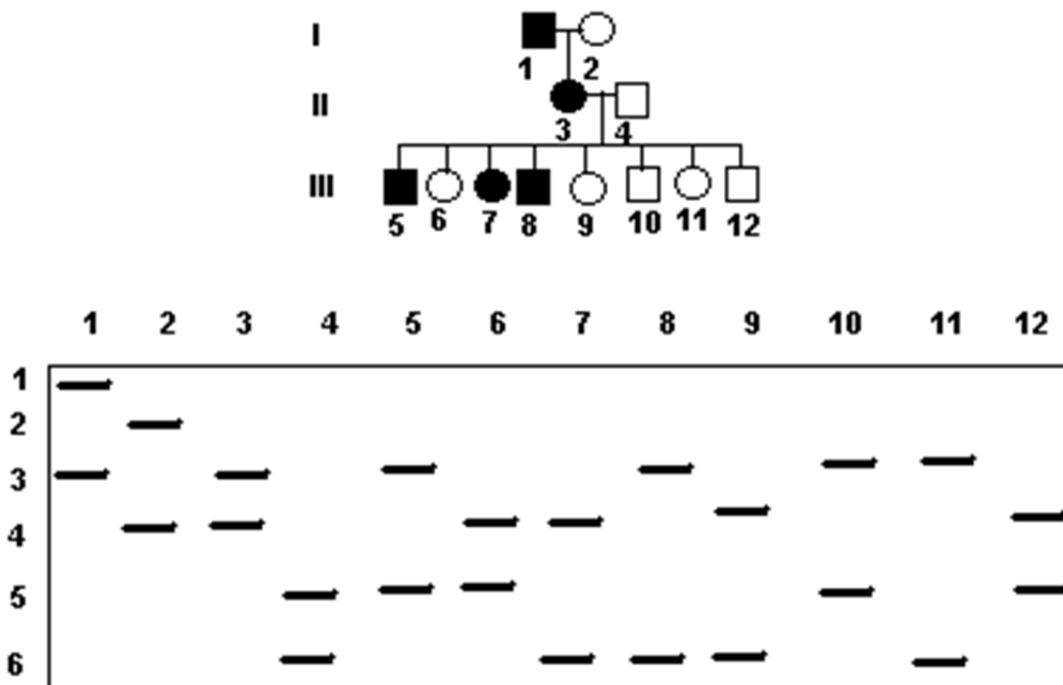


Рис. 15. Пример микросателлитного анализа в пределах 3 поколений одной семьи

Вверху – генеалогия. Условные обозначения:

- – наличие наследственной патологии у ♂;
- – отсутствие наследственной патологии у ♂;
- – наличие наследственной патологии у ♀;
- – отсутствие наследственной патологии у ♀.

I, II, III – поколения. 1, 2, 3... нумерация членов семьи.

Внизу – электрофореграмма, на которой представлены результаты фракционирования микросателлитной ДНК у всех пациентов, в пределах одного локуса. Чем ниже полоска, тем короче микросателлит. В данном случае представлено 6 аллелей одного микросателлита, но генотипы, по данному микросателлиту, совпадают только в двух случаях:

6 и 12, 5 и 10.

Естественная изменчивость длин микросателлитных повторов так высока, что может быть использована для однозначной идентификации личности и выяснения родственных связей. Считается, что для этого достаточно 13 микросателлитных локусов. Микросателлиты также применяются в популяционной генетике для оценки потоков генов между популяциями и выяснения их демографической истории.

Следует заметить, что до самого недавнего времени основной проблемой, сдерживавшей применение микросателлитов, была сложность дизайна новых пар праймеров. Это была долгая процедура с неопределенным исходом, поскольку после успешной амплификации искомого локуса требовалось еще и удостовериться в том, что он достаточно полиморфен в природных популяциях. Оказалось, что это – довольно редкое свойство, поскольку многие из простых повторов оказались защищены стабилизирующим отбором. В среднем же один потенциальный микросателлитный маркер приходится на менее чем 10000 пар нуклеотидов.

В настоящее время для разработки новых микросателлитных маркеров можно воспользоваться библиотеками, получаемыми после полногеномного секвенирования, и специальными программами, предназначенными для этой цели, такими как *microsat explorer*.

2.7. ГИПОТЕЗА МОЛЕКУЛЯРНЫХ ЧАСОВ

Определение времени расхождения генетических линий – один из важнейших этапов филогенетического анализа, нужный для того, чтобы «привязать» эволюционные процессы к определенным этапам развития Земли и определить скорость их развития. В случае если для анализа используют сравнение последовательностей нуклеиновых кислот, оценка скорости эволюционных процессов сводится к оценке средней скорости фиксации замен нуклеотидов или, после построения филогенетического дерева, оценке времени события по сумме длин ветвей, ведущих от узла события к современности.

Вычисление времени эволюционных событий основано на предположении о том, что чем раньше разделились эволюцион-

ные ветви, тем большим числом нуклеотидных замен будут различаться последовательности одного и того же гена у современных потомков далекого общего предка. Это предположение было выдвинуто в начале 60-х гг. XX в. и получило название «гипотезы молекулярных часов». Оно было выдвинуто параллельно двумя группами исследователей на основании обнаруженной ими прямой корреляции между палеонтологическими данными о существовании общего предка и разницы в аминокислотных последовательностях глобинов, выделенных из представителей соответствующих видов.

Последовательность действий при применении гипотезы молекулярных часов такова.

На первом этапе выбирают группу исследуемых видов таким образом, чтобы она включала в себя, помимо интересующих исследователя, несколько видов, об эволюции которых есть информация, полученная другими методами. Это могут быть палеонтологические данные о том, что в слоях определенного возраста предки исследуемых видов уже существовали как отдельные линии (т. е. они разделились до того) или между ареалами видов есть непреодолимый географический барьер, время возникновения которого известно (например, пролив для сухопутных организмов или перешеек – для водных).

1. Выбирают достаточно полиморфную последовательность ДНК, такую, чтобы между интересующими видами было достаточно много (обычно – больше десятка) нуклеотидных замен.

2. Расшифровывают или находят в геномном банке интересующие последовательности ДНК.

3. Строят филогенетическое дерево.

4. Отмечают на этом дереве геологические или палеонтологические события, о которых шла речь в п. 1, и строят калибровочную кривую, на которой по горизонтали отложены точки известных событий, а по вертикали – число нуклеотидных замен, оцененное по филогенетическому дереву. В случае отсутствия таких калибровочных событий используют литературные данные о скорости накопления замен в данном гене у ближайших групп организмов, однако этот подход весьма не надежен.

5. Используют эту кривую как калибровочную, находя на ней времена интересующих событий, используя число нуклеотидных замен между ними, определенное по филогенетическому дереву. Подробный пример содержится в практических заданиях настоящего пособия.

Теоретическое исследование гипотезы молекулярных часов показало, что она может соответствовать действительности в тех случаях, когда большинство мутаций достоверно не сказываются на жизнеспособности их носителей ни в гетеро-, ни в гомозиготном состоянии. Следует напомнить, что под жизнеспособностью понимается изменение вероятности оставить плодовитое потомство.

Скорость фиксации нейтральных мутаций (замещения в популяции одного или нескольких аллелей одним аллелем) зависит от размера популяции, длительности поколения (т. е. времени, проходящего от зиготы до половозрелого состояния) и многих других параметров. Оценки скорости фиксации замен колеблются от $4 \cdot 10^{-11}$ до $15 \cdot 10^{-11}$ на нуклеотид в год. Скорость молекулярных часов может заметно отличаться даже у близкородственных организмов.

В общем случае использование молекулярных часов для определения возраста тех или иных событий представляет собой непростую задачу, которая выходит за рамки настоящего пособия. Вместе с тем относительно несложно (очень приблизительно) оценить время дивергенции видов. Для этой цели используют калибровки молекулярных часов, основанные на датах известных геологических событий или на палеонтологической летописи. *Например*, появление Панамского перешейка, наступление и отступление ледников, появление и исчезновение сухопутного перехода через Берингов пролив («Берингов мост».)

Разные гены эволюционируют с разной скоростью, и поэтому их последовательности пригодны для решения эволюционных проблем разного масштаба. *Например*, один из самых медленно изменяющихся белков – инсулин, у человека и рыбы отличается только одной аминокислотной заменой. Использование этого гена для исследования эволюции позвоночных будет вряд ли продуктивно. С другой стороны, С-пептид инсулина – один из наиболее эволюционно лабильных белков. Поэтому сравнение его последовательностей у близкородственных видов оказалось полезно при решении ряда проблем эволюции и систематики.

В целом ассортимент генов, пригодных для решения эволюционных задач, относительно невелик и различается для разных групп организмов. В случае растений это, обычно, ген, кодирующий 18S рРНК, либо ген *rbcL*, кодирующий в хлоропластном геноме большую субъединицу рибулезо-1,5 дифосфаткарбоксилазы (иногда для обозначения этого гена используется несколько устаревшее сокращение – RUBISCO). Несмотря на широкий разброс скоростей накопления замен в этом гене, он сейчас широко используется для эволюционных датировок. Скорости эволюции этого гена у разных групп растений приведены в работе [Extensive variation ... , 1992]. Эти скорости различаются от $4,4 \cdot 10^{-11}$ на нуклеотид в год в ветви, к которой принадлежит *Marshantia*, $4,4 \cdot 10^{-11}$ для *Arecaceae*, $11,3 \cdot 10^{-11}$ для *Poaceae*, $7,3 \cdot 10^{-11}$ для *Magnoliaceae* и $8,5 \cdot 10^{-11}$ – для остальных растений.

Гипотеза молекулярных часов используется как для исследования видообразовательных, относительно долговременных процессов (от десятков тысяч лет и более), так и для исследования коротких, главным образом – микроэволюционных процессов, занимающих от нескольких десятков до десятков тысяч лет и за редкими исключениями сводящихся к демографическим изменениям и миграциям. Для анализа долговременных процессов под скоростью молекулярной эволюции понимают среднюю **скорость полной фиксации одного из аллелей**, а для исследования коротких процессов – **среднюю скорость изменения частот (концентраций) аллелей в результате генетического дрейфа**, которые различаются принципиально.

В популяции соблюдаются все требования, необходимые для соблюдения закона Харди – Вайнберга, за исключением одного – популяции имеют ограниченный размер. Очевидно, что, как обсуждалось выше, в маленьких популяциях размах колебаний гораздо больше. В некоторых случаях кривые концентраций касаются верхней или нижней границы. Касание границы (см. рис 5, А) означает, что частота одного из аллелей стала равна нулю, т. е. он исчез из популяции. Соответственно, частота другого стала равна 1 (или, если выражать в процентах – 100 %). Именно это явление и называется полной фиксацией. Средний промежуток времени, необходимый для полной фиксации вновь

возникшего в результате мутации аллеля, и называется **скоростью молекулярной эволюции**.

Обычно длительность среднего промежутка времени между появлением и полной фиксацией мы можем определить только очень приблизительно. Причины этого состоят в том, что, во-первых, неизвестно, какое из альтернативных состояний окажется поглощено в ряду поколений. Более того, редкие мутации, не приносящие заметной пользы, имеют большие шансы быстро исчезнуть в ряду поколений. Во-вторых, вероятность значительных случайных колебаний частот аллелей в достаточно больших (тысячи особей) популяциях относительно мала. Именно эти обстоятельства объясняют очень низкую достоверность многих оценок времени эволюционных событий, получаемых с помощью молекулярных часов.

К счастью, для прикладных исследований биоразнообразия более важно другое использование молекулярных часов. Его проще проиллюстрировать с помощью рис. 5, С. В больших популяциях отклонения от ожидаемой в соответствии с законом Харди – Вайнберга равновесной частоты аллелей накапливаются со временем, но полные поглощения происходят гораздо реже. На приведенном результате моделирования таких явлений не отмечается совсем (см. рис. 5, С). Такая картина характерна для внутрипопуляционных процессов, протекающих за относительно короткие промежутки времени.

В течение короткого времени ожидать нескольких случаев полного поглощения трудно из-за их редкости. Поэтому для оценок времен относительно быстрых процессов используется средняя скорость накопления отклонений от равновесных частот, она же – скорость генетического дрейфа V_t . Она определяется формулой

$$V_t \approx pq \left(1 - \exp \left(-\frac{t}{2N_e} \right) \right),$$

где p и q – начальные частоты альтернативных аллелей, t – время, измеряемое в поколениях, а N_e – эффективный размер популяции. Именно скорость дрейфа используется для оценки скорости процессов, например, в популяциях промысловых рыб или при оценке эффективности или планировании адресных природоохранных мероприятий.

Прикладное значение скорости генетического дрейфа обусловлено тем, что его величину можно оценить, используя данные о генетическом разнообразии объекта исследования. Главной сложностью таких оценок оказывается неидеальность природных популяций, состоящая в том, что они, как правило, занимают большие площади и вероятность скрещивания между далеко обитающими друг от друга особями существенно меньше, чем между соседями, что означает нарушение условий Харди – Вайнберга. Это явление описывается так называемой скоростью перемешивания, которую также можно измерить, исходя из данных по генетическому разнообразию.

Частоты аллелей можно измерить напрямую. Существует множество изящных экспериментальных приемов для получения точных оценок, они либо рассматриваются в предыдущих параграфах, либо – в рамках смежных предметов. Сложнее дело обстоит с измерением эффективной численности, однако и ее можно оценить. Осталось оценить только либо скорость, либо – время. Приоритет в этом определяется конкретной биологической задачей. *Например*, если исследователя интересует, насколько опасен для экосистемы недавно появившийся экзотический вид, то, как правило, время его появления известно и можно определить, насколько быстро идут популяционные процессы.

С другой стороны, о времени начала каких-либо событий можно судить, сравнивая свои данные по генетическому разнообразию с хорошо изученными аналогичными примерами, для которых скорость дрейфа надежно определена. *Например*, для определения возраста происхождения цветковых растений были использованы нуклеотидные последовательности 83 генов, расположенных как в хлоропластном и митохондриальном геномах, так и в ядре. В основном эти гены кодируют ДНК, но среди них есть и гены рибосомных РНК различной локализации. Всего в анализ взято 644 вида, из которых 632 принадлежит к цветковым растениям. После проведения филогенетического анализа, в результате которого было получено филогенетическое дерево, объединившее все виды и основанное на последовательностях всех генов, на этом дереве отметили все точки, для которых есть палеонтологические данные (рис. 16). (Обычно палеонтологические данные со-

стоят в том, что выше некоторого датированного пласта осадков встречаются представители двух групп организмов, а ниже – они отсутствуют, а присутствуют некие примитивные общие предки, такие оценки часто не точны и противоречат друг другу). Среди палеонтологически датированных узлов были найдены 43, возраста которых не противоречили друг другу. Это означает, что степень генетических различий между всеми генами пропорциональна их палеонтологическим возрастам. Например, разделение *Fagus* и *Quercus* произошло после 48 млн лет. Средняя оценка для всех использованных генов составила 0,04 замены на нуклеотид за 100 млн лет или $4 \cdot 10^{-10}$ замены на нуклеотид за год.

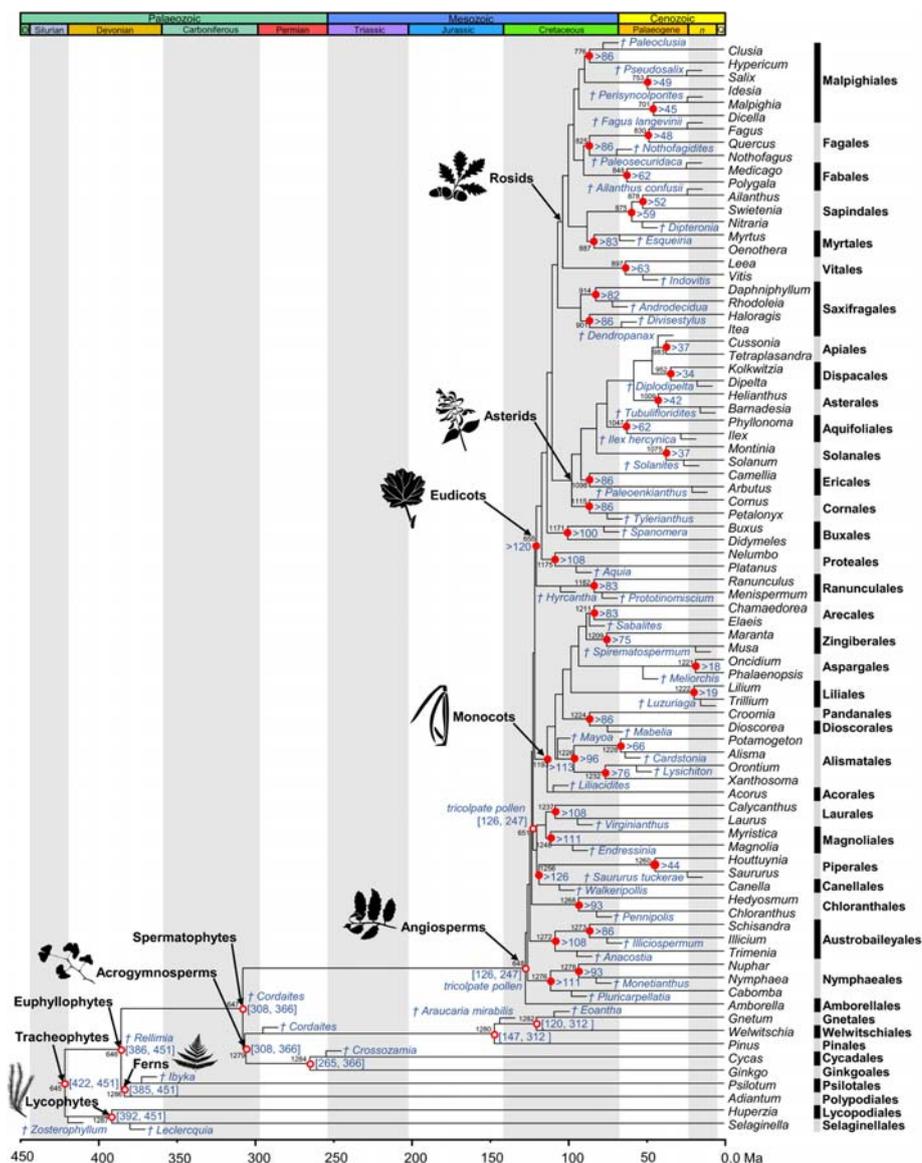


Рис. 16. Филогенетическое дерево покрытосеменных растений [Barba-Montoya, 2017]

Зная скорость накопления замен V и число замен на 100 нуклеотидов N , можно рассчитать время t , за которое накопилось такое количество, по формуле

$$t = N / V.$$

В результате этого установлено, что у *crown angiosperms* на 100 нуклеотидов приходится от 8,4 до 10,2 нуклеотидных замен. То есть для 10,2 замен на 100 нуклеотидов получится: $10,2/100/4 \cdot 10^{-10} = 255\ 106$ лет. Таким образом было определено, что *crown angiosperms* произошли 255–206 млн л. н., двудольные произошли 186–156 млн л. н., а время происхождения однодольных приходится на интервал 179–144 млн л. н.

2.8. АНАЛИЗ ДЕМОГРАФИЧЕСКИХ ПРОЦЕССОВ И МИГРАЦИИ

Миграция между популяциями одного вида часто представляет значительный интерес. С точки зрения генетики значение имеет только та миграция, которая оставляет следы в генофонде популяций, обменивающихся особями. Естественно, бывает и миграция, при которой скрещивание между представителями разных популяций пренебрежимо мало. Тогда изучение миграционных процессов сводится лишь к идентификации образцов. В этом разделе рассмотрены те миграционные процессы, которые влияют на частоты аллелей. В генетике популяций они называются потоками генов.

С количественной стороны миграцию между двумя группами особей можно охарактеризовать с помощью **коэффициента миграции (net coefficient of migration) m** . Эта величина определяется как разница между числом прибывших и убывших индивидуумов за год на 1000 особей. При исследовании биоразнообразия мы не можем оперировать такими точными статистическими данными. Поэтому для оценки миграции используется величина, подобная коэффициенту миграции, но не абсолютная численность особей, а эффективная, т. е. соотносящаяся с реальной так же, как эффективная численность популяции N_e соотносится с физической численностью популяции.

На качественном уровне представляется очевидным, что двусторонняя миграция между двумя популяциями, различающимися по частотам аллелей, должна приводить к выравниванию этих частот, а за счет притока новых особей – к уменьшению коэффициента инбридинга I . Последствия миграции можно поэтому описать количественно с помощью меры генетической дифференциации F_{ST} . Эта величина равна доле, которая приходится на различия между представителями разных популяций по сравнению с внутривидовыми различиями. Для измерения степени различия используются, как правило, молекулярные маркеры – последовательности нуклеиновых кислот или различия между микросателлитами. Ее вычисляют многие компьютерные программы, и измерена она для многих случаев. Внутри одного клона $F_{ST} = 0$, для абсолютно различных групп особей $F_{ST} = 1$, а в подавляющем большинстве случаев $F_{ST} < 0,25$. Связь между F_{ST} и коэффициентом миграции описывается уравнением

$$F_{ST} = \frac{1}{4N_e m + 1},$$

которое позволяет, используя данные о генетическом разнообразии, оценить число мигрантов:

$$m = (1 - F_{ST})/4F_{ST}N_e.$$

Эта величина представляет большой интерес, однако следует помнить, что она может серьезно искажать действительную картину миграционных потоков из-за того, что предполагает их симметрию.

В настоящее время существует несколько программ, которые оценивают асимметричные потоки генов, но они используют весьма сложный математический аппарат, и мы не будем рассматривать их использование.

В отличие от демографии населения, для оценок в популяционной генетике генетическая оценка коэффициента миграции относится не к 1000 физических особей, а к 1000 членов «эффективной популяции», поэтому его следует умножить на отношение реальной численности к эффективной, тогда эта величина окажется гораздо больше. Об этом важно помнить при работе с материалами, например, природоохранных органов, полагающихся на методы оценки биоразнообразия, основанные на популяционной генетике.

Все приложения генетики к исследованиям демографических процессов основаны на сравнении количественных характеристик генетического разнообразия. С этой целью анализируют несколько генетических маркеров (чем больше, тем лучше) и используют *D*-критерий Таджимы (Tajima's *D*).

В постоянно растущей популяции чем больше становится особей, тем больше вероятность возникновения новых мутаций. Следовательно, возрастает доля недавно возникших мутаций, у которых возможности для широкого распространения не было. Следовательно, наборы нуклеотидных последовательностей, полученных из такой популяции, будут содержать много позиций, где только одна последовательность содержит замену (такие позиции называются *синглетами*).

В вымирающей популяции наблюдается противоположная картина. Её особи прошли или проходят через снижение численности особей, которое могло быть резким и длительным. Такие популяции почти не содержат синглетов, в них увеличена доля аллелей, отстоящих друг от друга на несколько мутационных шагов, но промежуточные звенья исчезают в результате дрейфа.

Для дифференциации этих сценариев развития популяций используется критерий Таджимы, построенный на использовании двух величин: гаплотипического разнообразия *S* и нуклеотидного разнообразия π .

1. **Гаплотипическое разнообразие *S*** определяется долей переменных положений

Например, если в 100 нуклеотидных последовательностей есть 1 положение, в котором есть нуклеотидные замены, то $S = 0,01$, если положение аденина и гуанина будет 1 из 10 или 5 из 10, то это рассматривается как 1 вариант замены.

2. **Нуклеотидное разнообразие** определяется долей переменных положений и количеством замен в каждом положении, которое рассчитывают по формуле

$$\pi = \sum_{ij} x_i x_j \pi_{ij} = 2 \sum_{i=2}^n \sum_{j=1}^{i-1} x_i x_j \pi_{ij},$$

где n – общее число последовательностей. Варианты последовательностей: i и j , x_j – доля последовательностей типа i от общего

числа последовательностей, x_j – доля последовательностей типа j от общего числа последовательностей, d_{ij} – доля нуклеотидов, по которым эти варианты отличаются.

Например, если положение аденина и гуанина будет 1 из 10 или 5 из 10, а все остальные нуклеотиды – одинаковые, то:

в первом случае $\pi = 2 \cdot 1/10 \cdot 9/10/20 \cdot 1 = 0,009$;

а во втором $\pi = 2 \cdot 5/10 \cdot 5/10 \cdot 1/20 = 0,025$.

Второй набор данных оказался разнообразнее.

Когда популяция, численность которой была постоянна, переходит к росту или к вымиранию, соотношение гаплотипического разнообразия S и нуклеотидного разнообразия π будет разным. При увеличении численности особей в популяции гаплотипическое разнообразие S растёт быстрее, чем нуклеотидное, а при вымирании популяции оно быстрее сокращается.

Критерий D Таджимы можно рассчитать с помощью программы dnaSP [<http://www.ub.edu/dnasp/>] и др. При использовании критерия D Таджимы популяция в стабильном состоянии выглядит сходной с популяцией в состоянии регресса (табл. 4). Сложности при интерпретации признаков расширяющихся и вымирающих популяций заставляют обычно использовать несколько разных критериев, значения которых вычисляют из тех же данных те же программы.

Таблица 4

Интерпретация критерия Таджимы

| Значение D | Вывод 1 | Вывод 2 |
|--------------|--|--|
| $D = 0$ | Наблюдаемое генетическое разнообразие равно ожидаемому | Популяция имеет постоянный размер, нет свидетельств давления отбора |
| $D > 0$ | Дефицит редких аллелей (синглетов) | Популяция подверглась сильному вымиранию или оказалась под действием балансирующего отбора |
| $D < 0$ | Избыток синглетов | Увеличение численности популяции, последствия фиксации выигрышного аллеля |

Пример недавнего бутылочного горлышка

На севере Шотландии при анализе генетического полиморфизма сосны было обнаружено неожиданно **высокое генетическое разнообразие** $\pi_{tot} = 0,0078$, сравнимое и даже превышающее разнообразие континентальных популяций того же вида, **так как** северо-шотландские сосны находятся на самом краю ареала. Значение критерия Таджимы оказалось высоким ($D = 0,58$), в отличие от всех других европейских популяций, в частности в Испании ($D = -0,539$). То есть ранее условия среды соответствовали центральным областям ареала для этих сосен и были близки к оптимальным, но относительно недавно они стали неблагоприятными. В результате большая часть популяции вымерла.

Пример расширяющейся популяции

У чаек Эрмана *Larus heermanni*, обитающих в Калифорнийском заливе, на фоне **нуклеотидного разнообразия** ($\pi_{tot} = 0,00083$) значение критерия Таджимы отрицательно ($D = -2,467$). Следовательно, численность их популяции в настоящее время и на протяжении большей части четвертичного периода должна была быстро расти [Ruiz, Velarde, Aguilar, 2017].

Практикум

Занятие 1

АНАЛИЗ БИОРАЗНООБРАЗИЯ НА ОСНОВАНИИ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ДНК

План работы

1. Определить ген на основании заданной последовательности (выбранной из прил. 1, 2) при помощи GENBANK [https://www.ncbi.nlm.nih.gov/] (рис. П1–П5).

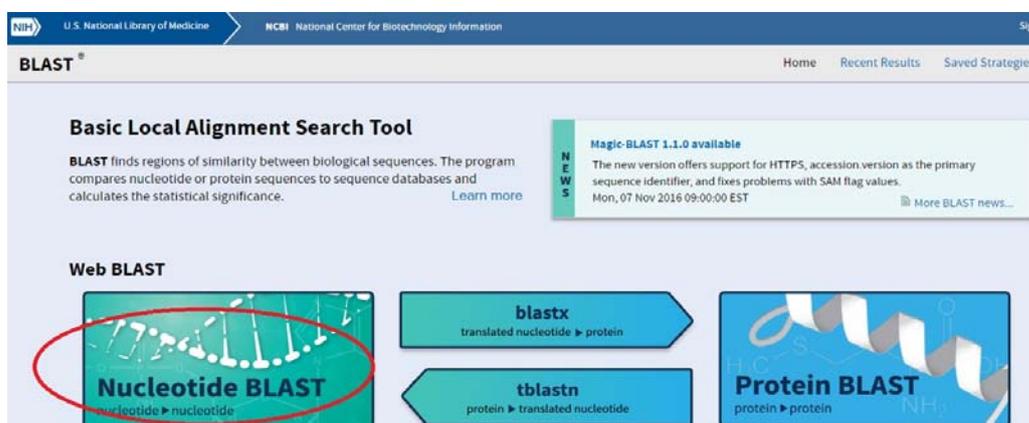


Рис. П1

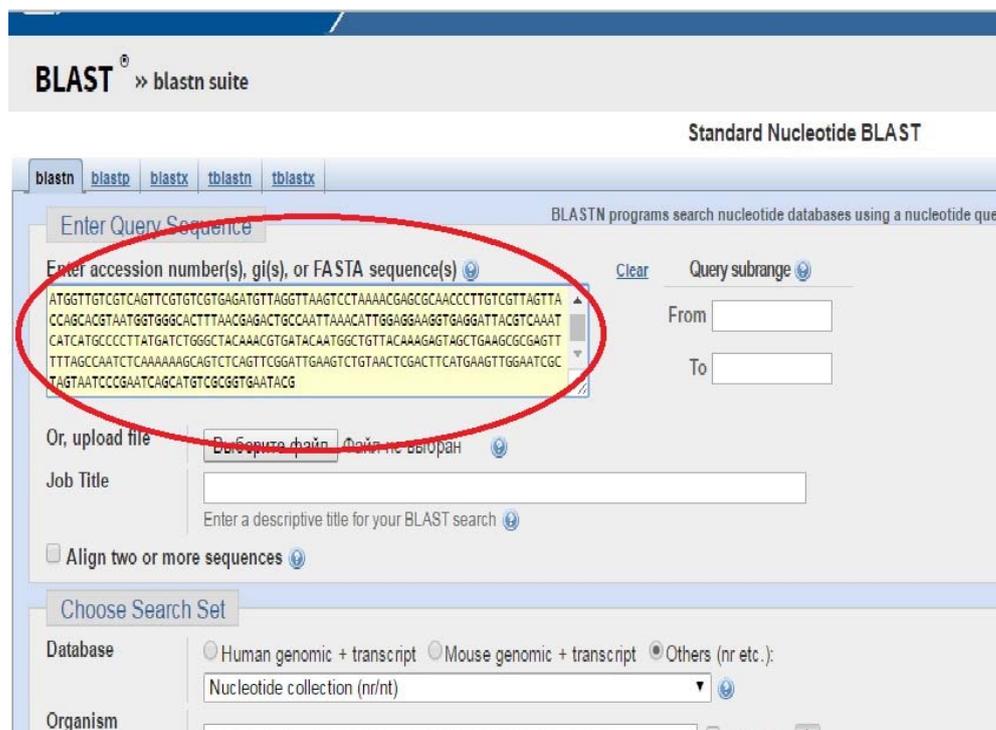


Рис. П2

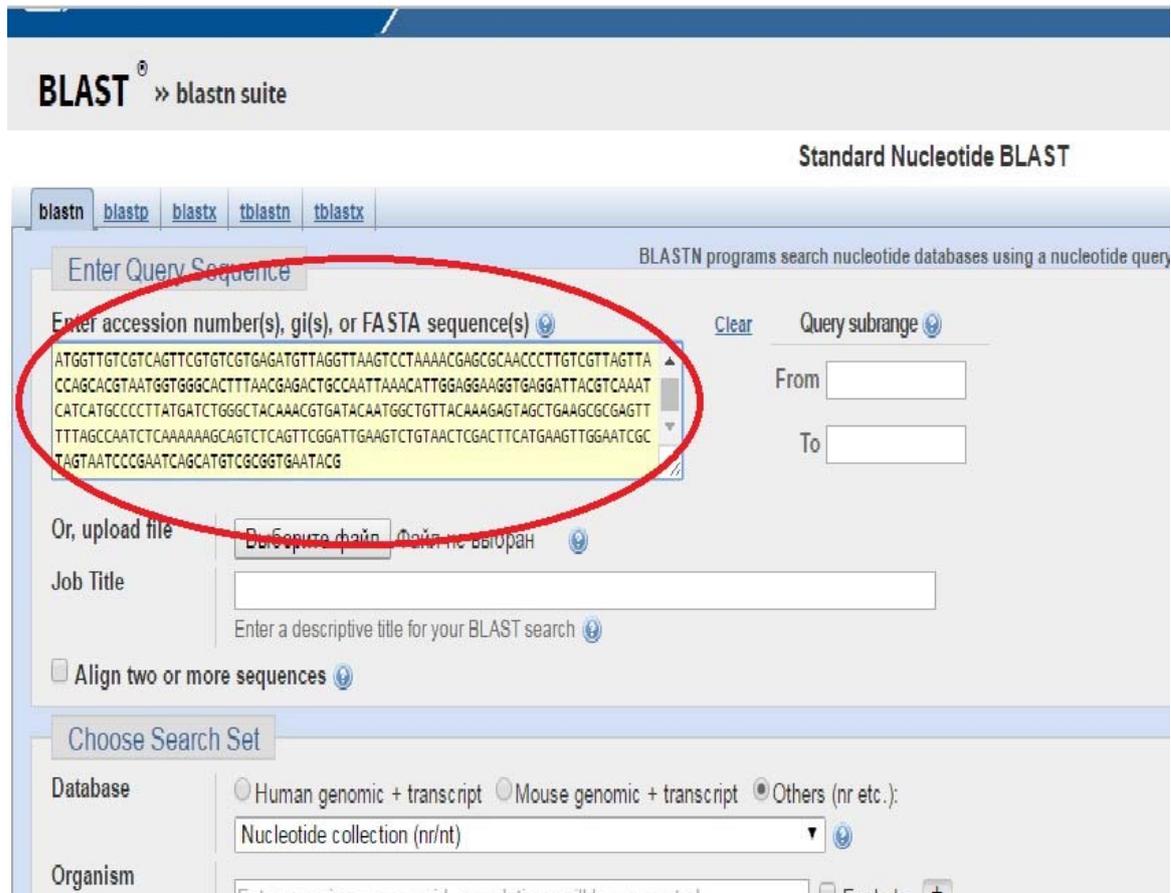


Рис. П3

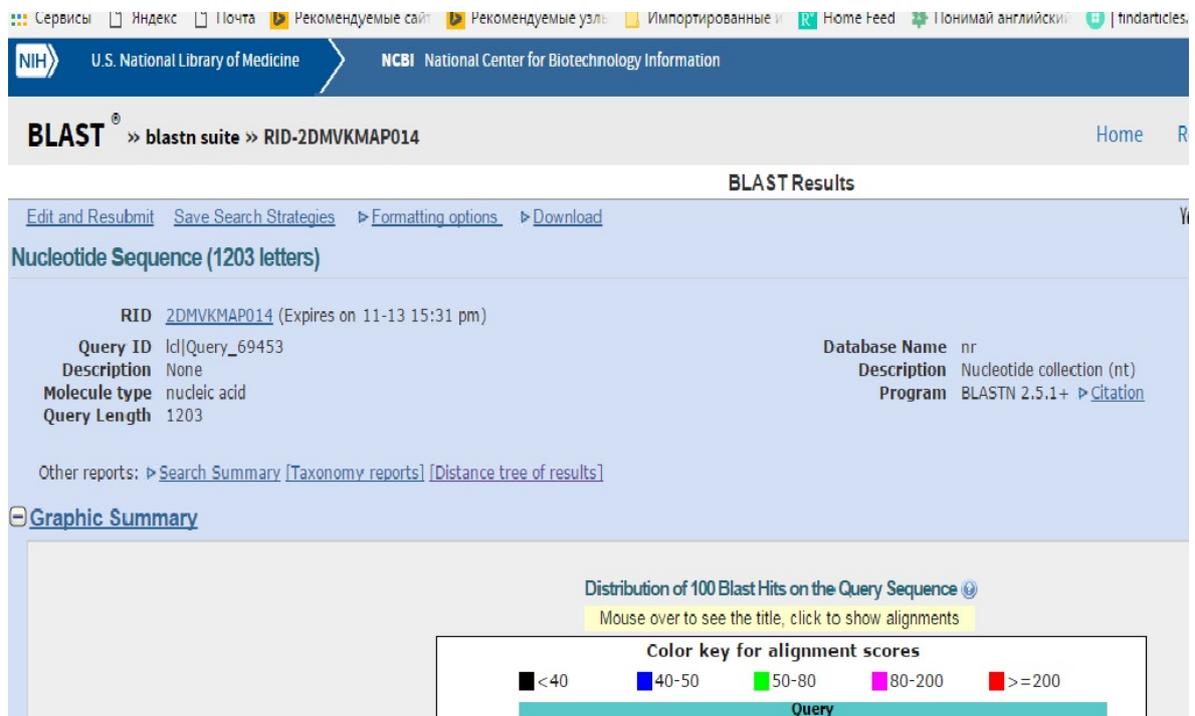


Рис. П4

'Magnolia grandiflora' phytoplasma strain Magn1 16S ribosomal RNA gene, partial sequence
Sequence ID: [KJ093444.1](#) Length: 1203 Number of Matches: 1

Range 1: 1 to 1203 GenBank Graphics

▼ Next Match ▲ Previous Match

| Score | Expect | Identities | Gaps | Strand |
|-----------------|---|-----------------|------------|-----------|
| 2222 bits(1203) | 0.0 | 1203/1203(100%) | 0/1203(0%) | Plus/Plus |
| Query 1 | AGCTGGATAGGAATAAAAAGGCATCTTTTATTTAAAAGACCTTTTTCGGAAGGTATGCTT | 60 | | |
| Sbjct 1 | AGCTGGATAGGAATAAAAAGGCATCTTTTATTTAAAAGACCTTTTTCGGAAGGTATGCTT | 60 | | |
| Query 61 | AAAGATGAGCTTGGCCACATTAGTTTGTGGTGGGGTAATGGCCTACCAAGACGATGAT | 120 | | |
| Sbjct 61 | AAAGATGAGCTTGGCCACATTAGTTTGTGGTGGGGTAATGGCCTACCAAGACGATGAT | 120 | | |
| Query 121 | GTGTAGCTGGACTGAGAGGTCGAACAGCCACATTGGGACTGAGACACGGCCCAAACCTCT | 180 | | |
| Sbjct 121 | GTGTAGCTGGACTGAGAGGTCGAACAGCCACATTGGGACTGAGACACGGCCCAAACCTCT | 180 | | |
| Query 181 | ACGGGAGGCAGCAGTAGGGAATTTTCGGCAATGGAGGAACTCTGACCGAGCAACGCCGC | 240 | | |
| Sbjct 181 | ACGGGAGGCAGCAGTAGGGAATTTTCGGCAATGGAGGAACTCTGACCGAGCAACGCCGC | 240 | | |
| Query 241 | GTGAACGATGAAGTATTTTCGGTATGTAAGTTCTTTTATTGAAGAAGAAAAATAGTGGA | 300 | | |
| Sbjct 241 | GTGAACGATGAAGTATTTTCGGTATGTAAGTTCTTTTATTGAAGAAGAAAAATAGTGGA | 300 | | |
| Query 301 | AAAATATCTTGACGTTATTCAATGAATAAGCCCCGGCTAACTATGTGCCAGCAGCCGCG | 360 | | |
| Sbjct 301 | AAAATATCTTGACGTTATTCAATGAATAAGCCCCGGCTAACTATGTGCCAGCAGCCGCG | 360 | | |
| Query 361 | GTAAGACATAGGGGGCGAGCGTTATCCGGAAATTATTGGGCGTAAAGGGTGCGTAGGCGGT | 420 | | |
| Sbjct 361 | GTAAGACATAGGGGGCGAGCGTTATCCGGAAATTATTGGGCGTAAAGGGTGCGTAGGCGGT | 420 | | |
| Query 421 | TAGGAAAGTCTATAATTTAATTTTCAGTGCTTAACGCTGTCTTGTATAGAAACTACCTTG | 480 | | |
| Sbjct 421 | TAGGAAAGTCTATAATTTAATTTTCAGTGCTTAACGCTGTCTTGTATAGAAACTACCTTG | 480 | | |
| Query 481 | ACTAGAGTTAGATAGAGGCAAGCGGAATTCATGTGTAGCGGTAAAATGTGTAATATAT | 540 | | |

Рис. П5

2. При помощи программы BLAST (<https://blast.ncbi.nlm.nih.gov/>) найти группу наиболее похожих последовательностей 20–30 штук.

3. Скачать эти последовательности в форматах FASTA и GENBANK.

4. Обнаружить максимально близкую последовательность и определить видовую принадлежность организма, из которого она была получена.

5. Проанализировать таксономическое положение и близость данного таксона при помощи кладограммы (рис. П6).

6. Сохранить результаты проделанной работы в электронном виде.

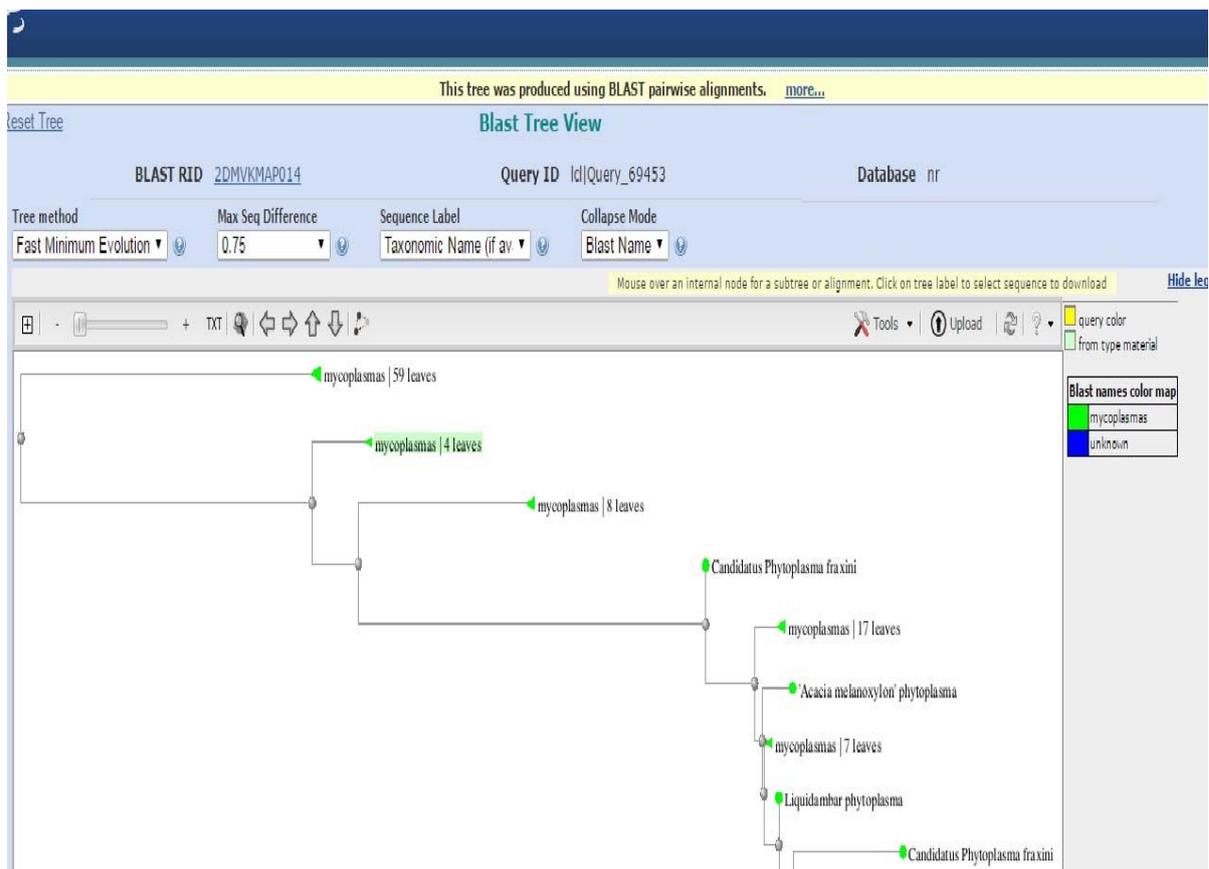


Рис. П6

Контрольные вопросы

1. Монотипическая концепция вида, суть и недостатки.
2. Политипическая концепция вида.
3. Биологическая концепция вида.
4. Какая таксономическая единица реально существует в природе: тип, вид, род, класс, семейство? Почему?

Занятие 2

МУТАЦИИ КАК ЭЛЕМЕНТАРНЫЙ ЭВОЛЮЦИОННЫЙ МАТЕРИАЛ, РЕГУЛИРУЮЩИЙ БИОРАЗНООБРАЗИЕ

План работы

1. Ознакомиться с файлом в формате GENBANK.

При помощи сайта <http://www.tolweb.org/tree> охарактеризовать систематическое положение на древе жизни.

2. Выровнять последовательности при помощи программы Clustal X (<http://www.clustal.org/omega/>).

2.1. Алгоритм выравнивания последовательностей.

Собрать последовательности таксонов (от 7 до 30 шт.) при помощи приложения «Блокнот» (рис. П7).

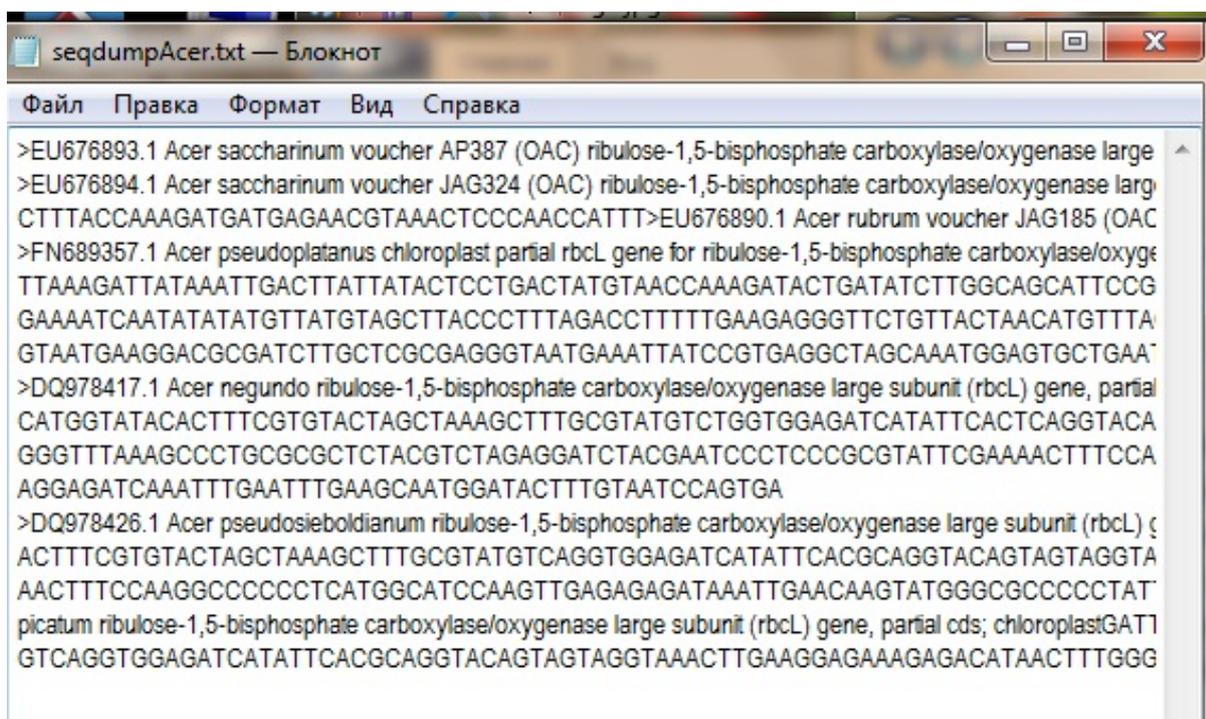


Рис. П7

3. Отредактировать список последовательностей таким образом, чтобы было ясно название таксона. Оставить название рода и вида в формате *Drosophila viridis* (пробелов быть не должно, идентификатор должен быть короче 35 символов) (рис. П8).

4. Открыть файл с последовательностями при помощи программы Clustal X2.1 (рис. П9 и П10).

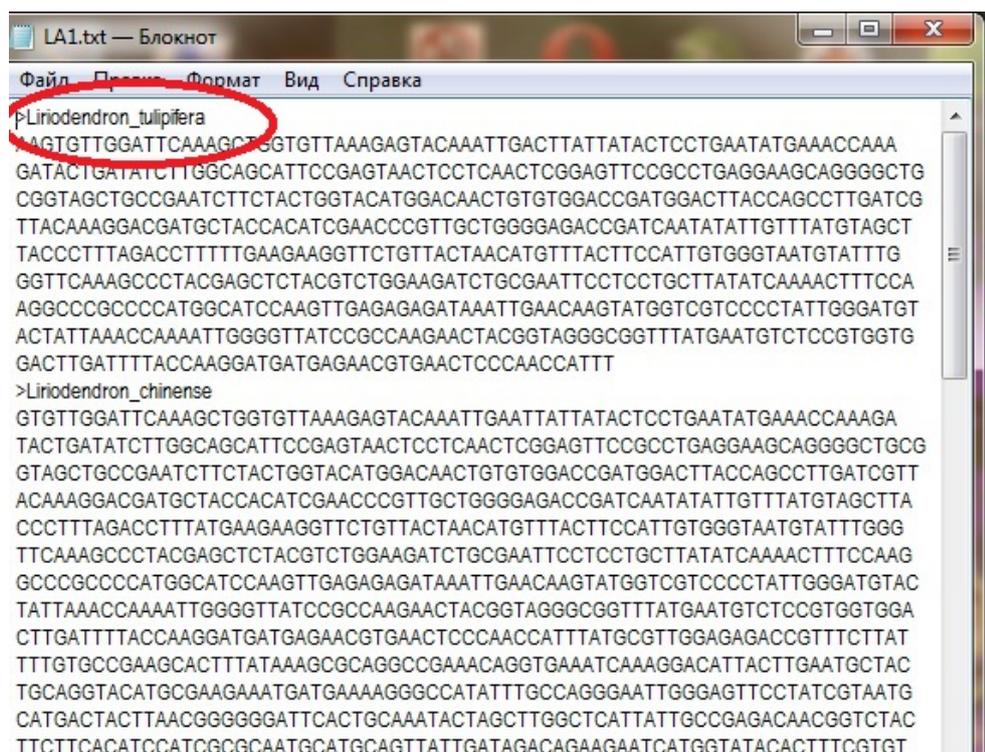


Рис. П8

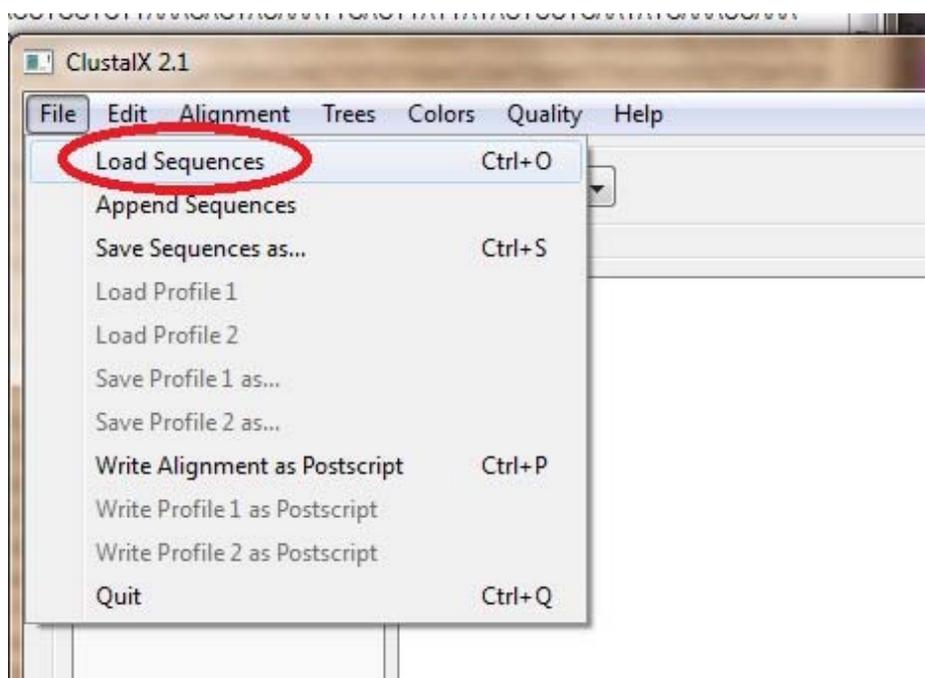


Рис. П9

5. Выровнять последовательности в программе Clustal X2.1 и сохранить их с расширением aln, т. е. в родном формате этой программы (рис. П11 и П12).

6. Сохранить результаты проделанной работы в электронном виде.

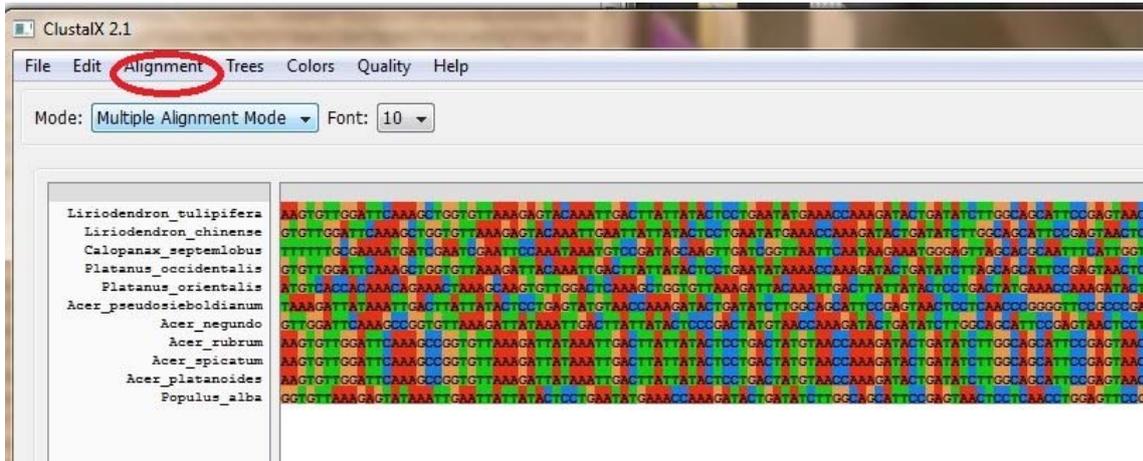


Рис. П10

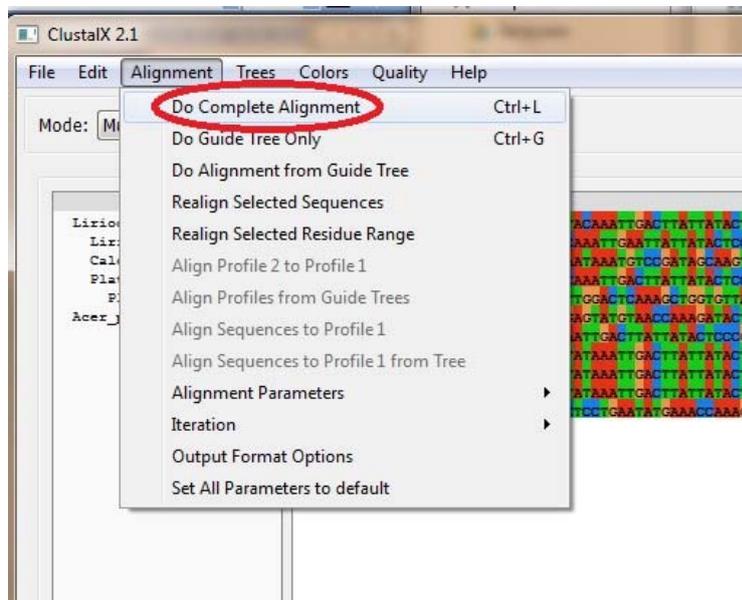


Рис. П11

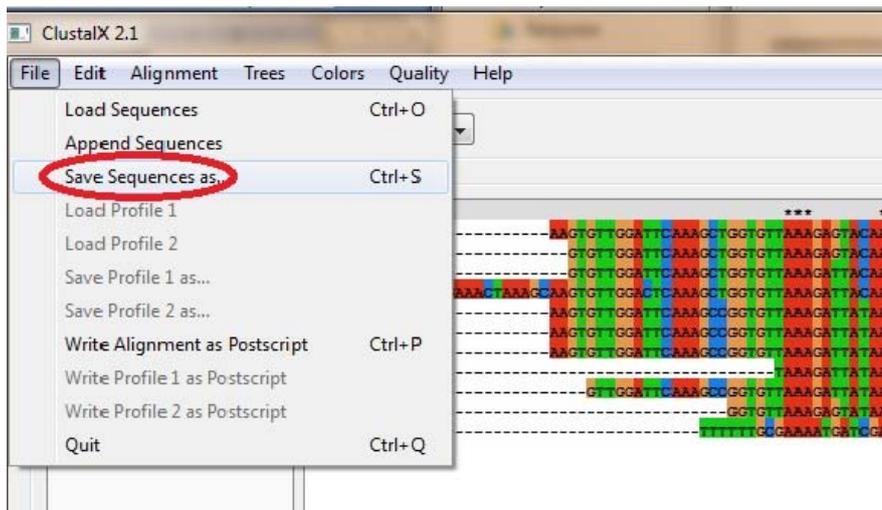


Рис. П12

Контрольные вопросы

1. Мутационная изменчивость.
2. Типы мутаций.
3. Общие свойства мутации.
4. Распространение мутаций в пределах ареала вида.
5. Эволюционное значение мутаций.
6. Закон Харди – Вайнберга, условия его соблюдения и отклонений от него.
7. Методы выявления генетического разнообразия.
8. К какому типу изменчивости относятся мутации? Проаргументировать выбранный вариант.
 - а) наследственная изменчивость
 - б) средовая изменчивость
 - в) модификационная изменчивость
 - г) онтогенетическая изменчивость
 - д) возрастная изменчивость

Занятие 3

АНАЛИЗ ГАПЛОТИПИЧЕСКОГО И НУКЛЕОТИДНОГО РАЗНООБРАЗИЯ

План работы

1. Построить филогенетическое дерево при помощи программ Phym1 и Seaview (рис. П13–П15).

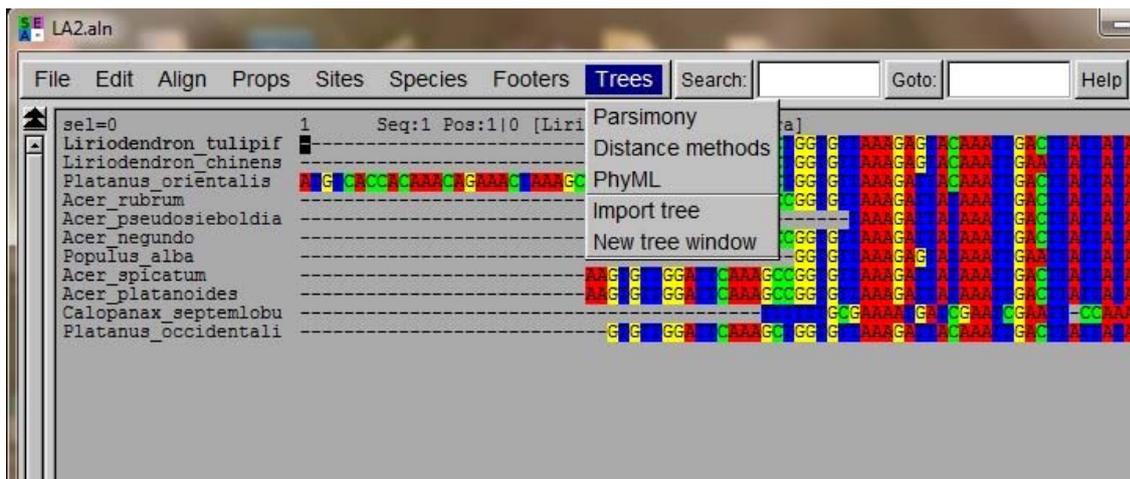


Рис. П13

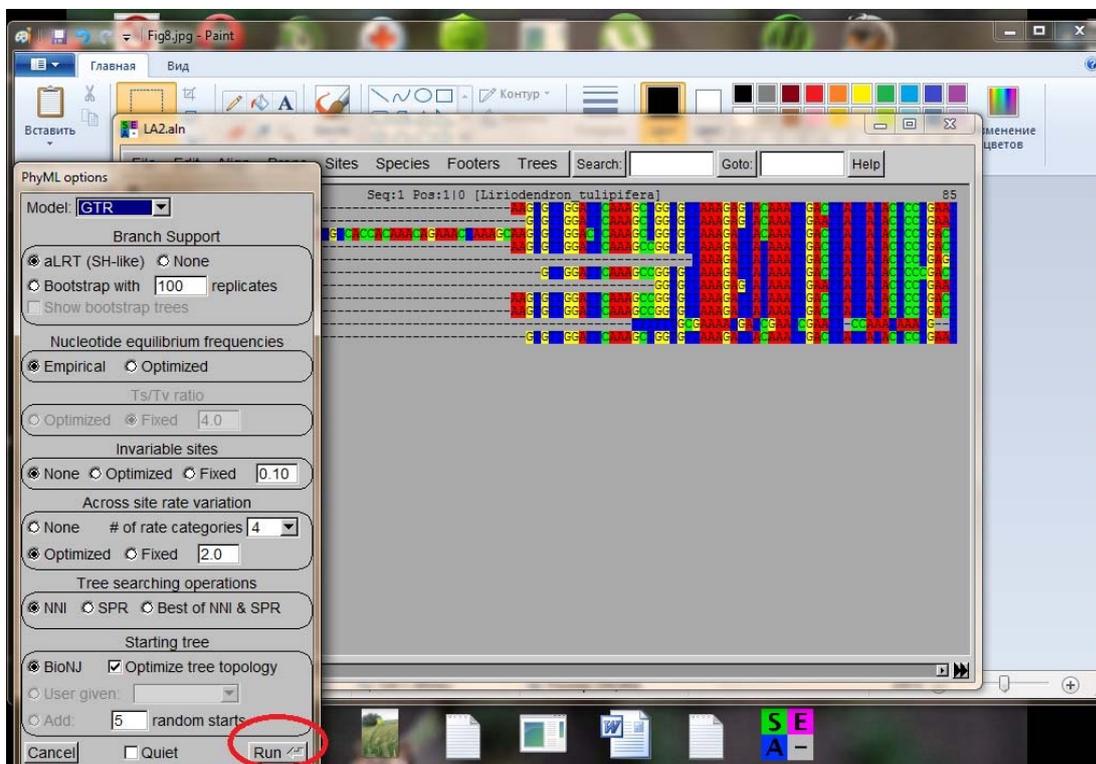


Рис. П14

PhyML ln(L)=-3799.9 2269 sites GTR 4 rate classes

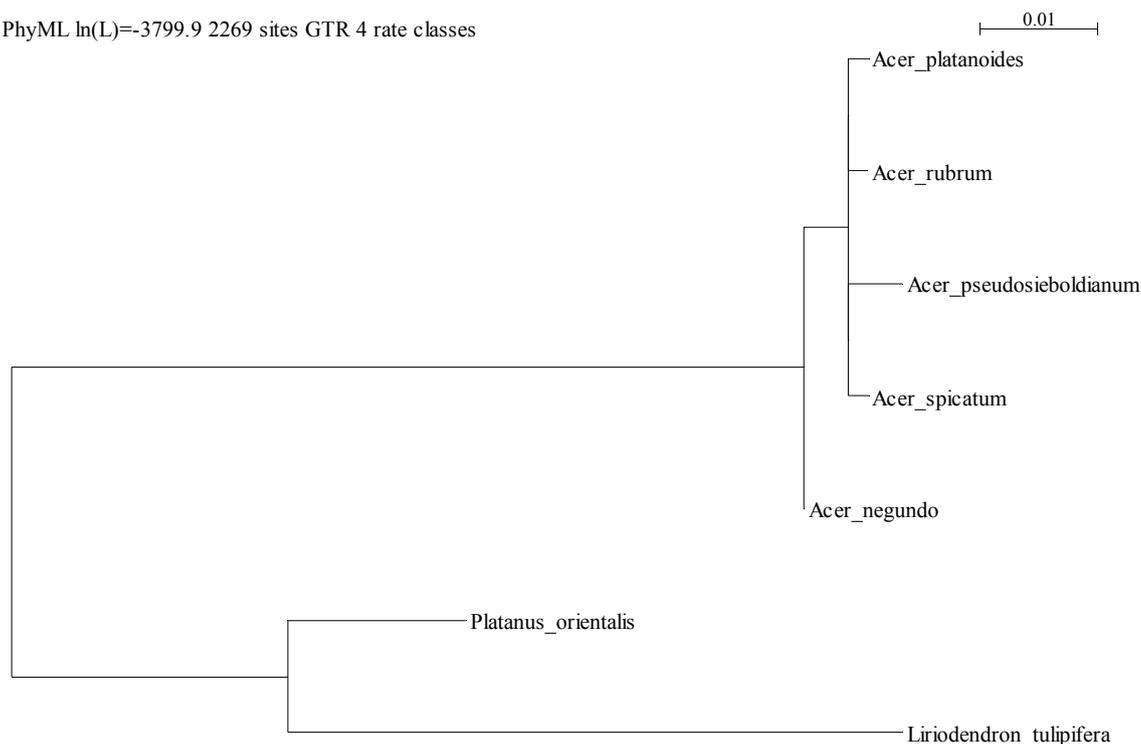


Рис. П15

2. Охарактеризовать вариабельность нуклеотидных последовательностей с помощью программы DnaSP (<http://www.ub.edu/dnasp/DnaSPHelp.pdf>), используя следующие показатели:

- число полиморфных сайтов – характеризующее размах изменчивости в рамках набора данных;
- общее число мутаций;
- G+C содержание – характеризует адаптацию;
- число гаплотипов;
- разнообразие гаплотипов;
- изменчивость гаплотипического разнообразия.

Нуклеотидное разнообразие (per site), P_i : 0.16897¹

Стандартное отклонение P_i : 0.06062

Среднее число нуклеотидных различий, k : 73.67143

Theta (в последовательности) от Eta: 138.14229

Theta (на сайт) от Eta: 0.31684

Тесты на нейтральность:

D-критерий Таджимы

D: -1.93250

¹ Здесь и далее точки в качестве десятичного разделителя используются в распечатках компьютерных программ и листингах команд.

Статистическая значимость: *, $P < 0.05$
Fu и Li's D* статистический тест: -1.66441
 $P > 0.10$
Fu и Li's F* статистический тест: -2.04466
Fu's Fs статистика: 0.234

При трактовке результатов вычислений следует принять во внимание, что при значениях нуклеотидного разнообразия $\rho_i > 0.05$ (как в рассматриваемом примере), можно утверждать, что набор последовательностей принадлежит видам, которые давно репродуктивно изолированы, поэтому следующие нижеупомянутые тесты на нейтральность не имеют существенного значения. Тем не менее то обстоятельство, что все они подтверждают гипотезу нейтральности, указывает на то, что не удалось обнаружить следов действия движущего отбора на исследуемый ген в пределах рассматриваемого набора видов. Однако это не исключает, что действие движущего отбора не удастся обнаружить с помощью других методов, специально предназначенных для анализа белок-кодирующих нуклеотидных последовательностей.

Соотношение нуклеотидного и гаплотипического разнообразия в условиях отсутствия отбора даёт представление о том, в какой демографической фазе находится рассматриваемая группа организмов. Если гаплотипическое разнообразие существенно больше нуклеотидного, то исследователи имеют дело с процветающей, активно эволюционирующей группой (растущей). Если же имеет место обратная ситуация, то следует предполагать, что исследуемая группа угнетена и представляет собой остаток бывшего разнообразия (т. е. регрессирует).

Остальные три теста проверяют наличие следов отбора в наборе данных и применимы, главным образом, для анализа образцов в пределах одной популяции.

То есть эти критерии применяются для проверки, наличия избытка редких мутаций (у расширяющейся популяции) или их недостатка (у вымирающей). Отрицательные значения во всех случаях указывают на избыток редких мутаций, а положительные – на недостаток.

Критерии Таджимы D и F_u and L_i 's D^* представляют собой разные математические способы оценки достоверности отличий имеющегося соотношения гаплотипической изменчивости и нуклеотидного разнообразия от ожидаемого. Они применимы для популяции, при условии, что исследуемые гены нейтральны, а численность – постоянна. Разница между тестами состоит в их меньшей или большей чувствительности к определенным популяционным сценариям (см. рис. П1 и П16, П17).

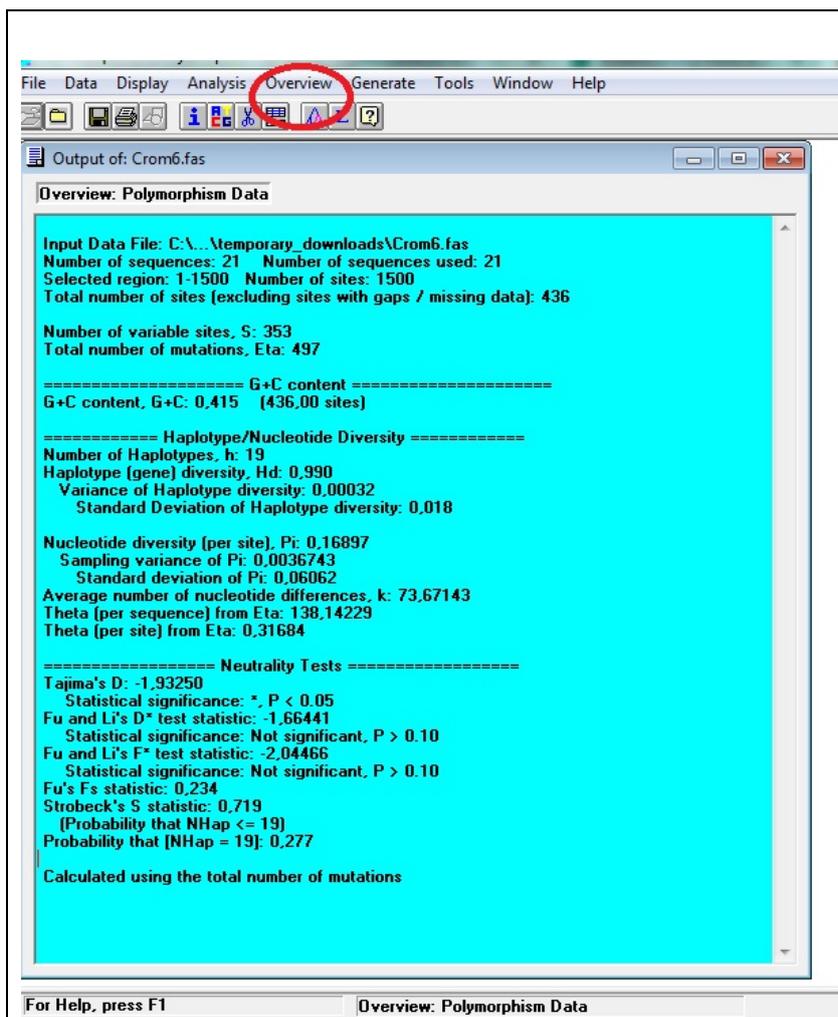


Рис. П16

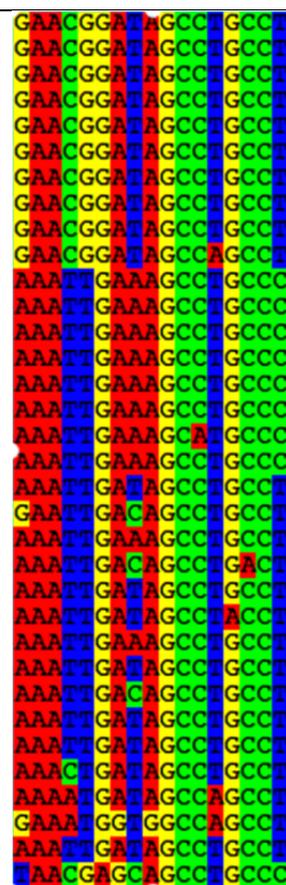


Рис. П17. Пример информативного положения последовательностей.

Столбцы 6, 11, 13 и 14 – содержат редкие мутации, первый столбец – пример информативной замены

3. Сохранить результаты проделанной работы в электронном виде.

Контрольные вопросы

1. Постепенное видообразование.
2. Мгновенное видообразование.
3. Эволюционная роль межвидовой гибридизации. Сетчатое видообразование.
4. Методы выявления механизмов видообразования.
5. Формы безреальной внутривидовой изменчивости.
6. Основные молекулярно-генетические методы анализа биоразнообразия.
7. Метод штрихкодирования.
8. Метагеномный анализ, суть и область применения
9. Микросателлитный анализ, суть и область применения.

ЗАНЯТИЕ 4

АНАЛИЗ РЕЗУЛЬТАТОВ ФИЛОГЕНЕТИЧЕСКИХ ПОСТРОЕНИЙ И МОЛЕКУЛЯРНЫЕ ЧАСЫ

План работы

1. Охарактеризовать филогенетические взаимоотношения последовательностей при помощи программ Phyml и Seaview на основании дерева, построенного на прошлом занятии, и оценить достоверность топологии (порядка ветвления) при помощи апостериорных вероятностей (рис. П18).

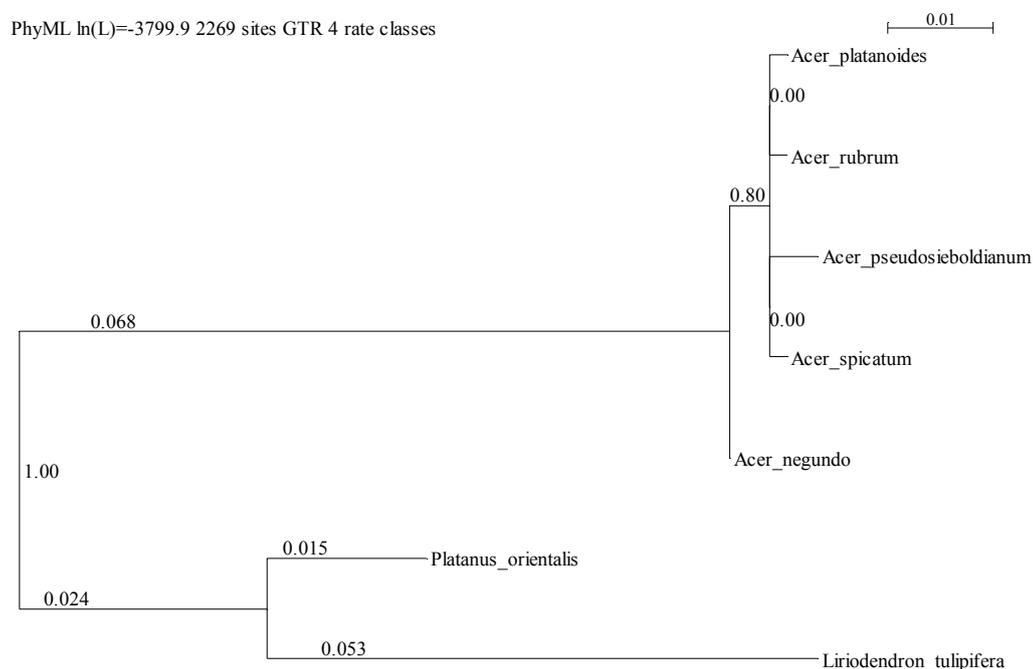


Рис. П18

Апостериорные вероятности на дереве ставятся на ветви, которая ведет от внутреннего узла к общему предку. Потомки этого узла называются сестринскими видами. Математический смысл апостериорной вероятности по сути состоит в том, что два потомка данного узла действительно являются сестринскими. Хорошей статистической поддержкой считаются апостериорные вероятности более 0.7. Максимальное их значение – 1 (рис. П19).

Seaview. Включен просмотр апостериорных вероятностей, значения: < 0.5 указывают на слабую поддержку ветвей, < 0.8 – указывают на умеренную поддержку ветвей.

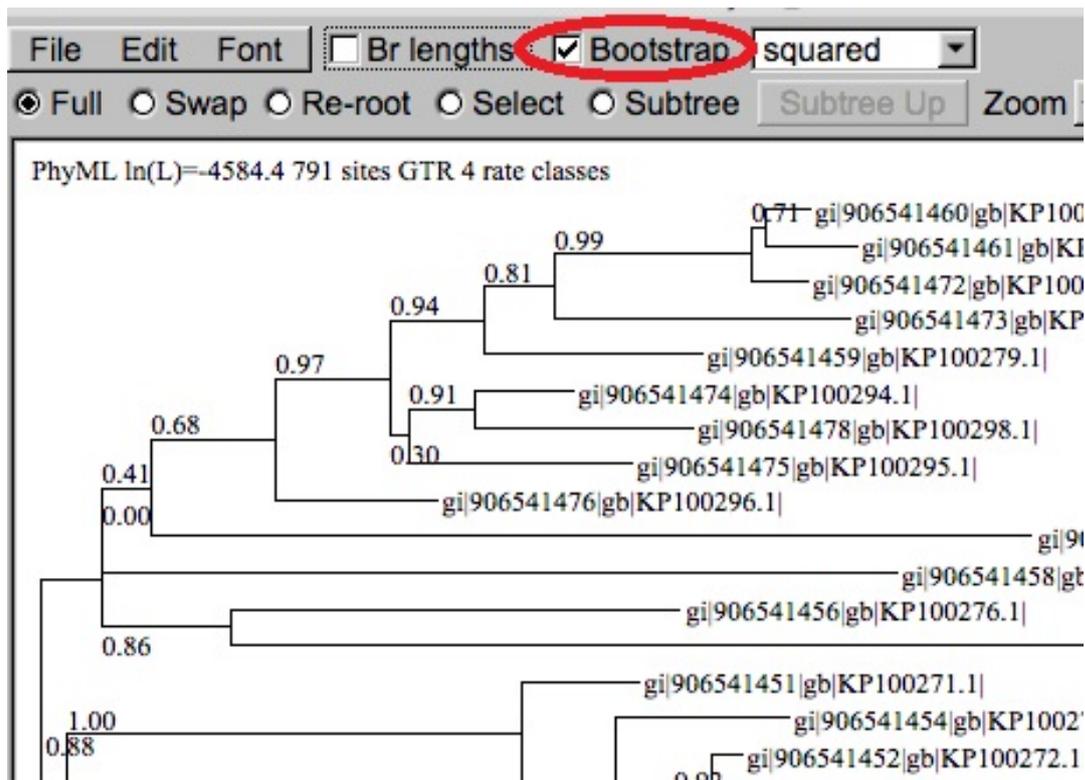


Рис. П19. Фрагмент филогенетического дерева в окне просмотра деревьев

2. Рассмотрим гипотетический пример, поясняющий использование молекулярных часов. Предположим, что мы расшифровали фрагмент ДНК длиной 100 нуклеотидов у гена *rbcL* из растений, время существования общего предка которых нам хочется определить. В статье Буске с соавторами [Extensive variation in evolutionary ... , 1992] находим, что для этих растений скорость накопления замен составляет $5 \cdot 10^{-11}$ замен на 1 нуклеотид в год. Оказывается, что фрагменты ДНК отличаются всего на один нуклеотид. Это означает, что расстояние (различие) между этими фрагментами, выраженное в нуклеотидных заменах, составит 1/100 замены (1 замена на 100 нуклеотидов). Для того чтобы определить время после расхождения последовательностей, разделим это расстояние на то, что накапливается за 1 год.

$$\text{Время} = 10^{-2} / 5 \cdot 10^{-11} = 0.2 \cdot 10^9 = 2 \cdot 10^8 \text{ лет.}$$

Но этот срок следует разделить пополам, поскольку после разделения различия накапливались в двух ветвях параллельно и расстояние от современной последовательности до общего предка должно быть половиной от расстояния между двумя современными последовательностями.

Современные программы построения филогенетических деревьев выражают мутационные расстояния в «долях замен». С этой целью можно использовать программу Phylml.

На рис. П21, П22 над ветвями помещены числа. Они означают долю мутаций между двумя узлами, т. е. число замен, деленное на длину последовательности. При этом расстояние указывается от узла до узла, поэтому делить на два, как в предыдущем примере, его не следует.

Пример определения возраста по дереву

- Строим дерево с помощью Phylml.
- Сохраняем дерево в Seaview командой Save rooted tree.
- Закрываем Seaview и открываем Mesquite (рис. П20).

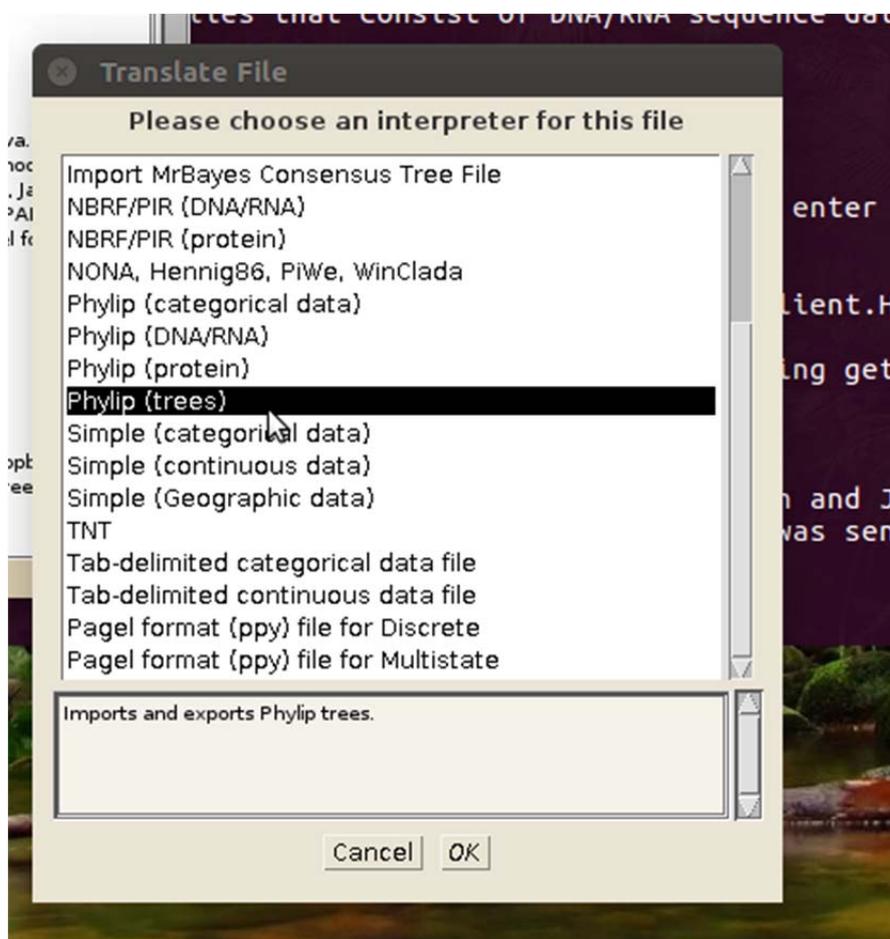


Рис. П20

Открыть файл с деревом в ньюикском формате (см. рис. П21):
Taxa&Trees → New tree window → With trees from source

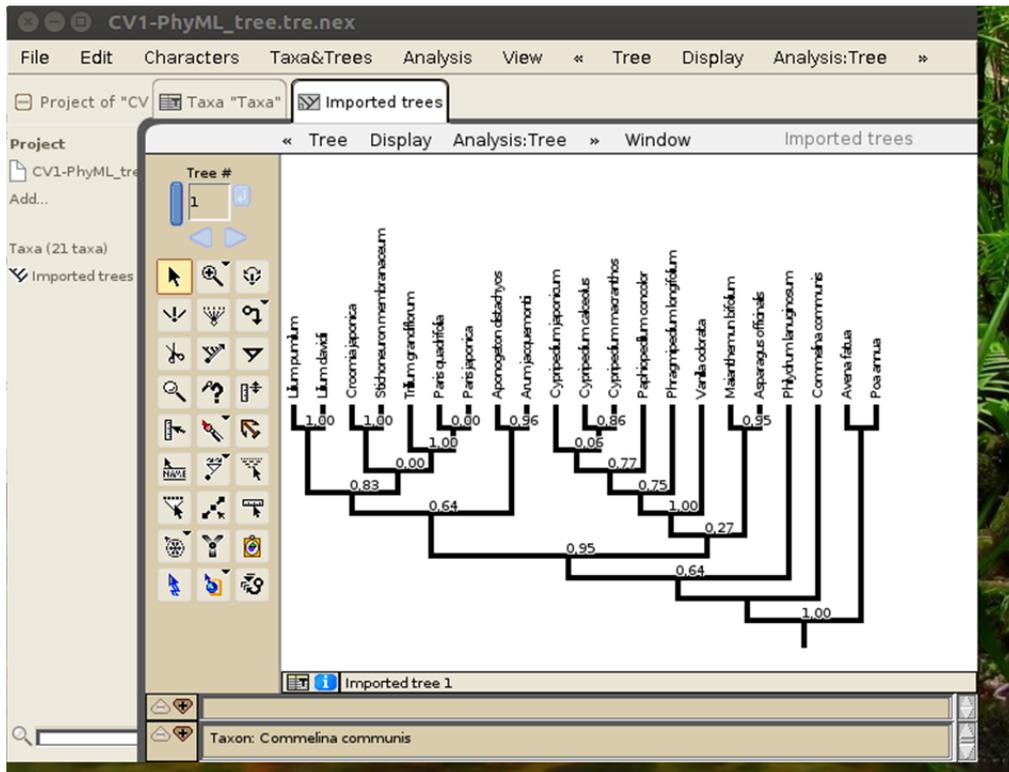


Рис. П21

- Display – поворачиваем дерево направо (orientation->right).
 - Отмечаем галочкой Display → Branch length labels.
- Появляются синие числа, означающие долю замен до узла (см. рис. П22).

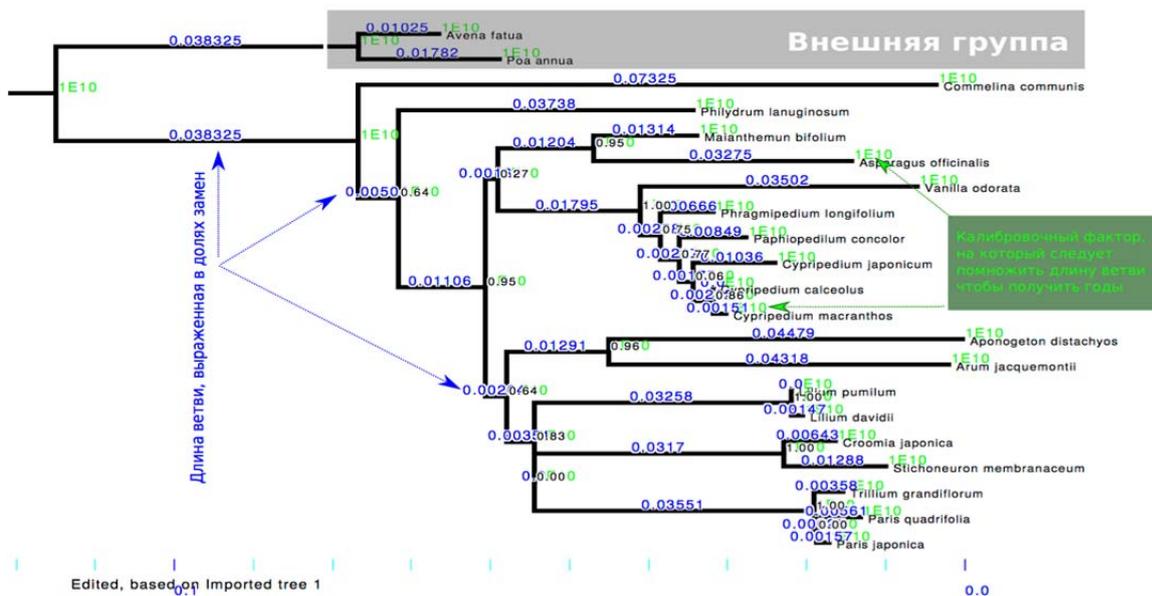


Рис. П22

На длину секвенированного фрагмента ДНК расстояние делить уже не надо. Обратите внимание, что ветви, ведущие от современных видов к их общим предкам, разной длины. Это может быть следствием различий в скорости эволюции и/или артефактом метода построения дерева.

- Отмечаем *Branches proportional to length* в меню *Display*.

Получаем дерево и синие длины ветвей множим на калибровочный фактор.

Формула для расчета возраста:

$$T = S/v,$$

где T – время в годах, от момента существования общего предка; v – скорость молекулярной эволюции, т. е. сколько в среднем замен на нуклеотид фиксируется в год; S – длина ветви, выраженная в долях замен (она может быть дробной из-за поправок).

Например, расстояние до общего предка с ближайшим сестринским видом у *Cypripedium macranthos* составляет 0.00151. Применяем скорость молекулярной эволюции $2 \cdot 10^{-10}$, пригодную для орхидей, и получаем оценку длины этой ветви примерно 7,5 млн лет. Это – очень приблизительная оценка, ошибка которой может составить порядок величины. Необходимо также помнить о приблизительном характере определения времен по деревьям с ветвями столь неравной длины.

В данном случае выбор скорости эволюции был сделан произвольно по литературным данным. Поскольку обосновывать выбор точной скорости довольно сложно, в последнее время широкое распространение получило относительное описание возраста, например «5 % от возраста расхождения однодольных и двудольных».

3. Построить дерево, учитывающее географию распространения таксонов при помощи программы *Mesquite*:

<http://mesquiteproject.org/>;

<https://mesquiteproject.wikispaces.com/Additional+Mesquite+Packages>;

<http://mesquiteproject.org/packages/cartographer/>.

4. В электронном виде сохранить результаты проделанной работы.

Контрольные вопросы

1. Факторы, лимитирующие распространение вида.
2. Классификация изолирующих механизмов.
3. Генетика изолирующих механизмов. Последствия вторичных нарушений репродуктивных барьеров между сестринскими видами. Принцип конкурентного исключения.
4. Нарушения изолирующих механизмов.
5. Методы выявления изолирующих механизмов.

Занятие 5

ПРЕДСТАВЛЕНИЯ РЕЗУЛЬТАТОВ АНАЛИЗА ПОСЛЕДОВАТЕЛЬНОСТЕЙ ДНК

План работы

1. При помощи программы Seaview представить дерево в разных форматах (циркулярное, радиальное и пр.) (рис. П23 и П24).

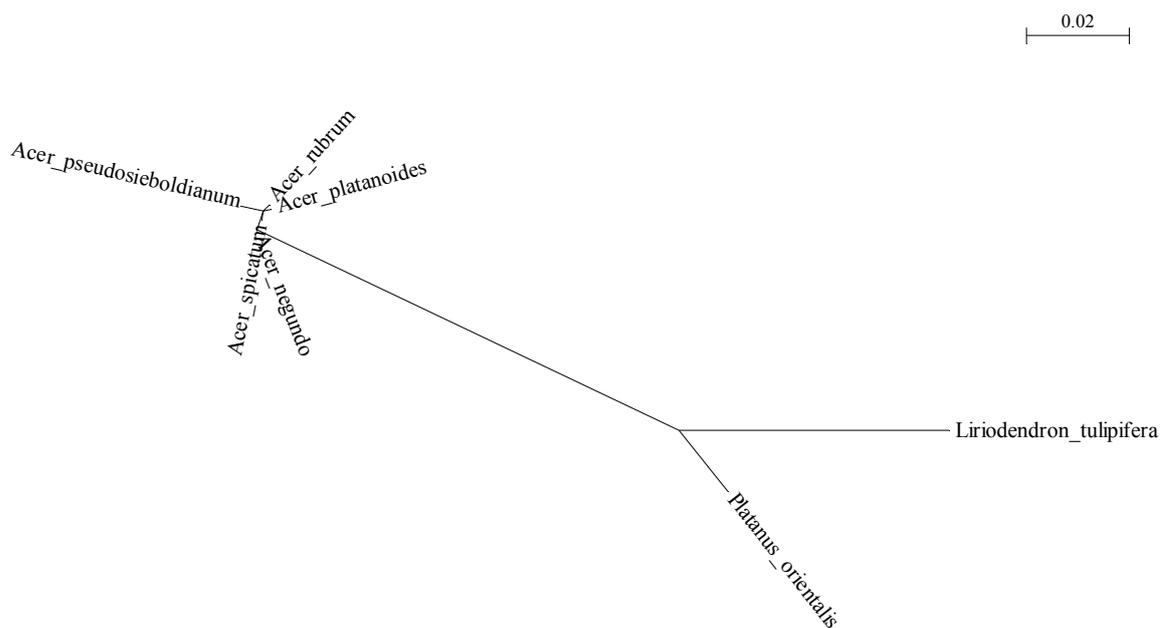


Рис. П23

2. Ознакомьтесь с методикой составления ньюикской нотации деревьев.

Ньюикская нотация предназначена для описания дерева с помощью текстовой строки. Для представления результатов анализа последовательностей ДНК можно использовать линейные скобочные записи деревьев, т. е. представление деревьев в виде строк, содержащих символы, помечающие узлы дерева, а также открывающие и закрывающие круглые скобки. Линейная запись деревьев может быть представлена при помощи скобок.

Например, предположим, что у нас есть четыре вида: s1, s2, s3 и s4. Степень их родства изображают при помощи скобок: $((s1,s2),s3),s4$; или ветвей дерева (рис. П25).

PhyML ln(L)=-3799.9 2269 sites GTR 4 rate classes

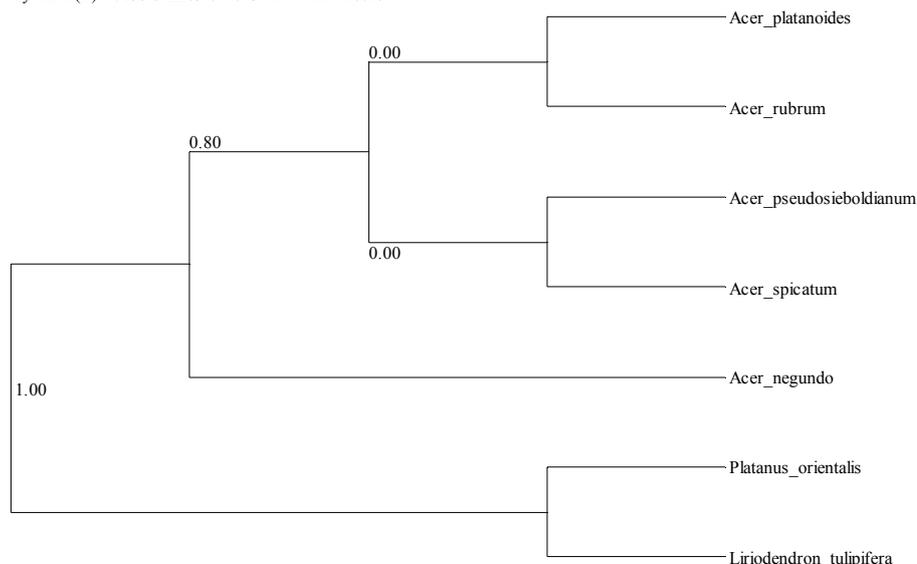


Рис. П24

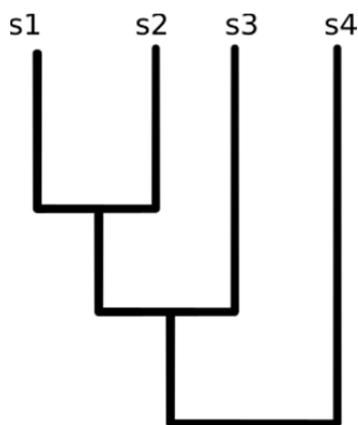


Рис. П25

Если требуется указать длину ветвей, то ее приводят при помощи двоеточия:

$((s1:0.1,s2:0.12):0.3,s3:0.5):0.6,s4:0.7):0.0;$

То есть самые близкие друг другу виды будут заключены в максимальное количество скобок, а самый внешний из них (s4) – будет закрыт только одной скобкой. Сколько скобок открыто, столько должно быть и закрыто. В конце записи ставят точку с запятой.

Расстояния после закрывающих скобок – это длины внутренних ветвей. К этой записи можно добавить еще и величины статистической поддержки, например апостериорные вероятности. Их помещают без знаков препинания сразу за закрывающими скобками:

(((s1:0.1,s2:0.12)0.8:0.3,s3:0.5)0.9:0.6,s4:0.7)0.5:0.0;

При этом в названиях видов не рекомендуется использовать знаки препинания, пробелы и прочие небуквенно-цифровые символы.

3. Составить таблицу, в которую включить информацию о последовательностях, взятых из генбанка (табл. П1).

Таблица П1

Информация о последовательностях

| № | Название вида | Код последовательности в генбанке | Авторы последовательности | Публикация и электронная ссылка на неё |
|---|---------------|-----------------------------------|---------------------------|--|
| 1 | | | | |
| 2 | | | | |

4. В электронном виде сохранить результаты проделанной работы.

Контрольные вопросы

1. Причины географической изменчивости.
2. Значение географической изменчивости.
3. Причины экологической изменчивости.
4. Значение экологической изменчивости.
5. Методы анализа экотопов.
6. Что понимают под «ньюикской нотацией»? Привести пример.
7. Популяционная структура вида.
8. Изменение генотипического состава популяций как элементарное эволюционное изменение.
9. Биологические свойства вида.
10. Дивергенция и видообразование.
11. Классические и генетические методы анализа биоразнообразия.

Занятие 6

ИСПОЛЬЗОВАНИЕ МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКИХ МЕТОДОВ ДЛЯ АНАЛИЗА ГЕОГРАФИЧЕСКОЙ ИЗМЕНЧИВОСТИ

План работы

1. Определить время расхождения клад при помощи Mesquite (<http://mesquiteproject.org/>).

Установить пакет Cartographer в Mesquite

Загрузить архив пакета в директорию (папку), куда автоматически попадают и другие загружаемые из интернета файлы («Загрузки» или «Downloads»). На странице: <http://mesquiteproject.org/packages/cartographer/download.html> перечислены все детали, касающиеся этого пакета, главная из которых – требования к системе. В данном случае необходимо, чтобы был установлен Mesquite версии старше 3.0 и Java версии старше 1.2. Сам пакет расположен по адресу:

<https://github.com/MesquiteProject/Cartographer/releases>.

Открыв эту страницу в браузере, надо выбрать загружаемый файл в зависимости от операционной системы. Для Windows он будет иметь расширение zip, для Linux и MacOS X – .tgz.

Распаковать загруженный файл. В результате появится папка (директория) с названием cartographer. Найти в своей системе папку Mesquite_folder. Запустить Mesquite и убедиться, что активированы все пакеты, как указано на рис. П26.

Использование пакета Cartographer. Выбрать картосхему, на которую будут нанесены данные. Источники карт: <http://mesquiteproject.org/packages/cartographer/sources.html>.

Определить проекцию карты. В настоящее время это должна быть одна из: Albers Equal-Area Conic, Lambert Azimuthal Equal-Area, Lambert Conformal Conic, Mercator Orthographic, Equirectangular, Transverse Mercator (это могут быть yandex maps или Google maps), желательно в формате jpg. с высоким разрешением.

Вставить координаты о местонахождении видов в новую матрицу (рис. П27 и П28), связанную (linked) с остальными видами, и приписать ей тип Geographic Data.

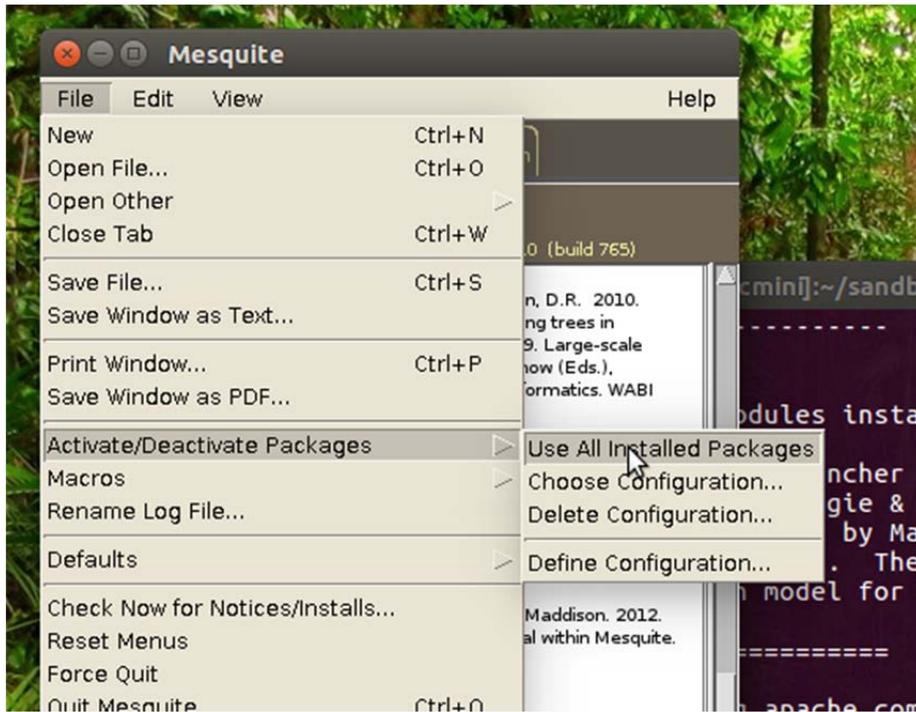


Рис. П26

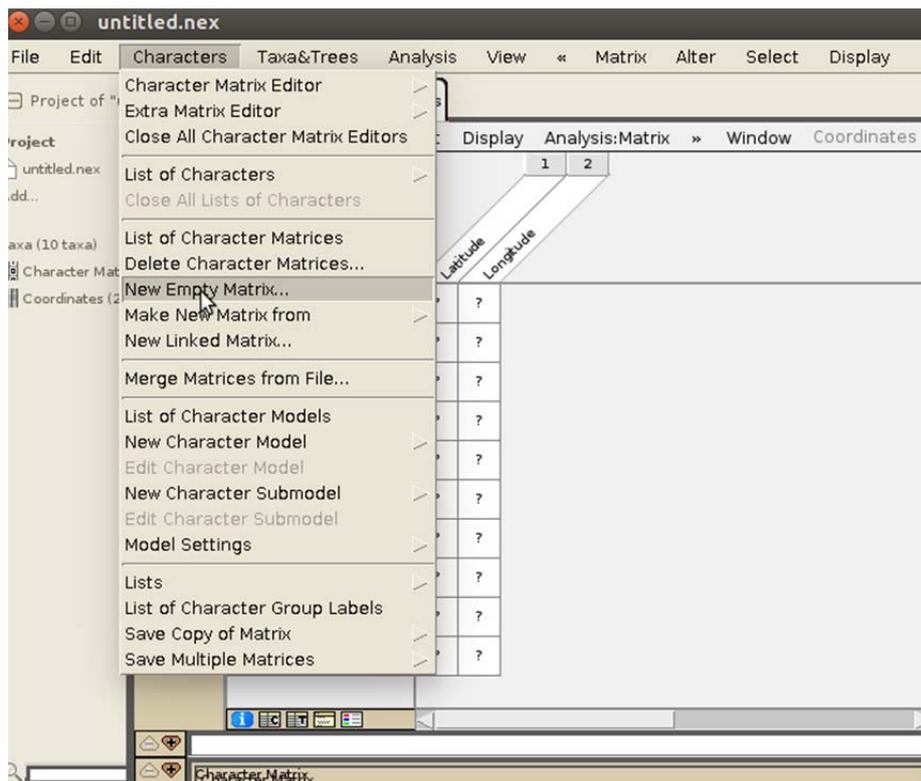


Рис. П27

Построить кладограмму по координатам и сделать её географическую привязку к скачанной картосхеме при помощи координат.

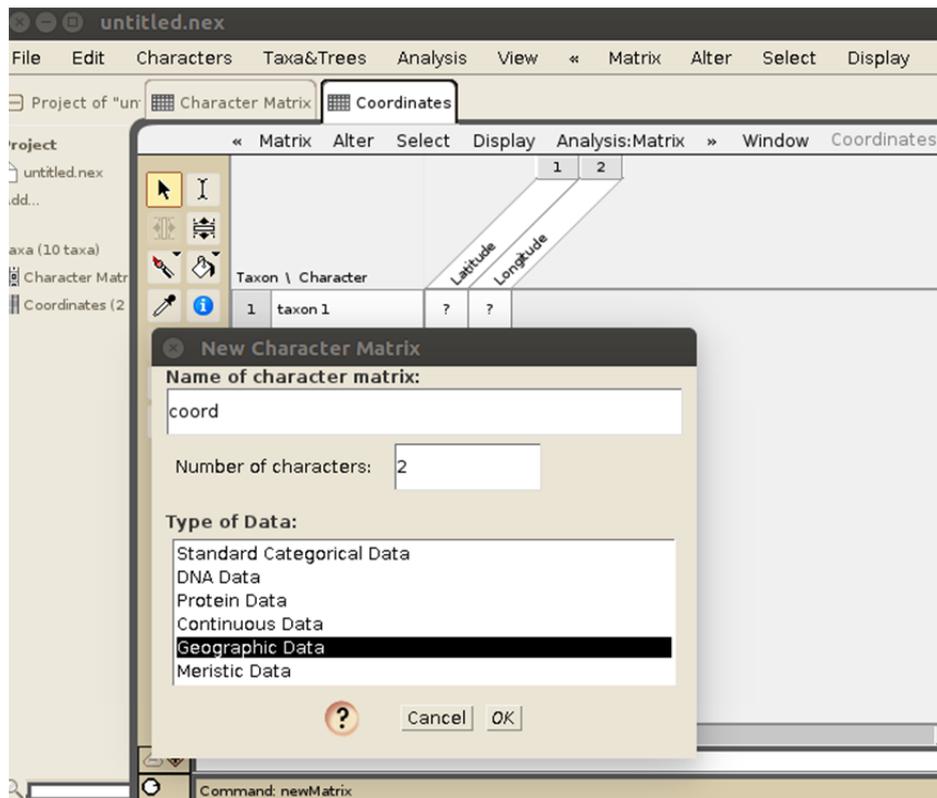


Рис. П28

3. В электронном виде сохранить результаты проделанной работы.

Контрольные вопросы

1. Адаптации.
2. Популяционные волны.
3. Модификационная изменчивость.
4. Методы выявления адаптаций.
5. Дать определение понятия ареал.
6. Типы ареалов
7. В чём суть и отличия макро- и микрофилогенеза.
8. Какие методы изучения темпов эволюции вам известны?
9. Какие формы эволюции вы могли бы назвать?
10. Что такое «критерий максимальной экономии» и в каких случаях его следует использовать?
11. Какие состояния признака называют апоморфными, а какие плезиоморфными?

Занятие 7

АНАЛИЗ БИОРАЗНООБРАЗИЯ ПРИ ПОМОЩИ ПОСТРОЕНИЯ ЭВОЛЮЦИОННЫХ СЦЕНАРИЕВ

План работы

1. Составить матрицу морфологических признаков и построить кладограмму, используя критерий максимальной экономии при помощи программы Mesquite [<http://mesquiteproject.org/>], и выявить апоморфные и плезиоморфные состояния признаков (рис. П29) с учётом топологии дерева.

Путь: Taxa&Trees → Make new trees bloc from → Tree search → Mesquite heuristic (рис. П30).

| Taxon \ Character | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|-------------------------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|
| 1 Lilium pumilum | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 Lilium davidii | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 Paris japonica | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 Trillium grandiflorum | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 Arum jacquemontii | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 Asparagus officinalis | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 Commelina communis | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 Vanilla odorata | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 Poa annua | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 Cypripedium calceol. | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 Cypripedium macrant. | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 Cypripedium japonicu | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 Paphiopedilum conoc | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 Phragmipedium longi | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 Paris quadrifolia | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 Aconogonon distachvc | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Рис. П29

В качестве критерия для оценки качества дерева можно использовать разные величины. На рис. П30 представлен выбор принципа максимальной экономии (parsimony) – пожалуй, наиболее распространенный.

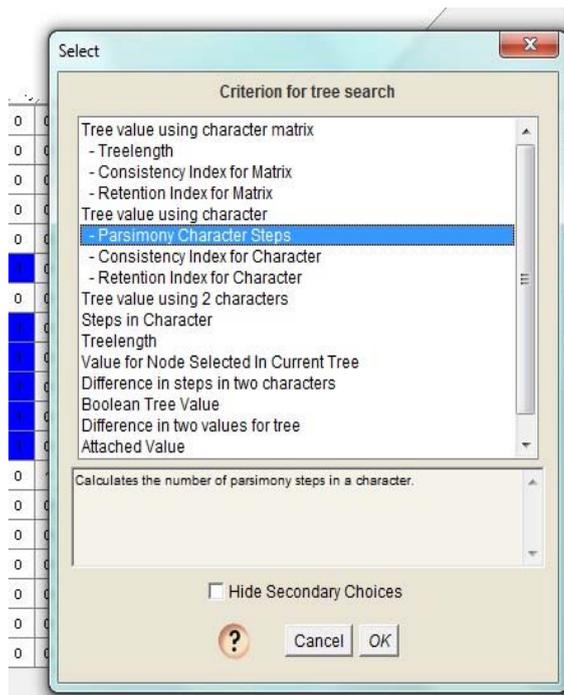


Рис. П30

После построения дерева можно раскрасить его ветви в соответствии со значением исследуемого признака. Для этого требуется восстановить наиболее вероятные состояния признака у предков, как показано на рис. П31, с помощью команды Analysis: Tree → Trace character history (см. рис. П31).

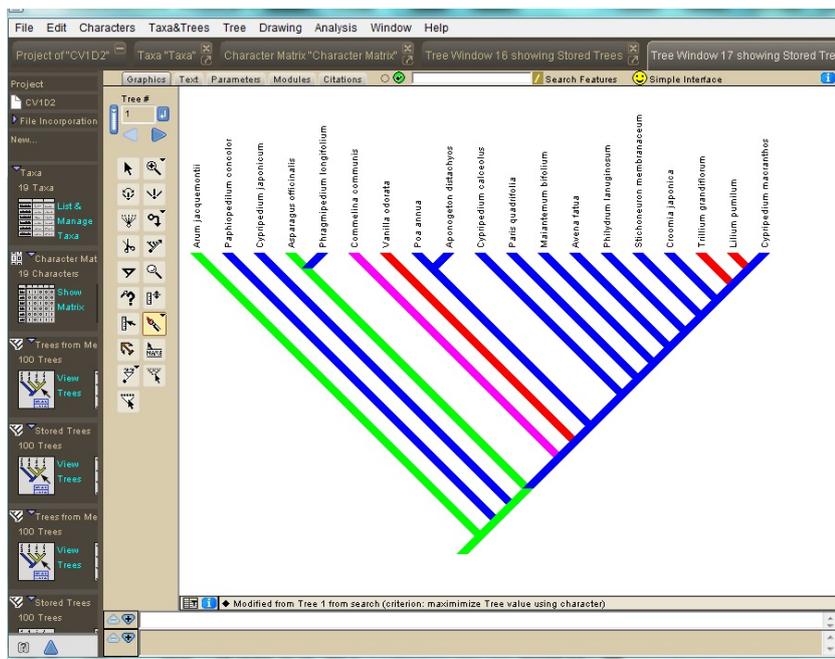


Рис. П31

В качестве метода для вычисления предковых состояний при исследовании морфологических признаков следует выбирать Parsimony Character States.

2. Составить выводы о проделанной работе и о биологической роли организма.

3. Представить результаты проделанной работы в электронном виде.

Контрольные вопросы

1. Для чего используется критерий максимального правдоподобия?

2. Для чего используется критерий максимальной экономии?

3. Что показывают бутстрепные поддержки?

4. Что показывают апостериорные вероятности?

5. На основании каких критериев может быть проведена оценка статистической достоверности эволюционного сценария?

6. Какие состояния признаков называют апоморфными, а какие плезиоморфными?

7. Что представляет собой ньюикская нотация деревьев?

Занятие 8

МЕТАГЕНОМНЫЙ АНАЛИЗ

План работы

1. Установить программу для анализа метагеномных данных

Самый распространенный на сегодня инструмент анализа метагеномных данных – программа *mothur*, разрабатываемая и поддерживаемая группой программистов под руководством Патрика Шлосса (Dr. Patrick Schloss) в Отделе микробиологии и иммунологии университета штата Мичиган [Schloss, 2009]. Эта программа позволяет проводить самые разнообразные операции над наборами метагеномных данных, и ее можно использовать на компьютерах с любой из распространенных операционных систем – Linux, MacOS X и Windows. Сайт программы: https://mothur.org/wiki/Main_Page. Здесь можно получить программу и справочные материалы на английском языке. Исходный код программы находится в открытом доступе, если его компилировать на том компьютере, на котором и предполагается использовать программу, то ее скорость существенно повысится. Кроме того, можно воспользоваться программами для 32- и 64-битных Windows, доступных по адресу: <http://www.applied-maths.com/download/mothur-1251-windows-executables-zip>. Там же содержатся и инструкции по его установке. Файлы данных, которые нам понадобятся в работе, находятся на <https://github.com/mothur/mothur/releases/tag/v1.39.5>.

На этом занятии мы проведем сравнительный анализ метагеномных данных, полученных в процессе программы сбора образцов Мирового океана во время экспедиции GOS (из 14 точек). Описание экспедиции можно найти по адресу:

<http://www.jcvi.org/cms/research/projects/gos/overview/>.

Упражнение основано на оригинальном примере, который можно найти на https://mothur.org/wiki/Marine_community_analysis.

Предполагается, что программа *mothur* установлена на компьютере и находится в рабочем состоянии.

2. Загрузить реальные данные о последовательностях, которые доступны по адресу <https://mothur.org/w/images/1/14/GOS.zip> (8,1 мегабайта), и распаковать файл в отдельную папку (директорию).

Образцы отобраны в следующих точках:

GS011 – эстуарий;

GS012 – эстуарий;

GS013 – прибрежный участок в Северной Атлантике;

GS014 – прибрежный участок в Северной Атлантике;

GS018 – прибрежный участок в Карибском море;

GS019 – открытое море, Карибское море;

GS020 – пресный водоем (озеро Gutan);

GS022 – открытый океан, восточные тропики;

GS023 – открытый океан, восточные тропики;

GS027 – прибрежный участок, Галапагосские острова;

GS028 – прибрежный участок, Галапагосские острова;

GS033 – гиперзасоленный участок на Галапагосских островах;

GS121 – открытый участок Индийского океана;

GS122a – открытый участок Индийского океана.

Последовательности собраны в файле `gos_proxy_seq.fasta`. Архив содержит также файл `gos.group`, в котором указано, какая последовательность относится к какой станции, а также `gos.names`, в котором перечислены идентичные последовательности.

Последовательности были получены с помощью высокопроизводительного секвенирования амплифицированных фрагментов бактериальной 16S рибосомной ДНК длиной примерно 560 нуклеотидов (отклонения в длине – за счет коротких делеций или вставок).

3. На следующем этапе необходимо выровнять последовательности. В нашем примере их более 1500, и для каждой из них нам потом потребуются найти ближайшую последовательность среди последовательностей, определенных ранее другими авторами. Для этой цели используют совместное выравнивание исследуемого набора и одной из стандартных баз данных, в нашем случае – GreenGenes:

а) перейти, используя какой-либо из стандартных браузеров, по адресу: <http://greengenes.lbl.gov/cgi-bin/nph-index.cgi> и выбрать

иконку Align: .

b) по кнопке «Choose File» (выбрать файл)

выбрать: gos_proху_seq.fasta;

c) в поле «Batch size for NAST» поставить «50»;

d) изменить значение в поле «Minimum Length» на 600;

e) ввести адрес своей электронной почты (чтобы на него могли прислать результаты выравнивания, которое может занять некоторое время) и запустить процесс кнопкой «Process FASTA File».

Не закрывайте страницу, пока не получите e-mail от Todd DeSantis (!).

Он будет содержать результаты выравнивания. В письме также будет файл в формате Excel с информацией о качестве последовательностей.

4. Построение матрицы дистанций между последовательностями

Филогенетический анализ, так же, как и другие способы анализа тысяч последовательностей, можно проводить только с помощью так называемых дистантных методов. Они несколько менее точны, чем обычные, однако позволяют использовать очень большие объемы данных. В любом случае требуется получить матрицу дистанций, т. е. таблицу, в которой по вертикали и по горизонтали находятся названия отдельных последовательностей (номера), а на пересечениях – оценка ожидаемого числа мутаций, которое требуется для перехода от одной последовательности к другой. Число мутаций оценивают в соответствии с выбранной моделью молекулярной эволюции, которая придает разным заменам разный вес (например, очевидно, что транзиции происходят легче и чаще, чем трансверсии). Поэтому не удивительно, что расстояния между последовательностями – дробные числа.

Для получения матрицы дистанций можно использовать онлайн-сервис GreenGenes, но можно воспользоваться возможностями mothur, запустив его в консоли или в случае Windows – из командной строки:

```
mothur> dist.seqs(fasta=gos.fasta,output=lt)
```

В такой форме программа построит матрицу дистанций в формате PHYLIP. Обратите внимание на то, что текст в скобках

не содержит пробелов, это важная особенность программы, которую необходимо учитывать в работе, причём даже использование содержащих пробелы имен файлов или папок приводит к ошибкам.

Особенность программы `mothur` состоит в том, что она не спрашивает пользователя, как назвать те файлы, которые она создает в процессе работы. Вот и в результате успешного выполнения приведенной выше команды у вас появится довольно большой файл:

```
gos.phylip.dist.
```

5. Кластеризация последовательностей

На этом этапе мы используем матрицу попарных генетических расстояний между последовательностями для того, чтобы разбить их на «операциональные таксономические единицы». Этот термин используется в филогенетическом анализе для того, чтобы избежать безосновательных утверждений о таксономическом статусе организмов, составляющих сообщество. Такие утверждения в любом случае особенно рискованны, когда речь идет о микроорганизмах, поскольку даже смысл понятия «вид» в приложении к ним является в настоящее время предметом бурной дискуссии. Выход из этой ситуации на сегодняшний день – в процедуре «биннинга» (от англ. *binning* – раскладывание по урнам). Согласно этой процедуре, организмы, различия между которыми не превышают, например, 0.1, принадлежат к одной и той же ОТЕ. Размер такой градации зависит от структурированности сообщества и определяется объемом выборки. В программе `mothur` он задается с помощью параметра `cutoff`.

Кластеризация осуществляется с помощью команды:

```
mothur> cluster(phylip=gos.phylip.dist,cutoff=0.1)
```

В результате выполнения этой команды появляются 3 файла:

```
gos.phylip.an.sabund,  
gos.phylip.an.rabund,  
gos.phylip.an.list.
```

Последний файл используем для того, чтобы определить, как «таксоны» распределены между образцами:

```
mothur > make.shared(list=gos_p.phylip.an.list,group=gos.groups)
```

Выполнение этой команды создает файл:

gos_p.phylip.an.shared

который и содержит эту информацию.

6. Проиллюстрировать результат разными способами

Для 3–4 групп/образцов хорошей формой представления материала является диаграмма Венна, которую можно построить с помощью команды: `mothur>venn(groups=...)` (рис. П32).

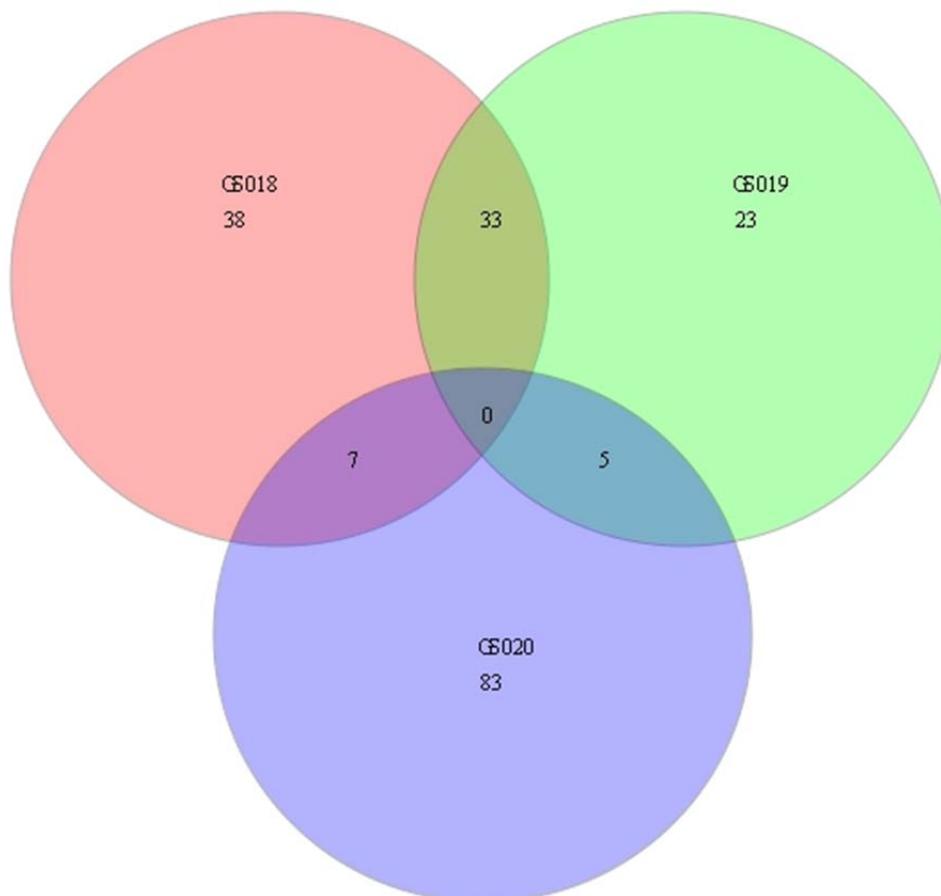


Рис. П32. Диаграмма Венна

Круги представляют образцы, а цифры в них – соответствующее число последовательностей.

Mothur создает графические файлы с расширением: SVG, которые можно редактировать во многих популярных редакторах векторной графики и просматривать в окне браузера.

В данном случае:

GS018 – прибрежный участок в Карибском море;

GS019 – открытое море, Карибское море;

GS020 – пресный водоем (озеро Gutan).

Другая форма представления – тепловая карта:

`heatmap.bin(scale=log2, label=0.03)` (рис. П33).

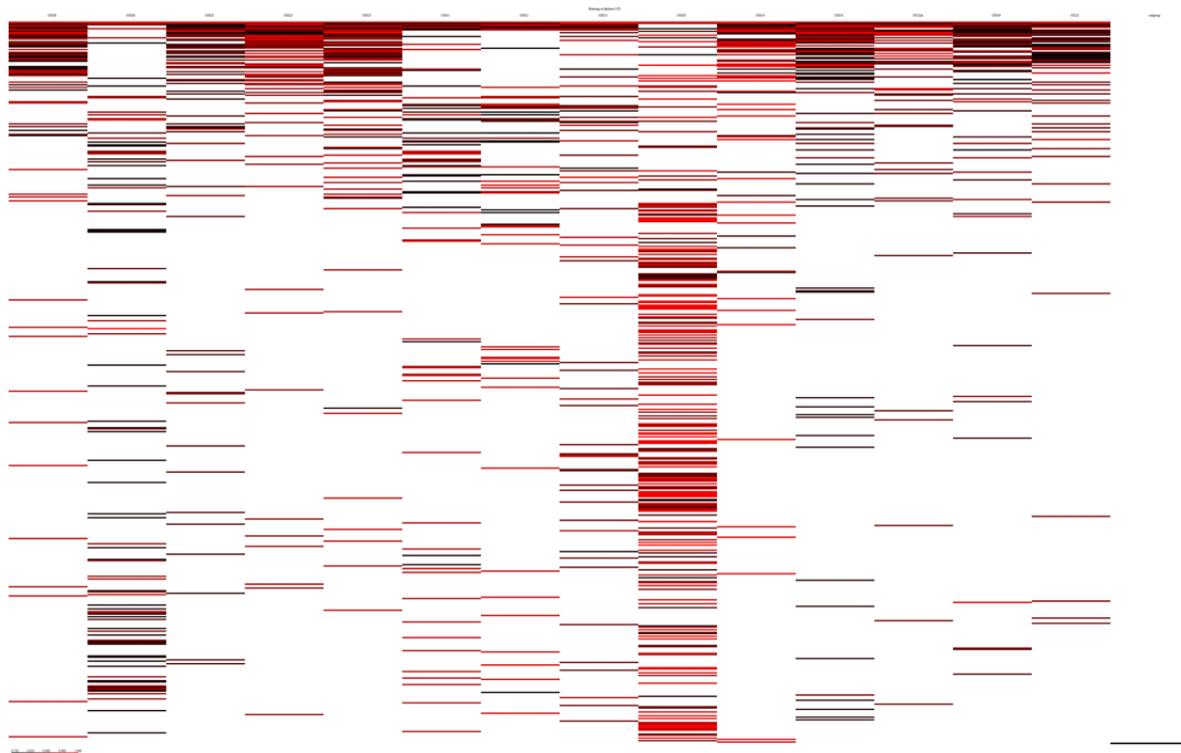


Рис. П35. Тепловая карта

Колонки представляют образцы, а горизонтальные черточки – таксоны. Чем больше идентичных последовательностей содержит образец, тем темнее черточка таксона.

Кластеризовать образцы на основании сходства организмов, входящих в соответствующие сообщества:

`tree.shared(calc=thetayc, groups=GS028-GS020-GS023-GS022-GS027-GS011-GS012-GS013-GS033-GS014-GS018-GS122a-GS019-GS121)` (рис. П34).

На основании данного состава этого кластера можно сделать вывод, что исходной группой были образцы GS033 с гиперзасоленного участка на Галапагосских островах, от них произошли образцы GS020 из пресного водоёма (озеро Gutan) и

GS121 – открытый участок Индийского океана;

GS011 – GS012 – эстуарий;

GS013 – GS014 – прибрежный участок в Северной Атлантике;
 GS018 – прибрежный участок в Карибском море;
 GS019 – открытое море, Карибское море;
 GS022 – восточная часть тропической части Тихого океана;
 GS023 – открытый океан, восточные тропики;
 GS027 – GS028 – прибрежный участок, Галапагосские острова.

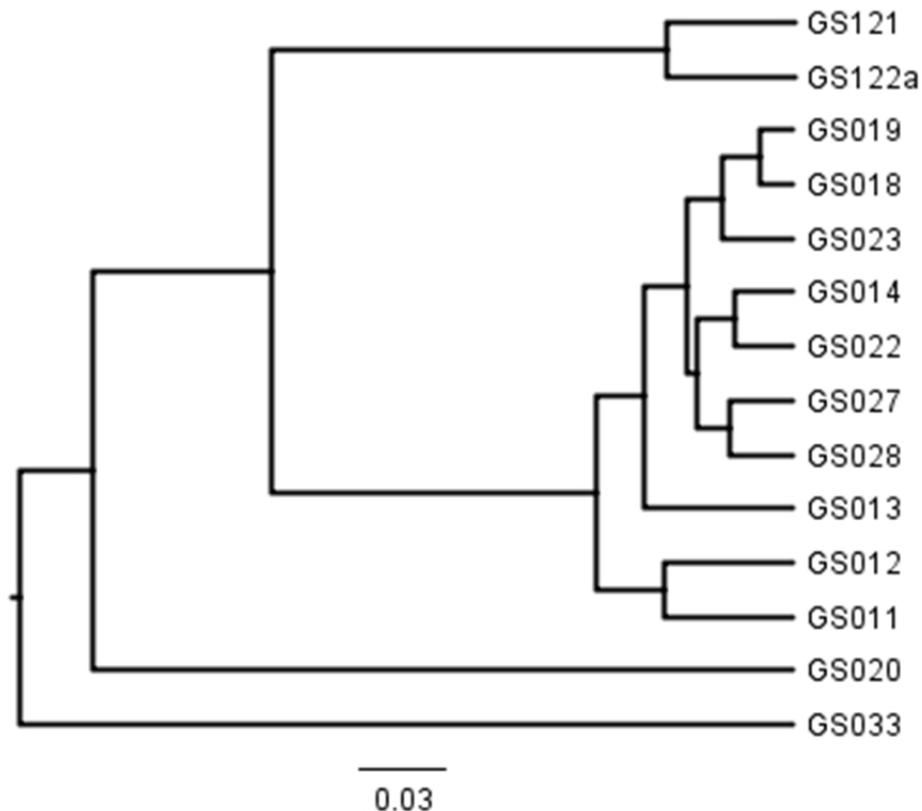


Рис. ПЗ4. Кластер степени сходства сообществ

Контрольные вопросы

1. Что такое «операциональные таксономические единицы»?
2. Методы выявления изолирующих механизмов.
3. Причины и значение географической изменчивости.
4. Причины и значение экологической изменчивости.
5. Популяционные волны.
6. Модификационная изменчивость.
7. Методы выявления адаптаций.
8. Методы анализа экотопов.
9. Методы изучения темпов эволюции.
10. Формы отбора.

Занятие 9

АНАЛИЗ БИОРАЗНООБРАЗИЯ ПРИ ПОМОЩИ МИКРОСАТЕЛЛИТНЫХ ПОВТОРОВ ДНК

План работы

Анализ полиморфизма микросателлитов в экологии чаще всего используется для решения проблем, связанных с описанием распространения каких-либо организмов и попытками проанализировать историю формирования современной ситуации. В частности, микросателлиты используются для выяснения популяционной структуры видов. Для этой цели существует много программ, соответственно – много форматов для входной информации и еще больше программ, позволяющих преобразовывать один формат в другой.

Это занятие полностью проводится в среде языка R – одного из наиболее развитых современных языков сверхвысокого уровня (наряду с Python и пр.). Этот язык содержит большое количество специализированных подгружаемых пакетов, которые подгружаются по команде пользователя. Часть из этих пакетов предназначена для решения популяционно-генетических и экологических задач с использованием самой разнообразной информации, в том числе – и результатов исследования полиморфизма микросателлитных маркеров (в терминах формальной генетики – кодоминантных маркеров). Мы воспользуемся пакетом *adegenet*, который наряду со средствами анализа хорошо приспособлен для подключения географической информации и имеет хорошие средства изображения результатов (Jombart T.(2008) *adegenet: a R package for the multivariate analysis of genetic markers*. *Bioinformatics*24: 1403-1405. doi: 10.1093/bioinformatics/btn129).

Предполагается, что R у пользователя уже установлен, если это не так, то следует установить его с сайта: <http://www.r-project.org>.

1. Запустить R и загрузить *adegenet* командой
>library(adegenet)

Начинающим пользователям следует также сразу же открыть страницу подсказок (откроется в браузере)

```
>help.start()
```

(если **adegenet** не установлен, то это можно сделать с помощью команды

```
>install.packages("adegenet", deps=T)
```

Установится сам пакет, данные и пакеты, необходимые для полноценной работы **adegenet**.

2. **Загружаемые данные** – микросателлитные генотипы 704 коров из Франции и Африки, принадлежащие к двум видам – *Bos taurus primigenius*, *Bos taurus indicus* и 15 группам. Для каждого образца было определено по 30 микросателлитных локусов:

```
>data(microbov)
```

Структуру набора данных можно просмотреть, набрав в командной строке R:

```
>microbov
```

На экране отобразится структура базы данных, которая хранится в виде текстового файла. Этот «шаблон» можно использовать для оформления собственных данных, когда это потребуется.

3. Сгруппировать набор данных, подразделив их на генетические кластеры. Следует заметить, что даже при наличии предварительной информации об ожидаемом числе групп генетически обоснованное разбиение сложно предсказать.

Вначале в разделе `tab` заменить отсутствующие данные (N/A) на средние арифметические:

```
X<-tab(microbov, freq=TRUE, NA.method=«mean»)
```

Затем делаем матрицу генетических расстояний между образцами:

```
>D<-dist(X)
```

После этого кластеризуем образцы командой:

```
h1<-hclust(D, method=«complete»)
```

Отображается дерево образцов на группировку ветвей на кластеры (названия, в данном случае, существенного значения не имеют):

```
plot(h1,labels=FALSE) (рис. П35)
```

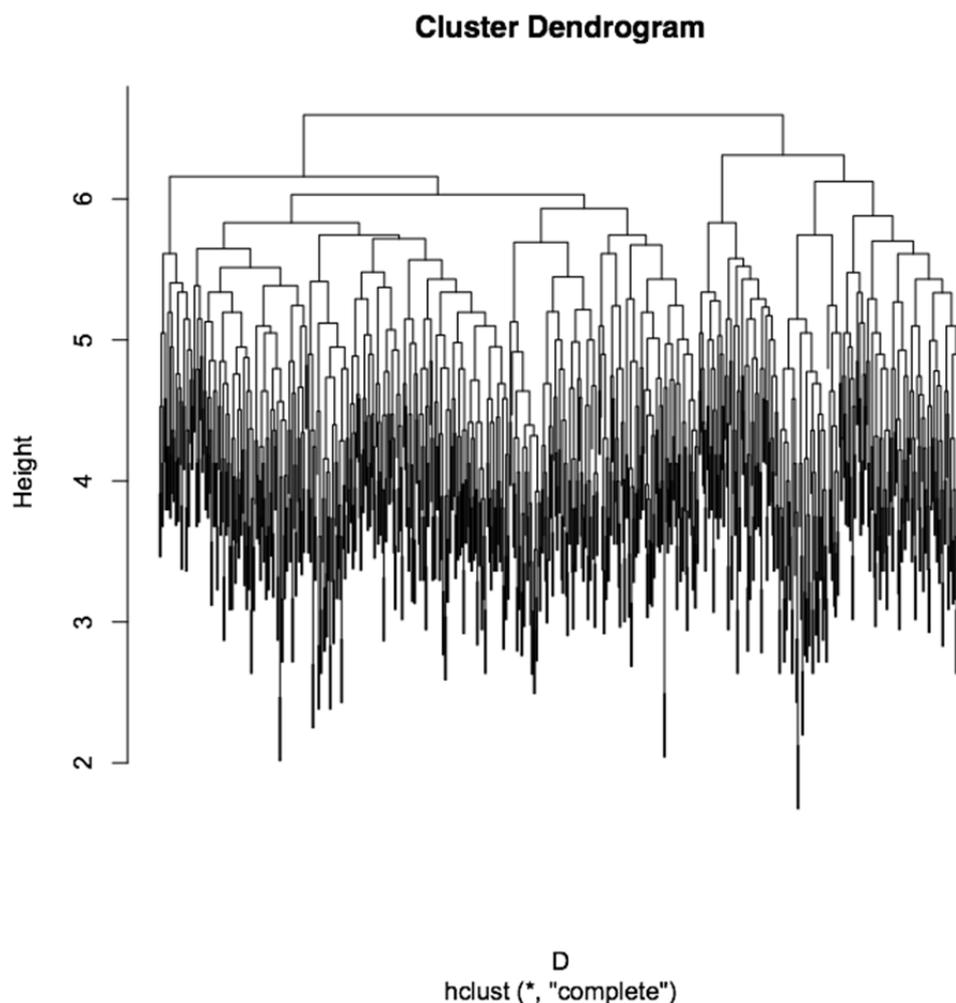


Рис. П35. Результат кластерного анализа

Получившееся дерево может служить наглядной демонстрацией проблем поиска естественных групп, на которые можно было бы разделить исследуемые организмы.

В ходе кластерного анализа задаётся уровень, который пересекает определённое количество ветвей, при этом каждая ветвь, пересекающая этот уровень, считается родоначальником своего кластера. Все потомки этой ветви являются членами соответствующего кластера (т. е. оказавшимися ниже черты).

Определить число групп и дать команду программе найти определённый уровень и провести классификацию при помощи команды:

```
>grp<-cutree(h1,k=2)
```

Для этого ввести переменную, присвоив ей имя: grp,

в качестве ее значения сообщить ей результат разбиения дерева h1 на 2 части.

Посмотреть результат можно при помощи команды:

```
>head(grp,10)
```

В результате получается список из первых десяти образцов, с указанием их принадлежности к группе:

```
AFBIBOR9503 AFBIBOR9504 AFBIBOR9505 AFBIBOR9506 AFBIBOR9507 AFBIBOR9508 AFBIBOR9509 AFBIBOR9510 AFBIBOR9511 AFBIBOR9512
```

```
1 1 1 1 1 1 1 1 1 1
```

Функция table позволяет строить таблицы соответствия.

Пример. Таксономическая принадлежность приводится в колонке spe;

распределение образцов по двум генетическим группам можно посмотреть при помощи команды

```
>table(grp, other(microbov)$spe);
```

результат выполнения:

```
grp BI BT
```

```
1 100 131
```

```
2 0 473
```

т. е. представители вида BI (*Bos taurus indicus*) встречаются только в группе 1, а вида BT (*Bos taurus primigenius*) – в основном в группе 2, но частично и в группе 1.

Полученные результаты являются основанием для гипотез о том, как оба подвида расселялись из района своего происхождения.

1. После разделения на два подвида (BT и BI) они заселили Африку, и только один из них попал во Францию.

2. Вид BT попал во Францию через Африку.

Проверить совпадение нашей классификации и происхождения образца следует с помощью команды:

```
>table(grp, other(microbov)$coun)
```

(происхождение указано в колонке coun).

Получается таблица:

```
grp AF FR
```

1 231 0

2 0 473

Данное разделение соответствует месту отбора образцов.

Теперь можно посмотреть, как по ним распределяются представители разных групп (в нашем анализе их 15). Это можно сделать с помощью функции

`table.value:`

1. Дерево разбить на 15 кластеров:

```
>grp1<-cutree(h1,k = 15)
```

2. Одной командой строится таблица, которую следует передать в качестве аргумента функции: `table,value`.

При помощи команды:

```
>table.value(table(pop(microbov),      grp1),      col.lab      =  
paste(«grp»,1:15))
```

значения в скобках – `tab` обозначает генотипы;

`col.lab` – обозначение групп, которое пишут в заголовках колонок (которых 15).

В результате получается таблица, состоящая из квадратиков, площадь которых пропорциональна числу случаев, когда принадлежность к данной группе, полученной путем кластеризации, соответствует определенной группе образцов.

Очевидно, что соответствие не абсолютное.

В состав библиотеки `adegenet` входят несколько функций кластеризации организмов на основании их генотипов. Помимо микросателлитов можно использовать полиморфные морфологические признаки, аллоферменты и однонуклеотидные замены («снипы», SNP).

Как видно на рис. ПЗ6, для каждого из них существует свой характерный спектр частот микросателлитных аллелей. Чем больше разных аллелей присутствует в генофонде группы, тем менее она репродуктивно изолирована от других. Вместе с тем знание особенностей генофондов отдельных групп позволяет контролировать скрещивания и, соответственно, может быть полезным инструментом для селекционеров.

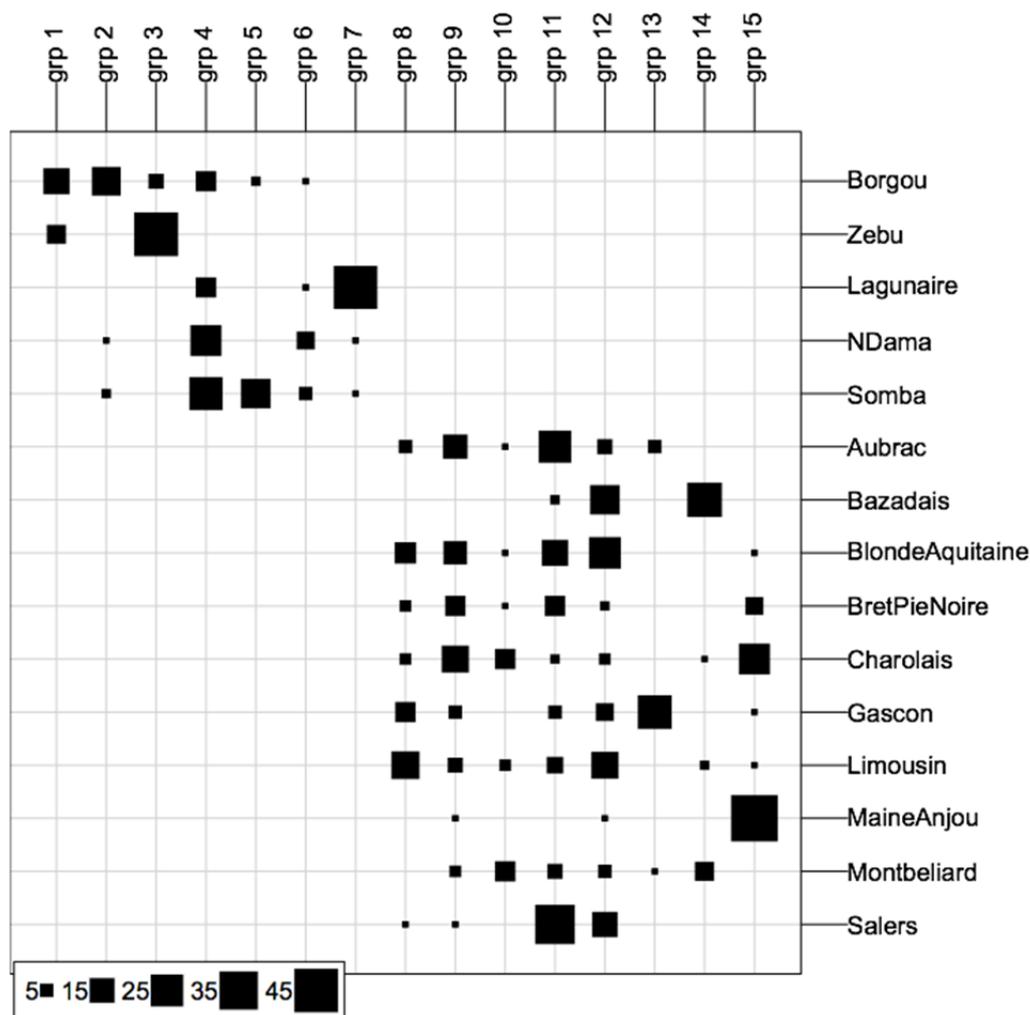


Рис. П36. Диаграмма соответствия между кластерами, выявленными микросателлитным анализом и породами КРС

Разбиение на 15 групп по степени генетических различий показало, что почти все они полиморфны по микросателлитным аллелям. Наиболее генетически однородными группами из них являются Zebu в Африке и Maine Anjou во Франции.

На основании данного примера кластеризации можно выбрать одну из двух гипотез, сформулированных ранее.

Во-первых, при использовании разбиений на 2 кластера в одном случае – по таксономической принадлежности, а в другом – по географическому принципу, мы увидели, что в последнем случае выявляется точное соответствие разбиению по генетическим расстояниям.

Если бы была верна вторая гипотеза, то следовало бы ожидать появление африканских генотипов среди французских образцов; так как этого не происходит, то результаты проведённого анализа свидетельствуют о независимом расселении коров в Африке и во Франции.

Контрольные вопросы

1. Типы мутаций.
2. Общие свойства мутации.
3. Распространение мутаций в пределах ареала вида.
4. Эволюционное значение мутаций.
5. Закон Харди – Вайнберга.
6. Методы выявления мутаций.
7. Формы видообразования.
8. Пути мгновенного видообразования.
9. Эволюционная роль гибридизации.
10. Методы выявления механизмов видообразования.
11. Метод максимального правдоподобия.
12. Факторы, лимитирующие распространение вида.
13. Классификация изолирующих механизмов.
14. Генетика изолирующих механизмов.
15. Роль изолирующих механизмов.
16. Нарушения изолирующих механизмов.

ЗАКЛЮЧЕНИЕ

В заключение следует обратить внимание читателя на то, что примеры использования методов молекулярно-генетического анализа позволяют продвинуться в решении вопросов, касающихся механизмов формирования и сохранения современного биоразнообразия, и структурировать сведения, которые получает исследователь в ходе своей работы. Нам остаётся пожелать нашим читателям не забывать, что «односторонний взгляд не способствует пониманию сути вещей» и поэтому желательно всячески избегать узкой специализации и не недооценивать те методы, которые использовались в ходе истории формирования смежных областей биологии. Морфология, на которой было сосредоточено внимание исследователей времён Гёте, на сегодняшний день не стала менее ценной для понимания биоразнообразия, так как именно она является инструментом адаптации организмов, так же как не утратили своего значения достижения классической генетики, объясняющие механизмы наследования признаков. Математические закономерности, сформулированные Фишером, Райтом и пр., по-прежнему важны при анализе изменчивости признаков, а разработанные ими методы математической статистики преобразили биологию и смежные ей науки. Овладевая новыми методами анализа и воспринимая новые идеи, мы оказались в ситуации, когда объём информации растёт с огромной скоростью, но полезно задаться вопросом: в какой мере увеличение объёма перерабатываемой информации приближает нас к пониманию сути проблем биоразнообразия?

ЗАДАНИЯ ДЛЯ КОНТРОЛЯ УМЕНИЙ И НАВЫКОВ СТУДЕНТОВ

1. На основании имеющихся последовательностей при помощи GENBANK определить ген.
2. При помощи программы BLAST найти группу наиболее похожих последовательностей.
3. Определить организм на основании заданной последовательности.
4. Охарактеризовать систематическое положение на древе жизни на основании заданной последовательности.
5. Построить дерево при помощи программ Phylml и Seaview.
6. Охарактеризовать вариабельность последовательностей.
7. При помощи программы Seaview представить дерево в разных форматах (циркулярное, радиальное и пр.).
8. Определить время расхождения клад при помощи Mesquite.
9. Построить кладограмму на основании географии распространения таксонов при помощи программы.
10. Выявить апоморфные и плезиоморфные состояния признаков с учётом критерия максимальной экономии на основании имеющейся кладограммы.

СЛОВАРЬ ТЕРМИНОВ

Ампликон – фрагмент ДНК с праймерами на флангах, полученный в результате амплификации с помощью полимеразной цепной реакции.

Ареал (от лат. *area* – *площадь, пространство*) – это территория (акватория), на которой распространен вид и проходит свой полный цикл развития.

Эндемик – вид, который встречается только в одном месте.

Космополиты – виды, которые встречаются на большей части обитаемых областей Земли.

Первичный ареал – территория, на которой происходит становление вида.

Аутбридинг – скрещивание неродственных организмов.

Инбридинг – близкородственное скрещивание.

Вид – это самый младший таксономический ранг, основная единица (категория) биологической классификации.

Подвид (географическая раса, экотип) – наиболее крупная таксономическая единица внутри вида, которая представляет собой группу свободно скрещивающихся особей, характеризующихся одним или несколькими наследуемыми признаками, имеющая свой внутривидовой ареал и приспособление к определенным экологическим условиям (климатическим, эдафическим, фитоценоотическим и пр.).

Клон – группа особей, происходящих от одной особи в результате бесполого размножения.

Экады – формы ненаследуемой (модификационной) изменчивости, приуроченные к определенным местообитаниям и носящие приспособительный характер.

Популяция – сообщество потенциально скрещивающихся особей в данной местности.

Внутривидовая изменчивость – свойство особей одного вида, отличающихся друг от друга по своим морфологическим, анатомическим признакам, биологическим или физиологическим и биохимическим особенностям, продолжительности онтогенеза и времени смены фаз развития и пр.

Скорость дрейфа – средняя скорость исчезновения аллелей в результате дрейфа генов.

Сорт представляет собой популяцию организмов, искусственно созданных человеком и обладающих определёнными наследственными особенностями (морфологическими, физиологическими и пр.), **проявляющимися при определённых условиях среды** (культивирования, агротехники и пр.).

Движущий отбор происходит, если оптимум значения признака сдвигается в сторону одного из крайних значений.

Дизруптивный отбор происходит, если возникают два оптимума, один вблизи максимального, другой – вблизи минимального значений признака.

Стабилизирующий отбор отсекает крайние формы независимо от того, в какую сторону от оптимума произошло отклонение.

Дрейф генов – случайное изменение частот аллелей в популяции, не обусловленное влиянием отбора.

Принцип Красной королевы – «чтобы оставаться на прежнем месте, нужно очень быстро бежать». Для того чтобы продолжать существовать в своей экосистеме, вид должен постоянно эволюционировать.

Метагеномный анализ – идентификация видового состава биотопа на основании нуклеотидных последовательностей стандартных локусов (штрихкодов) в ДНК, экстрагированной из природных образцов так, как если бы она была получена из одного организма.

Микросателлитный анализ – анализ на основании использования очень быстро эволюционирующих молекул ДНК, включающих короткие повторы коротких (от одного до восьми нуклеотидов) мономеров, расположенные вплотную друг за другом.

Штрихкодирование – идентификация организмов с помощью одного или нескольких стандартных фрагментов ДНК, уникальных для своих носителей.

PCR – полимеразная цепная реакция, или амплификация ДНК.

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Андреев А. В. Оценка биоразнообразия, мониторинг и экосети / А. В. Андреев. – Кишинев : Biotica, 2002. – 168 с.

Биоразнообразие и динамика экосистем: информационные технологии и моделирование / М. Г. Сергеев [и др.]; отв. ред.: В. К. Шумный [и др.]; Рос. акад. наук [и др.]. – Новосибирск : Изд-во Сиб. отделения РАН, 2006. – 646 с. – (Интеграционные проекты СО РАН ; вып. 7).

Бродский А. К. Биоразнообразие : учебник / А. К. Бродский. – М. : Академия, 2012. – 207 с.

Майр Э. Популяции, виды и эволюция / Э. Майр. – М. : Мир, 1974. – 460 с.

Павлинов И. Я. Аспекты биоразнообразия / И. Я. Павлинов. – М. : Тов-во науч. изд. КМК, 2016 – 831 с.

Протасов А. А. Биоразнообразие и его оценка. Концептуальная диверсиконология / А. А. Протасов. – Киев : Наукова думка, 2002. – 105 с.

Тахтаджян А. Л. Жизнь растений. В 6 т. Т. 1 / А. Л. Тахтаджян. – М. : Просвещение, 1974.

Тимофеев-Ресовский Н. В. Краткий очерк теории эволюции / Н. В. Тимофеев-Ресовский, Н. Н. Воронцов, А. В. Яблоков. – 2-е изд. – М. : Наука, 1969. – 408 с.

РЕКОМЕНДУЕМЫЕ ВЕБ-РЕСУРСЫ

<http://www.mobot.org/MOBOT/Research/APweb/welcome.html>

<https://www.ncbi.nlm.nih.gov/pubmed>

<http://www.tolweb.org/tree>

<http://www.barcodeoflife.org/>

<http://evolution.genetics.washington.edu/popgen/popg.html>

http://nature.air.ru/biodiversity/book3_1_5.html

<http://ecoregions2017.appspot.com>

<http://mesquiteproject.wikispaces.com>

<https://mesquiteproject.wikispaces.com/Additional+Mesquite+Packages>

<http://mesquiteproject.org/packages/cartographer/>

<http://www.ub.edu/dnasp/>

<http://www.ub.edu/dnasp/DnaSPHelp.pdf>

<http://mesquiteproject.org/>

<http://www.barcodeoflife.org/>

<http://www.tolweb.org/tree>

Seaview: <http://doua.prabi.fr/software/seaview>

Gouy M., Guindon S., Gascuel O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building // *Molecular Biology and Evolution*. 2010. Vol. 27(2). P. 221–224.

Phyml: <http://www.atgc-montpellier.fr/phyml/>.

Guindon S., Dufayard J.F., Lefort V. et al. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0 // *Systematic Biology*. 2010. Vol. 59(3). P. 307–321.

Clustal: <http://www.clustal.org/omega/>.

Sievers F., Wilm A., Dineen D. G. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega // *Molecular Systems Biology*. 2011. Vol. 7. P. 539.

ПРИЛОЖЕНИЯ

Приложение 1

1. AGCTGGATAGGAATAAAAGGCATCTTTTTATTAAAAGAC-
CTTTTTTCGGAAGGTATGCTTAAAGATGAGCTTGCGCCACATTAGTTT-
GTTGGTGGGGTAATGGCCTACCAAGACGATGATGTGTAGCTGGACTGA-
GAGGTCGAACAGCCACATTGGGACTGAGACACGGCCCAAACCTCCTAC-
GGGAGGCAGCAGTAGGGAATTTTCGGCAATGGAGGAAACTCTGAC-
CGAGCAACGCCGCGTGAACGATGAAGTATTTTCGG-
TATGTAAAGTTCTTTTTATTGAAGAAGAAAAAATAGTGGA AAAAC-
TATCTTGACGTTATTCAATGAATAAGCCCCGGCTAACTATGTGCCAG-
CAGCCGCGGTAAAGACATAGGGGGCGAGCGTTATCCG-
GAATTATTGGGCGTAAAGGGTGCCTAGGCGGTTAGGAAAGTC-
TATAATTTAATTTTCAGTGCTTAAACGCTGTCTTGTTATAGAACTACCTT-
GACTAGAGTTAGATAGAGGCAAGCGGAATTCATGTGTAGCGG-
TAAAATGTGTAAATATATGGAGGAACACCAGAAGCGTAGGCGGCTT-
GCTGGGTCTTGACTGACGCTGAGGCACGAAAGCGTGGGTAG-
CAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAG-
TACTAAGTGTCGGGATAAAAACCTCGG-
TACTGAAGTTAACACATTAAGTACTCCGCCTGAGTAGTACGTAC-
GCAAGTATGAAACTTAAAGGAATTGACGGGACTCCGCACAAGCGGTG-
GATCATGTTGTTAATTCGAAGATACACGAAAAATCTTACCAGGTCTT-
GACATGCTCTGCAAAGCTATAGAAATATAGTG-
GAGGTTATCAGGGACACAGGTGGTGCATGGTTGTCGTCAG-
TTCGTGTCGTGAGATGTTAGGTTAAGTCCTAAAACGAGCGCAACCCTT-
GTCGTTAGTTACCAGCACGTAATGGTGGGCACTTTAACGA-
GACTGCCAATTAACATTGGAGGAAGGTGAGGATTAC-
GTCAAATCATCATGCCCTTATGATCTGGGCTACAAACGTGATA-
CAATGGCTGTTACAAAGAGTAGCTGAAGCGCGAG-
TTTTTAGCCAATCTCAAAAAGCAGTCTCAGTTCGGATTGAAGTCTG-
TAACTCGACTTCATGAAGTTGGAATCGCTAGTAATCCCGAATCAG-
CATGTCGCGGTGAATACG

2. TTTTTACATTTAAATTATGTGTCAGATATACTAA-
TACCCACCCCATCCATCTGGAAATATTGGTTCAAACCTCTTCGC-
TACTGGGTGAAAGATGCCTCTTCTTTGCATTTATTACGATTCTTTTTC-
TATGAGTATCAGAATTGTAATAGTCTTATTACTCCAACCTCCAAA-
GAAATCCATTTCCATTGTTTCAAAAAGGAATCAAAGATTATTCTT-
GTTTCTATATAATTCTTATGTATCTGAATAC-
GAATCTATCTTCATTTTTCTTTGTAACCAATCTTCTCATTTAC-
GATCAACATCTTCTGGAACCTTTTTTGAGCGAATATATTTT-
TATGGAAAAATAAACATCTTGTAGAAGTCTTTTCTAATGAT-

TTTCCGGCCGTCCTATGGTTGTTCAAA-
GATCCGTTTCATGCATTATGTTAGGTATCAAGGAAAATCAATTCTGG-
CATCAAAGGTACGCCTCTTCTGCTGAATAAATGGAAATATTACCTT-
GTCAATTTTTGGCAATGTCATTTTTACGTATGGTCTCAACCAG-
GAAGGATCTATATAAAC-
CAATTATCCAATCATTCCCTCGACTTTCTGGGCTATCTTTCGAG-
TGTGTGGCTAAATCCTTCAGTGGTACGGAGTCAAATGCTA-
GAAAATTCATTTATAATCGATAATGTTATTAAGAAGTTCGA-
TATCATAGTTCCAATTATTCCTCTAATTAGATCATTGGCTAAA-
GCTAAATTTTTGTAACGTAGTAGGGGATCCTATTAGTAA-
GCCGGTCTGGGCCGATTCATCAGATTCTGATATTATCGACAGATTT-
GGGCGCATATGCAGAAATATTTCTCATTATCACAGCG-
GATCCTCAAAAAAAAAAAGT

3. ATGTATGTAAATGGCAGAATTTTAAGGA-
TATTAATAATTTAAAAAGGATGGATTTTCGGCAACAAAACCTCCTA-
TATCCGCTACTCCTTCAGGAGTATATTTACTCACTT-
GCTCATGATCATAACTTCAATAGTTTTTTTTTTTACAAACCTGTG-
GAAATTTTTGGTTATGACAATAAATCTAGTTTAGTACTTGTGAAAC-
GTTTAATTAATCGAATGTATCAACAGAATTCTTTGAT-
TTCTTCGGTTAATTATTCTAACCAAAAAGGATTTTGGGGGCACAA-
GAATTTTTTTTTTCTAATTTTTATTCTCAAATGGTATCAGAAGGTTTT-
GGAGTCATTCTGGAAATTCCATTCTCGTCCCAATTAGTATCTTCTCTT-
GAAGAAAAAAAAAATACCAAAATCTCAGAATTTAC-
GATCTATTCATTCAATATTTCCCTTTTTAGAGGACAAAATTTTTACATTT-
GAATTATGTGTCAGATCTACTAATACCCCATCCCATCCATCTG-
GAAATCTTGGTTCAAATCCTGCAATGCTG-
GATCAAGGATGTTCCCTTCTTGCATTTATTGCGATTGCTTTTCCACGAA-
TATCATTATTTAATAGTCTCATTACTTCAAAAAAAAAAGCATTAC-
GCCTTTTCAAGAATAAAGAAAAGATTCCTTTGGTTCCTATA-
TAATTCTTATGTATATGAATGCGAATATCTATTCCATTTTCTTCG-
TAAACAGTCTTCTTATTTACGATCAACATCTTCTGGAGTGTTTCTT-
GAGCGAACACATTTCTATGTAAAAATAGAACATCTTATAGTAG-
TGTGTTGTAATTCTTTTCATAGGATCC-
TATGCTTTCTCAAGGATCCTTTCATGCATTATGTTTCGATATCAAGGAAAA
TATCAAGGAAAA-
GCAATTCTGGCTTCAAAGGGAACTCTTATTCTGATGAA-
GAAATGGAAATTTTCATCTTGTTAATTTTTGGCAATCTTATTTGCAC-
TTTTGGTCTCAACCGTATAGGATCCATATAAAGCAATTATACAAC-
TATTCCTTCTCTTTTCTGGGGTATTTTTCAAGTGTACTAGAAAATCATT-
GGTAGTAAGAAATCAAATGCTAGAGAATTCATTTCTAATAAA-
TATTATGACTAAGAAATTAGATACCATAGCCCCAG-
TTATTTCTCTTATTGGATCATTGTGCGAAAGCTCAATTTTGTACTG-
TATTGGGCCATCCTATTAGTAAACCGATCTGGACCGATTTATCGGAT-

TCTGATATTCTTGATCGATTTTGCCGGATATGTAGAAATCTTT-
GTCGTTATCACAGCGGATCCTCAAAAAAAAAAAGTTTTGTATCG-
TATAAAGTATACTTCGATTTTCGTGTGCTAGAACTTTGGCGCG-
GAAACATAAAAGTACAGTACGCACTTTTATGCGAAGATTAGGTTCCG-
GACTATTAGAAGAATTCTTTATGGAAGAA-
GAAGTTCTTTCTTTAATCTTCCCTCAAAAAAATCCCTTTTCCCTTACAC-
GGATCACATAGAGATCGTATTTGGTATTTGGA-
TATTATCCATATCAATGATCTGGTGGATCATTTCATGA

4. GTTATATCGATCGATGTATAGAATTCTAATGGCAAATGGGTCCA
TAAAATCATCAAATCAATTAGACTATGAT-
TTAAGTCCTTTTTTTCTTCTTCCTTCCTGAAAAATAAAAAAATCATTT-
GTACTCATAACTCAAGTTGGATAACTTTTAAATAGCTCAAA-
GAAAAATCTTTAGCCATTTATTGAGTGGTCTTTAATCCCCTTTTTGTTT-
GTCTCATTTAAACCTATTTTGATTCTTCAGTCTGATCCAG-
TTATTGAGACAATTGAAAGGGTGTTCCTTGTCTGG-
GATCCTTTATTTTTTGTGTTTGAATCATTGGGTTTAGA-
CATTACTTCGGTGATCCTTAATCCTTTCAAACGGCAGCAACAT-
ACCTTTTTTTTTGATTTCTTCTATCAAAGAATCATATGAACAGTTGAT-
TCCCGCGTGATACACTTTGTATCGAAAGAG-
TTTTATCAATTCTAACAAATTTAATTTTCACTTTTTTTGATTGGAACTT-
GCTCGAATTGGATCCTTTGGATTTCTATATCGAAGATATACTTAC-
GAAGTTGTTCCAATTTATTGATTGGCATTAAACCCTAGATCTTT-
GCTCCCGAGAAATGAATCAATACTTTCTACTCGAGCTC

5. GAATTAAGAATTCTCACAATAACAAGGTCTACTCGATATGGATT
AGCCGTGACTCAAACCTTTGACAAAAATCTTTCCAAAAAATATTCCAC-
TAATATTATCTATTATCAACGAATTAAGTTCGTTATTAGAC-
CATGCGTTTCATTCACCAAATACATCATTATTGTATACTCTTTGA-
TATACATGGCGCAACCCAACCCAATCCTCTTCAAGTTTATAATTGAATT-
GATCCGCTTCTTTTTTTGAATATAAAAAATCTAAATACTAAGAAAAAA-
GCTACTTGACAGTAGTATATGTTGTATATGTAAATCCTA-
GATGTGAAAATAGGCATAATTCATCCATGAAAAAGATA-
TAAAACAAAAATAGGCACTATAACTATAAATAACTACTAACTA-
TATTATAACTATAAACTATAATATAAAATATATAAAATATAAA-
TAAAAGAAAAATAAATAGAATAAATAACTAACTATAACTATACTA-
TATAGAATATATATAATATAAAATATAAATCTATATTCTATATATAA-
TAGAAATTTCTAAATTCTAAATCTAGAAAATATAAAATAAA-
TAGAATTCCAAATGACAAATGCAATAAAAAAATTTTTGGAATTGGTT-
GAACTTGAAAATTGACTCATTCAAATGAAATGAGTAAACGAG-
TGAATTTGATTCGGTTGGGTAGTACCAACTAAATCGAG-
TGCTAATTCCATTTATTTCTTATTGAATTAACCGATCAACTTGC-
TATCGGATATTTATTTTCGATTTCGATTTCGGTACTA-
TAGTACTATTTCAATCTCCAATCCAAAAAAA

6. TCGAGAATCGATTGAGACACCGCGAACCTGTAAACGGATGATA
CCGTGTCGGGCAGGCGTTATGCCCGCCAACTCGGGACCTCG-
CATCGTGTCCGCGGCTGCCTTCGAGCGTTTCGGGCACGATTTGGGGGG-
GACGAACGAAACCCCGGCACGGCCTGTGCCAAGGAACATATGTCAG-
GACGGACGCTCGTCAATGCCTCAGTGGTGGGGCGACGTTTCGCTCTC-
TATCTATACGACTCTCGGCAACGGATATCTCGGCTCTCG-
CATCGATGAAGAACGTAGCGAAATGCGATACTTGGTGTGAATT-
GCAGAATCCCGTGAACCATCGAGTCTTTGAACGCAAGTT-
GCGCCCGAGGCCTTTCGGTTGAGGGCACGCCTGCCTGGGCGTCAACCCTT-
GCCTGCCTGGGCGTCAACCCTT-
GTTTCGCTCTGTGCCCATGCTCTTTCGGGGGCGGTCATGGATGCGGA-
GATTGGCCCTCCGTGCCTCGTGTGCGGGCGGGCTTAA-
GCGCGGGCTGTCGGCGTCGGGATGGGCACGACGAGTGGTGGACGGAG-
CACCAGCAGGATGTTGTGGTCCCCCGTCGCCTTAAGGGGCTCAAGA-
GACCCGGACTAGGCGAGCCGTGCTCCGTAAGAG-
GAGGGCGAGCCGTCTCGCAAGGCGACCCAGGTC

7. TTTCCGTAGGTGAACCTGCGGAAGGATCATTGTCGTGACCCTTA
AACAAAACAGACCGCGAACGAG-
TCCCTCGTGCCGCCGAGCTTCGGCTCGGCATT-
AGGTCCCCGAGCTCCGTCCCGGGGCG-
GAGGGGCCACAACAGAACCCACGGCGCCTTAGGGCGTCAAGGAACAC-
TTATATTGCCTTGCTCGGCG-
GAGCGGTCGGCCTGCCTTCCGCTCCCCGCGCAGCGATGA-
TATCTTAATCCACATGACTCTCGGCAACGGATATCTCGGCTCTCG-
CATCGATGAAGAACGTAGCAAATGCGATACTTGGTGTGAATT-
GCAGAATCCCGCGAACCATCGAGTTTTTTGAACGCAAGTT-
GCGCCCGAGGCCTTCTGGCCGAGGGCACGTCTGCCTGGGCGTCAAC-
GCCAAAAGACACTCCCAACCCACCCAAGAGGGGGAGGGAC-
GTGGTGTGGCCCCCGCGCCGAGGGCGTGGTGGGCCGAAGTT-
GGGGCTGCCGGCGAATCGTGTGCGGGCACAGCACGTGGTGGGCGACAC-
TCAGTTGTTCTCGGTGCAGTGCCCCGGCAGTCGGCCGGCG-
CATCGGCCCTAAGGACCCATGGAGCACCGCAGCGCATCGCCGCTCG-
GACCGCGACCCAGGTCAGTCGGGACTACCCGCTGAGTTTAA

Приложение 2

1. Ген SSU (16S) rRNA Genbank: KJ093444 – *Magnolia grandiflora* L. (семейство Magnoliaceae) – Магнолия крупноцветковая.
2. Ген: матураза (*matK*) Genbank: JQ895054 – *Astilbe chinensis* (Maxim.) Franch. & Sav. (семейство Saxifragaceae) – Астильбе китайская.
3. Ген: матураза (*matK*) Genbank: AY557208 – *Cypripedium calceolus* L. (семейство Orchidaceae) – Венерин башмачок настоящий.
4. Ген: Genbank: AB095570 16SrRNA, последовательность *Hydrangea paniculata* Siebold (семейство Hydrangeaceae) – Гортензия метельчатая.
5. Ген: Genbank: AB23732 rbcL-atpB intergenic spacer *Hemerocallis middendorffii* Trautv. & C.A. Mey. (семейство Hemerocallidaceae) – Лилейник Миддендорфа.
6. Genbank: AF074470 Фрагмент ДНК содержит internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and 25S ribosomal RNA gene, partial sequence – *Lilium debile pensylvanicum* – Лилия пенсильванская.
7. Genbank: JN544339 Фрагмент ДНК содержит: 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence – *Miscanthus purpurascens* Andersson (семейство Poaceae) – Верник краснеющий.

ИСПОЛЬЗОВАННАЯ ЛИТЕРАТУРА

- Булыгин Н. Е.* Дендрология. 2-е изд., перераб. и доп. Л. : Агропромиздат. Ленингр. отд-ние, 1991. 352 с.
- Вавилов Н. И.* Географическая изменчивость растений // Науч. слово. 1928. № 1. С. 23–33.
- Жуковский П. М.* Ботаника. 5-е изд., перераб. и доп. М. : Колос, 1982. 623 с.
- Комаров В. Л.* Избранные сочинения. М. ; Л. : Изд-во АН СССР, 1951. Т. 7, ч. 1. 508 с.
- Ламарк Ж. Б.* Философия зоологии // Избр. произведения. В 2 т. Т. 1. М. : Изд-во АН СССР, 1955. С. 171–778.
- Лобаишѐв М. Е.* Генетика : учеб. пособие. М. : Просвещение, 1970. 431 с.
- Майр Э.* Популяции, виды и эволюция. М. : Мир, 1974. 460 с.
- Тахтаджян А. Л.* Флористические области Земли. Л. : Наука, 1978. 247 с.
- Тимофеев-Ресовский Н. В.* Генетика, эволюция, значение методологии в естествознании. Екатеринбург : Токмас-Пресс, 2009. 144 с.
- Шмальгаузен И. И.* Биологические основы возникновения наземных позвоночных // Изв. АН СССР. Сер. Биол. 1957. № 1. С. 3–30.
- Barba-Montoya J. A.* Bayesian molecular clock dating and the divergence times of angiosperms and primates : Dr. sci. diss. – UCL (University College London), 2017.
- Extensive variation in evolutionary rate of *rbcL* gene sequences among seed plants / J. Bousquet, S. H. Strauss, A. H. Doerksen, R. A. Price // PNAS. 1992. Vol. 89. N 16. P. 7844–7848. doi:10.1073/pnas.89.16.7844
- Testing the Efficacy of DNA Barcodes for Identifying the Vascular Plants of Canada / T. W. A. Braukmann, M. L. Kuzmina, J. Sills, E. V. Zakharov, P. D. N. Hebert // PLoS one. 2017. Vol. 12, N 1. P. e0169515.
- Chen K., Pachter L. Bioinformatics for whole-genome shotgun sequencing of microbial communities // PLoS Comput. Biol. 2005. P. 106-112.
- PLoS ONE / Chen S. (ed.). 2017. Vol. 12(1): e0169515. doi: 10.1371/journal.pone.0169515.
- Darwin C. On the Origin of Species (1882) // Books. Public Domain 8. URL: <https://itunes.apple.com/ru/book/on-the-origin-of-species/id395536758?mt=11>
- Dobzhansky T. Genetics of the Evolutionary Process. N. Y. : Columbia University Press, 1970.
- Doyle J.J., Dickson E. Preservation of Plant Samples for DNA Restriction Endonuclease Analysis // Taxon. 1987. Vol. 36. N 4. P. 715–722.
- Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products // J. Handelsman, M. R. Rondon, S. F. Brady, J. Clardy, R. M. Goodman // Chemistry&Biology. 1998. Vol. 5, N 10. P. 245–249.

Handelsman J. Metagenomics is not enough // DNA and cell biology. 2008. Vol. 27. P. 219–221.

Jombart T. adegenet: a R package for the multivariate analysis of genetic markers // Bioinformatics. 2008. Vol. 24. P. 1403–1405

Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities // Applied and environmental microbiology. 2009. Vol. 75. N 23. P. 7537–7541.

Medawar P. B. An Unresolved Problem in Biology. Lewis, 1952.

Nuchprayoon S., Junpee A., Poovorawan Y. Random amplified polymorphic DNA (RAPD) for differentiation between Thai and Myanmar strains of *Wuchereria bancrofti* // Filaria journal. 2007. Vol. 6, N 1. P. 6.

Ruiz E. A., Velarde E., Aguilar A. Demographic history of Heermann's Gull (*Larus heermanni*) from late Quaternary to present: Effects of past climate change in the Gulf of California // The Auk. 2017. Vol. 134, N 2. P. 308–316.

Schaffner S., Sabeti P. Evolutionary adaptation in the human lineage // Nature Education. 2008. Vol. 1(1). P. 14.

Schmalhausen I. I. Factors of Evolution: The Theory of Stabilizing Selection. Blakiston, Philadelphia, 1949.

Stebbins G. L. Self fertilization and population variability in higher plants // Am Nat. 1957. Vol. 91. P. 337–354.

The Angiosperm Phylogeny Group. An ordinal classification for the families of flowering plants // Missouri Botanical Garden Press Annals of the Missouri Botanical Garden. 1998. Vol. 85, N 4. P. 531–553. doi 10.2307/2992015.

Toulmin Stephen. The Philosophy of Science: An Introduction. London : Hutchinson. Amer. ed. New York : Rinehart, 1953.

High genetic diversity at the extreme range edge: nucleotide variation at nuclear loci in Scots pine (*Pinus sylvestris* L.) in Scotland // W. Wachowiak, M. J. Salmela, R. A. Ennos, G. Iason, S. Cavers // Heredity. 2011. Vol. 106, N 5. P. 775–787. doi:10.1038/hdy.2010.118.

Wright S. I., Kalisz S., Slotte T. Evolutionary consequences of self-fertilization in plants // Proc Biol Sci. 2015. 280 (1760). P. 20130133. doi: 10.1098/rspb. 2013.0133.

<http://www.efloras.org/> FNA Vol. 3 Betulaceae

Учебное издание

**АКТУАЛЬНЫЕ ПРОБЛЕМЫ
СОВРЕМЕННОЙ ГЕНЕТИКИ**
**ГЕНЕТИЧЕСКИЕ МЕТОДЫ
АНАЛИЗА БИОРАЗНООБРАЗИЯ**

Составители:

Щербаков Дмитрий Юрьевич
Харченко Виктория Евгеньевна

ISBN 978-5-9624-1588-8

Редактор *А. Н. Шестакова*
Дизайн обложки: *П. О. Еришов*

Темплан 2018. Поз. 29
Подписано в печать 03.05.2018. Формат 60x90 1/16
Уч.-изд. л. 5,0. Усл. печ. л. 7,8. Тираж 100 экз. Заказ 58

ИЗДАТЕЛЬСТВО ИГУ
664074, г. Иркутск, ул. Лермонтова, 124