**Лекция 1**

 **Задачи молекулярной эволюции как науки. Определение и гипотезы молекулярной эволюции.**

*Цель занятия:* Ознакомление обучающихся с целями и задачами молекулярной эволюции.

*Молекулярная эволюция* – это наука, изучающая изменения генетических макромолекул (DNA, RNA, белков) в процессе эволюции, закономерности и механизмы этих изменений, а также реконструирующая эволюционную историю генов и организмов.

*Молекулярная эволюция исследует:*

- нуклеотидные последовательности (DNA, RNA) как носители генетической информации;

- белковые последовательности;

- структуры белковых молекул;

- геномы организмов.

*Преимущества* данного подхода:

1) состав DNA - четыре типа нуклеотидов - универсален для любых групп организмов (бактерии, растения, животные и др.);

2) изменения DNA в ходе эволюции регулярны, поэтому могут быть описаны математически (в том числе при сравнении DNA филогенетически отдаленных организмов);

3) геномы организмов содержат значительно больше филогенетической информации, чем морфологические признаки.

*Молекулярная эволюция как наука* включает в себя два направления исследований. *Ее целями* являются изучение законов изменения наследственной информации в живых системах, включая доклеточные и клеточные формы жизни, и изучение истории развития жизни на Земле, установление родственных отношений между формами жизни, их генетическим материалом— филогении форм жизни (phylogeny). Первое направление включает в себя изучение частоты и других характеристик эволюционных изменений в макромолекулах — нуклеиновых кислотах и белках, а также механизмов и причин, определяющих эти изменения.

Вторая область исследований молекулярной эволюции, направленная на реконструкцию эволюционного процесса развития форм жизни на Земле, установление родственных связей между формами жизни, создание их эволюционной классификации, называется молекулярной филогенетикой (phylogenetics). Традиционно изучение пребиотической эволюции, происхождения жизни, также понимается как задача молекулярной эволюции.

Однако механизмы пребиотической эволюции принципиально отличаются от законов эволюции живых систем, требуют специального рассмотрения и, как правило, описываются отдельно. Достижения молекулярной биологии с момента открытия строения нуклеиновых кислот позволили изучать эволюционные связи между организмами путем сравнения их нуклеотидных последовательностей.

***Гипотеза молекулярных эволюционных часов*** (*L. Pauling, E. Margoliash*): для каждого данного белка темп молекулярной эволюции примерно постоянный во всех эволюционных линиях. Если последовательность белка эволюционирует с постоянной скоростью, она может быть оценена временем расхождения видов. Это аналог датирования геологических образцов по радиоактивному распаду. Другими словами, в пределах каждого набора гомологичных последовательностей частота замен постоянна.

***Теория нейтральной молекулярной эволюции*** (*M. Kimura*) предполагает, что большинство изменений DNA не связаны с дарвиновским отбором. Значительно чаще фиксация мутаций происходит в результате случайного дрейфа генов и является селективно нейтральной (или слабо отрицательной). Селективная элиминация вредных мутаций и случайная фиксация селективно нейтральных или слабо отрицательных мутаций происходит в ходе эволюции гораздо чаще, чем положительный дарвиновский отбор благоприятных мутаций.

***Теория направленного мутационного давления*** (*N. Sueoka*) выделяет в качестве основной причины генных мутаций фактор, обусловленный повышенной частотой возникновения и фиксации замен А и Т на Г и Ц относительно частоты возникновения и фиксации замен Г и Ц на А и Т (ГЦ-давление или АТ-давление). Наиболее вероятными причинами возникновения мутационного давления являются ферментативное и спонтанное дезаминирование нуклеотидов и возникновение ошибок в процессе репликации и репарации DNA.

С использованием данных теорий и их интеграции разрабатывается большое количество методов молекулярной эволюции и филогенетики с последующим определением их эффективности при помощи компьютерного моделирования.

Молекулярная эволюция включает:

- эволюцию макромолекул - изучает типы и скорости изменений, происходящих в генетическом материале (DNA, RNA), а также образованных на его основе белков и механизмы, ответственные за эти изменения;

- молекулярную филогению - изучает эволюционную историю макромолекул и организмов, получаемую на основе изучения нуклеотидных и аминокислотных последовательностей.

Для продолжения работы требуется

# Контрольные вопросы:

1. Генная инженерия.
2. Цели и задачи генной инженерии.
3. История развития технологий и генной инженерии.

**Лекция 2**

 **Развитие геномных исследований.**

Цель занятия: Ознакомление обучающихся с развитием геномных исследований.

Первым был полностью секвенирован геном бактериофага Φ-X174; (5368 нуклеотидов) в 1977 году (Sanger et al., 1977), еще до разработки основных современных методов секвенирования ДНК, первой модификацией метода Сэнгера. Необходимость получения данных о последовательностях генов привела к быстрому совершенствованию методов секвенирования ДНК, а это, в свою очередь, дало возможность секвенировать не только отдельные гены, но и протяженные районы генома.

Первым полностью секвенированным геномом свободноживущего организма стал геном патогенной бактерии – Haemophilus influenzae (1,8 Mб; 1995 год) (Fleischmann et al., 1995). В последующие несколько лет количество полностью секвенированных бактериальных геномов начало стремительно расти. Однако по-настоящему ключевым событием в формировании геномики явился проект секвенирования генома человека.

Проект Геном человека (англ. The Human Genome Project, HGP) — международный научно-исследовательский проект, главной целью которого было определить последовательность нуклеотидов, которые составляют ДНК, и идентифицировать 20—25 тыс. генов в человеческом геноме. Этот проект называют крупнейшим международным сотрудничеством, когда-либо проводившимся в биологии.

Проект начался в 1990 году, под руководством Джеймса Уотсона под эгидой Национальной организации здравоохранения США. В 2000 году был выпущен рабочий черновик структуры генома, полный геном — в 2003 году, однако и сегодня дополнительный анализ некоторых участков ещё не закончен. Частной компанией Celera Corporation был запущен аналогичный параллельный проект, завершённый несколько ранее международного. Основной объём секвенирования был выполнен в университетах и исследовательских центрах США, Канады и Великобритании. Кроме очевидной фундаментальной значимости, определение структуры человеческих генов является важным шагом для разработки новых медикаментов и развития других аспектов здравоохранения.

Развитие проекта геном человека стало возможно не только благодаря совершенствованию биохимических методик, но и благодаря появлению более мощной вычислительной техники, которая позволила работать с огромными массивами данных. Протяженность геномов у живых организмов подчас измеряется миллиардами пар оснований. Например, объём генома человека составляет порядка 3 млрд пар оснований. Самый крупный из известных (на начало 2010 года) геномов принадлежит одному из видов двоякодышащих рыб (примерно 110 млрд пар).

На выполнение этой задачи было потрачено более полутора десятилетий и более 3 млрд долларов. Международный проект стартовал в 1990 г., через некоторое время после этого корпорацией Celera был aнонсирован аналогичный частный проект, обошедшийся гораздо дешевле – в 300 млн долларов, но частично использовавший результаты международного. Первые приблизительные результаты обоих проектов были опубликованы практически одновременно в 2001 г. (International Human Genome…, 2001; Venter et al., 2001).

Однако, официально завершенным международный проект считается после опубликования первичной структуры последней из секвенированных хромосом – хромосомы I (Gregory et al., 2006), и результаты его выполнения многократно окупили все затраты.

Во-первых, были созданы и отработаны технологии экспериментального секвенирования полных эукариотических геномов, состоящих из миллиардов пар нуклеотидов, что позволило радикально ускорить последующие геномные проекты.

Во-вторых, были разработаны совершенно новые биоинформатические алгоритмы и программы, позволяющие оперировать огромными объемами данных и проводить сборку протяженных последовательностей из относительно коротких секвенированных фрагментов. С дальнейшим развитием и удешевлением технологий секвенирования появилась возможность начать большое число новых геномных проектов, и в результате это привело к взрывному росту объема получаемых первичных данных.

Современная геномика, в первую очередь, ориентирована на анализ данных больших геномных проектов. Более простым их типом является секвенирование бактериальных геномов, поскольку бактериальный геном относительно невелик и, как правило, содержит малое число повторяющихся последовательностей (Gregory, 2005). На начало 2014 г. в базе данных GOLD зарегистрировано более 32 000 проектов по секвенированию индивидуальных геномов бактерий. Из них полноценно законченными являются только 2 711, основная масса – более 60 % – находится на различных этапах выполнения, а 9 591 – т. е. почти треть, брошена на промежуточной стадии (permanent draft).

**Вопросы для самоконтроля:**

1. История развития геномных технологий.

2. Проект геном человека

3. Современное состояние и перспективы развития геномных технологий.

**Лекция 3**

**Мутации. Нуклеотидные замены. Транзиции и трансверсии.**

*Цель занятия:* Ознакомление обучающихся с видами мутации: нуклеотидные замены, транзиции и трансверсии.

В процессе репликации нуклеиновых кислот могут происходить ошибки. В результате дочерний геном будет отличаться от родительского генома. Ошибки, происходящие при репликации генома, называют мутациями (mutations). У многоклеточных организмов мутации могут происходить как в половых, так и в соматических клетках. Соматические мутации не передаются от поколения к поколению, и, таким образом, исключены из эволюционного процесса.

Мутации могут быть классифицированы в соответствии с числом затронутых ими нуклеотидов (длиной мутации) и типами происходящих при этом событий. Мутации могут затрагивать только один нуклеотид — точечные мутации (point mutations) или несколько соседних нуклеотидов.

В кодирующих и не кодирующих участках нуклеиновых кислот нуклеотидные замены могут быть разделены на транзиции (transitions) и трансверсии (transversions).

**Вопросы для самоконтроля:**

1. Мутации.

2. Нуклеотидные замены.

3. Транзиции и трансверсии.

**Лекция 4**

**Молекулярная эволюция.**

**Исследование механизмов изменения геномов.**

*Цель занятия:* Ознакомление обучающихся с целями и задачи молекулярной эволюции. Исследование механизмов изменения геномов.

Молекулярная филогенетика — способ установления родственных связей между живыми организмами на основании изучения структуры полимерных макромолекул — ДНК, РНК и белков. Результатом молекулярно-филогенетического анализа является построение филогенетического дерева живых организмов.

Близкое родство между живыми организмами обычно сопровождается большой степенью сходства в строении тех или иных макромолекул, а молекулы не родственных организмов сильно различаются между собой. Молекулярная филогения использует такие данные для построения филогенетического древа, которое отражает гипотетический ход эволюции исследуемых организмов. Возможность анализировать и подробно изучать эти молекулы появилась только в последние десятилетия XX века.

Молекулярная филогенетика оказала сильнейшее влияние на научную классификацию живых организмов. Методы работы с макромолекулами стали доступны биологам самых различных специальностей, что привело к лавинообразному накоплению новой информации о живых организмах. На основании этих данных старые предположения об эволюции живых организмов пересматриваются. Описывают новые группы, в том числе, выделяемые только на основе молекулярно-филогенетических данных.

**Вопросы для самоконтроля:**

1. Молекулярная эволюция.

2. Методы, основанные на анализе генетических дистанций.

**Лекция 5**

**Методы секвенирования ДНК.**

**Сравнительная характеристика NGS и NNGS секвенирования.**

*Цель занятия:* Ознакомление обучающихся с методами секвенирования ДНК. Сравнительная характеристика NGS и NNGS секвенирования.

Секвенирование нового поколения (англ. next generation sequencing, NGS) — группа методов определения нуклеотидной последовательности ДНК и РНК для получения формального описания её первичной структуры. Технология методов секвенирования нового поколения позволяет «прочитать» единовременно сразу несколько участков генома, что является главным отличием от более ранних методов секвенирования. NGS осуществляется с помощью повторяющихся циклов удлинения цепи, индуцированного полимеразой, или многократного лигирования олигонуклеотидов. В ходе NGS могут генерироваться до сотен мегабаз и гигабаз нуклеотидных последовательностей за один рабочий цикл.

Первая концепция секвенирования была предложена Сенгером в 1977 году. Технология получила название «метод обрыва цепи». В том же году Максам и Гилберт предложили альтернативный метод, получивший название «метод химической деградации[англ.]» — в его основе лежит расщепление меченого по одному концу фрагмента ДНК под действием специфических реагентов. Определение нуклеотидной последовательности проводится методом электрофореза в полиакриламидном геле с последующей авторадиографией. Необходимость в массовом, качественном и быстром секвенировании стимулировала многочисленные модификации и всевозможные улучшения этих методов. В той или иной степени изменениям подверглись практически все составляющие этого процесса. Переломной точкой развития технологии стало появление ПЦР (середина 1980-х) и автоматизация основных этапов «чтения» ДНК, давшие начало методам секвенирования следующего поколения. Платформы для методов нового поколения основываются на распараллеливании процесса «чтения» ДНК, и таким образом за один прогон работы секвенатора можно определить первичные структуры нескольких участков генома. Секвенаторы нового поколения стали значительно дешевле и гораздо эффективнее своих предшественников. На сегодняшний день производительность некоторых секвенаторов измеряется уже сотнями миллиардов пар оснований, что, например, позволяет подобным приборам сканировать индивидуальный геном человека всего за несколько дней

**Вопросы для самоконтроля:**

1. Методы секвенирования ДНК.
2. Сравнительная характеристика NGS и NNGS секвенирования.

**Лекция 6**

**Молекулярная филогенетика: цели и задачи.**

*Цель занятия:* Ознакомление обучающихся с целями и задачами молекулярной филогенетики.

Молекулярная филогенетика — способ установления родственных связей между живыми организмами на основании изучения структуры полимерных макромолекул — ДНК, РНК и белков. Результатом молекулярно-филогенетического анализа является построение филогенетического дерева живых организмов.



Филогенетическое дерево, построенное с помощью методов молекулярной филогенетики

Близкое родство между живыми организмами обычно сопровождается большой степенью сходства в строении тех или иных макромолекул, а молекулы не родственных организмов сильно различаются между собой. Молекулярная филогения использует такие данные для построения филогенетического древа, которое отражает гипотетический ход эволюции исследуемых организмов. Возможность анализировать и подробно изучать эти молекулы появилась только в последние десятилетия XX века.

Молекулярная филогенетика оказала сильнейшее влияние на научную классификацию живых организмов. Методы работы с макромолекулами стали доступны биологам самых различных специальностей, что привело к лавинообразному накоплению новой информации о живых организмах. На основании этих данных старые предположения об эволюции живых организмов пересматриваются. Описывают новые группы, в том числе, выделяемые только на основе молекулярно-филогенетических данных.

**Контрольные вопросы:**

1. ГИ растений и животных: объекты, основные направления и задачи.

2. Биотехнология в животноводстве: состояние и актуальные проблемы.

3. Вопросы биоэтики в биотехнологии животных.

4. Регулирование выпуска генетически модифицированных организмов.

**Лекция 7**

**Гипотеза молекулярных эволюционных часов (L. Pauling, E. Margoliash).**

*Цель занятия:* Ознакомление обучающихся с Гипотеза молекулярных эволюционных часов (L. Pauling, E. Margoliash).

Гипотеза молекулярньгх эволюционных часов (L. Pauling, E. Margoliash): для каждого данного белка темп молекулярной эволюции примерно постоянный во всех эволюционных линиях. Если последовательность белка эволюционирует с постоянной скоростью, она может быть оценена временем расхождения видов. Это аналог датирования геологических образцов по радиоактивному распаду. Другими словами, в пределах каждого набора гомологичных последовательностей частота замен постоянна.

Молекулярные часы (англ. molecular clock, иногда gene clock, evolutionary clock) — метод датирования филогенетических событий (расхождений видов или других таксонов), основанный на гипотезе, согласно которой эволюционно значимые замены мономеров в биомолекулах происходят с практически постоянной скоростью (molecular clock hypothesis). Обычно для подобных вычислений используются нуклеотидные последовательности ДНК и аминокислотные последовательности белков.

Скорость мутаций может быть неравномерной и различается для разных видов, из-за чего метод дает лишь приблизительные результаты.

Гипотеза молекулярных часов была выдвинута в 1962 году при анализе аминокислотных последовательностей гемоглобина и цитохрома С Эмилем Цукеркандлем[англ.] и Лайнусом Полингом. Они отметили, что количество аминокислотных различий в гемоглобине растет линейно со временем, которое оценивалось по фоссилиям[1]. Они обобщили наблюдение и пришли к выводу, что скорость эволюционного изменения каждого белка приблизительно постоянна.

В 1963 году Эмануэлем Марголиашем[англ.] был обнаружен феномен «генетической эквидистантности» (genetic equidistance), заключающийся в независимости эволюции аминокислотных последовательностей в белках и морфологической эволюции[2]:

Полезной проверкой важной роли времени как главного фактора в накоплении изменчивости в цитохроме C должно быть сравнение аминокислотных последовательностей гомологичных белков, выделенных из видов, о которых известно, что они на протяжении длительных периодов времени не претерпевали морфологических изменений, и из быстро изменяющихся видов.

Работы этих трёх ученых привели к постулированию гипотезы в начале 1960-х.

**Контрольные вопросы:**

1. Молекулярные часы.

2. Выдвижение теории и её развитие.

**Лекция 8**

**Теория нейтральной молекулярной эволюции (M. Kimura).**

*Цель занятия:* Ознакомление обучающихся с теорией нейтральной молекулярной эволюции (M. Kimura).

 Согласно нейтральной теории молекулярной эволюции, подавляющее число мутаций носит случайный характер, а «выживание наиболее приспособленных» — не только не единственный, но и даже не наиболее распространенный способ эволюции видов.

Эта теория кажется очень простой, но ее влияние на генетику, экологию и эволюционную биологию трудно переоценить.

Не совсем нейтральная теория

Процесс случайных изменений вариантов генов (аллелей) в популяции называется дрейфом генов. Хотя сегодня дрейф генов считается главной движущей силой эволюции, до 1960-х годов биологи объясняли разнообразие действием отбора: вредоносные гены, которые препятствуют продолжению рода, со временем исчезают (отрицательный отбор); а полезные гены, способствующие умножению потомства, сохраняются (позитивный отбор) — всё согласно принципу естественного отбора, изложенному Чарлзом Дарвином и Альфредом Расселом Уоллесом.

Затем секвенирование белков показало намного больше наследственных изменений в пределах популяций, чем ожидалось. Некоторые ученые засомневались, что все эти гены одновременно подвергались отбору.

В 1968 году знаменитый генетик Мотоо Кимура предложил иное объяснение, которое позже получило название нейтральной теории молекулярной эволюции. Кимура считал, что большинство генетических изменений не увеличивают и не уменьшают шансы видов на выживание и произведение потомства. Соответственно, разнообразие обусловливается преимущественно не отбором, а волей случая.

**Контрольные вопросы:**

1. Теория нейтральной молекулярной эволюции.

3. История возникновения нейтральной теории и ее развитие.

**Лекция 9**

**Проведение эволюционного анализа.**

*Цель занятия:* Ознакомление студентов с проведением эволюционного анализа.

Генетическая информация в живых системах представлена молекулами нуклеиновых кислот, последовательностями нуклеотидов. При репликации этих молекул могут происходить ошибки, мутации, в т ом числе — замены, вставки и делеции нуклеотидов. В результате дочерние геномы будут отличаться от родительских.

В результате всех этих событий имеющаяся у живых систем генетическая информация будет различной, их нуклеотидные последовательности будут отличаться друг от друга.

Реконструкцией родственных отношений между формами жизни на основании анализа их генетических последовательностей, установлением законов эволюционных изменений наследственной информации занимается молекулярная филогенетика — ключевой раздел молекулярной эволюции.

**Контрольные вопросы:**

1. Дивергенция между эволюционными линиями.

2. Проведение эволюционного анализа.

**Лекция 10**

**Естественный отбор и неодарвинизм.**

*Цель занятия:* Ознакомление студентов с проведением эволюционного анализа.

Любые мутации в родительских геномах не изменяют приспособленности (fitness, adaptation) дочерних геномов к окружающей среде — характеристик, сводящихся в конечном счете к плодовитости родителей и их потомства. Соответственно мы подразумевали, что все мутации имеют одинаковые, 100%-ные, вероятности быть переданными дочерним геномам и закрепиться в популяции потомков.

Однако сформулированная Дарвином теория естественного отбора исходит из того, что вероятность закрепления (отбора) некоего наследственного признака в популяции зависит от качеств этого признака — от того, насколько он способствует преимущественному выживанию и преимущественному производству плодовитого потомства особью, этим признаком обладающей.

**Контрольные вопросы:**

1. Теория естественного отбора.

2. Неодарвинизм.

**Лекция 11**

**Методы построения филогенетических деревьев**

**в молекулярной филогенетике.**

*Цель занятия:* Ознакомление обучающихся с методами построения филогенетических деревьев.

Методы построения филогенетических деревьев в молекулярной филогенетике. Существует большое количество методов построения филогении на основании молекулярных данных. Их можно подразделить на два типа:

Методы, использующие оценку генетических дистанций

Методы, использующие анализ дискретных признаков

Методы, основанные на анализе генетических дистанций

Данная группа методов базируется на данных о генетических дистанциях. Общий принцип заключается в попарном сравнении объектов и построении матрицы дистанций, которая затем используется для построения филогенетического дерева.

**Контрольные вопросы:**

1. Методы, использующие оценку генетических дистанций.

2. Методы, использующие анализ дискретных признаков.

**Лекция 12**

**Анализ митохондриальной ДНК.**

*Цель занятия:* Ознакомление обучающихся с анализом митохондриальной ДНК.

Митохондриальные заболевания клинически представляют собой гетерогенную группу состояний, развивающихся в связи с нарушением работы ферментов дыхательной цепи митохондрий. Нарушения могут быть связаны как с мутациями, появляющимися в геноме человека, так и с мутациями в митохондриальной ДНК (мтДНК). Наиболее часто наблюдаются повреждения мтДНК, которые характеризуются протяженными делециями и дупликациями, а также точечными мутациями (чаще всего 3243A>G, 3460G>A, 8344A>G, 11778G>A, 14484T>C).

Митохондриальные заболевания могут манифестировать в любом возрасте. При некоторых митохондриальных заболеваниях поражается только один орган (наследственная оптическая нейропатия Лебера), однако в большинстве данная группа состояний имеет системные проявления, чаще всего связанные с поражением нервной и мышечной системы. Довольно часто клинические проявления митохондриальных заболеваний могут подпадать под специфические синдромы: MELAS-синдром (митохондриальная миопатия, энцефалопатия, лактат-ацидоз, инсультоподобные эпизоды), MERRF-синдром (миоклоническая эпилепсия, ассоциированная с порванными красными волокнами), прогрессирующая наружная офтальмоплегия, синдром Кернса-Сейра, синдром Лея, нейрогенная слабость с атаксией и пигментным ретинитом. Однако довольно много клинических случаев митохондриальных заболеваний не подпадают ни под одну из перечисленных категорий, имея перемежающийся спектр проявлений и симптомов, что возможно объясняется гетероплазмией мутаций в мтДНК (вариация количества «мутантных» митохондрий в одной клетке).

**Контрольные вопросы:**

1. Анализ митохондриальной ДНК.

2. Митохондриальные заболевания.

**Лекция 13**

**Филогенетическое дерево.**

*Цель занятия:* Ознакомление обучающихся с построением филогенетического дерева.

Филогенетическое дерево (эволюционное дерево, дерево жизни) — дерево, отражающее эволюционные взаимосвязи между различными видами или другими сущностями, имеющими общего предка.

Вершины филогенетического дерева делятся на три класса: листья, узлы и (максимум один) корень. Листья — это конечные вершины, то есть те, в которые входят ровно по одному ребру; каждый лист отображает некоторый вид живых организмов (или иной объект, подверженный эволюции, например, домен белка). Каждый узел представляет эволюционное событие: разделение предкового вида на два или более, которые в дальнейшем эволюционировали независимо. Корень представляет общего предка всех рассматриваемых объектов. Ребра филогенетического дерева принято называть «ветвями».

Идея «дерева» появилась в ранних взглядах на жизнь, как на процесс развития от простых форм к сложным. Современные эволюционные биологи продолжают использовать деревья для иллюстрации эволюции, так как они наглядно показывают взаимосвязи между живыми организмами.

*Типы филогенетических деревьев*

Укоренённое дерево — дерево, содержащее выделенную вершину — корень. Укоренённое дерево можно считать ориентированным графом, поскольку на нём имеется естественная ориентация — от корня к листьям. Каждый узел укоренённого дерева отвечает последнему общему предку нижележащих листьев дерева. На иллюстрации показано укоренённое филогенетическое дерево, окрашенное в соответствии с трёхдоменной системой живых организмов.

Неукоренённое дерево не содержит корня и отражает связь листьев без предполагаемого положения общего предка. Необходимость рассматривать неукоренённые деревья возникает из-за того, что часто связи между узлами восстановить легче, чем направление эволюции. На иллюстрации показано неукоренённое филогенетическое дерево. Наиболее достоверным методом для превращения неукоренённого дерева в укоренённое (для этого надо либо объявить корнем один из узлов, либо разбить одну из ветвей на две, выходящие из корня) является использование «внешней группы» видов — достаточно близких к интересующему нас набору видов (для достоверного восстановления топологии дерева для объединённого множества видов), но в то же время заведомо являющихся отдельной группой. Иногда положение корня можно угадать, исходя из каких-либо дополнительных знаний о природе изучаемых объектов (видов, белков, и так далее).

Укоренённое и неукоренённое филогенетическое дерево может быть бифуркационным или небифуркационным, а также маркированным или немаркированным. В бифуркационном дереве к каждому узлу подходят ровно три ветви (в случае укоренённого дерева — одна входящая ветвь и две исходящие). Таким образом, бифуркационное дерево предполагает, что все эволюционные события состояли в происхождении от предкового объекта ровно двух потомков. К узлу небифуркационного дерева могут подходить четыре и более ветви. Маркированное дерево содержит названия листьев, тогда как немаркированное просто отражает топологию.

Дендрограмма — общий термин, обозначающий схематическое представление филогенетического дерева.

Кладограмма — филогенетическое дерево, не содержащее информации о длинах ветвей.

Филограмма — филогенетическое дерево, содержащее информацию о длинах ветвей; эти длины представляют изменение некой характеристики, например, число мутаций в каком-либо гене.

Хронограмма — филограмма, длины ветвей в которой представляют эволюционное время.

**Контрольные вопросы:**

1. Филогенетическое дерево.

2. Типы филогенетических деревьев.

**Лекция 14**

**Международные базы генетических данных.**

*Цель занятия:* Ознакомление обучающихся с международными базами генетических данных.

GenBank — база данных, находящаяся в открытом доступе, содержащая все аннотированные последовательности ДНК и РНК, а также последовательности закодированных в них белков. GenBank поддерживается Национальным центром биотехнологической информации США (NCBI), входящего в состав Национальных Институтов Здоровья в США, и доступен на бесплатной основе исследователям всего мира. GenBank получает и объединяет данные, полученные в разных лабораториях, для более чем 100 000 различных организмов.

GenBank — архивная база данных, то есть ответственность за содержимое каждой записи несут создатели этой записи, которыми, как правило, являются экспериментаторы, определившие данную последовательность. GenBank вместе с банками EMBL и DDBJ входит в консорциум INSDC (http://insdc.org/), осуществляющий регулярный обмен данными между этими тремя архивами аннотированных нуклеотидных последовательностей.

Релиз GenBank происходит каждые два месяца и доступен с сайта по протоколу FTP. Заметки о выпуске для текущей версии GenBank предоставляют подробную информацию о выпуске и уведомлениях о предстоящих изменениях в GenBank. Также доступны примечания к выпуску предыдущих версий GenBank.

В настоящий момент база GenBank, помимо последовательностей отдельных генов, содержит много данных, полученных с помощью современных методов секвенирования ДНК и автоматического аннотирования последовательностей. Существует несколько разделов GenBank, посвящённых данным высокопроизводительного секвенирования.

Genomes – специальный раздел для хранения полных геномов. Созданы руководства по аннотации полных геномов прокариот и эукариот.

WGS (Whole genome shoutgun) – проекты по сборке неполных геномов, хромосом прокариот или эукариот, главным образом, секвенированных методом дробовика. В GenBank аннотация проектов WGS необязательна, однако NCBI располагает специальным пайплайном для аннотации прокариотических геномов. Существует список доступных WGS-проектов.

TPA (Third Party Annotation) – представляет собой базу данных экспериментальных или выведенных из уже имеющихся данных результатов, аннотация которых не произведена автором из первичных данных, а определена по косвенным. Записи TPA делятся, соответственно, на две категории:

experimental – аннотация последовательностей подтверждена экспериментальным доказательством в «мокрой» лаборатории.

inferential – аннотация последовательностей сделана путём умозаключения из доступной информации. При этом непосредственно молекула нуклеиновой кислоты или её продукт(ы) не являлись предметами прямых экспериментов.

TSA (Transcriptome Shotgun Assembly sequences) – последовательности транскриптомов, полученные путём секвенирования методом дробовика. Данный раздел содержит данные, собранные из последовательностей, размещённых в NCBI Trace Archieve, Sequence Read Archive и разделе GenBank EST. Отдел TSA представляет собой один из самых быстрорастущих разделов GenBank.

**Контрольные вопросы:**

1. GenBank — архивная база данных.

2. Разделы GenBank.

**Лекция 15**

**Молекулярная эволюция и филогенетика.**

*Цель занятия:* Ознакомление обучающихся с принципами молекулярной эволюции и филогенетика.

Молекулярная эволюция (англ. molecular evolution) — наука, изучающая процесс изменения последовательностей мономеров в биополимерных молекулах в живых организмах, а именно ДНК, РНК и белков[1]. Молекулярная эволюция опирается на принципы эволюционной биологии, молекулярной биологии и популяционной генетики. Задача молекулярной эволюции состоит в объяснении закономерностей таких изменений. Молекулярная эволюция занимается механизмами накопления изменений молекулами, и механизмами закрепления этих изменений в популяциях, а также проблемами видообразования.

Филогене́тика или филогенети́ческая система́тика — область биологической систематики, которая занимается выявлением и прояснением эволюционных взаимоотношений среди разных видов жизни на Земле, как современных, так и вымерших. Эволюционная теория утверждает, что сходство тех или иных особей или видов часто указывает на общее происхождение или общего предка. Потому взаимоотношения, установленные филогенетической систематикой, часто описывают эволюционную историю видов и их филогенез, исторические взаимоотношения между ветвями организмов или их частей, например, их генов. Филогенетическая таксономия, являющаяся ответвлением, но не логическим продолжением филогенетической систематики[1], занимается классификацией групп организмов согласно степени их эволюционных отношений.

Основателем систематики, области науки, которая занимается классификацией живых организмов и взаимоотношениями между компонентами живого, считается Карл Линней. Однако только в конце 1950-х годов немецкий энтомолог Вилли Хенниг высказал идею, что систематика должна отображать известную эволюционную историю так близко, как только возможно[2]. Так был основан подход к систематике, который он назвал филогенетической систематикой. Противники Хеннига пренебрежительно называли его последователей «кладистами»[3], из-за акцента на признание только монофилетичных групп или клад. Однако кладисты быстро приняли это название как полезный термин, и кладистический подход начал преобладать в систематике. Противоположностью филогенетической систематики является фенетика.

**Контрольные вопросы:**

1. Молекулярная эволюция.

2. Филогенетика.