

Филогенетические деревья

ФББ, IV семестр,
весна 2025
С.А.Спирин

«...великое Дерево Жизни заполняет земную кору своими мертвыми и сломанными ветвями и покрывает поверхность вечно ветвящимися и прекрасными побегами»

Ч. Дарвин

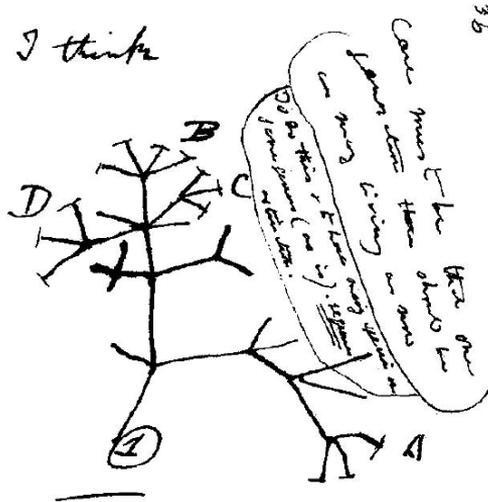


Рисунок из тетради «Notebook B»
Чарльза Дарвина, 1837

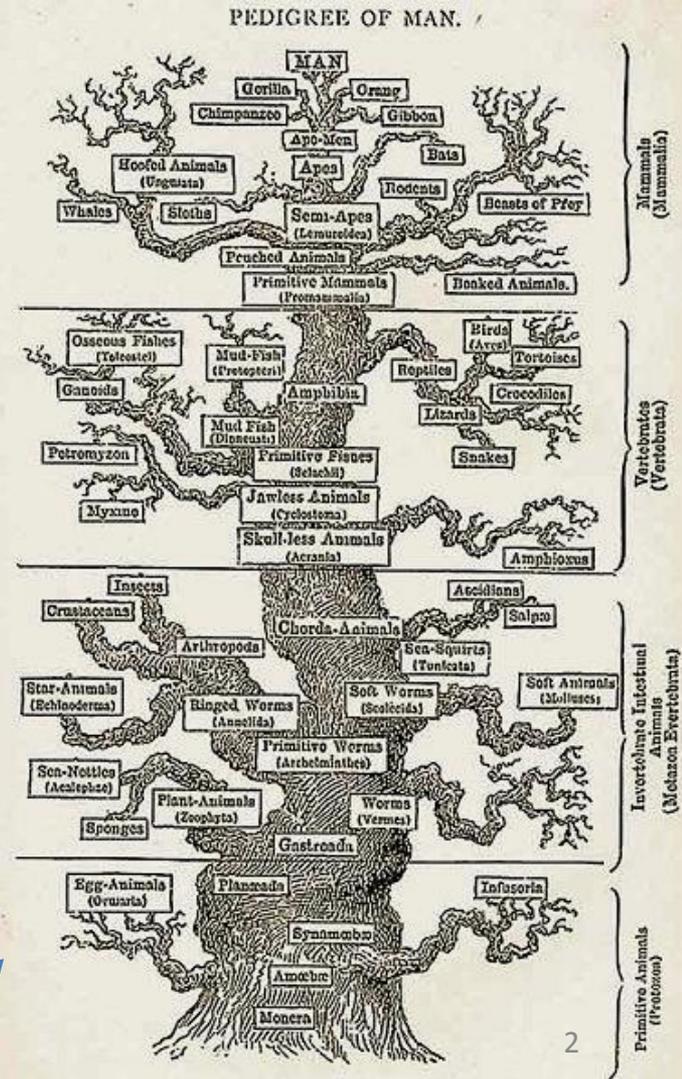
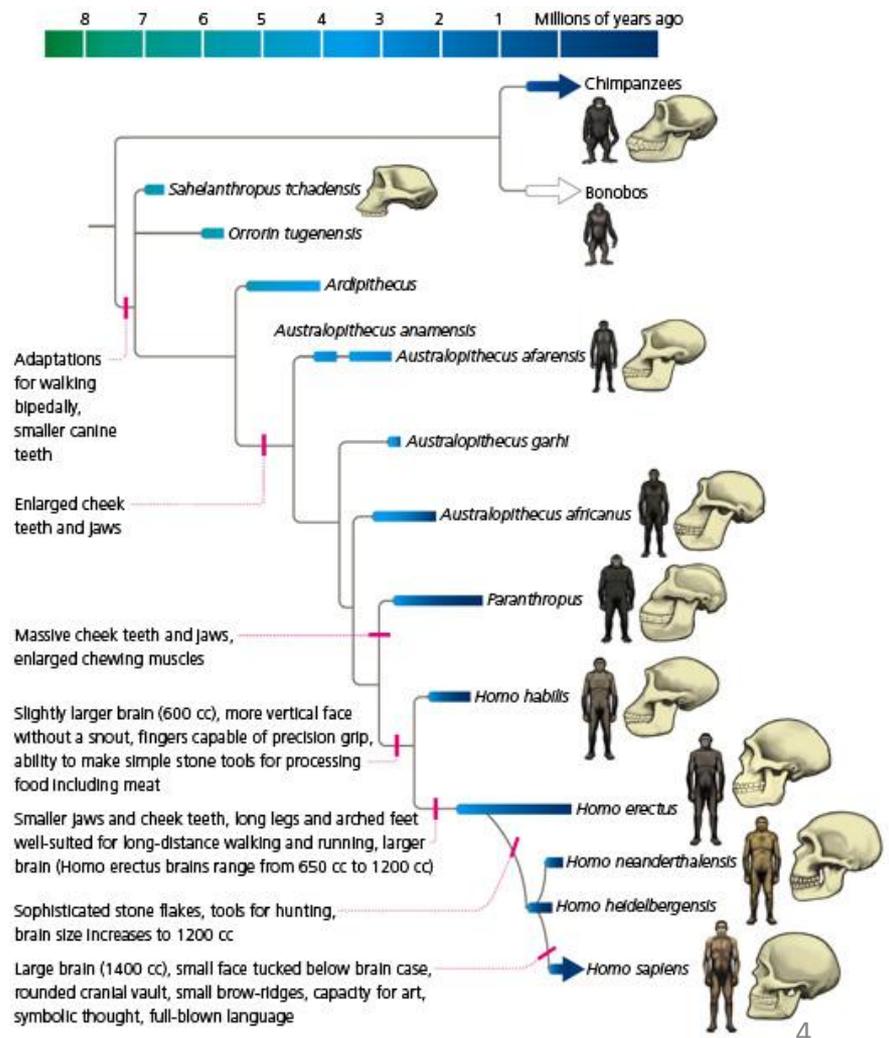
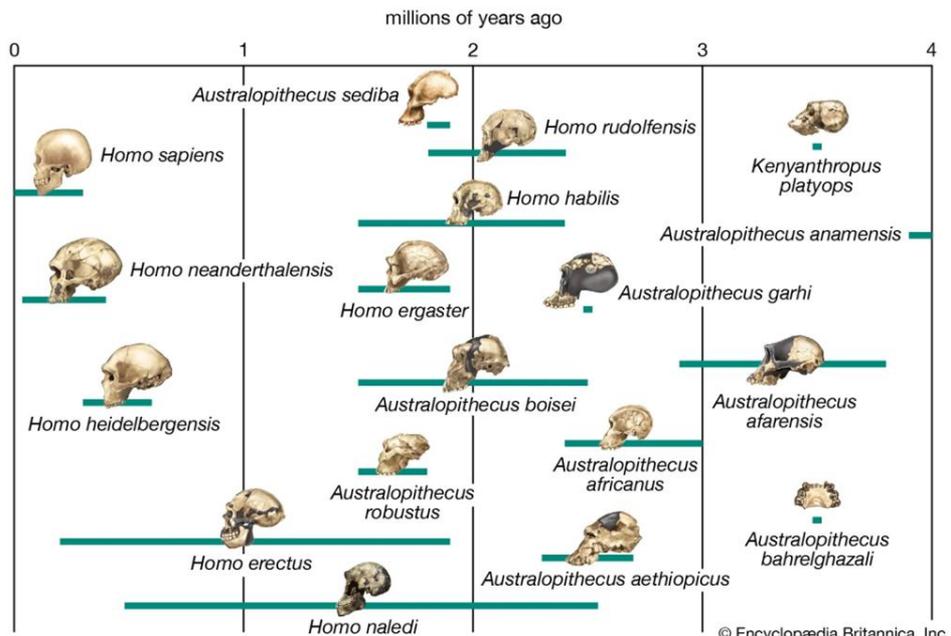


Рисунок из книги Эрнеста Геккеля «Эволюция человека», 1879

По каким данным можно строить дерево?

До эпохи массового секвенирования ответ был очевидным: по фенотипу, т. е. по морфологическим признакам

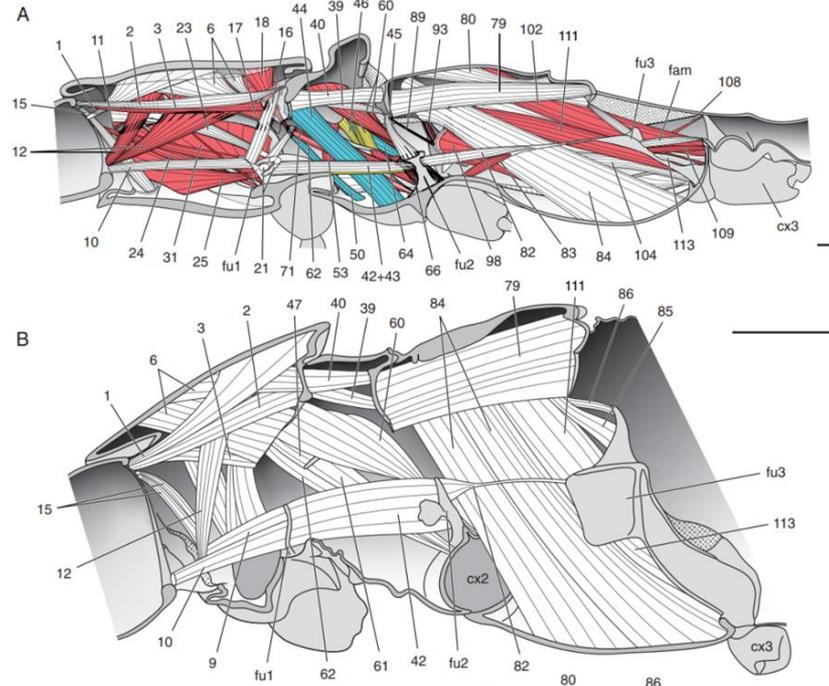
Эволюция человека восстанавливается по особенностям черепов (и других костей)



Морфологические признаки жуков

190 признаков

Comparison of the thoracic muscular system of representatives of: (A) Archostemata (Tetrphalarus, Ommatidae; modified from Friedrich et al., 2009); (B) Polyphaga (Helophorus, Helophoridae; redrawn from Beutel & Komarek, 2004)



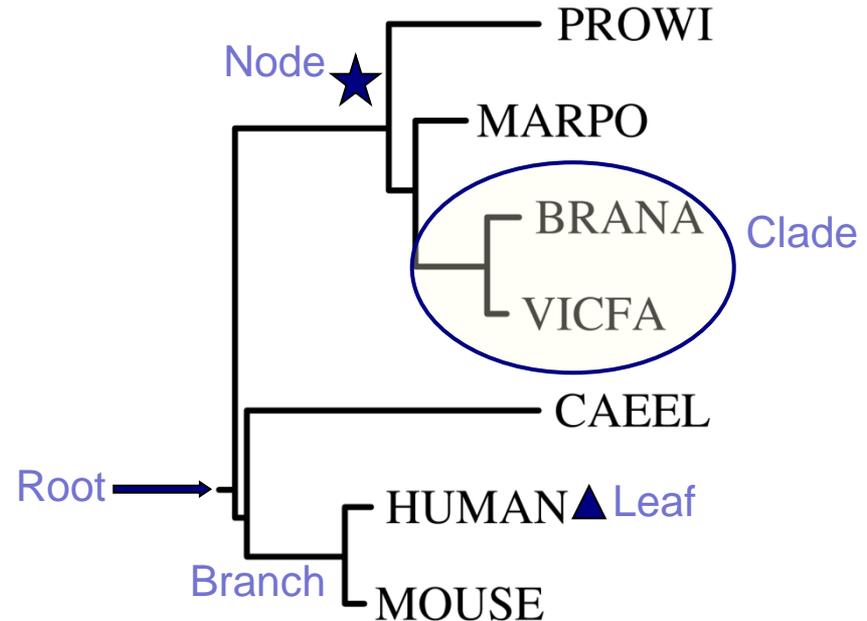
Матрица признаков

https://morphobank.org/index.php/projects/viewMatrix/matrix_id/30408/streaming/0/project_id/4162

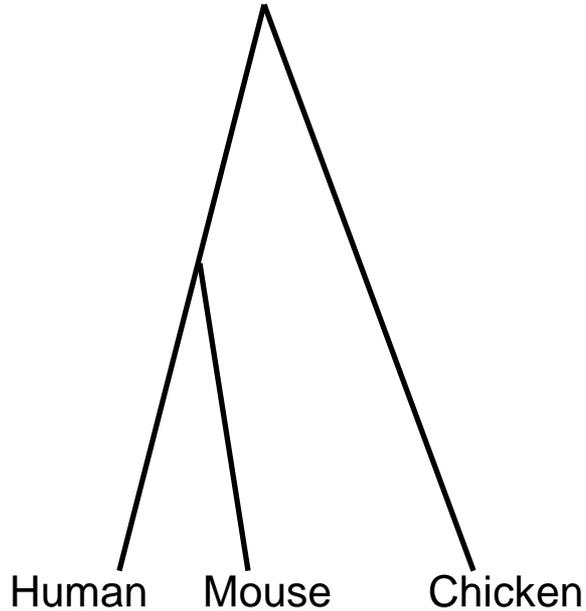
	[1] size of primary larva	[2] dorsal endocarina	[3] larval tentorium	[4] caudal tentorial arms	[5] orientation of the head	[6] head shape	[7] deep dorsal and ventral posteromedian emargination	[8] hemispherical projection of head capsule	[9] posterior tentorial grooves	[10] larval gula	[11] retinula cells
[1] <i>Zorotypus</i>	0	0	0	0	0	0	0	0	0	0	-
[2] Raphidiidae	0	0	0	0	1	0	0	0	1	1	1
[3] <i>Nevrorthus</i>	0	0	0	0	1	0	0	0	1	1	1
[4] <i>Myrmeleon</i>	0	0	0	0	1	0	0	0	0	0	1
[5] <i>Sialis</i>	0	0	0	0	0	0	0	0	0	0	1
[6] Corydalinae	0	0	0	0	1	0	0	0	1	1	1
[7] Cupedidae	0	1	0	0	1	1	1	0	0	0	-
[8] Ommatidae	0	1	0	0	1	1	1	0	0	0	?
[9] <i>Lepicerus</i>	0	0	0	0	0	1	0	0	0	0	?
[10] <i>Ytu</i>	0	0	0	0	0	1	0	0	0	0	?

Описание структуры дерева (терминология)

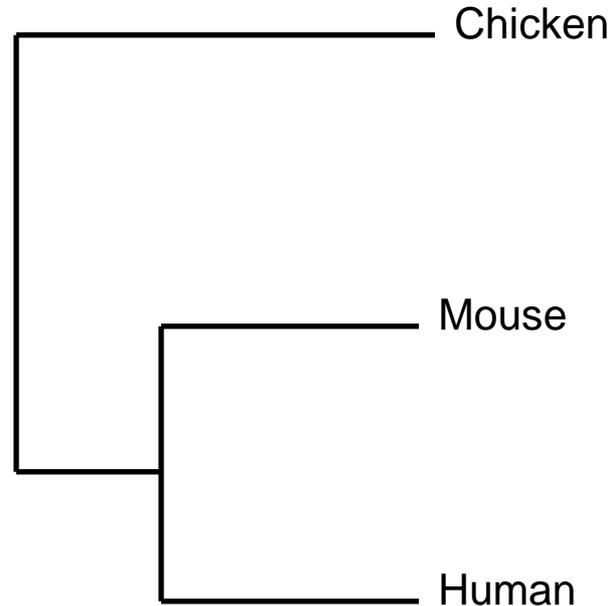
- **Узел (node)** — точка разделения предковой последовательности (вида, популяции) на две независимо эволюционирующие. Соответствует внутренней вершине графа, изображающего эволюцию.
- **Лист (leaf)** — реальный (современный) объект; внешняя вершина графа.
- **Ветвь (branch)** — связь между узлами или между узлом и листом; ребро графа.
- **Корень (root)** — гипотетический общий предок.
- **Клада (clade)** — группа организмов, представляющая собой всех потомков некоторого общего предка.



Филогенетическое дерево

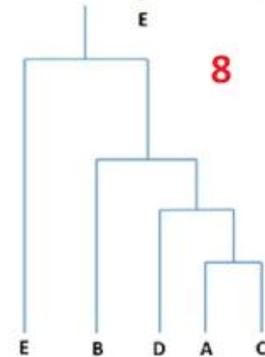
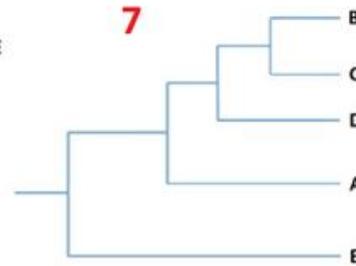
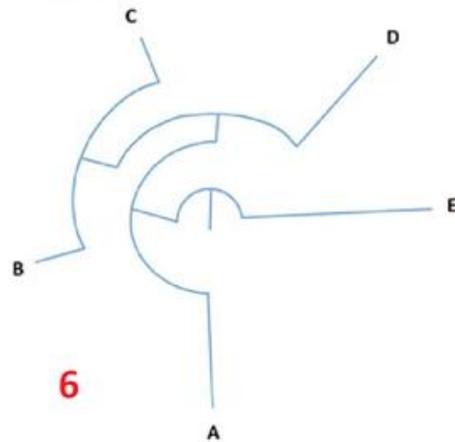
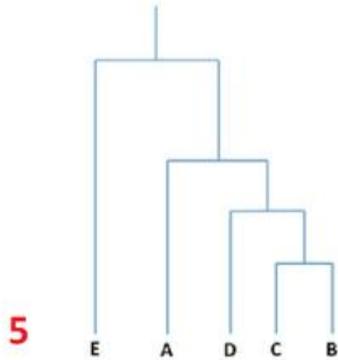
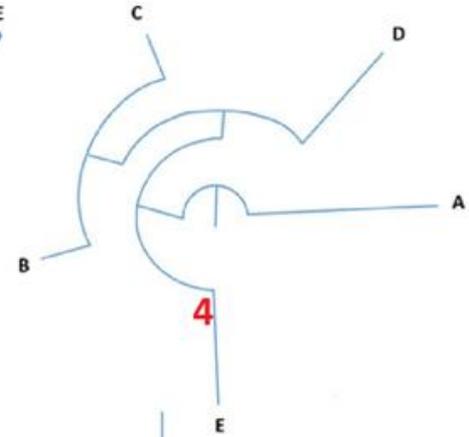
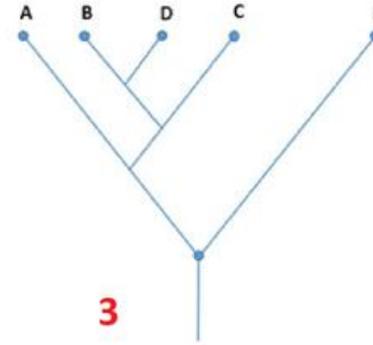
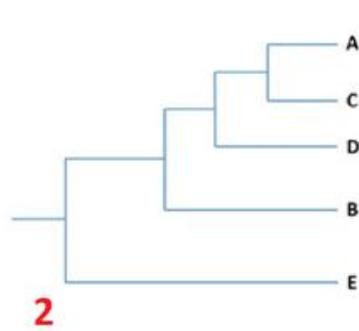
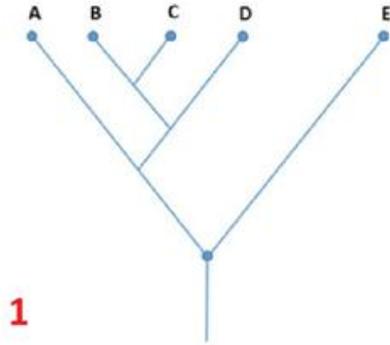


Угловая форма



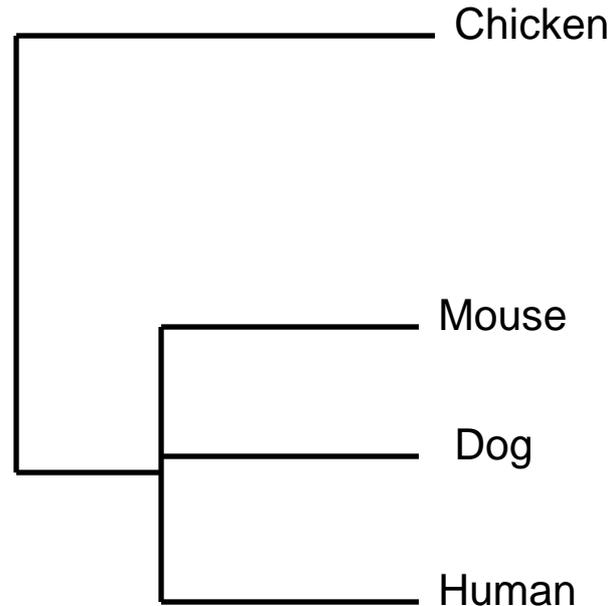
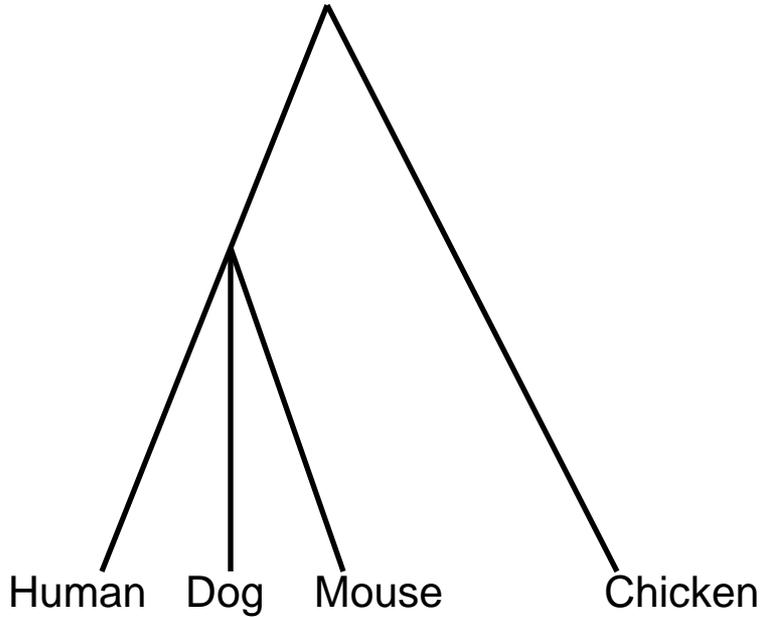
Прямоугольная форма

Упражнение



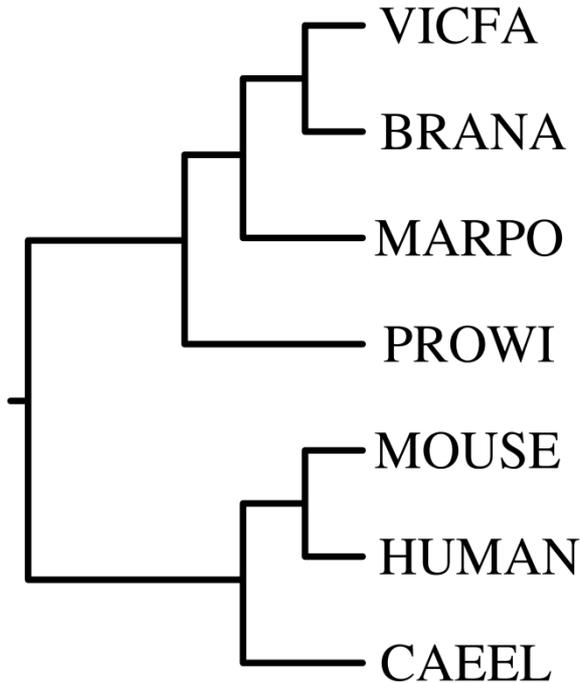
Какие деревья на самом деле одинаковые?

Небинарное дерево



Оно же “неразрешённое” — данных не хватает, чтобы разрешить все узлы

Скобочная формула



Newick Standard:

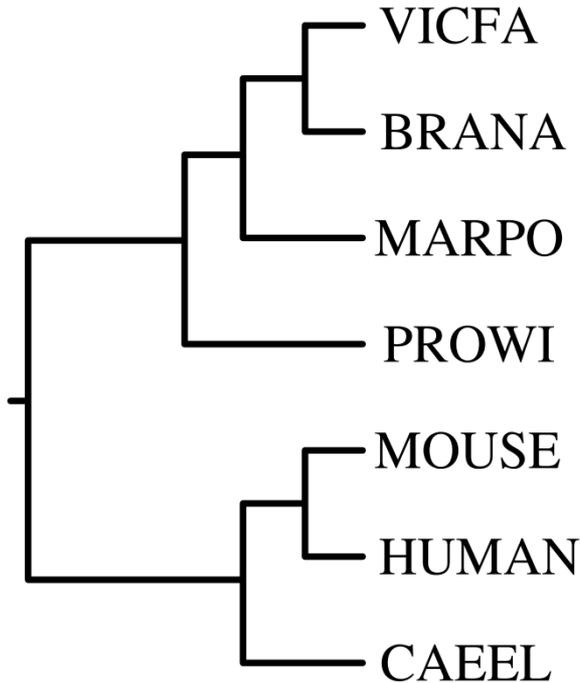
(((VICFA, BRANA), MARPO), PROWI), ((MOUSE, HUMAN), CAEEL));

“The reason for the name is that the second and final session of the committee met at Newick’s restaurant in Dover, and we enjoyed the meal of lobsters.”

Joseph Felsenstein, <http://evolution.genetics.washington.edu/phylip/newicktree.html>

Каждая пара скобок, кроме внешней, отвечает узлу, внешняя пара — корню.

Скобочная формула с длинами ветвей

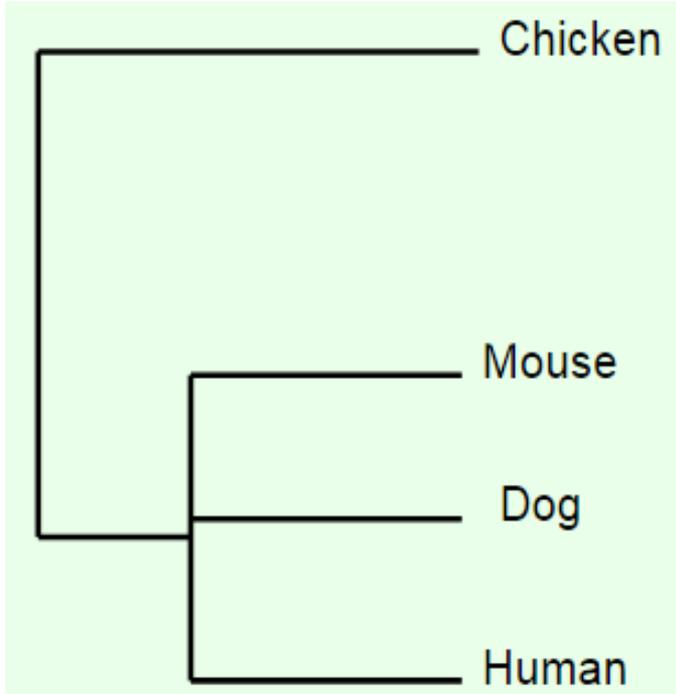


Newick Standard:

```
(((VICFA:3, BRANA:3):3, MARPO:6):2, PROWI:8):7, ((MOUSE:3, HUMAN:3):3, CAEEL:6):9);
```

Каждая пара скобок, кроме внешней, отвечает узлу. Число после двоеточия означает длину ветви, идущей из данного листа или узла в вышестоящий узел или корень

Скобочная формула: неразрешённое дерево



(Chicken, (Mouse, Dog, Human));

Каждая пара скобок, кроме внешней, отвечает узлу.

Клада vs таксоны

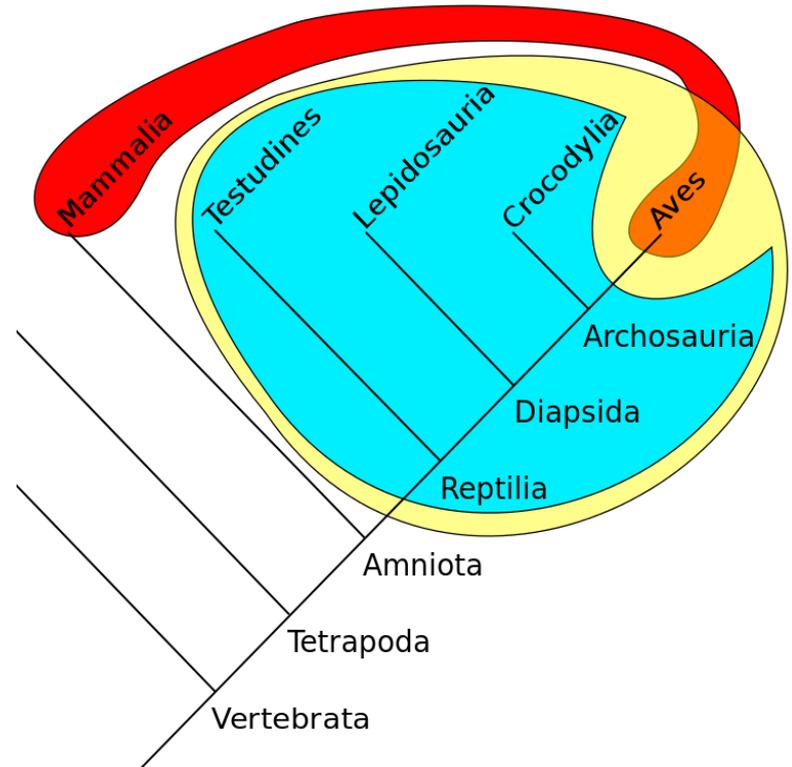
Клада — группа организмов, представляющая собой всех потомков некоторого общего предка (Last common ancestor, LCA)

Таксон не всегда соответствует кладе!

Монофилетический таксон включает LCA группы организмов и всех его потомков

Парафилетический таксон включает LCA, но не всех его потомков

Полифилетический таксон не включает общего предка ВСЕХ представителей таксона

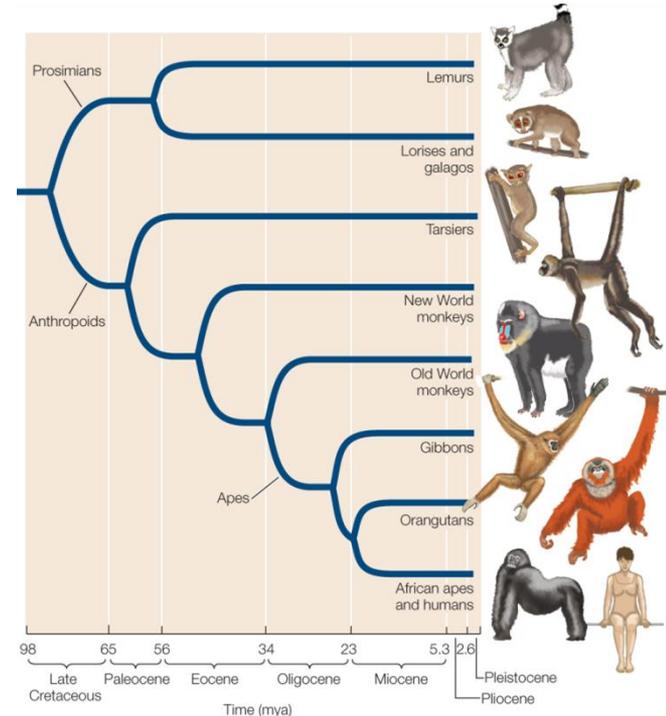
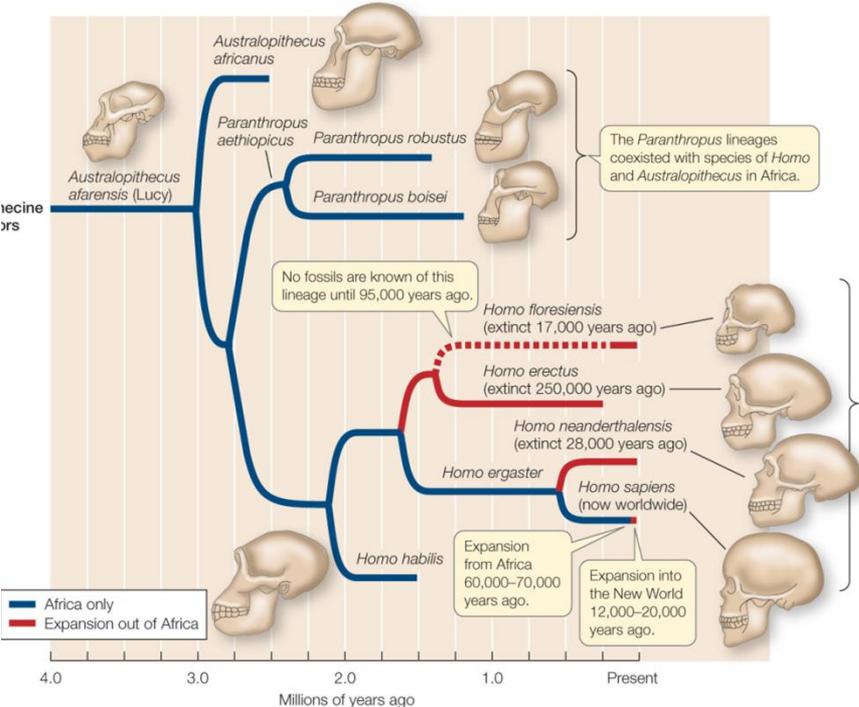


Что по горизонтальной оси у дерева?

И что отражают длины ветвей?

Иногда время

Почему на правом дереве все листья на одном уровне, а на левом нет?

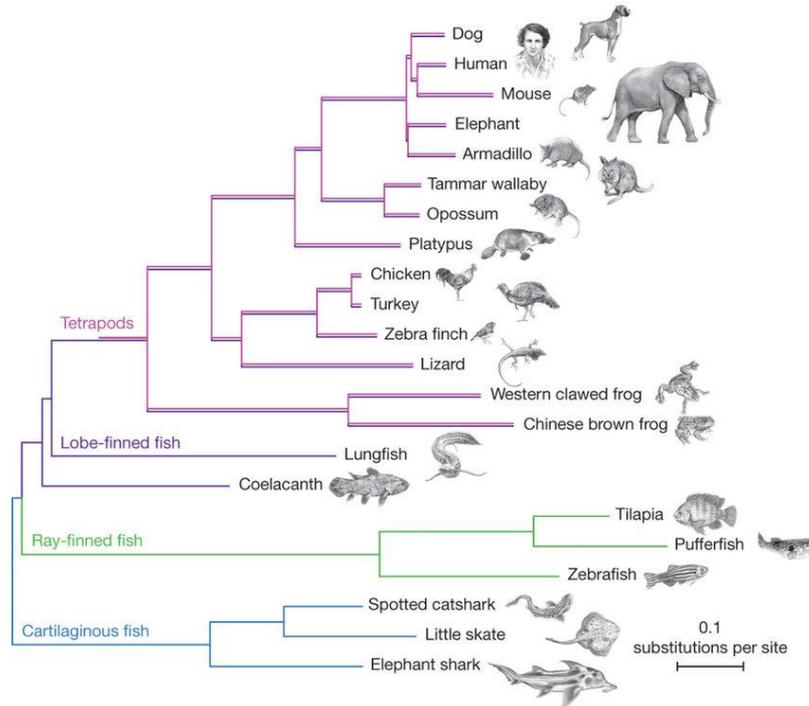


Что по горизонтальной оси у дерева?

И что отражают длины ветвей?

Иногда эволюционное расстояние

которое надо как-то оценивать



<https://www.nature.com/articles/nature12027>

По каким данным можно строить дерево?

1. По внешним признакам (по фенотипу);
2. По секвенированным последовательностям:
 - по гену,
 - по нескольким генам,
 - по геномам,
 - по последовательностям белков

Эволюция геномов

Фенотип определяется генотипом

Основой эволюции организмов и видов являются изменения геномов

- **Наследственность**

- Геном потомка похож на геном предка

- **Изменчивость**

- Геном потомка имеет отличия от генома предка:
 - ✓ комбинация генов родителей (у эукариот при половом размножении),
 - ✓ изменения генома клетки (у прокариот при горизонтальном переносе генов),
 - ✓ **мутации при репликации и репарации ДНК**

- **Отбор**

- Разные генотипы меняют свою распространенность в популяции с течением времени
- Последовательности многих участков генома (в частности, почти все белок-кодирующие участки) испытывают действие отбора. Вредные мутации элиминируются, закрепляются лишь безвредные и полезные мутации

Точечные замены в гене

... AATCCGTCAAGTCTA...

... Asn Pro Ser Ser Leu ...

1) “молчащая”(синонимическая)мутация

... AATCCGTCGAGTCTA...

... Asn Pro Ser Ser Leu ...

2) замена остатка на близкий по свойствам

... AATCCGACAAGTCTA...

... Asn Pro Thr Ser Leu ...

3) замена остатка на остаток с иными свойствами

... AATCCGTCAAGACTA...

... Asn Pro Ser Arg Leu ...

Судьба мутации

Бактерия разделилась, и у одного из потомков произошла мутация.
(ошибка репликации, или повреждение ДНК и ошибка репарации).

Что будет с потомством мутанта? Увидим ли мы эту мутацию, если отсеквенируем 1 000 000 бактерий этого штамма через 10 лет?

Потомство бактерии

В благоприятных условиях бактерия может делиться каждый час.
Сколько будет бактерий через 24 часа? А через год???

Потомство бактерии

В благоприятных условиях бактерия может делиться каждый час. Сколько будет бактерий через 24 часа? А через год????

Ответ: примерно столько же, сколько сейчас.

Численность подавляющего большинства популяций **постоянна** (по крайней мере на отрезках времени порядка лет) – погибает примерно столько же, сколько рождается.

Современная популяция человека – исключение!

Если члены популяции генетически идентичны, то вероятность оставить потомство для всех **одинакова**.

Следствие: математическое ожидание числа потомков одной бактерии через достаточно большой промежуток времени равно 1.

Судьба нейтральной мутации

Предположим, что мутация **нейтральна** = никак не влияет на матожидание числа потомков (таких мутаций довольно много).

Мутация произошла и передаётся потомкам мутанта. Значит, в популяции появился новый **полиморфизм**. У данного варианта генома есть **частота** (сначала очень маленькая).

Что произойдет с частотой через несколько поколений?

Судьба нейтральной мутации

Предположим, что мутация **нейтральна** = никак не влияет на матожидание числа потомков (таких мутаций довольно много).

Мутация произошла и передаётся потомкам мутанта. Значит, в популяции появился новый **полиморфизм**. У данного варианта генома есть **частота** (сначала очень маленькая).

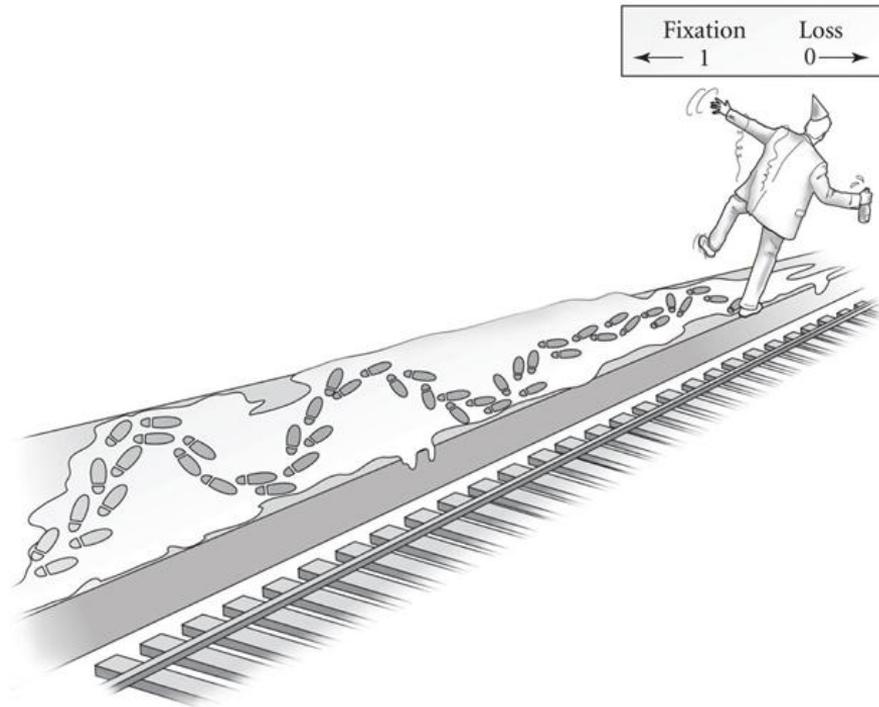
Что произойдет с частотой через несколько поколений?

Ответ: частота либо немного возрастёт, либо немного упадёт.
То и другое примерно равновероятно.

Что произойдет с частотой в итоге?

Случайное блуждание

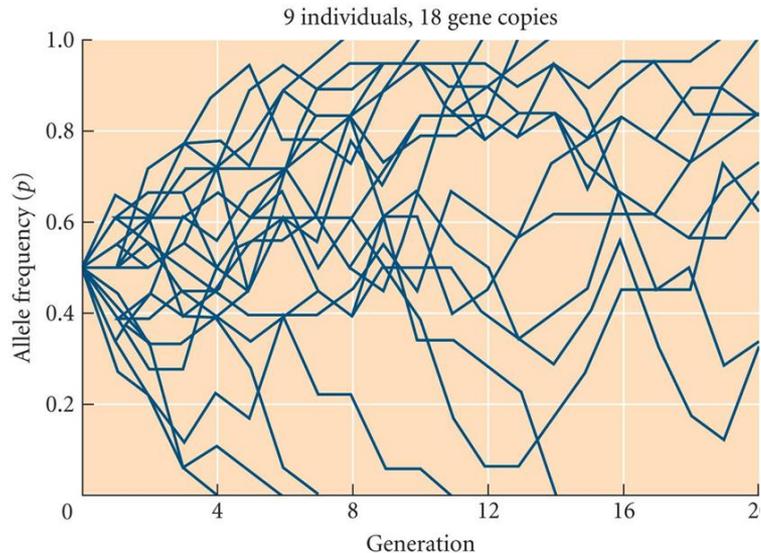
На тротуаре стоит пьяный и каждые несколько секунд делает шаг либо направо, либо налево, случайно выбирая направление. Что в итоге с ним произойдет?



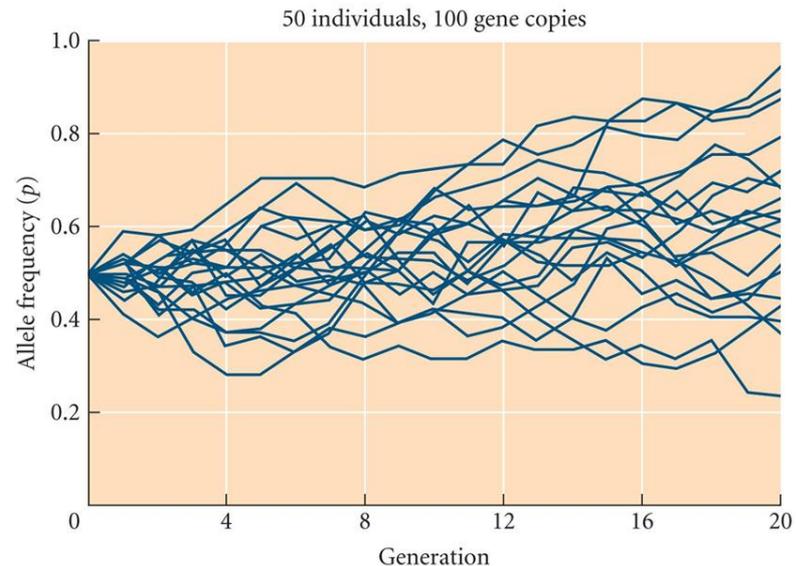
Случайное блуждание

Частота любого нейтрального полиморфизма постоянно колеблется случайным образом. Это называется «генетический дрейф».

Математическая модель такого процесса называется «случайное блуждание».



Результаты симуляций

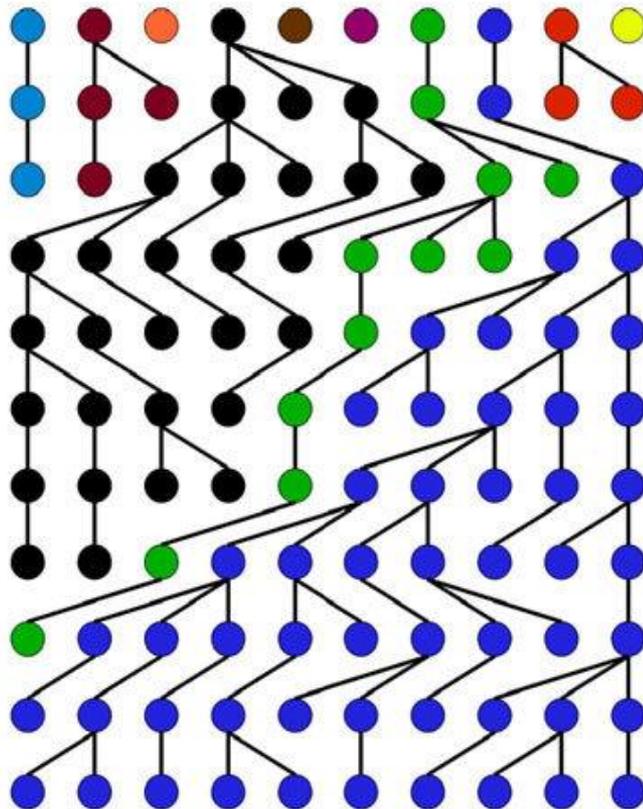


Генетический дрейф

Когда частота генетического варианта достигает 100% или 0%, процесс её изменения прекращается.

За исторически короткое время любой нейтральный вариант либо исчезает из популяции, либо закрепляется в ней!

В результате генетического дрейфа происходит закрепление новых нейтральных (или почти нейтральных) вариантов.



Накопление нейтральных мутаций

Вероятность закрепиться для новой нейтральной мутации очень мала, но не 0.

Организмов в популяции много, мутаций в них происходит тоже много (примерно 10^{-8} на п.н. на поколение – каждая сотая новорождённая бактерия несёт новую мутацию). Значительная доля мутаций нейтральна.

Итог: геномы независимых популяций начинают различаться, чем дальше, тем больше – в них независимо накапливаются нейтральные мутации.

С какой скоростью накапливаются нейтральные мутации в разных ветвях?

Накопление нейтральных мутаций

Вероятность закрепиться для новой нейтральной мутации очень мала, но не 0.

Организмов в популяции много, мутаций в них происходит тоже много (примерно 10^{-8} на п.н. на поколение – каждая сотая новорождённая бактерия несёт новую мутацию). Значительная доля мутаций нейтральна.

Итог: геномы независимых популяций начинают различаться, чем дальше, тем больше – в них независимо накапливаются нейтральные мутации.

С какой скоростью накапливаются нейтральные мутации в разных ветвях?

Если скорость **возникновения** мутаций не изменилась, то с одинаковой!

⇒ «молекулярные часы»

А если мутация не нейтральна?

Каждому варианту генома можно сопоставить его «приспособленность»

f = матожидание числа потомков организма с таким геномом (через какой-то фиксированный промежуток времени).

В подавляющем большинстве случаев новая мутация порождает либо нейтральный вариант ($f = 1$) либо вредный ($f < 1$).

Вредный вариант тоже начинает «блуждать», но вероятность «шага вверх» оказывается меньше вероятности «шага вниз». Это очень сильно уменьшает вероятность закрепления – тем сильнее, чем меньше f , и тем сильнее, чем больше популяция.

Явление невозможности закрепления вредной мутации называется **стабилизирующий отбор** или же **отрицательный отбор**.

Положительный отбор

Если вдруг $f > 1$, то вероятность закрепления мутации возрастает во много раз. Процесс закрепления полезных мутаций называется **положительным отбором**.

Собственно, полезных мутаций так мало именно потому, что большинство возможных полезных мутаций уже закрепились.

Обычно полезные мутации начинают появляться в заметном количестве только при изменении условий жизни организмов – например при появлении нового источника пищи или новой опасности или попадании части популяции в другой климат.

Положительный отбор

Если вдруг $f > 1$, то вероятность закрепления мутации возрастает во много раз. Процесс закрепления полезных мутаций называется **положительным отбором**.

Собственно, полезных мутаций так мало именно потому, что большинство возможных полезных мутаций уже закрепились.

Обычно полезные мутации начинают появляться в заметном количестве только при изменении условий жизни организмов – например при появлении нового источника пищи или новой опасности или попадании части популяции в другой климат.

Как выглядит дерево, одна из ветвей которого долго находилась под положительным отбором?

Кто строит деревья по последовательностям?

- **Специалисты по систематике и филогенетике**

находят ортологичные последовательности, строят деревья по ним, на основании этих деревьев судят о филогении организмов

- **Исследователи, изучающие семейства белков со схожей функцией**

дерево семейства в сочетании с деревом видов помогает судить о происхождении специализированных функций, предсказывать специфичность и т. д.

- **Эпидемиологи**

исследуют происхождение штаммов вирусов и бактерий

- **Исследователи злокачественных опухолей**

важно понимать, как и когда возникают раковые клетки с опасными свойствами (например, способностью к метастазированию, устойчивостью к препаратам и т.п.)

- ...

Филогения семейства белков

(organelle zinc finger editing factor family)

