

Филогенетические деревья

ФББ, 4 семестр, весна 2026

Алешина Юлия Александровна

По материалам С.А.Спирина, Д.В.Дибровой, Е.Д. Рюминой, А.В. Алексеевского

Представление эволюционной истории организмов в виде дерева

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

«Происхождение видов»,
Ч. Дарвин

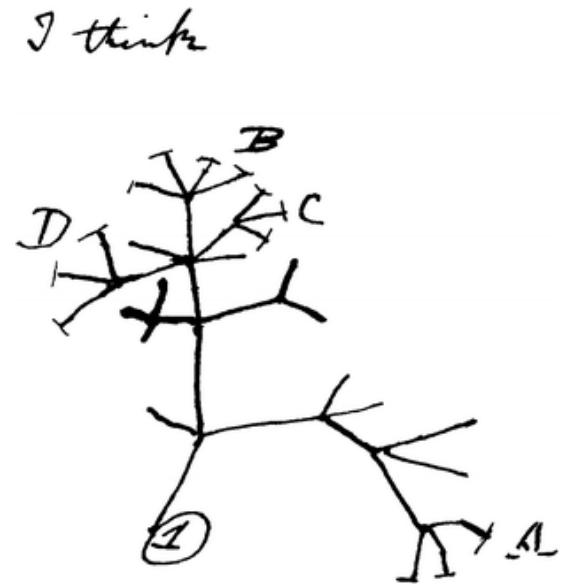


Рисунок из тетради «Notebook B»
Чарльза Дарвина, 1837 г.

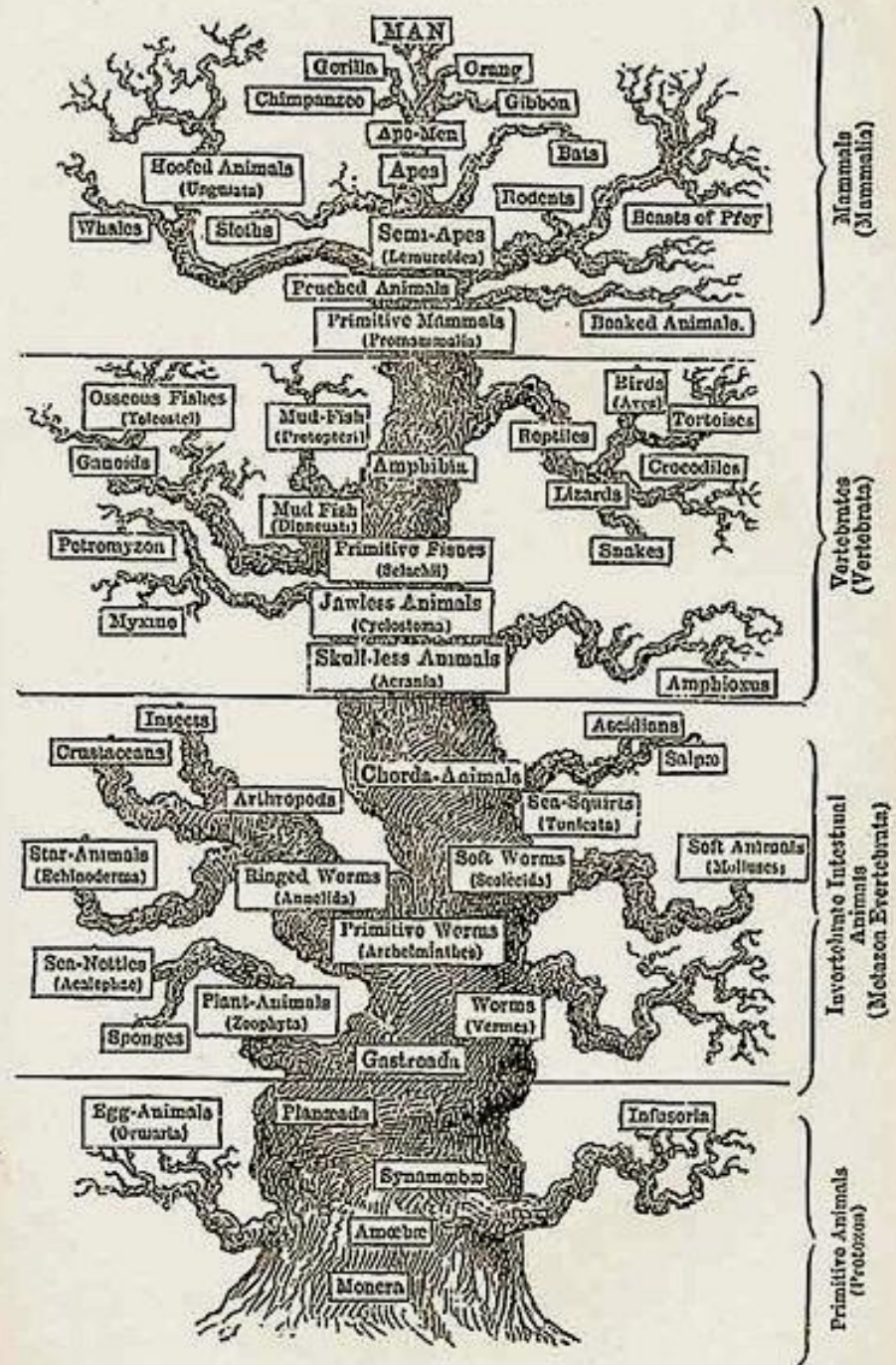
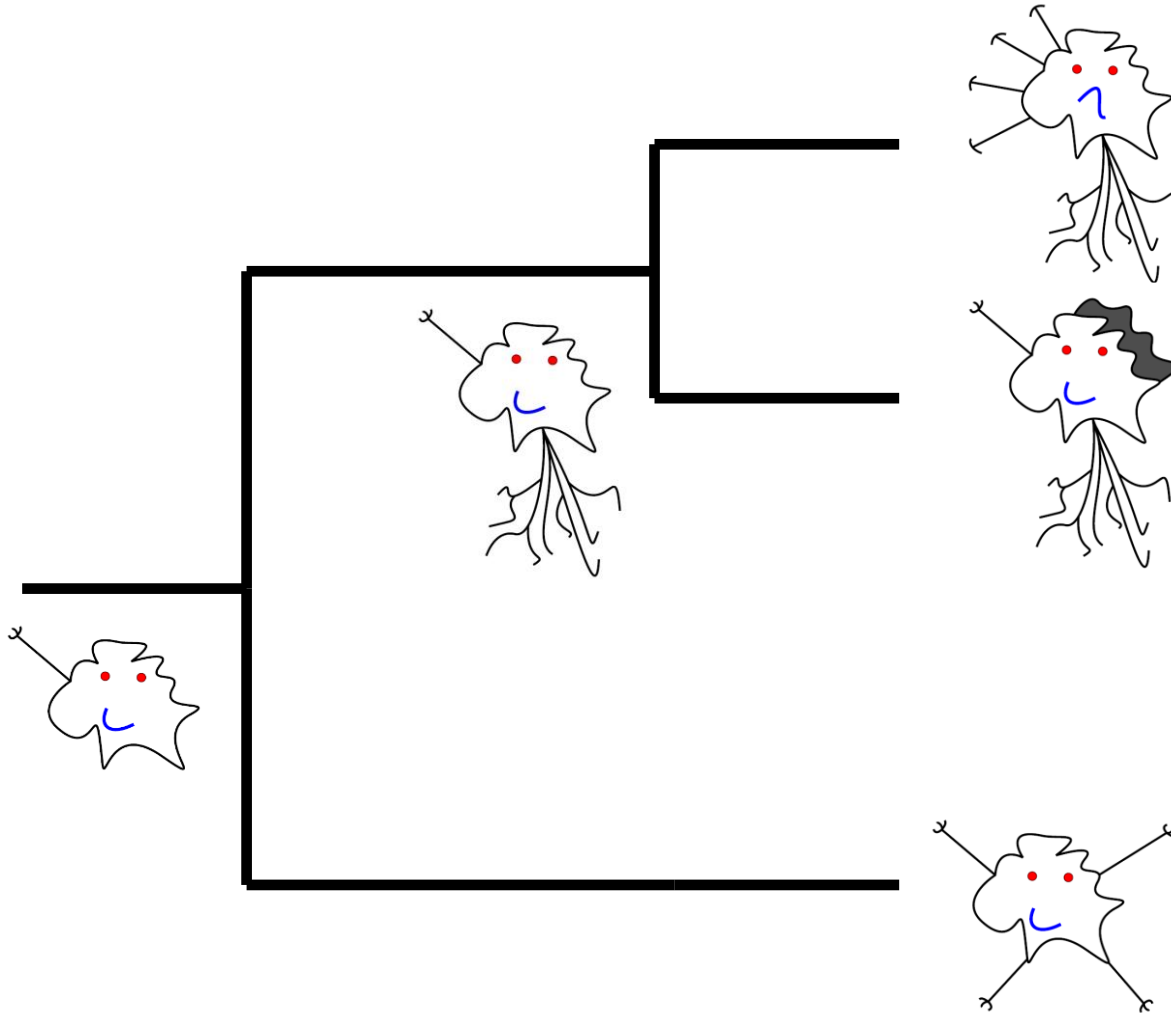


Рисунок из книги Эрнеста Геккеля «Эволюция человека», 1879

По каким данным можно строить дерево?

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

Реконструкция эволюции по характерным признакам



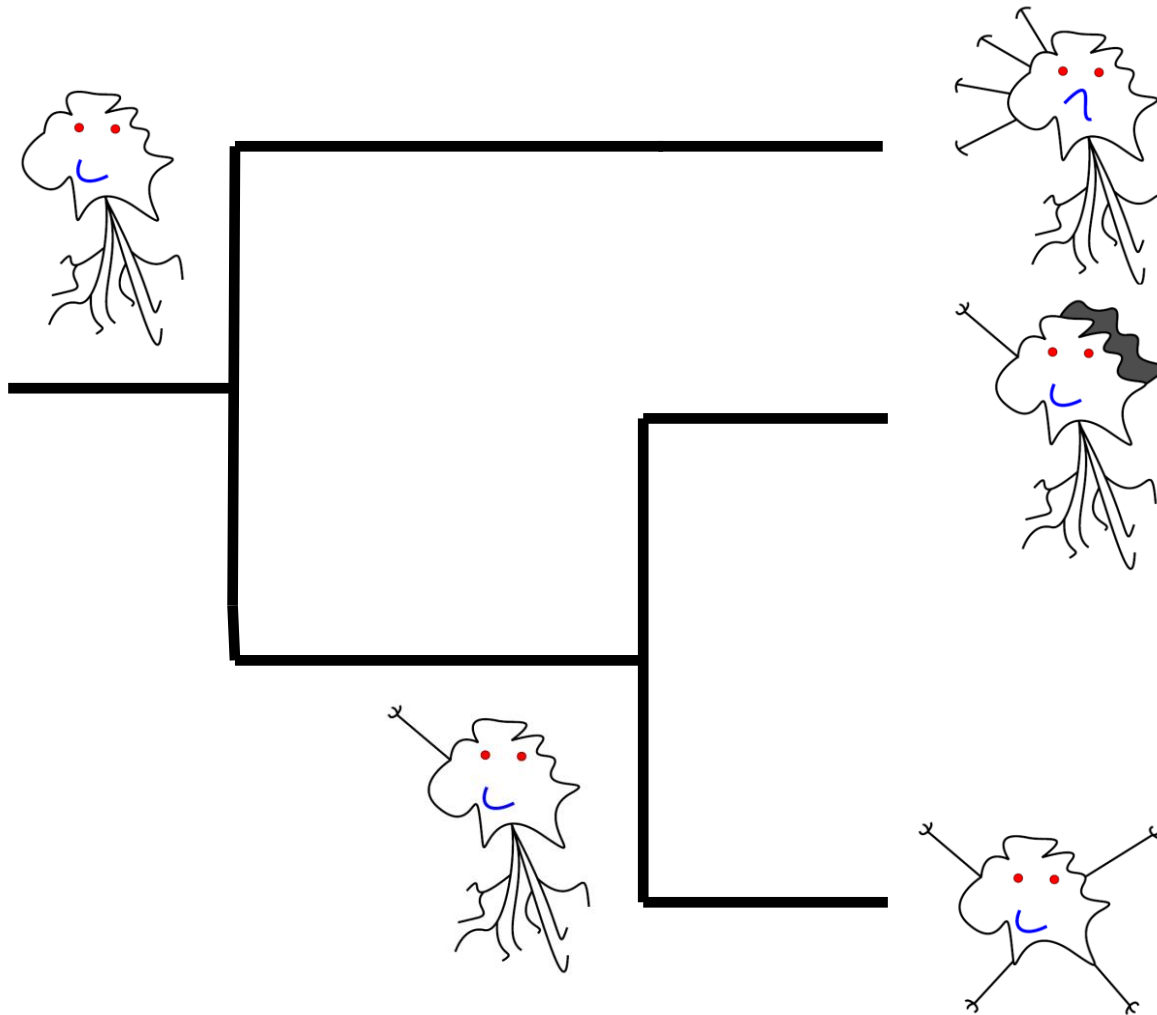
Современные
организмы



Реконструкция
эволюции

Дерево
отображает
ВОЗМОЖНЫЙ
ХОД ЭВОЛЮЦИИ

Альтернативный вариант дерева – и тоже обоснованное признаками



Современные
организмы



Реконструкция
эволюции

Дерево
отображает
ВОЗМОЖНЫЙ
ХОД ЭВОЛЮЦИИ

Система характерных признаков (characters) – первый шаг на пути к дереву

! У признака должны быть четко оговоренные варианты (часто встречаются бинарные признаки «есть» / «нет», но может быть и 3+ вариантов; есть вариант «не известно»)

| Признак | Жук 1 | Жук 2 | Жук 3 | ... |
|---------------------------|-------------------------|-------------------|------------|-----|
| ... | ... | ... | ... | ... |
| 51. Местообитание личинки | Наземное или полуводное | ? (неизвестно) | Водное | ... |
| ... | ... | ... | ... | ... |
| 57. Ориентация головы | Ортогнатная | Прогнатная | Прогнатная | ... |
| ... | ... | ... | ... | ... |
| 66. Число антенномеров | > 13 | 11 | < 11 | ... |
| ... | ... | ... | ... | ... |

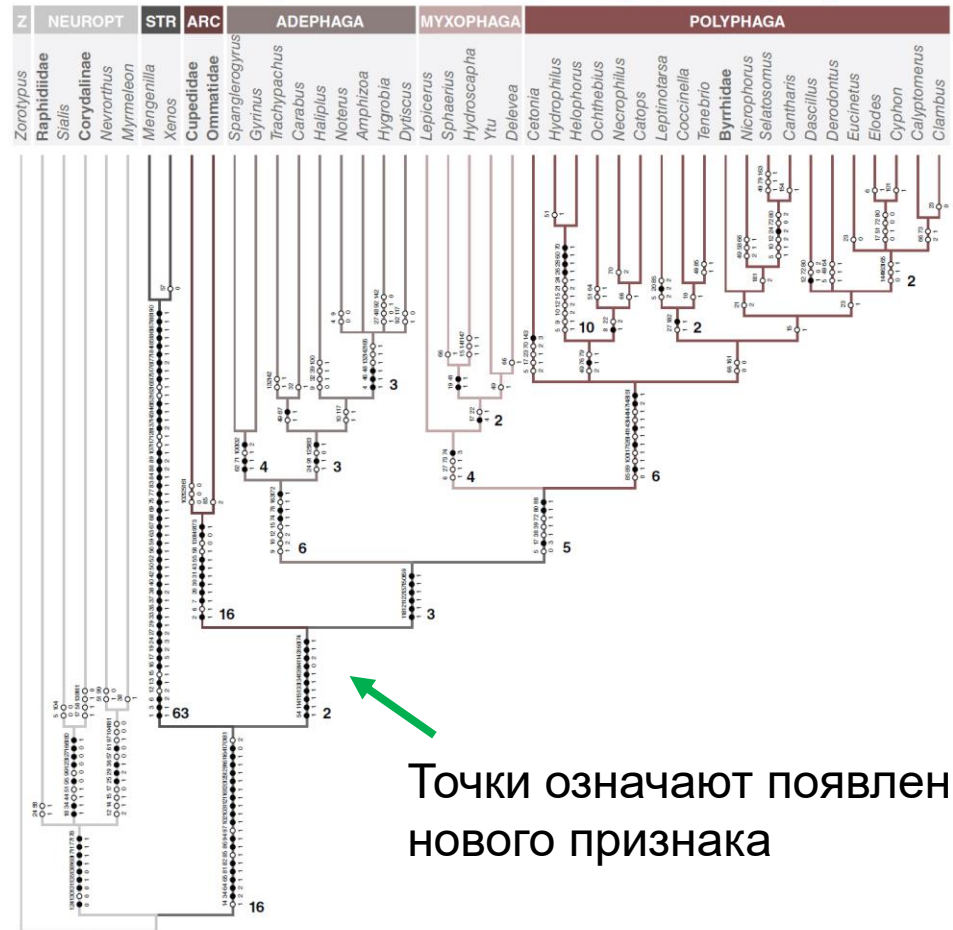
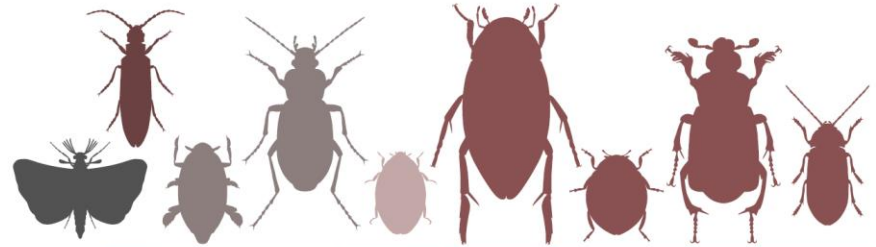
Beutel R.G. *et al.* The phylogeny of Coleoptera (Hexapoda) – morphological characters and molecular phylogenies. *Systematic Entomology* 44(1):75–102, 2018.



В этой статье описано 190 признаков жуков!

Реконструкция филогенетических отношений между видами

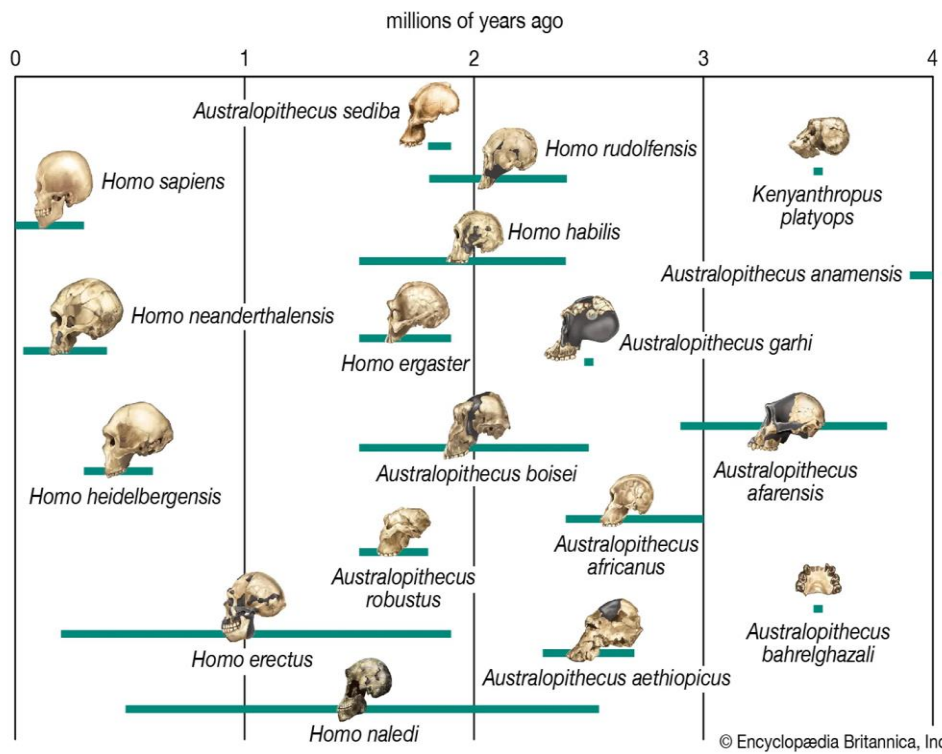
• Для построения деревьев используют принцип максимальной **экономии** (parsimony):
 Скорее правильно дерево, которое позволит описать имеющиеся данные при минимально необходимом числе эволюционных событий



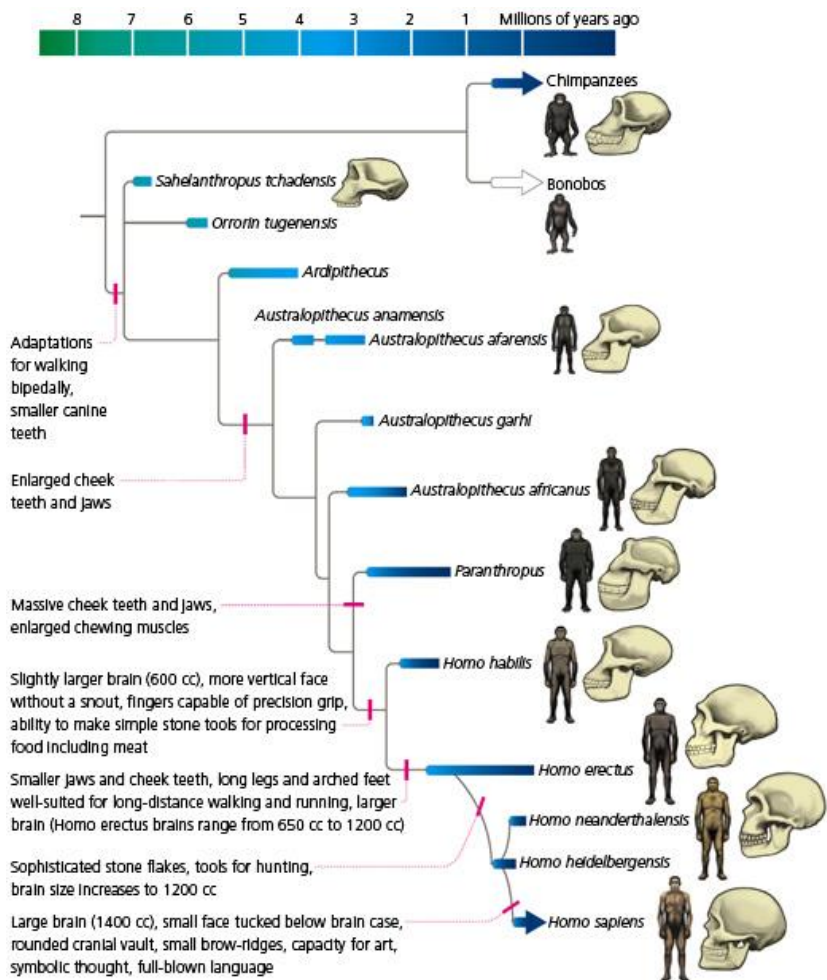
Beutel R.G. et al. The phylogeny of Coleoptera (Hexapoda) – morphological characters and molecular phylogenies. *Systematic Entomology* 44(1):75–102, 2018.

Точки означают появление нового признака

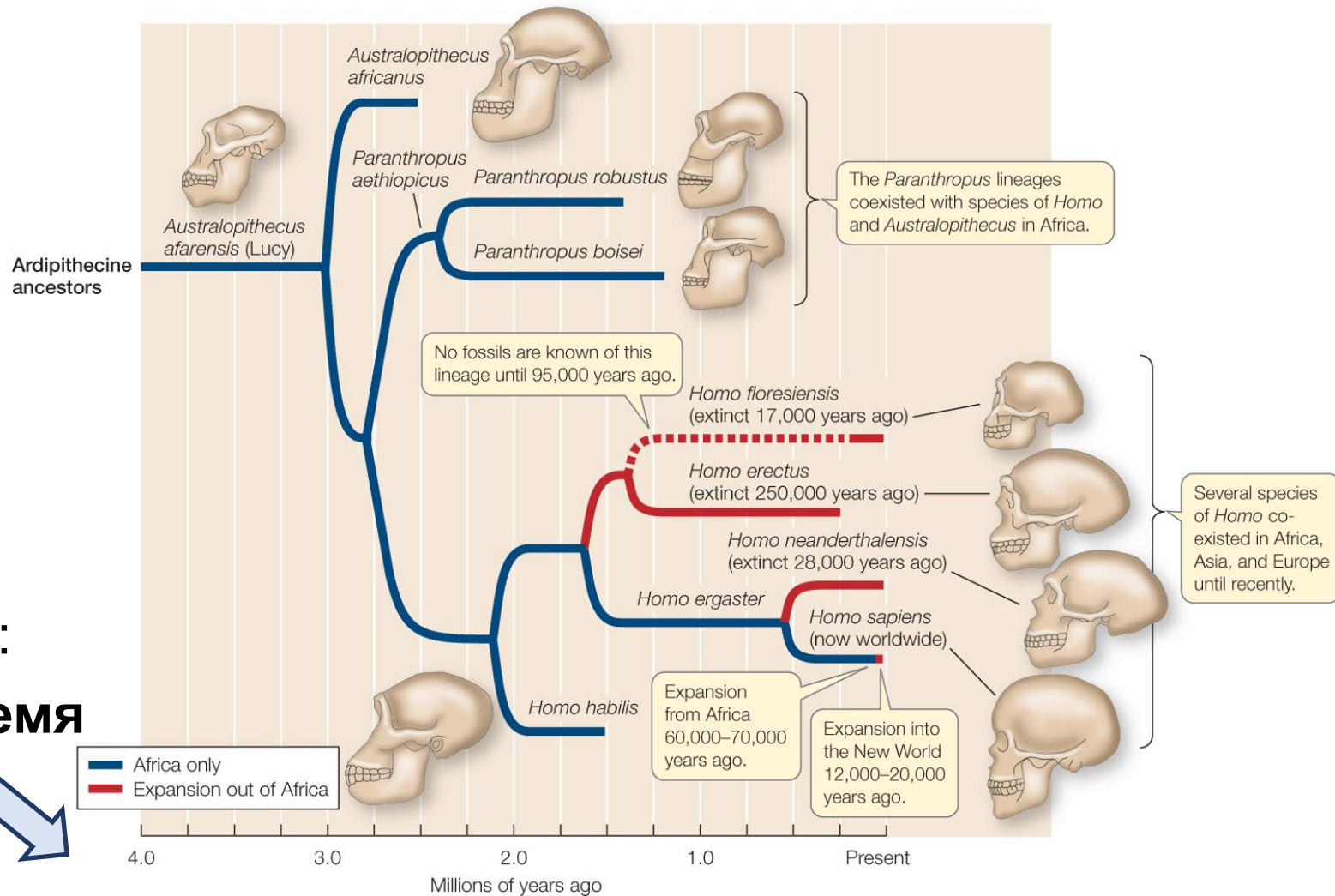
Эволюция человека восстанавливается по особенностям черепов (и других костей)



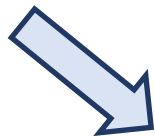
© Encyclopædia Britannica, Inc.



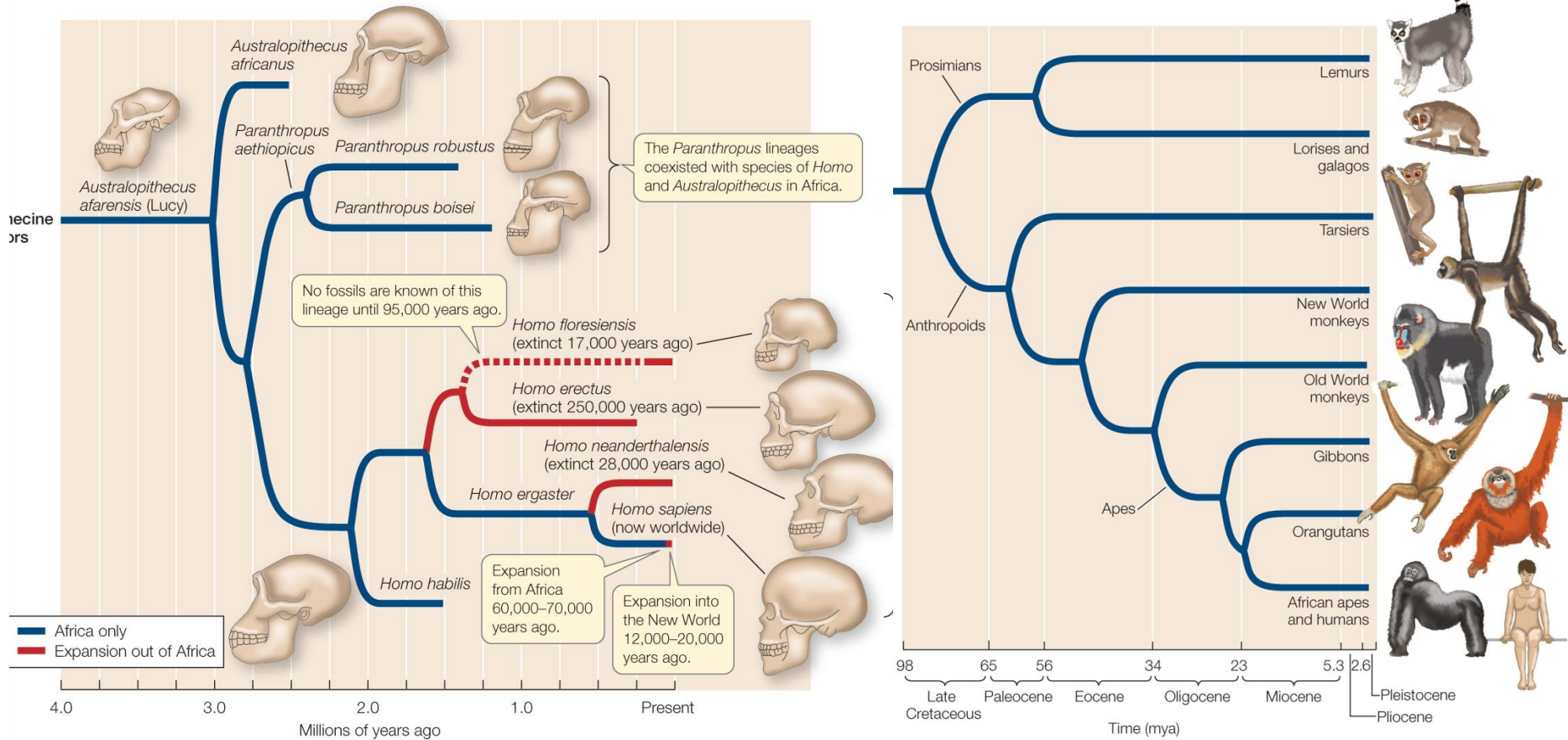
Что по горизонтальной оси у дерева?



Вариант 1:
по оси время



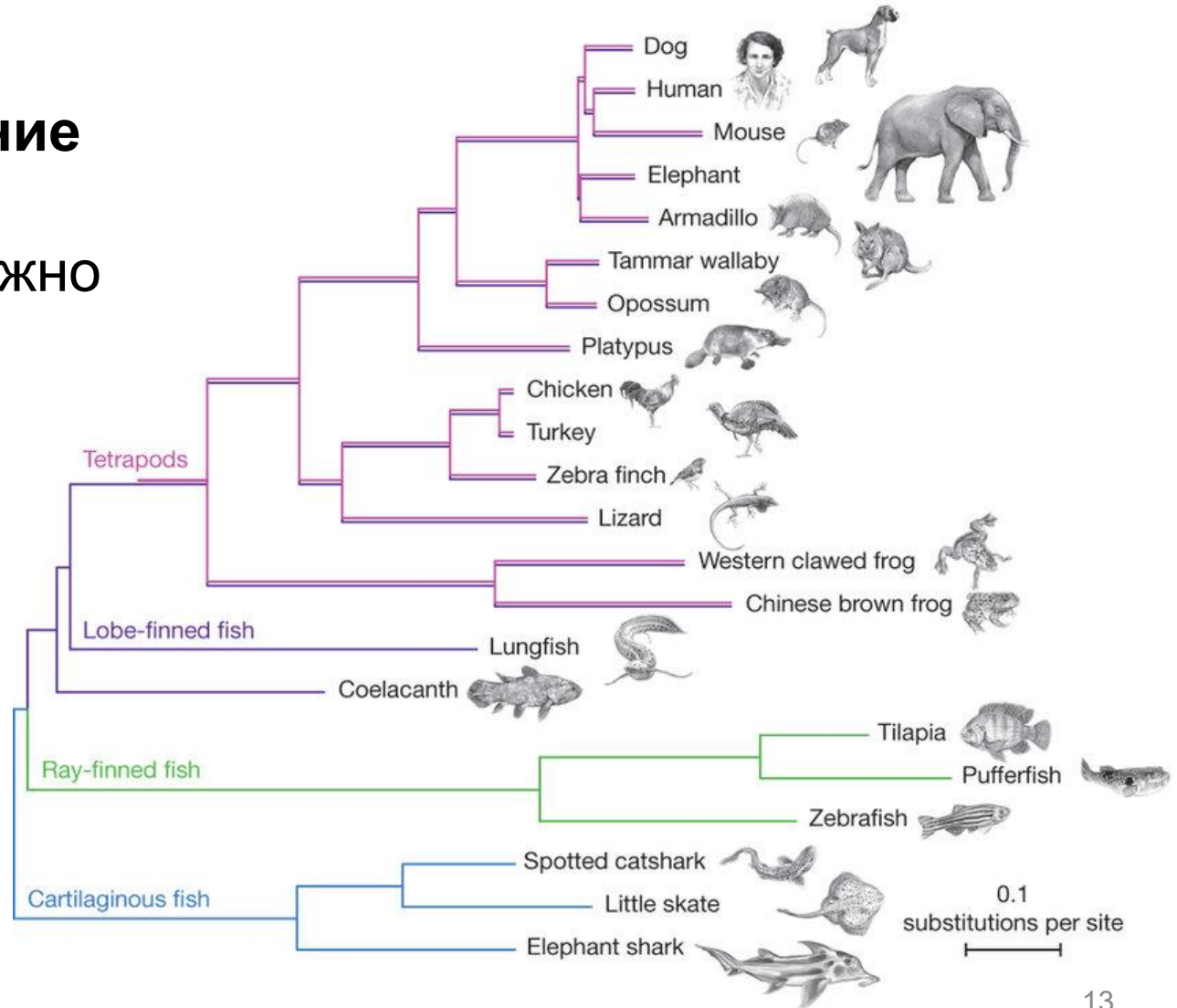
Почему на правом дереве все листья на одном уровне, а на левом нет?



https://digfir-published.macmillanusa.com/pol2e/asset/img_ch23/c23_fig55.html
https://digfir-published.macmillanusa.com/pol2e/pol2e_ch23_8.html

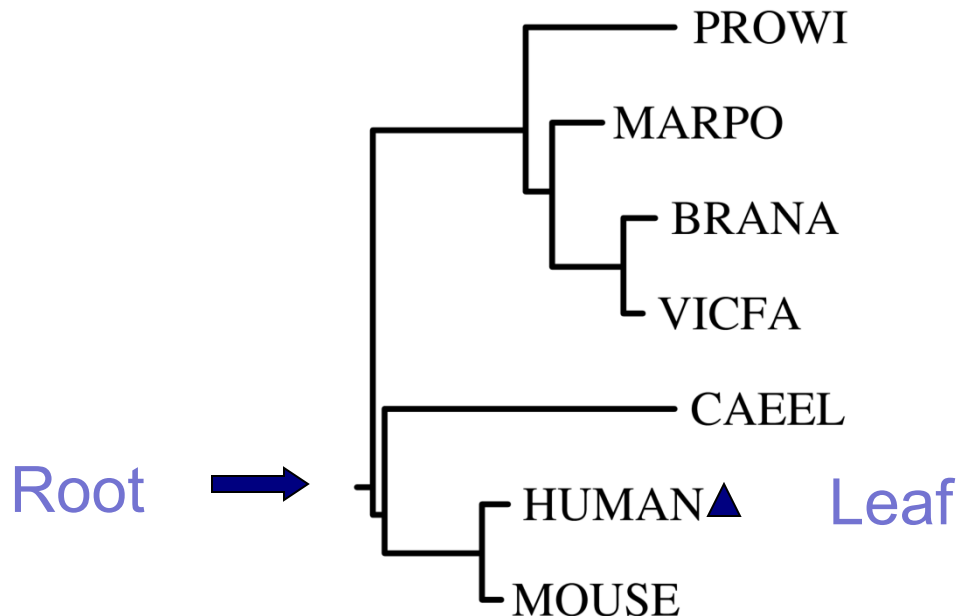
Что по горизонтальной оси у дерева?

Вариант 2:
по оси **расстояние**
между видами
(которое еще нужно
оценить)



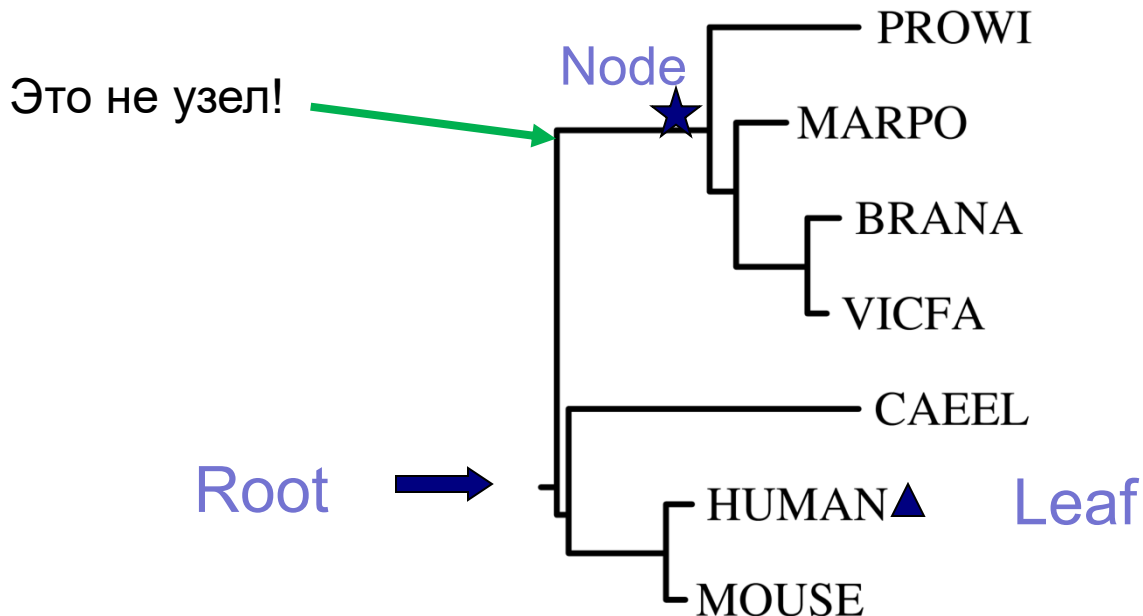
Описание структуры дерева (терминология)

- Лист (leaf) — реальный (современный) объект; внешняя вершина графа.
- Корень (root) — гипотетический общий предок.



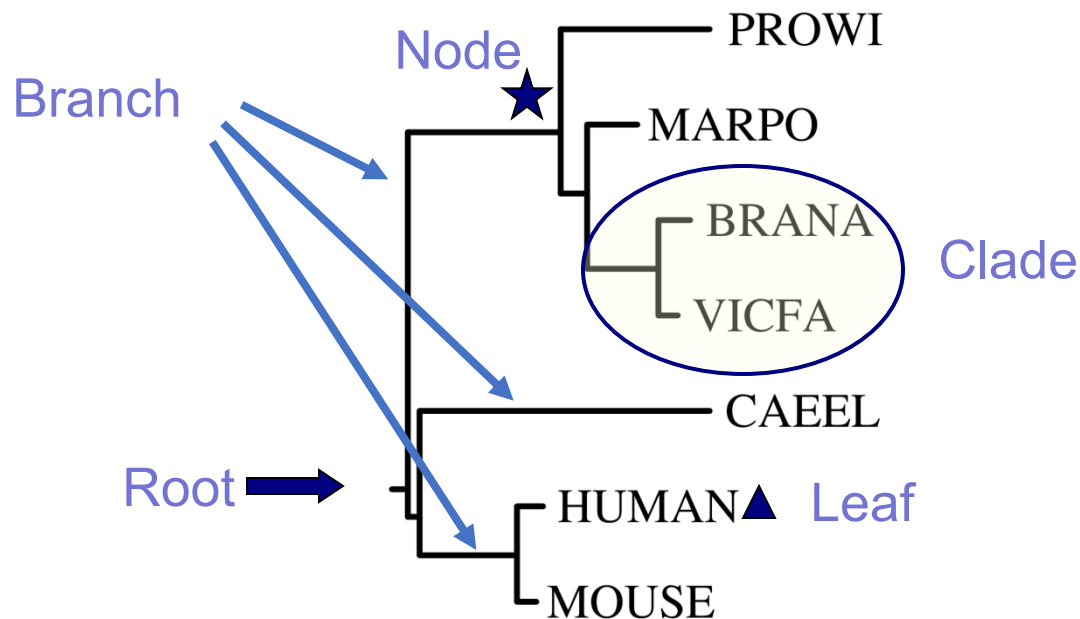
Описание структуры дерева (терминология)

- Узел (node) — точка разделения предковой последовательности (вида, популяции) на две независимо эволюционирующие. Соответствует внутренней вершине графа, изображающего эволюцию.



Описание структуры дерева (терминология)

- **Ветвь (branch)** — связь между узлами или между узлом и листом; ребро графа.
- **Клада (clade)** — группа организмов, представляющая собой всех потомков некоторого

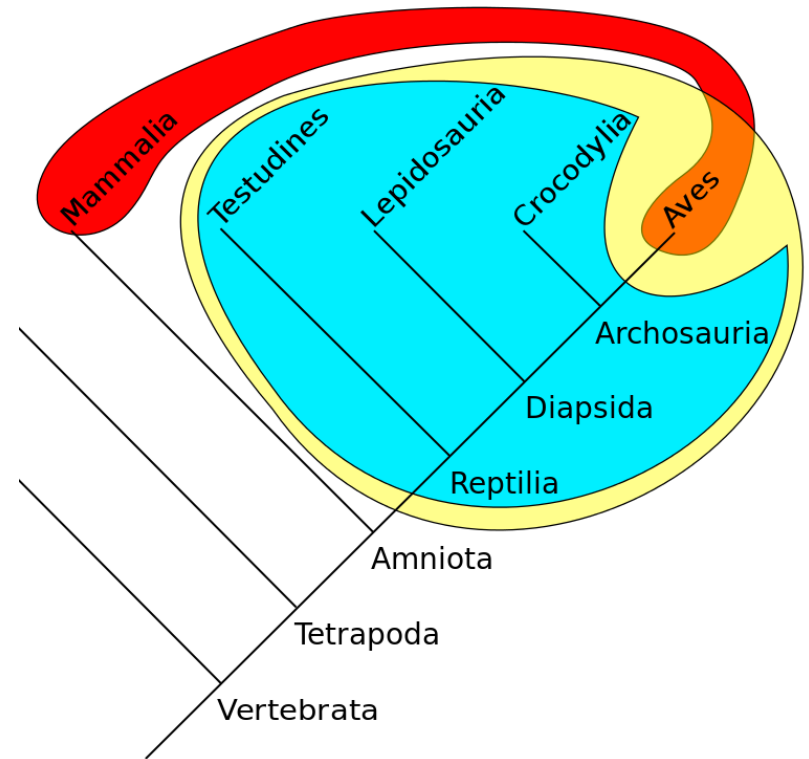


Клады и таксоны

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

Таксон (группа) 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.

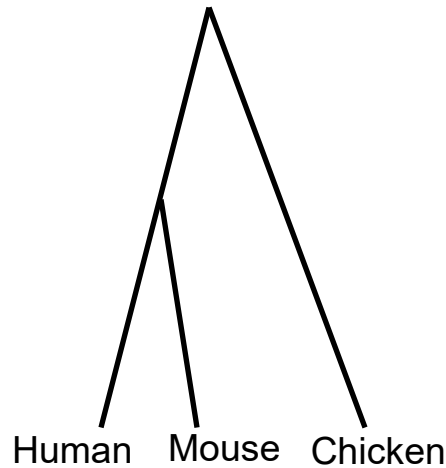


Sauropsida

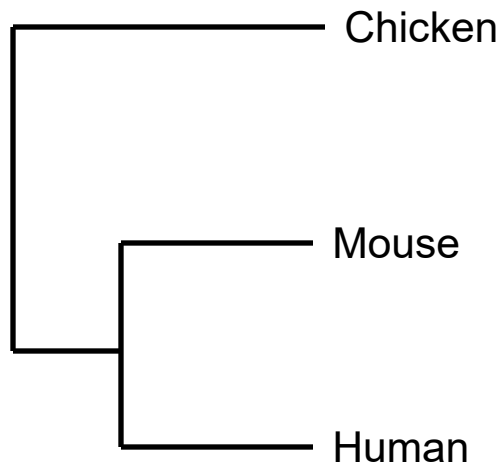
Рептилии

Теплокровные

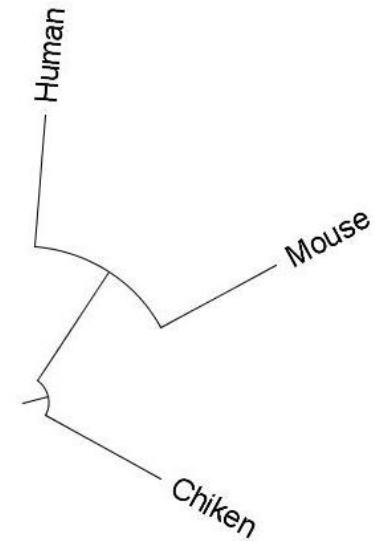
Виды представления филогений



Угловая форма

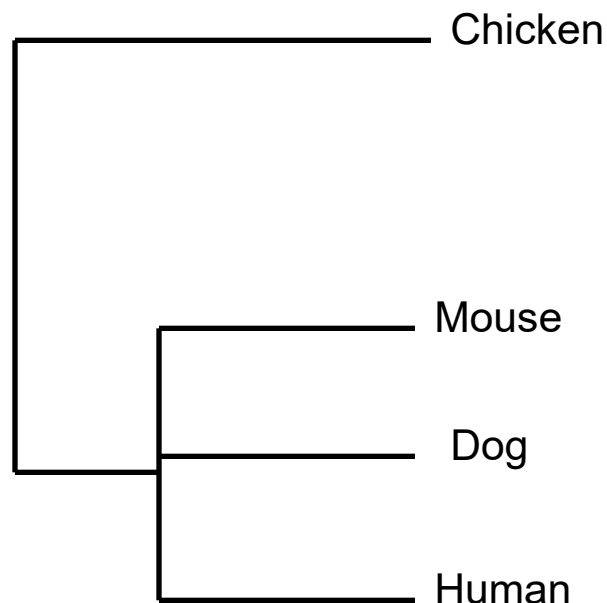
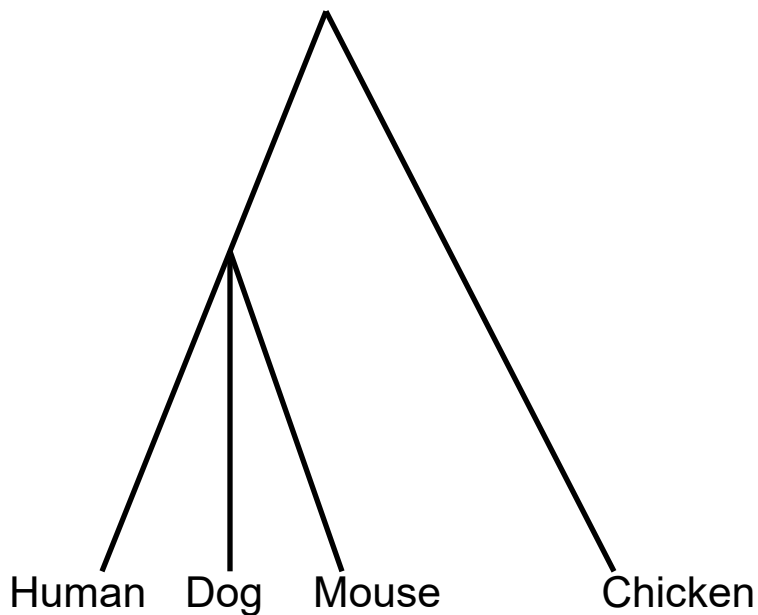


Прямоугольная форма



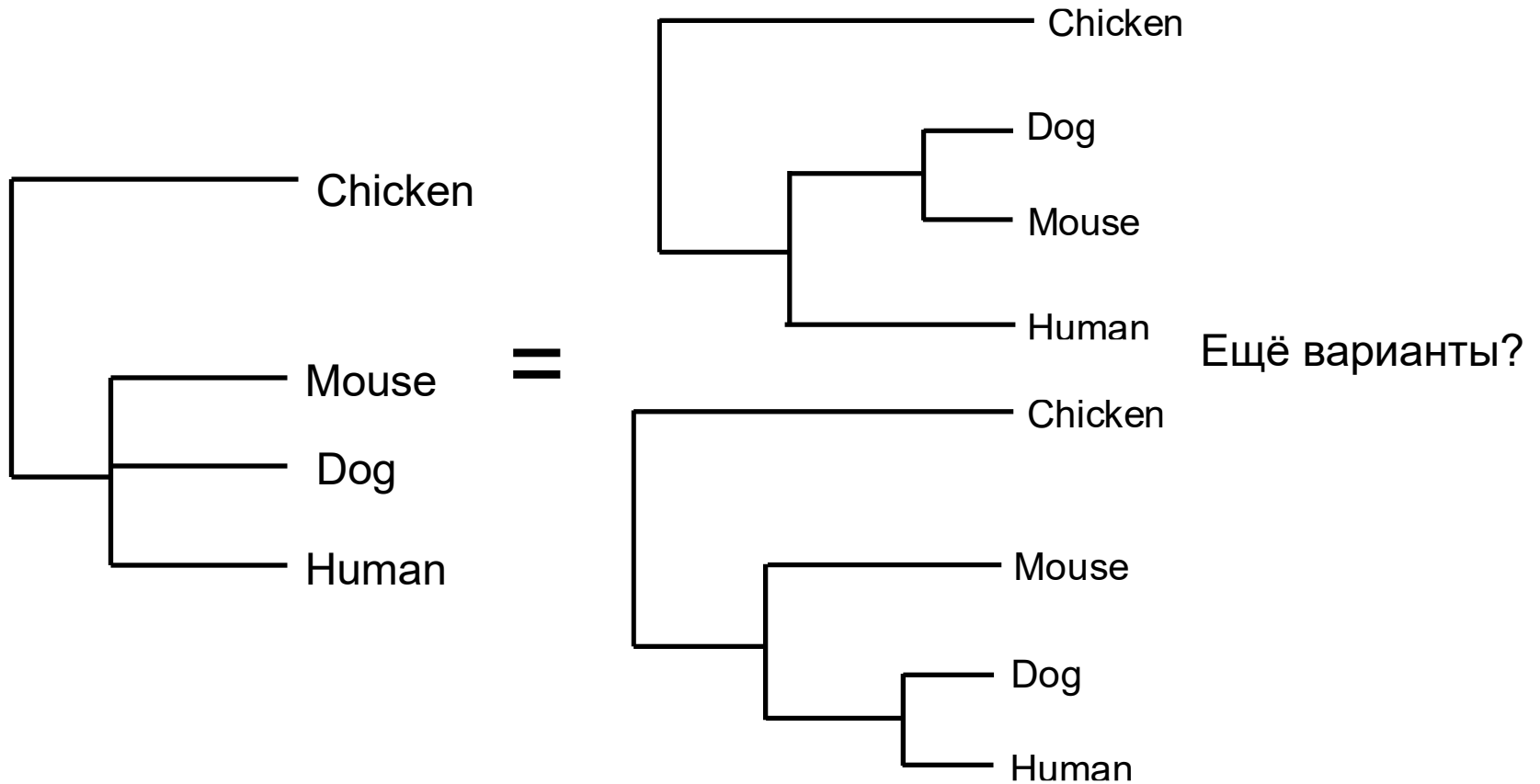
Круговая форма

Небинарное дерево

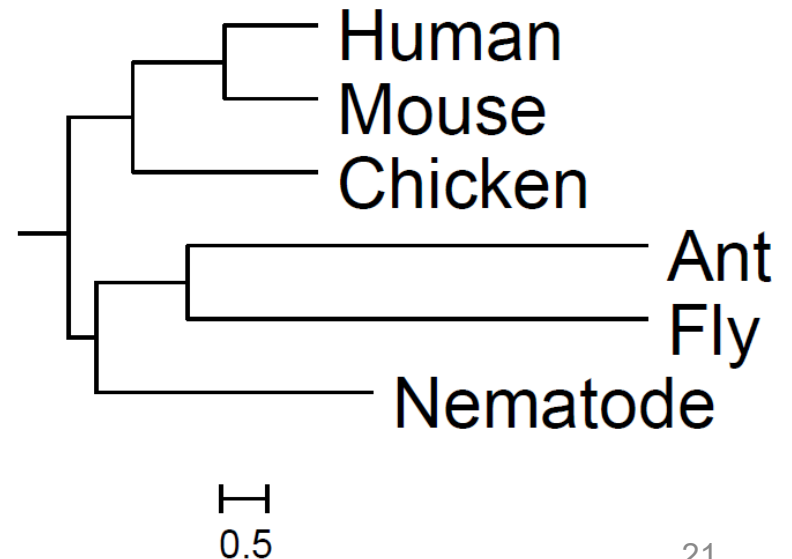
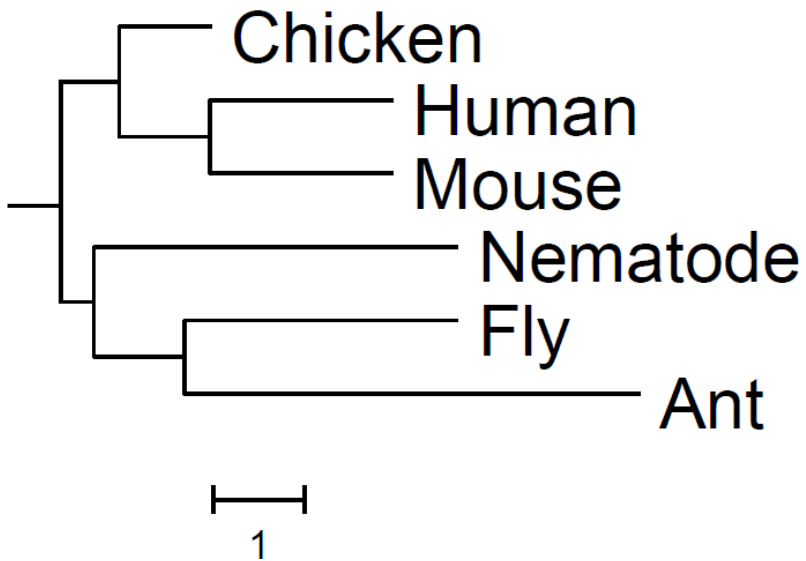
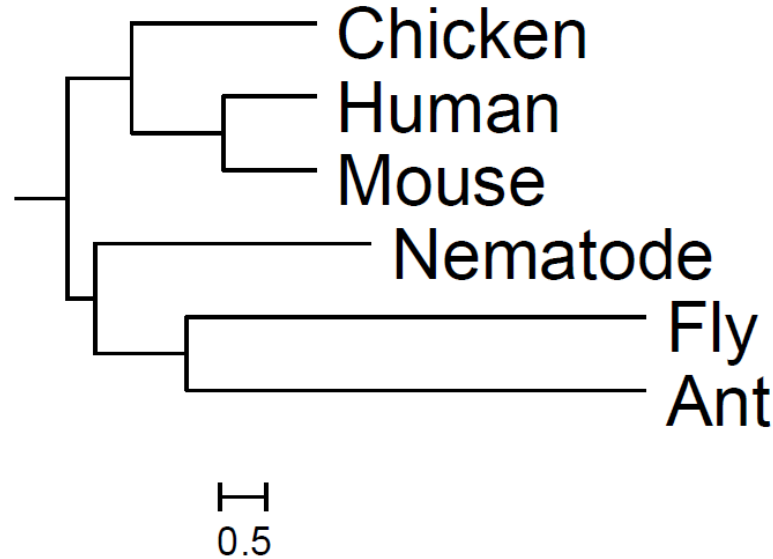


Оно же “неразрешённое” — данных не хватает, чтобы разрешить все узлы

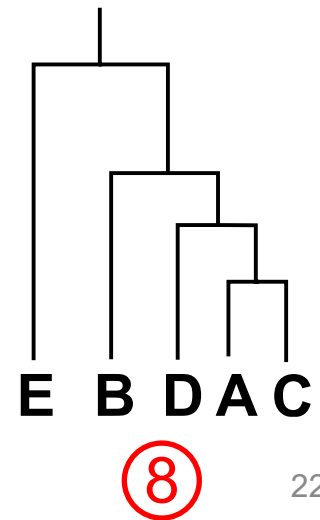
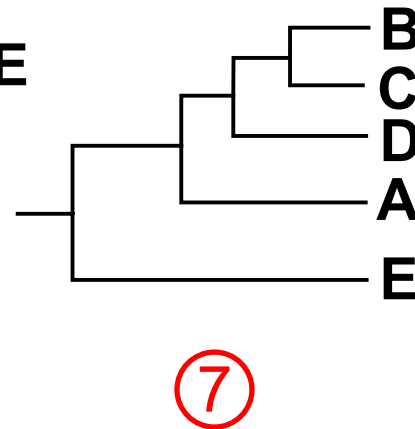
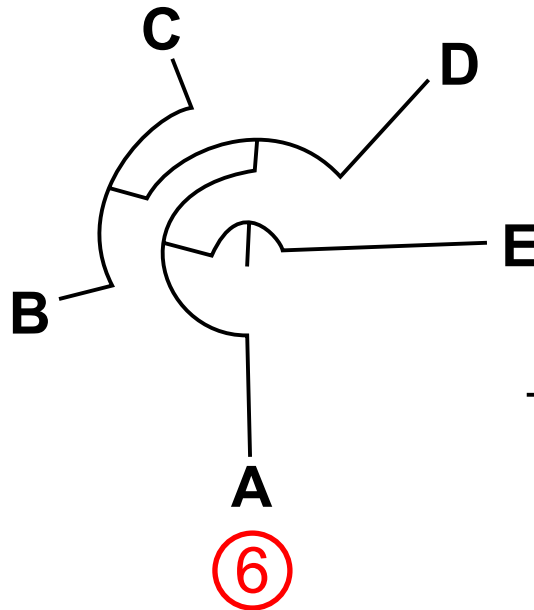
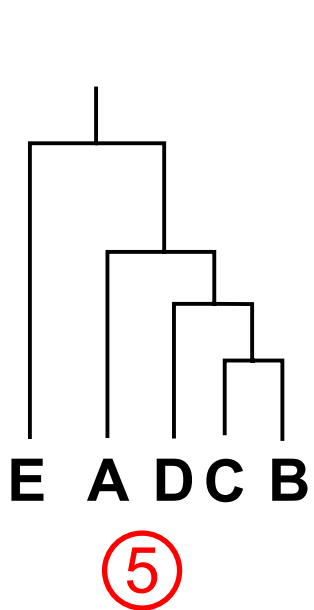
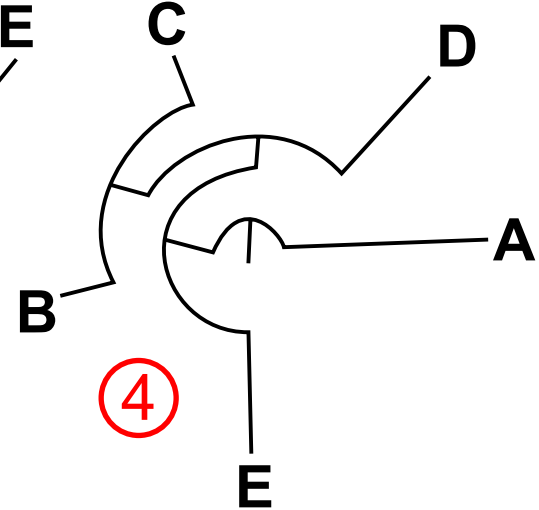
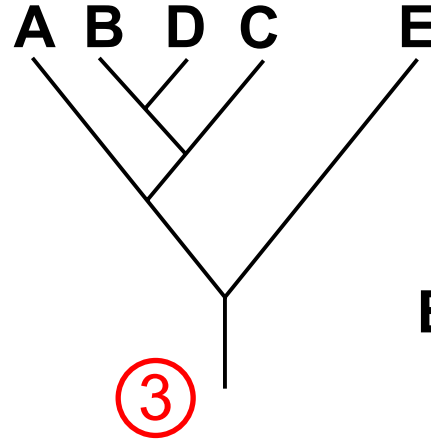
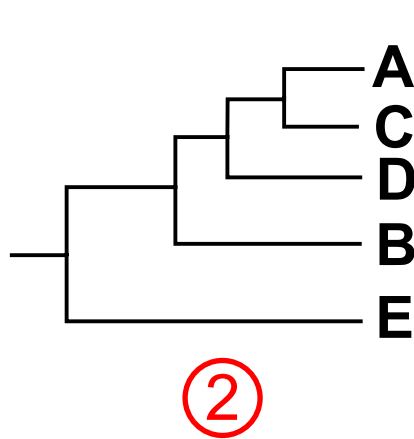
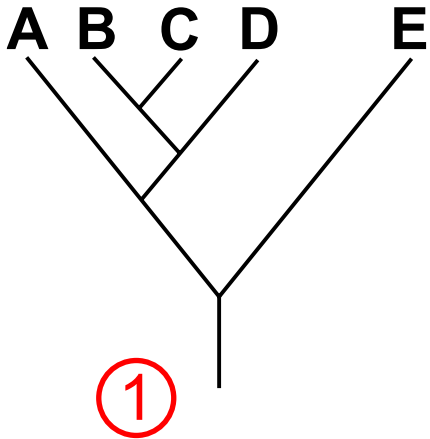
Небинарное дерево следует понимать как множество возможных «разрешений»



Топология дерева

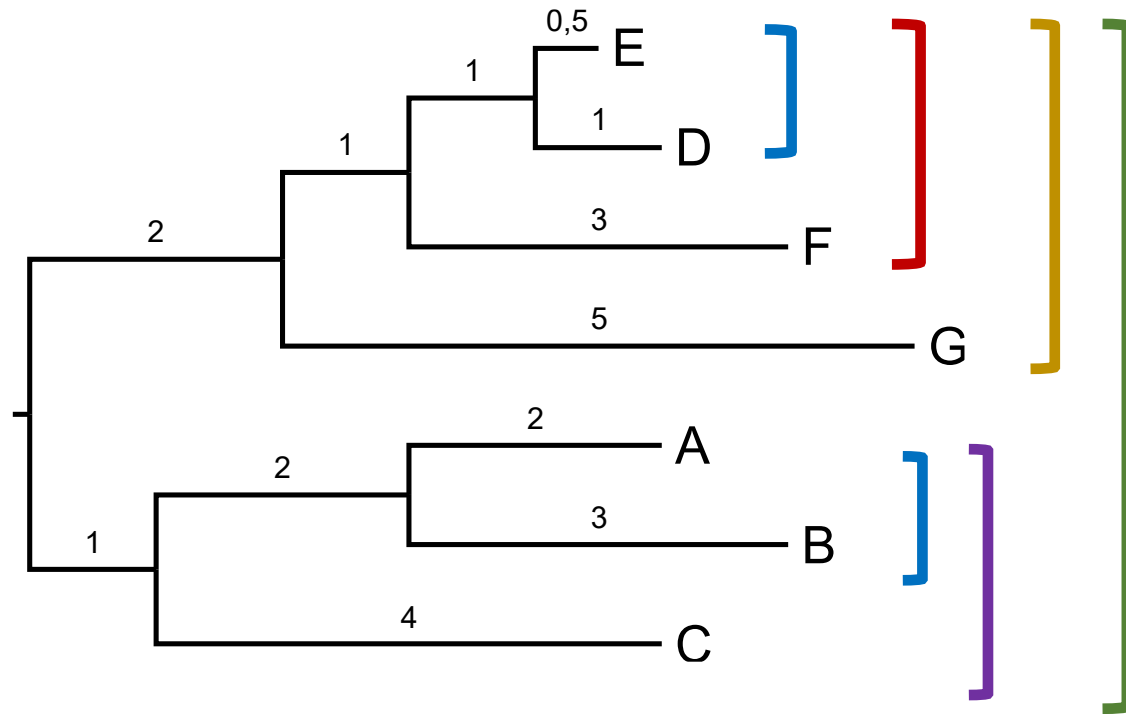


Какие деревья одинаковые?



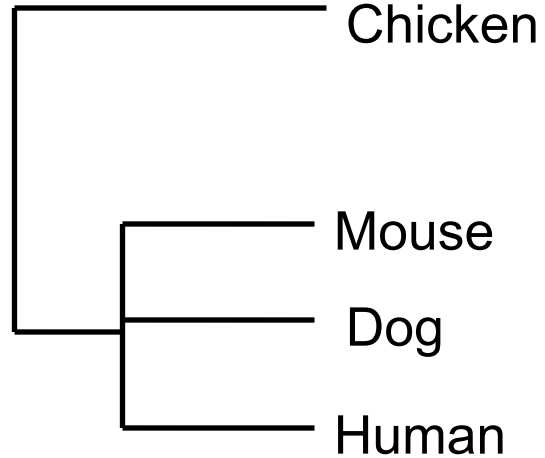
Скобочная формула дерева в формате Newick

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



$((((A:2,B:3):2,C:4):1, (((E:0.5,D:1):1,F:3):1,G:5):2);$

Скобочная формула неразрешенного дерева



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

По каким данным можно строить дерево?

- По внешним признакам организма
- Последовательности биополимеров:
 - по гену,
 - по нескольким генам,
 - по геномам,
 - по последовательностям белков

Молекулярная филогенетика

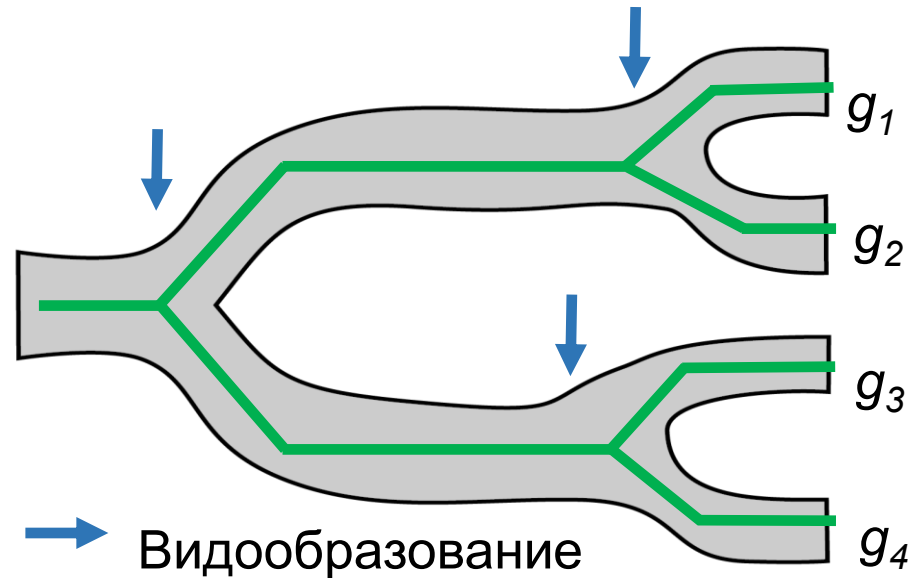
- Молекулярная филогенетика — способ установления родственных связей между живыми организмами на основании изучения структуры полимерных макромолекул — ДНК, РНК и белков.

Молекулярная филогения

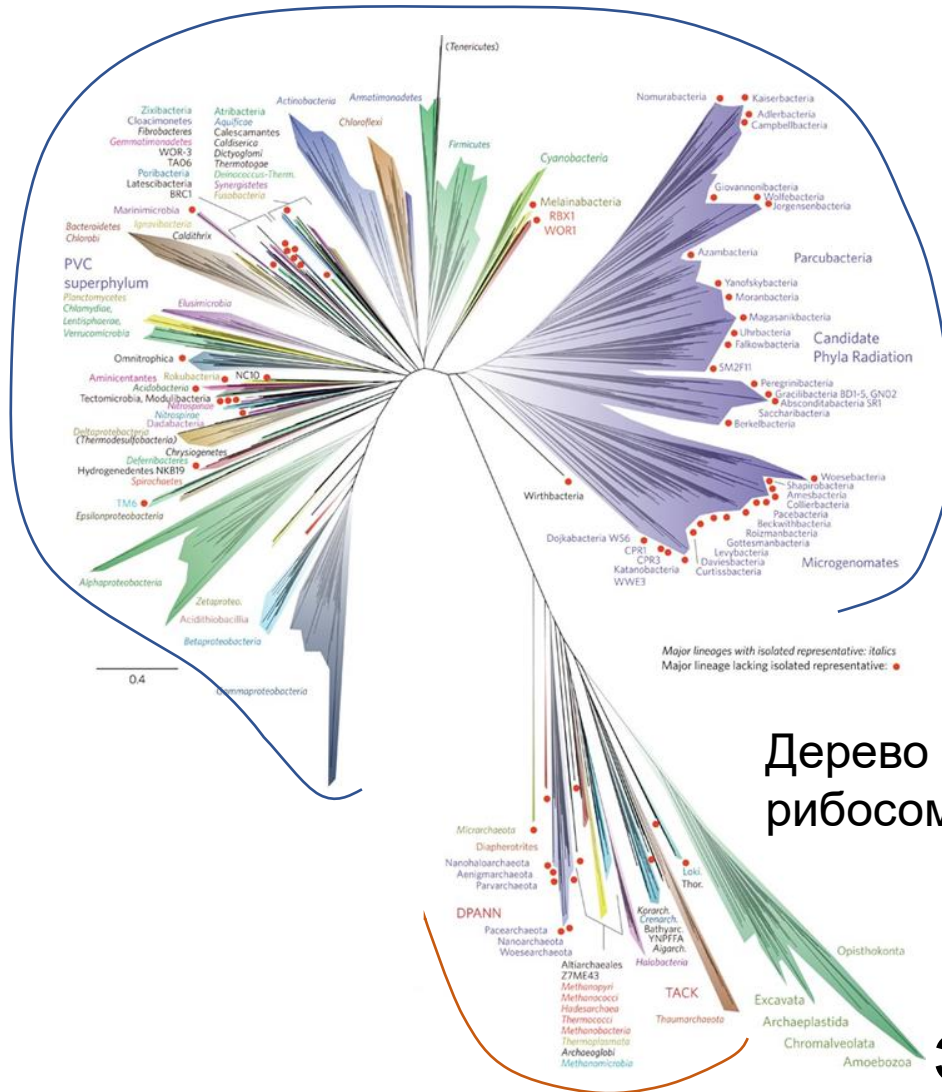
Для изучения эволюции организмов используют **филогенетические маркеры** – гены, эволюция которых отражает эволюцию живых организмов.

Такие гены должны быть редко подвержены:

- потере в ходе эволюции
- горизонтальному переносу генов
- дупликациям



Например: ген 16S рРНК, гены рибосомальных белков, ген цитохрома *b* (у эукариот)



Бактерии

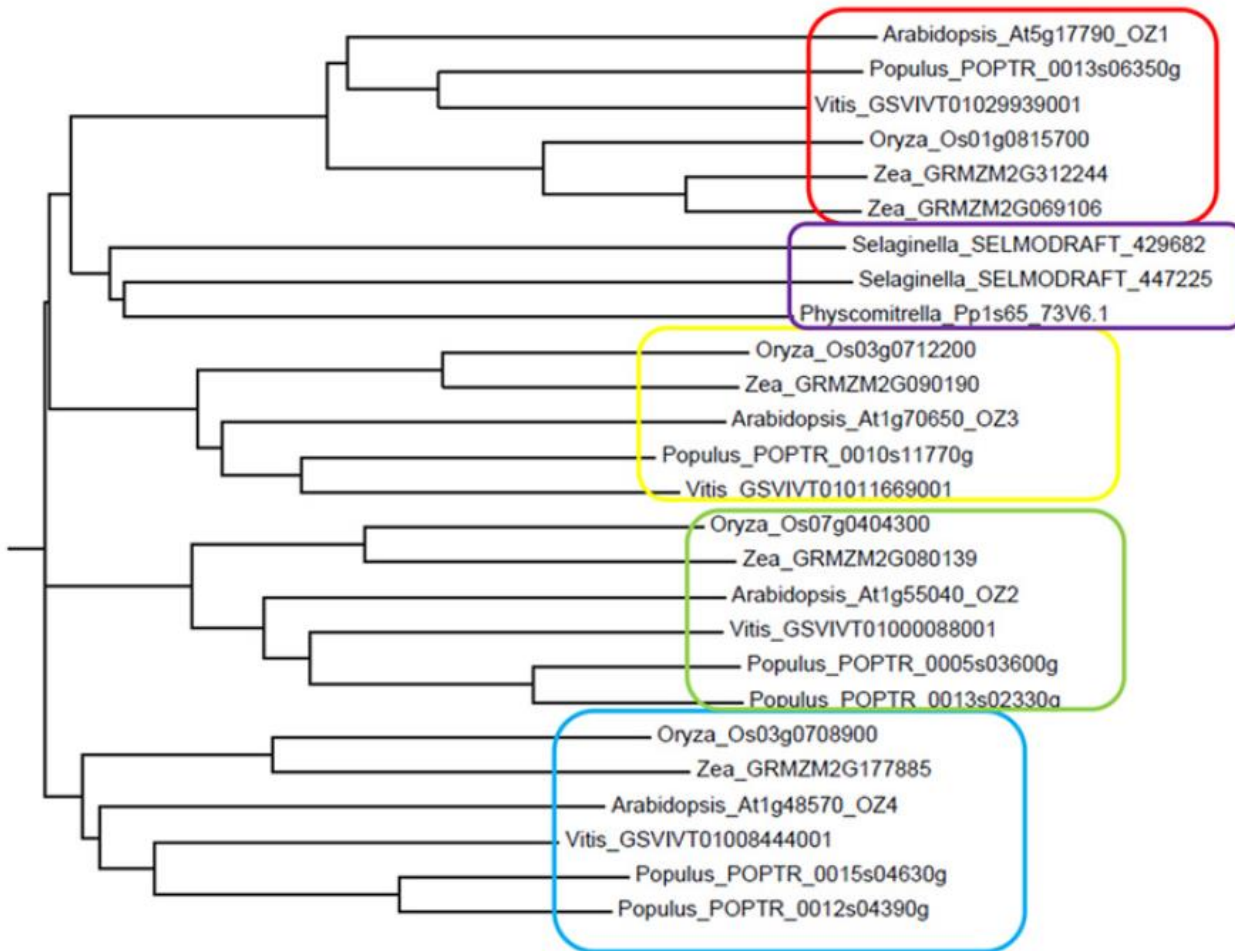
Дерево построено по
рибосомальным белкам

Эукариоты

Археи

При построении дерева может стоять задача анализа эволюции самих белков

Филогенетическое дерево белков семейства OZ (organelle zinc finger editing factor family)



«We were able to identify putative orthologs for all 4 genes in poplar, grape, rice, and maize. Moss and Selaginella, which exhibit chloroplast RNA editing, also encode OZ-like proteins but with a lower similarity. We could not detect proteins similar to the OZ family in Chlamydomonas or Volvox, where editing does not occur».

Три фактора эволюции

Наследственность

ДНК (и белки) за счет процессов репликации и репарации остаются практически неизменными и передаются потомкам

Изменчивость

- У эукариот при половом размножении происходит **комбинация генов родителей**
- У прокариот часто происходит **горизонтальный перенос генов**
- За счет мутаций ДНК (и белки) могут **изменяться**

Отбор

Разные генотипы меняют свою распространенность в популяции с течением времени. Таким образом, вредные мутации из популяции элиминируются, **закрепляются лишь безвредные и полезные мутации**

Классификация мутаций в кодирующих последовательностях ДНК

... ААТССГТСААГТСТА...
... Asn Pro Ser Ser Leu ...

Замена, вставка или делеция одного нуклеотида – *точечная мутация*.

❖ **Синонимичная** (амк остаток не меняется)

... ААТССГТС**Г**АГТСТА...
... Asn Pro Ser Ser Leu ...

❖ **Миссенс (missense)**: меняет амк остаток

✓ на близкий по свойствам

✓ с отличными свойствами

... ААТССГ**А**СААГТСТА...
... Asn Pro **Thr** Ser Leu ...

... ААТССГТСААГ**А**СТА...
... Asn Pro Ser **Arg** Leu ...

❖ **Нонсенс (nonsense)**: заменяет кодон

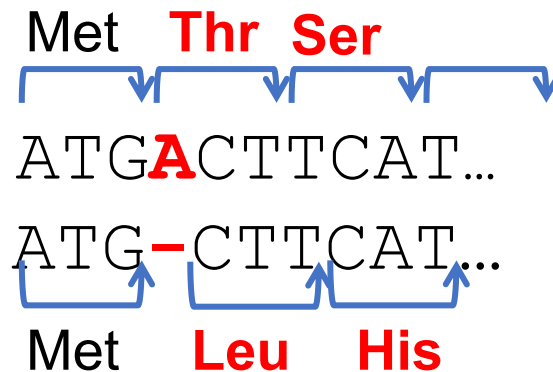
остатка на стоп-кодон

... ААТССГТ**Г**ААГТСТА...
... Asn Pro **Stop** Ser Leu ...

Классификация мутаций в кодирующих последовательностях ДНК

- Сдвиг рамки (**frameshift**): вставка или делеция размера, не кратного трём

Результатом являются совсем другие аминокислоты после мутации и, как правило, стоп-кодон сравнительно недалеко (в среднем через 21 триплет после мутации)



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 секвенируем 1 000 000 бактерий этого штамма через 10 лет?

Судьба нейтральной мутации

Предположим, что мутация **нейтральна** = никак не влияет на мат. ожидание 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 мат. ожидание 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).

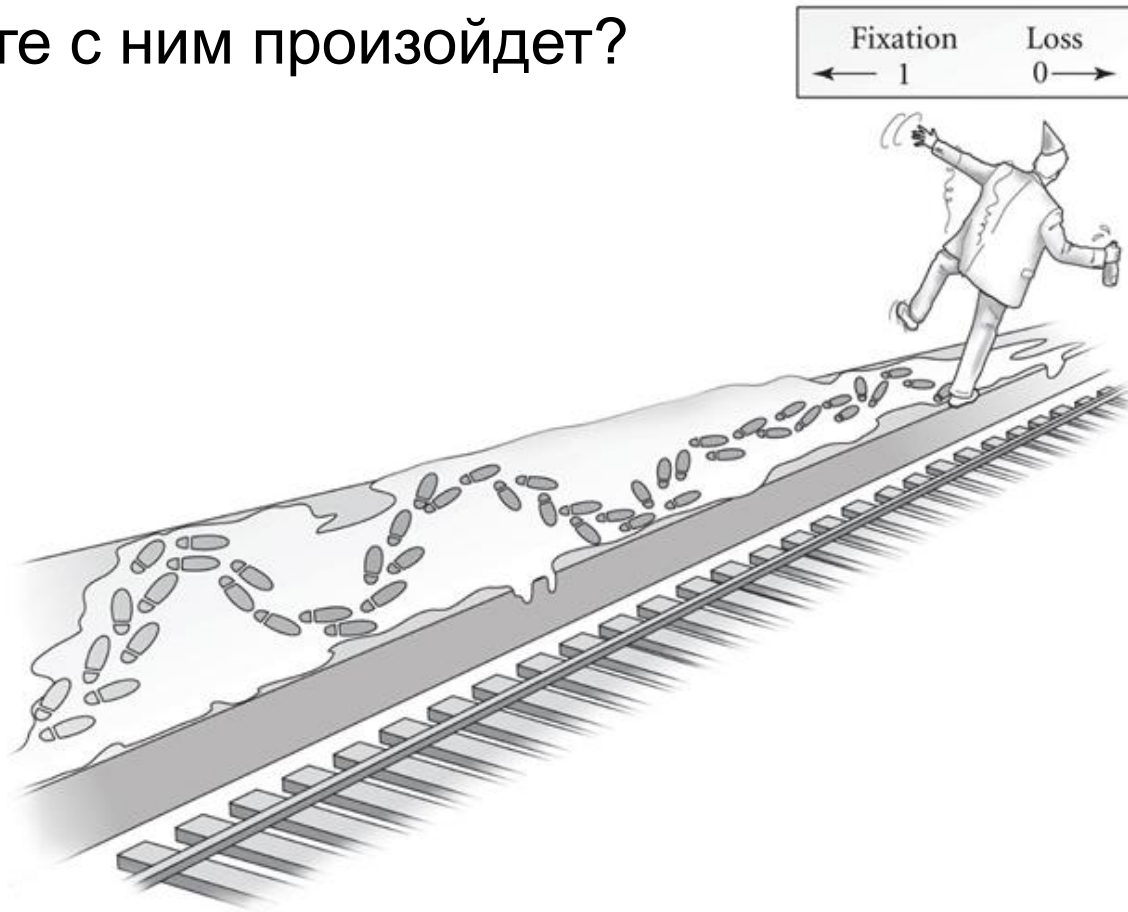
Что произойдет с частотой такого варианта генома через несколько поколений?

Ответ: частота либо немного возрастет, либо немного упадет. То и другое примерно равновероятно.

Судьба нейтральной мутации через много поколений

На тротуаре стоит пьяный и каждые 10 сек. делает шаг либо направо, либо налево, случайно выбирая направление.

Что в итоге с ним произойдет?

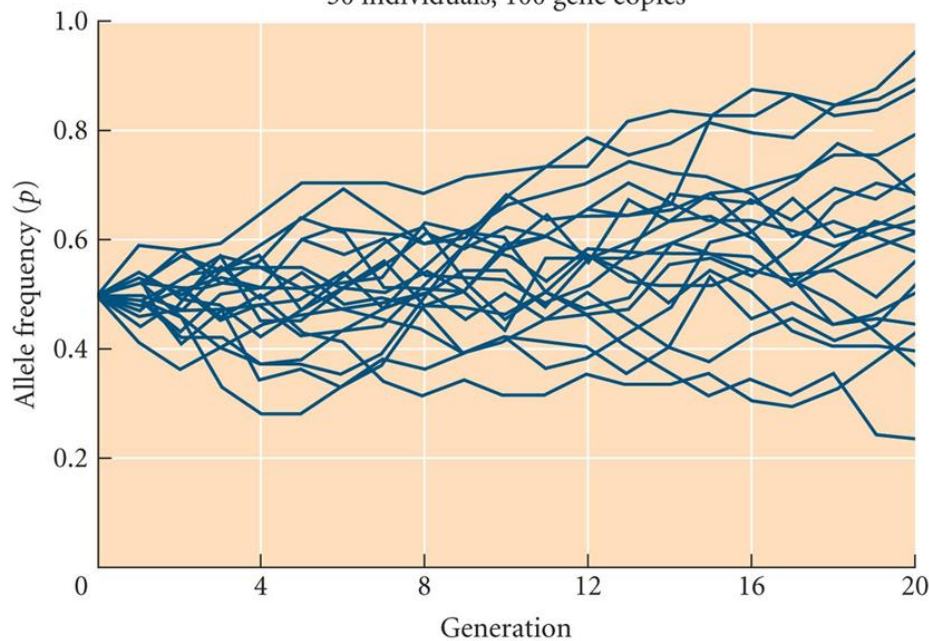


Судьба нейтральной мутации через МНОГО ПОКОЛЕНИЙ

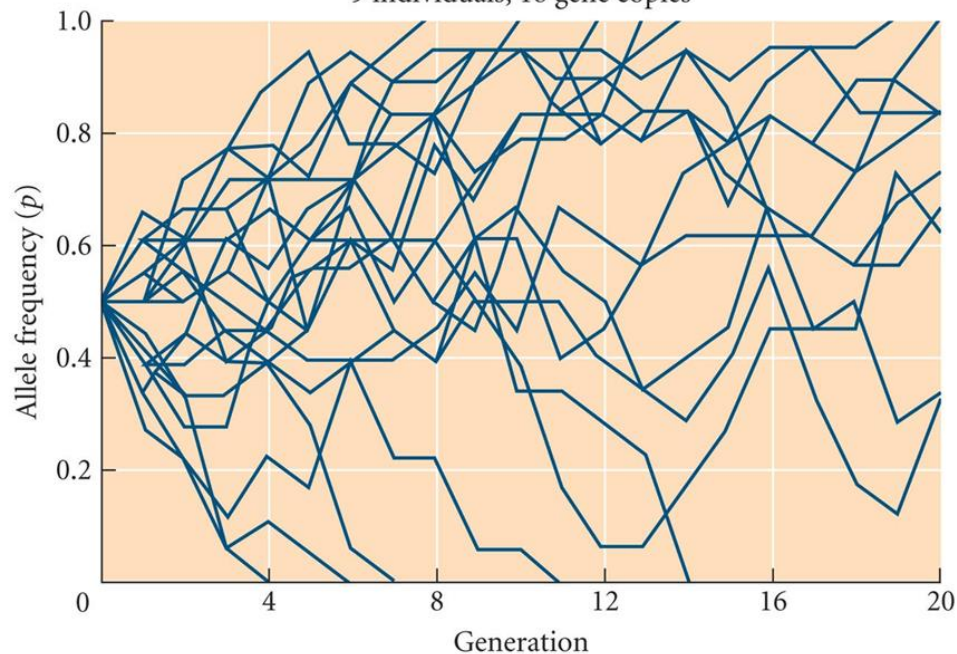
Частота любого нейтрального полиморфизма постоянно колеблется случайным образом. Это называется **«генетический дрейф»**.

Математическая модель такого процесса называется **«случайное блуждание»**.

50 individuals, 100 gene copies

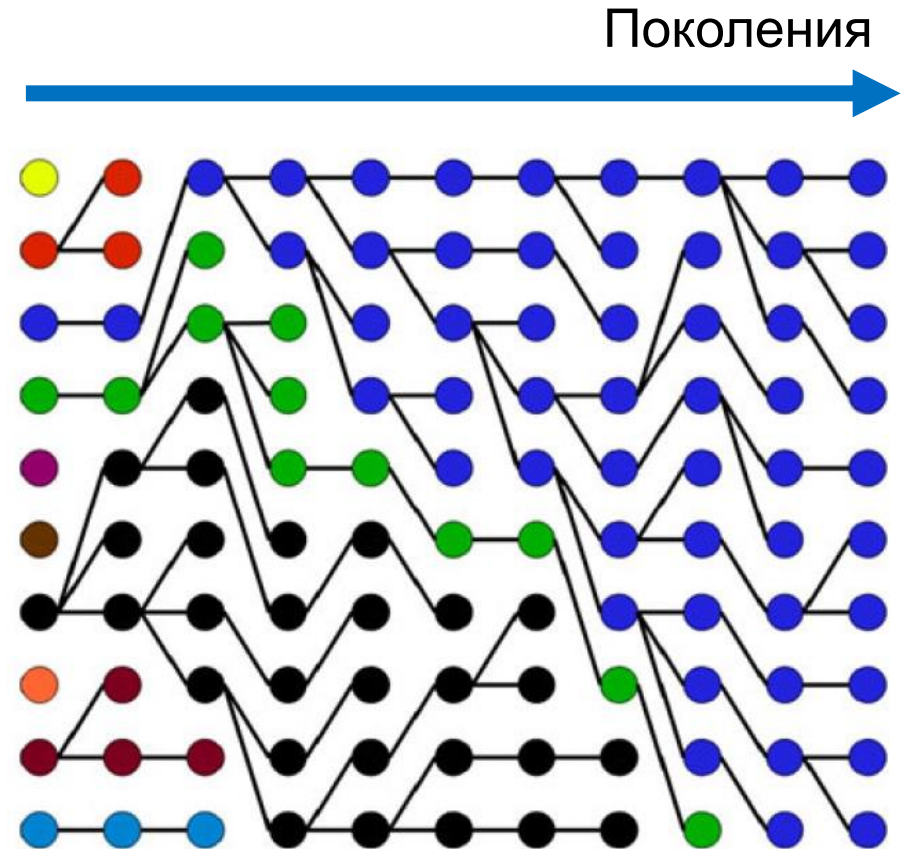


9 individuals, 18 gene copies



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

- 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.
- Вероятность закрепиться для новой нейтральной мутации очень мала, но не 0.
- Итог: геномы независимых популяций начинают различаться, чем дальше, тем больше – **в них независимо накапливаются нейтральные мутации.**
- **С какой скоростью в этом случае накапливаются мутации в разных ветвях?**

Накопление нейтральных мутаций

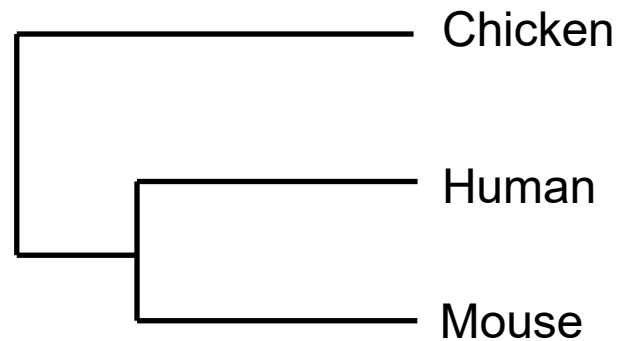
- Организмов в популяции много, мутаций в них происходит тоже много (примерно 10^{-8} на п.н. на поколение – каждая сотая новорождённая бактерия несёт новую мутацию). Значительная доля мутаций нейтральна.
- Вероятность закрепиться для новой нейтральной мутации очень мала, но не 0.
- Итог: геномы независимых популяций начинают различаться, чем дальше, тем больше – **в них независимо накапливаются нейтральные мутации.**
- **С какой скоростью в этом случае накапливаются мутации в разных ветвях?**

Накопление нейтральных мутаций

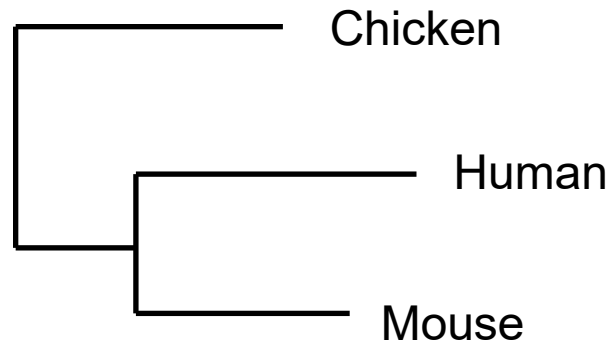
- Организмов в популяции много, мутаций в них происходит тоже много (примерно 10^{-8} на п.н. на поколение – каждая сотая новорождённая бактерия несёт новую мутацию). Значительная доля мутаций нейтральна.
- Вероятность закрепиться для новой нейтральной мутации очень мала, но не 0.
- Итог: геномы независимых популяций начинают различаться, чем дальше, тем больше – **в них независимо накапливаются нейтральные мутации.**
- **С какой скоростью в этом случае накапливаются мутации в разных ветвях?**
- Если мутации нейтральны, а скорость возникновения мутаций не изменилась, то с одинаковой! \Rightarrow «молекулярные часы»

Длины ветвей на дереве могут отражать число мутаций

«Молекулярные часы» работают



Но скорость эволюции может быть неравномерна, тогда получается такое дерево



А если мутация не нейтральна?

- Каждому варианту генома можно сопоставить его «приспособленность» f = мат. ожидание 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**.

А если мутация не нейтральна?

- 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?**

Положительный отбор порождает длинные ветви на дереве

- Дерево строилось по объединенному выравниванию 13 ортологичных белков.
- Даже на дереве по последовательностям видна длинная ветвь!



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.)
- ...

Востребованность

Упоминание некоторых филогенетических программ в аннотациях статей, вышедших после 1 января 2025 (по PubMed)

- RAxML 26 раз
- IQ-Tree 45 раз
- MrBayes 23 раза
- BEAST 113 раз
- Neighbor-joining 126 раз
- UPGMA 97 раз
- Maximum parsimony 83 раза