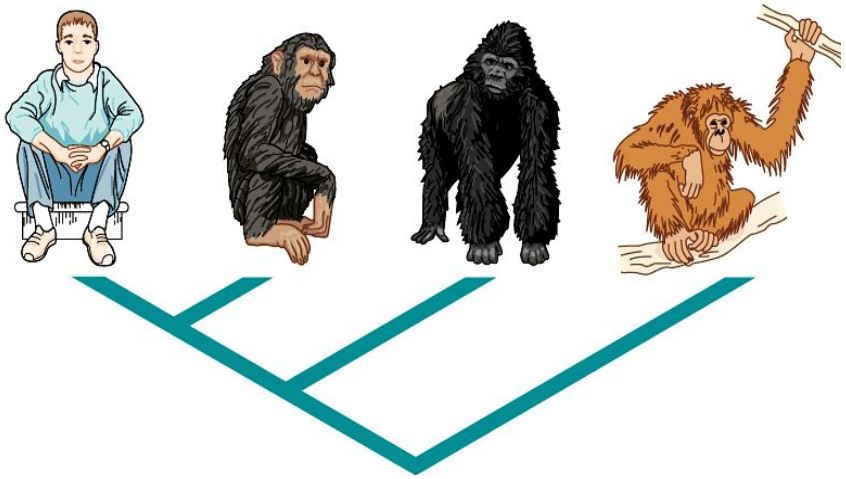
Молекулярная эволюция и филогенетика
Лекция 1

(Введение, основные понятия и основные свойства филогенетических деревьев)



Лухтанов Владимир Александрович

- кафедра энтомологии СПбГУ
- отделение кариосистематики Зоологического института РАН
- Тел. +7(960) 2758672
- lukhtanov@mail.ru

V. Lukhtanov

- **University of Florida:** curator at the Lepidoptera collection (Palearctic butterflies)
  http://www.flmnh.ufl.edu/mcguire/profiles/lukhtanov.htm

- **Harvard University:** associated researcher at Museum of Comparative Zoology
  http://www.mcz.harvard.edu/Departments/Entomology/people.html

CV

- http://www.entomology.bio.pu.ru/lukhtanov_front.htm
- http://www.bio.spbu.ru/staff/id110_val.php

- Прогресс в науке связан с появлением и развитием новых методов научного анализа
- Сравнительный метод (XVIII и XIX век)

Gregor Mendel
(1822-1884)

генетический анализ (1865, начало XX века)

выявление генетических связей между особями и популяциями

Однако за рамками генетического (гибридологического) анализа остается гигантский круг проблем и биологических объектов:

- генетические связи существуют не только между особями и популяциями, они есть и между видами и эволюционными линиями

- Но как эти связи выявить, если скрещивания между отдаленно родственными организмами невозможны?

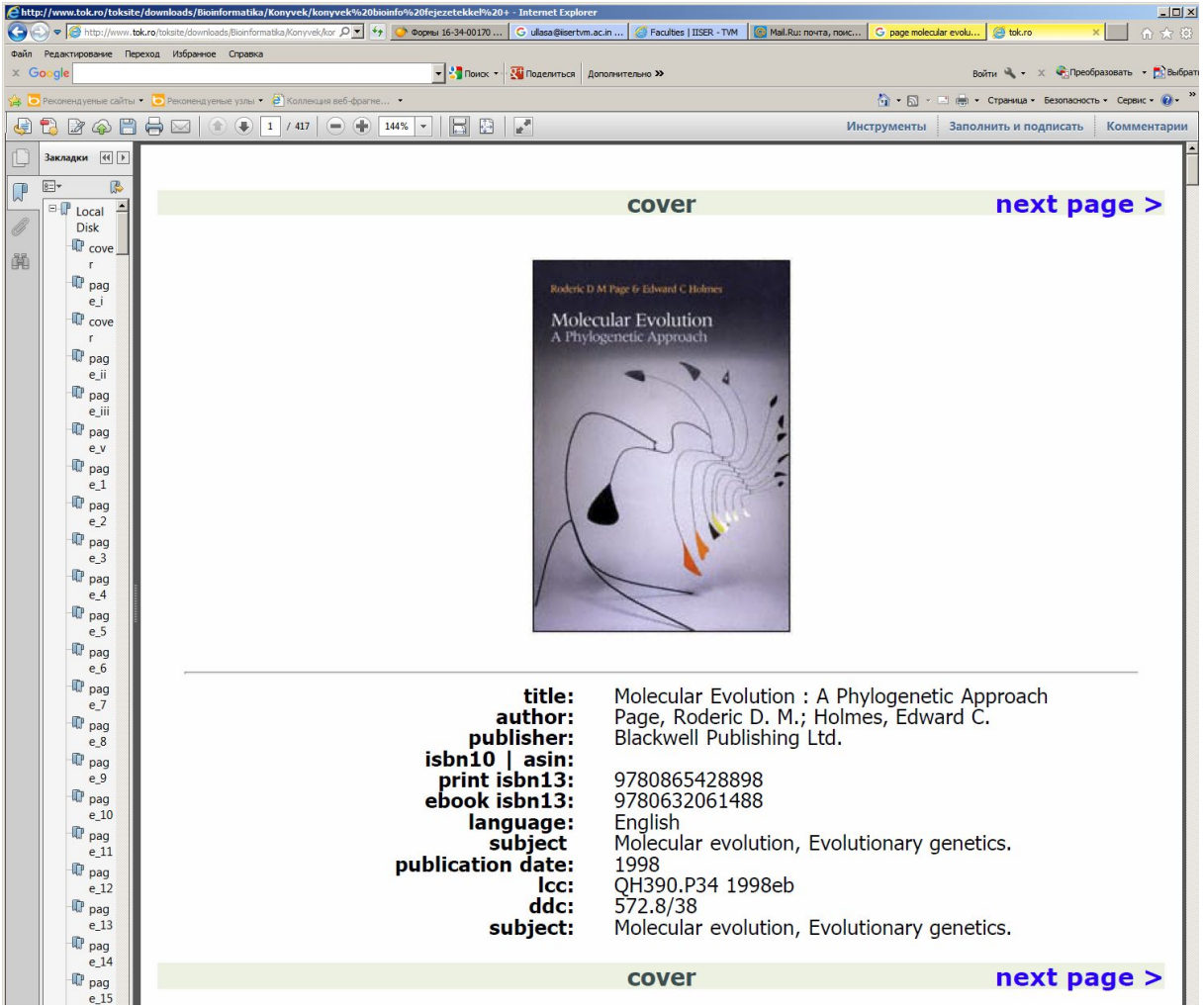
Филогенетический анализ:

выявление генетических (генеалогических) связей между видами
и эволюционными линиями более высокого ранга

В чем отличие этого курса от других подобных?

- Поваренные книги (cookbooks)
- Серьезные монографии

- http://www.tok.ro/toksite/downloads/Bioinformatika/Konyvek/konyvek%20bioinfo%20fejezetekkel%20+%20bioinfo%20konyvek/halado/Molecular%20Evolution%20A%20Phylogenetic%20Approach.pdf



Структура курса

- Лекции
- Практические занятия по освоению программ филогенетического анализа
- Самостоятельная работа: филогенетический анализ собственных (или самостоятельно извлеченных из ГенБанка) молекулярных данных
- Зачет: обсуждение самостоятельной работы + вопросы по курсу

Задача курса:

не просто рассказать, **что** можно сделать с использованием молекулярных признаков,

но и показать, **как** это можно сделать

вообще

и своими руками в частности

- ДНК хранит информацию о плане строения организма,
  - и эту информацию организм использует для построения самого себя.
- Прочесть ее не очень трудно – есть буквы (нуклеотиды), и есть генетический год, который однозначно определяет связь между ДНК и белком
- Хотя в реальности все сложнее, но суть остается прежней: мы читаем эту запись в ДНК как книгу.

- Но эта не единственная информация, хранящаяся в ДНК и других макромолекулах, есть еще информация об истории организмов, об их прошлом.
- Но извлечь ее гораздо труднее, более того – извлечь ее можно только при сравнении нескольких объектов
- В курсах генетики и молекулярной биологии говорилось об информации первого типа (кодирование белков)
- В филогенетике – расшифровка информации второго типа – об истории.

Эволюционная биология – историческая наука.
Общая методология всех исторических наук: использованием вещественных свидетельств (археология, палеонтология) + документы (записи).
В биологии до недавнего времени документов не было, пока нее научились читать ДНК

Документы (**ДНК**) + **Материальные свидетельства (ископаемые формы)**

Реконструкция истории

- - Что можно делать с использованием ДНК?
- – Расшифровывать эволюционную историю организмов!

Расшифровка истории на разных уровнях

The time will come, I believe, though I shall not live to see it, when we shall have fairly true genealogical trees of each great kingdom of Nature.

Charles Darwin

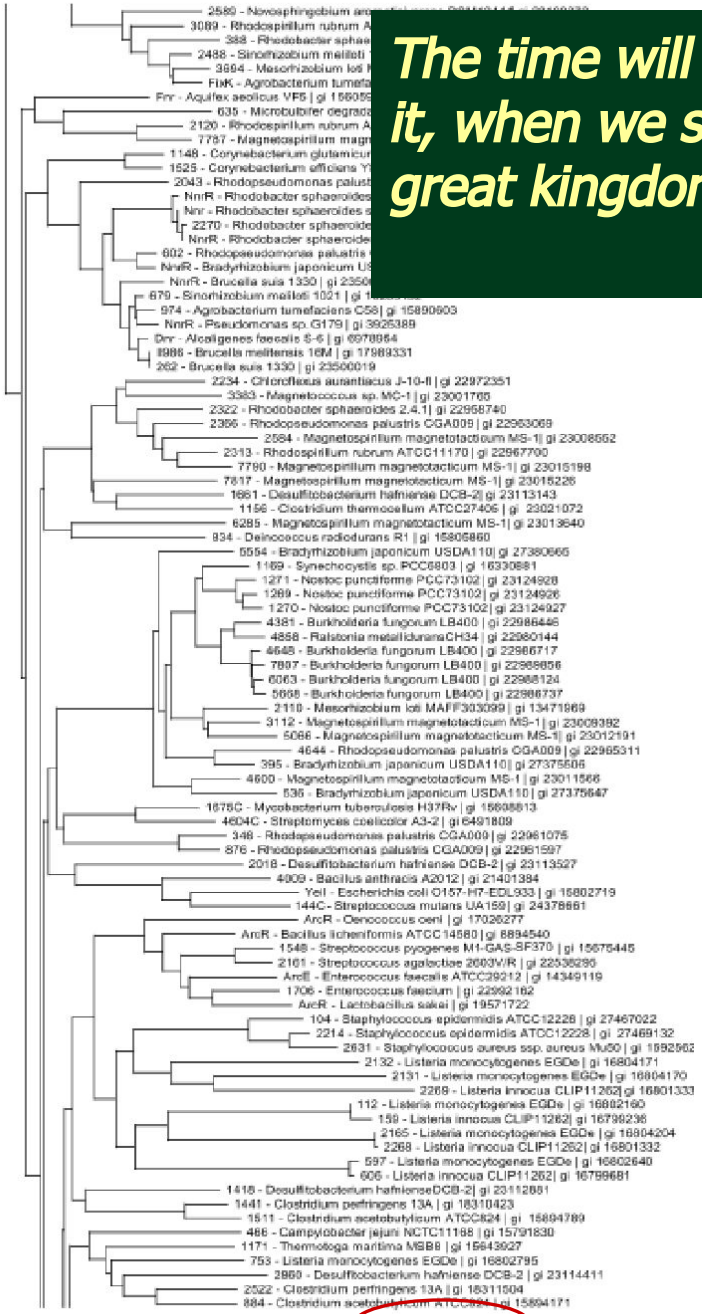
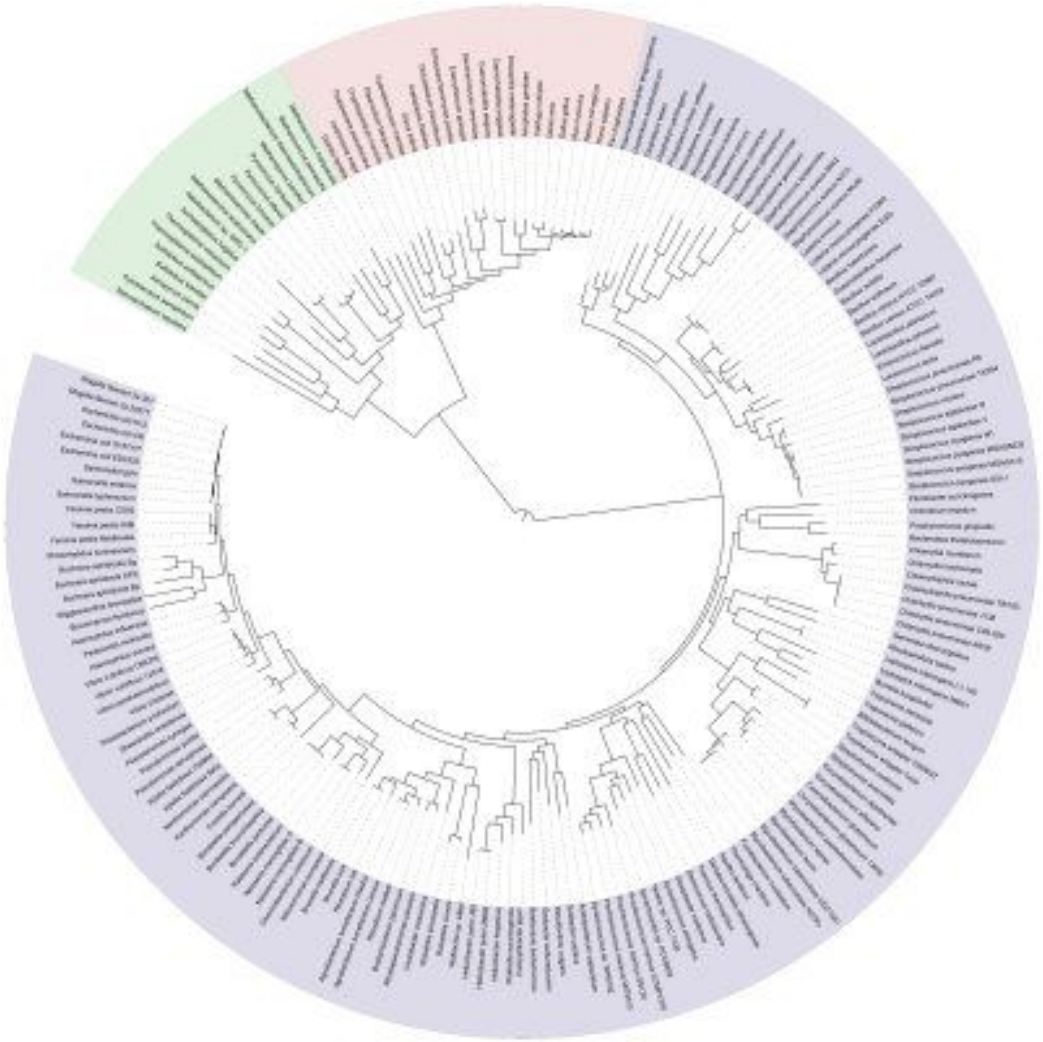
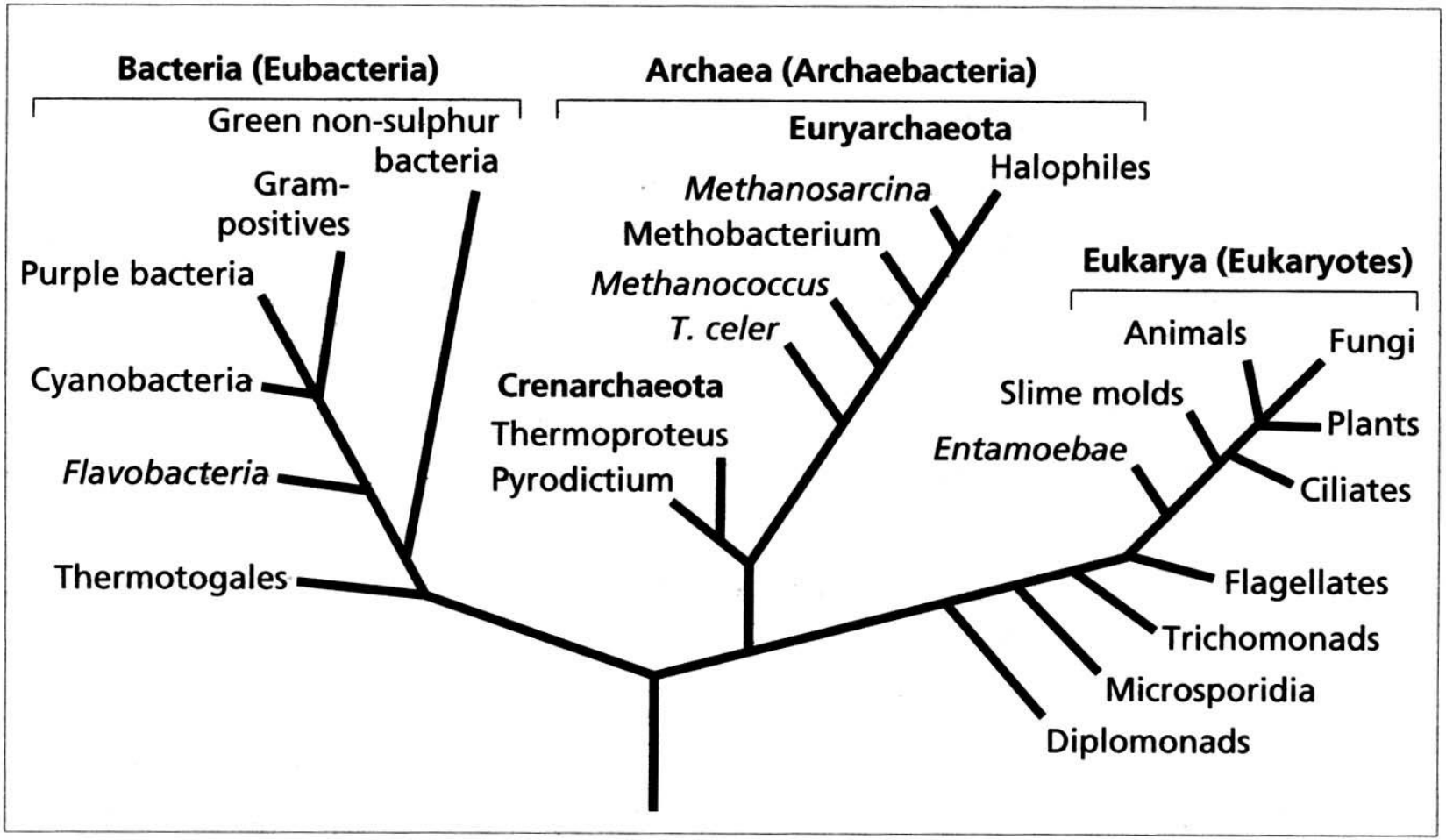


Fig. 1 (*Continued*).



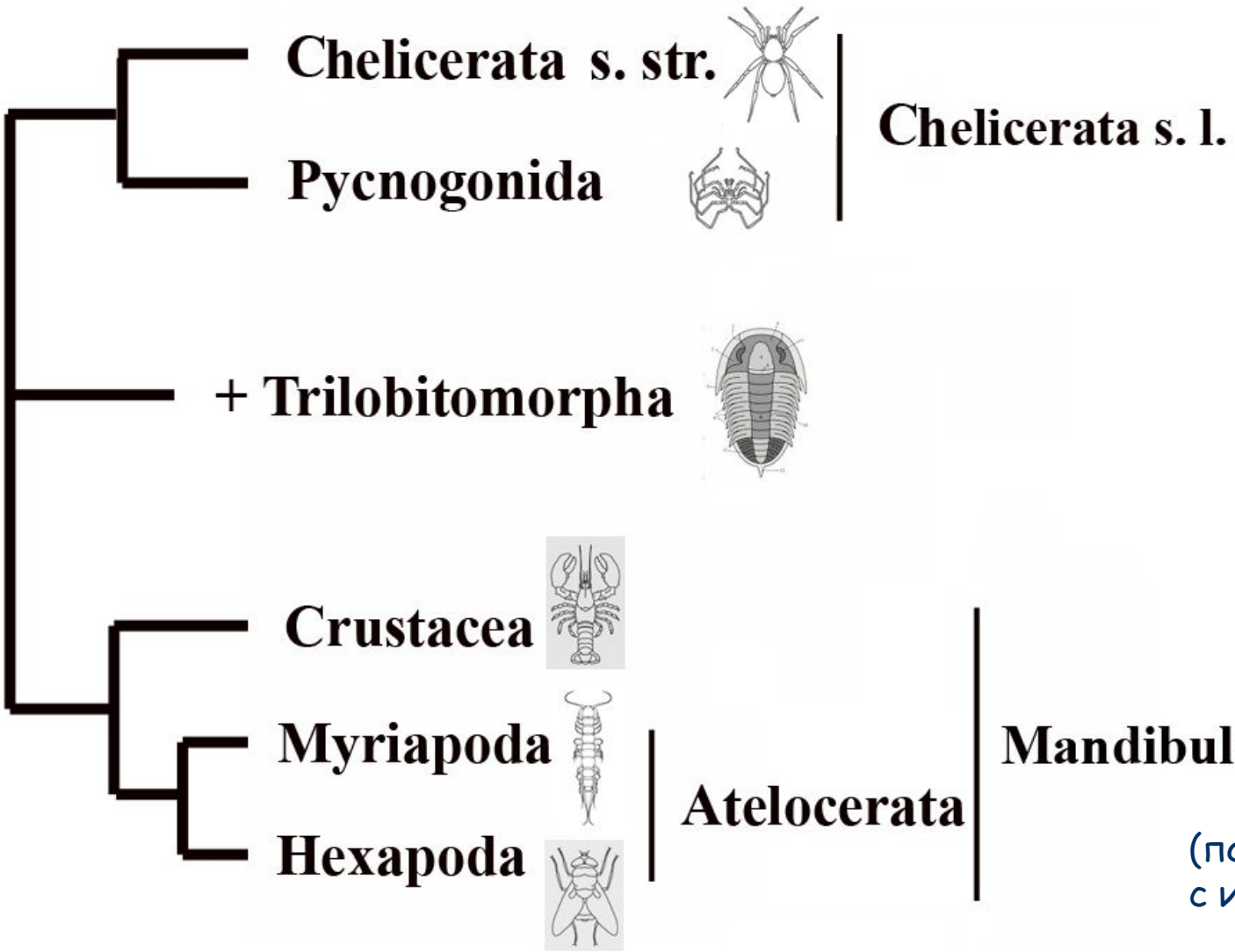
Филогенетические отношения между живыми организмами (на основании анализа рибосомальной РНК)



Phylogenetic relationships between members of the three domains of cellular life—the Archaea, Bacteria and Eukarya, based on rRNA.

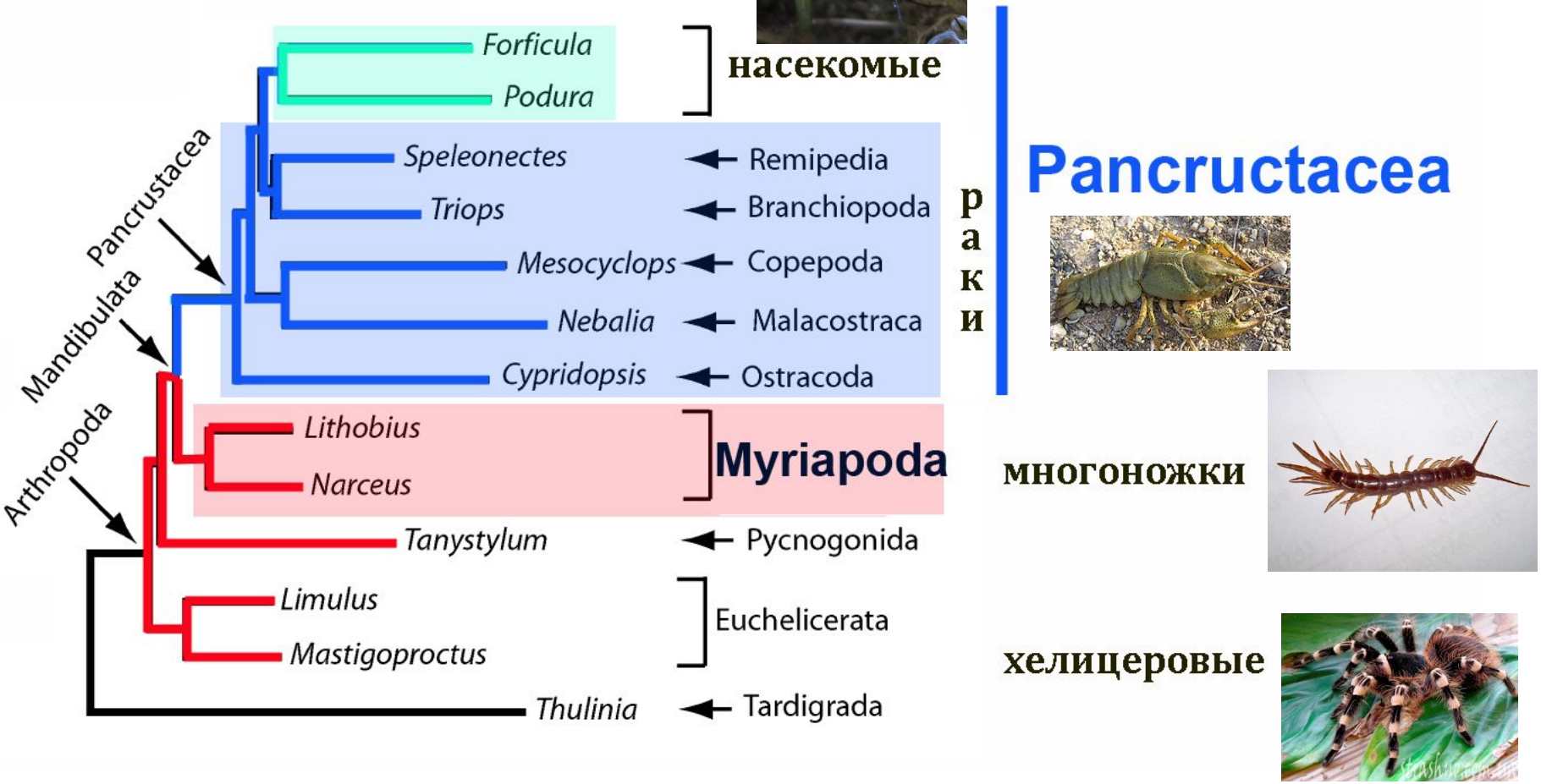
Представление о филогении членистоногих, которое недавно считалось классическим:

насекомые (Hexapoda) и многоножки (Myriapoda) – сестринские группы

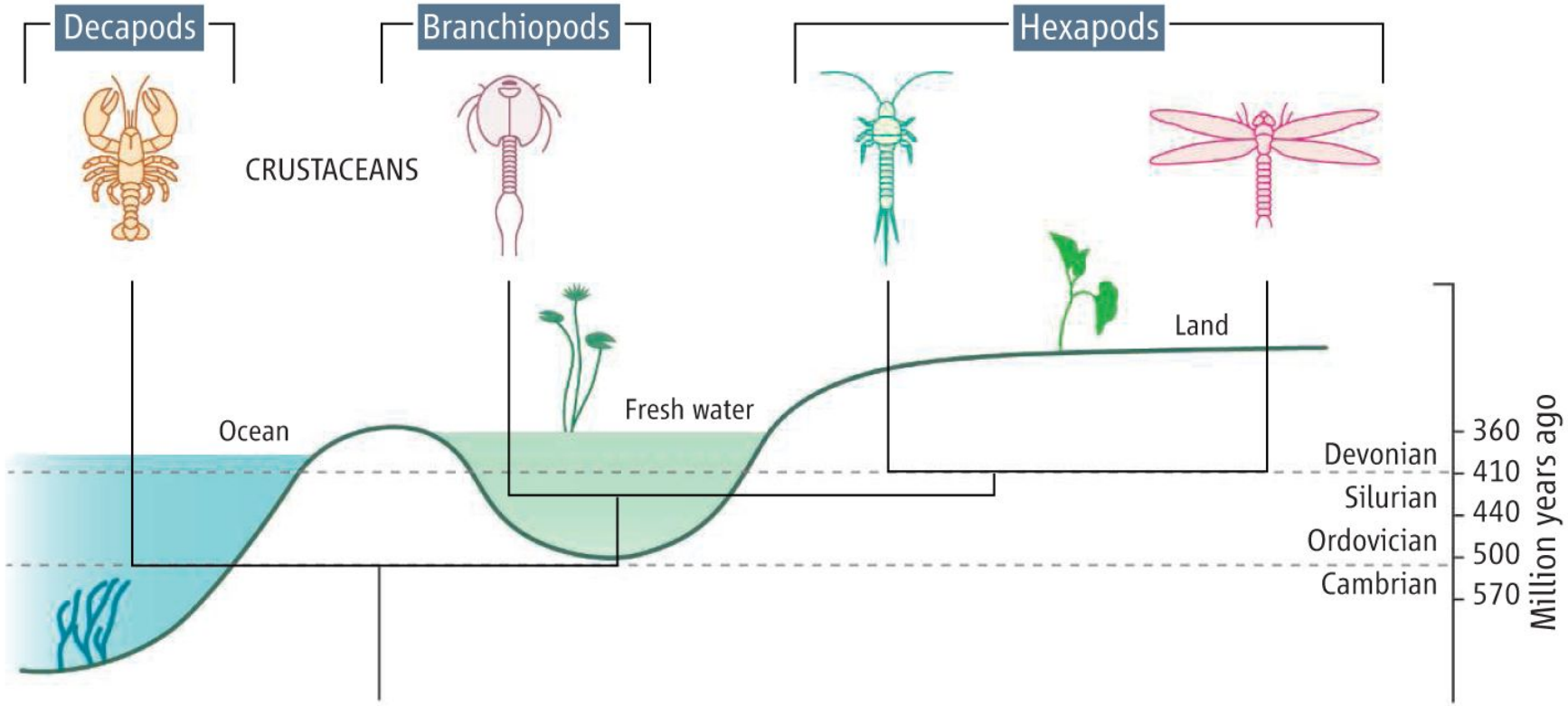


(по: Клюге, 2000, с изменениями)

Молекулярная филогения членистоногих (Regier et al., 2008



Происхождение насекомых:
Морские раки >пресноводные раки>насекомые



Hexapod evolution. The last common ancestor of hexapods and crustaceans (branchiopods, specifically) may have originated in freshwater during the Late Silurian, giving rise to extant freshwater dwelling branchiopods (fairy shrimps, water fleas, and tadpole shrimps) and insects. This hypothesis accounts for the missing fossil record of branchiopods and hexapods before the Devonian.

даман

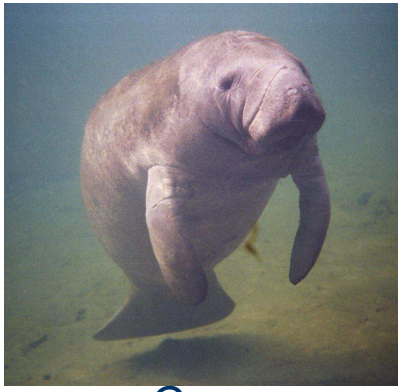
Клада Afrotheria выделяется на основании молекулярных данных

Figure 14.16 Phylogeny of the mammals This tree is from Murphy, W. J. et al. (2001). Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science* 294: 2348–2351.

Трубкозуб

Прыгунчики

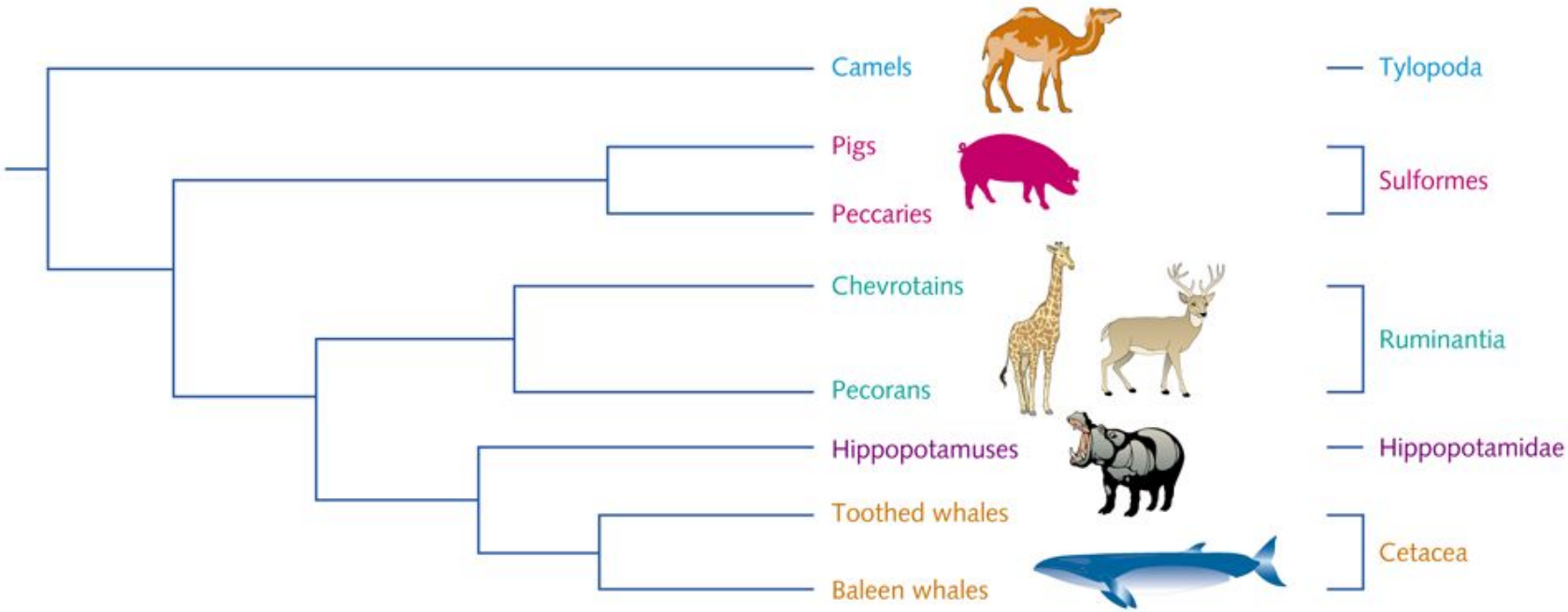
Златокроты



Сирены

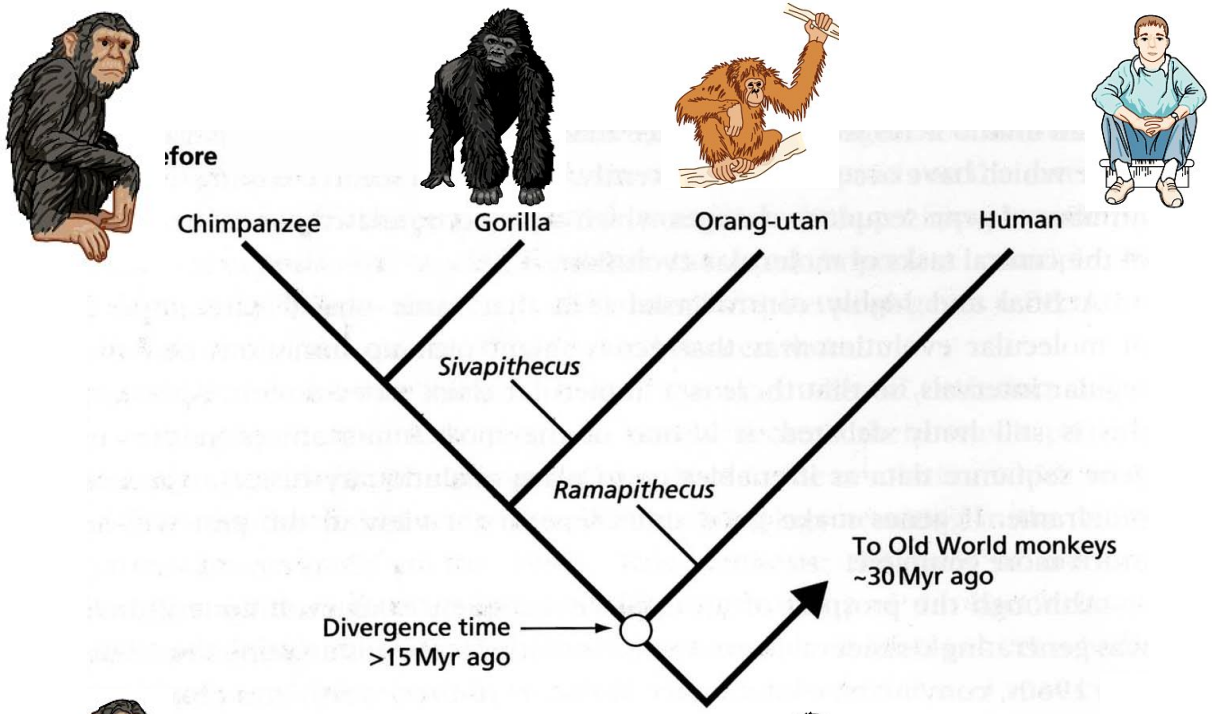
Даманы

Молекулярная систематика отряда китопарнокопытных

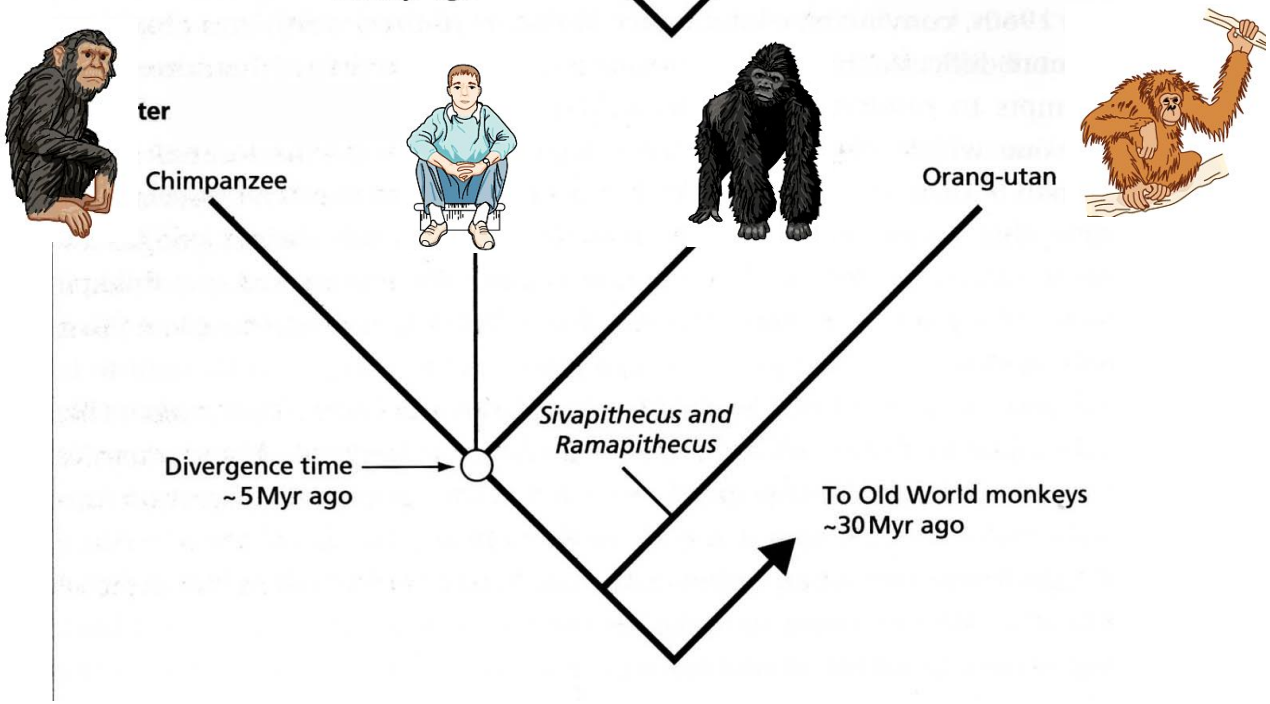


The closest relatives of whales are hippopotamuses, and whales are clearly ungulates—highly modified for life in the sea. (Reproduced from Nikaido et al. (1999) Phylogenetic relationships among cetartiodactyls based on insertions of short and long interpersed elements. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 10261–6. National Academy of Sciences, USA.)

Происхождение человека

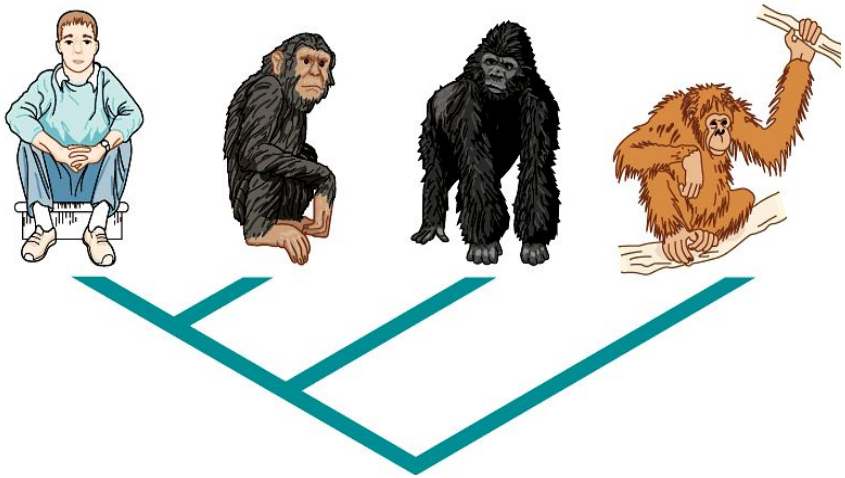


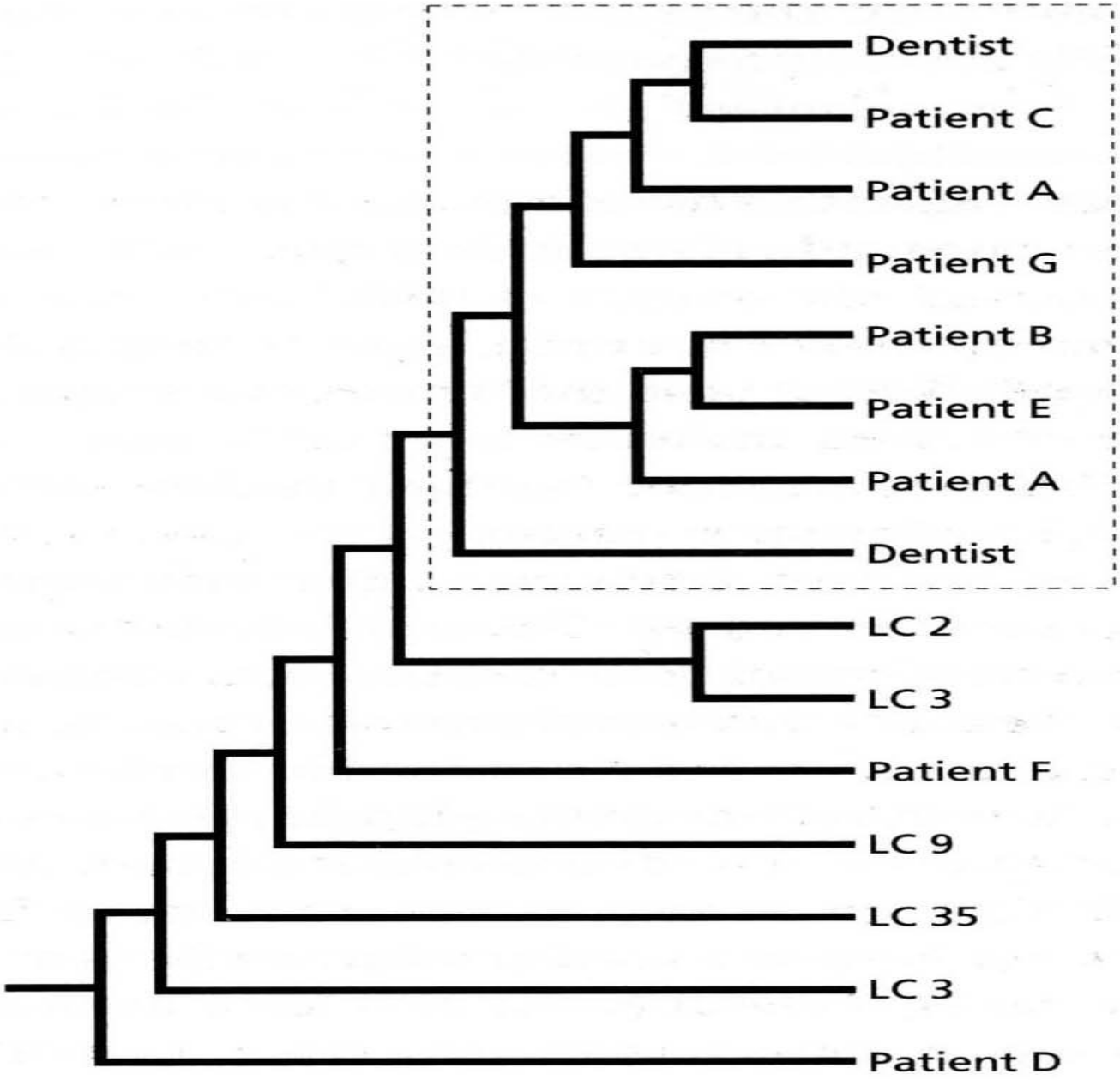
До появления молекулярных данных



После появления первых молекулярных данных

Молекулярная эволюция и филогенетика





История одной ВИЧ-инфекции (история флоридского стоматолога) (*Science*, 1992, 256:1165-1171)

Молекулярные филогении

- В систематике
- В молекулярной биологии
- В эволюционной биологии
- Систематика – молекулярная филогенетика ее естественная часть
- Молекулярная биология - молекулярная филогенетика также ее естественная часть, если говорить об эволюции макромолекул
- Эволюционная биология - молекулярная филогенетика также ее естественная часть, костяк для любой эволюционной идеи

- Но и
- Экология, этология… да и все остальные разделы биологии
- Медицина, ветеринария, криминалистика и многое другое

Основные понятия филогенетики и свойства филогенетических деревьев: термины, определения, аксиомы и некоторые теоремы

Деревья (trees)

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

Ч. Дарвин

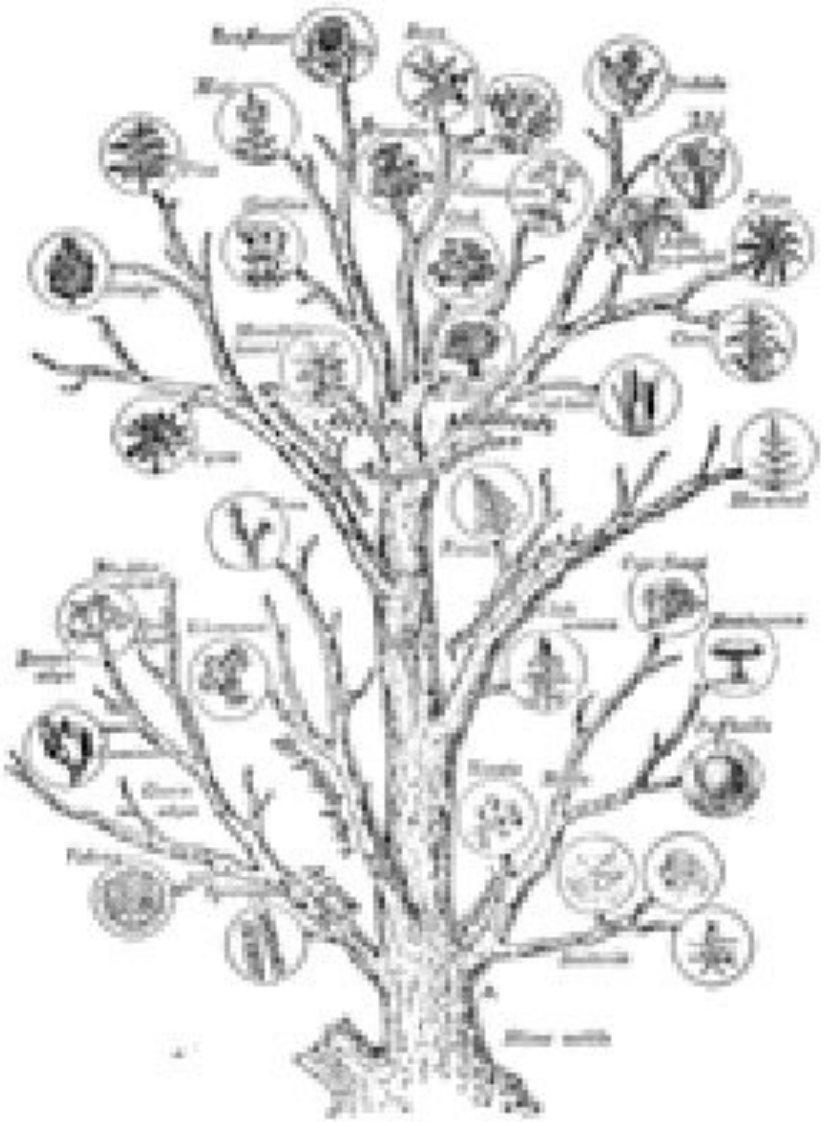
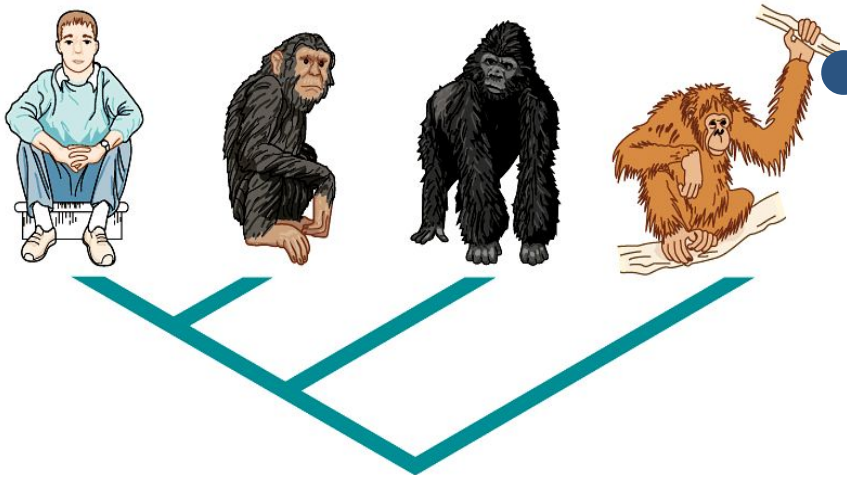


Fig. 19. Genealogical tree of plant life

This diagram is intended to suggest the common origin of all plant forms, with the constant progressive departure from ancestral types, now in one direction and now in another, like the branching of a tree. Lower and higher more nearer to or further from the original types. The closer together two forms are on a given branch, the more closely related they are considered (cf. Fig. 19)

Задача построения филогенетического дерева

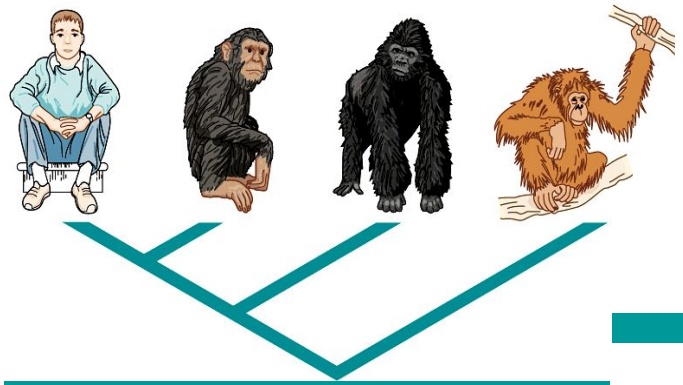


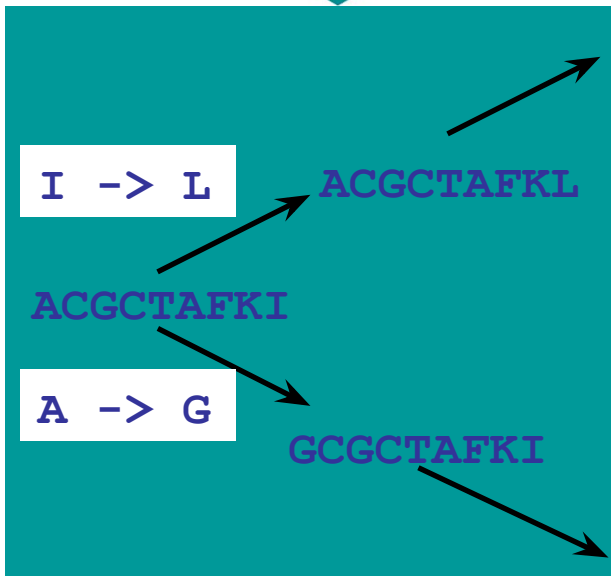
- **<u>Биологические задачи</u>:**
  - **сравнение 3-х и более объектов**
    (кто на кого более похож ....)
  - **реконструкция эволюции**
    (кто от кого, как и когда произошел…)

- **<u>Математическая задача</u>** – задача комбинаторной оптимизации: на основе [биологических] данных получить дерево, возможно лучше согласованное с этими данными.

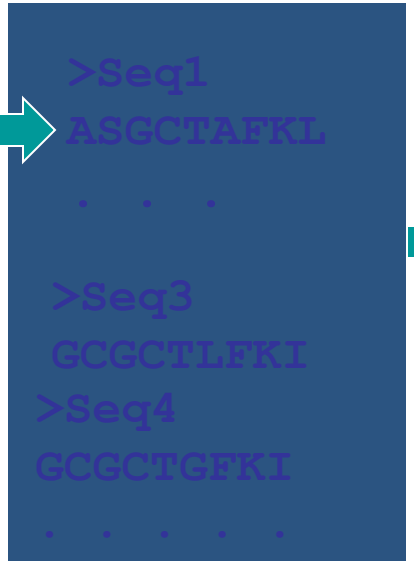
Реальные события : **Признаки:** **Построенное дерево**

эволюция в природе или в лаборатории, компьютерная симуляция

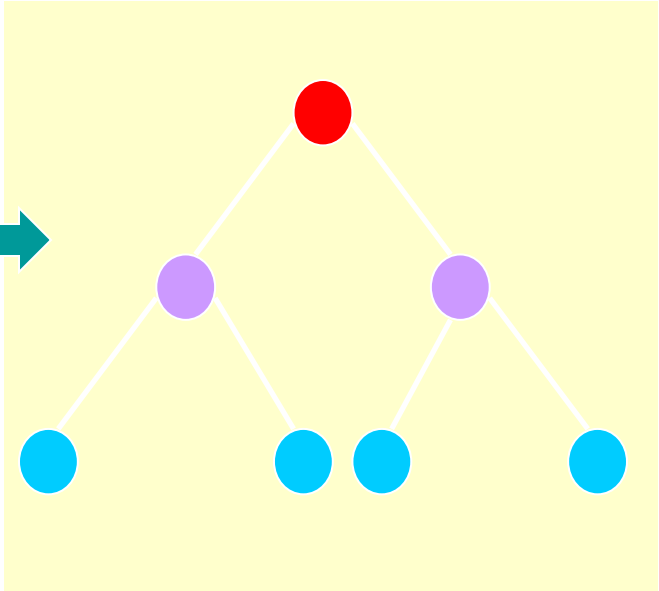




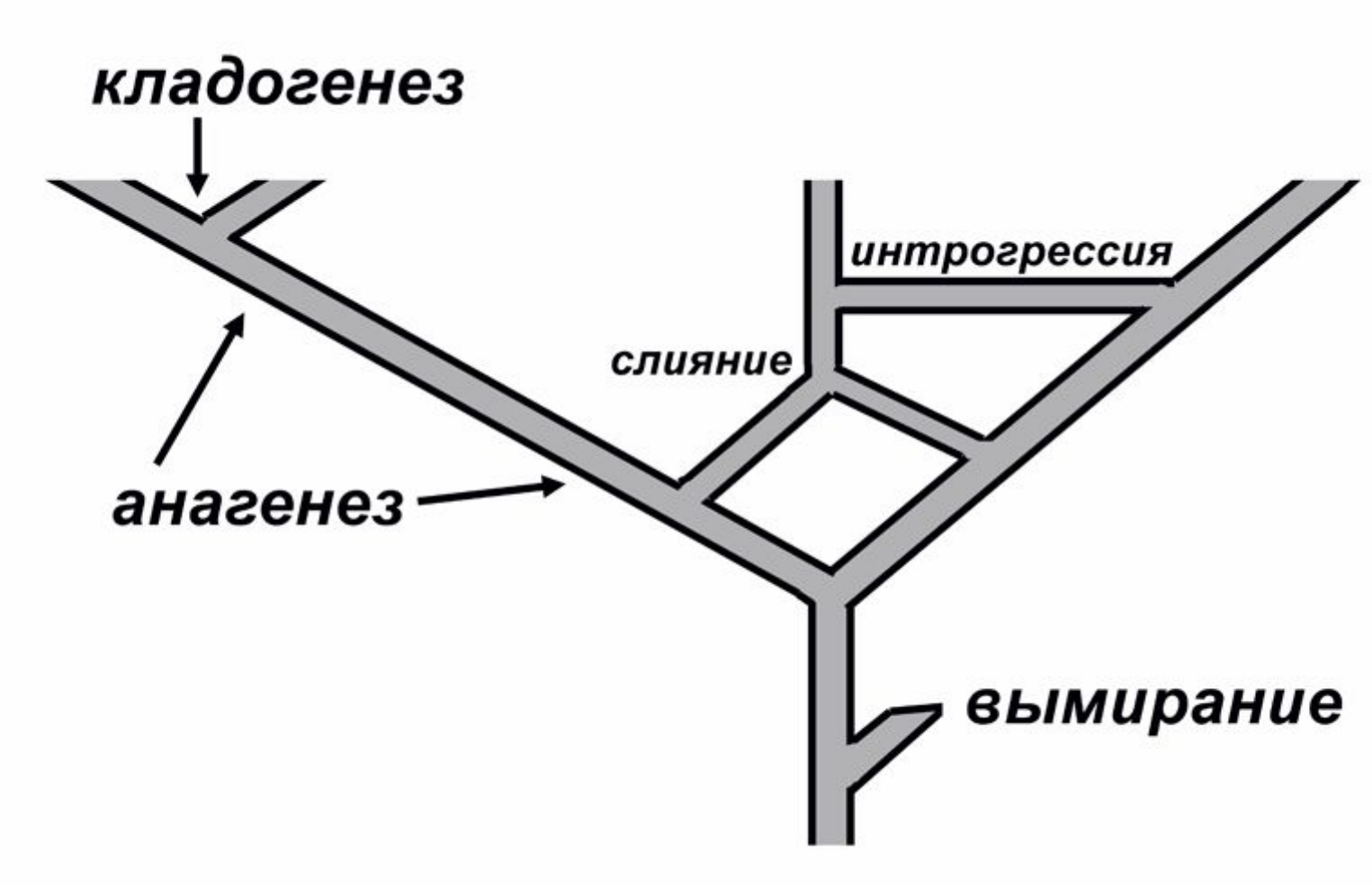
например, а.к. последовательности или количество щетинок



древовидный граф, ***вычисленный на основе данных***, может отражать или не отражать реальные события



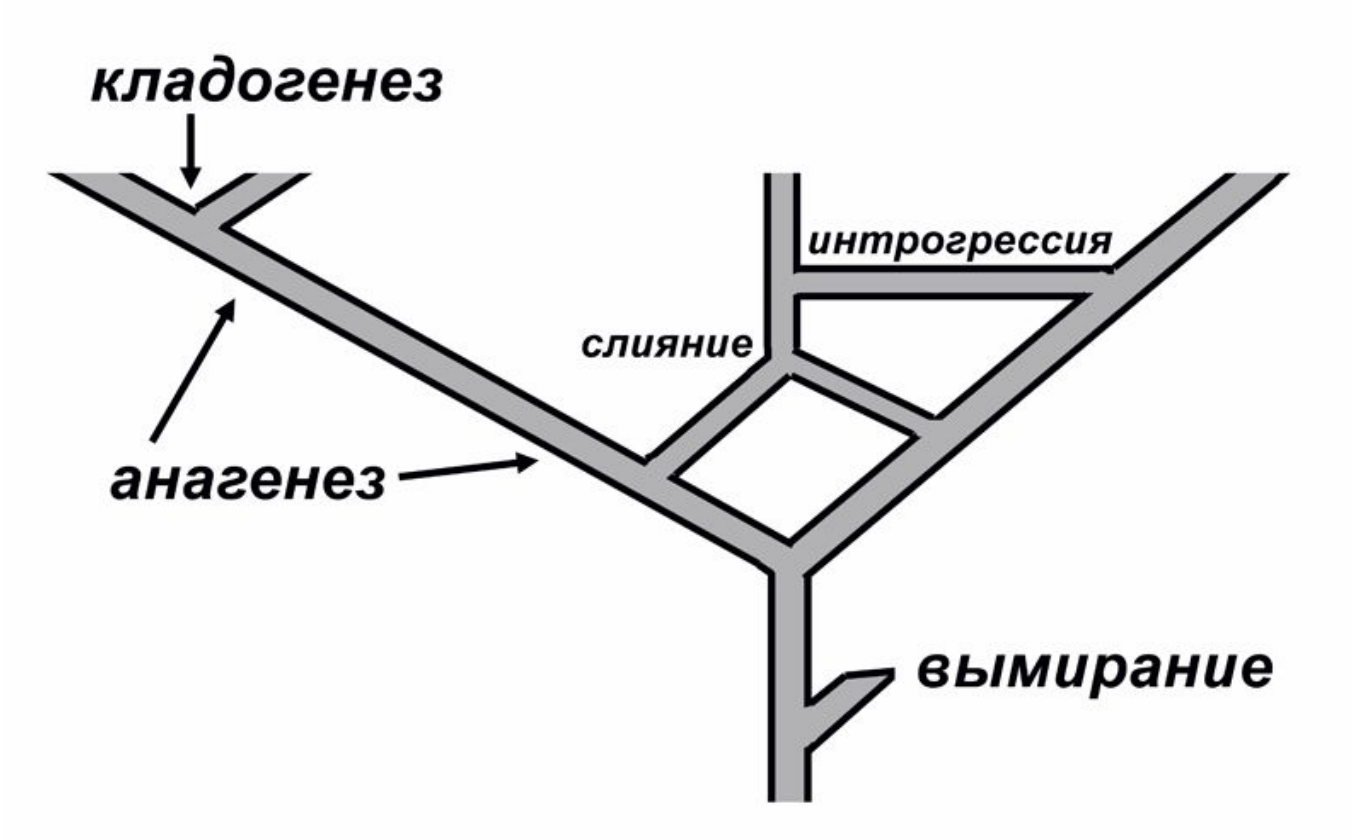
Типы филогенеза



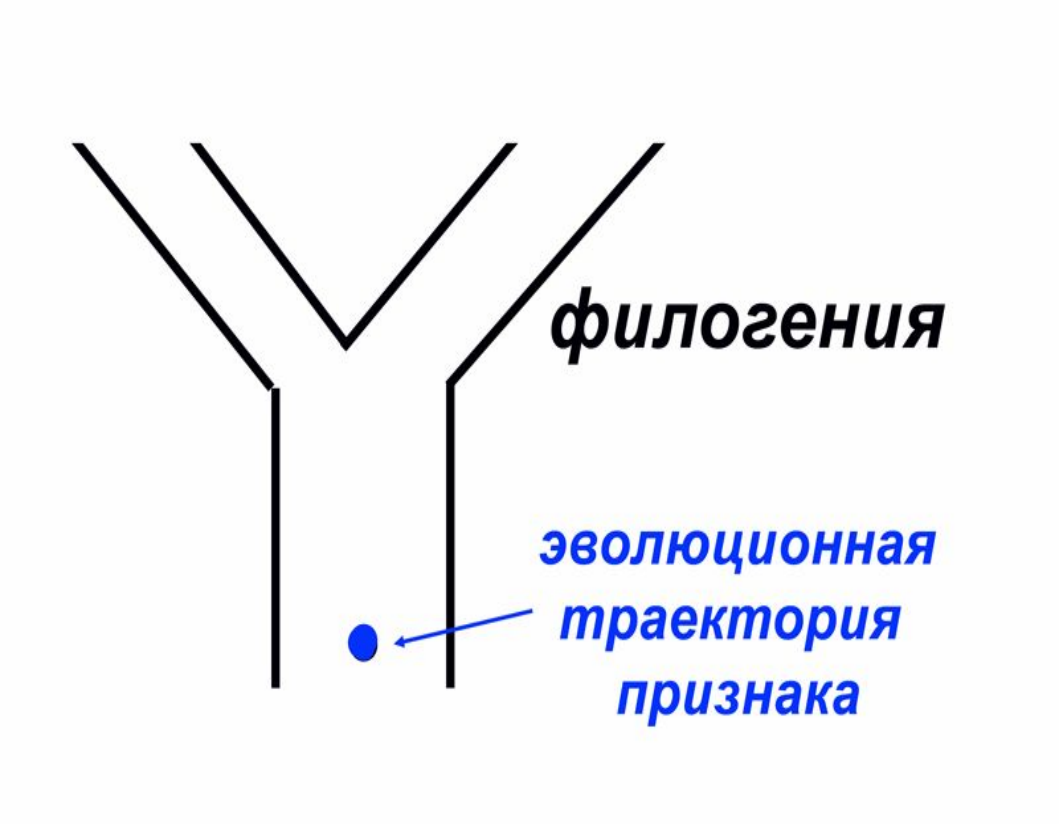
- В ходе филогенеза возникают и меняются признаки. Признаки – это следы филогенеза, которые мы можем использовать для реконструкции филогении

Аксиома филогенетики:

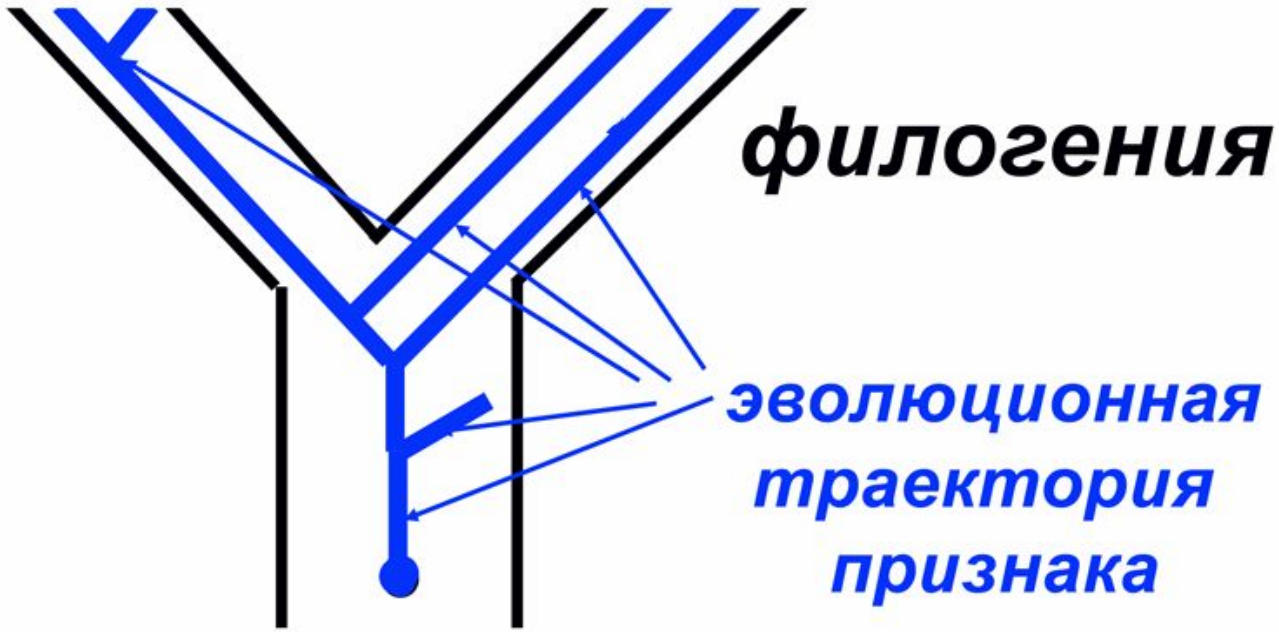
Филогения *таксонов и* ***траектории*** *эволюционных трансформаций* ***признаков****, характерных для этих таксонов,* ***конгруэнтны***



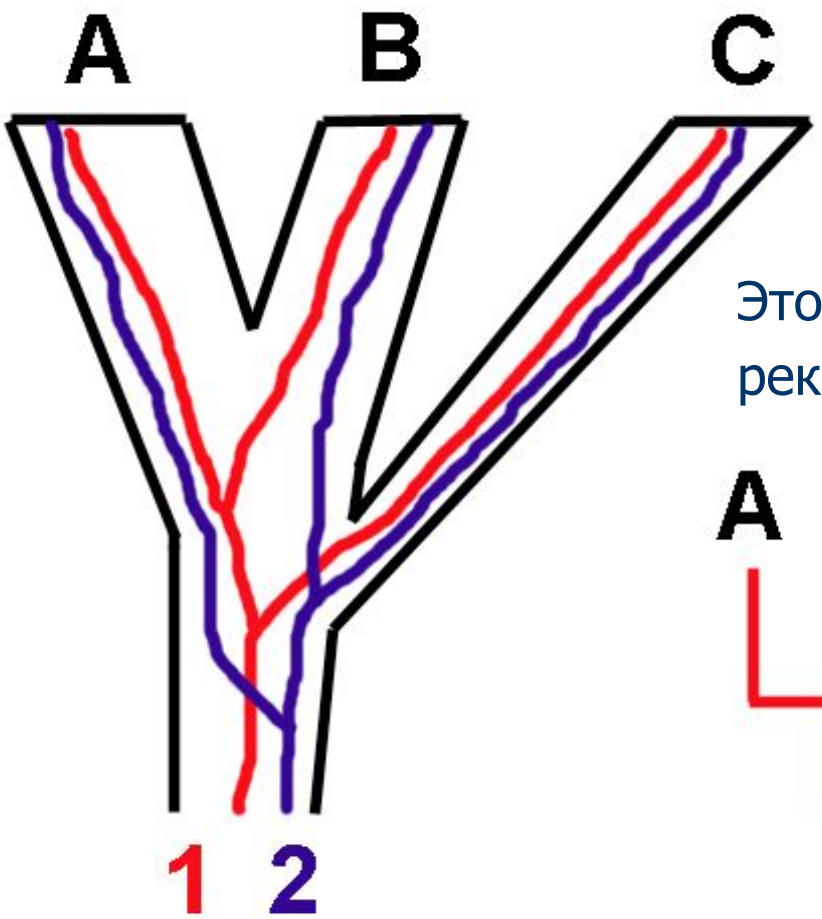
Носители нового признака могут погибнуть, и эволюционная траектория признака останется в виде точки



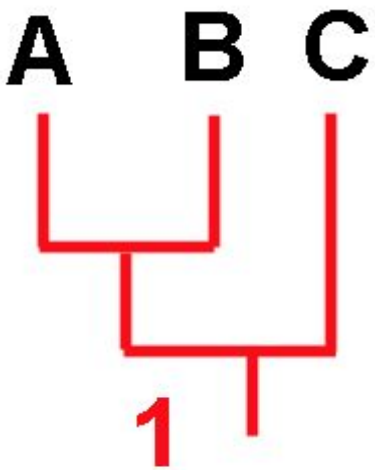
Филогения *таксонов* ***и траектория*** *трансформаций* ***признака*** *не тождественны, однако они конгруэтны, поскольку траектория остается внутри филогении*

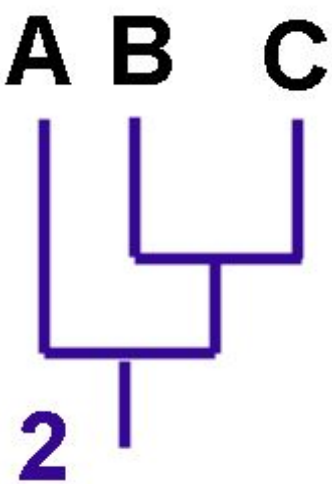


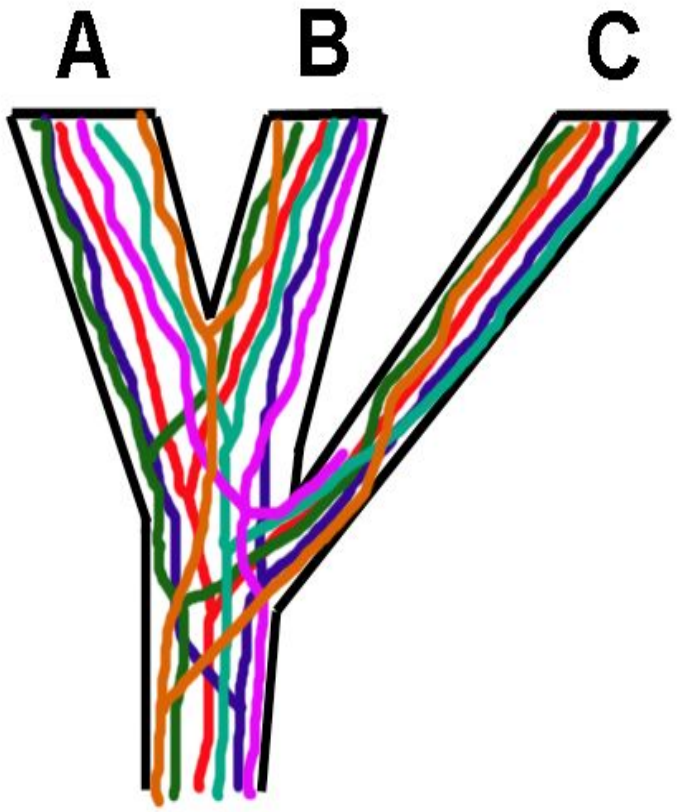
Филогенетические траектории признаков (**1** и **2**) не вполне совпадают с реальной филогенией **(A+B)+C**



Это приводит к разным реконструкциям одной филогении

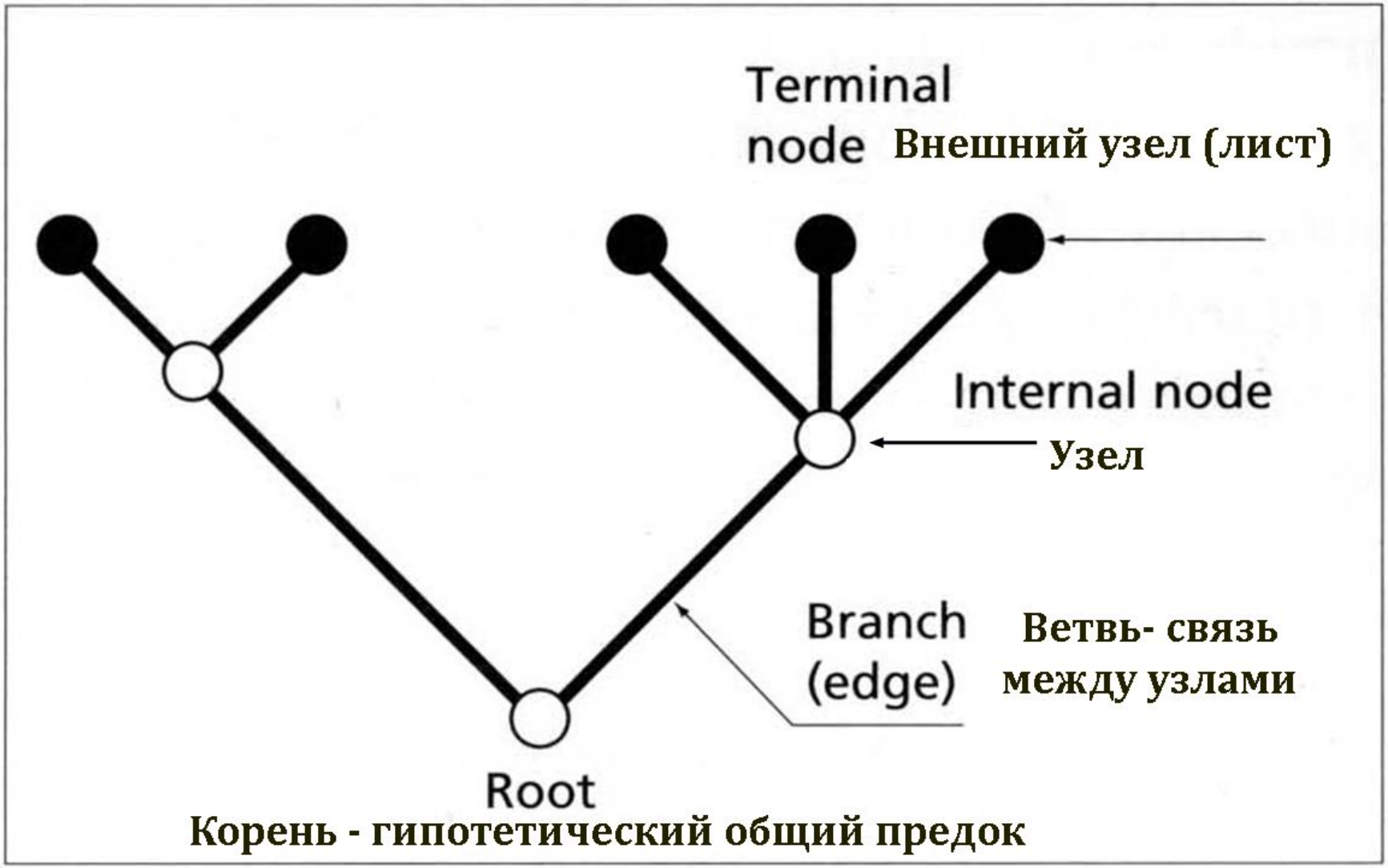




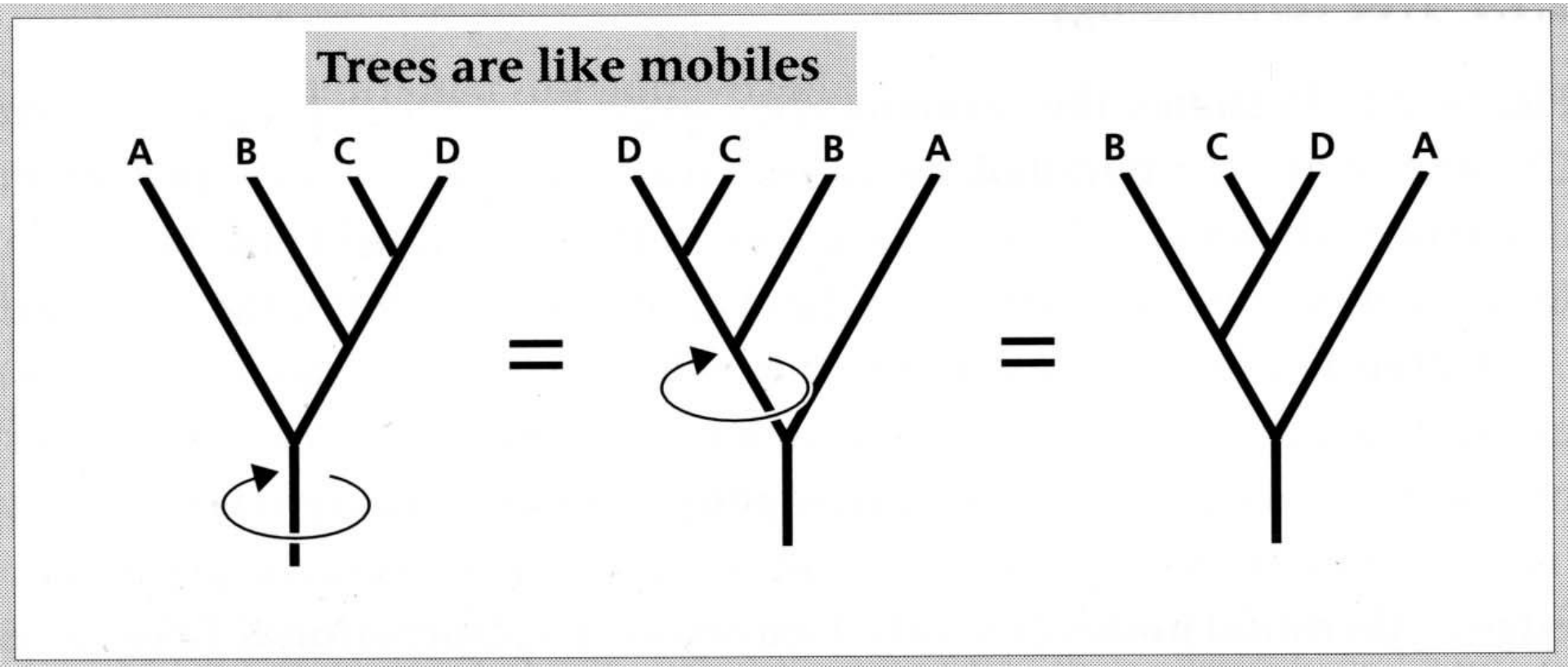


- Поскольку траектории отдельных признаков не выходят за пределы филогении, то они конгруэнтны, и восстановлении филогении по их совокупности возможно

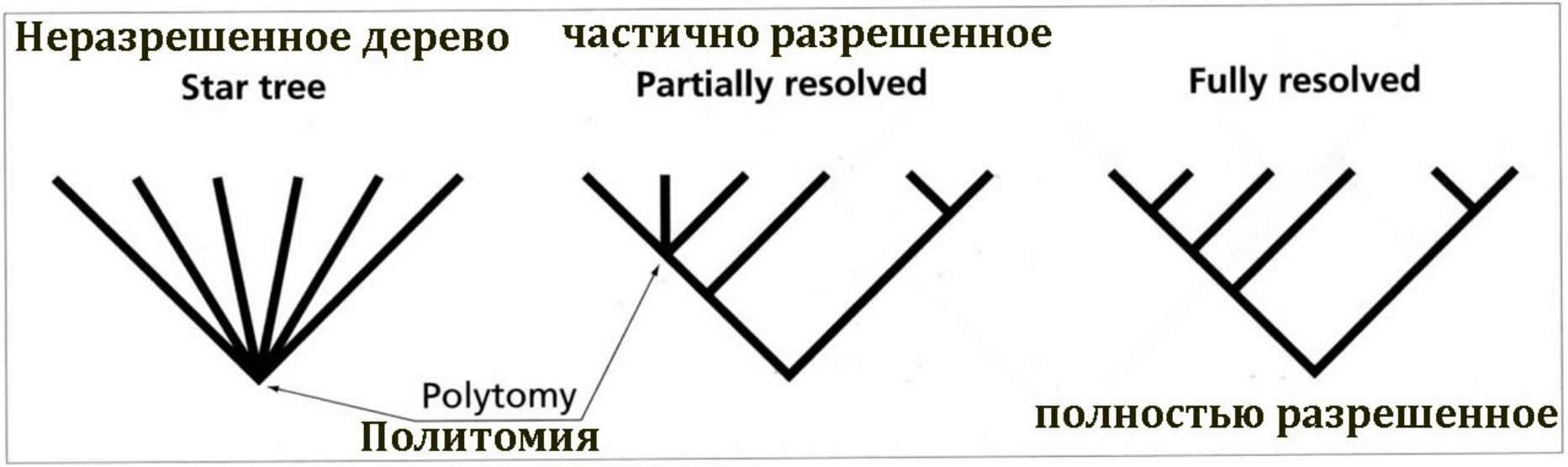
Дерево: основная терминология



Деревья (и их части) можно вращать в районе любого узла

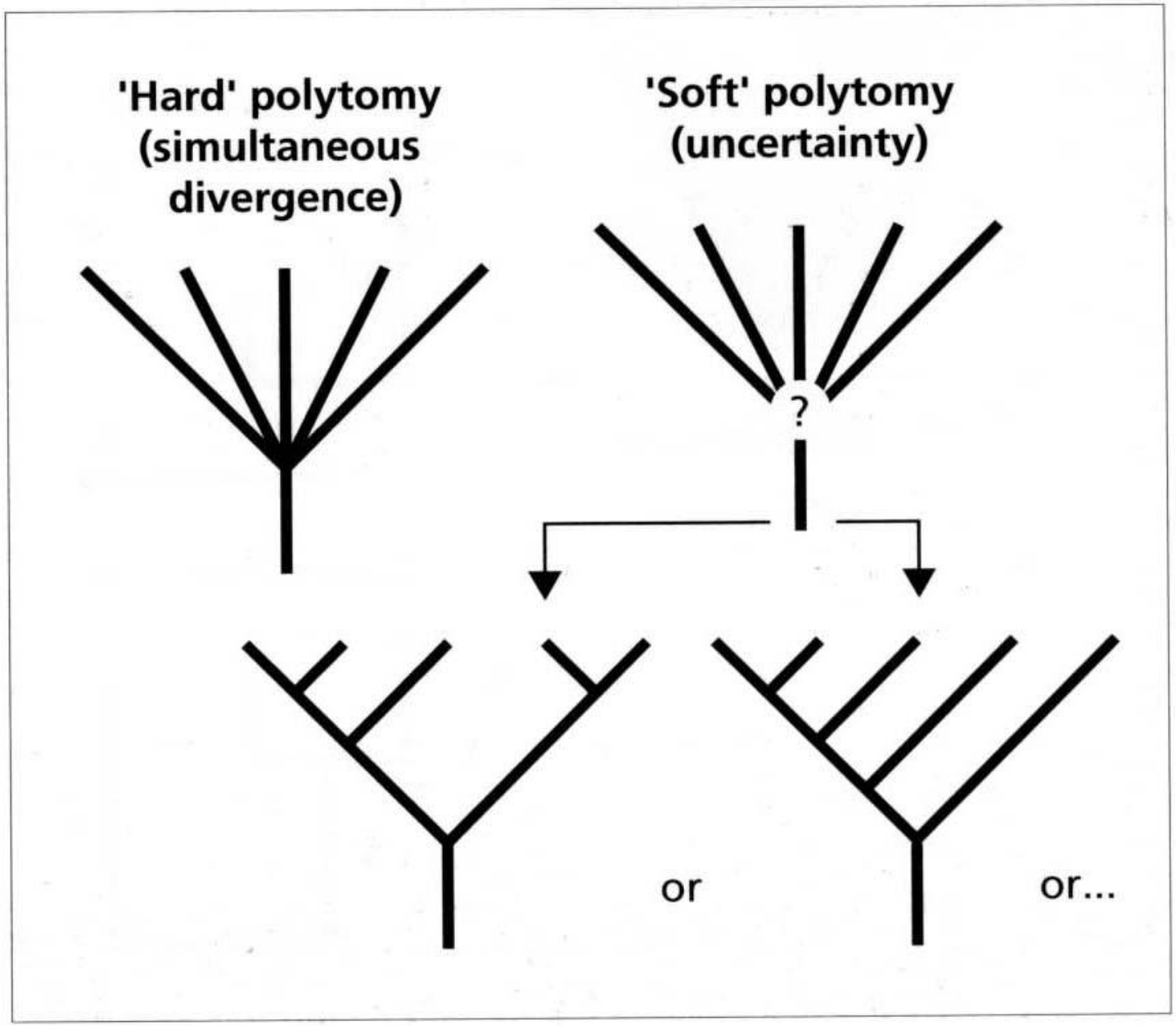


“Разрешенная” и “неразрешенная” филогения



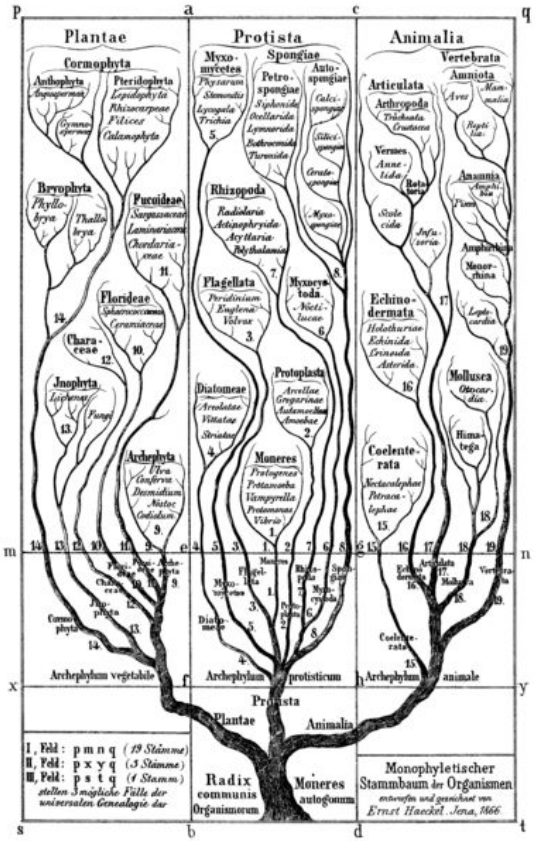
Three trees showing various degrees of resolution, ranging from a complete lack of resolution (star tree) to a fully resolved tree. Any internal node with more than two immediate descendants is a polytomy.

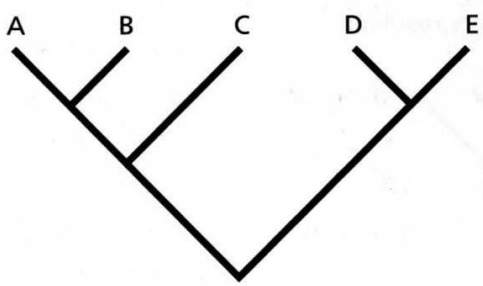
Что такое “частично разрешенная филогения?”

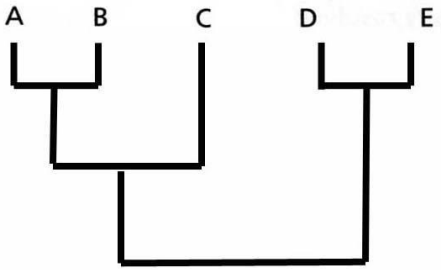


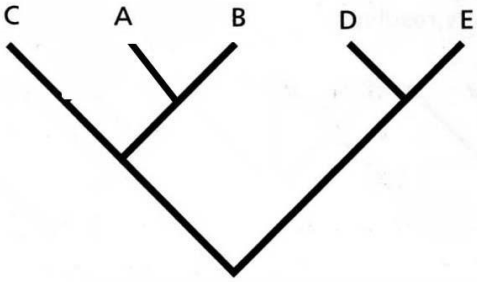
Политомия бывает “жесткая” и “мягкая”

Представление филогении в виде дерева







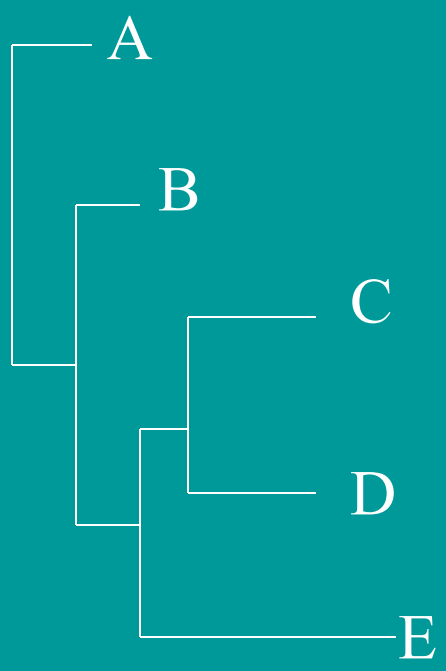


(((A.B).C).(D,E))

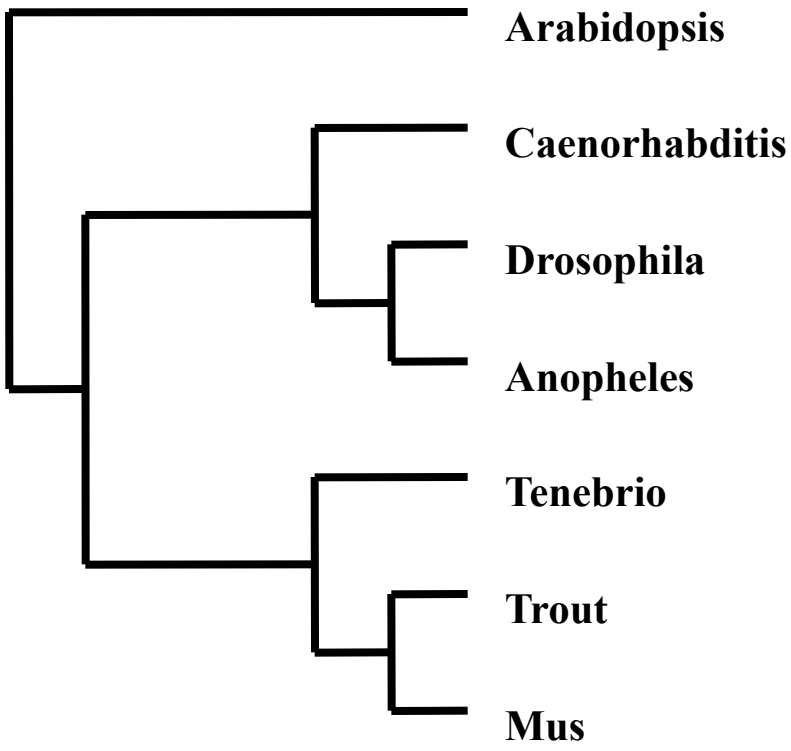
Графические и буквенное представления одного и того же дерева

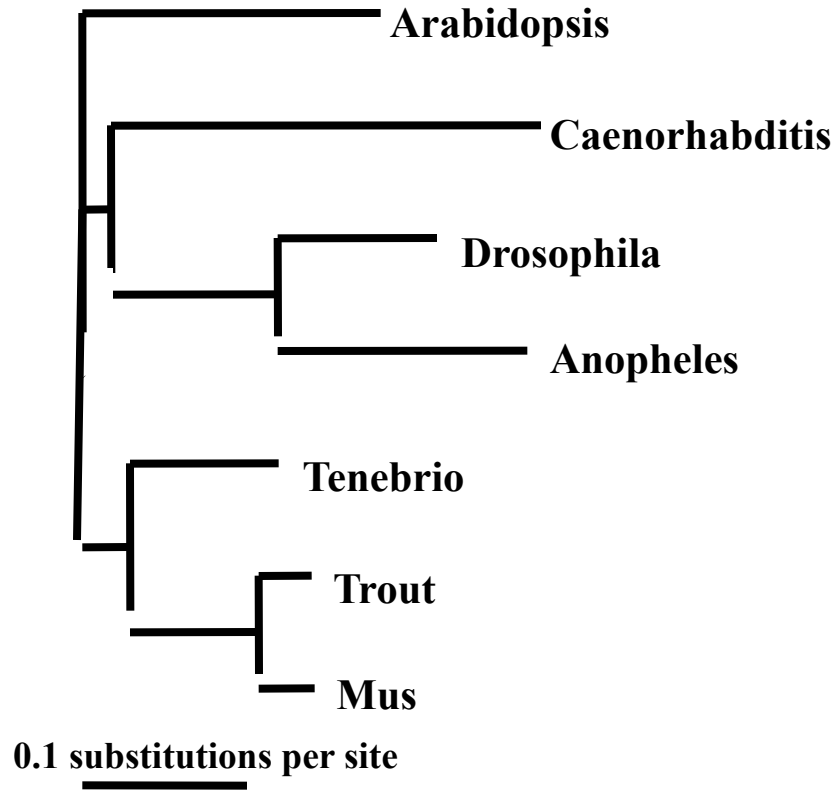
- Топология дерева — только листья, узлы, (корень) и связывающие их ветви

  (топология не зависит от способа изображения дерева)



Два изображения одной и той же топологии





Кладограмма:

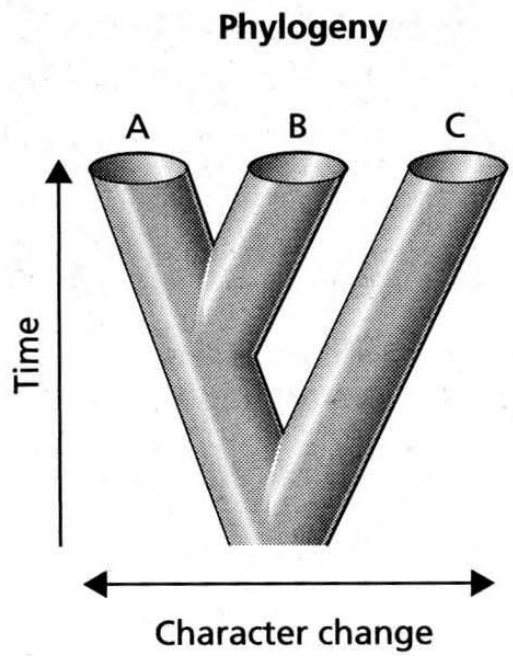
представлена только топология, длина ветвей игнорируется

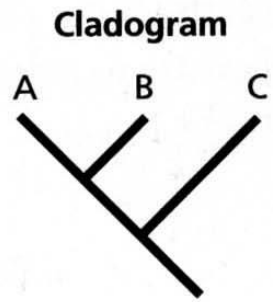
Филограмма:

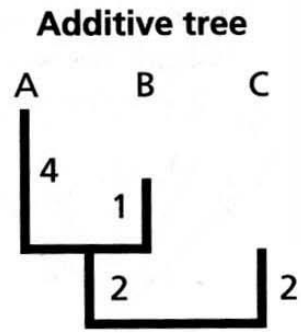
длина ветвей пропорциональна эволюционному расстоянию между узлами

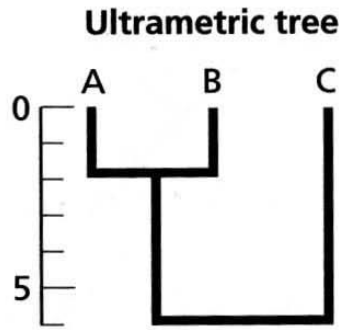
Как измерить расстояние?

Филогения и ее представления в виде кладограммы, аддитивного дерева и ультраметрического дерева

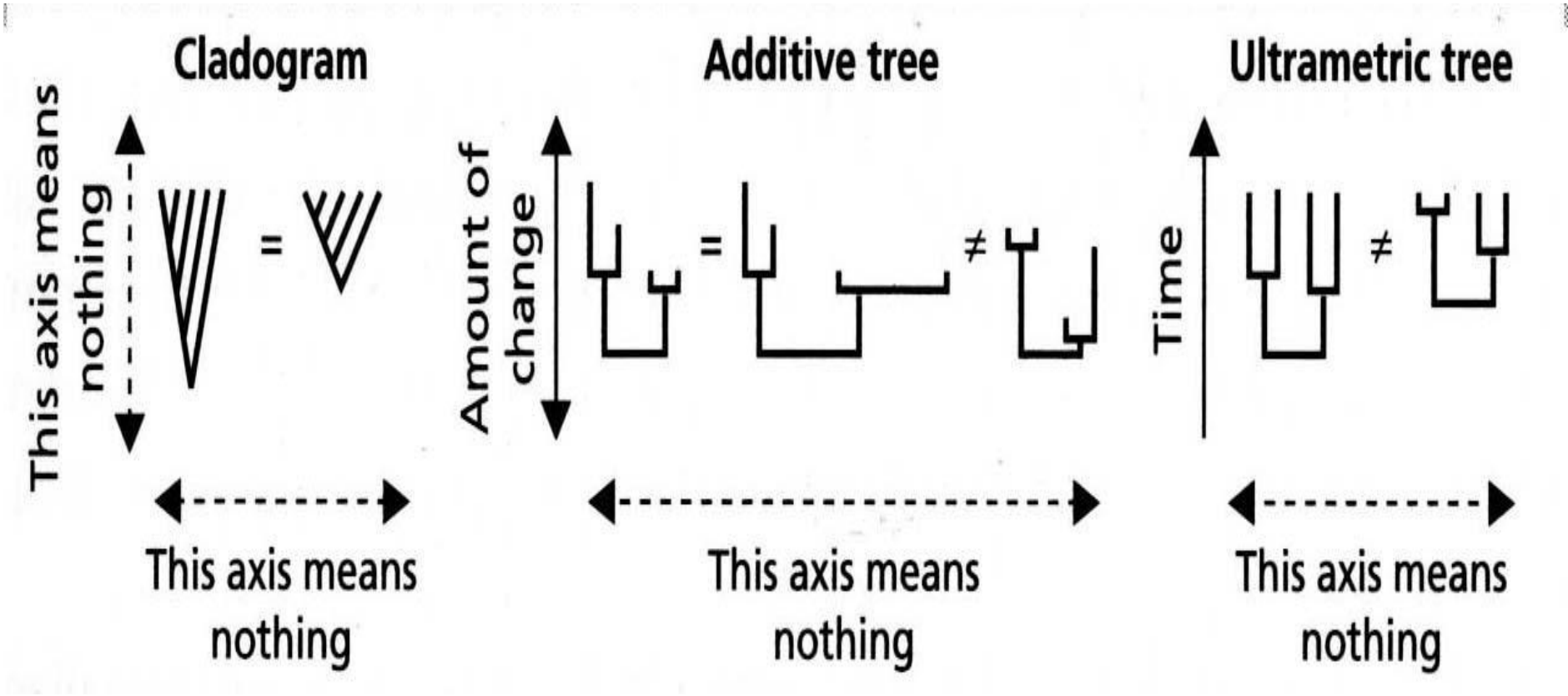


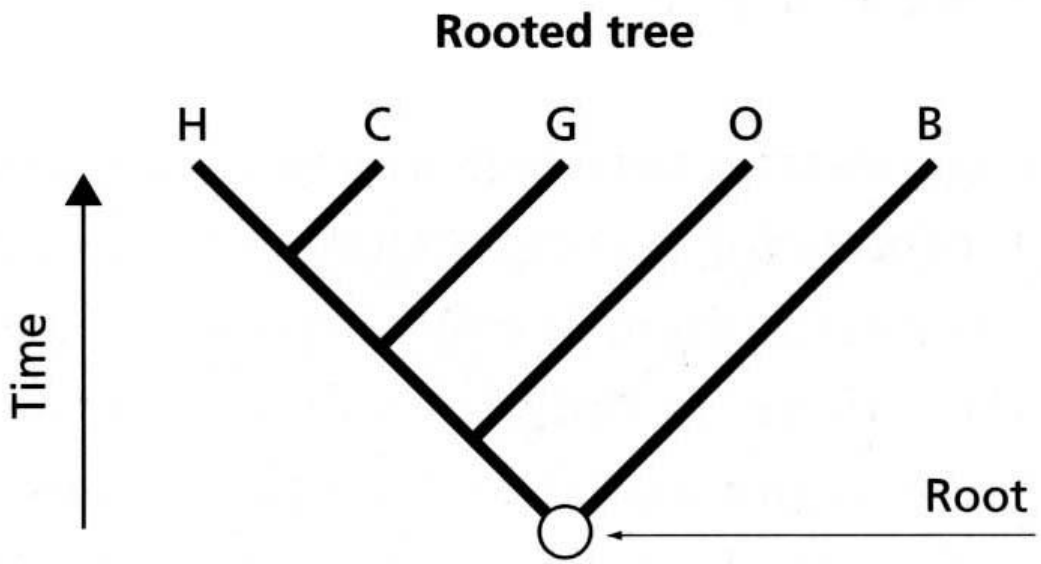




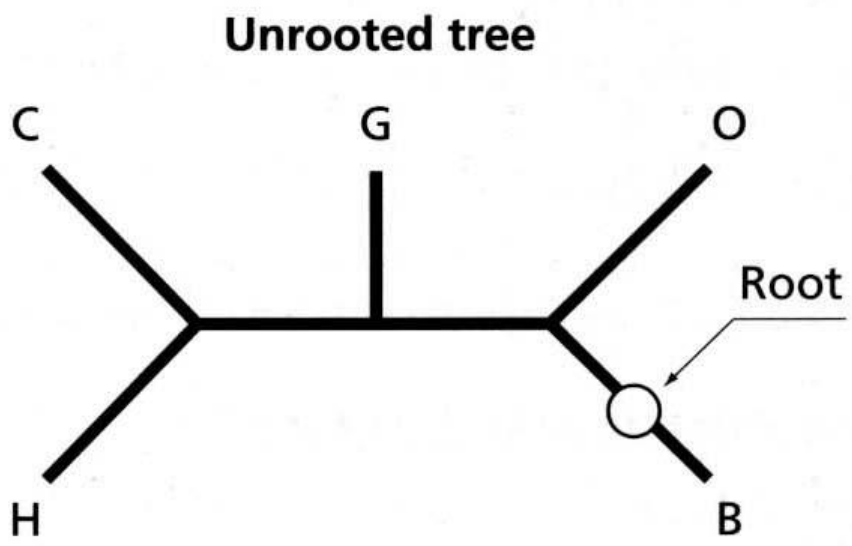


Что означают горизонтальная и вертикальная оси?

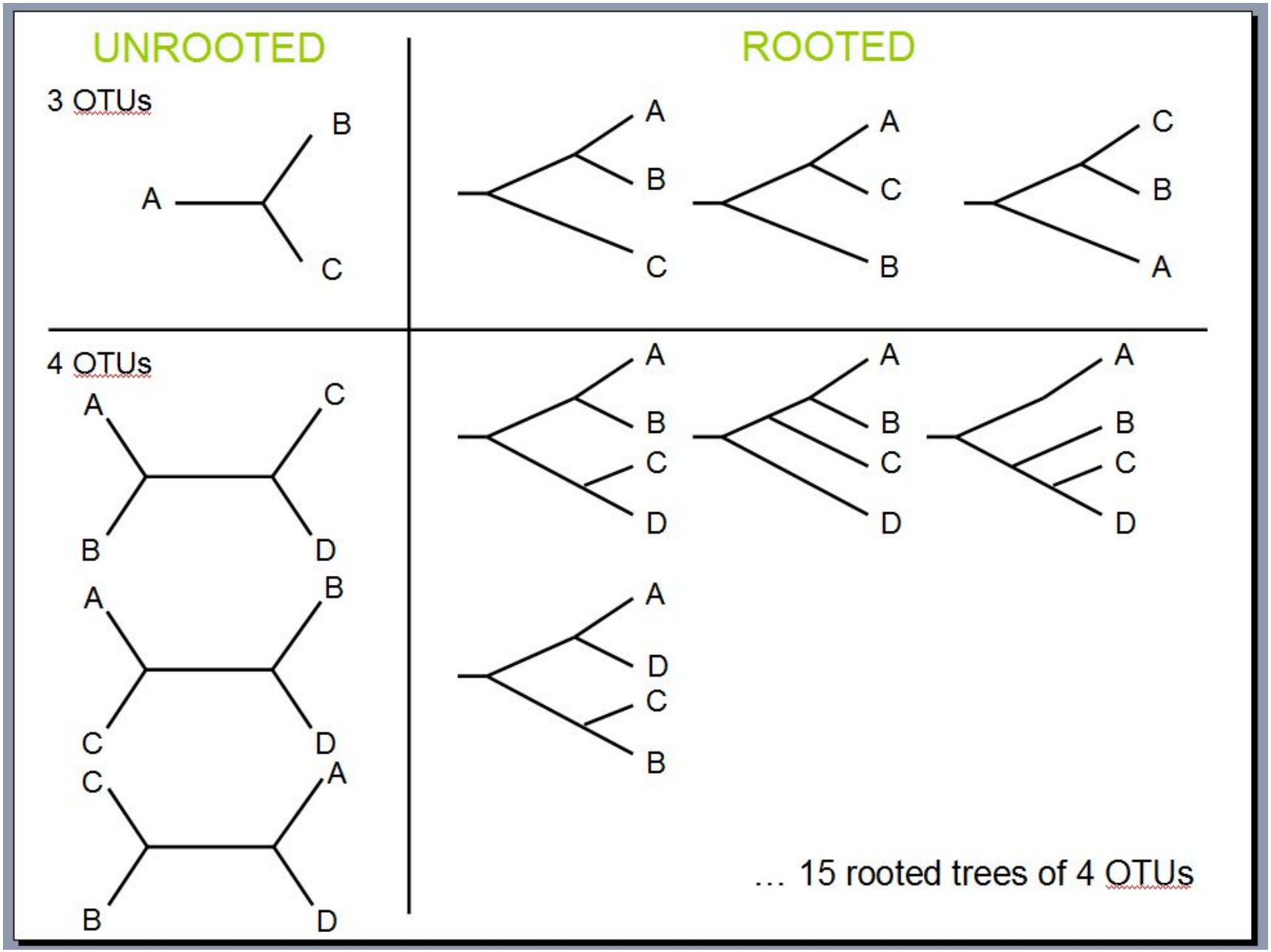




Укорененное дерево



Неукорененное дерево


UNROOTED
ROOTED
3 OTUs
A
B
C
A
B
C
A
C
B
B
A
C
B
A
4 OTUs
A
B
C
D
A
B
C
D
A
B
C
D
A
B
C
D
A
D
C
B
A
B
C
D
… 15 rooted trees of 4 OTUs

Число таксонов (сиквенсов) и число топологий, которые можно получить на их основе

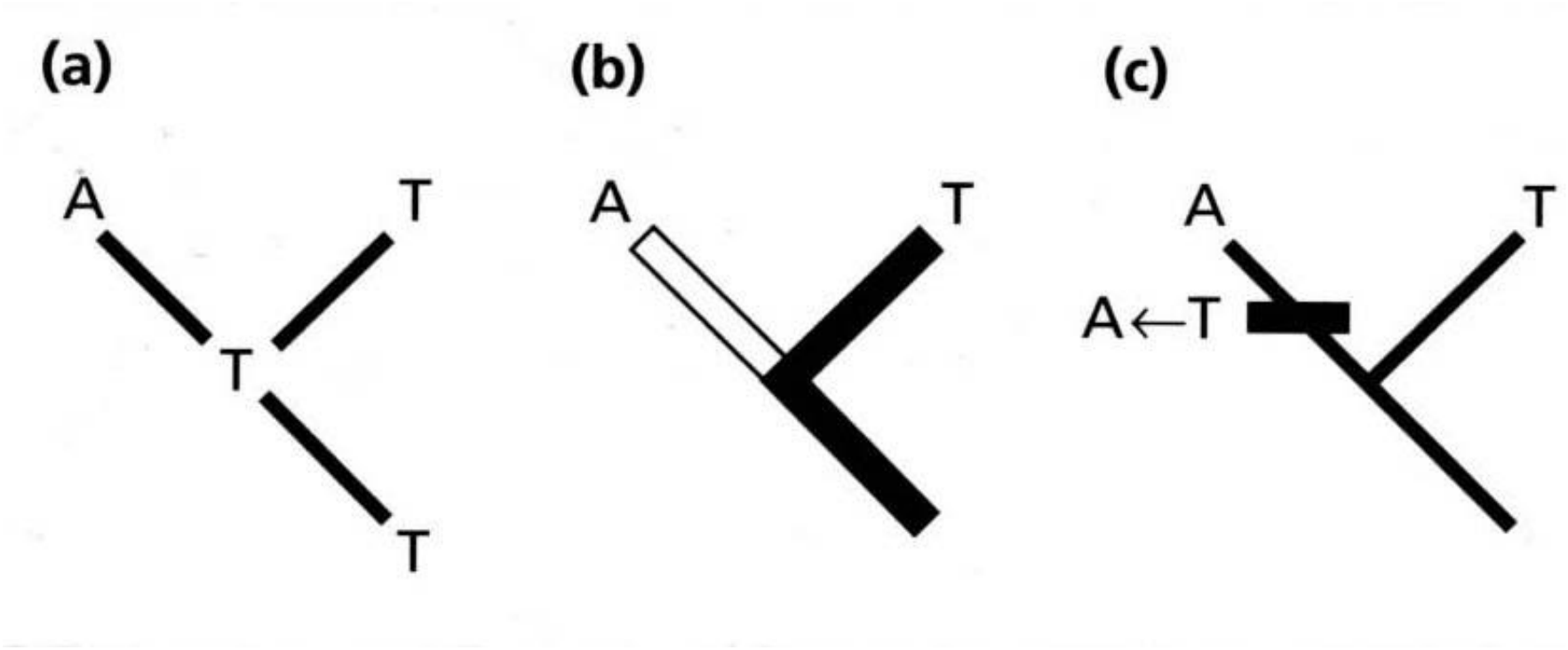
Number of sequences	Number of unrooted trees	Number of rooted trees
2	1	1
3	1	3
4	3	15
5	15	105
6	105	945
7	945	10 395
8	10 395	135 135
9	135 135	2 027 025
10	2 027 025	34 459 425

Numbers of unrooted and rooted trees for 2–10 sequences.

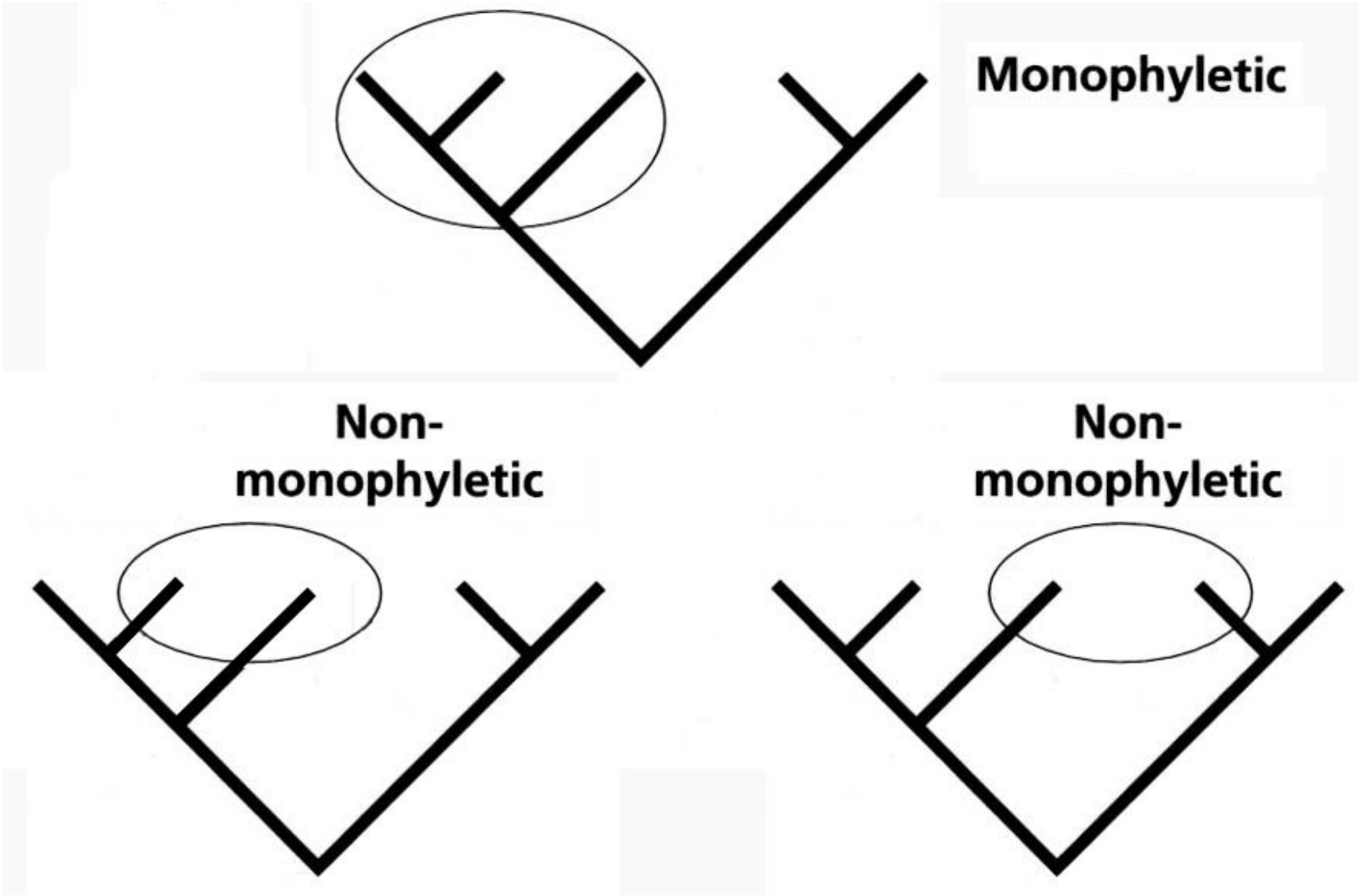
Сколько из них правильных ?

Если число таксонов равно n, существует
(2n-3)!!
разных бинарных укоренных деревьев.
(2n-3)!! – это нечто вроде факториала, но учитываются только нечетные числа

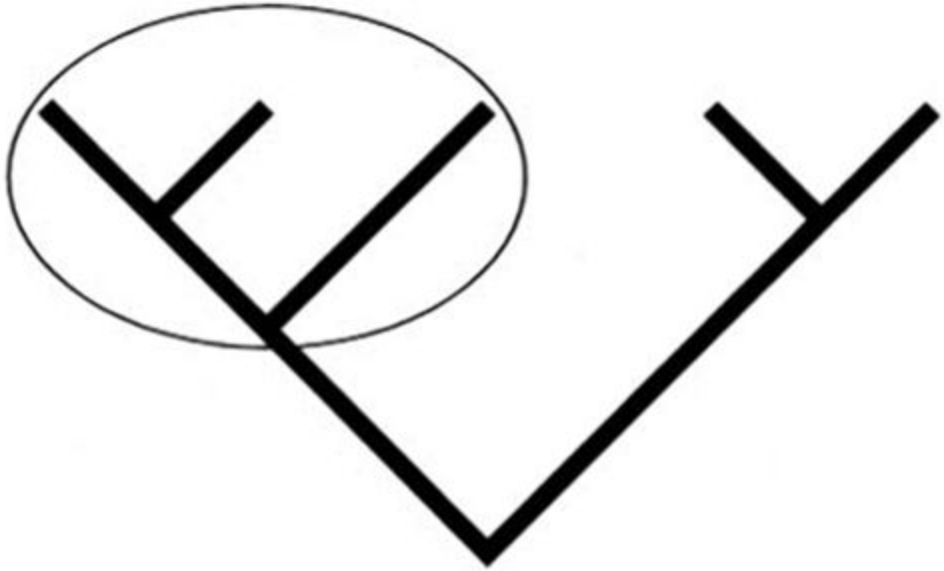
Три способа изображения эволюционных изменений на дереве



Части дерева: группы и клады

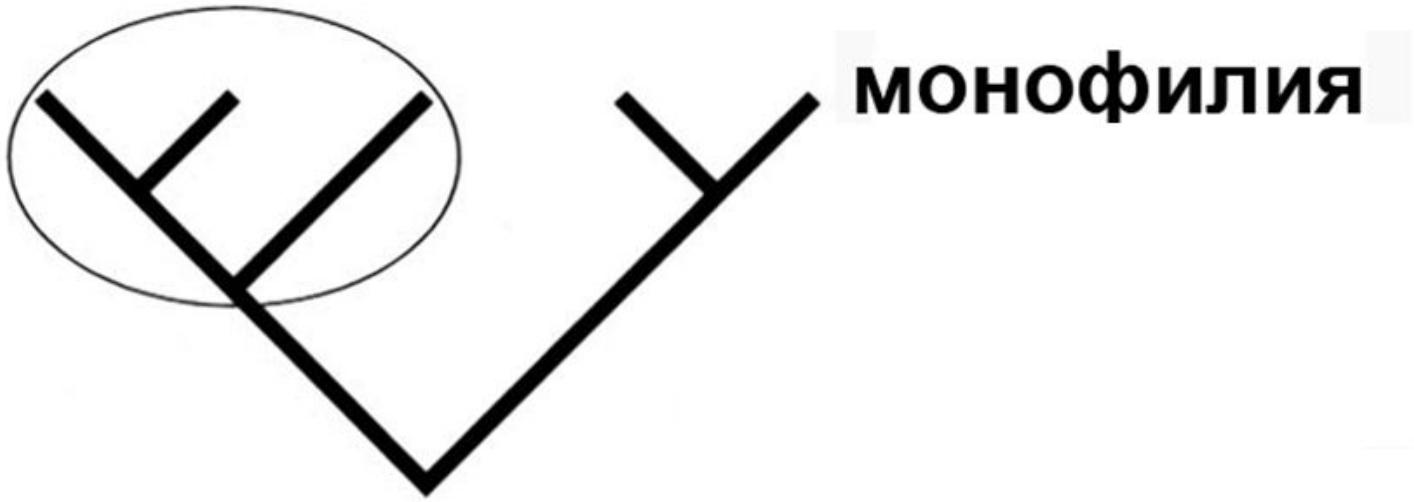


Монофилетическая группа называется кладой



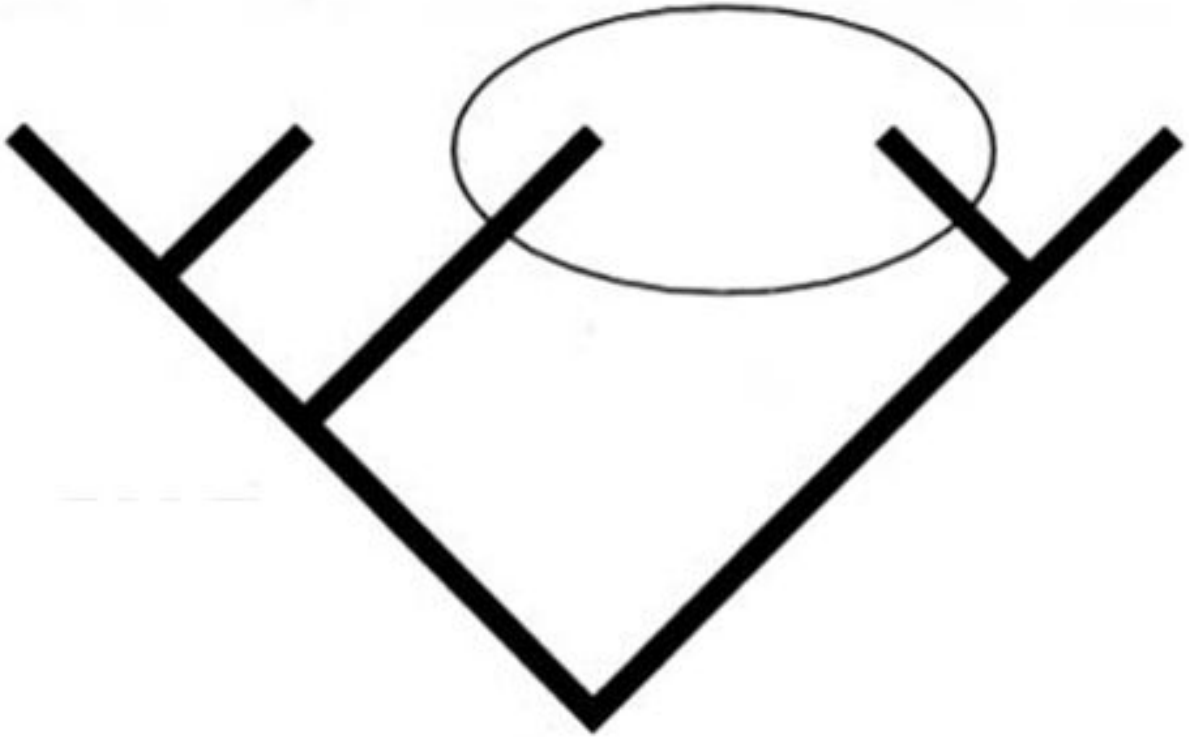
монофилия

Предковый таксон и **ВСЕ** его потомки



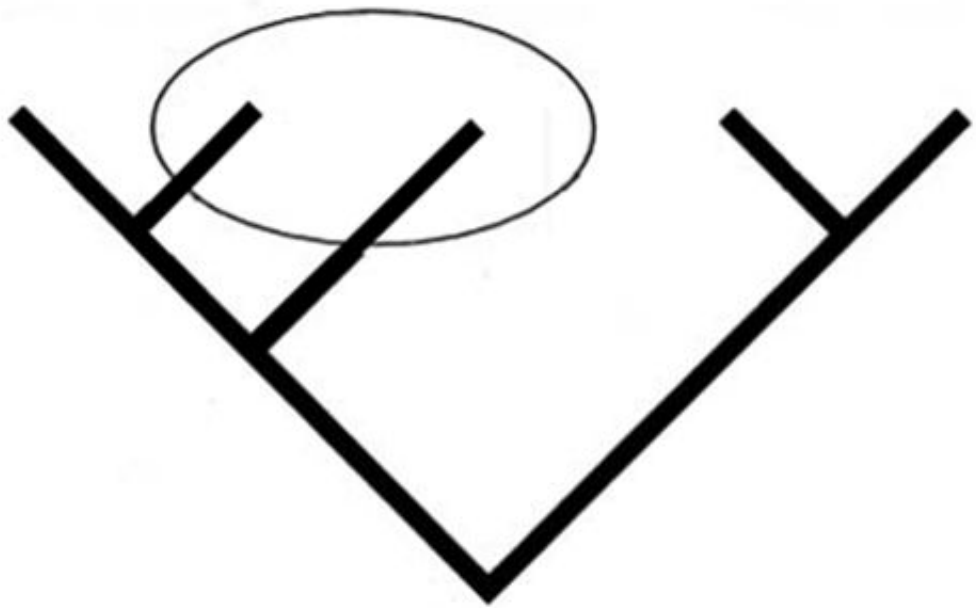
Монофилетический таксон = клада = ветвь = ветка = линия

полифилия



Произвольное объединение двух и более терминальных таксонов

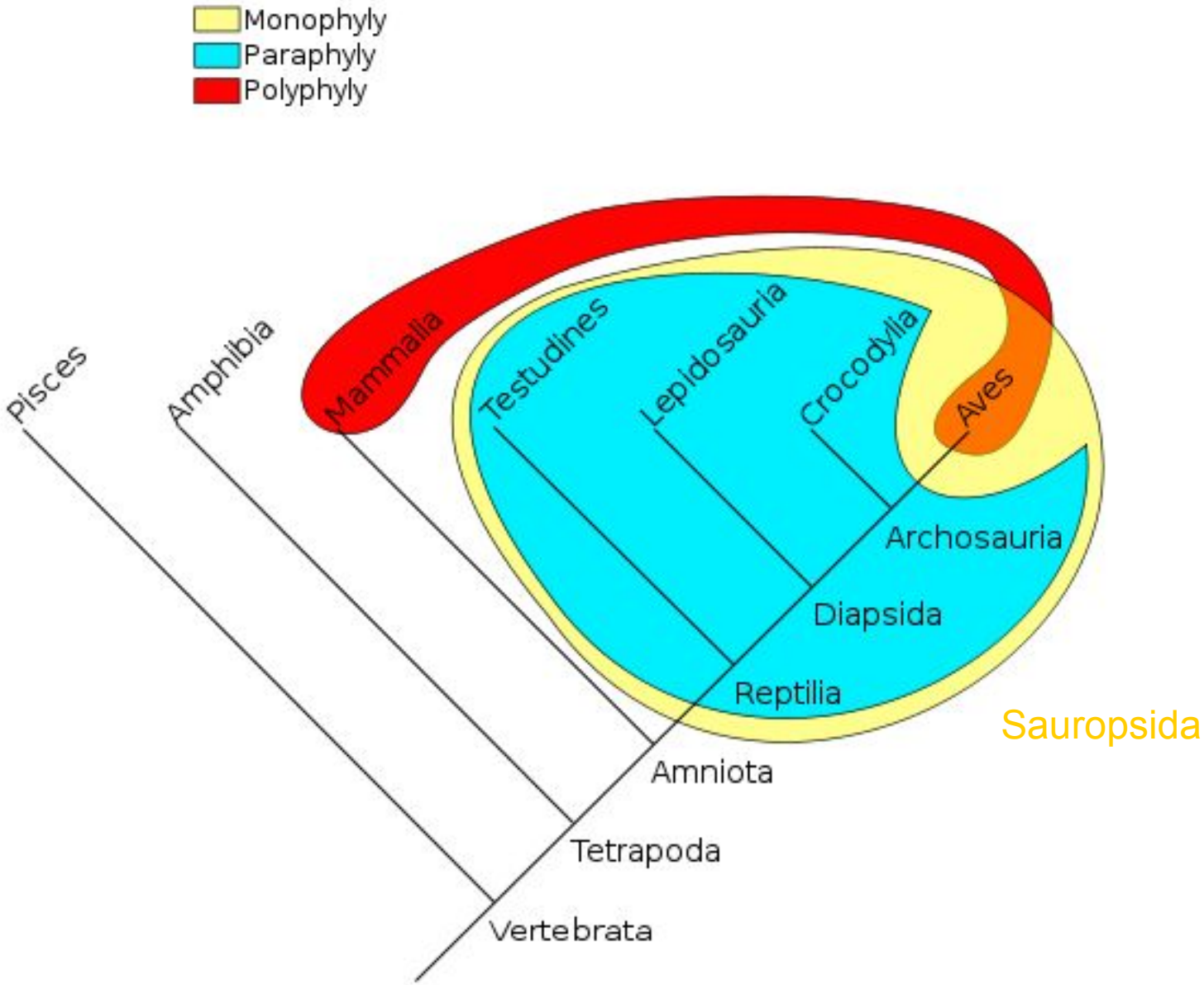
парафилия



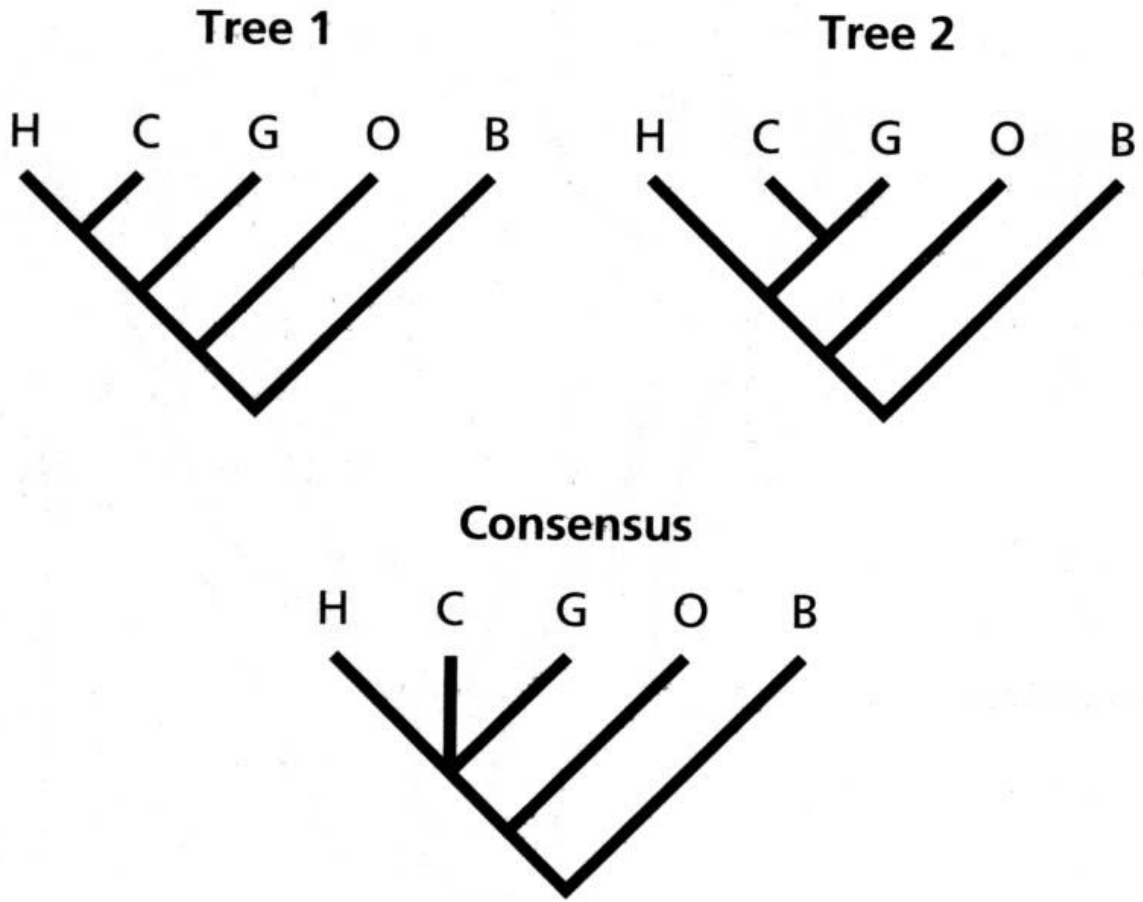
Парафилия: из явно монофилетической группы вычитаем один или более таксонов

Одни считают, что это частный случай полифилии
Другие – неполноценная монофилия

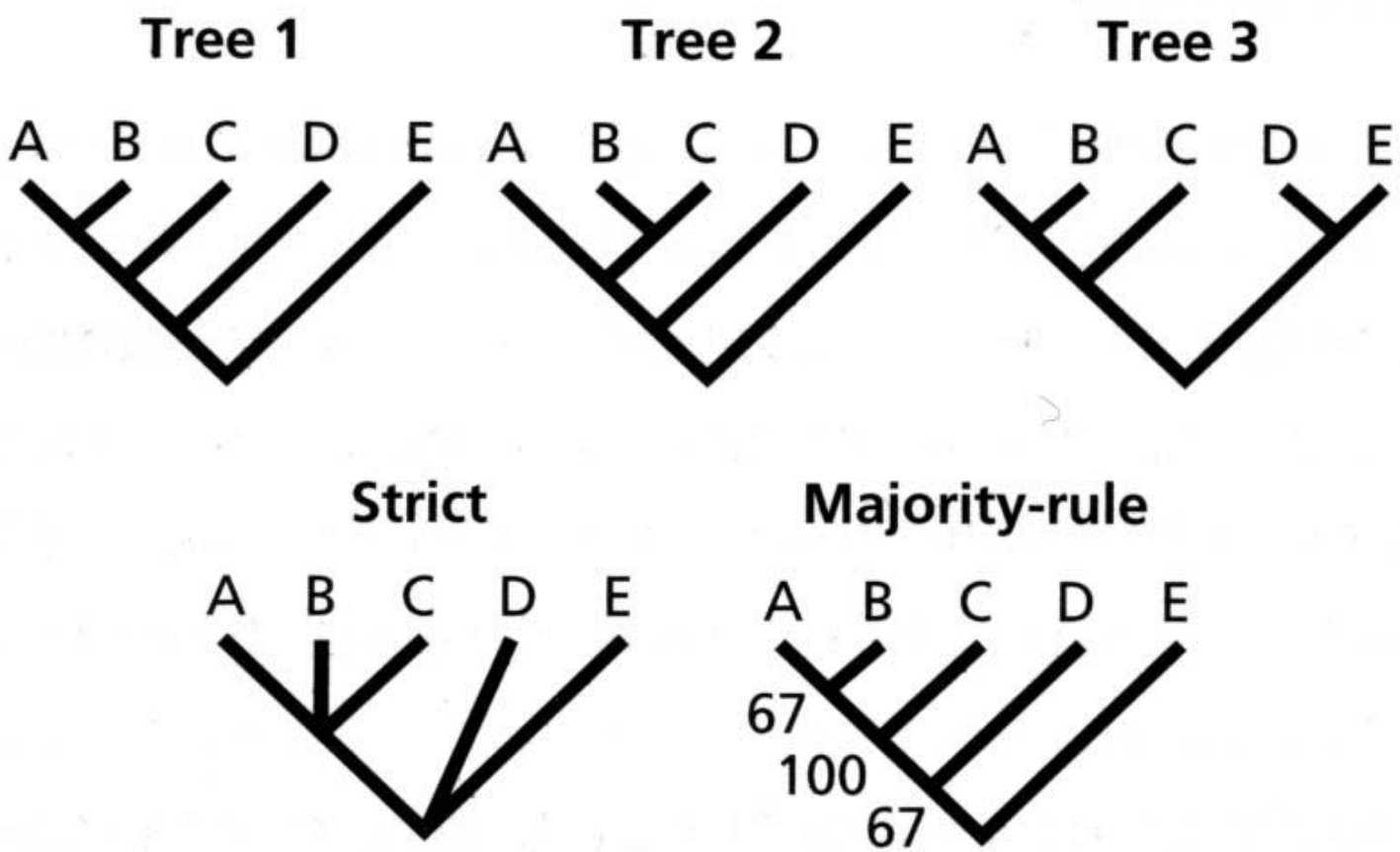
Ветка – это естественная часть дерева
Обламанная ветка не вполне естественна


Monophyly
Paraphyly
Polyphyly
Pisces
Amphibia
Mammalia
Testudines
Lepidosauria
Crocodylia
Aves
Archosauria
Diapsida
Reptilia
Amniota
Tetrapoda
Vertebrata
Sauropsida

Консенсусные деревья



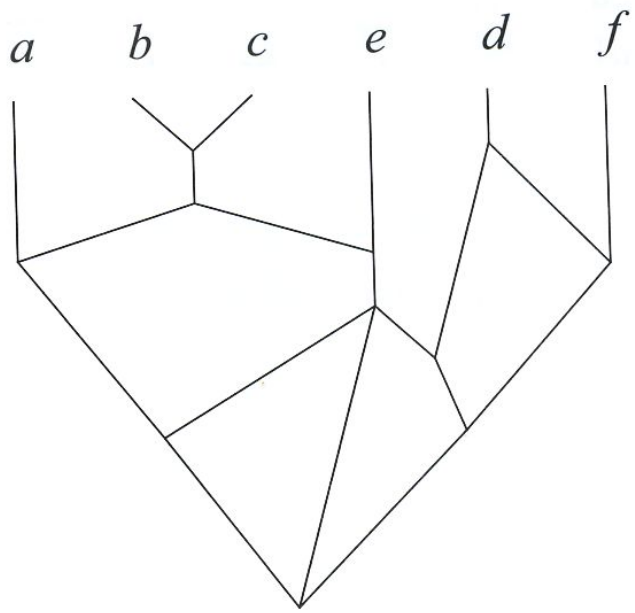
H – human; C – chimpanzee; G – gorilla; O – orangutan; B - gibbon



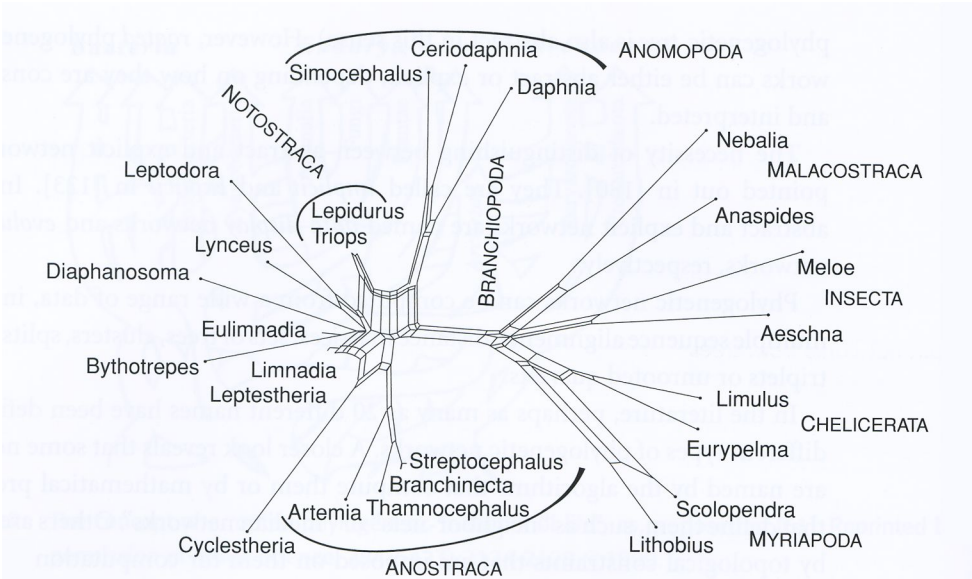
Строгий консенсус и консенсус по правилу большинства

H – human; C – chimpanzee; G – gorilla; O – orangutan; B - gibbon

- Укорененная сеть



Неукорененная сеть



A split network on 25 species of Branchiopoda and outgroups, computed from 18S rDNA