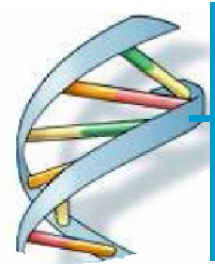


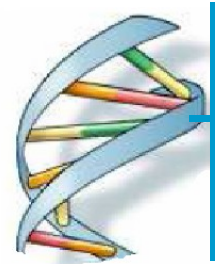


Филогенетична реконструкция



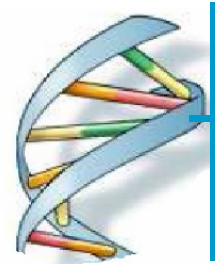
Филогенетичен анализ

- В биологията, филогенетиката (phylogenetics) е изучаването на еволюционната свързаност между различни групи от организми (например, видове, популации).
- Компютърната филогенетика е прилагането на изчислителни алгоритми, методи и програми за филогенетични анализи.
- Целта е да се представи филогенетично дърво, което представлява хипотеза за еволюционните прадеди на набор от гени, видове или други таксони.



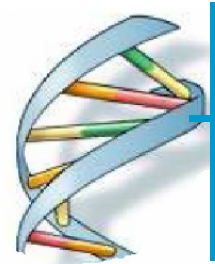
Филогенетичен анализ

- Например, тези техники са използвани за изследване на дървото на семейството на видовете „hominid“ (към който спада и човека) и отношенията между специфични гени, споделяни от много видове организми.
- Съвременната молекулярна филогенетика използва **нуклеотидните секвенции, кодирани гени или аминокиселинови секвенции** на кодираните протеини като основа за класификация.



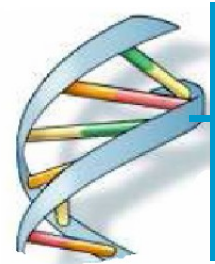
Филогенетичен анализ

- Много форми на молекулярната филогенетика са тясно свързани с тях и са широко използвани за последователно подреждане и конструиране на филогенетични дървета, които се използват за класифициране на еволюционните взаимоотношения между съответните гени, представени в геноми на различни видове.
- Филогенетичните дървета, построени на основата на изчислителни методи не могат перфектно да възпроизведат еволюционното дърво, което представлява историческите връзки между видовете.



Филогенетично дърво

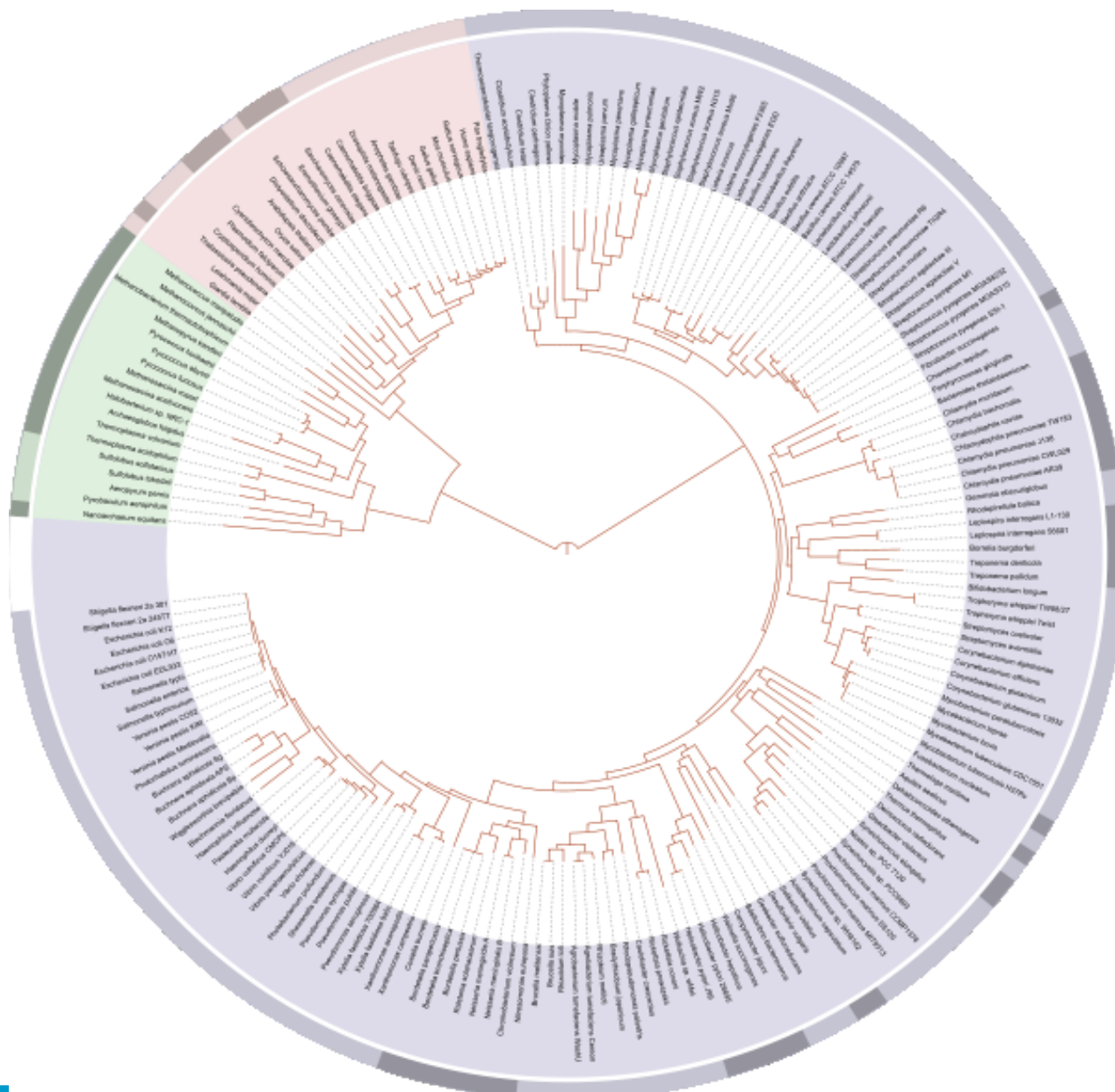
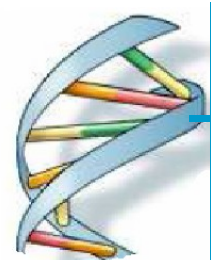
- Филогенетично дърво или еволюционно дърво е дърво, показващо еволюционните взаимоотношения между различните биологични видове или други същества, за които се счита, че имат общ прародител.
- Във филогенетичното дърво всеки възел представлява най-новият общ прародител на потомството, както и върховете на дължините в някои дървета съответстват на временни оценки.
- Всеки възел се нарича ***таксономична единица***. Вътрешните възли обикновено са наречени хипотетично таксономични единици (HTUs), тъй като те не могат да бъдат пряко наблюдавани.

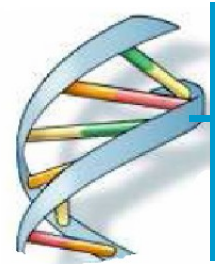


История

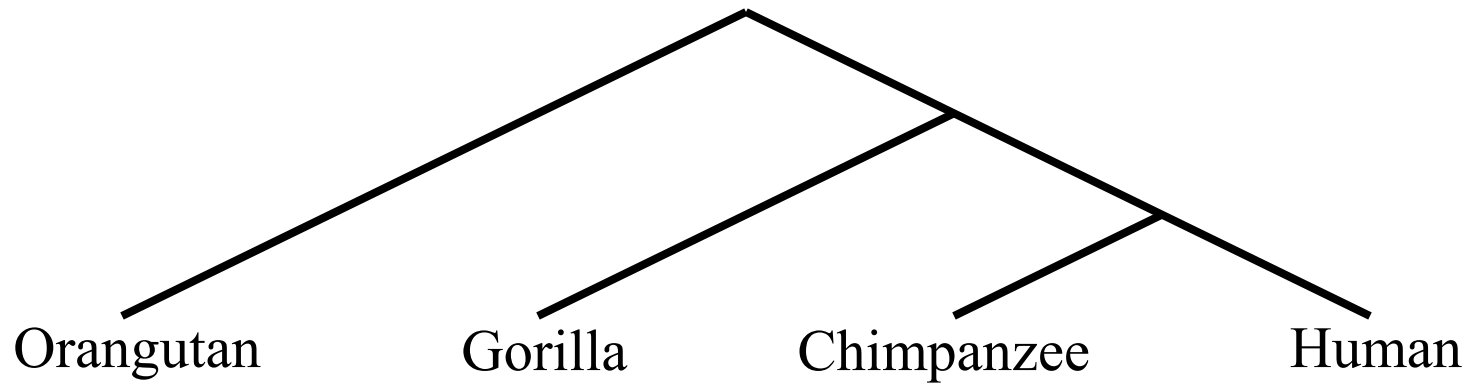
- Въпреки, че идеята на "дървото на живота" е възникнала от древните понятия за стълбата като прогресия от по-ниските до по-високи форми на живот (например великата верига на битието), Charles Darwin (1859) първи илюстрира и популяризира понятието „еволюционно дърво“ в неговата книга "Произход на видовете".
- Повече от един век по-късно, биолозите все още използват дървовидни диаграми, показващи еволюцията, тъй като цветната аналогия внушава ефективно идеята, че съществата се създават чрез адаптивни и случайни принципи на разделянето на предните. С течение на времето, класификацията на видовете е станала по-динамична.

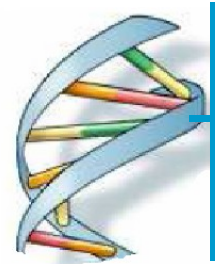
Филогенетично дърво на живота



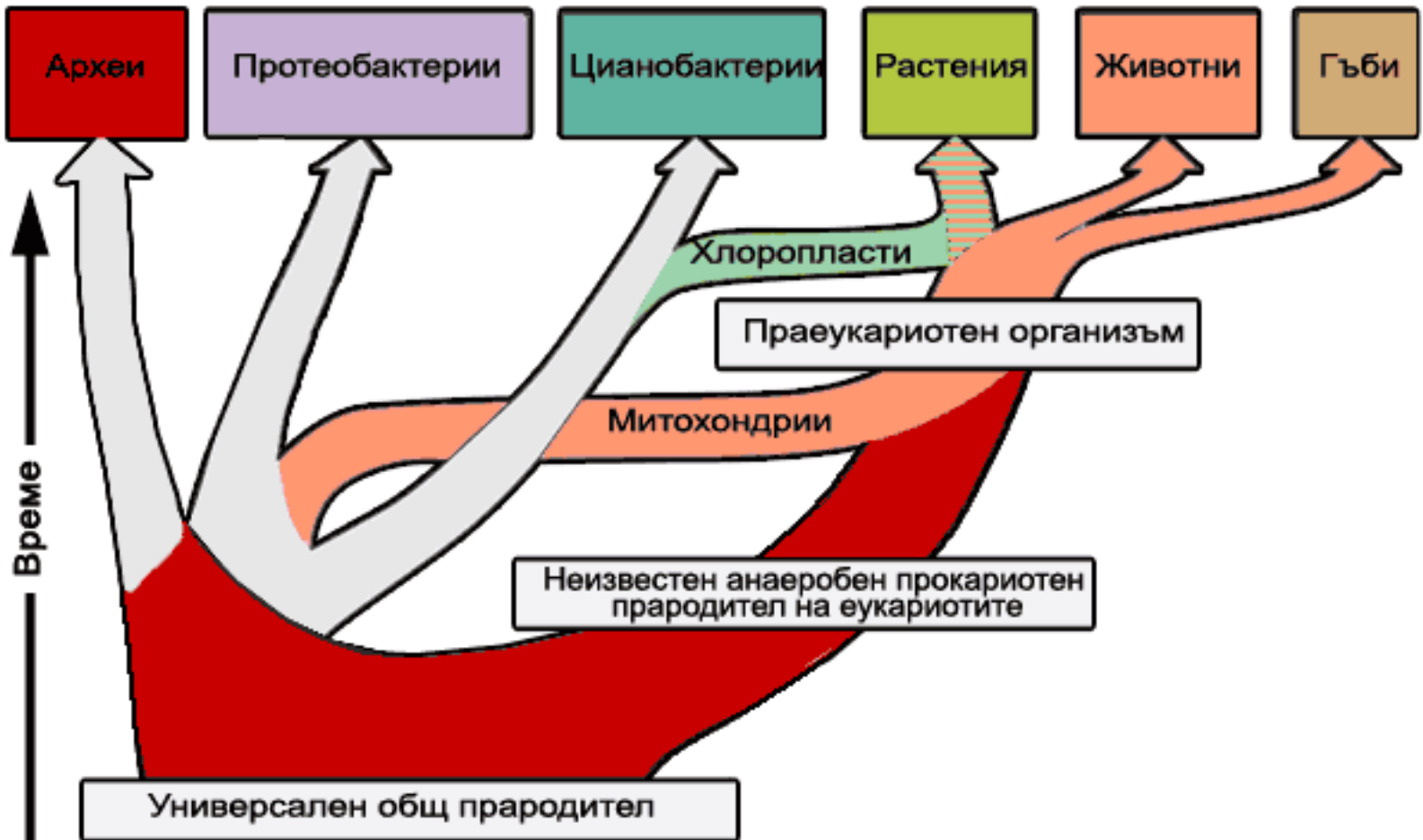


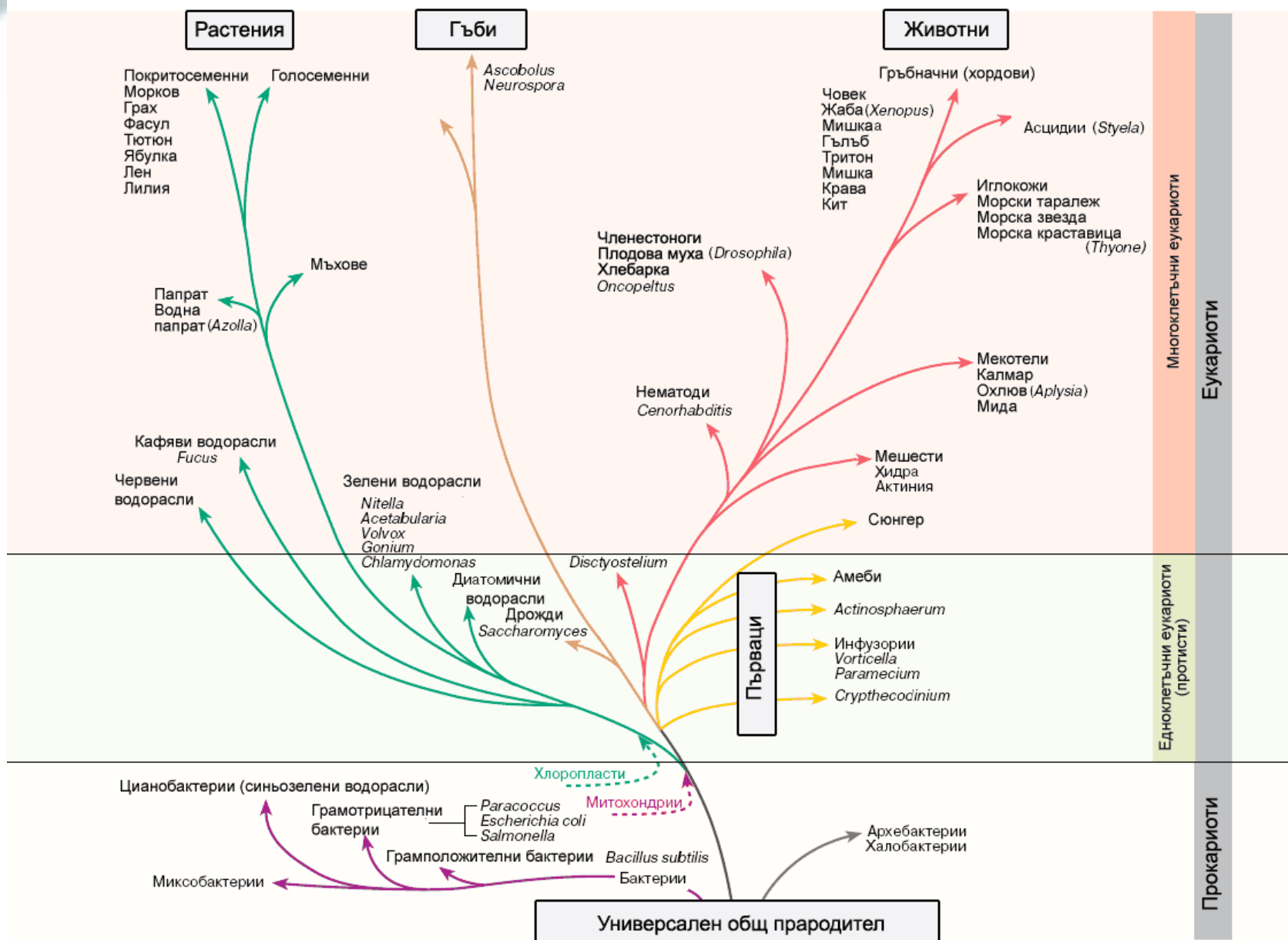
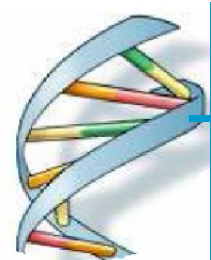
Филогенетика

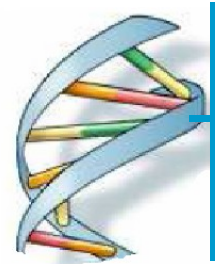




Еволюционно дърво на живота

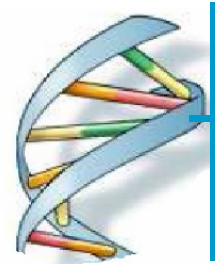






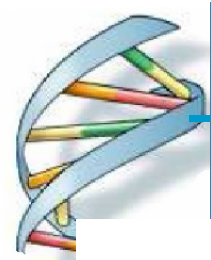
Видове филогенетични дървета

- Филогенетичните дървета, генерирани от изчислителната филогенетика, могат да бъдат или с корен (rooted) или без (unrooted), в зависимост от входните данни и алгоритмите, които се използват.
- Филогенетично дърво с корен е насочено дърво с уникален възел, съответстващ на общ прародител на всички организми в листата на дървото.
- Дървета без корен илюстрират свързаност на листните възли, без да правят предположения за общи прадеди. Дърветата без корен винаги могат да бъдат генерирани от дървета с корени (с пропускане на корена), корена не може да се създаде от дърво без корен без някои средства за идентифициране на прадеди.

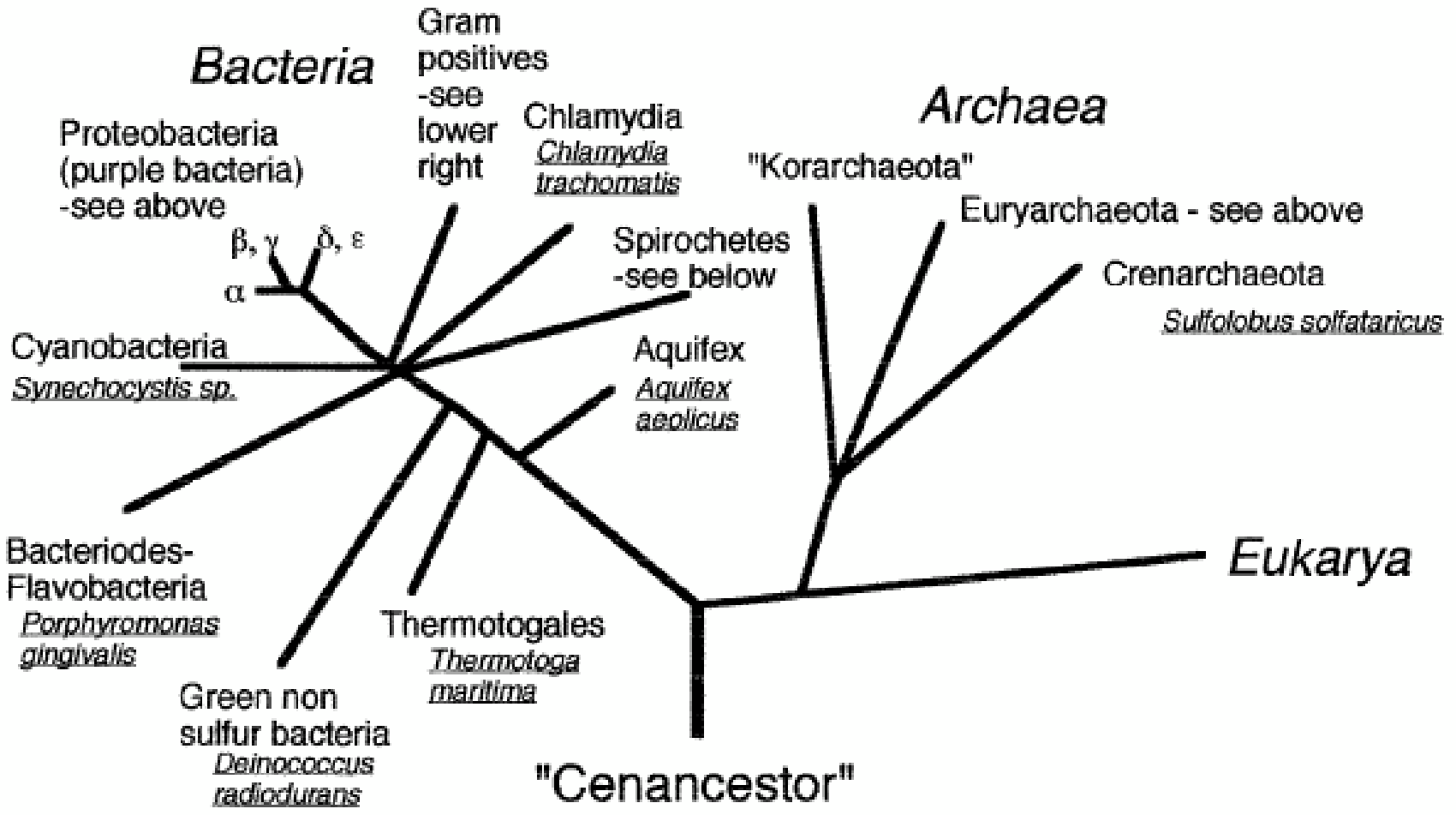


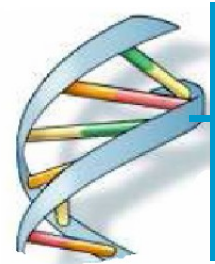
Видове филогенетични дървета

- Дендограм (dendrogram) е общ термин за диаграмно представяне на филогенетично дърво.
- Кладограм (cladogram) е оформено дърво чрез кладистични методи. Този вид дърво представлява само разклонен модел, т.е. дължината на неговите клонове не представлява време.
- Филограм (phylogram) е филогенетично дърво, което изрично представлява броя на характерните промени чрез дължината на своя клон.
- Хронограм (chronogram) или ултраметрично (ultrametric tree) дърво е филогенетично дърво, което чрез дължината на своите клонови представлява еволюционното време.

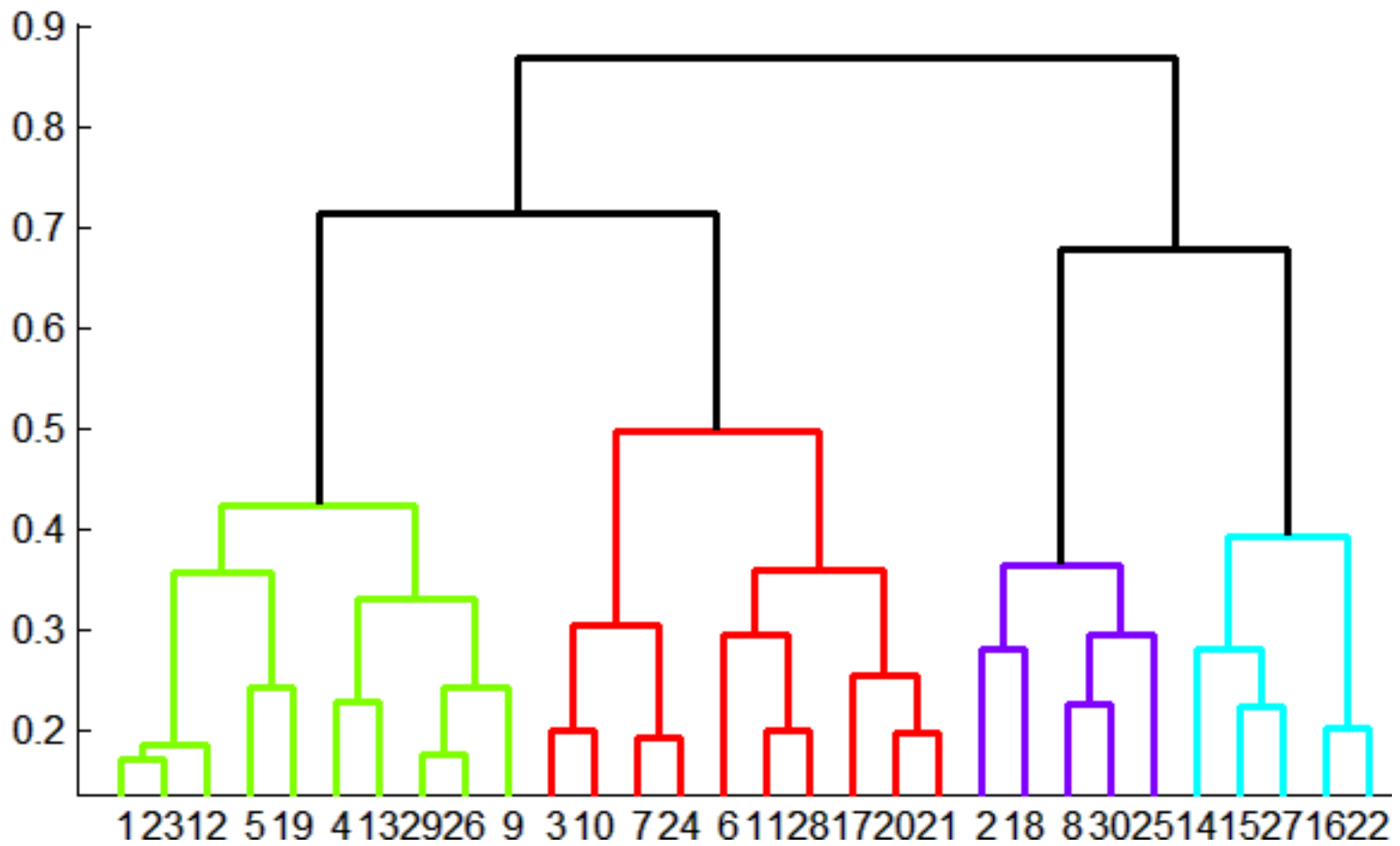


Филогенетично дърво на живота

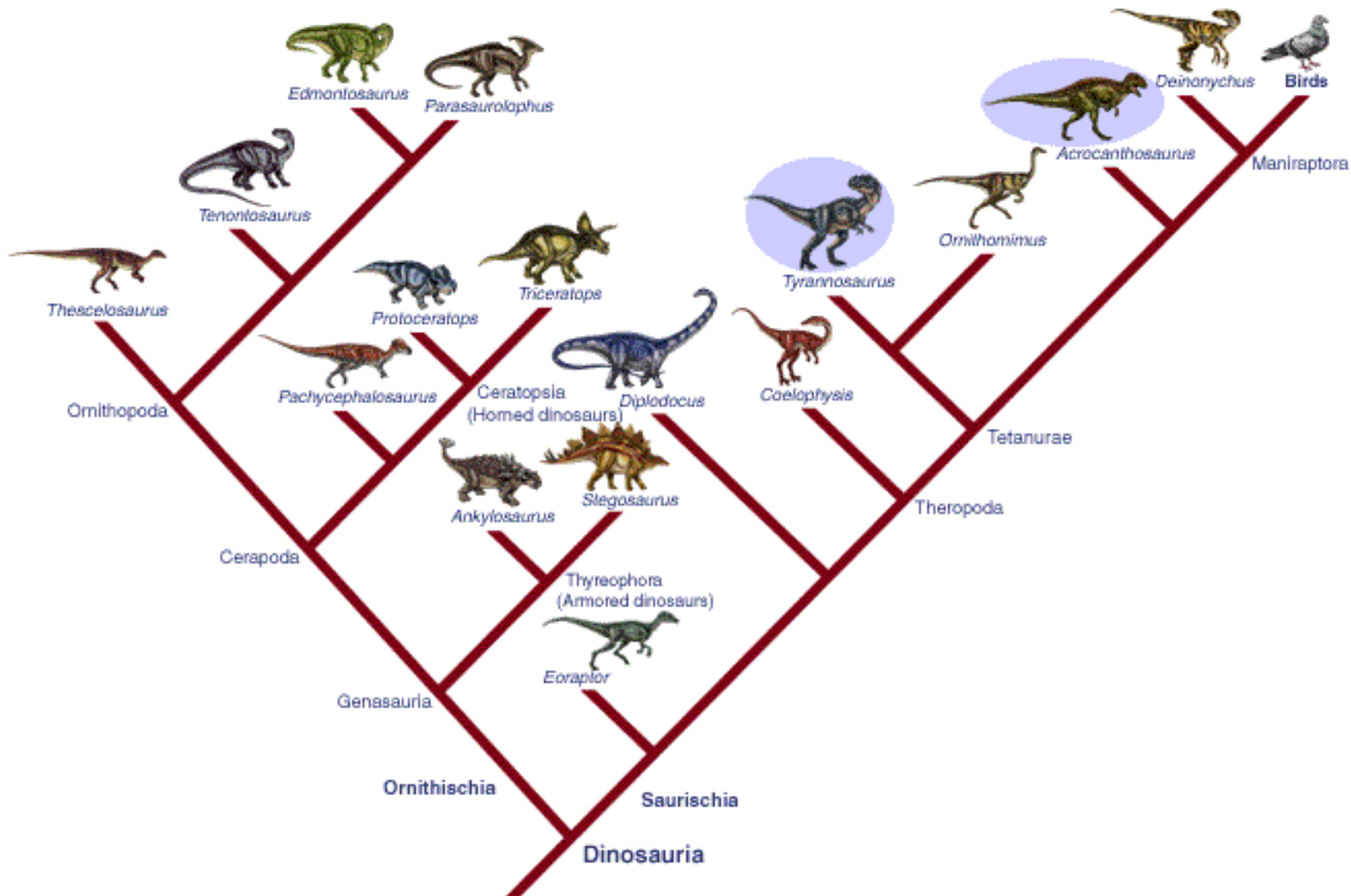


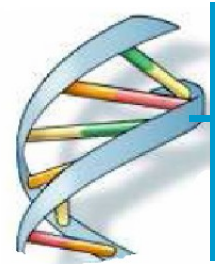


Дендограм

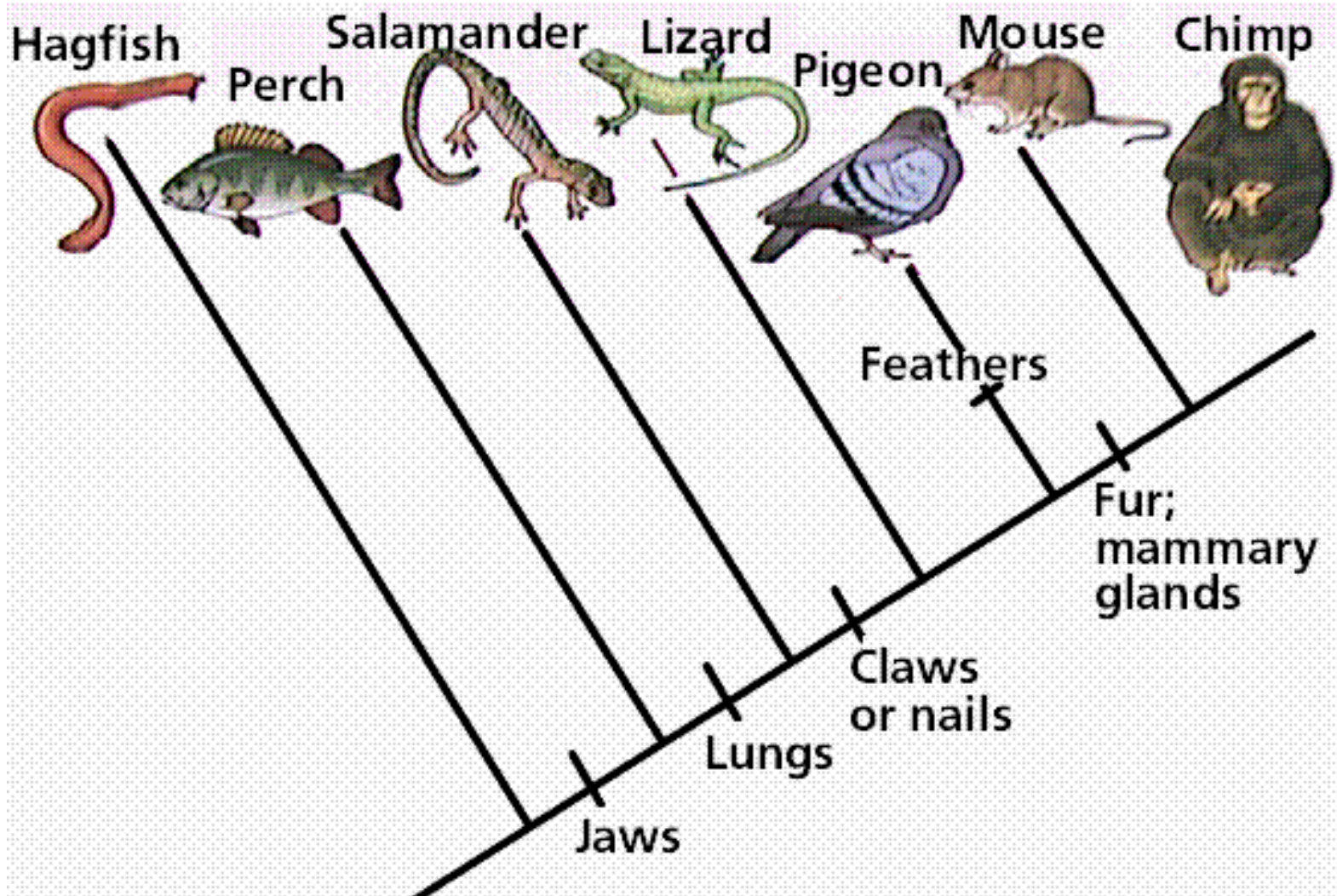


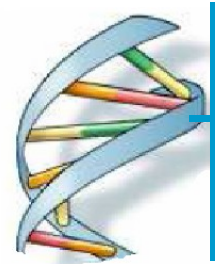
Кладограм





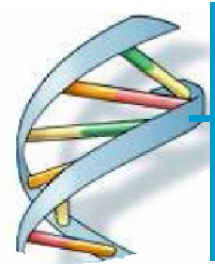
Кладограм





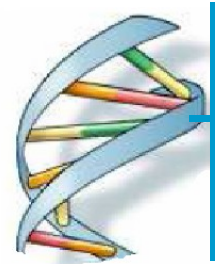
Построяване на филогенетични дървета

- Филогенетичните дървета се конструират чрез филогенетични изчислителни методи.
- Дистанционно-матричните методи като „neighbor-joining“ или „UPGMA“, които изчисляват генетичните разстояния от множество подравнявания на секвенциите (multiple sequence alignments), най-лесно се изпълняват, но не позовават еволюционен модел.



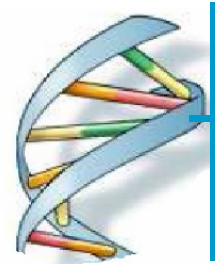
Построяване на филогенетични дървета

- Много методи за подравняване на секвенции, като ClustalW, също така създават дървета с помощта на алгоритми, които се основават на разстояние, за създаване на дървото.
- Методът на максималната пестеливост (maximum parsimony) е друг метод за изчисляване на филогенетичните дърветата, но предполага имплицитен модел на развитие (т.е. пестеливост).



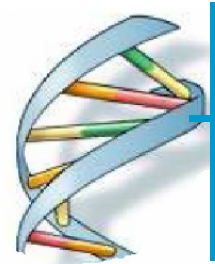
Построяване на филогенетични дървета

- По-съвременните методи използват оптимален критерий на максимална подобност, често в рамките на алгоритъма Bayesian, и прилагат определен модел на еволюцията за оценка на филогенетичното дърво.
- Използвайки много от тези техники, определянето на оптималното дърво е със сложност NP, така че се използва евристичното търсене и методите за оптимизация, за да се идентифицира сравнително добро дърво.



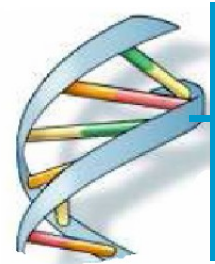
Построяване на филогенетични дървета

- Методите за построяване на филогенетичното дърво могат да бъдат оценени въз основа на няколко критерия:
 - ефективност (колко време трябва за да се изчисли дървото, колко памет е необходима?)
 - мощност (дали използването на данни е полезно, или се губи информацията?)
 - последователност (ще се срещне ли същия отговор няколкократно, ако всеки път дава различни данни за проблема на същия модел?)
 - устойчивост (справят ли се добре с нарушения на предположенията на извършените модели?)
 - „Фалсификация“ (дали ни предупреждава, когато не е добре да се използва, т.е. когато предположенията са нарушени?)



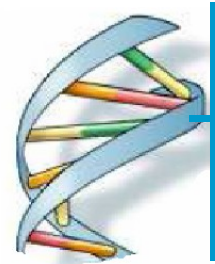
Ограничения при построяването

- Филогенетичните дървета, които са построени на базата на секвенции на гени за различни видове, могат да осигурят еволюционен поглед, но въпреки това те имат важни ограничения.
- Не е задължително да представят еволюционната история. Данните, на които се основават могат да се дискутират.

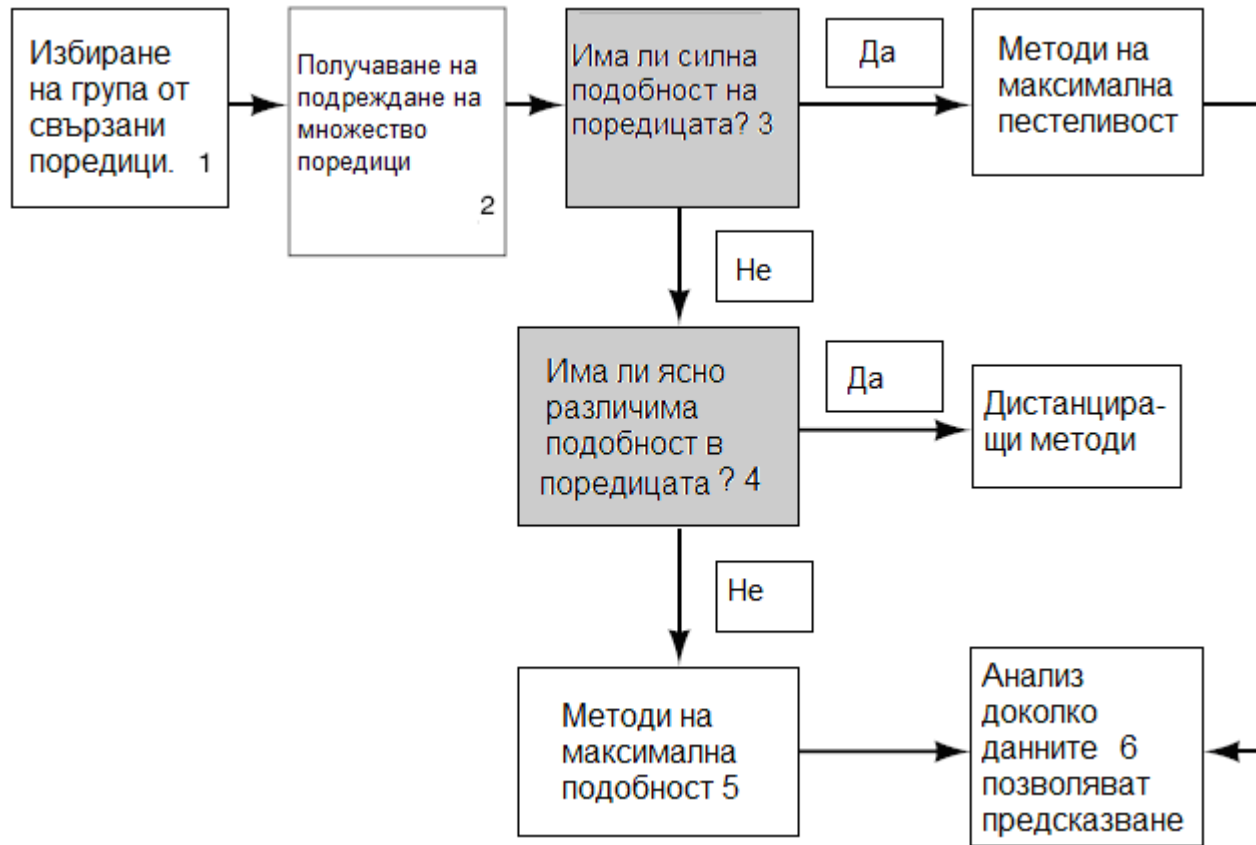


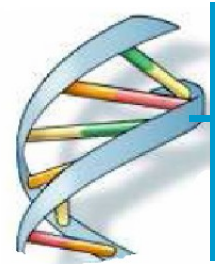
Ограничения при построяването

- Когато в едно дърво са включени изчезващи видове, те са крайни възли, тъй като е малко вероятно, те да са преки предци на съществуващи видове.
- Скептицизъм трябва да се прилага, когато изчезващите видове са включени в дърветата, които са основани изцяло или частично на данни за ДНК последователност.
- Това е така, защото малко полезна "древна ДНК" е запазена около 100000 години. Не могат да се запазят достатъчно дълъги ДНК секвенции за използване във филогенетичният анализ и все още не е възстановен никакъв материал от над 1 милион години.



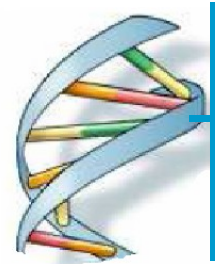
Методи за филогенетична реконструкция





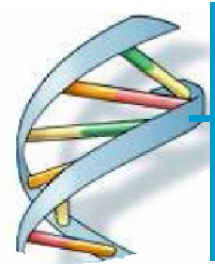
Методи за построяване на филогенетични дървета

- Избраните поредици могат да са ДНК или белтъчни: за всеки тип се използват различни програми и програмни настройки.
- РНК поредиците се анализират с методи на ковариация и чрез анализиране на промените във вторичната структура. Избраните поредици трябва да се наредят една с друга по целите си дължини, в противен случай трябва да имат обща група образци, които да осигурят силен признак за еволюционна свързаност.



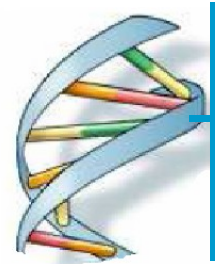
Методи за филогенетична реконструкция

- Нарезждането на двойките поредици не трябва да има голям брой празнини, които са очевидно необходими за нареждане на идентични или свързани символи. Филогенетичният анализ трябва да се извършва само над части от поредици, които могат да бъдат добре наредени.
- Най-общо, филогенетичните методи анализират запазените райони, представени във всички поредици.



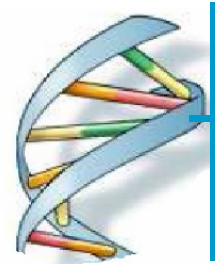
Методи за филогенетична реконструкция

- При анализ по метода максимална пестеливост най-добри резултати се получават, когато сумата от вариациите по всички двойки поредици е подобна (да няма много различаващи се поредици) и когато вариациите са малко.
- Трябва да има ясно мнозинство от определени остатъци в някои колони на нареждането, но и някакво отклонение.
- Тези по-общи остатъци се взимат за представяне на по-ранна група поредици, от които другите са се разклонили. Ако има твърде много вариация, ще има твърде много възможни връзки с предшественици.



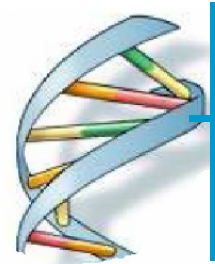
Методи за филогенетична реконструкция

- Чрез използване на дистанциращите методи се възможно да се прогнозира еволюционно дърво, когато е налице вариация между поредиците (някои поредици са по-подобни от други), и при умерени вариации.
- Броят на променените позиции в нареждането между две поредици, разделени на общия брой съвпаднали позиции е разстоянието между поредиците.
- С увеличаване на разстоянието са необходими корекции за отклонения от промени между поредиците.
- Поредици с този тип отклонение могат също да са подходящи за филогенетичен анализ с методите на максимална подобност.
- Дистанциращите методи могат да се ползват с голям брой поредици.

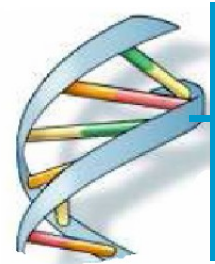


Методи за филогенетична реконструкция

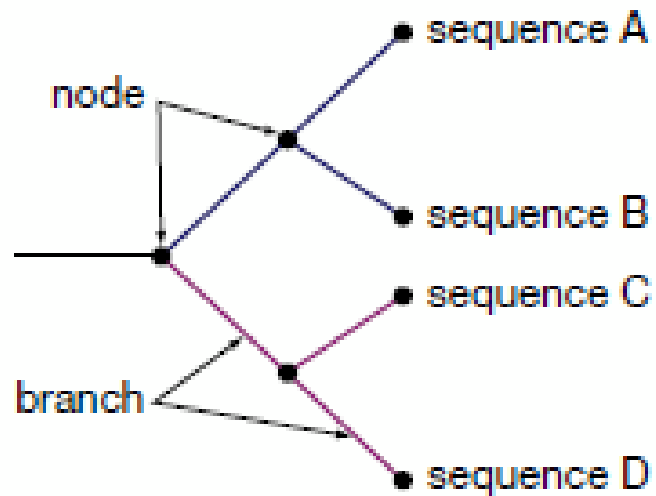
- Методите за максимално съседство могат да се използват за всякаква група свързани поредици, но са особено практични, когато поредиците са по-вариращи.
- Тези методи са изчислително тежки, като сложността на изчисленията нараства с броя на поредиците, понеже трябва да се пресметне вероятността на всяко отделно дърво.
- Предимство на тези методи е, че те осигуряват еволюционни модели за обяснение на отклонението в поредиците.



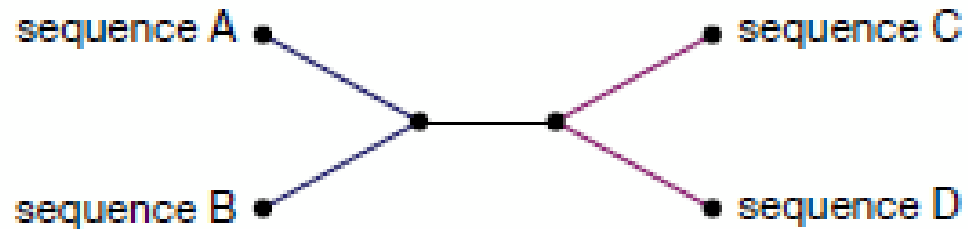
- Данните в наредените колони на многократните поредици се изчисляват отново, за да се провери колко добре се поддържат клоните на еволюционното дърво.

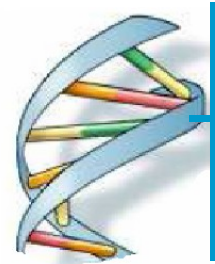


A. Rooted tree

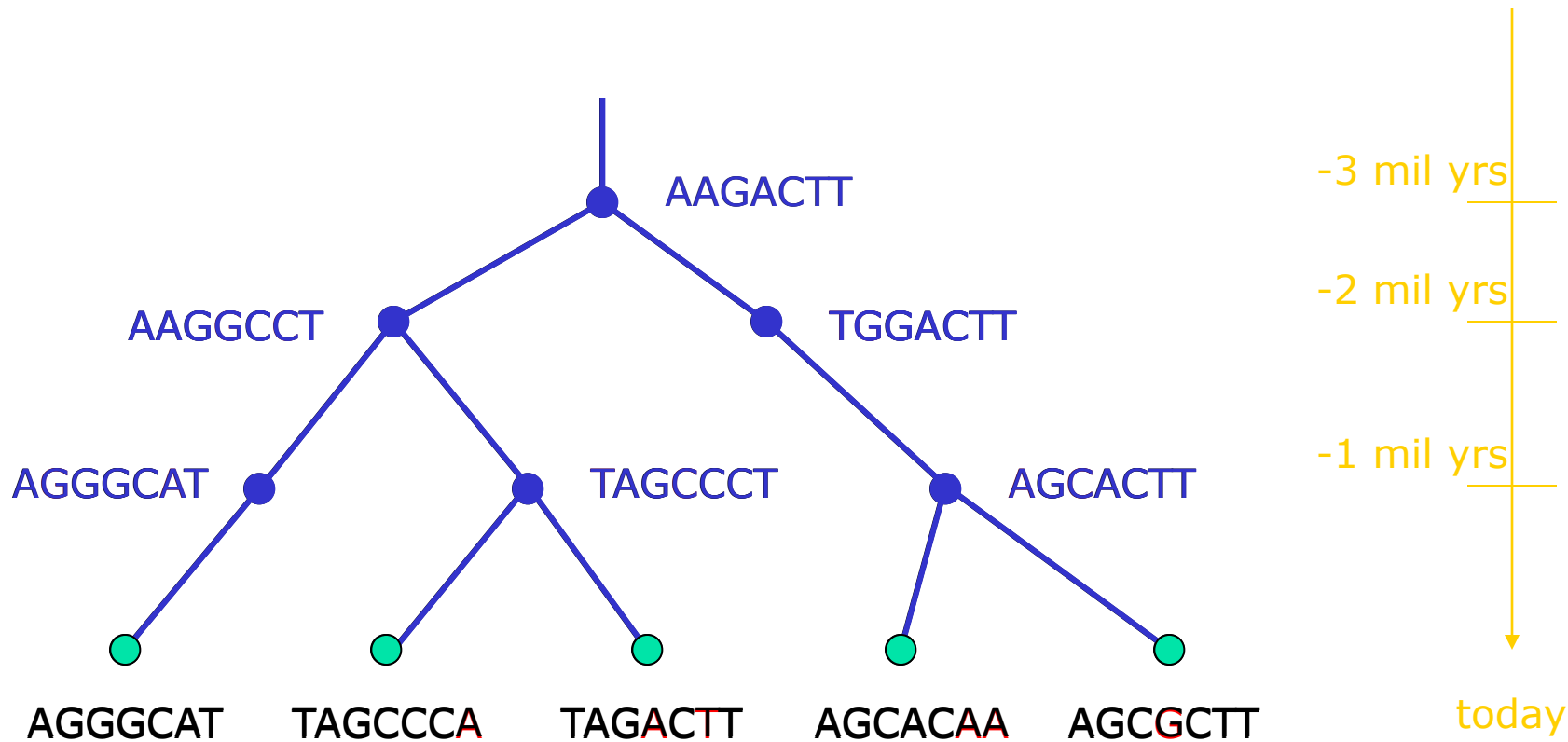


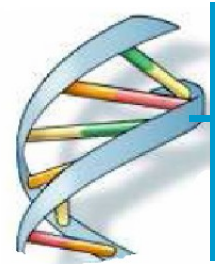
B. Unrooted tree





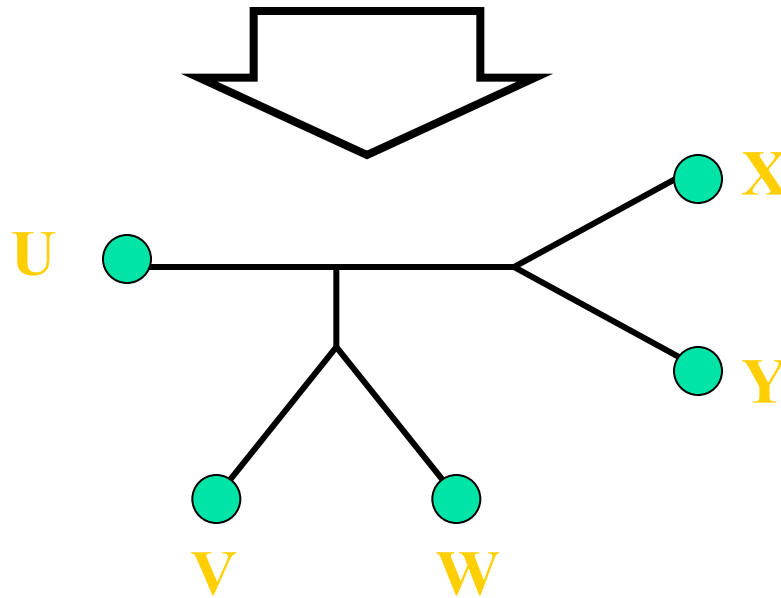
Еволюция на DNA

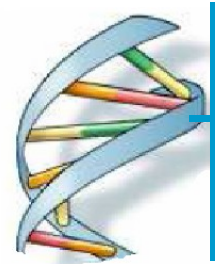




Филогенетична реконструкция

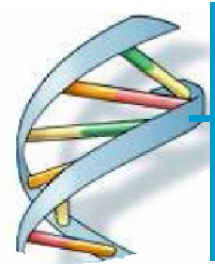
U V W X Y
AGGGCAT TAGCCCA TAGACTT TGCACAA TGCAGCTT





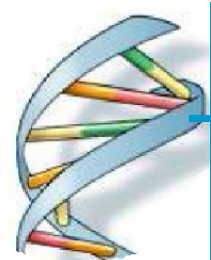
Стъпки при филогенетичен анализ

- Събиране на данни
- Подравняване на последователности
- Възстановяване филогенетиката на няколко подравнявания - често получаване на голям брой дървета
- Избор на консенсус (оценка на надеждни компоненти на еволюционната история)
- Анализ на дървото

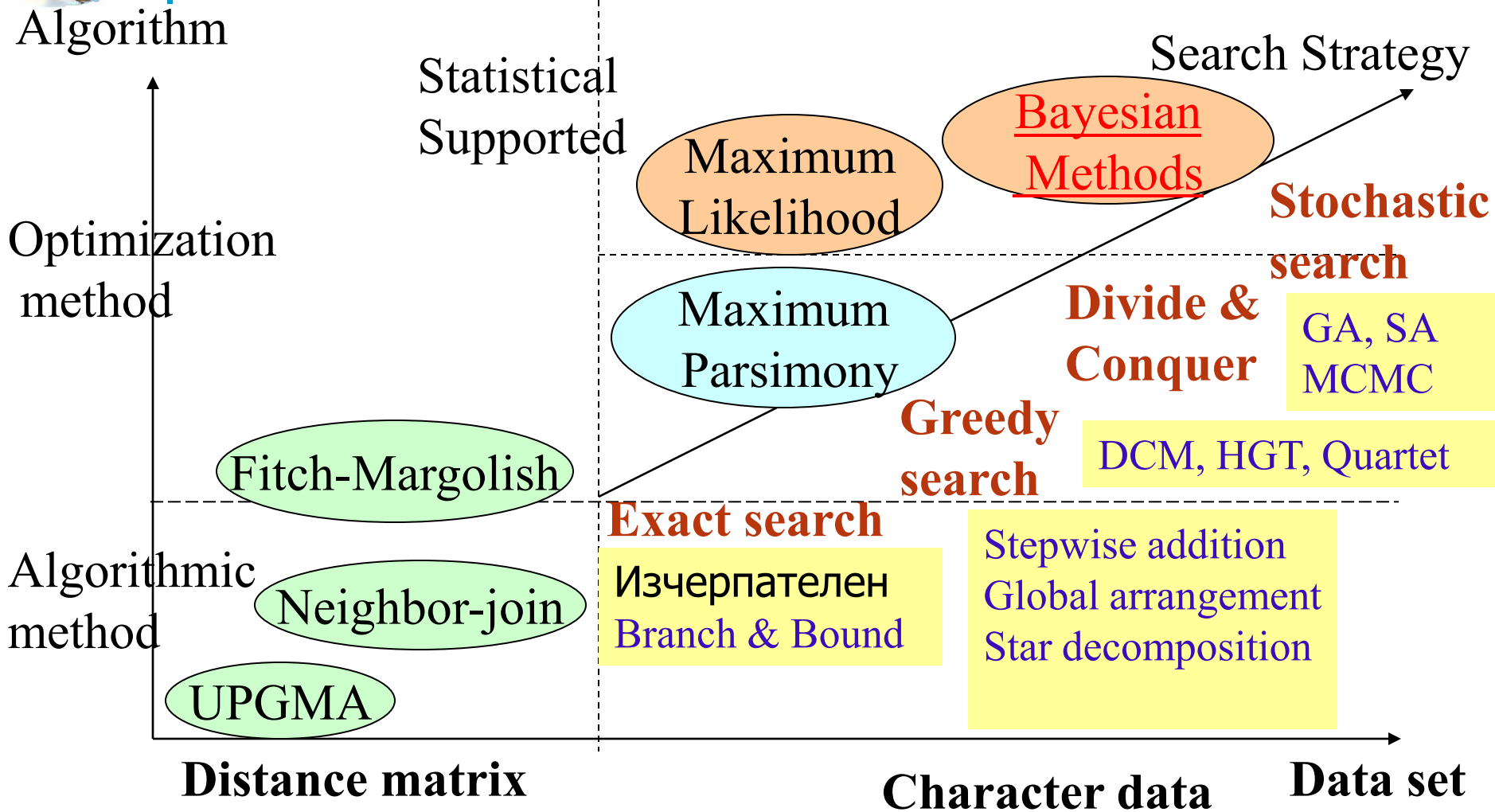


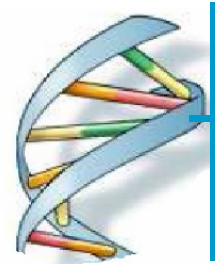
Изследване на алгоритми

- Евристики за NP-тежки проблеми при филогенетична реконструкция
- Компактно представяне на дървета
- Производителност на методите за филогенетична реконструкция при стохастични модели на еволюцията
- Подравняване на гени
- Distance-based реконструкция
- Еволюция на гените
- Множествено подравняване на секвенции



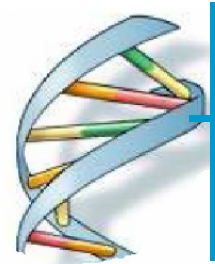
Филогенетични методи





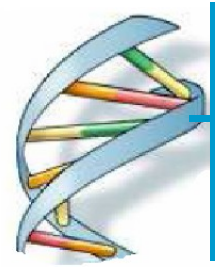
Критерии за производителност

- Време за изпълнение.
- Пространство.
- "Топологични точност" по отношение на вярно построеното дърво. Обикновено получено чрез симулации.
- Точност по отношение на определен критерий (например дължина на дървото или правдоподобност на резултат) на реални данни.



MEGA – Molecular Evolutionary Genetic Analysis Software

- MEGA - Molecular Evolutionary Genetics Analysis – свободно достъпен софтуер за създаване на дендограми и филогенетични дървета с използване на протеинови секвенции. Разработен от [Pennsylvania State University](http://www.megasoftware.net/)
<http://www.megasoftware.net/>
- MEGA е интегрирано средство за автоматично и ръчно подравняване на секвенции, изграждане на филогенетични дървета, извличане на web базирани бази данни, оценка на молекулярната еволюция и тестване на еволюционни хепотези.
- MEGA е Windows базирано приложение.



PHYLIP (PHYLogeny Inference Package)

- Свободно достъпен пакет от програми за изграждане на филогенетични дървета.
- Съдържа 35 програми, наличен е сорс код, на езика С и прекомпилирани изпълними програми за Windows и Linux операционни системи.
- Автор е [Joseph Felsenstein](#), професор в Department of Genome Sciences and the Department of Biology at the University of Washington, Seattle

