

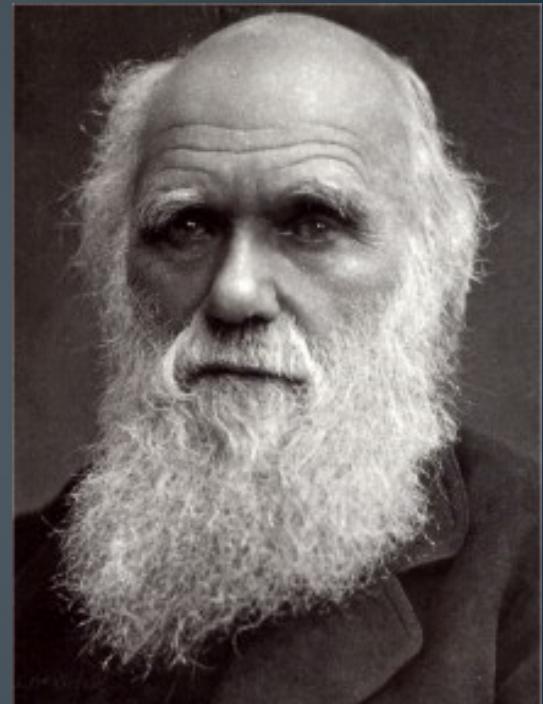
# Základy fylogenetiky

Přednáška 9

ABIN – 2017/18

# Evoluce

- Evoluce je řízena:
  - **diversitou** – různí jednici v populaci organismu představují varianty stejné základní šablony genomu,
  - **mutací** – DNA sekvence může být změněna jednonukleotidovou mutací SNP, delecí nebo inzercí nukleotidu, nebo chromozomálními přestavbami,
  - **selekcí** – různé mechanismy způsobují přenos mutací do dalších generací.



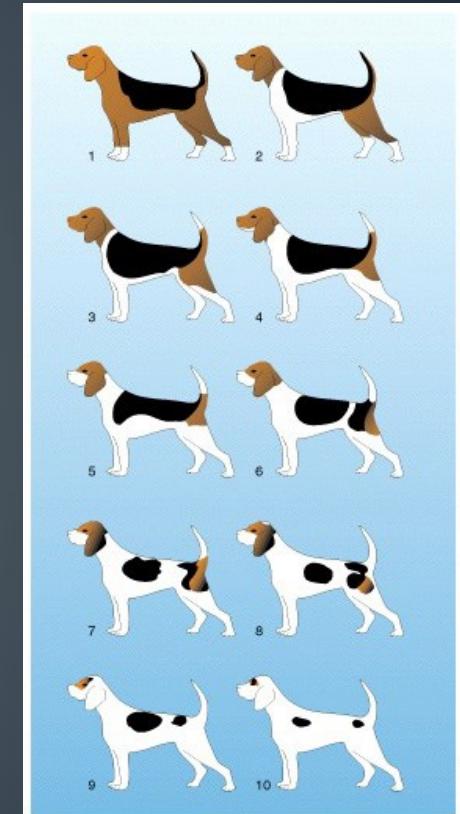
Charles Robert  
Darwin (1809–82)

# Fenotypová variabilita

- Fenotypová variabilita je jednou z univerzálních vlastností přírodních populací.
- Mezi jedinci jakékoliv populace existují různé fenotypy ve většině znaků.

Fenotyp je ovlivněn:

- genotypem,
- epigenezí (ovlivnění exprese, methylace DNA),
- fyzikálně-chemickými vlivy prostředí.

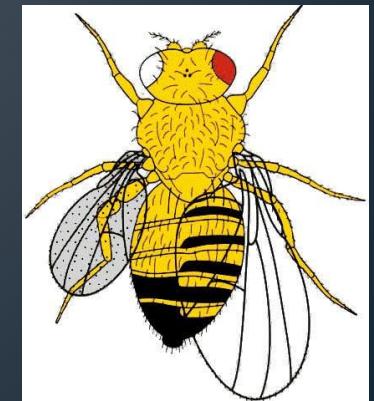


# Epigenetika

- Studium dočasných a lokálních změn genové exprese během života jedince, které nejsou způsobeny změnou nukleotidového složení DNA.
- Vypínání/zapínání genů ovlivněno vnějšími faktory.
- Epigenetické změny mohou být dědičné mitoticky i meioticky.
- Př.: různé tělní buňky – mají identický genom, ale fenotypově se liší díky epigenomu (které geny jsou aktivní a které neaktivní).
- Různá vývojová stádia organismu (larva, kukla, dospělec), pohlavní dimorfismus, atd.

# Mutace

- Mutacemi se nazývají změny ve struktuře genetického materiálu, kdy se:
  - mění smysl genetické informace, a zároveň
  - nejsou porušena syntaktická pravidla jejich zápisu.
- V opačném případě se jedná o tzv. poškození DNA.



# Mutace

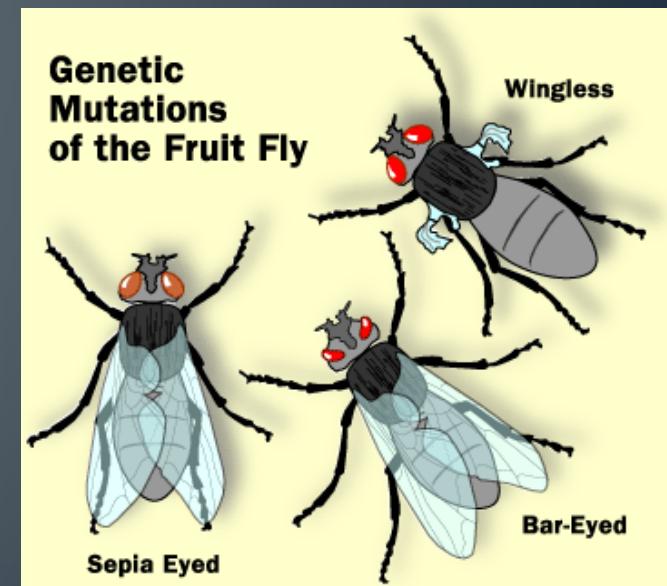
- Mutace patří k systematickým evolučním procesům, které mění genové (alelové) četnosti v populacích.
- Pojem mutace se používá pro všechny genetické změny, které mohou přejít na potomky.
- K mutacím dochází ve velmi malé četnosti.

# Příčiny vzniku mutací

- Spontánní – během replikace může dojít k nesprávnému zařazení nukleotidu. Malá pravděpodobnost a existence reparačního systému.  
(Př. frekvence u člověka je  $10^{-10}$  na nukleotid a buněčný cyklus.)
- Indukované jsou způsobené vnějšími vlivy (mutageny), které mohou být fyzikální (např. radiace) nebo chemické.

# Mutace

- Mutace lze rozdělovat z hlediska vlivu na biologickou zdatnost jejich nositele na:
  - mutace pozitivní (užitečné),
  - mutace negativní (škodlivé),
  - mutace selekčně neutrální.
- Mutace ovlivňující morfologické a behaviorální znaky jsou obvykle selekčně významné (tj. pozitivní nebo negativní).

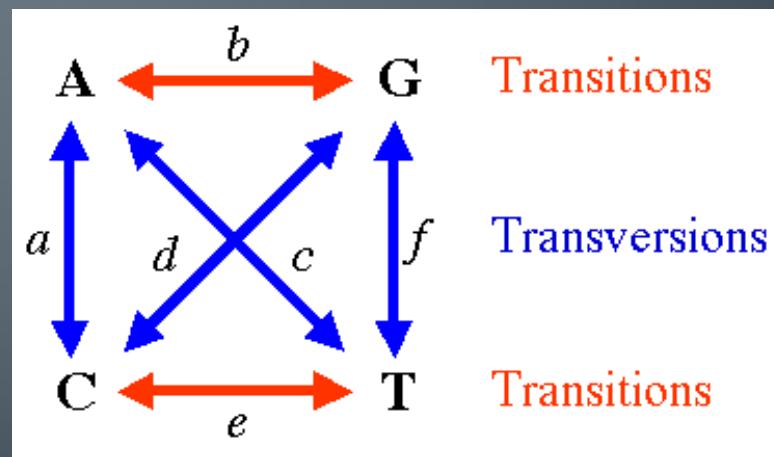


# Typy mutací

- Podle fyzické povahy **mutace rozdělujeme** na:
  - bodové,
  - na úrovni úseků DNA (řetězcové),
  - chromozomové (rozsáhlé translokace, fúze či rozpad celých chromozomů),
  - genomové (méně nebo více kopií chromozomů, polyploidizace).

# Bodové mutace

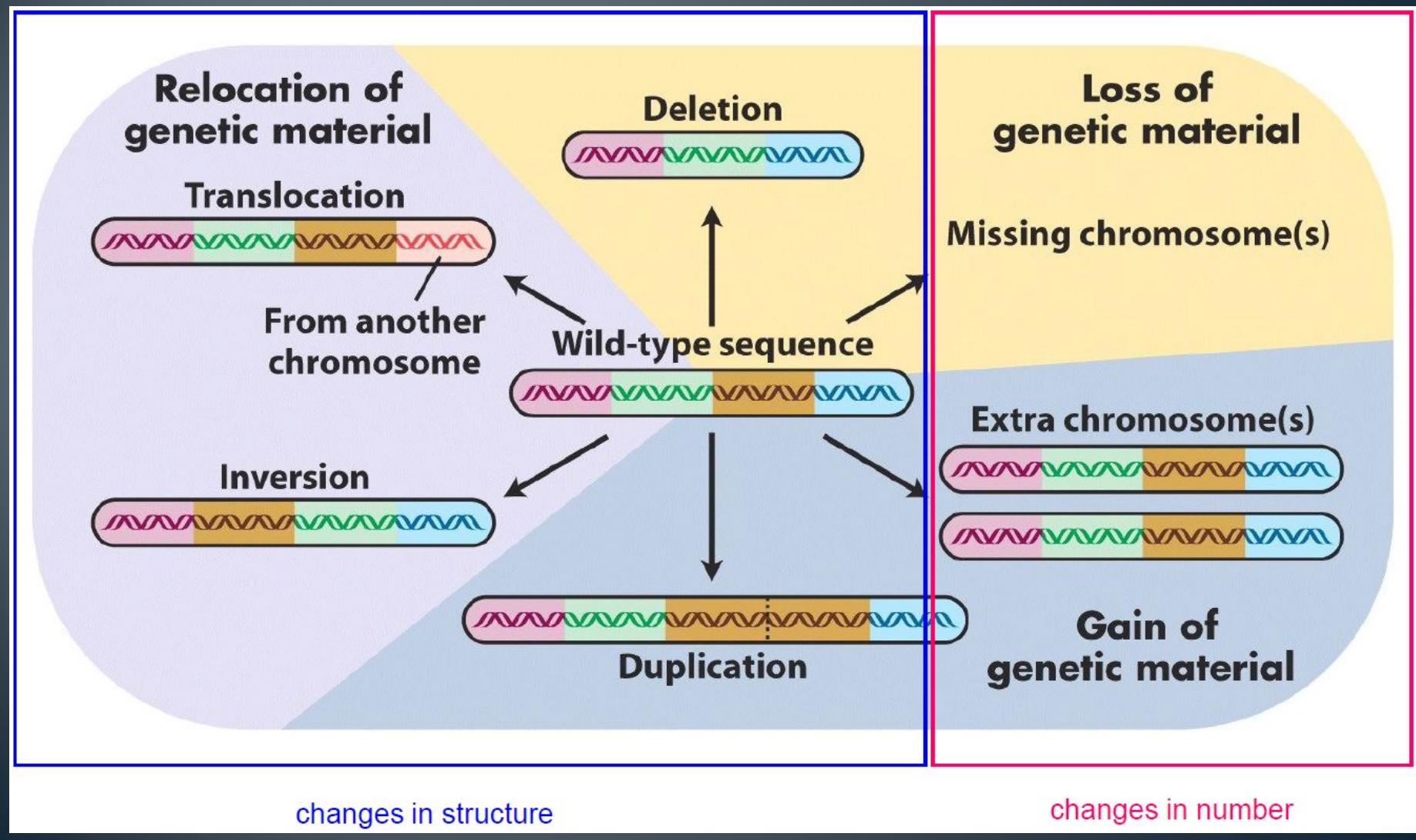
- Bodové mutace:
  - substituce (tranzice, transverze), delece a inzerce.
- Tranzice: záměna nukleotidu nukleotidem stejného typu (purin – purin, pyrimidin - pyrimidin).
- Transverze: záměna pyrimidinu purinem nebo naopak.
- Delece a inzerce: lokální změna počtu nukleotidů.



# Vliv bodových mutací

- Pokud dochází k mutacím **v úseku kódujícím protein**, rozdělují se na:
  - **synonymní mutace** (beze změny smyslu, záměna za kodon kódující stejnou aminokyselinu),
  - **nesynonymní mutace** (se změnou smyslu, záměna aminokyseliny).
  - **nesmyslné mutace** (předčasné ukončení translace proteinu – vznik stop kodonu) = poškození.

# Řetězcové x genomové mutace



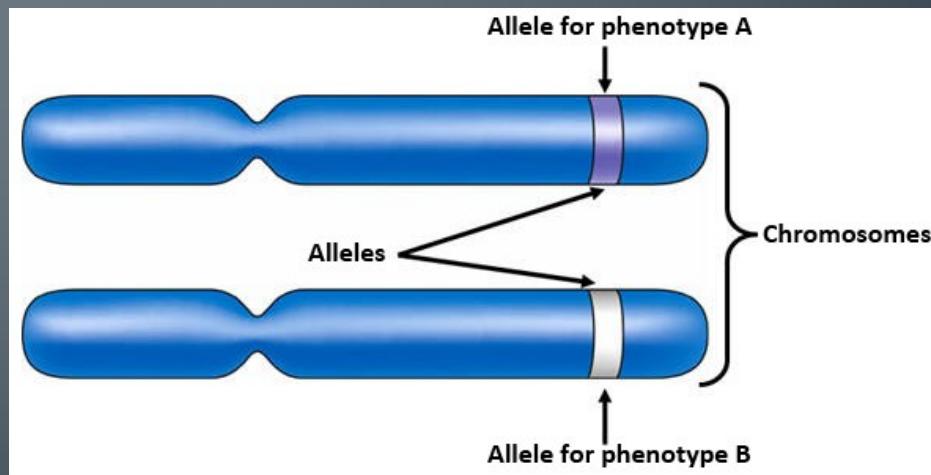
# Řetězcové mutace

- Rozlišujeme:
  - delece,
  - inzerce,
  - duplikace,
  - transpozice
  - translokace,
  - inverze.

výchozí stav	<table border="1"><tr><td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td></tr><tr><td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td></tr></table>	a	b	c	d	e	f	g	a	b	c	d	e	f	g				
a	b	c	d	e	f	g													
a	b	c	d	e	f	g													
tandemová duplikace	<table border="1"><tr><td>a</td><td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td></tr></table>	a	a	b	c	d	e	f	g										
a	a	b	c	d	e	f	g												
inzerce	<table border="1"><tr><td>a</td><td>z</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td></tr></table>	a	z	b	c	d	e	f	g										
a	z	b	c	d	e	f	g												
delece	<table border="1"><tr><td>a</td><td>b</td><td>d</td><td>e</td><td>f</td><td>g</td></tr></table>	a	b	d	e	f	g												
a	b	d	e	f	g														
transpozice	<table border="1"><tr><td>a</td><td>a</td><td>b</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td></tr><tr><td></td><td></td><td></td><td></td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td></tr></table>	a	a	b	b	c	d	e	f	g					c	d	e	f	g
a	a	b	b	c	d	e	f	g											
				c	d	e	f	g											
reciproká translokace	<table border="1"><tr><td>d</td><td>e</td><td>f</td><td>d</td><td>e</td><td>f</td><td>g</td></tr><tr><td>a</td><td>b</td><td>c</td><td>a</td><td>b</td><td>c</td><td>g</td></tr></table>	d	e	f	d	e	f	g	a	b	c	a	b	c	g				
d	e	f	d	e	f	g													
a	b	c	a	b	c	g													
inverze	<table border="1"><tr><td>a</td><td>b</td><td>f</td><td>e</td><td>d</td><td>c</td><td>g</td></tr></table>	a	b	f	e	d	c	g											
a	b	f	e	d	c	g													

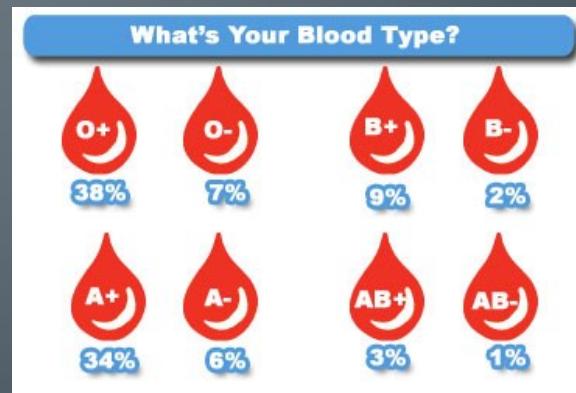
# Genetická variabilita

- Genetickou variabilitou způsobující fenotypovou diverzitu se zabývá genetika populací.
- Existence genetické variability je nutnou podmínkou evolučních změn.
- Ve většině přírodních populací existuje genetická variabilita ve formě mnohanásobných alel u mnoha genů.



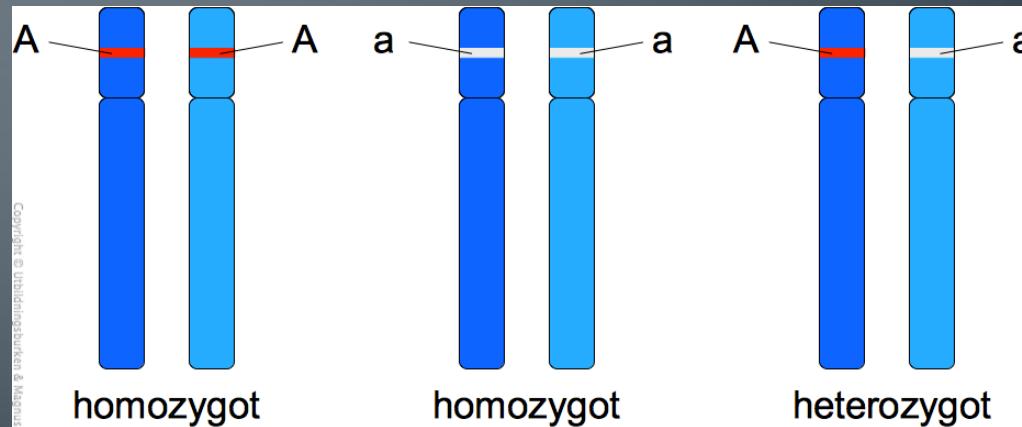
# Alely a polymorfismus

- Alela je konkrétní forma genu v jednom lokusu, může mít více variant v populaci.
- Lokus je název části chromozomu.
- Při existenci více alel v populaci se jedná o polymorfismus (polymorfní lokus).
- Př: polymorfismus v genu RHD určující lidský Rh faktor: tento gen se vyskytuje v alelách Rh+ a Rh-.



# Homozygoti a heterozygoti

- Mnoho eukaryotních organismů je **diploidních** (mají dva páry chromozomů = dvě kopie každého lokusu).
- Jsou-li dvě sekvence na odpovídajících lokusech kopie stejné alely, je organismus na tomto lokusu **homozygotní**.
- Při odlišných alelách je organismus na tomto lokusu **heterozygotní**.



# Vznik a uchování genetické variability

- Genetický polymorfismus vznikl především pomocí **mutací** a přírodnímu **výběru**.
- **Polymorfní geny** poskytují u diploidních organismů možnost vzniku velkého množství genotypů.
- Při  $n$  alelách genu může existovat:
  - $n$  homozygotů,
  - $n(n - 1)/2$  heterozygotů,
  - celkem  $n(n + 1)/2$  genotypů.

Příklad:

alely A, B, C

3 homozygoti: AA, BB, CC

3 heterozygoti: AB, AC, BC

6 genotypů celkem

# Vznik a uchování genetické variability

- V případě velkého množství polymorfních genů bude u diploidních organismů existovat velké množství **vícenásobných heterozygotů** (heterozygotů v různých genech).
- $n$ -násobný heterozygot může vytvořit  $2^n$  různých typů gamet.
- Odhaduje se, že u člověka je 6,7 % genů v heterozygotní konstituci.
- Z odhadovaných 23 500 genů člověka je asi 1 575 heterozygotních genů, ze kterých může teoreticky vzniknout  $2^{1575}$  ( $\sim 10^{480}$ ) různých typů gamet.
- Tento počet gamet nemůže vytvořit nejen jedinec, ale ani všichni lidé za dobu existence lidstva.
- Žádné dvě gamety proto nejsou shodné.

# Fylogenetika

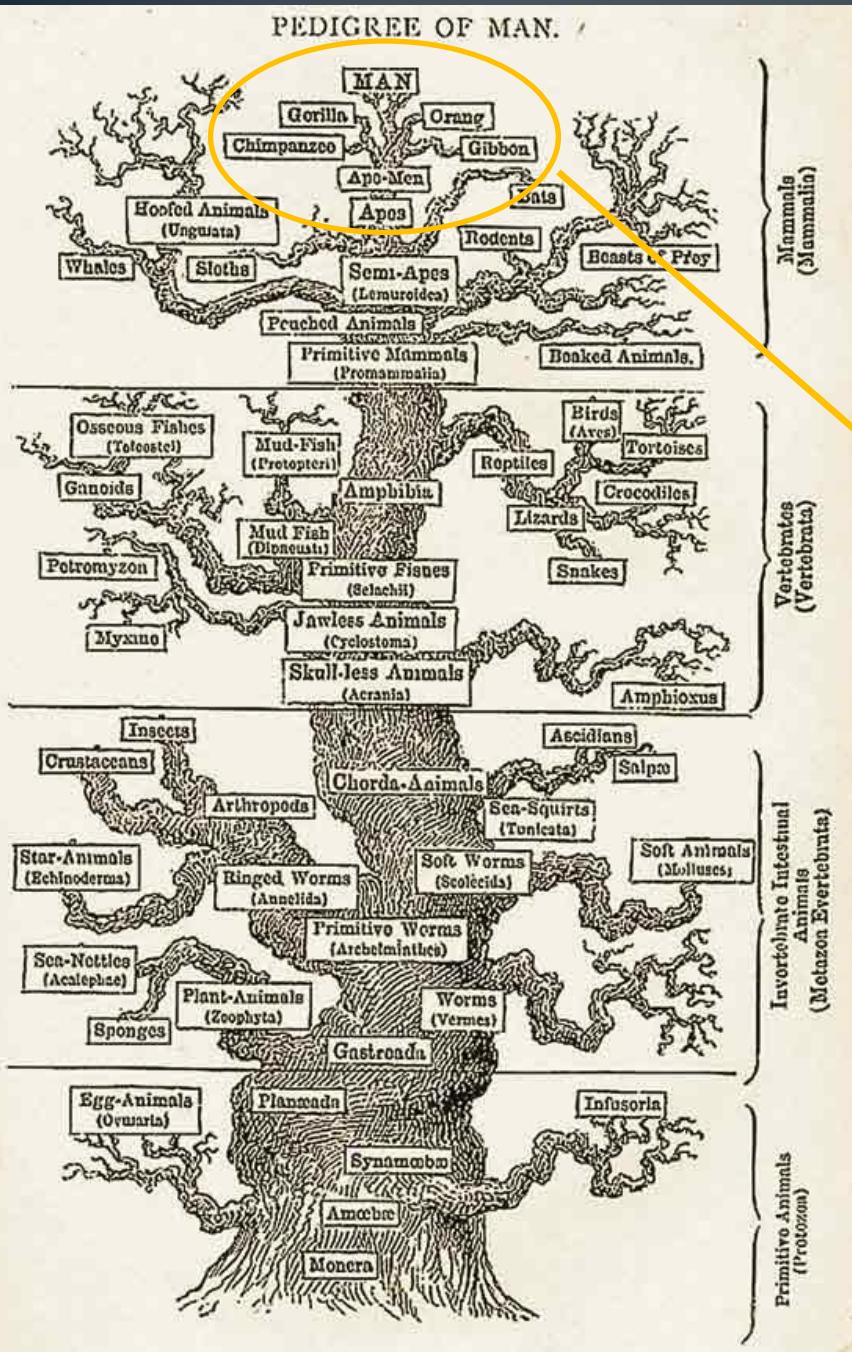
- Věda zabývající se studiem **příbuzenských** vztahů mezi organismy.
- Analyzuje vznik a vývoj jednotlivých vývojových linií = **taxonů**.
- **Rekonstruuje** pořadí a způsob větvení taxonů v průběhu evoluce.
- Vychází z předpokladu, že všechny známé druhy organismů žijící na Zemi vznikly v průběhu fylogeneze ze **společného předka**.

# Fylogenetický strom

- Fylogeneze (fylogenetický vývoj) je vývoj druhů organismů v historickém sledu ve smyslu evoluční teorie.
- Věda zkoumající fylogenezi se nazývá fylogenetika.
- Fylogeneze bývá popsána fylogenetickými stromy, které znázorňují posloupnost událostí vedoucí k formování dnešní množiny druhů.

# Historie fylogenetických stromů

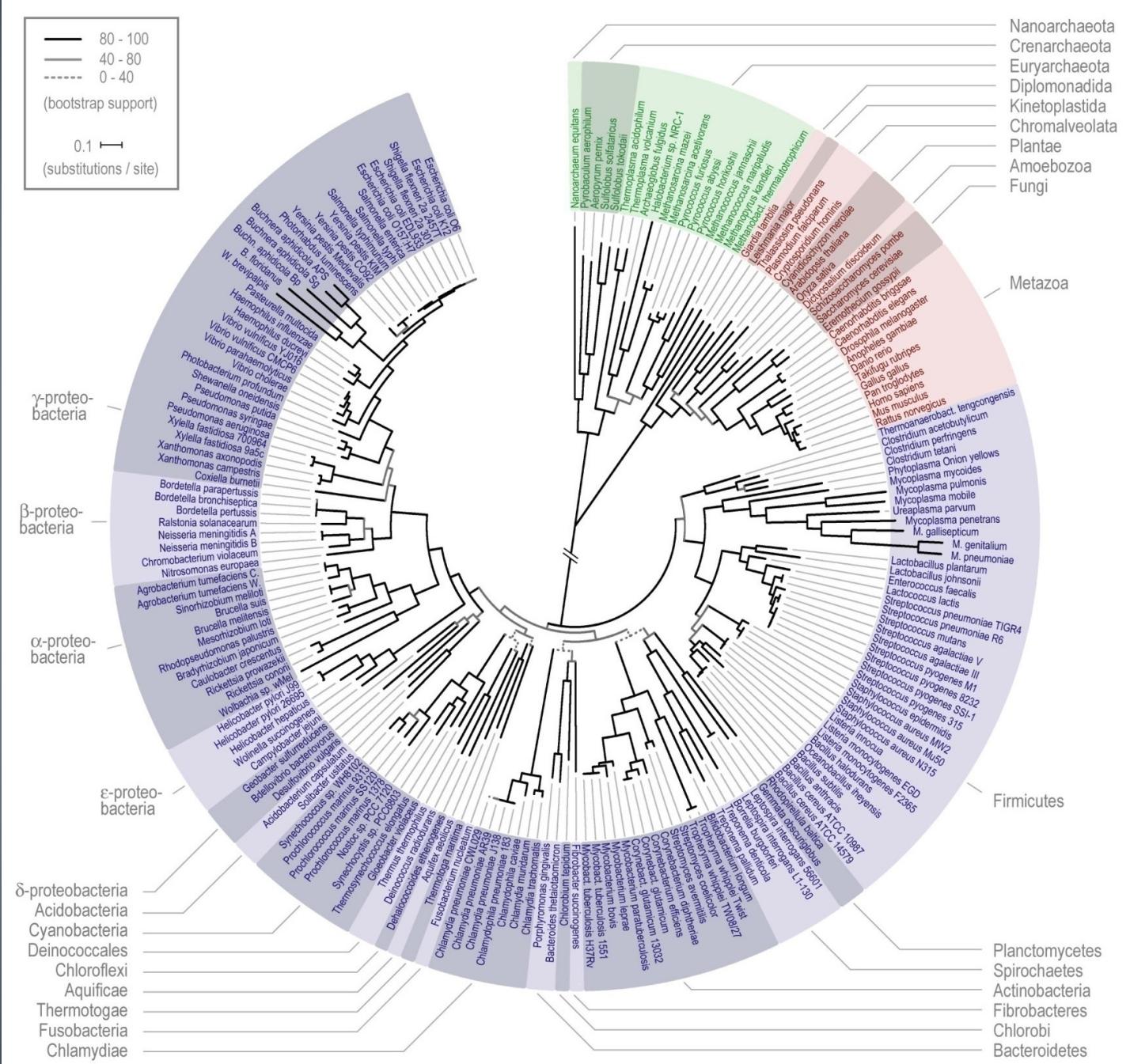
- Do 50. let minulého století byly fylogenetické stromy konstruovány experty na základě jejich zkušenosti - **subjektivní kritéria**.
- Poté byla hledána **objektivní kritéria** pro konstrukci fylogenetických stromů.



# Strom života – první rodokmen člověka



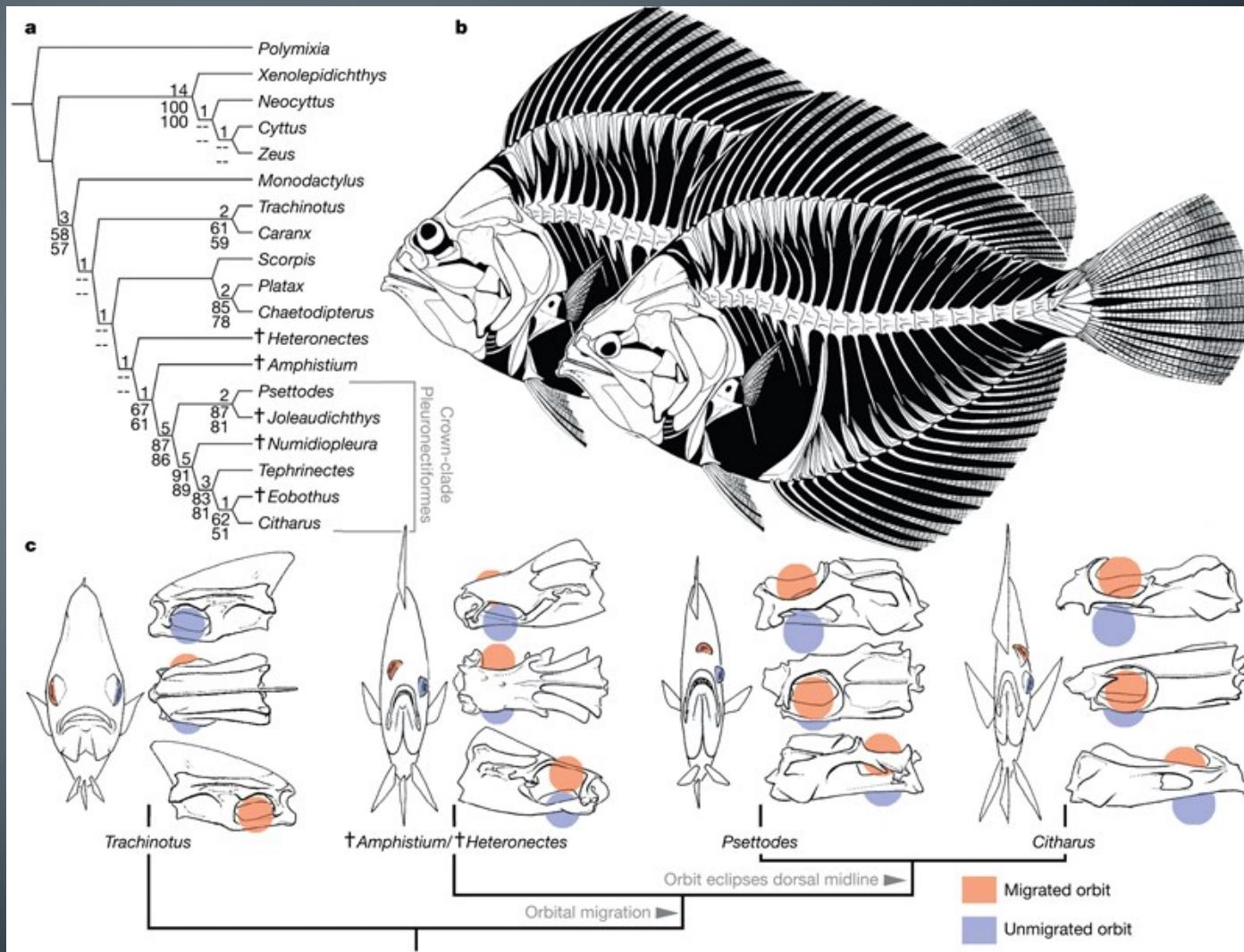
# Strom života – dnešní pohled



# Morfologické a molekulární znaky

- Klasická fylogenetická analýza vychází z **morfologických znaků**: počet nohou, délka nohou, srst, poloha očí, ...
- Dnešní metody využívají **molekulární znaky**:
  - DNA sekvence,
  - proteinové sekvence.

# Morfologická topologie



# Od sekvence k fylogenetickému stromu

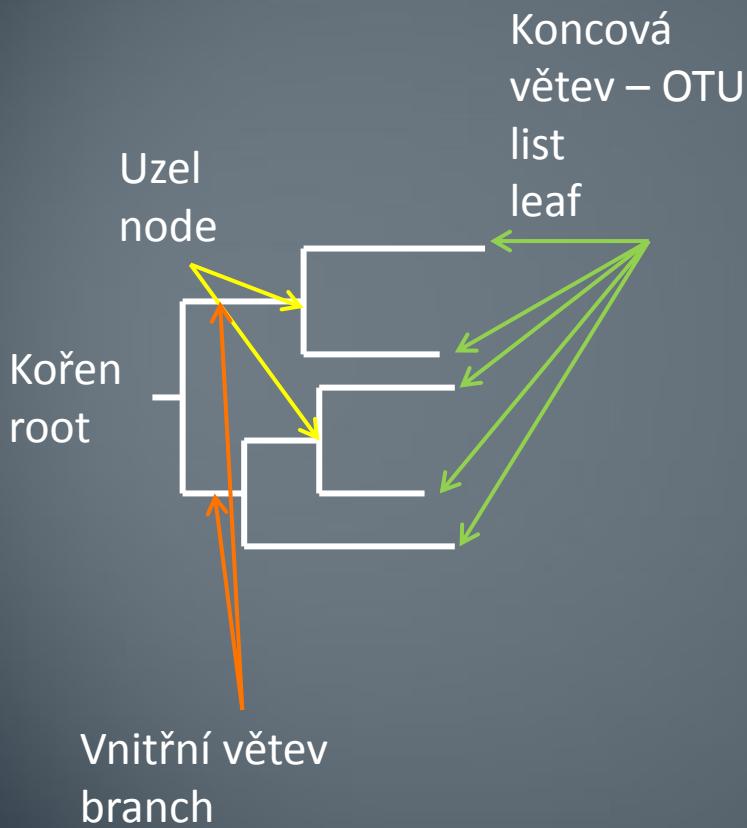
- potkan           **QEPGGLVVPPTDA**
- králík           **QEPGGMVVPPPTDA**
- gorila           **QEPGGLVVPPTDA**
- kočka           **REPGLVVPPTEG**



# Molekulární znaky

- Nepodléhají také selekčním tlakům prostředí.
- Jsou vzájemně **distinktní**.
- Mají **kvalitativní** charakter (nelze je spočítat jako např. počet končetin).
- Dají se **modelovat** matematickými algoritmy.
- Umožňují třídit i vzájemně nepodobné a nepříbuzné organismy.

# Fylogenetický strom

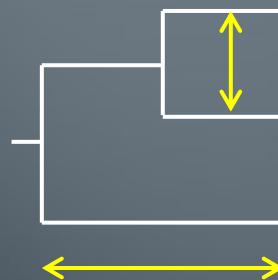


- Kořen stromu = společný prapředek.
- Koncová větev = OTU (Operační taxonomická jednotka) reprezentuje jednotlivé organismy (sekvence).
- Uzel = společný předek dvou OTU.
- Délka větví = doba trvání evoluce nebo míra rozdílnosti sekvencí.

# Stromové struktury

## Dendrogram (kladogram)

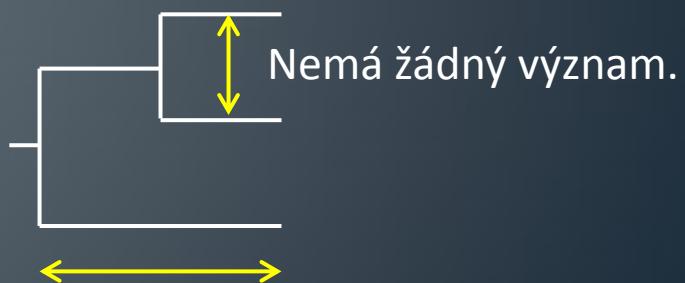
- znázorňuje pořadí a způsob větvení jednotlivých vývojových linií – topologii,
- délka větví ani úhel mezi nimi nenese žádnou informaci.



Nemá žádný význam.

## Fylogram x fenogram

- délka větví vyjadřuje dobu trvání jednotlivých druhů,
- nebo množství evolučních změn, ke kterým došlo v jednotlivých větvích.



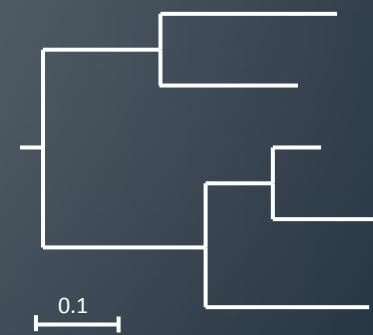
Délka trvání evoluce,  
množství evolučních  
změn.

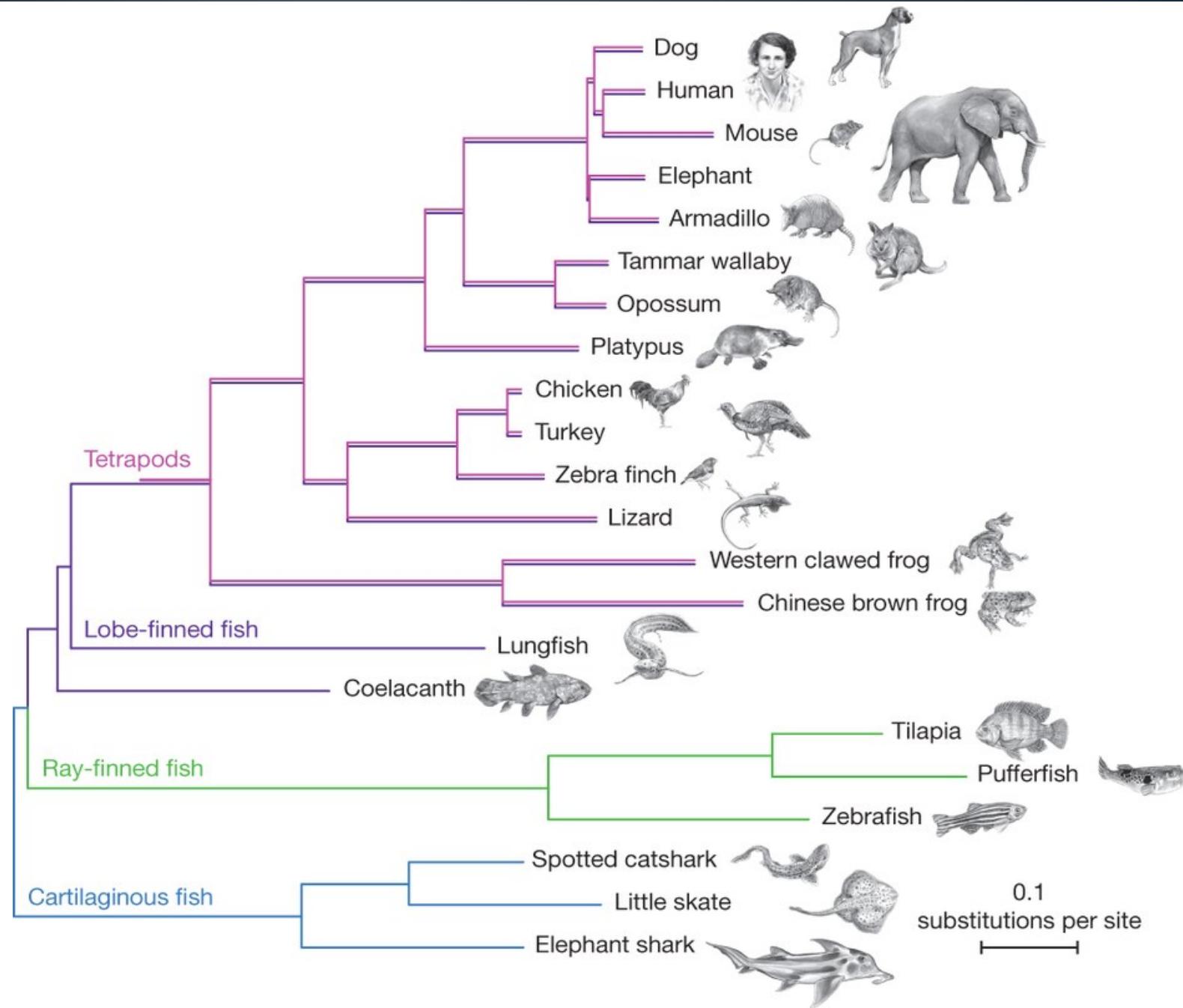
# Kořenový vs. nekořenový strom

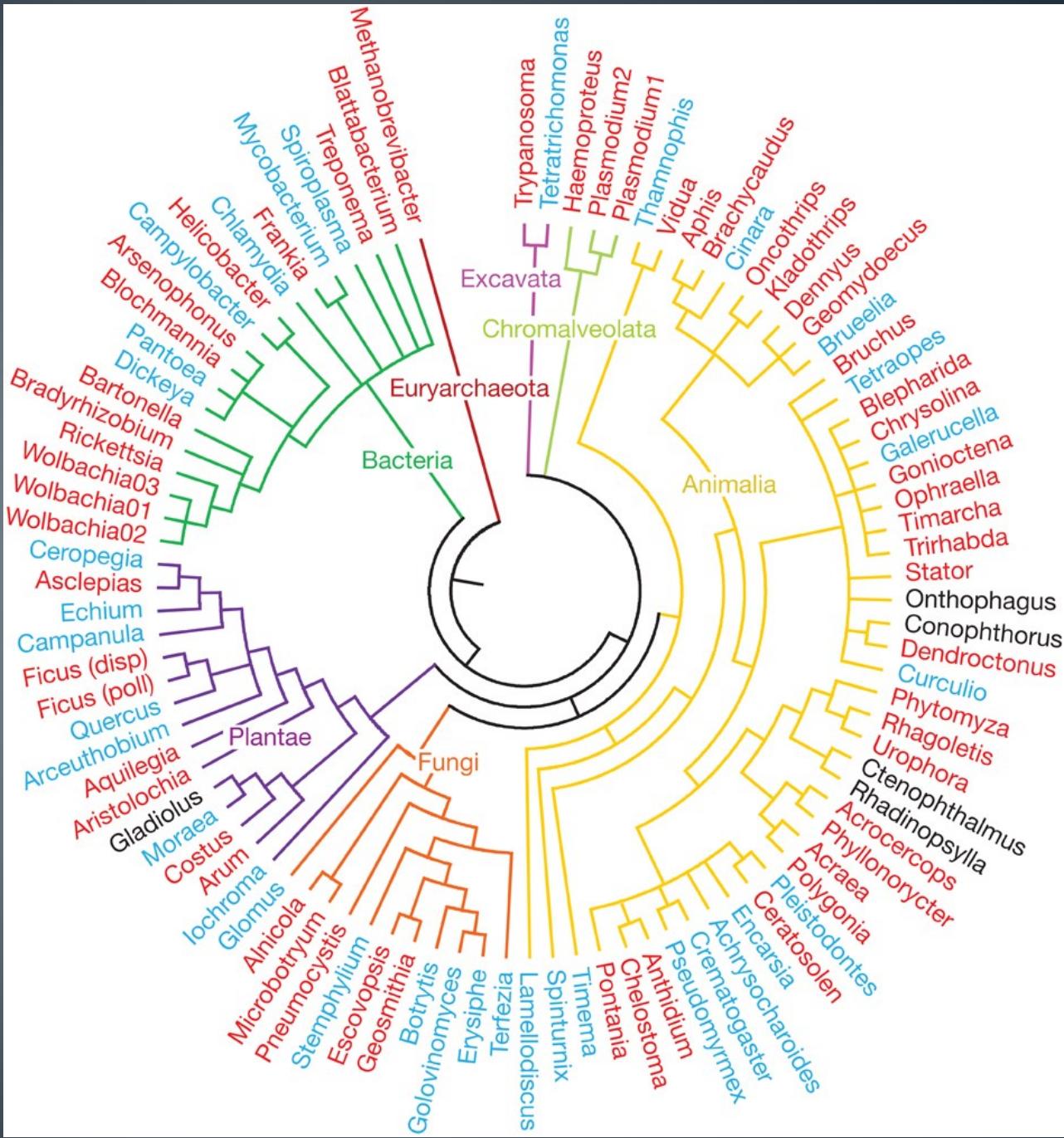
- Nekořenový strom zobrazuje vztahy mezi OTU bez specifikace společného předka.
- Kořenový strom stanovuje společného předka všech OTU.
- Z kořenového stromu lze udělat nekořenový vždy odstraněním společného předka.
- Z nekořenového stromu lze získat kořenový pouze někdy, často potřeba k původním OTU přidat další jedno OTU evolučně vzdálené od ostatních.

# Kořenový vs. nekořenový strom

nezakořeněný x zakořeněný strom







# Topologie kořenového stromu

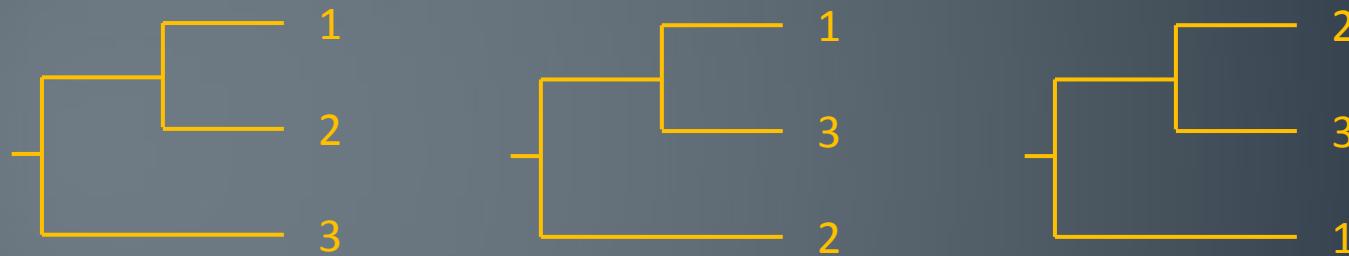
- Počet možných topologií (uspořádání větví) **bifurkovaného** (v uzlu vždy jen dvě větve) **kořenového stromu** je:

$$\frac{(2m-3)!}{2^{m-2}(m-2)!}$$

- kde  $m$  je počet OTU.
- Pro  $m = 4 \rightarrow 15$  topologií.
- Pro  $m = 10 \rightarrow 34\,459\,425$  topologií.

# Topologie kořenového stromu

- $m = 3 \rightarrow 3$  topologie



- Na umístění OTU v rámci jednoho uzlu nezáleží.



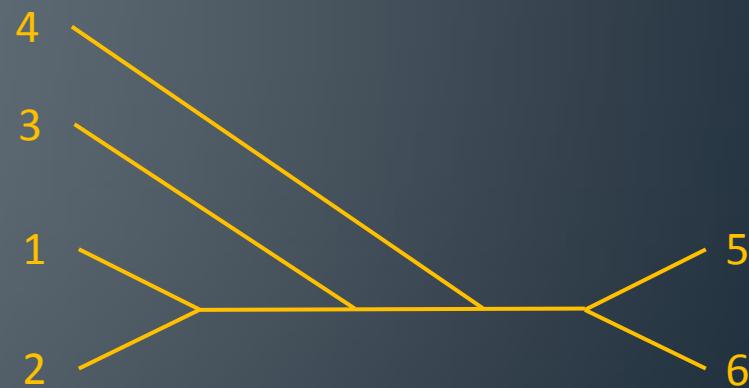
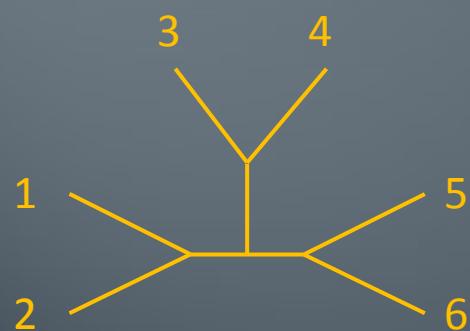
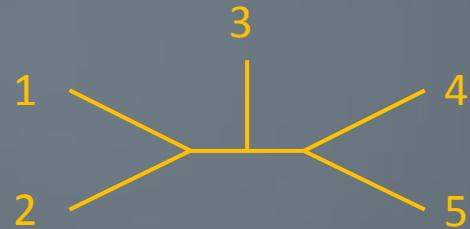
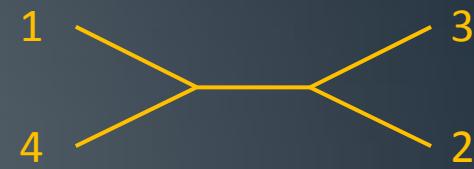
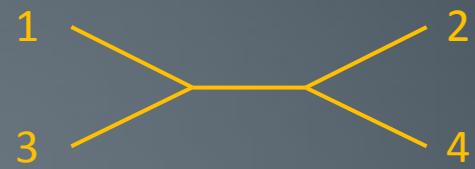
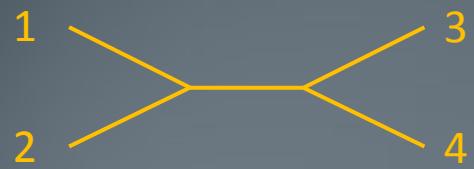
# Topologie nekořenového stromu

- Počet možných topologií **bifurkovaného** nekořenového stromu:

$$\frac{(2m-5)!}{2^{m-3}(m-3)!}$$

- Pro  $m = 4 \rightarrow 3$  topologie.
- Pro  $m = 10 \rightarrow 2\ 027\ 025$  topologií.
- Topologií nekořenového stromu o  $m$  OTU se rovná počtu topologií kořenového stromu o  $m-1$  OTU.

# Topologie nekořenového stromu



# Textový zápis stromů

- Newick formát – jednoduchý závorkový zápis, který je schopna zpracovávat většina softwarů.
- I pro multifurkované stromy (více než 2 větve v jednom uzlu).
- Běžně zápis (zleva) → (doprava) nebo (shora) → (dolů).



$(B,(A,C,E),D);$

ale mají stejný význam i libovolně uspořádané:

$(D,(C,A,E),B);$

$((A,E,C),B,D);$

...

# Newick formát

- Bez délek větví:

$((\text{Býk}, \text{Kůň}), (\text{Pes}, \text{Myš}));$



- S délkami větví:

$((\text{Býk}:0.045, \text{Kůň}:0.07):0.0825, (\text{Pes}:0.09, \text{Myš}:0.1):0.0275);$

- Popis vnitřních uzlů:

$((\text{Býk}:0.045, \text{Kůň}:0.07)\text{X}:0.0825, (\text{Pes}:0.09, \text{Myš}:0.1)\text{Y}:0.0275);$

# Metody konstrukce fylogramu

Dva možné přístupy:

## Znakové metody:

- každá pozice v sekvenci = digitální znak s jasně vymezenými možnými stavami,
- výpočet pravděpodobnosti změny stavu = substituce nukleotidu nebo AMK na konkrétní pozici,
- nutná znalost frekvence možných záměn.

## Distanční metody:

- výpočet celkové vzájemné podobnosti mezi sekvencemi,
- vzdálenost mezi dvěma sekvencemi je dána počtem změn,
- nutné zarovnání všech zkoumaných sekvencí,
- nutná znalost substitučních matic.

# Metody konstrukce fylogramu

- Existují tři základní metody (ze kterých se odvíjí ostatní) založené na principech:

- **minimální evoluce**

Minimalizace součtu délek větví v celém stromu podle vypočtených vzájemných evolučních vzdáleností.

- **maximální parsimonie** (úspory změn)

Vyhledání stromové struktury odpovídající nejmenšímu počtu mutací nutných k dosažení potřebných stavů.

- **vyhledání „nejlepšího možného“ stromu (maximum likelihood)**

Prohledáváme všechny možné konstrukce stromu a vyhledáme ten s největší evoluční pravděpodobností.

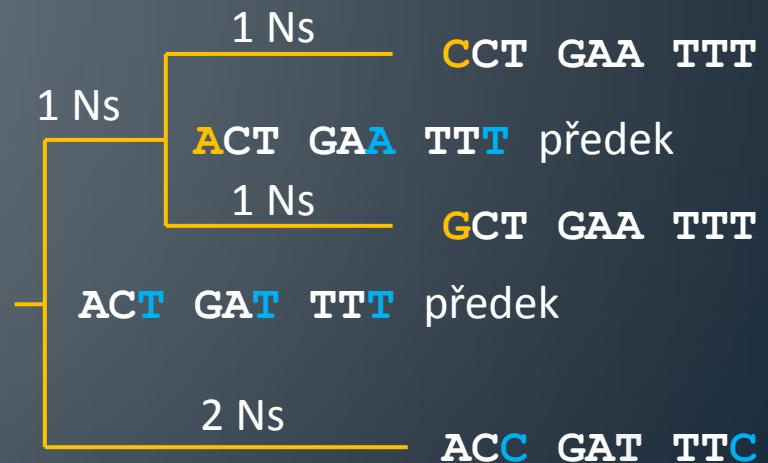
# Metoda maximální parsimonie

- Znaková metoda.
- Parsimonie ~ úspornost.
- Úvaha: evoluce probíhá nejúspornější cestou (**s minimálním počtem změn**).
- Na základě vstupních sekvencí se vytvoří všechny varianty stromů a z nich se vyhledá strom s nejmenším množstvím změn – parametr je **celková délka stromu**.
- Výpočetně náročná metoda.

# Délka stromu

- Počet kroků  $n_s$  (substitucí) nutných k vytvoření stromu.
- Potřeba určit společného předka sekvencí.
- Volíme strom s nejmenší délkou.
- Pokud má více stromů (topologií) stejnou délku:
  - zohledňujeme průběžný počet kroků,
  - a vytváříme konsenzuální strom.

Délka stromu = 5 substitucí

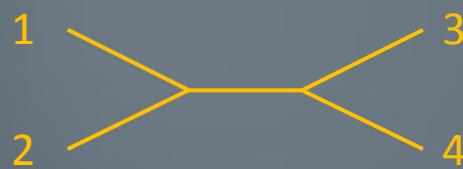


# Kroky maximální parsimonie

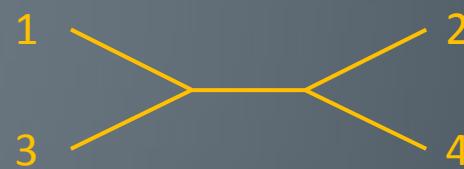
- Vytvoření všech možných topologií stromu pro daný počet OTU.
- Výpočet délky stromu pro každou možnou topologii.
- Strom s nejmenší délkou je výsledný.

# Všechny topologie

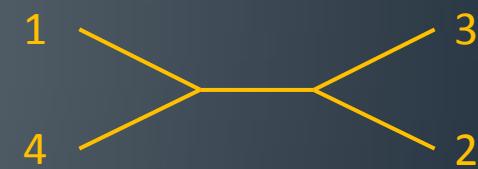
- Sekvence AACGTA, AATGTA, TATCAA, TATCGA.
- Máme 4 OTU → 3 různé topologie nekořenového stromu:



Topologie 1

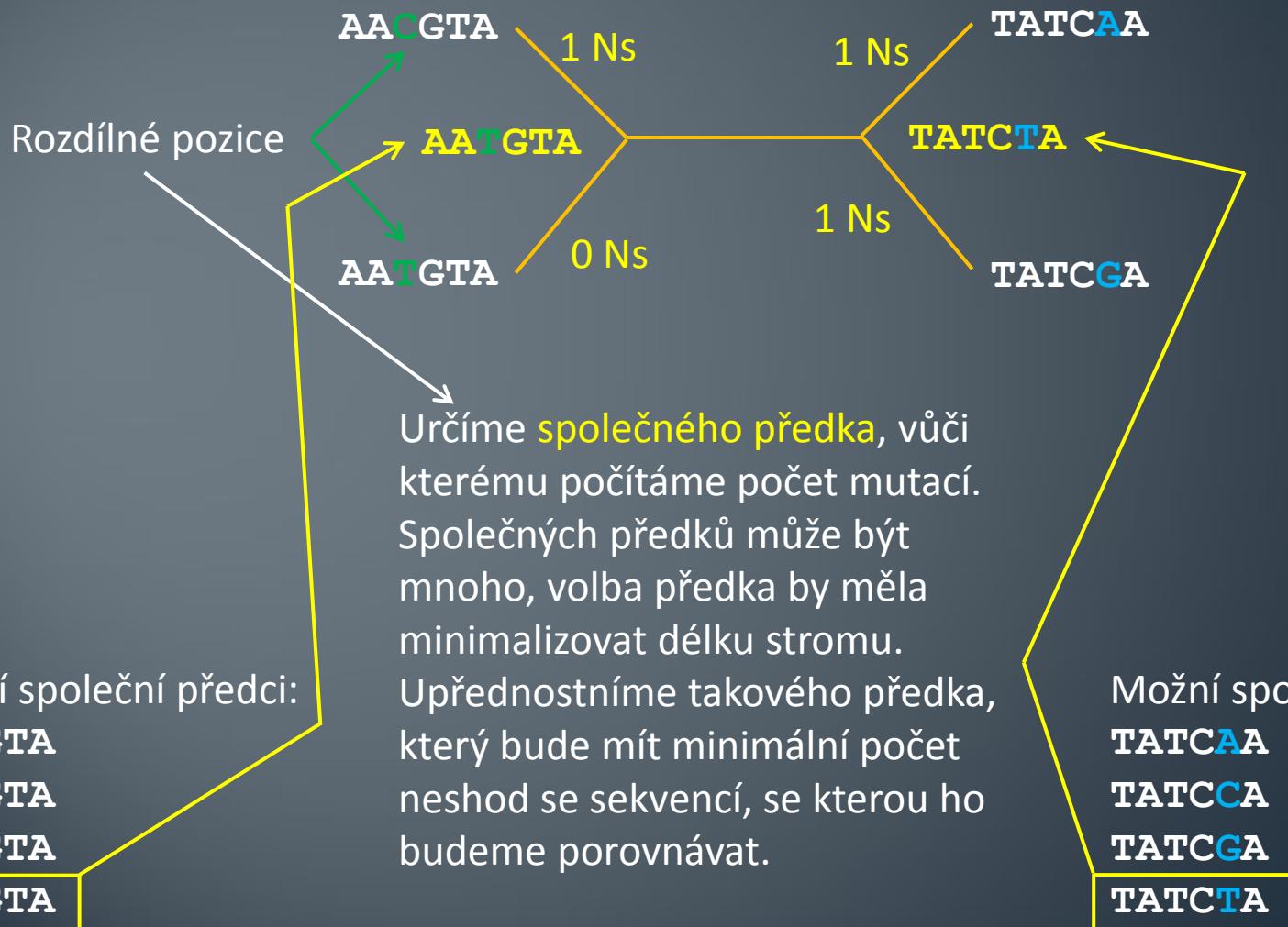


Topologie 2

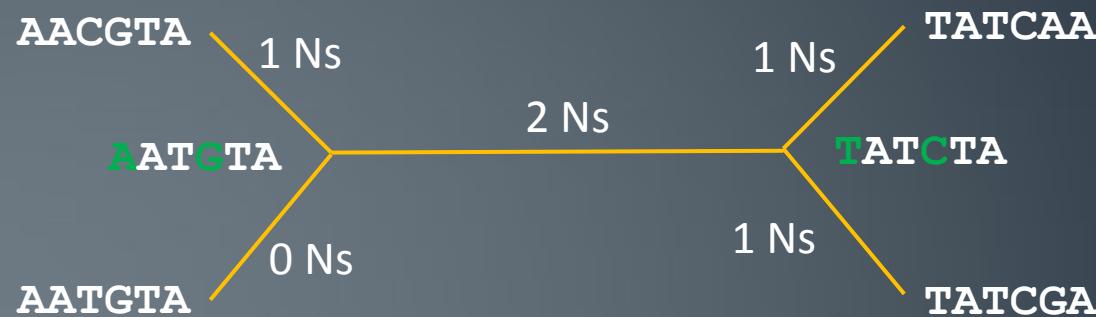


Topologie 3

# Topologie 1



# Topologie 1



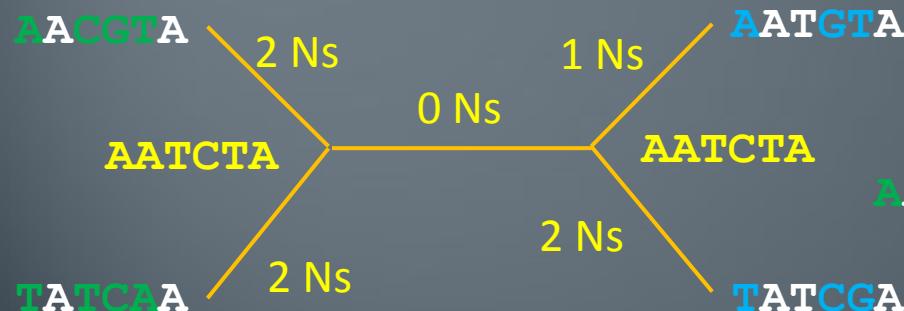
Máme porovnané všechny OTU,  
nyní porovnáváme určené předky  
v uzlech.

Oba předci se od sebe liší ve dvou  
pozicích = 2 Ns.

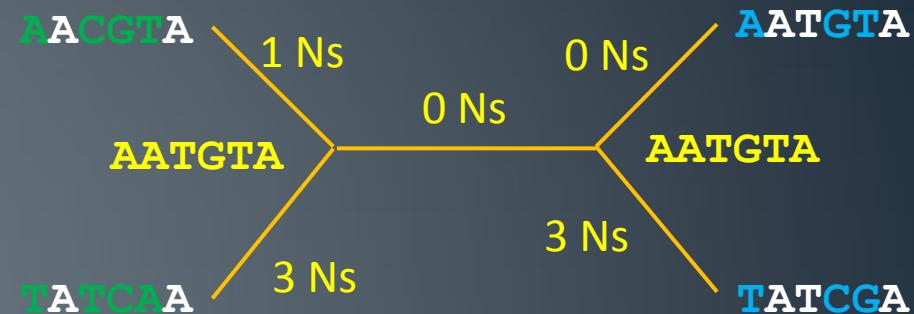
Celková délka stromu = 5 Ns

# Topologie 2

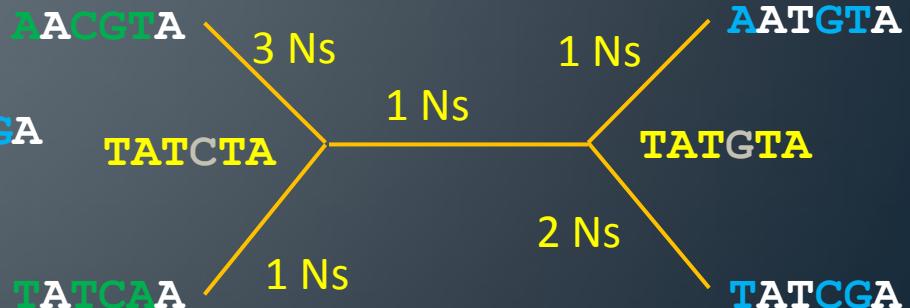
Velké množství potenciálních předků.  
Prakticky se projíždí všechny možné  
kombinace a hledá se ten, který  
minimalizuje délku stromu.



Délka stromu = 7 Ns



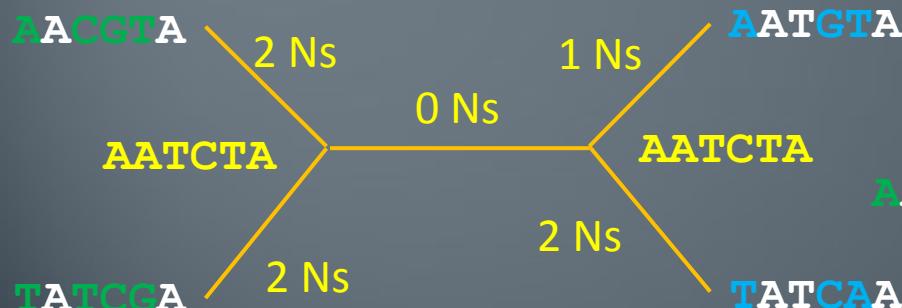
Délka stromu = 7 Ns



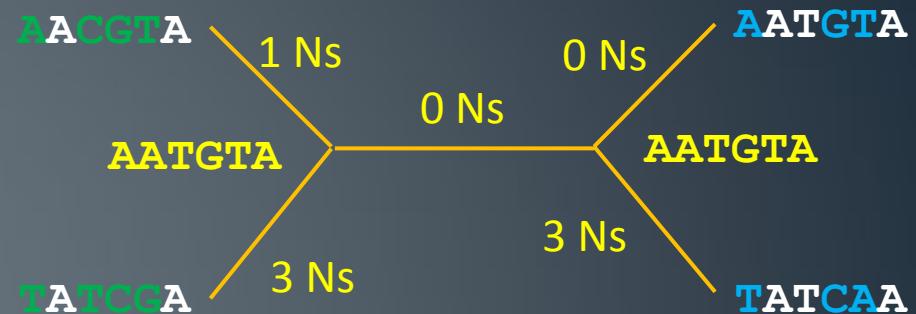
Délka stromu = 8 Ns

# Topologie 3

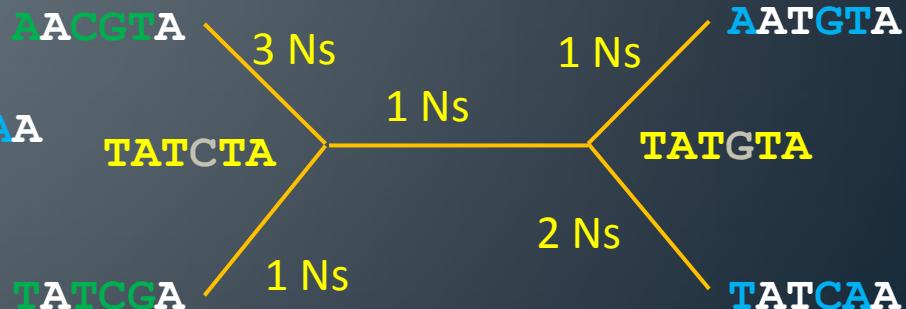
Velké množství potenciálních předků.  
Prakticky se projíždí všechny možné kombinace a hledá se ten, který minimalizuje délku stromu.



Délka stromu = 7 Ns



Délka stromu = 7 Ns



Délka stromu = 8 Ns

# Výsledný strom

Nejmenší délku stromu (5 Ns) měla topologie 1.



# Distanční metody

- Vytváří se **distanční matice** ze vzdáleností ( $p$ -distance) všech dvojic sekvencí.
- Počet prvků  $n$  matice je dán kombinací druhé třídy ( $k = 2$ , porovnáváme vždy 2 sekvence) bez opakování z  $N$  sekvencí:

$$n = \frac{N!}{k!(N-k)!}$$

- Pro  $N = 4$

$$n = \frac{4!}{2!(4-2)!} = 6$$

	Tur	Kůň	Myš	Pes
Tur	*			
Kůň	0,09	*		
Myš	0,27	0,26	*	
Pes	0,25	0,24	0,2	*

# Proporcionální vzdálenost

- *p-distance* je počet mutací v sekvenci proporcionálně ku délce sekvence.

$p = n_p/n$       kde:       $n_p$  .... počet mutovaných pozic dvou sekvencí  
                                 $n$  .... počet všech pozic

Př.:

MDPNCSAKAGV-CTCASSCKCKECKCTSCKKSCCS CCP-GCAKCAQGCICKGASEKCSCCA  
MD--CSCAKAGVSCTCASSCKCKDCKCTSCKASCC-SCPVGCAKCAQGCSCKGASEKCSCCA  
... | | ..... | ..... | ..... | ... | | ..| ..... | ..... | .....

$$n_p = 9$$

$$n = 62$$

$$p = 9/62 = 0,1452$$

# UPGMA

- Unweighter Pair Group Method with Arithmetic mean.
- Hierarchická shlukovací metoda z roku 1977.
- Pracuje s distanční maticí.
- Spojuje OTU na základě minimální vzdálenosti v distanční matici.
- Všechny párové vzdálenosti mají stejný vliv na tvorbu stromu.

# Kroky UPGMA

- Vyhledání minima v distanční matici.
- Výpočet délky větví pro dvojici sekvencí s minimální distancí.
- Přepočet matice distancí vůči spojeným sekvencím.
- Vyhledání minima v distanční matici.
- ...
- Opakujeme dokud nemáme ve stromu spojené všechny OTU.
- Délka větví vycházejících z jedno uzlu je shodná.

# Vyhledání minima – spojení OTU

	Tur	Kůň	Myš	Pes
Tur	*			
Kůň	<b>0,09</b>	*		
Myš	0,27	0,26	*	
Pes	0,25	0,24	0,2	*

Nalezené minimum

Délka sumy větví mezi oběma OTU = 0,09

Délka jedné větve od OTU do uzlu =  $0,09/2$

Přepočet matice průměrováním vůči  
spojeným OTU.



# Přepočet matice

Originální matice

	Tur	Kůň	Myš	Pes
Tur	*			
Kůň	0,09	*		
Myš	0,27	0,26	*	
Pes	0,25	0,24	0,2	*

Přepočtená matice

	Tur - Kůň	Myš	Pes
Tur-Kůň	*		
Myš	0,265	*	
Pes	0,245	0,2	*

$$(\text{Tur:Myš} + \text{Kůň:Myš})/2 = \text{Tur-Kůň:Myš}$$

$$(0,27+0,26)/2 = 0,265$$

$$(\text{Tur:Pes} + \text{Kůň:Pes})/2 = \text{Tur-Kůň:Pes}$$

$$(0,25+0,24)/2 = 0,245$$

Ostatní hodnoty zůstávají.

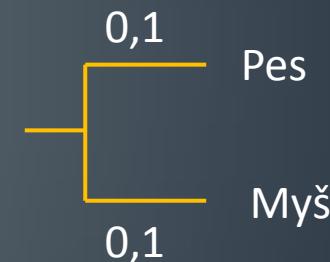
# Vyhledání minima – spojení OTU

	Býk - Kůň	Myš	Pes
Býk-Kůň	*		
Myš	0,265	*	
Pes	0,245	0,2	*

Nalezené minimum

Délka sumy větví mezi oběma OTU = 0,2

Délka jedné větve od OTU do uzlu = 0,2/2



Přepočet matice průměrováním vůči  
spojeným OTU.

# Přepočet matice

Originální matice

	Býk	Kůň	Myš	Pes
Býk	*			
Kůň	0,09	*		
Myš	0,27	0,26	*	
Pes	0,25	0,24	0,2	*

Přepočtená matice

	Býk-Kůň	Myš-Pes
Býk-Kůň	*	
Myš-Pes	0,255	*

Přepočet vždy z originální matice!

$$(Býk:Pes + Kůň:Pes + Býk:Myš + Kůň:Myš)/4 = Býk-Kůň:Pes-Myš$$
$$(0,27+0,26+0,25+0,24)/4 = 0,255$$

# Vyhledání minima – spojení OTU

	Býk-Kůň	Myš-Pes
Býk-Kůň	*	
Myš-Pes	0,255	*

Nalezené minimum

Délka sumy větví mezi oběma OTU = 0,2

Délka jedné větve od OTU do uzlu = 0,2/2

$$0,045 + X + Y + 0,1 = 0,255$$

$$0,045 + X = 0,255/2$$

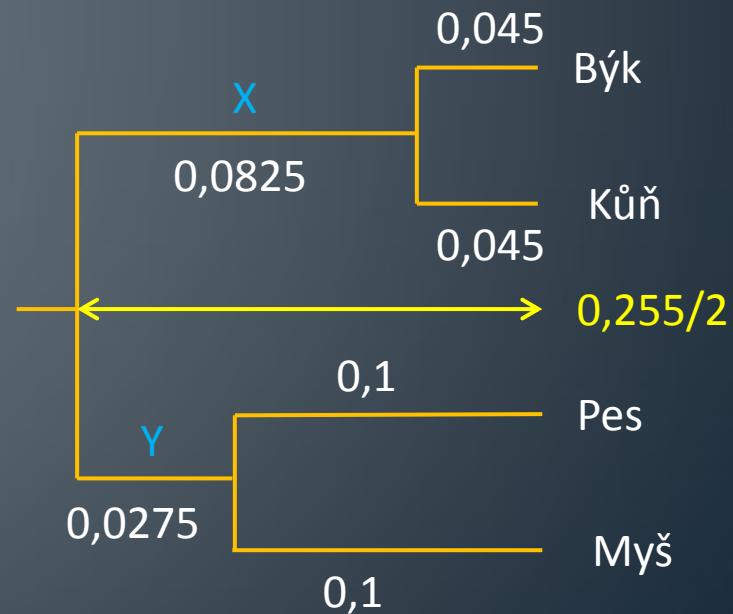
$$X = 0,255/2 - 0,045$$

$$X = 0,0825$$

$$0,1 + Y = 0,255/2$$

$$Y = 0,255/2 - 0,1$$

$$Y = 0,0275$$



# Přepočet z originální matice

	A	B	C	D	E	F	G
A	*						
B	19	*					
C	27	31	*				
D	8	18	26	*			
E	33	36	41	31	*		
F	18	1	32	17	35	*	
G	13	13	29	14	28	12	*

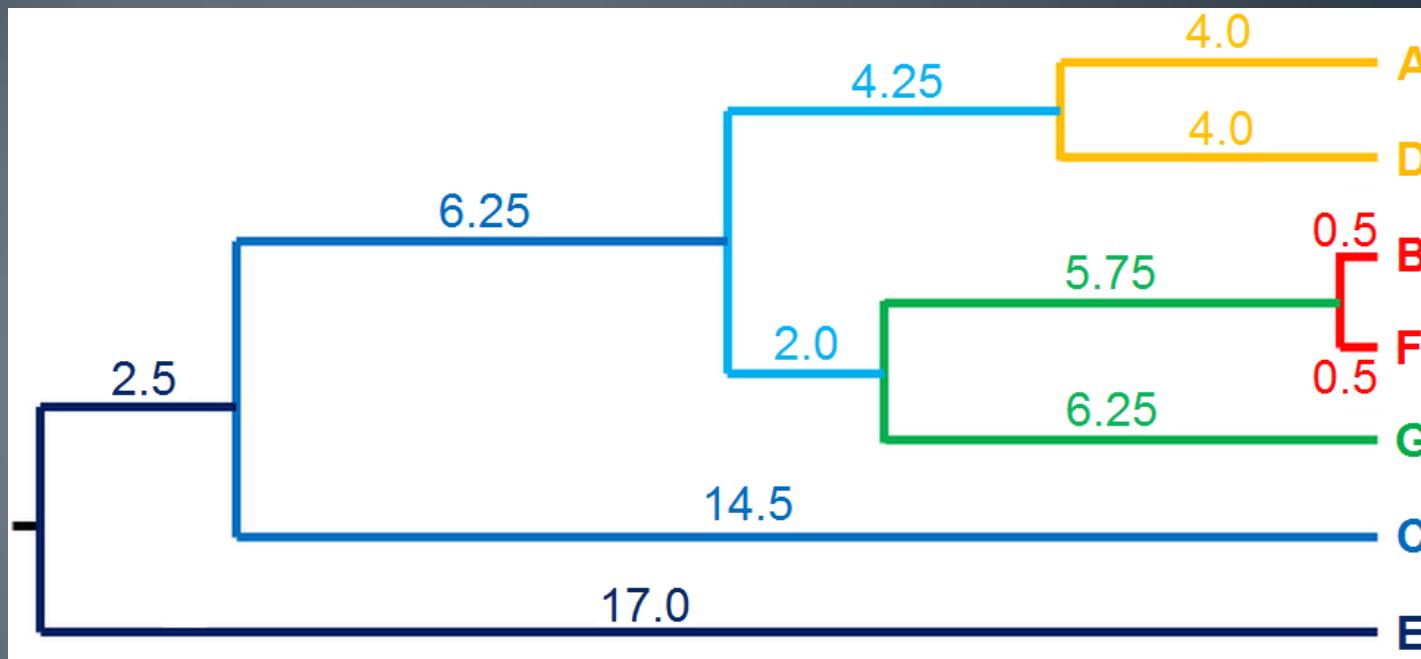
AD vůči BFG:

$$(AB+AF+AG+DB+DF+DG)/6 = \\ = (19+18+13+18+17+14)/6 = \mathbf{16,5}$$

	AD	BF	C	E	G
AD	*				
BF	18	*			
C	26,5	31,5	*		
E	32	35,5	41	*	
G	13,5	12,5	29	28	*

[AD] vůči [BF]G:

$$([AD][BF]+[AD]G)/2 = \\ = (18+13,5)/2 = \mathbf{15,75}$$



((X:0.3,Y:0.25):0.1,Z:0.2);

# Porovnání konstrukčních metod

## Distanční metody

- + Úsporný výpočetní algoritmus
- + Aplikace i na velké množství dlouhých sekvencí
- + Použití matematických modelů
- Minimální evoluce nemusí odpovídat minimální délce větví
- Nutné mnohočetné zarovnání
- Aproximace proporcionální vzdálenost  $\sim$  evoluční vzdálenost

## Znakové metody

- + Zpracovává fylogeneticky informativní mutace
- + Porovnává každou pozici zvlášť – bodové mutace
- + Evolučně nejvhodnější (nejúspornější řešení)
- Více možných řešení
- Velmi náročný algoritmus
- Nepoužitelný na velké množství dlouhých sekvencí

# Výsledek fylogenetické analýzy

- Rekonstrukční metody **vždy sestrojí fylogram** z jakýchkoli dat.
- Různé metody mohou dát **různé topologie**.
- Stromy lze hodnotit pomocí **statistických testů** pro:
  - ověření správnosti topologie,
  - srovnání dvou topologií,
  - ověření robustnosti stromu.

# Robustnost stromu

- Topologie stromu (uzlu) je dostatečně **stabilní** – nezmění se po odstranění jedné nebo více OTU.
- Topologie stromu (ne délka větví) závisí na celém objemu dat, ne jen na konkrétní pozici
  - šum, chyby v sekvenování, neviditelné mutace, apod.

# The Tree of Life

