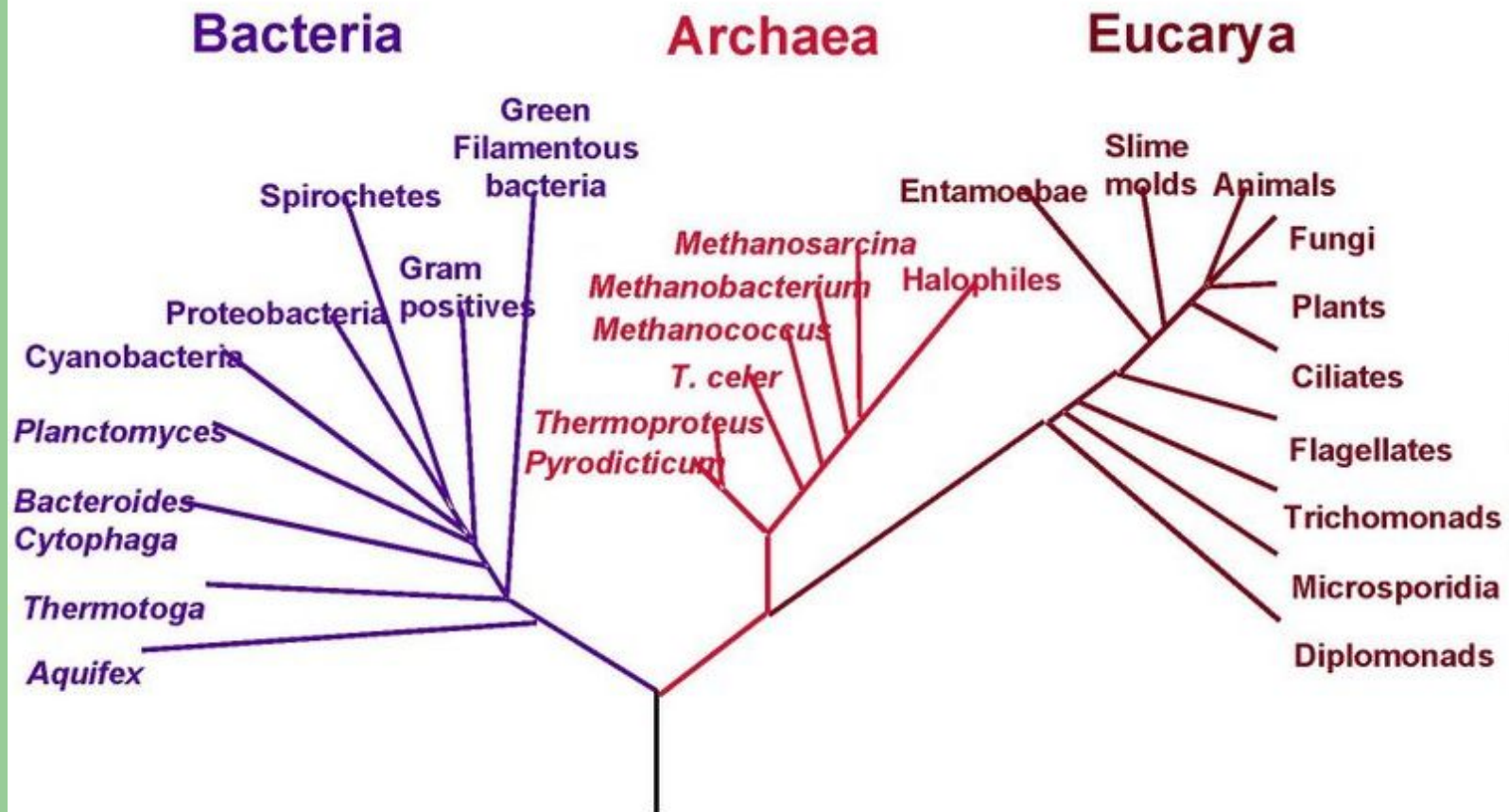




Filogenetiniai medžiai

Kas yra filogenetiniai medžiai?

Phylogenetic Tree of Life



Filogenija

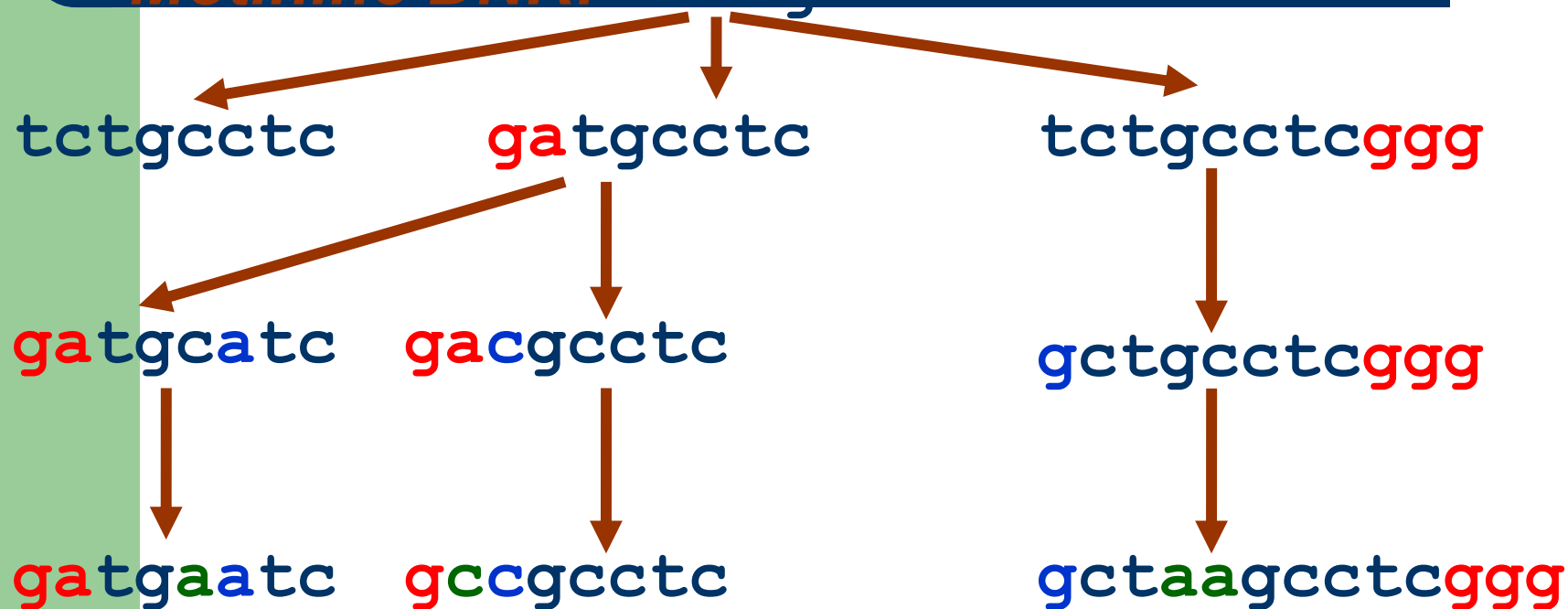
- Rodo tėvinius/palikuoninius sąryšius tarp genų ar organizmų
- Sąryšiai atrandami remiantis genotipais, o ne fenotipais

O kam to reikia?

- Siekiant suprasti organizmų istoriją
- Daugybiniai palyginiai (clustalw algoritmas)
- Genų funkcijos spėjimai
- ...suvaldyti infekcijoms.

Mutācijas - bioloģiskās evolūcijas esmē

Motīnā DNR:

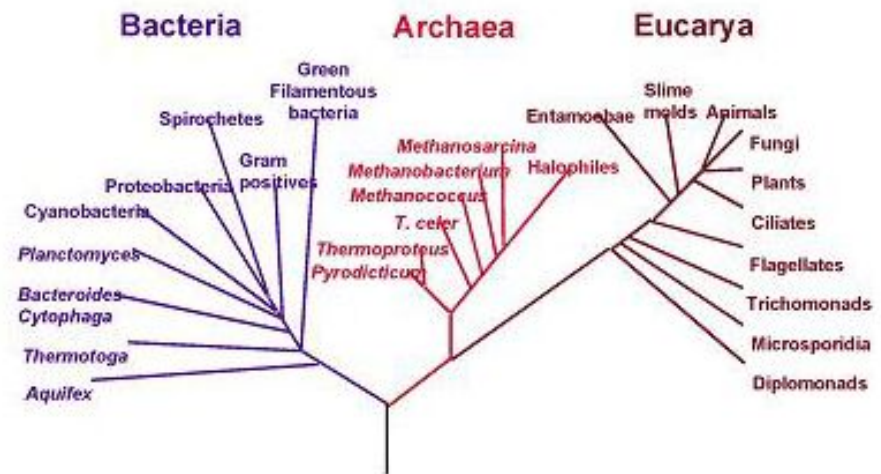


Dabartinās rūšys

Filogenetiniai medžiai

- Mazgas = taksonominis vienetas
 - Lapas = genas arba organizmas
 - Vidinis mazgas = numanomas protėvis
 - Dvišakis = dvi linijos
 - Daugiašakis = daugiau nei dvi linijos
- Šakos = kilmės sąryšis

Phylogenetic Tree of Life

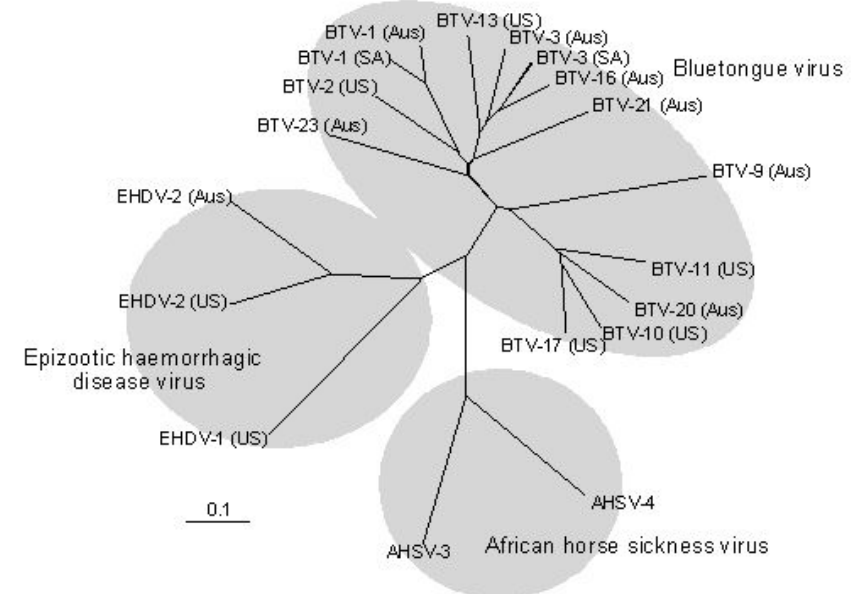
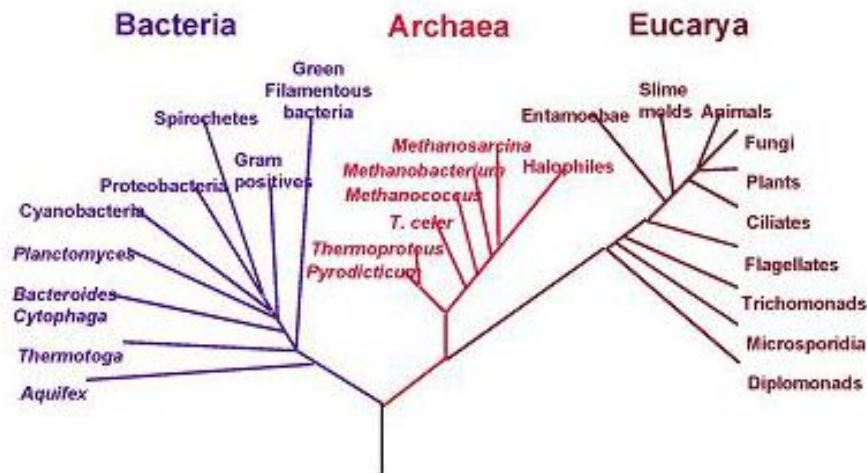


Filogenetiniai medžiai

- Turintys šakninį mazgą (rooted) = mazgas yra bendras protėvis
- Be šakninio mazgo (unrooted) = nesuteikia informacijos apie evoliucijos kryptį

Viruses of the family Reoviridae

Phylogenetic Tree of Life



Filogenetinių medžių tipai

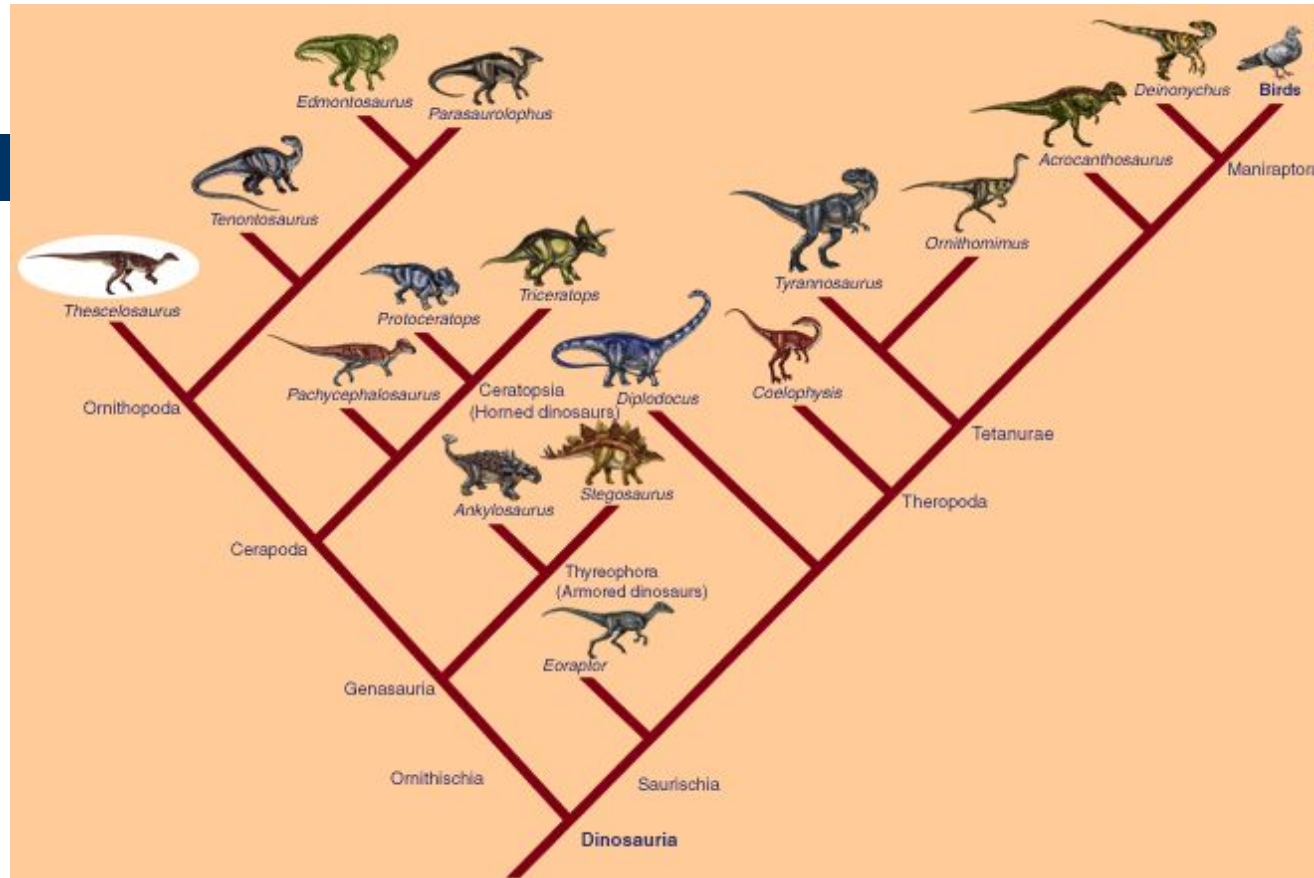
➤ Kladogramos

- Pavaizduoja rūšių evoliucinį sąryšį neatsižvelgiant į jų evoliucinį atstumą.

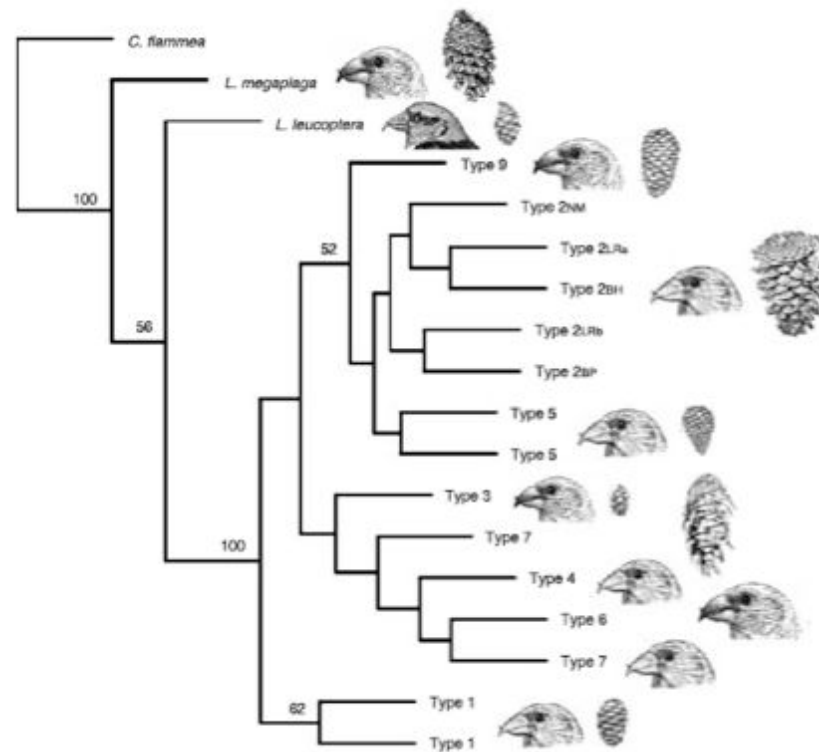
➤ Filogramos

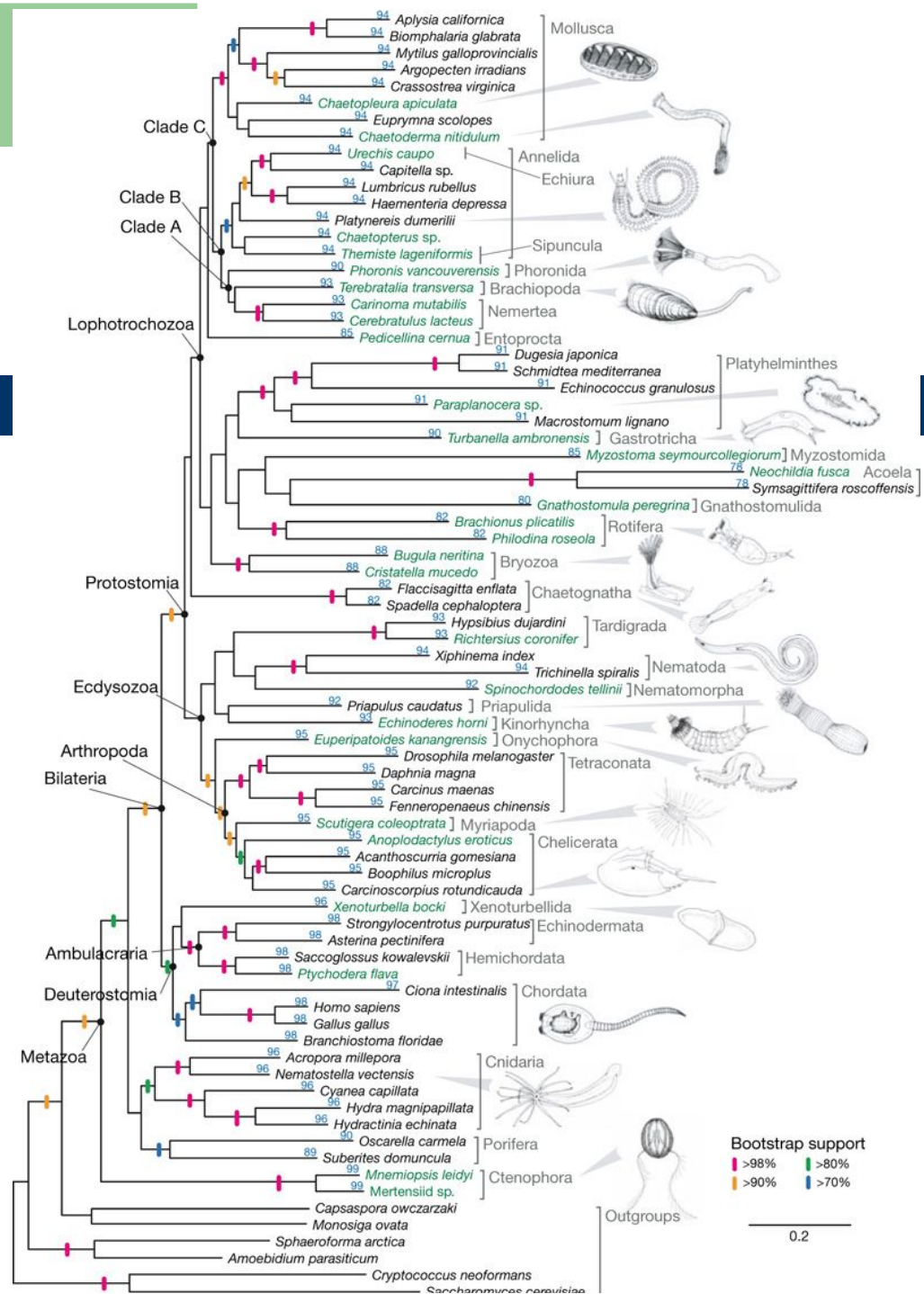
- Pavaizduoja rūšių evoliucinius sąryšius bei tikėtiną evoliucinį atstumą (remiantis atstumu tarp jų)

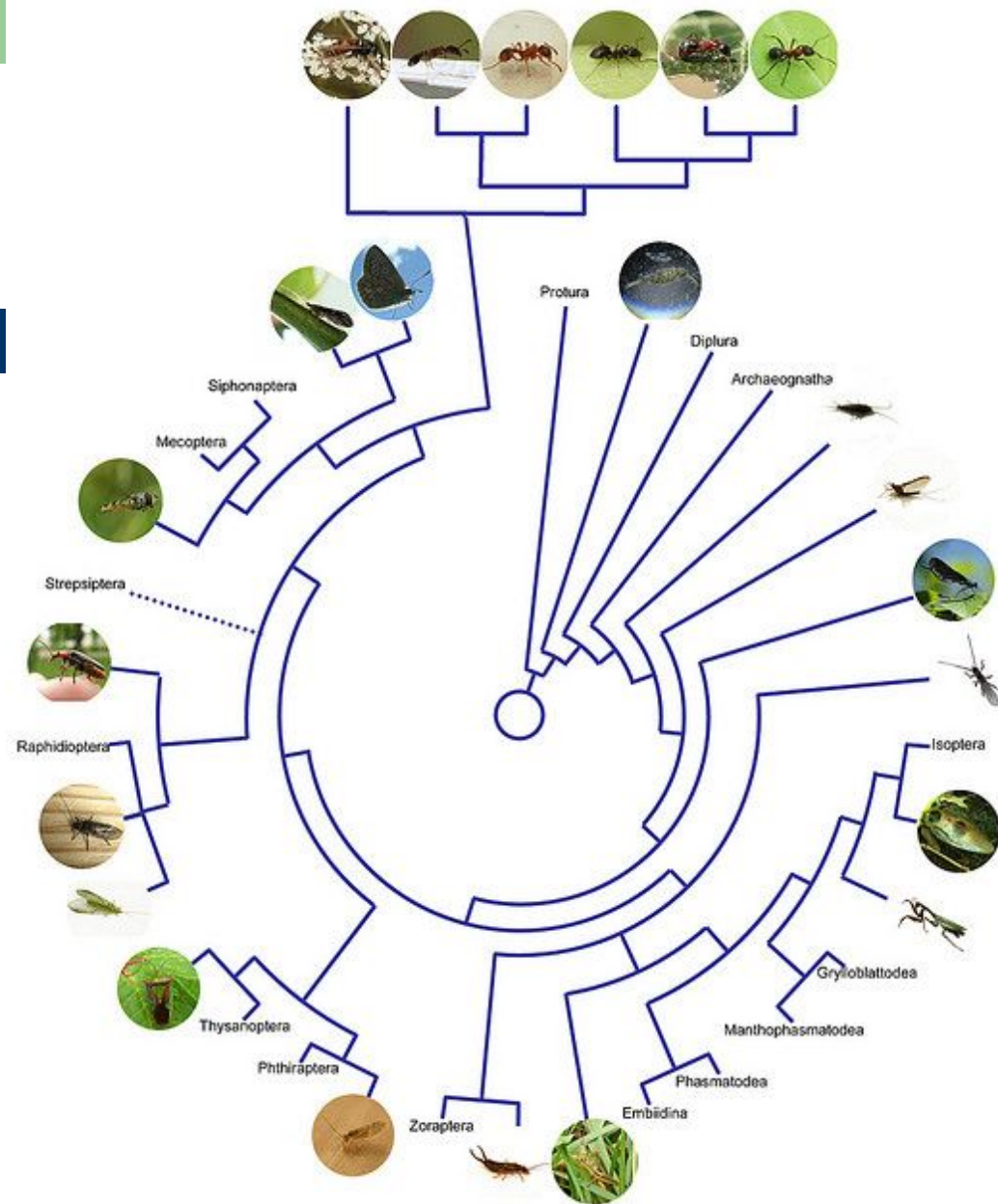
Kladograma

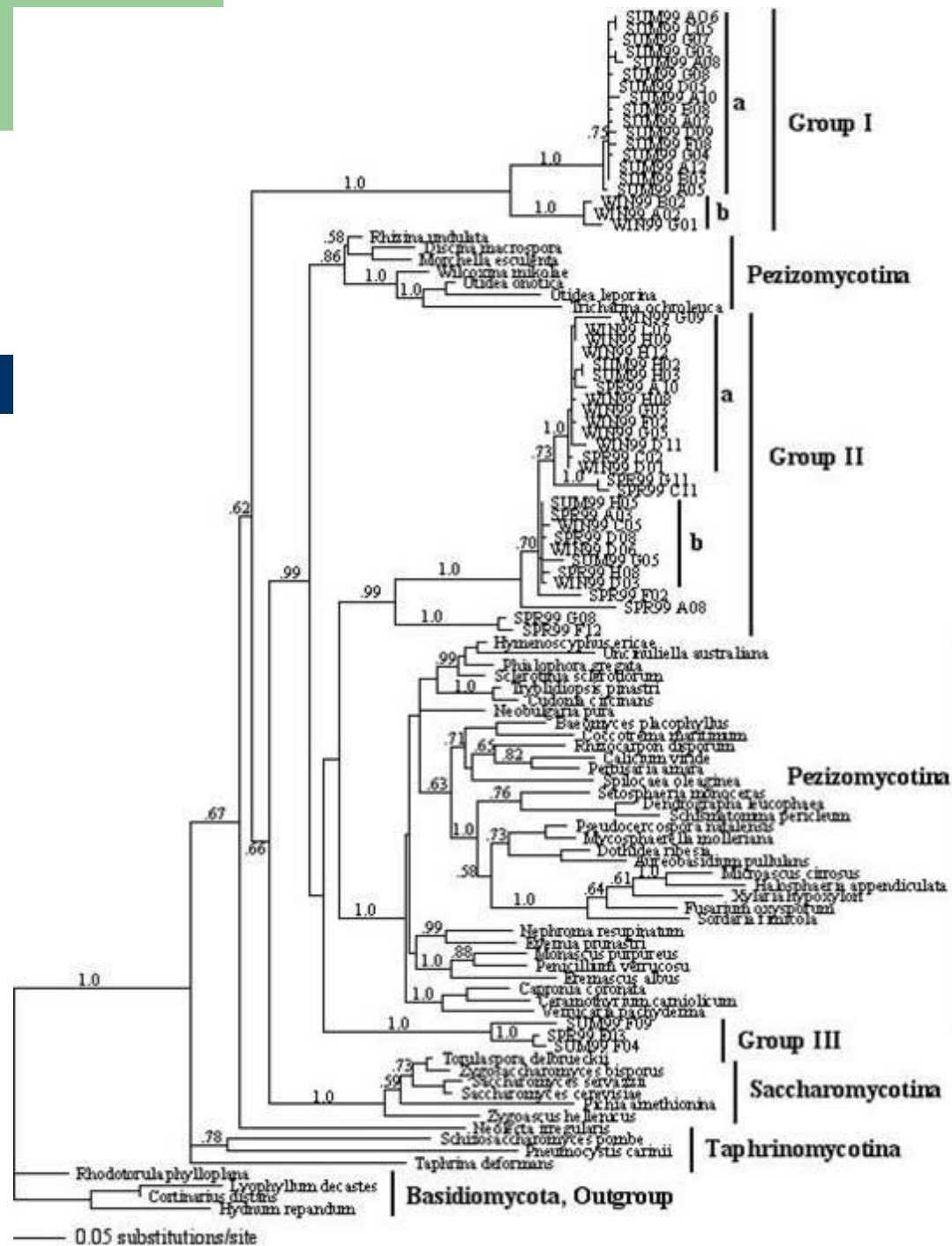


Filograma





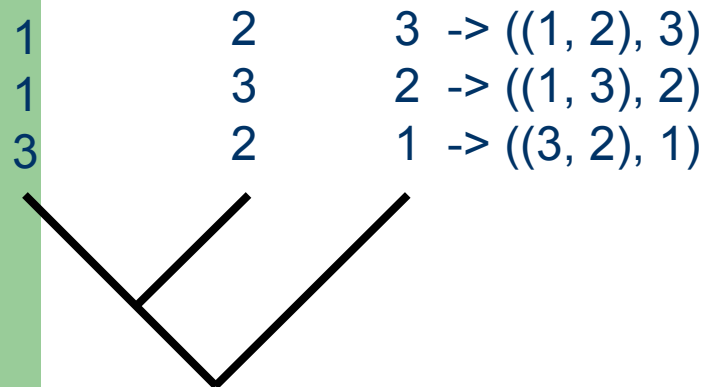




Filogenetiniai medžiai

- n = Mazgų skaičiai
- Rooted $\Rightarrow NR = (2n-3)!/2^{n-2}(n-2)!$
- Unrooted $\Rightarrow NU = (2n-5)!/2^{n-3}(n-3)!$

Newick'o formatas



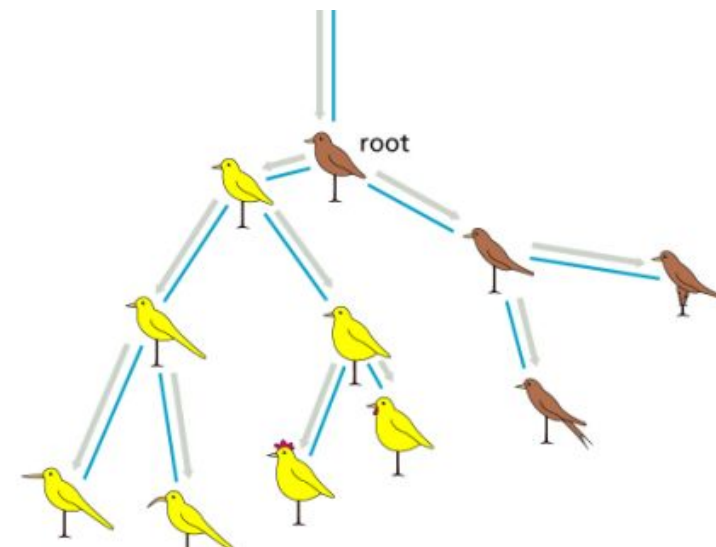
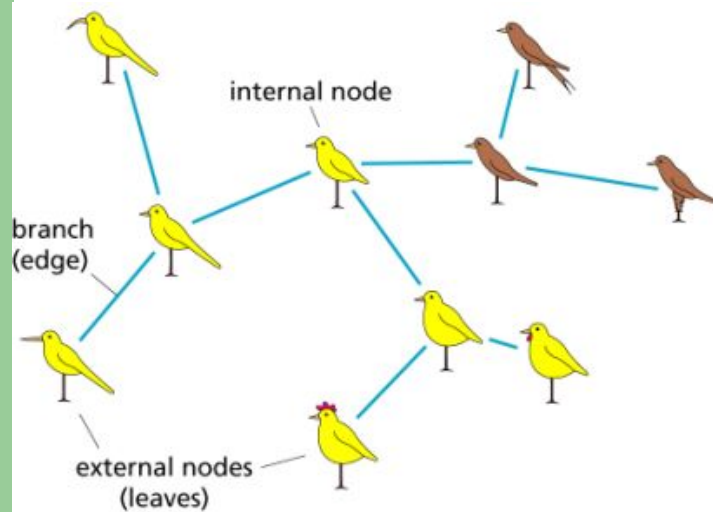
n	NR	NU
2	1	1
3	3	1
5	105	15
10	34×10^6	2×10^6
15	213×10^{12}	7×10^{12}
20	8×10^{21}	0.2×10^{21}

Filogenija

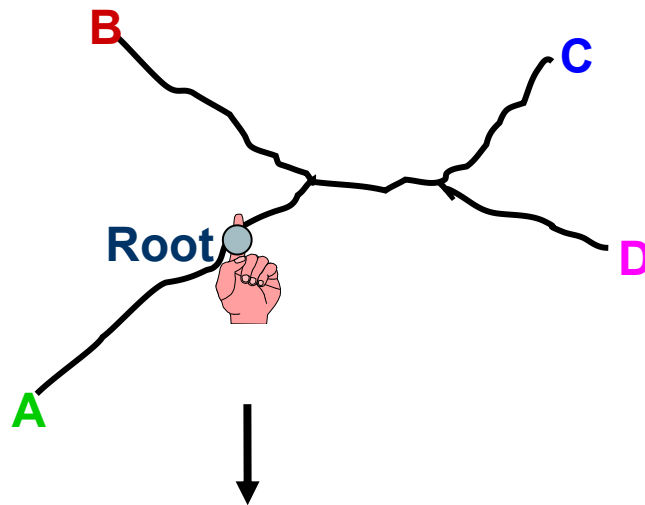
- Šakniniai ir bešakniai medžiai:
 - Dauguma filogenetinių metodų sukuria bešaknius medžius, nes jie remiasi skirtumais tarp sekų, bet negali nustatyti, kurie sekų pokyčiai atsirado anksčiausiai.

Filogenija

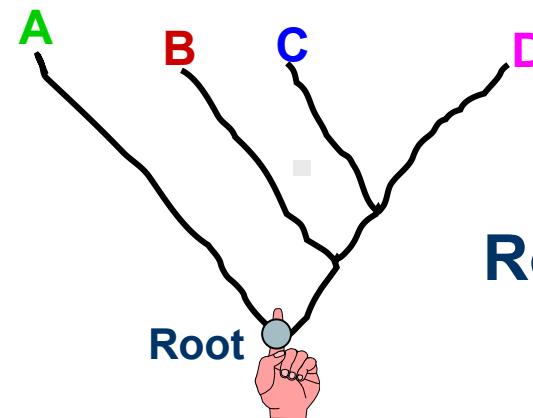
- Šakniniai ir bešakniai medžiai



Šakniniai ir bešakniai medžiai



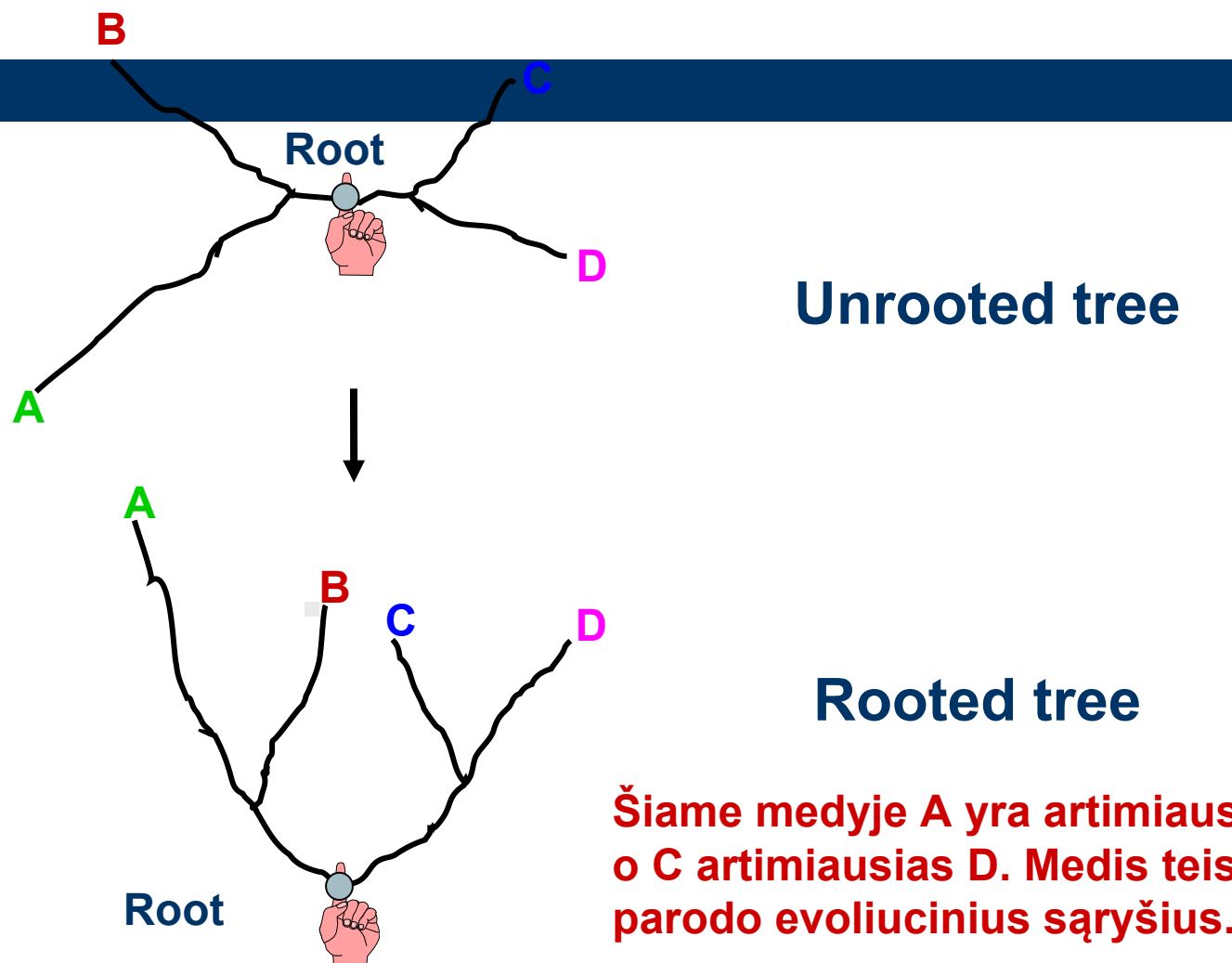
Unrooted tree



Rooted tree

Reikia atkreipti dėmesį, kad A
nebėra artimiausias B

Šakniniai ir bešakniai medžiai

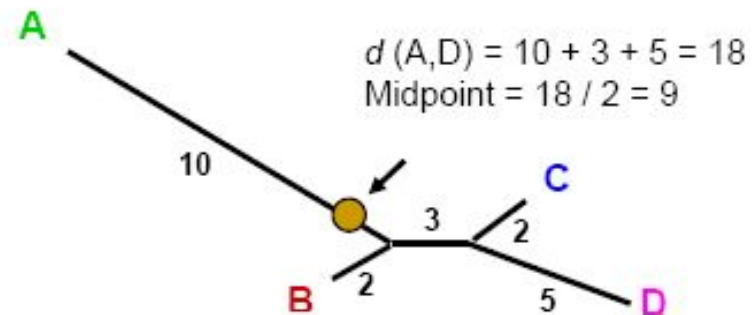
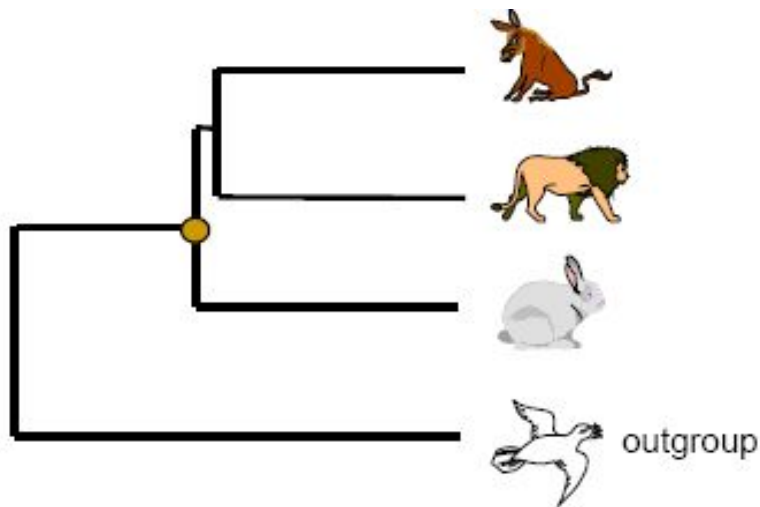


Filogenija

- Šakniniai ir bešakniai medžiai:
 - Du būdai priskirti šaknį bešakniam medžiui:
 - **Šalutinės grupės metodas** : įtraukti prie analizuojamų sekų seką, kuri kilo iš *a priori* žinomo organizmo(ų), kuris būtų artimas bendram protėviui, seką. Pvz. analizuojant beždžionių ir žmonių filogenezę kartu analizuoti kirstuko ir kurmio sekas.
 - **Laikytis molekulinio laikrodžio hipotezės**: laikyti, kad visų analizuojamų rūšių sekos tolo nuo bendro protėvio sekų (kaupėsi mutacijos) vienodu greičiu. Tokiu atveju šaknis turi būti ties viduriu kelio jungiančio tarp labiausiai nutolusių sekų filogenetiniame medyje. Keliai nuo šaknies iki visų medžio lapų turi būti vienodo ilgio.

Filogenija

- Šakniniai ir bešakniai medžiai:
 - Du būdai priskirti šaknį bešakniam medžiui:
Šalutinės grupės metodas Molekulinis laikrodis



Filogenija

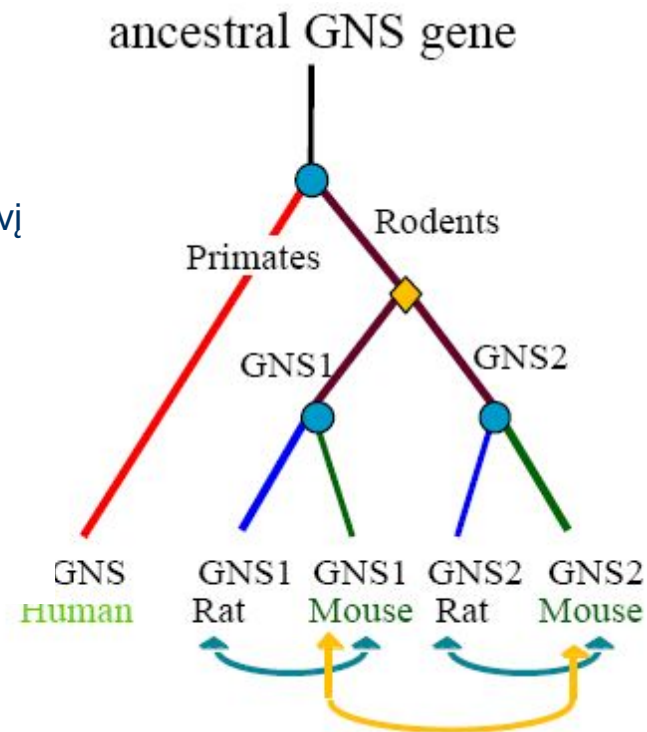
- Ortologija/ Paralogija:



Homologija: du genai turi bendrą protėvį

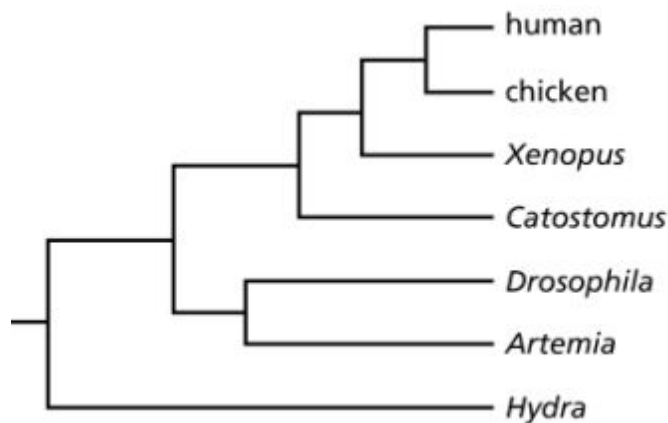
↔ Ortologija: genai yra ortologiniai jei atsirado susidarant rūšims.

↔ Paralogija: genai yra paraloginiai jei atsirado vykstant duplikacijai rūšies viduje

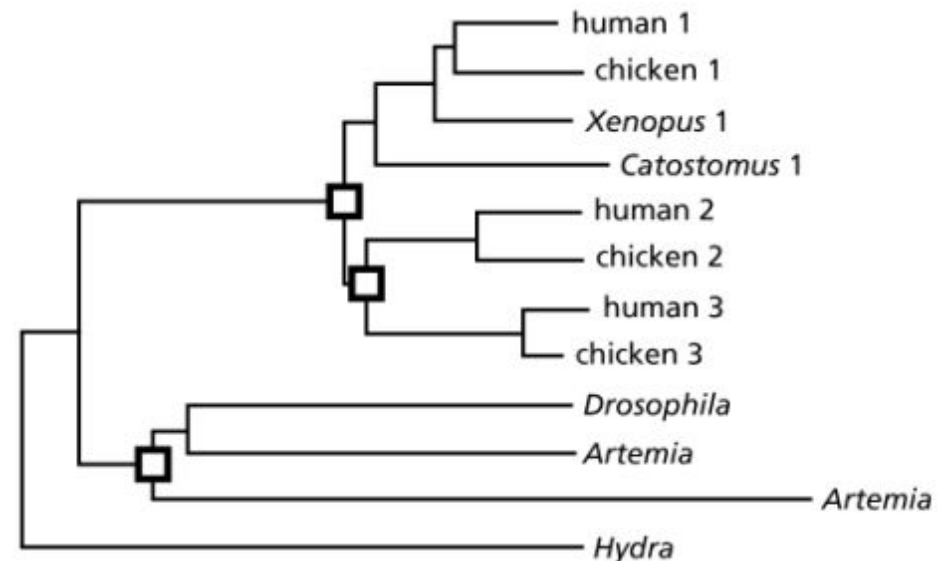


Filogenija

- Rūšių medis ir genų medis:



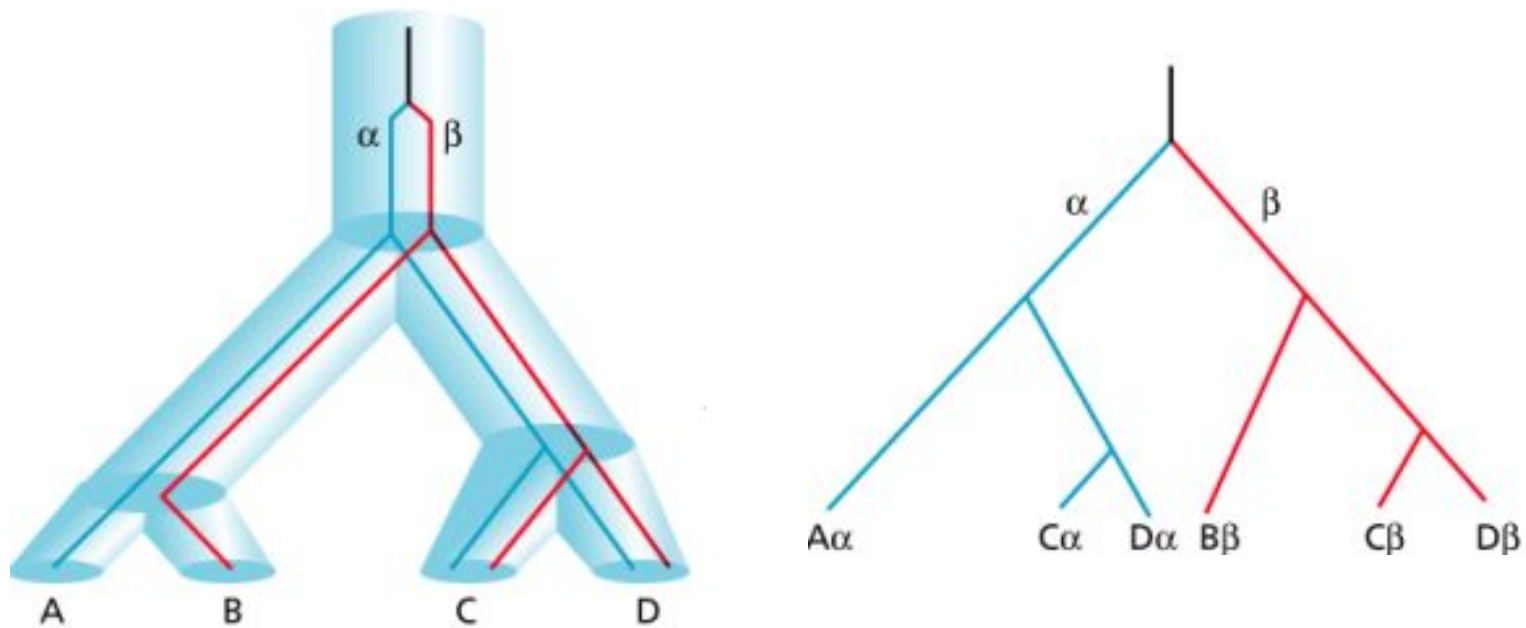
Keletos stuburinių medis



Na⁺-K⁺ jonų pompos genų medis.
Kvadratai žymi duplikaciją

Filogenija

- Rūšių medis ir genų medis:



A,B,C,D - rūšys. α, β - du variantai geno po duplikacijos. A,B rūšys turi tik vieną variantą, C,D - po abu geno variantus.
Primatų spalvinė rega...pavyzdys.

SARS: epidemija

SARS: epidemija

- * Vasario 28, 2003, Hanoi, Vietnamo prancūzų ligininė informavo PSO apie į gripą panašią infekciją.
- .
- * Dr. Carlo Urbani (PSO ekspertas) atvyko ir nustatė, kad tai buvo naujas ir neįprasta patogenas.
- * Kitas kelias dienas Dr Urbani rinko mėginius, dirbo dokumentavo išvadas ir organizavo pacientų karantiną.
- * Karščiavimas, sausas kosulys, dusulys, palaipsniui blogėjantis kvėpavimo nepakankamumas, mirtis dėl kvėpavimo nepakankamumo.

SARS: epidemija

- * Dr Carlo Urbani buvo pirmasis nustatęs SARS (sunkus ūmus respiracinis sindromas).
- * Per tris savaites Dr. Urbani ir penki kiti kitų sveikatos priežiūros specialistai ligoninėje mirė nuo SARS.
- * Iki kovo 15, 2003, PSO paskelbė pasaulinį perspėjimą, skelbdamas SARS grėsmę.

SARS: epidemija



Dr. Carlo Urbani (1956-2003)
PSO. Mirė gelbėdamas kitus.

**Hanoi, Vietnamo Prancūzų ligoninė vietnamo
prancūzų ligoninė, Kovas 2003**



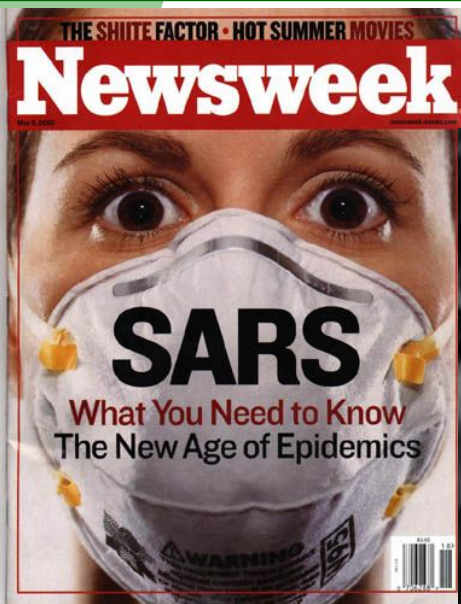
SARS epidemijos kilmė

- Ankstyviausi atvejai ką dabar vadinama SARS įvyko Lap. 2002 Guangdžou (Kinija) Guangdžou ligoninėje išplito 106 naujų atvejų .
- Gydytojas iš šios ligoninės aplankė Honkongą, Vas 21, 2003, ir gyveno 9 aukšte Metropole viešbutyje.
- Gydytojas susirgo ir mirė (diagnozuotas plaučių uždegimas)
- Daugelis lankytojų 9 aukšte Metropole Hotel tapo ligos platintojais.

Pasaulinės SARS epidemijos kilmė

- * Vienas iš lankytojų 9 aukšte Metropole viešbutyje buvo buvęs Amerikos verslininkas. Jis išvyko į Hanoi ir mirė vietos Vietnamo Prancūzų ligoninėje
- * Jis užsikrėtė 80 žmonių prieš mirtį
- * Kiti lankytojai 9 aukšte Metropole viešbutyje išplatino ligą Kanadoje, Singapūre ir JAV.
- * Iki 2003 metų balandžio pabaigoje, liga buvo aptikta 25 šalių visame pasaulyje. 4300 atvejų ir 250 mirčių.

SARS: epidemija



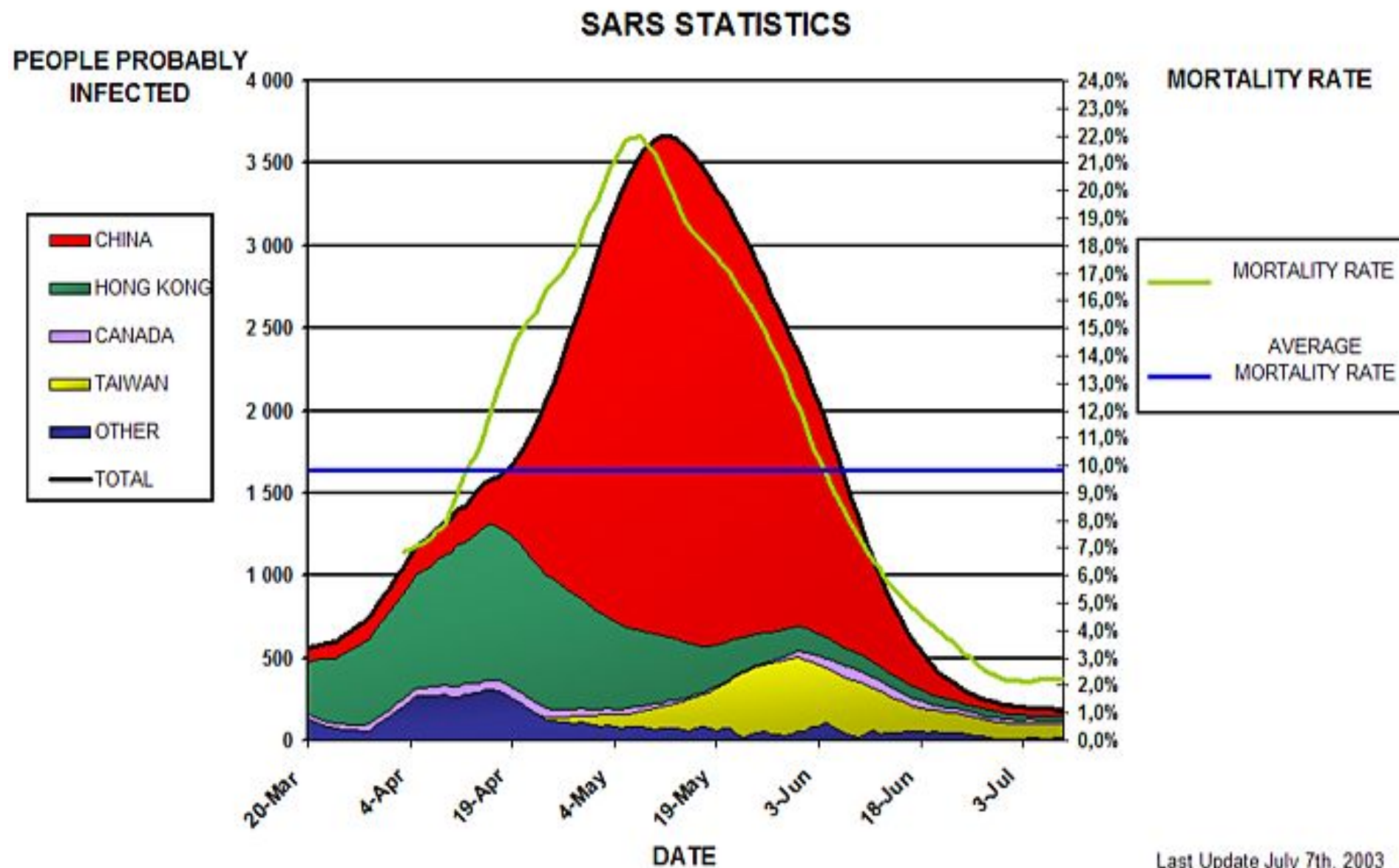
SARS panika Balandis-Birželis 2003



SARS korona virusas

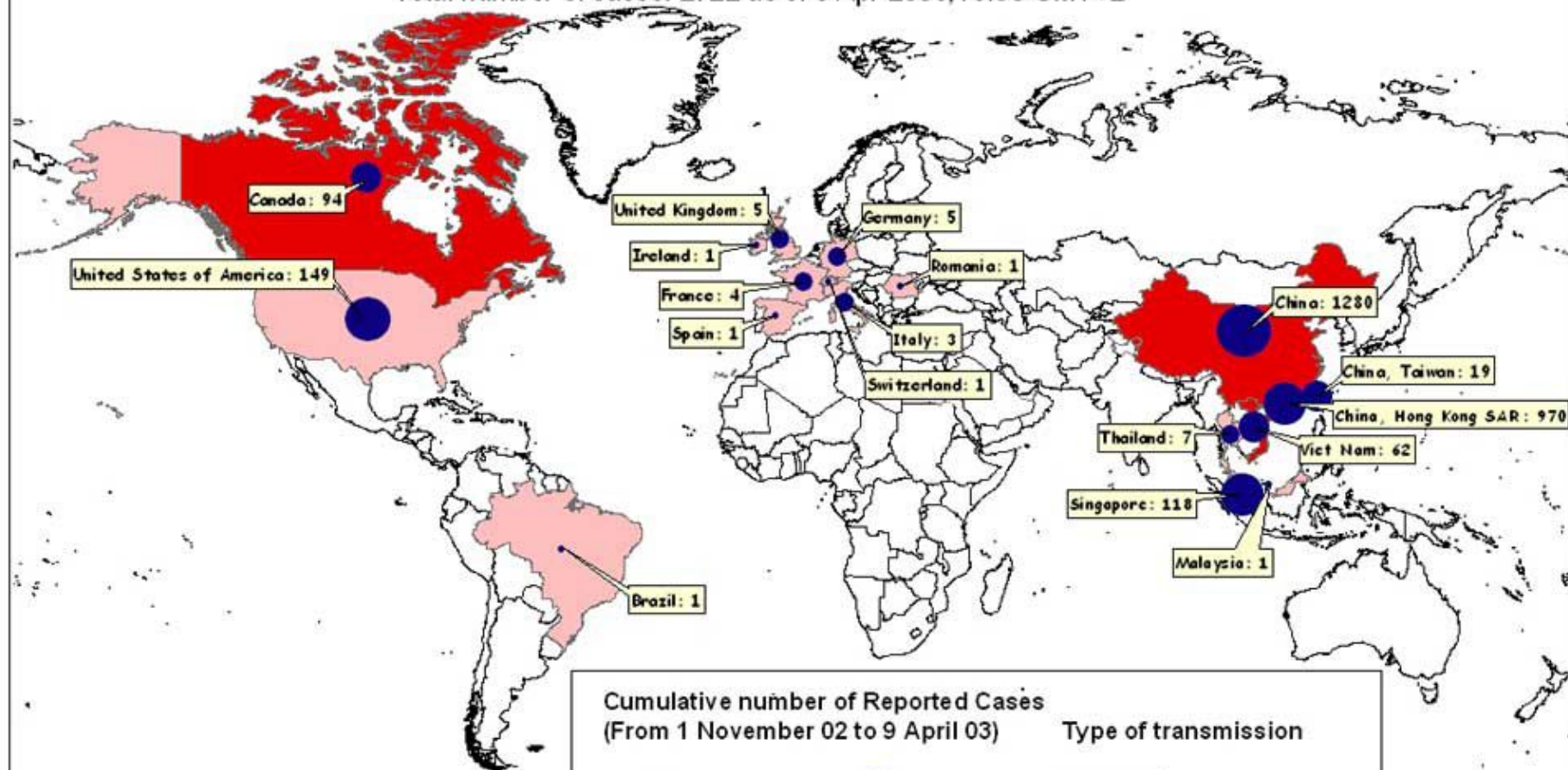
- * 2003 kovo pradžioje, PSO koordinavo tarptautinį tyrimą.
- * 2003 metų kovo pabaigoje, laboratorijos Vokietijoje, Kanadoje, JAV ir Honkonge nepriklausomai nustatytė naują virusą, kuris sukelia SARS.
- * SARS korona virusas (SARS-CoV) yra RNR virusas (toks yra ŽIV).
- * Korona virusai yra dažni tarp žmonių ir gyvūnų, sukelia ~ 25% visų viršutinių kvėpavimo takų infekcijų (pavyzdžiui, peršalimą).

SARS: epidemija



SARS : Cumulative Number of Reported Probable* Cases

Total number of cases: 2722 as of 9 Apr 2003, 15:00 GMT+2



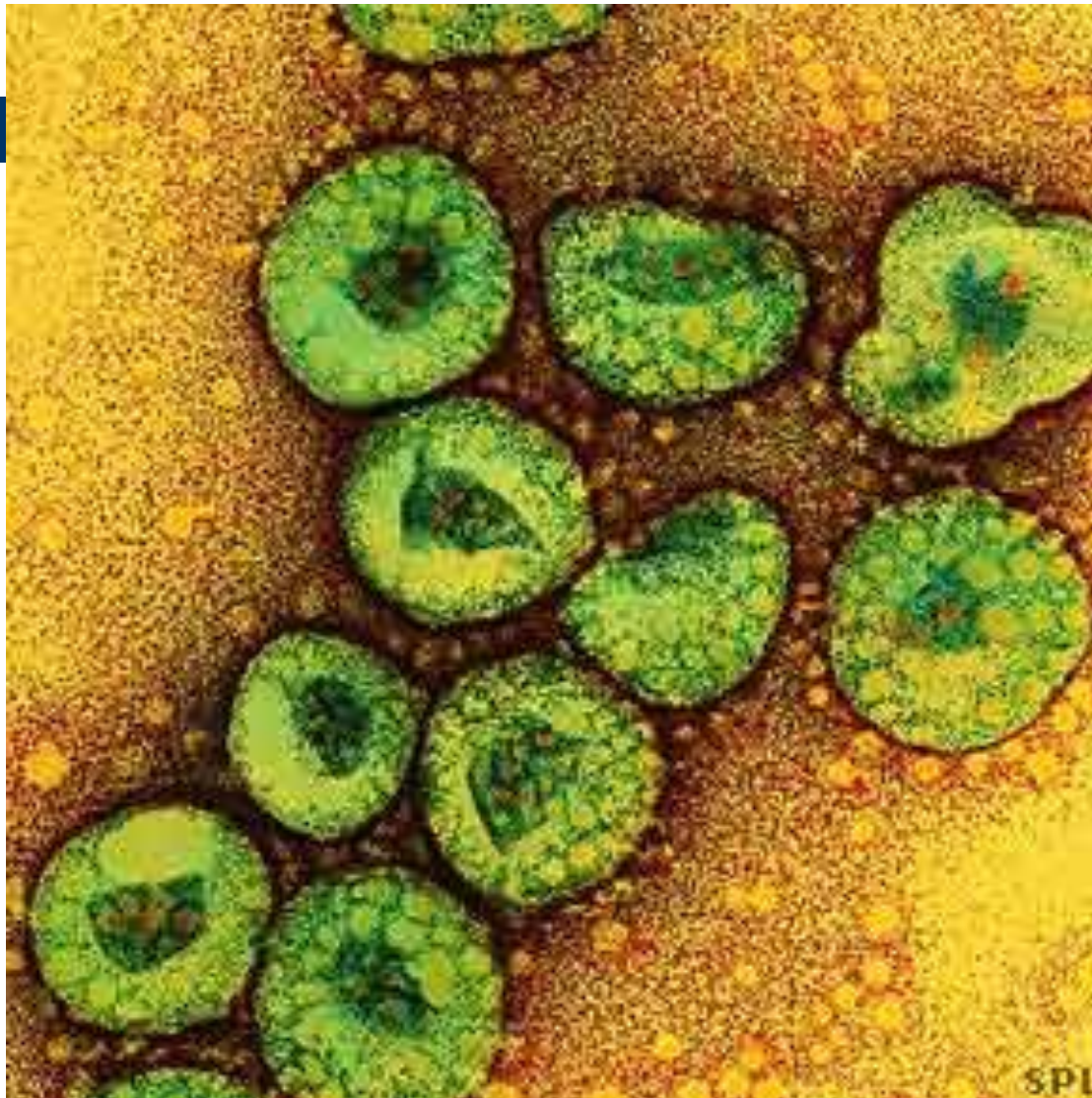
*Due to differences in the case definitions being used at a national level, probable cases are reported by all countries except the United States of America, which is reporting suspect cases under investigation.



The presentation of material on the maps contained herein does not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or areas or of its authorities, or concerning the delimitation of its frontiers or boundaries.

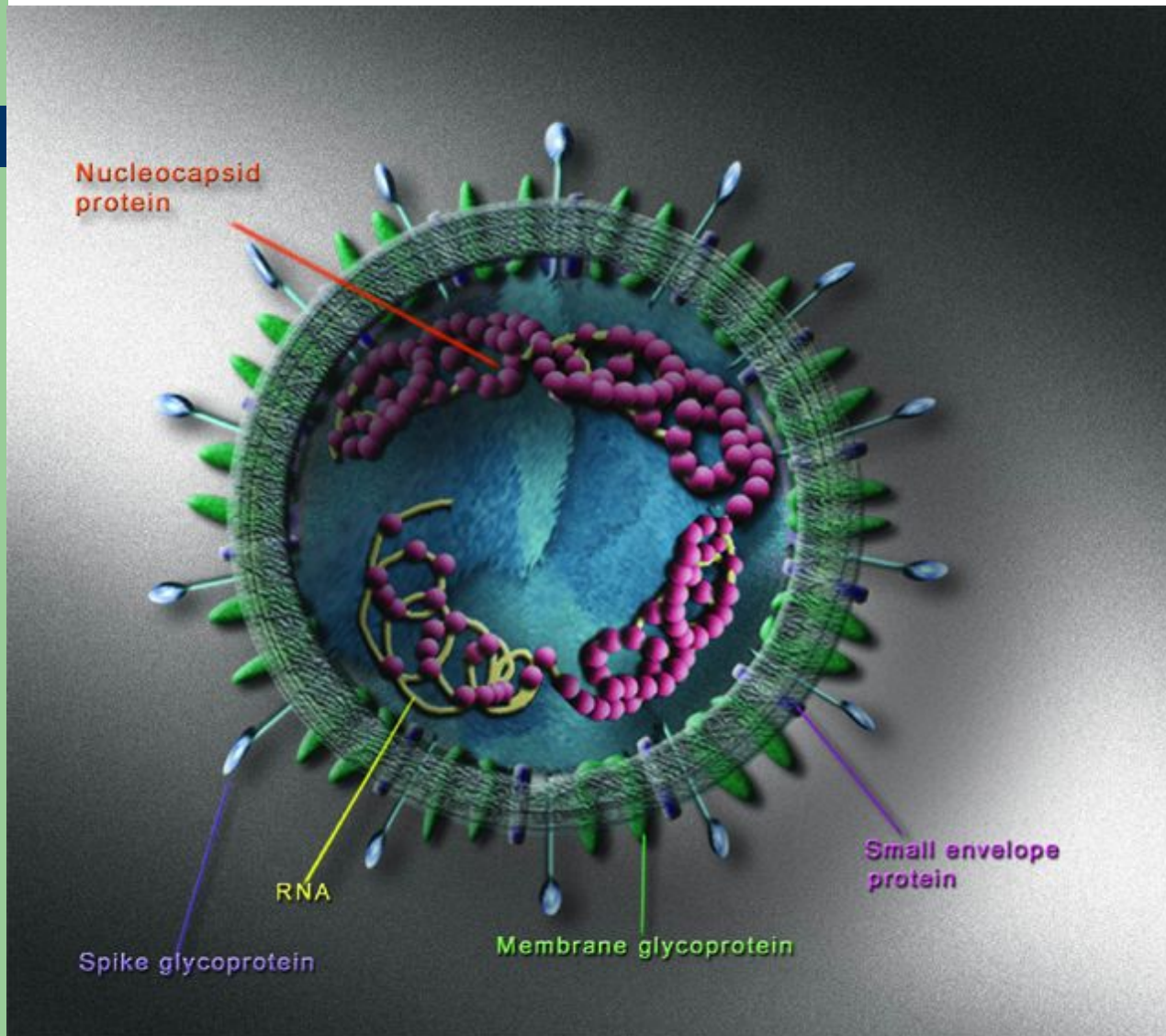
Data Source: World Health Organization
Map Production: Public Health Mapping Team
Communicable Diseases (CDS)
© World Health Organization, April 2003

SARS virusas



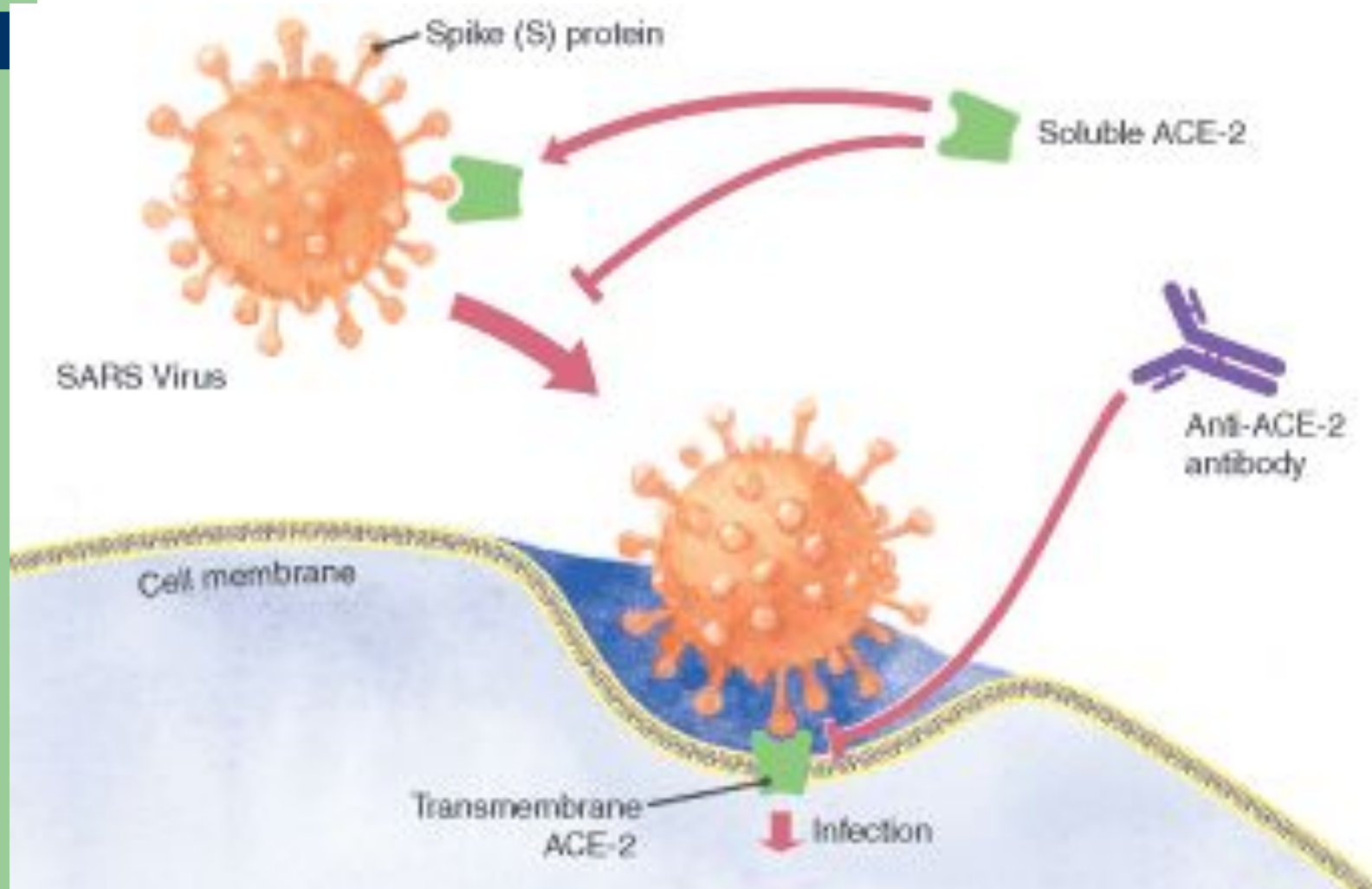
SARS: epidemija

SARS koronavirusas



SARS: epidemija

SARS koronavirusas



The SARS koronavirusas

- * 2003 balandžio mėn. Kanados laboratorija paskelbė visą SARS genomo seką.
- * Filogenetinė susijusių virusų analizė parodė, kad labiausiai susiję su juo yra koronavirusas iš palmių civetos.
- * Palmių civeta yra populiarus maisto produktas Guangdong provincijoje



SARS: epidemija



Palmių civeta

**Palmių civeta – maistas
kinijoje**



Filogenetinė SARS viruso analizė

- * Gegužę 2003, 2 straipsniai Science žurnale publikavo pilną genomo seką.
- * Genomas 29,751 bp ilgio.
- * Esmingai skiriasi nuo žmogaus koronavirusų.
- * Skiriasi nuo paukščių koronavirusų - nesusijęs su paukščių gripu

Filogenetinė SARS viruso analizė

- * Kokio tipo virusas sukelia infekciją?
- * Kas yra infekcijos šaltinis?
- * Kada ir kur virusas peržengė rūšių barjerą? Kokios mutacijos buvo esminės peržengimui?
- * Kuriose vietovėse iš pradžių plito virusas?

Filogenetinių medžių apskaičiavimo algoritmai

Medį galime apskaičiuoti remiantis

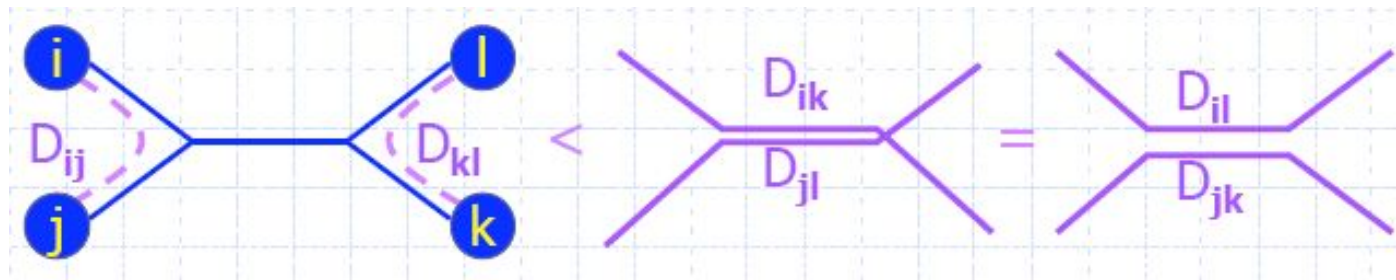
1. kiekybiniais įverčiais pagal atstumą ar panašumą tarp sekų atstumų matrica -> filogenetinis medis
Aptarsime šioje paskaitoje.

2. Remiantis kokybiniais aspektais, tokiais kaip labiausiai tikėtinas medis remiantis sekų simboliais (ateinančioje arba dar kitoje paskaitoje).

Filogenija

- **Adityvus medis:** Atstumų matrica atitikantį medį yra vadinama adityvia (pats medis yra adityvus)
- Tarkim matrica yra keturioms rūšims (indeksai i, j, k, l).
- Matrica bus adityvi tik tada jei kiekvienam indeksui iš i, j, k, l maksimumas ir mediana visų galimų porinių sumų bus identiška.

$$D_{ij} + D_{kl} < D_{ik} + D_{jl} = D_{il} + D_{jk}$$



UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA
 - Unweighted Pair – Group Method using arithmetic Averages.
 - Laiko, kad mutacijos vyksta pastoviu greičiu.
 - Laikoma, kad dvi sekos, tarp kurių atstumas yra mažiausias (mažiausiai besiskiria) išsiskyrė vėliausiai ir sudaro gali būti medyje apjungtos pirmiausiai.
 - Leidžia gauti medį su šaknimi.

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Atstumas tarp dviejų klasterių:
 - Turime N sekų
 - Klusteris X turi N_X sekų, klusteris Y turi N_Y sekų,
 - d_{XY} : evoliucinis atstumas tarp dviejų klasterių X ir Y

$$d_{XY} = \frac{1}{N_X N_Y} \sum_{i \in X, j \in Y} d_{ij}$$

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Kada toliau apjungiam klasterius X ir Y sudarydami klasterį Z:
 - Nereikia naudoti sekų ar sekų atstumų atnaujinant atstumų matricą ir į ją įtraukiant klasterį Z.
 - Apskaičiuojam atstumą tarp klasterio ir kitų klasterių (tokio kaip W)

$$d_{ZW} = \frac{N_X d_{XW} + N_Y d_{YW}}{N_X + N_Y}$$

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Pavyzdys:

d_{ij}	A	B	C	D	E	F
A	–	6	8	1	2	6
B		–	8	6	6	4
C			–	8	8	8
D				–	2	6
E					–	6

Atstumų matrica

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Pavyzdys:

d_{ij}	A	B	C	D	E	F
A	–	6	8	1	2	6
B		–	8	6	6	4
C			–	8	8	8
D				–	2	6
E					–	6



UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Pavyzdys:
 - A – D sekos apjungiamos į naują klasterį, tarkim V.
 - Turime modifikuoti atstumų matricą:
 - Apskaičiuoti kokie yra atstumai yra tarp likusių matricos komponentų:
 - V ir B,
 - V ir C,
 - V ir E,
 - V ir F.

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:

- Pavyzdys:

- A – D sekos apjungiamos į naują klasterį, tarkim V.
- Turime modifikuoti atstumų matricą:
- Koks yra atstumas tarp

- **V** ir **B** (**apskaičiuojame**)

d_{ij}	A	B	C	D	E	F
A	–	6	8	1	2	6
B		–	8	6	6	4
C			–	8	8	8
D				–	2	6
E					–	6

$$d_{VB} = \frac{N_A d_{AB} + N_D d_{DB}}{N_A + N_D} = \frac{1 * 6 + 1 * 6}{1 + 1} = 6$$

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:

- Pavyzdys:

- A – D sekos apjungiamos į naują klasterį, tarkim V.
- Turime modifikuoti atstumų matricą:
- Koks yra atstumas tarp
 - V ir C (**apskaičiuojame**)

d_{ij}	A	B	C	D	E	F
A	–	6	8	1	2	6
B		–	8	6	6	4
C			–	8	8	8
D				–	2	6
E					–	6

$$d_{VC} = \frac{N_A d_{AC} + N_D d_{DC}}{N_A + N_D} = \frac{1 * 8 + 1 * 8}{1 + 1} = 8$$

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:

- Pavyzdys:

- A – D sekos apjungiamos į naują klasterį, tarkim V.
- Turime modifikuoti atstumų matricą:
- Koks yra atstumas tarp

d_{ij}	A	B	C	D	E	F
A	–	6	8	1	2	6
B		–	8	6	6	4
C			–	8	8	8
D				–	2	6
E					–	6

- V ir E (apskaičiuojame)

$$d_{VE} = \frac{N_A d_{AE} + N_D d_{DE}}{N_A + N_D} = \frac{1 * 2 + 1 * 2}{1 + 1} = 2$$

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:

- Pavyzdys:

- A – D sekos apjungiamos į naują klasterį, tarkim V.
- Turime modifikuoti atstumų matricą:
- Koks yra atstumas tarp
 - V ir F (**apskaičiuojame**)

d_{ij}	A	B	C	D	E	F
A	–	6	8	1	2	6
B		–	8	6	6	4
C			–	8	8	8
D				–	2	6
E					–	6

$$d_{VF} = \frac{N_A d_{AF} + N_D d_{DF}}{N_A + N_D} =$$



UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Pavyzdys:
 - Nauja matrica

d_{ij}	A	B	C	D	E	F
A	–	6	8	1	2	6
B		–	8	6	6	4
C			–	8	8	8
D				–	2	6
E					–	6

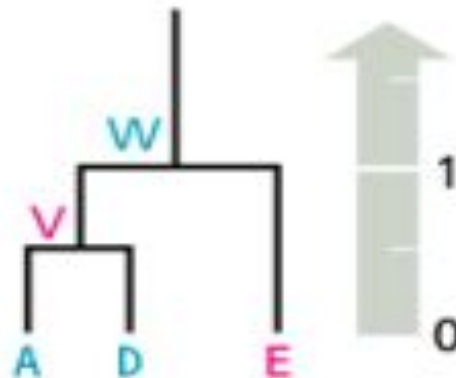


d_{ij}	B	C	E	F	V
B	–	8	6	4	6
C		–	8	8	8
E			–	6	2
F				–	6

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Pavyzdys:
 - Klasterizavimas pagal minimalų atstumą:

d_{ij}	B	C	E	F	V
B	–	8	6	4	6
C		–	8	8	8
E			–	6	2
F				–	6



UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:

- Pavyzdys:

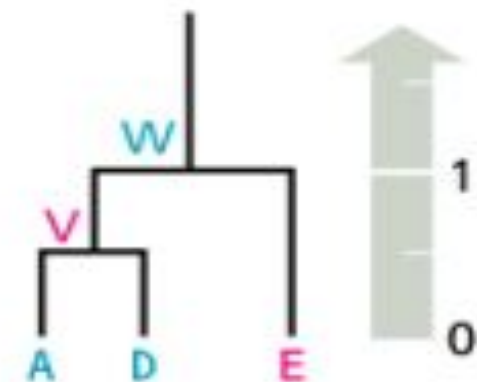
- V – E ajungiamos į naują klasterį, tarkim W.
- Modifikuojam atstumų matricą.
- Apskaičiuojam kokie atstumai yra tarp:

- **W** ir B,

- **W** ir C,

- **W** ir F.

d_{ij}	B	C	E	F	V
B	–	8	6	4	6
C		–	8	8	8
E			–	6	2
F				–	6



UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:

- Pavyzdys:

- V – E ajungiamos į naują klasterį, tarkim W
- Modifikuojam atstumų matricą.
- Koks atstumas yra tarp :
 - **W ir B (Apskaičiuojam),**

d_{ij}	B	C	E	F	V
B	–	8	6	4	6
C		–	8	8	8
E			–	6	2
F				–	6

$$d_{WB} = \frac{N_V d_{VB} + N_E d_{EB}}{N_V + N_E} = \frac{2 * 6 + 1 * 6}{2 + 1} = 6$$

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:

- Pavyzdys:

- V – E ajungiamos į naują klasterį, tarkim W
- Modifikuojam atstumų matricą.
- Koks atstumas yra tarp :
- **W ir C (Apskaičiuojam),**

d_{ij}	B	C	E	F	V
B	–	8	6	4	6
C		–	8	8	8
E			–	6	2
F				–	6

$$d_{WC} = \frac{N_V d_{VC} + N_E d_{EC}}{N_V + N_E} = \frac{2 * 8 + 1 * 8}{2 + 1} = 8$$

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:

- Pavyzdys:

- V – E ajungiamos į naują klasterį, tarkim W
- Modifikuojam atstumų matricą.
- Koks atstumas yra tarp :
 - **W ir F (Apskaičiuojam)**

d_{ij}	B	C	E	F	V
B	–	8	6	4	6
C		–	8	8	8
E			–	6	2
F				–	6

$$d_{WF} = \frac{N_V d_{VF} + N_E d_{EF}}{N_V + N_E} =$$



UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Pavyzdys:
 - Nauja matrica:

d_{ij}	B	C	E	F	V
B	–	8	6	4	6
C		–	8	8	8
E			–	6	2
F				–	6

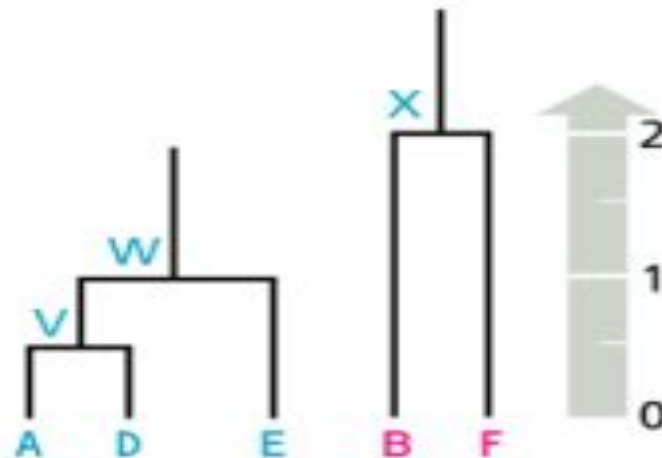


d_{ij}	B	C	F	W
B	–	8	4	6
C		–	8	8
F			–	6

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Pavyzdys:
 - klasterizuojam pagal minimalų atstumą:

d_{ij}	B	C	F	W
B	–	8	4	6
C		–	8	8
F			–	6



UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Pavyzdys:
 - F – B apjungiam į naują klasterį, tarkim X.
 - Turim modifikuoti atstumų matricą.
 - Koks yra atstumas tarp :
 - **W** ir **X**.

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Pavyzdys:
 - Koks yra atstumas tarp **W** ir **X** (Apskaičiuojame).

$$d_{WX} = \frac{1}{N_W N_X} \sum_{i \in W, j \in X} d_{ij} =$$

$$\frac{1}{N_W N_X} (d_{AB} + d_{AF} + d_{DB} + d_{DF} + d_{EB} + d_{EF}) =$$

$$\frac{1}{3 * 2} * (6 + 6 + 6 + 6 + 6 + 6) = 6$$

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Pavyzdys:
 - Nauja matrica:

d_{ij}	B	C	F	W
B	–	8	4	6
C		–	8	8
F			–	6

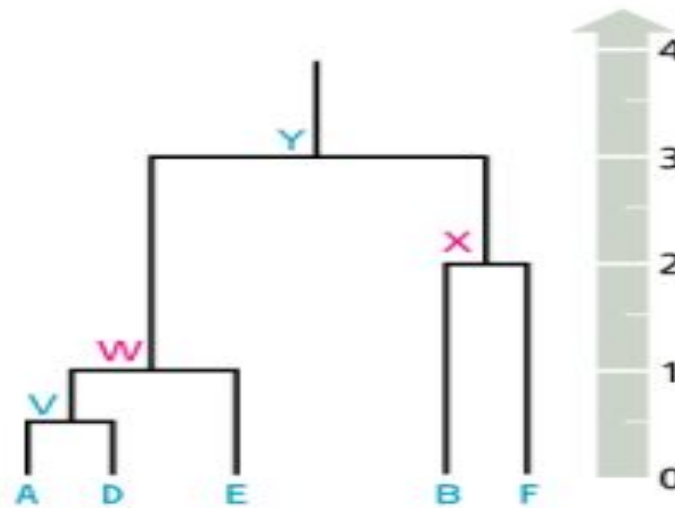


d_{ij}	C	W	X
C	–	8	8
W		–	6

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Pavyzdys:
 - klasterizuojam toliau pagal minimalų atstumą:

d_{ij}	C	W	X
C	–	8	8
W		–	6



UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Pavyzdys:
 - X – W apjungiam į naują klasterį, tarkim Y.
 - Turime modifikuoti atstumų matricą.
 - Turime apskaičiuoti atstumą tarp :
 - **Y ir C.**

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Pavyzdys:
 - Nauja matrica:

d_{ij}	c	w	x
c	-	8	8
w		-	6



d_{ij}	c	y
C	-	8

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Pavyzdys:
 - Koks yra atstumas tarp : **Y** ir **C** (Apskaičiuojame).

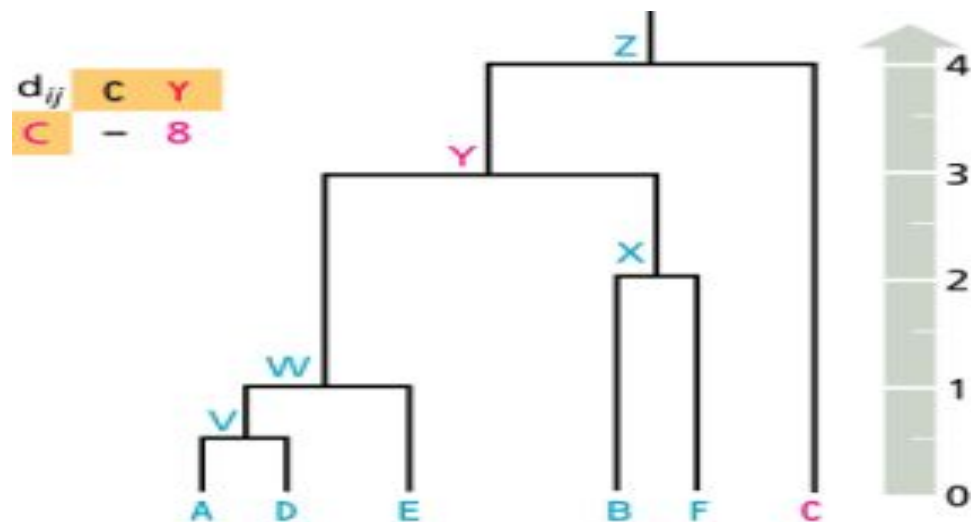
$$d_{YC} = \frac{1}{N_Y N_C} \sum_{i \in Y, j \in C} d_{ij} =$$

$$\frac{1}{N_Y N_C} (d_{AC} + d_{DC} + d_{EC} + d_{BC} + d_{FC}) =$$

$$\frac{1}{5 * 1} * (8 + 8 + 8 + 8 + 8) = 8$$

UPGMA

- Filogenetinių medžių kūrimo metodas UPGMA:
 - Pavyzdys:
 - toliau klasterizuojuame pagal atstumą:

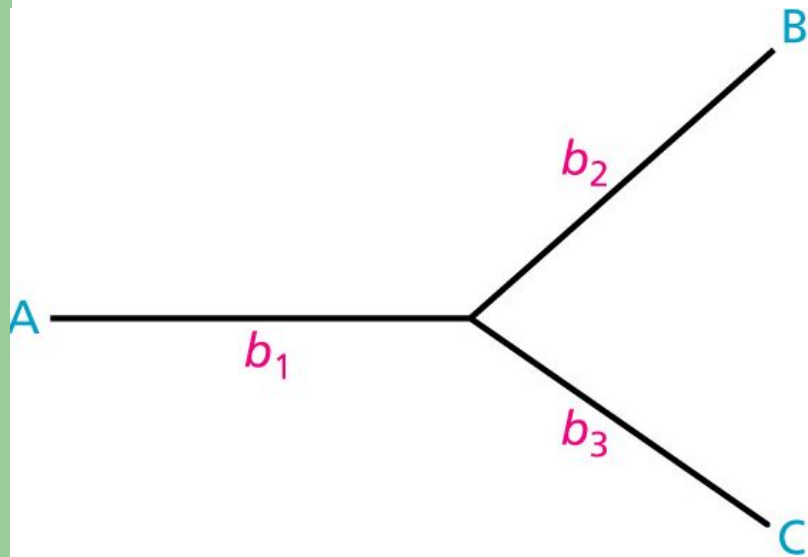


Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Netaikoma prielaida, kad mutacijų greitis yra pastovus.
 - Daroma prielaida, kad atstumai yra adityvūs.
 - Gaunamas bešaknis medis.

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Atstumai d_{ij} :



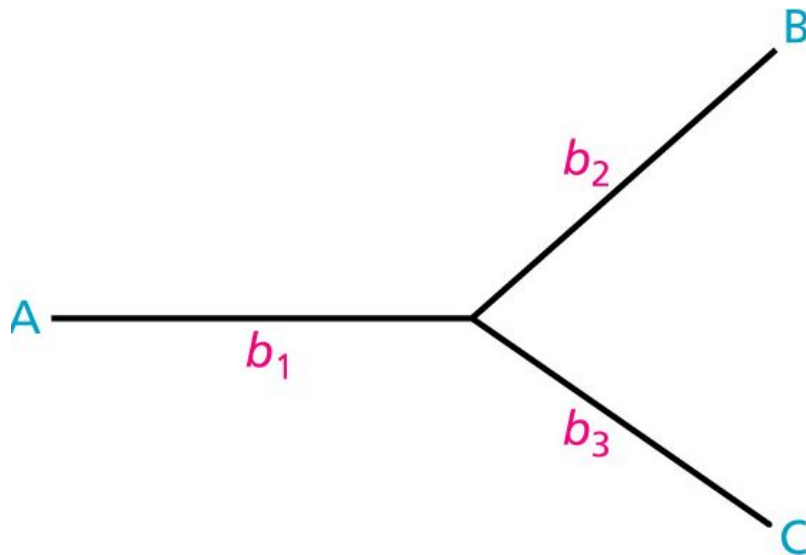
$$d_{AB} = b_1 + b_2$$

$$d_{AC} = b_1 + b_3$$

$$d_{BC} = b_2 + b_3$$

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - The šakų ilgius galime suskaičiuoti laikantis prielaidos, kad atstumai yra adityvūs



$$b_1 = \frac{1}{2}(d_{AB} + d_{AC} - d_{BC})$$

$$b_2 = \frac{1}{2}(d_{AB} + d_{BC} - d_{AC})$$

$$b_3 = \frac{1}{2}(d_{AC} + d_{BC} - d_{AB})$$

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - The atstumai tarp klasterių apibrėžiami kaip ir UPGMA atveju:

$$d_{XY} = \frac{1}{N_X N_Y} \sum_{i \in X, j \in Y} d_{ij}$$

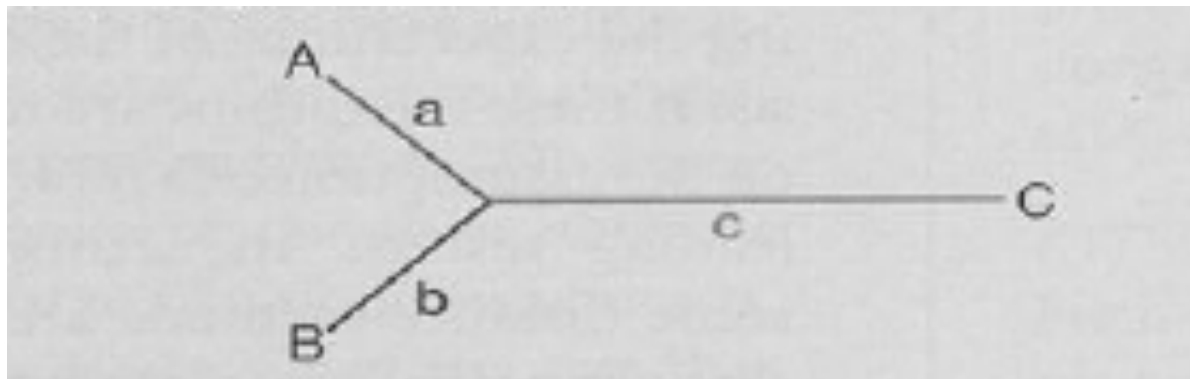
$$d_{ZW} = \frac{N_X d_{XW} + N_Y d_{YW}}{N_X + N_Y}$$

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :

- Pavyzdys trims rūšims:

Nubraižykite bešaknį medį su trim atšakomis einančiomis nuo bedro mazgo. Iš anksto žinant, kad A ir B yra artimesnės tarpusavyje nei su C, šakų ilgiai tarp jų turėtų būti trumpesni:

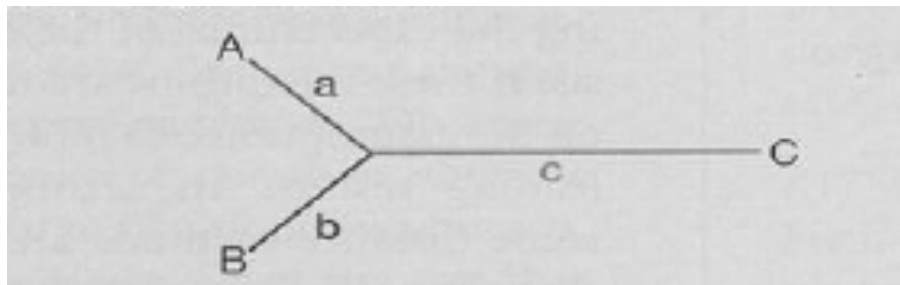


	A	B	C
A	–	22	39
B	–	–	41
C	–	–	–

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Pavyzdys trims rūšims:
a,b,c atstumus apskaičiuojame algeбриškai, remiantis žinoma atstumų matrica.

	A	B	C
A	–	22	39
B	–	–	41
C	–	–	–



$$A..B = a+b= 22 \quad (1)$$

$$A..C = a+c=39 \quad (2)$$

$$B..C= b+c=41 \quad (3)$$

$$(2)-(3), a-b=-2 \quad (4)$$

$$(1)+(4), 2a=20, a=10$$

įstatę į (1) ir (2), gaunam
 $b=12, c=29$

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:

	A	B	C	D	E
A		22	39	39	41
B			41	41	43
C				18	20
D					10
E					

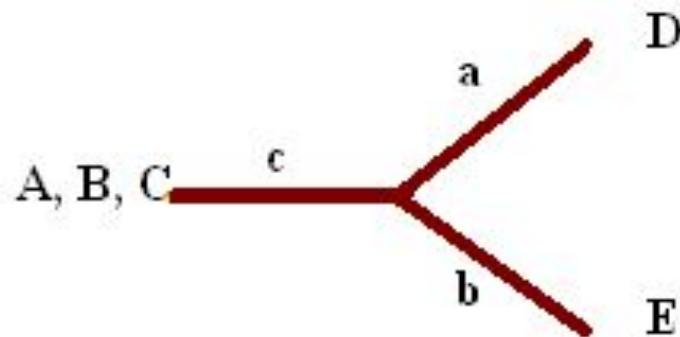
Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - **D ir E yra artimiausios sekos**

	A	B	C	D	E
A		22	39	39	41
B			41	41	43
C				18	20
D					10
E					

Fitch-Margoliash metodas:

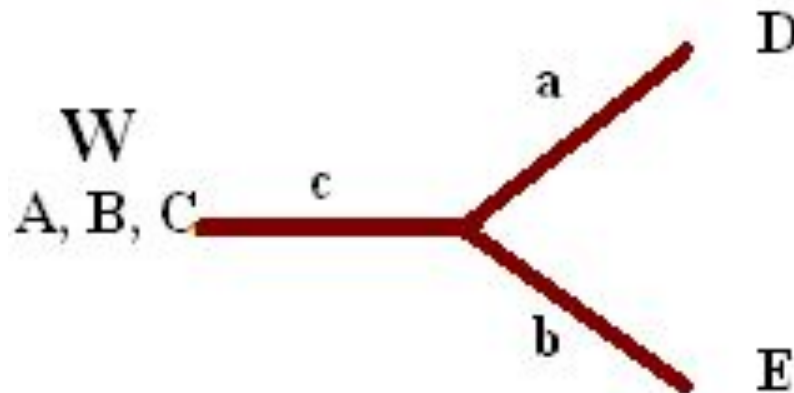
- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - **D ir E yra artimiausios sekos**



	A	B	C	D	E
A		22	39	39	41
B			41	41	43
C				18	20
D					10
E					

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - Laikome, kad $\{A, B, C\}$ yra W,



	A	B	C	D	E
A		22	39	39	41
B			41	41	43
C				18	20
D					10
E					

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - **Atstumas tarp W ir D:**

$$d_{WD} = \frac{1}{N_W N_D} \sum_{i \in W, j \in D} d_{ij} =$$
$$\frac{1}{N_W N_D} (d_{AD} + d_{BD} + d_{CD}) =$$
$$\frac{1}{3 * 1} * (39 + 41 + 18) = 33$$

	A	B	C	<u>D</u>	E
A		22	39	39	41
B			41	41	43
C				18	20
D					10
E					

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - **Atstumas tarp W ir E:**

$$d_{WE} = \frac{1}{N_W N_E} \sum_{i \in W, j \in E} d_{ij} =$$
$$\frac{1}{N_W N_E} (d_{AE} + d_{BE} + d_{CE}) =$$
$$\frac{1}{3 * 1} * (41 + 43 + 20) = 35$$

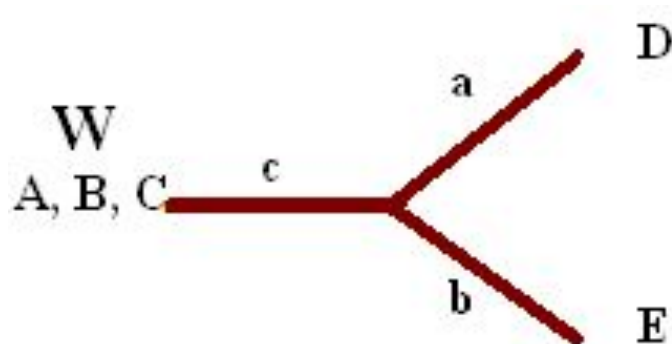
	A	B	C	D	<u>E</u>
A		22	39	39	41
B			41	41	43
C				18	20
D					10
E					

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :

- Antras pavyzdys:

- **Apskaičiuojam šakas a, b ir c:**



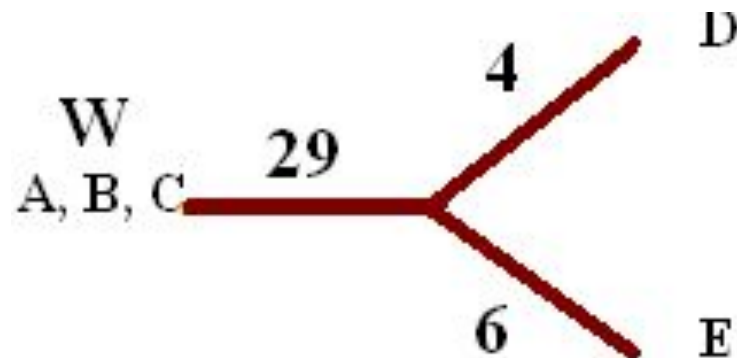
$$a = \frac{1}{2}(d_{WD} + d_{DE} - d_{WE}) = \frac{1}{2}(33 + 10 - 35) = 4$$

$$b = 10 - 4 = 6$$

$$c = 33 - 4 = 29$$

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - Šakos a, b ir c:



Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - **Atnaujina atstumų matricą:**

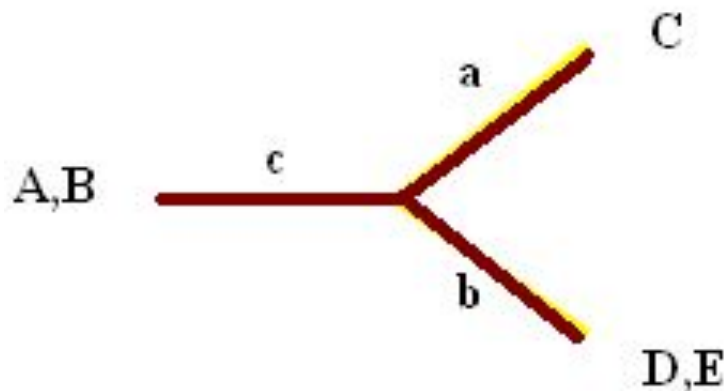
	A	B	C	D	E
A		22	39	39	41
B			41	41	43
C				18	20
D					10
E					



	A	B	C	{D,E}
A		22	39	40
B			41	42
C				19
{D,E}				

Fitch-Margoliash metodas:

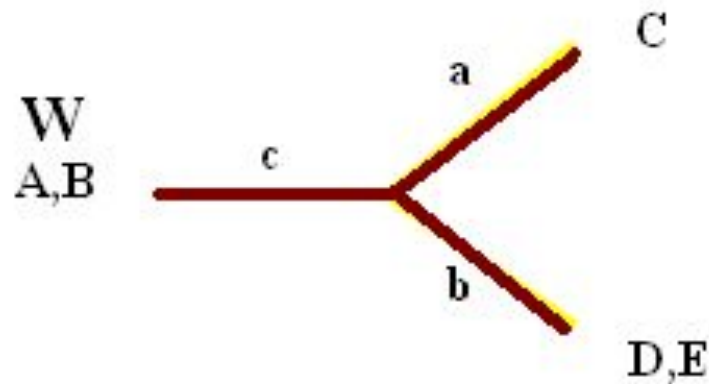
- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - **{D,E} ir C yra artimiausios sekos**



	A	B	C	{D,E}
A		22	39	40
B			41	42
C				19
{D,E}				

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - Laikome, kad $\{A, B\}$ yra W :



	A	B	C	{D,E}
A		22	39	40
B			41	42
C				19
{D,E}				

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - **Atstumas tarp W ir C:**

$$d_{WC} = \frac{1}{N_W N_C} \sum_{i \in W, j \in C} d_{ij} =$$

$$\frac{1}{N_W N_C} (d_{AC} + d_{BC}) =$$

$$\frac{1}{2 * 1} * (39 + 41) = 40$$

	A	B	C	D	E
A		22	39	39	41
B			41	41	43
C				18	20
D					10
E					

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - **Atstumas tarp W ir {D,E} ({D,E} žymime X):**

$$d_{WX} = \frac{1}{N_W N_X} \sum_{i \in W, j \in X} d_{ij} =$$
$$\frac{1}{N_W N_X} (d_{AD} + d_{AE} + d_{BD} + d_{BE}) =$$
$$\frac{1}{2 * 2} * (39 + 41 + 41 + 43) = 41$$

	A	B	C	D	E
A		22	39	39	41
B			41	41	43
C				18	20
D					10
E					

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - **Atstumas tarp C ir {D,E} (name {D,E} as X):**

$$d_{CX} = \frac{1}{N_C N_X} \sum_{i \in C, j \in X} d_{ij} =$$

$$\frac{1}{N_C N_X} (d_{CD} + d_{CE}) =$$

$$\frac{1}{1 * 2} * (18 + 20) = 19$$

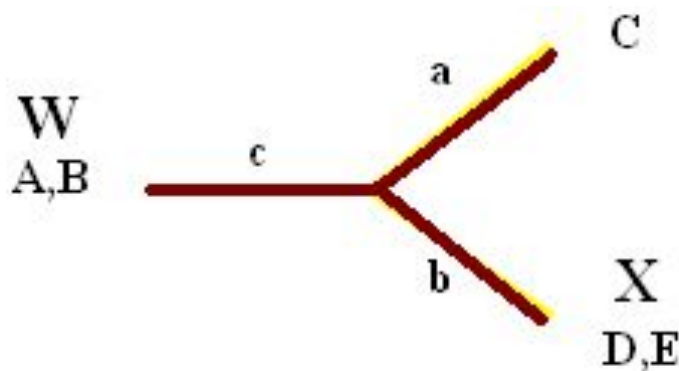
	A	B	C	D	E
A		22	39	39	41
B			41	41	43
C				18	20
D					10
E					

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :

- Antras pavyzdys:

- **Apskaičiuojam šakas a, b ir c:**



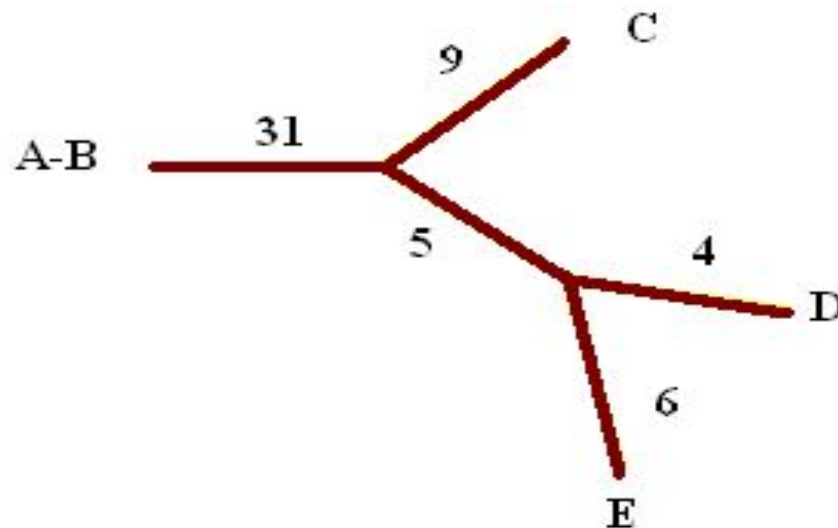
$$a = \frac{1}{2}(d_{WC} + d_{CX} - d_{WX}) = \frac{1}{2}(40 + 19 - 41) = 9$$

$$b = 19 - 9 = 10$$

$$c = 40 - 9 = 31$$

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - Šakos a, b ir c:



Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - **Atnaujiname atstumų matricą:**

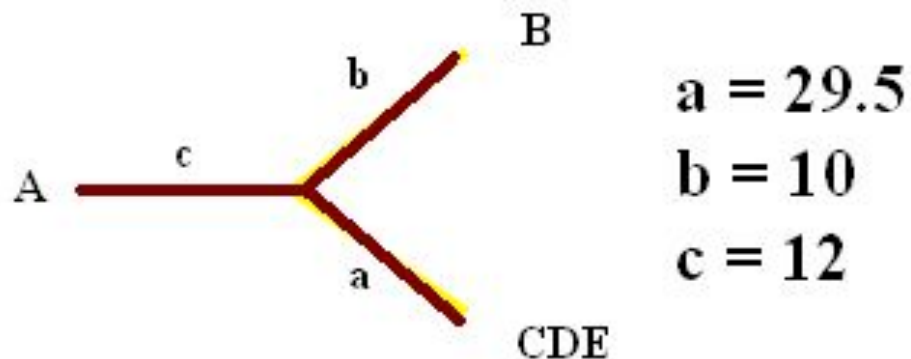
	A	B	C	{D,E}
A		22	39	40
B			41	42
C				19
{D,E}				



	A	B	{C,D,E}
A		22	39.5
B			41.5
{C,D,E}			

Fitch-Margoliash metodas:

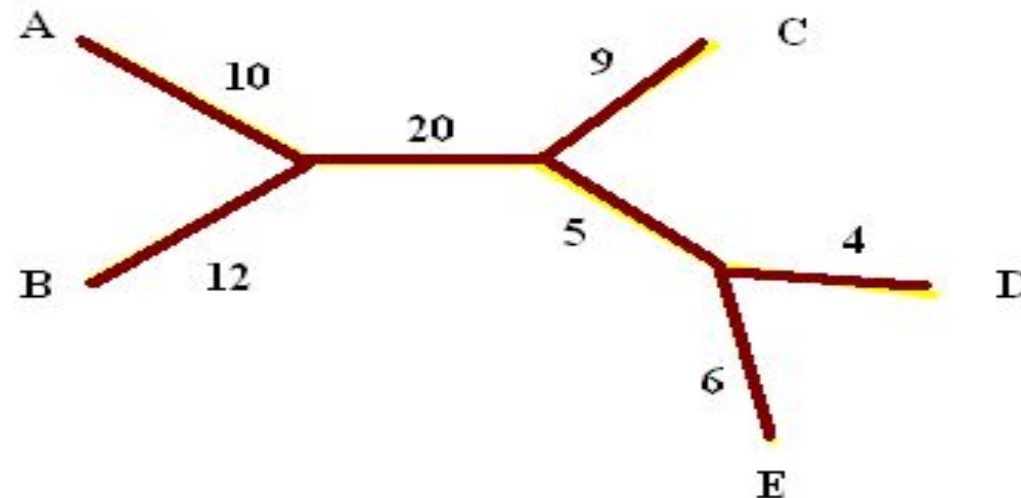
- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - Dabar mes esame klasikiniame pavyzdyje su tri sekomis (ankstesnis pavyzdys):



	A	B	{C,D,E}
A		22	39.5
B			41.5
{C,D,E}			

Fitch-Margoliash metodas:

- Filogenetinis medis Fitch-Margoliash metodu :
 - Antras pavyzdys:
 - **FINAL TREE:**



Neighbor-Joining metodas:

- Filogenetinis medis Neighbor-Joining metodu:
 - Tikras medis yra tas, kurio bendras šakų ilgis, S , yra trumpiausias.
 - Kaimynai: mazgų pora, kurie yra tik per vieną kitą mazgą.
 - Gaunamas bešaknis medis

Neighbor-Joining metodas:

- Filogenetinis medis Neighbor-Joining metodu:

- Algoritmas(Duotajai atstumų matricai $d_{i,j}$):

- Kartojame, kol 2 mazgai yra likę:

- Kiekvienam mazgui randame

$$U_i = \sum_{k=1}^N d_{i,k}$$

- Pasirinkite porą (i, j) su mažiausiu

$$\delta_{ij} = d_{ij} - \frac{U_i + U_j}{N - 2}$$

- Apjungite du mazgus i ir j į naują vidinį mazgą Y, ir raskite šakų ilgius

$$b_{iY} = \frac{1}{2} \left(d_{ij} + \frac{U_i - U_j}{N - 2} \right) \quad b_{jY} = d_{ij} - b_{iY}$$

- Atnaujiname atstumų matricą

$$d_{Yk} = \frac{1}{2} (d_{ik} + d_{jk} - d_{ij})$$

Neighbor-Joining metodas:

- Filogenetinis medis Neighbor-Joining metodu:
 - Pavyzdys:

		d_{ij}			
		B	C	D	E
A	5	4	9	8	
B		5	10	9	
C			7	6	
D				7	
E					

$$U_i = \sum_{k=1}^N d_{i,k} \quad \delta_{ij} = d_{ij} - \frac{U_i + U_j}{N - 2}$$

$$b_{iY} = \frac{1}{2} \left(d_{ij} + \frac{U_i - U_j}{N - 2} \right) \quad b_{jY} = d_{ij} - b_{iY}$$

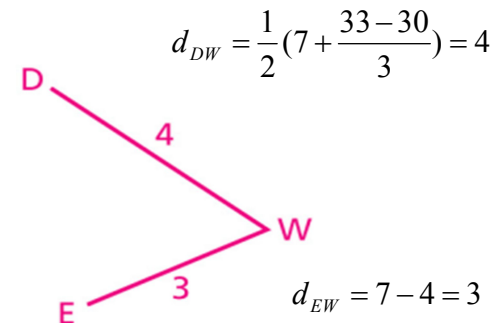
Neighbor-Joining metodas:

- Filogenetinis medis Neighbor-Joining metodu:
 - Pavyzdys:

STEP 1 (N = 5)

	d_{ij}				U_i	$3\delta_{ij}$				
	B	C	D	E		B	C	D	E	
A	5	4	9	8	26	-40	-36	-32	-32	A
B		5	10	9	29		-36	-32	-32	B
C			7	6	22			-34	-34	C
D				7	33				-42	D
E					30					E

D ir E yra kaimynai per vidinį mazgą W su atstumais nuo jo:



$$U_i = \sum_{k=1}^N d_{i,k} \quad \delta_{ij} = d_{ij} - \frac{U_i + U_j}{N - 2}$$

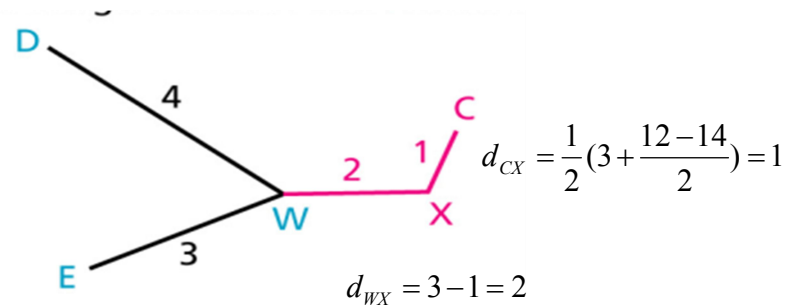
$$b_{iY} = \frac{1}{2} \left(d_{ij} + \frac{U_i - U_j}{N - 2} \right) \quad b_{jY} = d_{ij} - b_{iY}$$

Neighbor-Joining metodas:

- Filogenetinis medis Neighbor-Joining metodu:
 - Pavyzdys:

STEP 2 (N = 4)

	d_{ij}			U_i	$2\delta_{ij}$			
	B	C	W		B	C	W	
A	5	4	5	14	-20	-18	-18	A
B		5	6	16		-18	-18	B
C			3	12			-20	C
W				14				W



C ir W yra kaimynai per vidinį mazgą X su atstumais nuo jo:

$$U_i = \sum_{k=1}^N d_{i,k} \quad \delta_{ij} = d_{ij} - \frac{U_i + U_j}{N - 2}$$

$$b_{iY} = \frac{1}{2} \left(d_{ij} + \frac{U_i - U_j}{N - 2} \right) \quad b_{jY} = d_{ij} - b_{iY}$$

Neighbor-Joining metodas:

- Filogenetinis medis Neighbor-Joining metodu:

– Pavyzdys:

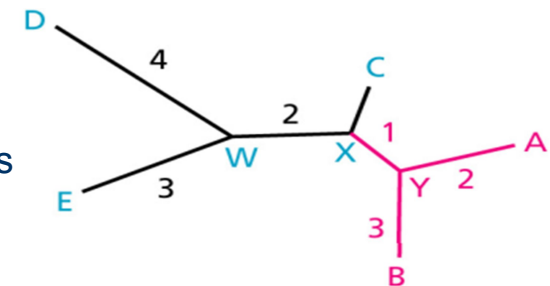
STEP 3 (N = 3)			d_{ij}	U_i	δ_{ij}	
	B	X				
A	5	3	8	-12	-12	A
B		4	9	-12		B
X			7			X

Trys alternatyvos (pasirenkam tuos, kurie yra su vidiniu mazgu).

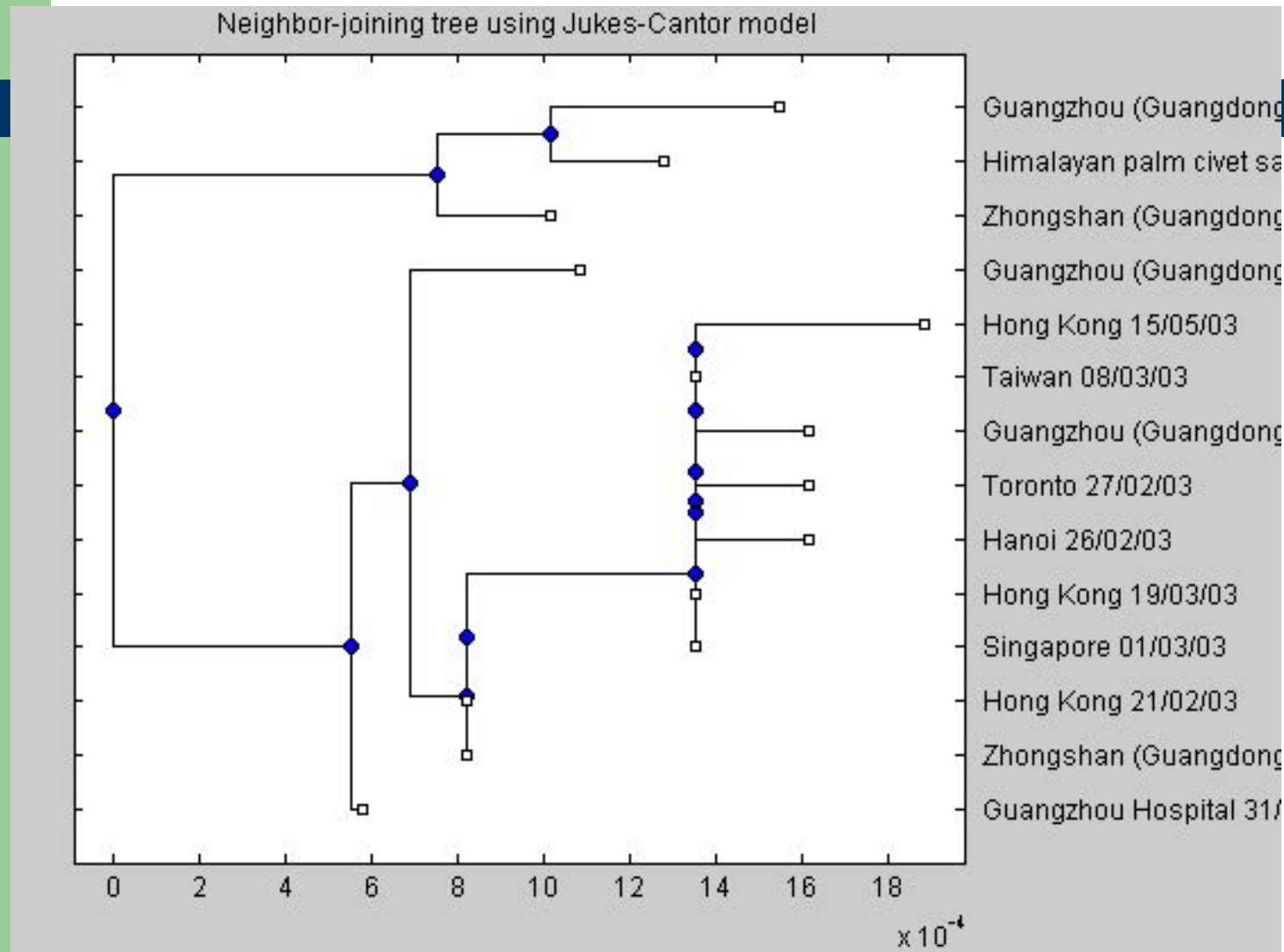
A ir X yra kaimynai per vidinį mazgą Y su atstumais $d_{AY}=2$ ir $d_{XY}=1$

B ir X yra kaimynai per vidinį mazgą Y su atstumais $d_{BY}=3$ ir $d_{XY}=1$.

Nesvarbu, kurį variantą bepasirinktume likęs atstumas d_{AY} arba d_{BY} bus rastas apskaičiuojant sekančią d_{ij} matricą.

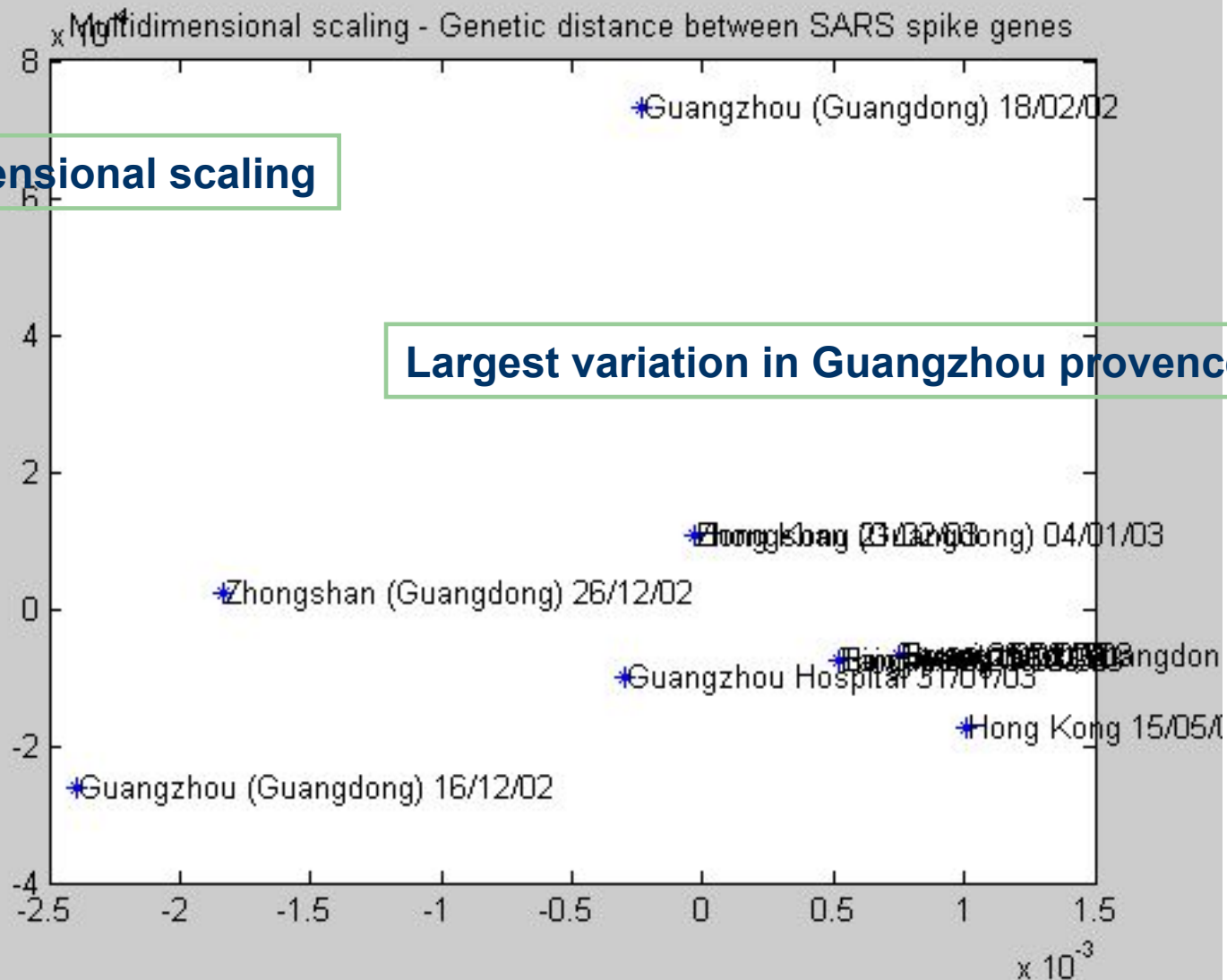


Neighbor-Joining metodu gautas SARS viruso medis



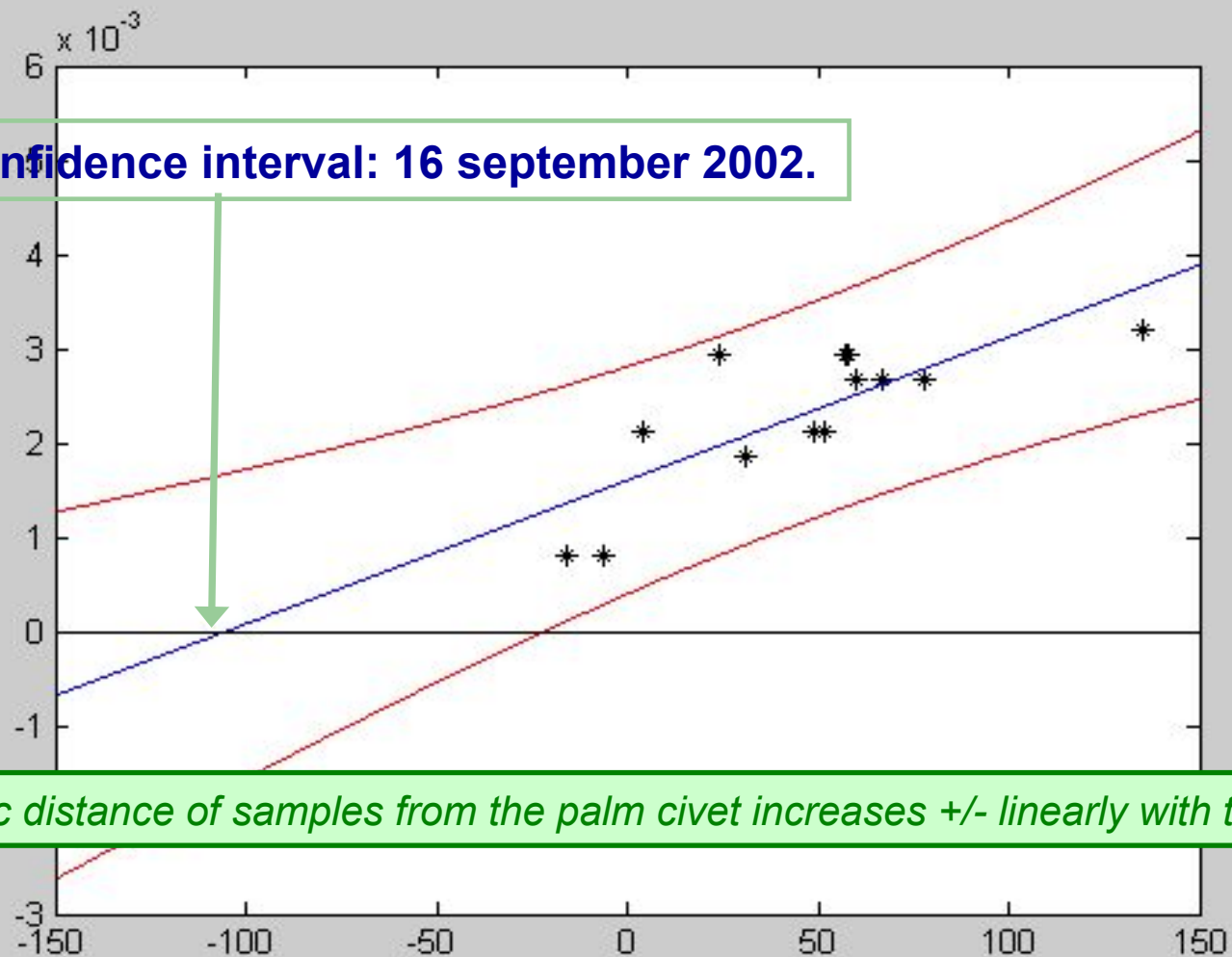
Kuris parodo kilmės vietą

multidimensional scaling



Atsiradimo vieta...

95% confidence interval: 16 september 2002.



The genetic distance of samples from the palm civet increases +/- linearly with time