

# Konstrukce fylogenetických stromů

Jakub Zárýbnický (xzaryb00@stud.fit.vutbr.cz)

March 24, 2020

## 1 Viacnásobné zarovnanie sekvencie cytochromu C a základná fylogenetická analýza

Stiahnite si aminokyselinové sekvencie nasledovných organizmov z NCBI a uložte ich do jedného multifasta súboru:

- Equus caballus (NP<sub>001157486.1</sub>)
- Homo sapiens (NP<sub>061820.1</sub>)
- Bos taurus (NP<sub>001039526.1</sub>)
- Gallus gallus (NP<sub>001072946.1</sub>)
- Canis lupus (AEP27248.1)
- Sus scrofa (NP<sub>001123442.1</sub>)
- Pan troglodytes (NP<sub>001065289.1</sub>)
- Taeniopygia guttata (NP<sub>001137145.2</sub>)
- Novosphingobium sp. PP1Y (WP<sub>013831814.1</sub>)

Sekvence importujte do programu MEGA a analyzujte:

- Kliknite File->Open a File/Session, How would you like to open this fasta file - kliknite Align, potom z menu Alignment exploreru vyberte Alignment -> Align by ClustalW (ak sedíte v sudom stĺpci), resp. Muscle (ak sedíte v lichom stĺpci). **Vyšiel susedovi rovnaký alignment?**
  - V niekoľkých pozíciách se Muscle a ClustalW líši, v umiestnení gaps

- Vizualne porovnajete jednotlivé zarovnané sekvencie. \*Ktorá zo sekvencií sa zarovnala vzhľadom k ostatným nesprávne/je vysoko divergentná vzhľadom k ostatným? Do akej skupiny organizmov patrí?\*

– *Novosphingobium* sp. PP1Y je jediná viditeľne odlišná, jedná sa o proteobakterii

- **V koľkých pozíciách sa líšia zvyšné sekvencie?**

– Zběžně jsem napočítal 14 pozic

- V Alignment exploreri vyberte v menu Data->Export alignment->Mega format a uložte.
- Zatvorte Alignment explorer (alignment už znova nemusíte ukladať)
- Vytvorte v programe Mega jednoduchú fylogenetickú analýzu - ak ste v sudom stĺpci, použite algoritmus NJ, ak ste v lichom stĺpci, použite Maximum Parsimony (Phylogeny->Construct NJ/Maximum parsimony).

Pre algoritmus **NJ** použite:

- Model/Method: Equal Input Model
- Gaps/Missing Data Treatment: Partial deletion

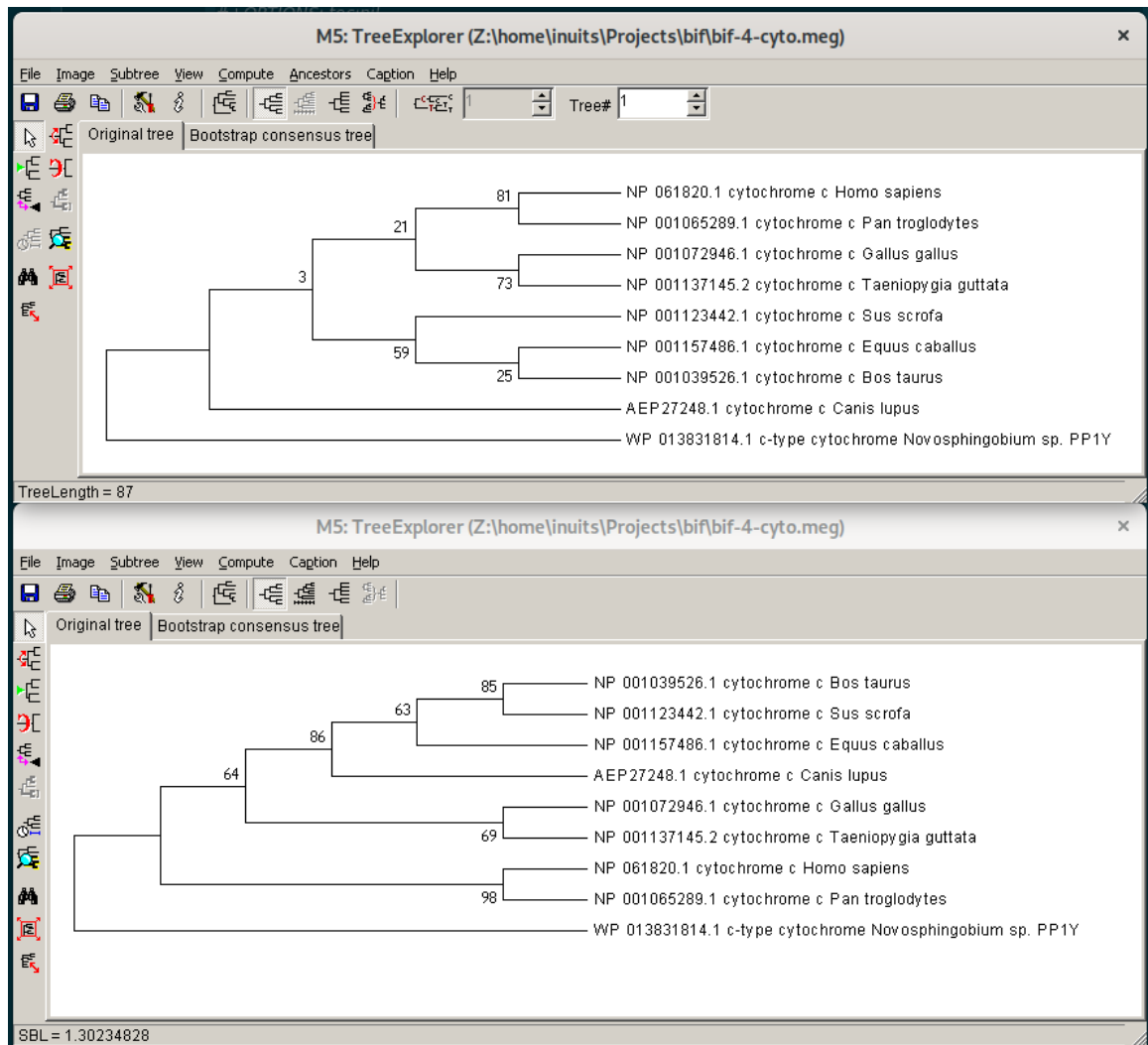
Pre algoritmus **Maximum Parsimony** použite:

- Gaps/Missing Data Treatment: Partial deletion

Ostatné parametre ponechajte implicitné. V TreeExploreri sa Vám zobrazí vygenerovaný fylogenetický strom. Zvoľte v menu TreeExploreri Subtree->Root a kliknite na vetvu *Novosphingobium*. Týmto strom "zakoreníte" a označíte Outgroup.

- Zhoduje sa topológia stromov vytvorených pomocou MP a NJ? V čom sa líšia? Diskutujte s Vašimi susedmi výsledky.
  - Původně jsem psal, že stromy jsou stejně, ale to musela být chyba. "Original tree" se liší značně - *Canis lupus* je podle MP nejméně podobný ostatním. "Bootstrap consensus tree" už je podobnější, ale pořád se liší ve stromu kopytnatců, kde prase je blíže buď koni (v MP), nebo turovi (v NJ).

- Vypátrajte české/slovenské názvy skúmaných organizmov. Korešponduje vzniknutý fylogenetický strom s ich fenotypom?
  - Ne úplne - kopytnatci jsou ve výsledné topologii správně blízko, ale jsou blíže ptákům (zebříčka a kur) než člověku nebo šimpanzovi.



## 2 Fylogenetická analýza človeka s príbuznými druhmi

Zostavte si jednoduchý dataset (multifasta) pozostávajúci z DNA sekvencií mitochondrial d-loop nasledovných druhov, ktoré si stiahnite z NCBI. Ako popis sekvencií v multifasta súbore použite české/slovenské názvy organizmov.

- Pongo abelii (FR717938.1) alias orangutan
- Pan troglodytes (GU136845.1) alias šimpanz
- Homo sapiens (HM009355.1) alias človek dnešného typu
- Gorilla graueri (AF050738.1) alias gorila
- Homo sapiens neanderthalensis (FM866397.1) alias neandertálec

Sekvence analyzujte pomocou programu MEGA nasledovne:

- Kliknite File->Open a File/Session, How would you like to open this fasta file - kliknite Align, potom z menu Alignment exploreru vyberte Alignment -> Align by ClustalW.
- V Alignment exploreri vyberte v menu Data->Export alignment->Mega format a uložte.
- Zatvorte Alignment explorer (alignment už znova nemusíte ukladať)
- Analyzujte zarovnanie pomocou fylogenetickej analýzy s využitím bootstrappingu. Ak sedíte v lichom stĺpci, použite Phylogeny -> Construct/Test Maximum likelihood tree, ak sedíte v sudom stĺpci, použite Phylogeny -> Construct/Test Neighbor-Joining tree a nastavte parametre:
  - Test of phylogeny: Bootstrap method
  - No. of Bootstrap Replications: 1000
  - Model/Method: Kimura 2-parameter model
  - Gaps/Missing Data Treatment: Partial deletion

Ostatné parametre ponechajte implicitné. V TreeExploreri sa Vám zobrazí vygenerovaný fylogenetický strom a consensus strom.

- Majú obidva stromy rovnakú topológiu? Čo prezrádzajú bootstrapové hodnoty? Čo je najpodobnejšie človeku? Majú na základe získaného stromu neandertálec a človek dnešného typu priameho spoločného predka?
  - Stejnou topológiu nemajú, ML a NJ se liší v zařazení orangutana (na společné větvi se šimpanzem, nebo s člověkem/neandertálcem).
  - Bootstrapová hodnota ukazuje, kolikrát byl takový podstrom sestavený při převzorkování z X iterací - vynechání jednoho sloupce a sestavení stromu z upravených dat.
  - Nejpodobnější člověku je neandertálec, nejspíše přímého předka mají.