3. FYLOGENETICKÁ ANALÝZA ŠELEM NA ZÁKLADĚ MITOCHONDRIÁLNÍHO GENU COX1

Cíl úlohy

Rychlost a frekvence mutací mitochondriálního genomu jsou vhodné pro vyhodnocení genetických polymorfismů ve vztahu s fylogenetickým vývojem a taxonomií na úrovni řádů, čeledí a podčeledí. V následující úloze vycházejte z hypotézy, že všichni vybraní zástupci šelem měli společného předka. Tudíž tuto skupinu druhů budeme chápat jako monofyletickou. Cílem úlohy je vypracování fylogenetické analýzy na základě metody maximální věrohodnosti (Maximal Likelihood Method – ML). Prvním důležitým krokem je správně provést alignment výchozích sekvencí genu Cox1, který je frekventovaně používán nejen pro fylogenetické analýzy, ale je rovněž aplikován pro DNA barcoding systém zacílený na identifikaci druhů. Vzhledem k tomu, že Cox1 je strukturní gen kódující protein, je vhodné zvolit takový typ alignmentu, který pracuje s kodóny. Analýzy mtDNA vychází z předpokladu homoplazmie a haplotypové analýzy. Haplotypy je nezbytné identifikovat vhodným programem. Cílem této úlohy je naučit pracovat studenty s programovým balíkem MEGA X, který umožňuje provedení několika typů alignmentu, výběru vhodného substitučního modelu pro metodu ML a následnou konstrukci a editaci stromu – fylogramu.

Vstupní data

Sekvence mtDNAve formátu FASTA, které byly získány z databáze NCBI. Sekvence představují vlákna s orientací 5'-3'.

Potřebné bioinformatické nástroje

- BioEdit 7.2 (<u>https://bioedit.software.informer.com/7.2/</u>)
- > DNA Sequence Polymorphism v6.12.03 (http://www.ub.edu/dnasp/downloadTv6.html)
- MEGA X (<u>https://www.megasoftware.net/dload_win_gui</u>)
- NCBI (<u>https://www.ncbi.nlm.nih.gov/</u>)

Návod na řešení úlohy

Vstupními daty pro fylogenetickou analýzu jsou kompletní sekvence mitochondriálního genu *Cox1* u 35 druhů šelem. Sekvence představují vlákna DNA s orientací 5′- 3′, která byla získána z mezinárodní bioinformatické databáze NCBI (<u>https://www.ncbi.nlm.nih.gov/</u>). Jako tak zvaný outgroup genotyp použijte sekvenci *Cox1* genu ježury (*Tachyglossus aculeatus*). Vstupní data jsou uložena ve formátu FASTA v příloze této úlohy (soubor COX1 šelmy před alignmentem.fas)

🎽 BioEdit Sequence Alignment Editor - [E:\Z	iklady bioinformatiky\Fylogenetická analýza\COX1 šelmy před alignmentem.fas]	
🎾 File Edit Sequence Alignment Vie	v Accessory Application RNA World Wide Web Options Window Help	-
🖻 🖸		
📮 📇 Courier New 💌 11 💌 🖪	36 total sequences	
Mode: Select / Slide Selection: Position:	Sequence Mask: None Start Numbering Mask: None ruler at:	
🔒 I D I D 🔒 сл 🕂 🕾 🗱		
	la <u>de construction de la construction de construction de</u>	(**** <u>*</u>
	$1 \\ 10 \\ 20 \\ 30 \\ 40 \\ 50 \\ 60 \\ 70 \\ 80 \\ 90 \\ 100 \\ 110 \\ 120 \\ 130 \\ 140 \\ 150 \\ 160 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170 \\ 170$	1
Canis latrans		100GCI
Canis lupus familiaris		ACCGC
Canis lupus lupus		ACTIC
Crocuta crocuta		ACTIGC
Felis catus		ACTIGC
Hypers hypers		ACCGC
Lutra lutra	at gtt cataga cogat gat tatt cocacea costa a gat att go caccett tact cot tatt cost gat cost cacce cost caccet a caccet cacet cac	ACCGC
Lutra sumatrana	atgiteatagacegatgactatteteceacaaaceataagatattggcaccetttatteggtgeatgggeaggatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggaatggeeggee	ACCGC
Lycaon pictus	atgttcattaaccgatgactgtcctccaccaaggtatctggtatctgtatctattattggagcatgggctggactcgccttgagtccccccaggtcagccaggtagtcaagtcaggcatgatcaaatttattggagcatggactgaactaggtcagccaggtagtcaagttatcaatttattggagcatggactgaactaggtcagccaggtagtcaagttatcaatttattggagcatggactgaactaggtcaggtagtcaagttatcaggtagtcaggtagtcaagttatcgagtagtcaggtagtcaagttatcgagtagtcaggtagtcaggtagtcaggtagtcaggtagtcaggtagtcaggtagtcaggtagtcaggtagtcaggtagtcaggtagtcaggtagtcaggtagtcaggtagtgagtagtagtagtagtagtagtagtagtagtag	ACCGC
Lvnx canadensis	atgiticataaaccgctgattattttcaactaatcacaaagatattggcactctctaccttttatttggtgcctgggccggatggtcatgctctcagccccggaccggactggtcaacctggtcaacctggtacattattaggagggcgaccagatttacaatgatcgtca	ACCGC
Lynx lynx	ATGTTCATAAACCGCTGACTAATTCAACTAATCACAAAGATATTGGCACCCTCTACCTTTTTTGGTGCCTGGGCCGGTATGGTGGGAACTGCTCCGAGCCCGAACTAGGTCAACCTGGCACCGCACCTAGGTCAGGACTAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAG	ACTGC
Lynx pardinus	atgitcataaaccgctgactaatcataatcataatcgtaatcggcactctctaccttattggcgcccgggtatagggactgctctctcagcctccgggccgaactgggtcaacctggtcaggtcaggtcaggtcaggtcaggtcagatcggatcggatcggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcaggtcagg	ACTGC
Lynx rufus	atgitcataaaccgctgactattttcaactaatcacaaagatatcggtactctttatctgtgtgccctgggcccggtatggtggggactccctagtcccgaactgggccaacctgggccaacctgggccaactactgggactagatcagatttacaatgatcagatttacaatgatcgggactgaccctgacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggactgaccctgacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggccaacctgggcaacctgggcaacctgggcaacctgggcaacctgggccaacctgggccaacctgggcaacctggggggaatggatcgggggggg	4CCGC
Martes flavigula	argitcataaarcgatgattattctccacaaagcaccacggcaccctttacctggcgccatgggcgcatgggcgctgcctattaagcccgaattgggcgcgactggccctggggagatgaccagatttatagtgggcactggcctattaagcctattaatgggggggg	ACCGC
Martes martes	argitcataaarcgatgattaitcicacaaaacacacggcactcititacittgccccatgagccggaataggccactgcattgagccattgcgtaitaartcgcgctattaggccactgcctgcattaggccactgcctattggatgaccaaattitaitggcgcactgcctattggatgaccaaattitaitggcgcactgactggccactggatggatgaccaaattitaitgggagatggatggatggatggatggatgga	1CCGC
Martes zibellina	argitcataatcGarGattattCtcCacAaatCacAaaGaCatCGGCAcCCTTTACTTGGCGCATGAGCCGGGATAGGGCACTGCATTGATCGCCCTGAATTGGTCACCTGGGCCCTCACCTGGGACGACGACGACGACGACGACGACGACGACGACGACGA	ACCGC
Meles anakuma		ACAGC
Meles leucurus		JCAGC
Meles meles		ACHCC
Mustela erminea		ACTGC
Mustela nivalis		ACTIGC
Mustela sibirica		ACTIGC
Panthera leo		ACCGC
Panthera pardus	atgitcataaaccgctgactattttcaaccaatcacaaagatattggaactctttacctatttggtgcctgggctggggatggtcctctcagtctcttaatccgggccgaactgggcacgaccacctgggacgaccaactttattataatgtagtcgtt	ACCGC
Parahvaena brunnea	argiteataaccegetgaetattteetaetaaccaetaagaeattgaecetttaecteetatttggtgeetgaggaetgegeetgeaetgegetgeaettggetgaeetgegetgeaettggetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgegetgeaetgeaetgegetgeaetgeaetgeaetgeaetgeaetgegetgeaetgegetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetgeaetg	ACCGC
Puma concolor	atgttcataaatcgctgactgttttcaactaatcataagatattggcactctttaccttctatttggtgcccgaactgctcgtaggaactgctccttagtcccgggccgaactaggtcaacctggcactactaggagatgatcaaatttataatgtggtcgtc	ACTGC
Tachyglossus aculeatus	atgttcattaatcgctgactaattttcaactaaccataagatattggtaccctctatcttctattcggtgcatggcgcacagccctcagtattctcattcgatccgaattaggccaaccaggctccctcttaggtgatgatcaaatttataacgttatcgtc2	ACAGC
Ursus americanus	atgitcataaatcgatgactgitctctacaaatcataaggatattggcactctttaccattctgttcggtgatggtgatggtactgctctcgcccttttaaggccggggtcaggcagg	ACTGC
Ursus arctos	atgitcataagccggtgattattctctacgaaccataaggacattgccgtgatcaggccggaatagggccctggccctcagccctttaggccggggcagggcaggggcggggtcaggggatgatcagatctacaatgtggggatgatcagatctacaatgtggggatgatcagatctacaatgtggggatgatcagatctacaatgtggggatgatcagatctacaatgtggggatgatcagatctacaatgtggggatgatcagatctacaatgtggggatgatcagatctacaatgtggggatgatcagatctacaatgtggggatgatcagatctagatctacaatgtggggatgatcagatgatcggggatgatcagatctacaatgtggggatgatcagatcagatctacaatgtggggatgatgatgatgatgatgatgatgatgatga	ACTGC
Ursus maritimus	atgitcataaaccggtgattattcitccgaacccgtaaagacattggcgcgatcatcggccatgagcccggaatagtgggcgctcagcccttitaattcgtgggcggggcg	ACTGC
Ursus thibetanus thibetanus	atgitcataaaccgatgactattcitcacaaatcataaagatattggcactctctattctactatcggtgcatgagccggaatagtactgctctccgcccttttgattcgtgccggaactaggccgggatgatcagatctacaatgtagtcgtgtcgtat	ACTGC
Vulpes lagopus	prefileateateateateateateateateateateateateate	ACCGC
Vulpes vulpes		ACCGC
Vulpes zerda	INTETTLATTAATUGATGATTATTUTUTATTAGUCAGAAGAATTAGGCGACGATCAGGCATTAGGCAGTGGCCCTAAGGCCGAATTAGGCGGCGCCCCTTATTAGGGGGGGG	SCCPC

2. Pro fylogenetickou analýzu použijte program MEGA X. Prvním krokem je zadání souboru se vstupními daty. Jedná se o soubor ve formátu FASTA, který byl představen na předchozím obrázku. Otevřete tento soubor z umístění na vašem počítači.

M	Molecul	ar Evolutionar	y Genetics A	Analysis								_	
File	Analy	sis Help											
-	Open A	File/Session		Ctrl+0									
	Open a	Recently Use	d File		II.	π	任			(Ξ)		Θ	•
	Edit a T	ext File			ANCE	DIVERSITY	PHYLOGENY	USER TREE	ANCESTORS	SELECTION	RATES	CLOCKS	DIAGNOSE
	Conver	t File Format f	to MEGA										NEXUS
+	Printer	Setup											
	Quit M	EGA		Ctrl+Q									(CB)
													TIMETREE
													·C.
													DATAMONKEY
REC	ENT PUBL	ICATIONS											
(0		33				٢	()	٥		0	ANALYZE	Γ ×1
HEL	.P DOCS	EXAMPLES	CITATION	REPO	RT BUG	UPDATES	MEGA LINKS	TOOLBAR	PREFERENCES			PROTOTYPE	

 Před vlastní analýzou je nutné provést mnohonásobné porovnávání analyzovaných sekvencí, pro které se mnohem častěji používá anglický termín Multiple Alignment. Po zadání vstupních dat je nutné potvrdit, že s daty budeme provádět právě tuto proceduru. Zadejte proto červeně onačenou možnost Align.

Molecul	ar Evolutionary	Genetics An	alysis				and and	26 28 12 % 4		_		×
File Analy	sis Help											
ALIGN	DATA	MODELS	DISTANCE		PHYLOGENY	USER TREE			RATES			OSE
RECENT PUBL	ICATIONS							How would yo Analy Align	u like to o ze or Align File Analyze	× e?		EE
HELP DOCS	EXAMPLES	JJ CITATION	REPORT BUG	UPDATES	MEGA LINKS	TOOLBAR	PREFERENCE	s		ANALYZE	ſ×	1

MX: Alignment Ex	plorer (COX	1 šelmy př	ed alignm	nentem.fas)													_		\times
Data Edit Search	n Alignme	nt Web	Seque	ncer Dis	play Hel	р														
	2 📭	T W	6 =	⇒ +	s 🗖	⊁∣	Ъ Х	Ø _G	+ 🔁	•	▶	Q	°,	٩						
DNA Sequences Tra	nslated Prote	ein Sequen	ces																	
Species/Abbrv	* * * * *	* *	* *	* * * *	* * *	* *	* * 1	* * *	* * *	* *	* *	* *	* *	*	* *	*	* *	* * *	* * *	* * ^
1. Canis_latrans	A T G T T	CATTA	ACC	ATGA	тт <mark>с</mark> тт (стст	АСТА	АТС	ACAA	A G A	T <mark>A</mark> T	TGGT	ACT	тт	TAT	СТАС	Т G Т	TTGO	G A G C	ATG
2. Canis_lupus_familiari	s <mark>A</mark> TGTT	САТТА	A C C <mark>G</mark>	G A T G A	с т <mark>с</mark> тт (стсс	АСТА	A A T C	A C A A	GGA	T <mark>A</mark> T	T G G T	ACT	• т т <mark>и</mark>	TAC	T T A C	СТАТ	тт <mark>сс</mark>	G A G C	A T G
3. Canis_lupus_lupus	A T G T T	CATTA	A C C <mark>G</mark>	GATGA	тт <mark>с</mark> тт (стсс	АСТА	A A T C	ACAA	GGA	TAT	T G G T	ACT	ТТ 🖊	ТАС	T T A C	TAT	тт <mark>сс</mark>	G A G C	A T G
4. Crocuta_crocuta	A T G T T	CATAA	A C C <mark>G</mark>	C T G A	тт <mark>а</mark> тт:	т т с т	ACCA	A A C C	ACAA	A <mark>G</mark> A	T <mark>A</mark> T	C G G C	ACO	сто	TAT	стсс	TAT	TT G G	з т <mark>с</mark> с	CTG
5. Felis_catus	A T G T T	CATAA	A C C <mark>G</mark>	GTGA	с <mark>т а</mark> т т :	Т Т С А		A A T C	ACAA	A <mark>G</mark> A	T <mark>A</mark> T	TGGT	ACT	с т 1	TAC	с ттт	T T <mark>a</mark> T	Т <mark>С</mark> СС	G T G C	CTG
6. Felis_silvestris	ATGTT			GTTGA	С Т А Т Т .	Т Т С А	АСТА		ACAA	AGA	TAT	T G G T	ACT	С Т 1	TAC	сттт	T T A T	T <mark>C</mark> GO	G T G C	CTG
7. Hyaena_hyaena	ATGTT			GCTGA	T T <mark>A</mark> T T '	т т с т	ACTA	ACC	A C A A	AGA	CAT	TGGC	ACO	стт	TAC	СТТС		TTGO	G T G C	TTG
8. Lutra_lutra	ATGTT		A C C G	GATGA	T T <mark>A</mark> T T (стсс	A C G A	A A C C	ATAA	A G A	TAT	TGGC		: C T T	TAC	стто		т <mark>с</mark> сс	G T G C	ATG
9. Lutra_sumatrana	ATGTT		A C C G	GATGA	CTATTO	C T C C	ACAA	1 A C C		A G A	TAT	TGG	ACO	: C T T	TAC		T T A T	т <mark>с</mark> сс	G T G C	ATG
10. Lycaon_pictus	ATGTT	CATTA	A C C <mark>G</mark>	GATGA	CTGTT	стсс	ACTA		ACAA	A G A	TAT	TGGT	ACT	СТС	TAT		T T A T	TTGO	G A G C	ATG
11. Lynx_canadensis	ATGTT	CATAA	A C C G	G C T G A	T T <mark>A</mark> T T '	Т Т С А			ACAA	AGA	TAT	TGGC		СТС	TAC		TAT	TTGO	G T G C	CTG
12. Lynx_lynx		CATAA	ACCO	G C T G A	C T A T T '					AGA	TAT	TGGC		сто	TAC		T T A T	TTGO	G T G C	CTG
13. Lynx_pardinus		CATAA		G C T G A						AGA	TAT		ACT	сто	TAC	СТТТ	TAT	TTGO	3 C G C	CTG
14. Lynx_rufus	ATGTT	CATAA		GCTGA					ACAA	AGA	TAT		AC		TAT			TCGC	G T G C	CTG
15. Martes_flavigula	ATGTT	САТАА	ATCO	GATGA	TTATT	стсс	ACAA	ATC	ACAA	AGA	TAT	CGGC	ACO	СТТ	TAC			TTGO	3 C G C	ATG
16. Martes_martes	ATGTT	САТАА	ATCO	GATGA	TTATT	стсс	ACAA		ACAA		CAT		AC	СТІ	TAC			TTGO	G C G C	ATG
17. Martes_zibellina	ATGTT	САТАА	ATCO	GATGA		стсс	ACAA	ATC	ACAA	AGA	CAT	CGGC	ACO	СТТ	TAC			TTGO	G C G C	ATG
18. Meles_anakuma	ATGTT	САТАА		GATGA	TTATT	гтсс	ACAA	ATC	ATAA	AGA	TAT	CGGC			TAC			TTGO	GTGC	ATG
19. Meles_leucurus	ATGTT	CATAA	ATCO	GATGA	TTATT	гтсс	ACAA			AGA	TAT		ACA		TAC			TTGO	GTGC	ATG
20. Meles_meles	AIGII	CATAA	ATCO	GATGA			ACAA			AGA			ACO						3160	ATG
21. Mustela_erminea	AIGII	CGTAA	ATCO	ATGA			ACTA		ACAA		CAT		ACO						3160	ATG
22. Mustela_nivalis	ATOTT	CATTA		ATGA			ACTA	ATC	ACAA	AGA			ACC		TAC			TTO		ATG
23. Mustela_putorius	ATOTT						ACT	ATC	ACAA	AGA	CAT		ACC		TAC			TTC		ATC
24. musicia_siomica	ATGTT	CATAA							ACAA		CAT			СТ	TAC			TTC		CTG
26. Panthera_leo	ATCTT	CATAA		ACTOA		ТТСА			ACAA		TAT			С Т Т	TAC	СТТС		TTO		CTC
27 Parahyaana brunos	ATGTT	CATAA				ттст								СТТ	TAC			TTCC		CTG
227. Paranyaena_prunne	ATCTT	CATAA					ACT				TAT				TAC			TTC		CTC
20. Puna_concolor	ATGTT	CATTA				ТТСА					TAT	TGGT		СТС	TAT			TCCC		ATG
20. Hreue americanus	ATGTT	CATAA		ATGA						AGA					TAC	CTTC	TOT	TCCC		AT GY
<	AIG	CATAA		AIGA			ACAA	A		AGA	A	000	AU		AC				9 0 0	>
Site #			- •	with	O w/o g	Japs														

4. Vstupní data, která byla zobrazena na začátku řešení úlohy v programu BioEdit, jsou načtena do programu MEGA X.

5. Pro alignment sekvencí použijte algoritmus MUSCLE a to variantu, která porovnává sekvence na základě kodónů. Tento algoritmus je možné použít, protože se jedná o kódující sekvence, kdy první triplet kóduje první aminokyselinu. Jedná o sekvence mitochondriálního genu, který je tvořen pouze exonovými (kódujícími) oblastmi. Parametry analýzy, které musíte zvolit, jsou vyznačeny červenými rámečky.

MX: Alignment Explor	rer (COX1 šelmy před alignmenten	n.fas)		- 🗆 X
Data Edit Search	Alignment Web Sequencer	Display	Help	
1 = 8 👑 🗊	Align by ClustalW		□ X I X 4 + ₽ 4 ▶ Q 9 9 9.	
	Align by ClustalW (Codons)			
DIVA Sequences Transi	Align by MUSCLE			
Species/Abbry	Alian by MUSCLE (Codons)		TTO TO TA CTAAT CAAAAAAAAAAATATTO OTAC TTTATATATATATATATATATATATATATATATATA	
2 Canis Junus familiaris A		_	TTCTCCACTAATCACAAAGATATTGGTACTTTATATCTACTGTTT	GGAGCATG
3. Canis lupus lupus A	Mark/Unmark Site	Ctrl+M	TTC TC CAC TAATCACAAGGATATTGGTAC TTTATACTTACTACTATT	GGAGCATG
4. Crocuta_crocuta A	Alian Marked Sites	Chelul	T T T T C T A C C A A C C A C A A A G A T A T C G G C A C C C T C T A T C T C C T A T T T	<u> </u>
5. Felis_catus A		CUI+L	TTTTCAACTAATCACAAAGATATTGGTACTCTTTACCTTTTATTC	G G T G <mark>C C T</mark> G
6. Felis_silvestris	Unmark All Sites		TTTTCAACTAATCACAAAGATATTGGTACTCTTTACCTTTTATTC	G G T G <mark>C C</mark> T G
7. Hyaena_hyaena A		-	TTTTCTACTAACCACAAAGACATTGGCACCCTTTACCTTCTATTT	G G T G <mark>C </mark> T T G
8. Lutra_lutra A	Delete Gap-Only Sites		T T C T C C A C G A A C C A T A A A G A T A T T G G C A C C C T T T A C C T T C T A T T C	G G T G C A T G
9. Lutra_sumatrana A	 Auto-Fill Gaps 		TTCTCCACAAACCATAAAGATATTGGCACCCTTTACCTTTATTC	GGTGCATG
10. Lycaon_pictus A			TTC TC CAC TAATCACAAAGATATTGGTACTCTGTATCTATTATTT	G G A G C A T G
11. Lynx_canadensis A	I G T T C A T A A A C C G C T		TTTTCAACTAATCACAAAGATATTGGCACTCTCTCCCTTTTATTT	
13 Lynx pardinus	IGTTCATAAACCGCT	ЗАСТА	TTTTCAACTAATCATAAAGATATCGGCACTCTCTACCTTTTATTT	
14. Lynx rufus A	GTTCATAAACCGCT	GACTA	T T T T C A A C T A A T C A C A A A G A T A T C G G T A C T C T T T A T C T T C T A T T C	GGTGCCTG
15. Martes_flavigula A	I G T T C A T A A A T C G A T C	GATTA	T T C T C C A C A A A T C A C A A A G A T A T C G G C A C C C T T T A C C T T T T A T T T	GGCGCATG
16. Martes_martes A	I G T T C A T A A A T C G A T	<mark>З А Т Т А</mark>	T T C T C C A C A A A T C A C A A G A C A T C G G C A C T C T T T A C C T T T T A T T T	G G C G C A T G
17. Martes_zibellina A	I G T T C A T A A A T C G A T (G A T T A	T T C T C C A C A A A T C A C A A A G A C A T C G G C A C C C T T T A C C T T T T A T T T	G G C G C A T G
18. Meles_anakuma A	I G T T C A T A A A T C G A T C	GATTA	TTTTCCACAAATCATAAAGATATCGGCACACTTTACCTTTTATTT	G G T G <mark>C</mark> A T G
19. Meles_leucurus A	I G T T C A T A A A T C G A T (GATTA	TTTTCCACAAATCATAAAGATATCGGTACACTTTACCTTTTATTT	G G T G C A T G
20. Meles_meles A	I G T T C A T A A A T C G A T (G A C T A	TTTTCCACAAACCATAAAGATATTGGTACCCTTTACCTTCTATTT	G G T G C A T G
21. Mustela_erminea	I G T T C G T A A A T C G A T C	GATTA	T T C T C C A C T A A T C A C A A G A C A T C G G C A C C C T T T A T C T C T T A T T C	GGTGCATG
22. Mustela_nivalis A	I G T T C A T T A A T C G A T C	GATTA	TTTTCCACTAATCACAAAGACATCGGCACCCTTTACCTCTTATTT	GGTGCATG
23. Mustela_putorius A	I G T T C A T A A A T C G A T		TTO TO CAUTAATCACAAAGACATCGGCACCCTCTACCTCTATTT	GETECATE
25 Panthera len A	GTTCATAAACCGCT	ЗАСТА	TTTTCAACCAATCACAAAGACATTGGAACTCTTTACCTTTATTT	G T G C C T G
26. Panthera pardus A	I G T T C A T A A A C C G C T (ЗАСТА	T T T T C A A C C A A T C A C A A G A T A T T G G A A C T C T T T A C C T T C T A T T T	GGTGCCTG
27. Parahyaena_brunne A	I G T T C A T A A C C C G C T (ЗАСТА	T T T T C T A C T A A C C A C A A G A C A T T G G C A C C C T T T A C C T C C T A T T T	<mark>д д т д с с</mark> т д
28. Puma_concolor A	I G T T C A T A A A T C G C T (<u>аст</u> б	T T T T C A A C T A A T C A T A A A G A T A T T G G C A C T C T T T A C C T T C T A T T T	<mark>д д т д с с т д</mark>
29. Tachyglossus_acule A	I G T T C <mark>A T T A A T C G C T</mark> (ЗАСТА	T T T T C A A C T A A C C A T A A A G A T A T T G G T A C C C T C T A T C T T C T A T T C	G G T G C A T G
30. Ursus americanus A	I G T T C A T A A A T C G A T (GACTG	TTCTCTACAAATCATAAAGATATTGGCACTCTTTACCTTCTGTTC	GGTGCATG
				>
Site # 1	💼 💿 with	0	w/o gaps	

File Analysis Help Analysis Help ALIGN DATA MODELS DISTANCE DISTANCE DIVER MUSCLE Alignment Options DIAGNOSE Gap Open 280 Gap Open 280 Gap Extend 0.00 Help Addattattisgitating tatting tatt	Molecular Evolutionary Genetics Analysis		_		Styly
Image: Stance Image: Stance ALIGN DATA MODELS DISTANCE DISTANCE DISTANCE Option Setting GAP PENALTIES Gap Open Gap Den Humon Kernon MEMORY/IFERATIONS Memory in MB Que Max Iterations IC Max Iterations IC Addatatt GG CACCCT Addatatt GG CACCCCT Adda	File Analysis Help				
Option Setting Gap Option Gap Open Gap Open 230 Gap Appendence 230 MEMORY/ITERATIONS Accent and the appendence Max Memory in MB 2048 Genetic Code Standard Accent at the accent accen	ALIGN DATA MODELS DISTANCE DIVERS	MUSCLE Alignment Options	•;) (Ξ) (II) (Θ) ×	DIAGNOSE	i 4 1 Q Q Q Q Q
GAP PENALTIES Gap Open Gap Den Hydrophobicity Multiplier 120 MEMORY/TERATIONS TIMETREE Max Memory in MB 2048 Genetic Code Standard Cluster Method (Iterations 1.2) UPGMA Cluster Method (Other Iterations) UPGMA Min Diag Length (Lambda) 24 Mask Cat Tot Go Cat Cot Cat Cat Cat Cat Cat Cat Cat Cat Cat Ca		Option	Setting		
Gap Open -2.90 Gap Extend 0.00 Hydrophobicity Multiplier 1.20 MEMORY/ITERATIONS Image: Construction of the constr		GAP PENALTIES		NEXUS	x x x x x x x x x x x
Gap Extend 0.00 Hydrophobicity Multiplier 1,20 MEMORY/ITERATIONS TIMETREE Max Memory in MB 2048 Max Iterations 16 AGG AT AT T G G T AC T C T AGG AT AT T G G T AC T C T AGG AT AT T G G T AC T C T AGG AT AT T G G T AC T C T AGG AT AT T G G T AC T C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C T C T AGG AT AT T G G C A C T C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T AGG AT AT T G G C A C C C T <t< td=""><td></td><td>Gap Open</td><td>-2,90</td><td></td><td>AAGATATTGGTACTTT</td></t<>		Gap Open	-2,90		AAGATATTGGTACTTT
Hydrophobicity Multiplier 1,20 MEMORY/ITERATIONS Max Memory in MB 2048 Max Memory in MB 2048 16 ADVANCED OPTIONS ADVANCED OPTIONS Genetic Code Standard Cluster Method (Iterations),2) UPGMA Cluster Method (Other Iterations) UPGMA Cluster Method (Other Iterations) UPGMA Min Diag Length (Lambda) 24 RECENT PUBLICATIONS RECENT P		Gap Extend	0,00		AGGATATTGGTACTTT
MEMORY/ITERATIONS Max Memory in MB 2048 Max Memory in MB 2048 Max Memory in MB 16 Max Memory in MB 16 ADVANCED OPTIONS As Gat At TIGGTACTCT Ad Gat At TIGGTACTCT As Gat At TIG		Hydrophobicity Multiplier	1,20		A G G A T A T T G G T A C T T T
Max Memory in MB 2048 Max Memory in MB 2048 Max Iterations 16 AGATATTGGTACTCT AGATATTGGTACTCT AGATATTGGTACCCT AGATATTGGCACCCT AGATATCGCCACCCT AGATATCGGCACCCT AGATATCGCCCACCCT AGATATCGCCCCACCCT AGATATCGCCCCACCCT AGATATCGCCCACCCT AGATATCGCCCACCCT AGATATCGCCCACCCT AGATATCGCCCACCCT AGATATCGCCCCACCCT AGATATCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		MEMORY/ITERATIONS		TIMETREE	AAGATATCGGCACCCT
Max Iterations 16 ADVANCED OPTIONS Genetic Code Standard Cluster Method (Iterations 1,2) UPGMA Cluster Method (Other Iterations) UPGMA Cluster Method (Other Iterations) UPGMA Min Diag Length (Lambda) 24 RECENT PUBLICATIONS Image: Construction of the processing of the procesing of the processing of the processing of the procesing of the pr		Max Memory in MB	2048		AAGATATTGGTACTCT
ADVANCED OPTIONS Genetic Code Cluster Method (Iterations 1,2) Cluster Method (Other Iterations) Min Diag Length (Lambda) Z Help Reset X Cancel OK AGA TATT GG CAC CCT AAGA TATT CG CAC CCT AAGA CAT CG CCAC CCCT		Max Iterations	16		AAGACATTGGCACCCT
Genetic Code Standard A A G A T A T T G G C A C C C T A A G A T A T T G G C A C C C T A A G A T A T T G G T A C T C T A A G A T A T T G G C A C C C T A A G A T A T T G G C A C C C T A A G A T A T T G G C A C C C T A A G A T A T T G G C A C C C T A A G A T A T T G G C A C C C T A A G A T A T T G G C A C C C T A A G A T A T T G G C A C C C T A A G A T A T T G G C A C C C T A A G A T A T T G G C A C C C T A A G A T A T T G G C A C C C T A A G A T A T T G G C A C C C T A A G A T A T T G G C A C C C T A A G A T A T T G G C A C C C T A A G A T A T T C G G C A C C C T A A G A T A T T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A C A C A C A C A C A C A C		ADVANCED OPTIONS			AAGATATTGGCACCCT
Cluster Method (Iterations 1,2) Cluster Method (Other Iterations) Min Diag Length (Lambda) 24 RECENT PUBLICATIONS RECENT PUBLICATI		Genetic Code	Standard	ATAMONKEY	A A G A T A T T G G C A C C C T
Cluster Method (Other Iterations) Min Diag Length (Lambda) 24 RECENT PUBLICATIONS RECENT		Cluster Method (Iterations 1,2)	UPGMA		AAGATATTGGTACTCT
Min Diag Length (Lambda) 24 A A G A T A T C G G C A C T C T A A G A T A T C G G C A C T C T A A G A T A T C G G C A C T C T A A G A T A T C G G C A C T C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A		Cluster Method (Other Iterations)	UPGMA		
RECENT PUBLICATIONS (? Help Reset (Cancel (C) OK A A G A T A T C G G T A C T C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T		Min Diag Length (Lambda)	24		AAGATATCGGCACTCT
RECENT PUBLICATIONS					AAGATATCGGTACTCT
RECENT PUBLICATIONS A A G A C A T C G G C A C T C T A A G A C A T C G G C A C T C T A A G A C A T C G G C A C T C T A A G A C A T C G G C A C T C T A A G A C A T C G G C A C T C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C C C T A G A C A T C G G C A C		2 Help Reset			A A G A T A T C G G C A C C C T
A GACATCGGCACCCT A A A GACATCGGCACCCT A A GACATCGGCACCCT A A A A A A A A A A A A A A A A A A A	RECENT PUBLICATIONS	i Help Keset			AAGACATCGGCACTCT
A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A C A C T A A G A T A T C G G T A A T C G G T A A T C G G T A A T C G G T A A T C G G C A C C C T A A G A T A T C G G T A A T C G G C A C C C T A A G A T A T C G G T A A T C G G T A A T C G G T A A T C G G C A C C C T A A G A T A T C G G T A A T C G G C A C C C T A A G A T A T C G G T A A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A T A T C G G C A C C C T A A G A A G A A C A T C G G C					AAGACATCGGCACCCT
HELP DOCS EXAMPLES CITATION REPORT BUG UPDATES MEGALINKS TOOLBAR PREFERENCES) (*) (*) (M	AAGATATCGGTACACT
A A G A C A T C G G C A C C C T 23. Mustela_pulo 24. Mustela_pulo 24. Mustela_pulo 23. Mustela_pulo 23. Mustela_pulo 24. Mustela_pulo 25. Mustela_pulo 25. Mustela_pulo 26. Mustela_pulo 27. Mustela_pulo 27. Mustela_pulo 27. Mustela_pulo 28. Mustela_pulo 29. Mustela_pulo 29. Mustela_pulo 29. Mustela_pulo 20. Mustela_pulo 20. Mustela_pulo 20. Mustela_pulo 20. Mustela_pulo 20. Mustela_pulo 20. Mustela_pulo 21. Mustela_pulo 22. Mustela_pulo 23. Mustela_pulo 24. Mustela_pulo 25. Mustela_pulo 26. Mustela_pulo 27. Mustela_pulo 27. Mustela_pulo 27. Mustela_pulo 28. Mustela_pulo 29. Mustela_pulo 29. Mustela_pulo 29. Mustela_pulo 20. Mustela_pul	HELP DOCS EXAMPLES CITATION REPORT RUG UPDAT	TES MEGALINKS TOOLBAR PREE	PROTOTYPE		AAGATATTGGTACCCT
23. Mustela_putorius A TGTTCATTAATCGATGATTATTCTCCACTAATCACAAAGACATCGGCACCCT					A A G A C A T C G G C A C C C T
	23. Mustela_puto 24. Mustela sihiri	Iustela_nivalis AIGIICAI		AATCACA	AAGACATCGGCACCCT
	25. Panthera_leo	Iustela sibirica ATGTTCAT	A A T C G A T G A T T A T T C T C C A C I	AATCACA	AAGACATCGGCACCCT

6. Pro analýzu ponechejte všechny implicitně nastavené parametry. Alignment spusťte zvolením nabídky OK.

7. Před zahájením analýzy je nezbytné potvrdit, že při analýze budou kompletně vymazány tak zvané Gapy. Gapy jsou mezery, které algoritmus vytvoří porovnáváním sekvencí. Jedná se v podstatě o nukleotidy nebo skupiny nukleotidů, které algoritmus detekuje pouze u některých porovnávaných druhů. Jedná se o inzerčně – deleční typ polymorfismů. Sekvence způsobující vznik gapů se nacházejí pouze u některých srovnávaných druhů. Proto je nelze využít k hodnocení substitučních změn napříč celým spektrem hodnocených druhů. Proto jsou ve fylogenetických analýzách obvykle oblasti gapů kompletně odstraněny.



8. Pokud alignment proběhl, je nezbytné uložit jeho výsledek do vašeho počítače, a to ve formátu FASTA. Zvolte takové jméno souboru, aby bylo patrné, že představuje data po alignmentu.

Molecular Evolutionary Genetics Analysis	$ \square$ \times _{Styly}
File Analysis Help	
	M MX: Alignment Explorer (COX1 šelmy po alignmentu.fas) — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — — _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _
ALIGN DATA MODELS DISTANCE	🗋 Create New Ctrl+Alt+N 🍃 🔳 🏪 🔸 🗋 🛠 🌇 🛪 🖓 🕂 🄁 🔍 🕨 🔍 🤤 🤤 🤤
	InL Open es Open a Recently Used File ************************************
	Phylogenetic Analysis CTACAGGAGGGCCCTCCCCCATATCACACTTGAAGAACCGACCTATGTGCTATCAAAATAA CTACACGGATGTCCTCCCCCCATATCACACCTTCGAAGAACCAACC
	Save Session Ctrl+S TTACATGGATGTCCCCCCTCCATACCACACATTTGAAGAGCCAACTTATGTGTTATTAAAATAA Export Alignment MEGA Format CCTCCGTACCACACATTTGAAGAGCCAACTTATGTGTTATTAAAATAA
	DNA Sequences Protein Sequences C TA CA CA GA TA CCA CA CA CA TA CCA CA CA TA CCA A CA TA CA A A TA A C TA CA CA GA TA CA CA CA TA CCA CA TA CCA CA TA CCA CA CA CCA C
	UC Translate/Untranslate Genetic Code TTACATGGATGCCCTCCCATATCACACATTTGAAGAGCCGACCTATGTGCTATCAAAATAA
RECENT PUBLICATIONS	CTACATGGATGCCCTCCTCCATATCACACATTCGAAGAACCAACC
	Reverse TTGCATGGGTGTCCCCCTCCATACCACATTCGAAGAACCAACC
HELP DOCS EXAMPLES CITATION REPORT BUG	C T G C AC G G A T G T C C T C C C C C T A C C A T A C A T T T G A A G A A C C A G C T T A T G T T T C T C T A A A A T A A Quit C T A C A T G G A T G T C C C C C T C C C T A C C A C A T T T G A A G A G C C A A C T T A T G T A C T A T T A A A T A A
HELP DOCS EX	29. 1achygiossus_acuectar FGAGFGACTCCACGGATGTCCACCGCCTTACCACCCCTTGAAGAACCGGTATACATTAAAATTTAA 30. Ursus_americanus_CATTGAGTGGCTACATGGATGTCCTCCTCCTCCATATCACACATTTGAAGAACCCGCTACGTTACGTTACACATTAAAATAA 31. Ursus_arctosCATTGAATGACTACATGGATGCCCCTCCTCCATATCACACATTCGAAGAACCTGCCTACGTCACACTAAAATAA

9. Analyzované sekvence mitochondriálního genu *Cox1* můžeme označit jako haplotypy. Mitochondriální genom je charakteristický matroklinní dědičností a homoplazmií. Pro fylogenetickou analýzu je potřeba pracovat s takovým datovým souborem, který bude obsahovat každý haplotyp pouze jedenkrát. Vzhledem k tomu, že ve vstupních datech je každý zoologický druh zastoupen pouze jedním jedincem, můžeme předpokládat, že každý zoologický druh představuje jeden haplotyp a sekvence haplotypů se tudíž vzájemně odlišují. Nicméně v souboru hodnocených šelem se nacházejí dvě domestikované formy (pes domácí a kočka domácí) včetně jejich nedomestikovaných předchůdců. V těchto případech je mnohem větší pravděpodobnost, že shodný haplotyp by se mohl vyskytovat jak u volně žijící, tak i u domestikované formy. Z těchto důvodů je nutné použít haplotypovou analýzu, a to pomocí programu DNA Sequence Polymorphism v6.12.03.



10. Vstupními daty pro haplotypovou analýzu je soubor se sekvencemi, u kterých byl proveden alignment pomocí programu MEGA X a algoritmu Muscle s využitím kodónů. Tento soubor jste vytvořili a uložili ve formátu FASTA v přechozích krocích analýzy. Otevřete tento program v prostředí programu DNA Sequence Polymorphism v6.12.03. Po načtení dat vám program potvrdí, že bylo načteno všech 36 sekvencí. Pro identifikaci haplotypů zvolte nabídku Generate – Haplotype Data File



V dalším kroku haplotypové analýzy program podá informaci o parametrech vlastního hodnocení. Ponechejte všechny implicitně nastavené hodnoty.
 Program podává informaci o počtu hodnocených sekvencí (36), o hodnocené oblasti (1. – 642. nukleotid). Nepolymorfní oblasti nejsou hodnoceny.
 Oblasti s výskytem gapů nejsou rovněž hodnoceny. Analýza bude pokračovat po zvolení nabídky OK.

💆 DNA Sequence Polymorphism - šelmy cox1.nex	
File Data Display Analysis Overview Generate	íools Window Help
🖻 🗖 🗐 🖉 💶 🔚 👪 📰 🛝 🗵 🔋	
Hanlotyne/DNA Sequences Data File Ontions	×
hiplotype biox bequeices bita hiel options	
Data Set: All Included Sequences (n = 36)	▼
Region to Analyze	Invariable Sites
From site: 1 to: 642	© Removed
From site. 1 to. 042	C Included
Sites with gaps/missing	Generate
Not considered	NEXUS Haplotype Data File
C Considered	O Arlequin Haplotype List
C Only gaps/missing are considered	C Roehl Data File (Network software)
L	

	e lools Window Help
Haplotype/DNA Sequences Data File. Options	X
Save Haplotype Data from File: selmy cox1.r	nex X
← → × ↑ 🕞 > Tento počítač > H	KINGSTON (E:) > Základy bioinformatiky > Fylogenetická analýza v 🗗 🖉 Prohledat: Fylogenetická an
Uspořádat 🔻 Nová složka	
A Název	Datum změny Typ Velikost
Rychly pristup	Hledání neodpovídají žádné položky.
Stažené souh	······································
Bokumenty #	
Chrázky 🖈	
Amplikonové se	
Obrázky * Amplikonové se Fylogenetická ar	
Obrázky Amplikonové se Fylogenetická ar Hrubá sekvenačí	
 Obrázky * Amplikonové se Fylogenetická ar Hrubá sekvenačí Základy bioinfor 	
 Obrázky * Amplikonové se Fylogenetická ar Hrubá sekvenačí Základy bioinfor OneDrive 	
 Obrázky * Amplikonové se Fylogenetická ar Hrubá sekvenačí Základy bioinfor OneDrive Tento počítač 	
 Obrázky * Amplikonové se Fylogenetická ar Hrubá sekvenači Základy bioinfor OneDrive Tento počítač 	
 Obrázky * Amplikonové se Fylogenetická ar Hrubá sekvenači Základy bioinfor OneDrive Tento počítač Název souboru: selmy 	
 Obrázky * Amplikonové se Fylogenetická ar Hrubá sekvenačí Základy bioinfor OneDrive Tento počítač Název soubort šelmy Uložit jako typ: NEXUS Format (*.nex) 	

12. Výsledky haplotypové analýzy uložte do vašeho počítače ve formátu NEXUS.

13. V tuto chvíli je haplotypová analýzy kompletní. Byl detekovaný shodný počet haplotypů (36) jako je počet hodnocených sekvencí. Můžeme tudíž tvrdit, že všechny haplotypy, které budou použity pro fylogenetickou studii, se vzájemně sekvenčně odlišují. Haplotypová diverzita HD je tudíž rovna 1,00.

DNA Sequence Polymorphism - šelmy cox1.nex	
File Data Display Analysis Overview Generate Tools Window Help	
🖻 🗖 📓 🛃 🔢 👪 🏢 🔼 🗵 🔋	
🖬 Output of: šelmy cox1.nex 📃 💷	×
Haplotype/DNA Sequences Data File	
Haplotype/DNA Sequences Data File Input Data File: CV Weiny cox1nex Number of sequences: 05 Selected region: 1-642: Number of stes: 642 Total number of stes: 642 Total number of mable stes: 642 Stess with alignment gaps: not considered Number of haplotypes. h: 36 Haplotype diversity. Hd: 1.0000 Hap. 1: 1[] Hap. 2: 1 [2] Hap. 3: 1 [3] Hap. 7: 1 [7] Hap. 7: 1 [7] Hap. 7: 1 [7] Hap. 1: 1 [1] Hap. 1: 1 [1] Hap. 1: 1 [1] Hap. 2: 1 [2] Hap. 1: 1 [1] Hap. 4: 1 [4] Hap. 7: 1 [7] Hap. 1: 1 [1] Hap. 1: 1 [1] Hap. 1: 1 [1] Hap. 1: 1 [1] Hap. 2: 1 [2] Ha	
Hap_4: I [Hap_4] Hap_5: 1 [Hap_5] Hap_6: 1 [Hap_6] Hap_7: 1 [Hap_7] Hap_8: 1 [Hap_8] Hap_9: 1 [Hap_10] Hap_10: 1 [Hap_10] Hap_11: 1 [Hap_11] Hap_12: 1 [Hap_12] Hap_13: 1 [Hap_13] Hap_15: 1 [Hap_15] Hap_16: 1 [Hap_16]	~

14. Nyní opusťte program DNA Sequence Polymorphism v6.12.03. a vraťte se zpět do programu MEGA X, ve kterém budete pokračovat ve fylogenetické analýze. Existuje velké množství algoritmů a strategií pro fylogenetické analýzy a konstrukce fylogenetických stromů. Jednotlivé metody se liší nejen svou sofistikovaností, spolehlivostí a věrohodností výsledků, ale rovněž i náročností na hardwarové vybavení nebo čas, po který analýzy probíhají. Pro kurz Základů bioinformatiky byla zvolena metoda maximální věrohodnosti (Maximum Likelihood Method). Spusťte znovu program MEGA X a otevřete vámi uložený datový soubor po alignmentu ve formátu FASTA. Po výběru souboru zvolte, že data budete analyzovat. Alignment jste již provedli v předešlých krocích analýzy.



15. Program MEGA X nyní ověří, jakým způsobem je použitý pro označování jednotlivých pozic u hodnocených sekvencí. Myší označte, že se jedná o nukleotidové sekvence. U ostatních parametrů ponechejte implicitně nastavené hodnoty a zadejte OK.

Molecular Evolutionary Genetics Analysis -	- 🗆 X
File Analysis Help	
ALIGN DATA MODELS DISTANCE DIVERSITY PHYLOGENY USER TREE ANCESTORS SELECTION RATES CLOC	
Input Data Options	NEXUS
Nucleotide Sequences Missing Data Protein Sequences	C
Pairwise Distance Alignment Gap Identical Symbol .	TIMETREE
? Help X Cancel	DATAMONKEY
file:///C:/Users/vejl/AppData/Local/MEGAX/MEGA10_10180924-x86_f	
RECENT PUBLICATIONS	
Image: Constraint of the second se	

16. V dalším kroku analýzy budete vyzváni k potvrzení, zdali se jedná o sekvenci DNA kódující proteiny. *Cox1* gen je strukturní gen kódující protein, a proto tuto nabídku potvrďte.

File Analysis Help Image: the state	0)	\bigcirc
	0)	
ALIGN DATA MODELS DISTANCE DIVERSITI PHILOGENT USER TREE ANCESTORS SELECTION RATES C	LOCKS	DIAGNOSE
Confirmation × Protein-coding nucleotide sequence data? Ves No	-	NEXUS TIMETREE
	LYZE	



17. Následně vyberte typ genetického kódu. Analyzujete mitochondriální gen *Cox1* u šelem – obratlovců (Vertebrate Mitochondrial).

18. Nyní jsou data načtena a program MEGA X má zadané informace, jak bude k sekvencím přistupovat při fylogenetické analýze. Přítomnost dat v paměti počítači označuje zeleně označená ikona. Během analýzy nikdy neklikejte na červeně označenou ikonu, která provede vymazání všech zadaných dat a tím i ukončení všech probíhajících analýz.



19. V předchozí části postupu bylo zmíněno, že pro analýzu použijeme metodu maximální věrohodnosti (Maximum Likelihood Method). Prvním krokem je volba vhodného substitučního modelu, který popisuje, jakým způsobem probíhaly substituční mutace, které ve finále mohly diverzifikovat historického společného předka na recentní druhy hodnocených šelem. V hlavní liště zvolte nabídku MODELS a konkrétní nabídku Find Best DNA/Protein Models (ML). Pro kontrolu budete dotázáni, zdali chcete skutečně analyzovat aktuálně otevřený datový soubor.



20. V dalším kroku analýzy ponechte všechny implicitně nastavené hodnoty s výjimkou Select Codon Position. Zde nechte označenou pouze hodnotu 1st, a to proto, že víte, že v hodnocených sekvencích představuje první kodon skutečně první 3 nukleotidy. Parametr Number of Threads se nastavuje automaticky podle počtu jader procesoru ve vašem počítači. Po zadání nabídky OK bude analýza pokračovat.



21. Nyní bude program testovat a hodnotit jednotlivé substituční modely (24)

Progress Details Status/Options Run Status Start time 22.01.2021 1 Status Making init Thread-2 HKY + 1 Thread-1 TN93 + G + < Analysis Options Tree to Use Statistical Substitution Mod	4:14:50		🗶 Stop	
Details Status/Options Run Status Start time 22.01.2021 1 Status Making init Thread-2 HKY + I Thread-1 TN93 + G + < Analysis Options Tree to Use Statistical Substitution Mod Su	4:14:50 ial tree		X Stop	
Status/Options Run Status Start time 22.01.2021 1 Status Making init Thread-2 HKY + I Thread-1 TN93 + G + Analysis Options Tree to Use Statistical Substitution Mod	4:14:50 ial tree			
Run Status Start time 22.01.2021 1 Status Making init Thread-2 HKY + 1 Thread-1 TN93 + G + Analysis Options Tree to Use Statistical Substitution Moo	4:14:50 ial tree			
Start time 22.01.2021 1 Status Making init Thread-2 HKY + 1 Thread-1 TN93 + G + < Analysis Options Tree to Use Statistical Substitution Moo	4:14:50 ial tree			-
Status Making init Thread-2 HKY + I Thread-1 TN93 + G + Analysis Options Tree to Use Statistical Substitution Mode Substitution Mode Substitution Substitution Mode Substitution	ial tree			
Thread-2 HKY + 1 Thread-1 TN93 + G + Analysis Options Tree to Use Statistical Substitution Moo				
Thread-1 TN93 + G + Analysis Options Tree to Use Statistical Substitution Moo				_
Analysis Options Tree to Use Statistical Substitution Mode	l			—,
Analysis Options Tree to Use Statistical Substitution Moo				>
Tree to Use Statistical Substitution Mod				
Statistical Substitution Mod		:	Automatic (Neighbor-joining tree)	^
Substitution Mod	Method	:	Maximum Likelihood	
	iel 15 Type		Nucleotide	
Data Subset to N	Jse	•	hadreoblac	
Gaps/Missing	y Data Treatment	:	Complete deletion	
Select Codor	n Positions	:	1st,2nd,3rd,Non-Coding	
Branch Swap	Filter	:	None	
System Resource				
Number of Th	Usage	:	2	

22. Výsledky hodnocení modelů zobrazí program tabulkovou formou, kdy nejvěrohodnější substituční model je vybrán na základě BIC skóre. Jako nejlepší byl vybrán substituční model TN93+G. Bližší odkazy na tento model jsou uvedeny pod tabulkou.

File Falls Vi	n Expert: Find Bi	est-Fit Substitutio	on iviodel (IVIL)																								_	U ^
	ew neip																											
Table Maximum Likelihood fits of 24 different nucleotide substitution models																												
lable. Max	Darameter	lihood fits o	1 24 differ	ent nucleo	tide su	(+C)	ition m	f(A)	#(T)	f(C)	f(C)		r(AC)	r(AC)	r(TA)		r(TC)	HCA)		r(CC)	dCA)		r(CC)					
	75	5205.813	4610 501	2234 441	(+) n/a	0.11	14.00	0.266	0.230	0.203	0.202	0.007	0.006	0.060	0.008	0 372	0.008	0.008	0.437	0.008	0.063	0.007	0.006					
TN02+C+I	76	5205,015	4019,001	2234,441	0.00	0,11	14,00	0,200	0,239	0,203	0,202	0,007	0,000	0,003	0,000	0,372	0,000	0,000	0,437	0,000	0,005	0,007	0,000					
070.0	70	5215,145	4021,022	-2234,193	0,09	0,12	14,01	0,200	0,239	0,203	0,292	0,007	0,000	0,072	0,008	0,309	0,009	0,008	0,434	0,009	0,000	0,007	0,000					
GIR+G	78	5223,486	4613,747	-2228,539	n/a	0,11	13,52	0,266	0,239	0,203	0,292	0,012	0,006	0,069	0,014	0,370	0,013	0,008	0,435	0,000	0,063	0,011	0,000					
K2+G+I	72	5229,676	4666,792	-2261,111	0,46	0,20	16,61	0,250	0,250	0,250	0,250	0,007	0,007	0,236	0,007	0,236	0,007	0,007	0,236	0,007	0,236	0,007	0,007					
GTR+G+I	79	5232,649	4615,101	-2228,207	0,10	0,12	13,27	0,266	0,239	0,203	0,292	0,013	0,006	0,073	0,014	0,366	0,013	0,008	0,430	0,000	0,066	0,011	0,000					
T92+G+I	73	5240,116	4669,423	-2261,418	0,46	0,20	16,47	0,252	0,252	0,248	0,248	0,007	0,007	0,233	0,007	0,233	0,007	0,007	0,238	0,007	0,238	0,007	0,007					
K2+G	71	5242,694	4687,621	-2272,533	n/a	0,11	16,57	0,250	0,250	0,250	0,250	0,007	0,007	0,236	0,007	0,236	0,007	0,007	0,236	0,007	0,236	0,007	0,007					
TN93+I	75	5252,700	4666,388	-2257,885	0,74	n/a	13,41	0,266	0,239	0,203	0,292	0,008	0,007	0,078	0,009	0,361	0,009	0,009	0,424	0,009	0,072	0,008	0,007					
T92+G	72	5252,916	4690,033	-2272,731	n/a	0,11	16,46	0,252	0,252	0,248	0,248	0,007	0,007	0,233	0,007	0,233	0,007	0,007	0,238	0,007	0,238	0,007	0,007					
GTR+I	78	5269,655	4659,916	-2251,623	0,74	n/a	11,45	0,266	0,239	0,203	0,292	0,016	0,005	0,078	0,018	0,357	0,015	0,007	0,420	0,000	0,071	0,012	0,000					
HKY+G+I	75	5279,423	4693,111	-2271.246	0.47	0.20	17.86	0.266	0.239	0.203	0.292	0.006	0.006	0.276	0.007	0.192	0.008	0.007	0.226	0.008	0.252	0.006	0.006					
HKY+G	74	5294 997	4716 494	-2283 946	n/a	0 11	17 54	0 266	0 239	0 203	0 292	0 007	0 006	0 276	0 007	0 192	0 008	0 007	0 226	0 008	0 251	0 007	0 006					
K2+I	71	5311 368	4756 295	-2306.870	0.78	n/a	12.87	0.250	0.250	0.250	0.250	0,009	0,009	0.232	0,009	0.232	0.009	0,009	0 232	0,009	0.232	0,009	0,009					
T02+1	72	5320.801	4757 018	2306 674	0.78	n/a	12,01	0.250	0.250	0.248	0.248	0,000	0,000	0,202	0,000	0,202	0,000	0,000	0,202	0,000	0.234	0,000	0,000					
19271	74	5320,001	4705 500	-2300,074	0,70	11/a	12,01	0,252	0,252	0,240	0,240	0,009	0,009	0,230	0,009	0,230	0,009	0,009	0,234	0,009	0,234	0,009	0,009					
HKY+I	74	5374,023	4795,520	-2323,459	0,78	n/a	12,97	0,200	0,239	0,203	0,292	0,009	0,007	0,270	0,010	0,188	0,011	0,010	0,221	0,011	0,240	0,009	0,007					
1N93	74	5750,357	51/1,854	-2511,626	n/a	n/a	10,50	0,266	0,239	0,203	0,292	0,010	0,008	0,080	0,011	0,352	0,012	0,011	0,413	0,012	0,073	0,010	0,008					
GTR	77	5782,689	5180,759	-2513,053	n/a	n/a	7,43	0,266	0,239	0,203	0,292	0,012	0,014	0,081	0,013	0,337	0,016	0,018	0,396	0,016	0,074	0,013	0,011					
JC+G+I	71	5783,341	5228,268	-2542,857	0,43	0,23	0,50	0,250	0,250	0,250	0,250	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083					
JC+G	70	5790,805	5243,542	-2551,501	n/a	0,12	0,50	0,250	0,250	0,250	0,250	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083					
JC+I	70	5838,838	5291,576	-2575,518	0,77	n/a	0,50	0,250	0,250	0,250	0,250	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083					
K2	70	5900,870	5353,608	-2606,534	n/a	n/a	10,33	0,250	0,250	0,250	0,250	0,011	0,011	0,228	0,011	0,228	0,011	0,011	0,228	0,011	0,228	0,011	0,011					
T92	71	5908,788	5353,715	-2605,580	n/a	n/a	10,32	0,252	0,252	0,248	0,248	0,011	0,011	0,226	0,011	0,226	0,011	0,011	0,230	0,011	0,230	0,011	0,011					
НКҮ	73	5959,334	5388,641	-2621,028	n/a	n/a	10,28	0,266	0,239	0,203	0,292	0,011	0,009	0,265	0,012	0,185	0,013	0,012	0,217	0,013	0,242	0,011	0,009					
JC	69	6393,253	5853,801	-2857.639	n/a	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083					
NOTE		Dist DIO and	(Dece)	2001,000			2,00	0,200	0,200	0,200	0,200	2,300	0,000	5,500	0,000	3,500	0,000	(1)	0,000	0,000	0,000	0,000	0,000	 due (lat) a	 		in alcolin a la	and the second se

NOTE -- Models with the lowest BIC scores (Bayesian Information Criterion) are considered to describe the substitution pattern the best. For each model, AICc value (Akaike Information Criterion, corrected), Maximum Likelihood value (*inL*), and the number of parameters (including branch lengths) are also presented (1). Mon-uniformiting of evolutionary rates among sites may be modeled by using a discrete Gamma distribution (+G) with 5 rate categories and by assuming that a certain fraction of sites are evolutionarily invariable (+I). Whenever and distribution may be modeled by using a discrete Gamma shape parameter and/or the estimated fraction of invariant sites are shown. Assumed or estimated values of transition/transversion bias (*R*) are shown for each model, as well. They are followed by nucleotide frequencies (*f*) and rates of base substitutions (*r*) for each nucleotide pair. Relative values of instantaneous *r* should be considered when evaluating them. For simplicity, sum of *r* values is made equal to 1 for each model. For estimating ML values, a tree topology was automatically computed. This analysis involved 36 nucleotide sequences. Codon positions included were 1st. All positions containing gaps and missing data were eliminated (complete deletion option). There were a total of 514 positions in the final dataset. Evolutionary analyses were conducted in MEGA X [2].

Abbreviations: TR: General Time Reversible; HKY: Hasegawa-Kishino-Yano; TN93: Tamura-Nei; T92: Tamura 3-parameter; K2: Kimura 2-parameter; JC: Jukes-Cantor./div-1. Nei M. and Kumes S. (2000). Molecular Evolution and Phylogenetics. Oxford University Press, New York. 2. Kumar S., Stecher G., Li M., Knyaz C., and Tamura K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35:1547-1549. 23. Po nalezení optimálního substitučního modelu přistupte ke konstrukci fylogenetického stromu. Na hlavní liště zvolte funkci PHYLOGENY a v další nabídce zvolte vybranou metodu Construct/Test Maximum Likelihood Tree.



24. V tuto chvíli jste dospěli k poslední fázi nastavení parametrů vlastní fylogenetické analýzy. Program MEGA X implicitně nastavuje test stromu, respektive test zařazení jednotlivých druhů do větví pomocí tak zvané Bootstrap metody, která se implicitně opakuje 500x. Zde doporučuji ponechat implicitní nastavení 500x. Je potřeba si uvědomit, že právě nastavení hodnoty opakování výrazně zpomaluje celou analýzu a v případě nedostatečného hardwarového vybavení bude celý chod počítače výrazně zpomalen a analýza může trvat až několik hodin. Pokud nastanou tyto problémy, doporučuji při opakovaném spuštění analýzy hodnotu opakování snížit například na 200.

V červeně označených polích je provedeno správné nastavení substitučního modelu TN93+G. Tyto parametry je nutné manuálně nastavit. U parametru Select Codon Position zkontrolujte, zda je správně nastavený 1. kodón (1st). Vlastní konstrukci stromu zahájíte stisknutím nabídky OK.

Molecul	ar Evolutionary /sis Help	/ Genetics Ar	alysis							_		×	· 17 · 1
ALIGN	DATA	MODELS	DISTANCE		PHYLOG	GENY USER TREE			N RATES			DSE	
T?	X Close Data					MX: Analysis Pret Phylogeny Recor	ferences			Catting	NEVUC		
									_	Setting		-	-
							Statistical M	lethod \rightarrow	Maximum Like	lihood			- 1
							PHYLOGEN	Y TEST					
							Test of Phyl	ogeny 🕂	Bootstrap meth	hod			
						No. of Bo	otstrap Replic	ations \rightarrow	500				
						SU	BSTITUTION N	NODEL					[
							Substitution	s Type $ ightarrow$	Nucleotide				
							Genetic Code	\cdot Table \rightarrow	Not Applicable	<u> </u>			- 1
RECENT PUB	LICATIONS						Model/M	lethod →	Tamura-Nei m	odel			_
					-	R/	ATES AND PAT	TERNS					
		(33			6	No of Discusto	Rates amon	g Sites –	Gamma Distrib	uted (G)			
HELP DOCS	EXAMPLES	CITATION	REPORT BUG	UPDATES	MEGA L	IND OF DISCIELE			5				
	Data	_			_	Gaps/Mis	sing Data Trea	tment \rightarrow	Complete dele	tion			
	Data					Site	Coverage Cuto	off (%) →	Not Applicable	2			- F
						Sel	ect Codon Po:	sitions \rightarrow	🗹 1st) 🗆 2n	d 🗌 3rd	Nonco	oding Si	tes
						TREE I	NFERENCE OP	TIONS					
						N	1L Heuristic M	lethod $ ightarrow$	Subtree-Prunir	ng-Regrafting	- Fast (SP	PR level	3) TI
0							Initial Tree f	for ML $ ightarrow$	Make initial tre	e automatica	lly (Defau	ılt - NJ/	BioNJ)
							Initial Tr	ee File \rightarrow	Not Applicable	2			
							Branch Swap	m Filter ightarrow	None				
						SYSTE	M RESOURCE U	JSAGE					
							Number of T	hreads \rightarrow	2				
							? Help	×	Cancel	⊘ ok			_

25. V tuto chvíli bude program zpracovávat data a konstruovat větve fylogenetického stromu tak, aby zařazení druhů do jednotlivých větví bylo provedeno na základě maximální věrohodnosti. Analýza bude probíhat v závislosti na hardwarovém vybavení několik desítek minut až hodin. Během analýzy budou na monitoru zobrazeny parametry, které byly pro tvorbu stromu nastaveny.

chcete udBat	Vei Power 10 store									
Molecular Evolutionary Ger	MX: Progress									
File Analysis Help	Progress									
\sim										
	Setting site coverage	(+)								
ALIGN DATA M	Details 🗶 Stop	DIAGNOSE								
	Status/Options									
Data	Run Status									
	Start time 22.01.2021 14:36:40									
	Status Bootstrapping Tree									
	Log Likelihood -2193,8634	TIMETREE								
	Replicate No. 3 of 500									
	< >									
· · · · · · · · · · · · · · · · · · ·	Andreis Onting									
		TAMONKEY								
	Statistical Method : Maximum Likelihood									
	Prylogeny Test Test of Phylogeny · Bootstran method									
	No. of Bootstrap Replications : 500									
	Substitution Model									
	Substitutions Type : Nucleotide									
RECENT PUBLICATIONS	Model/Method : Tamura-Nei model									
	Rates among Sites : Gamma Distributed (G)									
	No of Discrete Gamma Categories : 5									
	Data Subset to Use									
HELP DOCS EXAMPLES	Gaps/Missing Data Treatment : Complete deletion									
	Tree Inference Options									
	ML Heuristic Method : Subtree-Pruning-Regrafting - Fast									
	Initial Tree for ML : Make initial tree automatically (D									
	< >									

26. Po skončení analýzy je výsledný strom zobrazen pomocí podprogramu Tree Explorer, který umožní mino jiné grafické editace stromu. Hodnoty bootstrap replikátů jsou uvedeny u kořenů větví (kládů) a na následujícím obrázku jsou označeny modrým kruhem. V prostředí programu Tree Explorer je možné provést například takzvané zakořenění stromu na základě přítomnosti separované větve s outgroup druhu (*Tachyglossus aculeatus*).



27. Výstupem programu MEGA X je rovněž kompletní popis fylogenetické analýzy včetně odkazů na primární literární zdroje, kde jsou popsány algoritmy, na základě kterých byl vybírán optimální substituční model a konstruován fylogenetický strom.

Evolutionary analysis by Maximum Likelihood method

The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model [1]. The tree with the highest log likelihood (-2224,95) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0,1072)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 36 nucleitide sequences. All positions containing gaps and missing data were eliminated (complete deletion option). There were a total of 514 positions in the final dataset. Evolutionary analyses were conducted in MEGAX [2].

1. Tamura K. and Nei M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 10:512-526.
2. Kumar S., Stecher G., Li M., Knyaz C., and Tamura K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution 35:1547-1549.

Disclaimer: Although utmost care has been taken to ensure the correctness of the caption, the caption text is provided "as is" without any warranty of any kind. Authors advise the user to carefully check the caption prior to its use for any purpose and report any errors or problems to the authors immediately (www.megasoftware.net). In no event shall the authors and their employers be liable for any damages, including but not limited to special, consequential, or other damages. Authors specifically disclaim all other warranties expressed or implied, including but not limited to the determination of suitability of this caption text for a specific purpose, use, or application.

28. Vytvořený strom je možné rovněž exportovat v NEWICK formátu. Jedná se standardní formát, ve kterém jsou uložena kompletní data potřebná k otevření stromu a jeho další editace či analýzy. S formátem NEWICK pracuje většina programů zaměřených na fylogenetické analýzy. Pomocí tohoto formátu jsou exportovány konkrétní hodnoty analyzovaného stromu, jako jsou například délky větví (Branch Legths) nebo hodnoty bootstrapů (Bootstrap Values).



29. Po ukončení analýz je možné veškerá vstupní i výstupní data uložit formou projektu (session) ve formátu MEGA. Uložený projekt je možné opětovně otevřít a upravovat analýzy či provádět analýzy nové.

