MASARYKOVA UNIVERZITA
Přírodovědecká fakulta
RECETOX

Diplomová práce

Brno 2025 Michaela Havlíčková

MASARYKOVA
UNIVERZITA
PŘÍRODOVĚDECKÁ FAKULTA
RECETOX

Detekce selekce v genetických markerech používaných pro fylogenetickou rekonstrukci u obratlovců

Diplomová práce

Michaela Havlíčková

Vedoucí práce: Mgr. Markéta Harazim, Ph.D. **Brno 2025**

Bibliografický záznam

Autor:	Bc. Michaela Havlíčková Přírodovědecká fakulta, Masarykova univerzita RECETOX
Název práce:	Detekce selekce v genetických markerech používaných pro fylogenetickou rekonstrukci u obratlovců
Studijní program:	Matematická biologie a biomedicína
Studijní obor:	Biomedicínská bioinformatika
Vedoucí práce:	Mgr. Markéta Harazim, Ph.D.
Akademický rok:	2024/25
Počet stran:	$7 + 86$
Klíčová slova:	Fylogenetická rekonstrukce; Fylogenetický marker; Selekce; Přirozený výběr; Pozitivní selekce; Neutrální evoluce; PAML

Bibliographic Entry

Author:	Bc. Michaela Havlíčková Faculty of Science, Masaryk University RECETOX
Title of Thesis:	Signal for natural selection in genetic markers used for phylogenetic reconstuction in vertebrates
Degree Programme:	Computational biology and biomedicine
Field of Study:	Biomedical bioinformatics
Supervisor:	Mgr. Markéta Harazim, Ph.D.
Academic Year:	2024/25
Number of Pages:	$7 + 86$
Keywords:	Phylogenetic reconstruction; Phylogenetic marker; Selection; Natural selection; Positive selection; Neutral evolution; PAML

Abstrakt

Metody a modely běžně využívané při fylogenetické rekonstrukci implicitně předpokládají evoluční neutralitu použitých fylogenetických markerů. Působení selekce však porušuje předpoklady substitučních modelů. Přesto se selekční signál markerů před samotnou fylogenetickou analýzou běžně netestuje. V této práci jsem zkoumala působení selekce na 11 nukleárních a 3 mitochondriální fylogenetické markery běžně využívané u savců. Kódující sekvence markerů jsem získala z databáze NCBI pro téměř 200 zástupců sedmi řádů savců. Analýza byla provedena pomocí modelů založených na poměru nesynonymních a synonymních substitucí (ω) implementovaných v programu PAML. Zároveň jsem testovala signál pozitivní selekce a specifické rozdíly v působení selekce u devíti různých taxonů. Signál pozitivní selekce napříč všemi savci nebo alespoň u jednoho z taxonů byl detekován u 10 z 11 nukleárních markerů. Většina markerů, včetně těch mitochondriálních, zároveň vykazovala známky silné negativní selekce, což odpovídá očekávanému selekčnímu tlaku na zachování funkce kódovaného proteinu. Výsledky ukazují, že i často používané fylogenetické markery mohou být pod vlivem selekce, což může mít dopad na přesnost fylogenetické rekonstrukce. Zvážení selekčních vlivů by tak mělo být důležitou součástí výběru markerů pro fylogenetické analýzy.

Abstract

Methods and models commonly used in phylogenetic reconstruction implicitly assume the evolutionary neutrality of the utilized phylogenetic markers. However, the effect of selection violates the assumptions of substitution models. Despite this, selection signals in markers are rarely tested prior to phylogenetic analysis. In this thesis, I investigated the effects of selection on 11 nuclear and 3 mitochondrial phylogenetic markers commonly used in mammals. I obtained coding sequences of these markers from the NCBI database for nearly 200 representatives across seven mammalian orders. The analysis was conducted using models based on the ratio of nonsynonymous to synonymous substitutions (ω) implemented in the PAML software. I also independently tested for signal of positive selection and taxon-specific differences in natural selection pressures in 9 different taxa. A signal of positive selection across all mammals or in at least one of the taxa was detected in 10 out of 11 nuclear markers. Most markers, including mitochondrial ones, also showed signs of strong negative selection, consistent with the expected selective pressure to preserve protein function. These results indicate that even widely used phylogenetic markers may be subject to selection, which can affect the accuracy of phylogenetic reconstruction. Therefore, considering the influence of selection should be an important part of marker selection for phylogenetic analyses.

MUNI
SCI

MASARYKOVA UNIVERZITA
PŘÍRODOVĚDECKÁ FAKULTA
KOTLÁŘSKÁ 2, 611 37 BRNO
IČ: 00216224
DIČ: CZ00216224

ZADÁNÍ DIPLOMOVÉ PRÁCE

Akademický rok: 2023/2024

Ústav:	RECETOX
Studentka:	Bc. Michaela Havlíčková
Program:	Matematická biologie a biomedicína
Specializace:	Biomedicínská bioinformatika

Ředitel *ústavu* PřF MU Vám ve smyslu Studijního a zkušebního řádu MU určuje diplomovou práci s názvem:

Název práce:	Detekce selekce v genetických markerech používaných pro fylogenetickou rekonstrukci u obratlovců
Název práce anglicky:	Signal for natural selection in genetic markers used for phylogenetic reconstruction in vertebrates
Jazyk závěrečné práce:	čeština

Oficiální zadání:
Fylogenetická rekonstrukce vztahů taxonů předpokládá evoluční neutralitu použitého genetického markeru v rámci zkoumaného taxonu. Na základě literární rešerše student identifikuje geny běžně používané jako neutrální markery pro fylogenetickou rekonstrukci u obratlovců. Získá kódující sekvence těchto markerů reprezentující taxony obratlovců dostupné v nukleotidových databázích nebo celogenomových sekvencích. Otestuje selekční signál detekovatelný v rámci jednotlivých genů i taxonů a vyhodnotí vhodnost markerů pro fylogenetickou rekonstrukci u různých skupin obratlovců.

Literatura:

Thomson R. C., I. J. Wang, J. R. Johnson. 2010. Genome-enabled development of DNA markers for ecology, evolution and conservation. Molecular Ecology 19.11: 2184-2195

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution, 24.8: 1586-1591.

Vedoucí práce:	Mgr. Markéta Harazim, Ph.D.
Konzultant:	doc. Mgr. Natália Martínková, Ph.D.
Datum zadání práce:	17. 10. 2022
V Brně dne:	31. 1. 2024

Zadání bylo schváleno prostřednictvím IS MU.

Bc. Michaela Havlíčková, 6. 2. 2023
Mgr. Markéta Harazim, Ph.D., 8. 2. 2023
RNDr. Tomáš Pavlík, Ph.D., 9. 2. 2023

Poděkování

Chtěla bych poděkovat vedoucí své diplomové práce Mgr. Markétě Harazim, Ph.D. za odborné vedení a čas, který mi věnovala. Dále bych ráda poděkovala doc. Mgr. Natálii Martínkové, Ph.D. za konzultace a cenné rady při zpracování této práce.
Výpočetní zdroje byly poskytnuty projektem "e-Infrastruktura CZ" (e-INFRA CZ LM2018140) podporovaným Ministerstvem školství, mládeže a tělovýchovy České republiky.

Prohlášení

Prohlašuji, že jsem svoji diplomovou práci vypracovala samostatně pod vedením vedoucího práce s využitím informačních zdrojů, které jsou v práci citovány.

Brno 12. května 2025

. . . . . .
Michaela Havlíčková

Obsah

Seznam obrázků 3

Seznam tabulek 4

Seznam pojmů a zkratek 5

1 Úvod 6
- 1.1 Fylogenetická rekonstrukce 6
  - 1.1.1 Fylogenetický marker 7
- 1.2 Selekce 8
- 1.3 Metody detekce selekce 8
  - 1.3.1 Aproximační metody 9
  - 1.3.2 Metoda maximální věrohodnosti 10
- 1.4 Vliv selekce na fylogenetickou rekonstrukci 10
- 1.5 Cíle práce 12

2 Metodika 13
- 2.1 Výběr taxonů obratlovců 13
- 2.2 Identifikace fylogenetických markerů 13
- 2.3 Získání dat z NCBI 14
- 2.4 Alignment 15
  - 2.4.1 Příprava dat 15
  - 2.4.2 Prvotní alignment 16
  - 2.4.3 Filtrace 16
  - 2.4.4 Rafinace, finální úpravy a kontrola kvality alignmentu 17
  - 2.4.5 Extrakce markerů z alignmentů celých genů 18
- 2.5 Fylogenetická rekonstrukce 18
- 2.6 Výpočet selekčního signálu 19
  - 2.6.1 Modely 20
  - 2.6.2 Vstupní data 26
  - 2.6.3 Automatizace práce s codeml 28

3 Výsledky 30
- 3.1 Výběr taxonů obratlovců 30
- 3.2 Identifikace fylogenetických markerů 30

3.3 Získání dat z NCBI 31
3.4 Alignment 33
3.4.1 Extrakce markerů z alignmentů celých genů 34
3.5 Fylogenetická rekonstrukce 35
3.6 Výpočet selekčního signálu 38

4 Diskuze 44
4.1 Fylogenetická rekonstrukce 44
4.2 Detekce selekce 44

5 Závěr 48

Příloha 1 – Taxonomie analyzovaných druhů 49

Příloha 2 – Přístupová čísla NCBI 53

Příloha 3 – Výsledky kladového modelu C 66

Příloha 4 – Výsledky modelu větví-pozicí A 69

Seznam použitých zdrojů 73

Seznam obrázků

Obr. 1 Topologie výsledného fylogenetického stromu s barevně odlišenými řády savců. Geometrickými symboly jsou vyznačeny uzly společného předka dalších tří analyzovaných taxonů. Čísla představují podpory některých hlubokých uzlů stromu. 37

Seznam tabulek

Tab. 1 Parametry použitých kodonových substitučních modelů. Volné parametry jsou uvedeny bez závorky a jejich počet je uveden ve sloupci „#p“. Tabulka obsahuje také specifická nastavení jednotlivých modelů v kontrolním souboru codeml. 28

Tab. 2 Seznam všech devíti analyzovaných taxonů a jejich zastoupení, včetně šedě vyznačeného řádu Carnivora, který nebyl analyzován jako celek, ale rozdělený na dva podřády. 31

Tab. 3 Analyzované fylogenetické markery savců a šedě vyznačený kontrolní gen *TP53*. Používá-li se jako marker pouze konkrétní exon nebo exony, je u genu uvedeno, o které se jedná. Typy markerů „mt“ a „nc“ označují mitochondriální resp. nukleární geny. 32

Tab. 4 Přehled stažených sekvencí a chybějících dat. 33

Tab. 5 Délky alignmentů (bp) v jednotlivých krocích alignmentu. Předposlední sloupec představuje finální délku alignmentů celých genů a poslední sloupec část těchto alignmentů, představující fylogenetické markery. 34

Tab. 6 Výsledky pozicových modelů M1a a M2a. Počet stupňů volnosti LRT je 2. 40

Tab. 7 Výsledky pozicových modelů M7 a M8. Počet stupňů volnosti LRT je 2. . 41

Tab. 8 Signifikantní výsledky modelu cmC s rozdílem poměrů $|\Delta\omega| \geq 0,1$ pro jednotlivé taxony. Nulové modely nejsou uvedeny, pouze výsledky LRT testu s jedním stupněm volnosti. 42

Tab. 9 Signifikantní výsledky modelu bsA pro jednotlivé taxony. Nulové modely nejsou uvedeny, pouze výsledky LRT testu s jedním stupněm volnosti. . . 43

Seznam pojmů a zkratek

Alignment	Zarovnání dvou a více sekvencí pod sebe tak, aby si odpovídaly homologní pozice.
bp	Pár bází (base pair); jednotka délky genetické sekvence.
bsA	Kodonový substituční model větví-pozicí A (branch-site A model).
CDS	Protein kódující sekvence (coding sequence).
cmC	kladový model C (clade model C); jeden z kodonových substitučních modelů programu codeml.
Fylogenetický marker	Specifický gen či jeho část běžně využívaná pro fylogenetickou rekonstrukci.
Homologní sekvence	Sekvence se společnou evoluční historií.
kb	Kilobáze; 1000 párů bází (bp).
Kodon	Trojice nukleotidů kódujících jednu aminokyselinu.
LRT	Test poměru věrohodností (likelihood ratio test).
ML metoda	Metoda maximální věrohodnosti (maximum likelihood method).
mt	Mitochondriální.
nc	Nukleární (jaderný).
Nesynonymní substituce	Substituce způsobující změnu kódované aminokyseliny.
Substituce	Změna nukleotidové báze v genetické sekvenci za jinou.
Synonymní substituce	Substituce zachovávající kódovanou aminokyselinu.
Topologie	Způsob větvení fylogenetického stromu.

1 Úvod

1.1 Fylogenetická rekonstrukce

Fylogenetická rekonstrukce je proces, při kterém se snažíme na základě molekulárních nebo morfologických dat odhadnout evoluční historii a příbuzenské vztahy mezi organismy. Výsledkem rekonstrukce je fylogenetický strom, který graficky znázorňuje hypotézu o tom, jak se jednotlivé druhy v minulosti oddělovaly od posledního společného předka. Skutečné evoluční vztahy mezi organismy nemůžeme pozorovat přímo, protože se jedná o události, které se odehrály v minulosti a musí být zrekonstruovány na základě dostupných důkazů (Patwardhan et al., 2014). Fylogenetická rekonstrukce se nemusí zaměřovat pouze na jednotlivé organismy, ale i na druhy, geny či celé genomy, a je klíčovým nástrojem v evoluční biologii a systematice (Kapli et al., 2020; Naser-Khdour et al., 2019).

Rekonstrukci lze provézt na základě mnoha typů dat jako jsou molekulární sekvence (DNA, RNA, proteiny), genové frekvence nebo tradičně dle morfologických znaků (Felsenstein, 2004; Patwardhan et al., 2014). V dnešní době díky rychlému rozvoji sekvenačních technologií převažuje ve fylogenetice využití molekulárních dat. Tyto technologie umožňují rychle získat obrovské množství dat pro fylogenetickou analýzu (Young a Gillung, 2020).

V molekulární fylogenezi se vztahy mezi organismy nebo geny zkoumají porovnáváním homologních sekvencí DNA či proteinů (sekvence se společnou evoluční historií). Rozdíly mezi těmito sekvencemi odrážejí genetickou divergenci, ke které docházelo v důsledku molekulární evoluce v průběhu času. Porovnáním homologních molekul u různých organismů lze určit míru jejich podobnosti a na základě toho rekonstruovat hierarchii příbuzenských vztahů ve formě fylogenetického stromu (Patwardhan et al., 2014).

Existuje několik metod fylogenetické rekonstrukce, které se liší ve svém přístupu. Metody založené na vzdálenosti využívají matici genetických vzdáleností mezi dvojicemi sekvencemi. Genetická vzdálenost vyjadřuje míru odlišnosti dvou sekvencí. Vzdálenosti se počítají na základě odhadu počtu substitucí (změn) mezi sekvencemi za využití evolučních modelů, které berou v úvahu také možný výskyt více substitucí na jedné nukleotidové pozici. Strom lze na základě genetických vzdáleností vytvořit například pomocí metod shlukovací analýzy, jako je metoda nejbližšího souseda (Salemi et al., 2009).

Další skupina metod je založená na diskrétních znacích, nejčastěji alignovaných nukleotidových či peptidových sekvencích. Ty pracují přímo s mutacemi, které se v sekvencích akumulovaly a nedochází tak ke ztrátě informací jako v případě vzdálenostních metod. Důsledkem toho jsou metody přesnější než metody založené na vzdálenosti. Jednou z neznámějších je metoda maximální věrohodnosti, která hledá takový strom, který maximalizuje pravděpodobnost pozorování daných dat. Další z této skupiny metod je metoda maximální parsimonie, která hledá strom, který minimalizuje počet evolučních změn potřebných k vysvětlení daných dat a stojí na myšlence, že to nejjednodušší možné vysvětlení je to správné (Salemi et al., 2009; Patwardhan et al., 2014).

K modelování pravděpodobností výskytů jednotlivých substitucí se používají nejčastěji nukleotidové substituční modely.

Nukleotidové substituční modely mají tři základní předpoklady: (1) Pro každou nukleotidovou pozici platí, že změna z báze *i* na bázi *j* je **nezávislá** na bázi, která se na pozici vyskytovala před bází *i*. (2) Substituční rychlosti se v čase nemění – jsou **homogenní** pro celý strom. (3) Relativní frekvence bází jsou ve stavu ekvilibria, to znamená, že se v čase nemění (**stacionarita**). U časově reversibilních modelů, které bývají používány nejčastěji, se navíc předpokládá, že substituční rychlost nukleotidu *i* na *j* je stejná jako *j* na *i* (**reversibilita**). Většina metod fylogenetické rekonstrukce předpokládá také nezávislost pozic mezi sebou. Například metoda maximální věrohodnosti očekává nezávislou evoluci mezi všemi pozicemi, aby bylo možné spočítat věrohodnostní funkci vynásobením věrohodností jednotlivých pozic. Navíc předpokládá také nezávislost evoluce mezi jednotlivými vývojovými větvemi (Salemi et al., 2009; Naser-Khdour et al., 2019).

1.1.1 Fylogenetický marker

Fylogenetický marker je vhodný gen, jeho část, či jiná specifická DNA sekvence běžně využívaná za účelem fylogenetické analýzy.

Ne každá genetická sekvence je vhodným fylogenetickým markerem a použitelnost markerů se může lišit také taxon od taxonu. Ideální gen, který by mohl být potenciálním fylogenetickým markerem, by měl mít několik vlastností. Mělo by se jednat o gen, který se v genomu vyskytuje pouze v jediné kopii (single-copy), aby se předešlo nechtěnému sekvenování různých kopií genů u jednotlivých studovaných organismů. Vzhledem k tomu, že sekvence je třeba před fylogenetickou analýzou zarovnat, neměl by jejich alignment být příliš obtížný. Délky téhož genu se mohou u různých taxonů lišit vlivem inzercí a delecí, což může správný alignment značně ztěžovat. Dalším faktorem může být také dostupnost primerů potřebných k amplifikaci markeru. Je také důležité, aby měl marker optimální substituční rychlost. Alespoň pro vzdáleně příbuzné organismy jsou preferované pomalu se vyvíjející geny s dostatkem konzervovaných pozic, ale zároveň i těch variabilních. U příliš rychle mutujícího genu může dojít po opakovaných substitucích k saturaci (opětovnému výskytu původního nukleotidu) (Patwardhan et al., 2014). Při analýze blízce příbuzných organismů je ale naopak nutné zvolit gen s větší rychlostí mutací, aby byl dostatečně fylogeneticky informativní (Cruickshank, 2002).

Kvůli struktuře protein kódujících genů mohou být variabilní většinou pouze třetí kodonové pozice bez toho, aniž by došlo ke změně kódované aminokyseliny. Mutace měnící aminokyselinu mají mnohem větší pravděpodobnost, že budou narušovat správnou funkci genem kódovaného proteinu. Důsledkem toho se snižuje šance jejich výskytu (Cruickshank, 2002).

Nekódující sekvence, jako například introny nukleárních genů, nepodléhají funkčním a strukturálním omezením proteinů, a mají tak tendenci mutovat rychleji. Z toho důvodu je většinou možné použít je pouze pro fylogenetickou rekonstrukci blízce příbuzných druhů (Cruickshank, 2002).

Ribozomální RNA (rRNA) je považována za ideální fylogenetický marker, protože je univerzální pro všechny organismy a je tvořena jak variabilními, tak vysoce konzervovanými doménami (Patwardhan et al., 2014).

Mitochondriální geny se používají především pro fylogenezi na druhové úrovni. Jedná se například o geny pro podjednotky cytochrom *c* oxidázy I a II (*COX1* a *COX2*) nebo

cytochrom *b* (*CYTB*) (Patwardhan et al., 2014).

Z nukleárních genů jsou nejvhodnější ty s dlouhými konzervovanými exony. Pro fylogenezi savců byly jedněmi z prvních markerů například první exon Retinol vázajícího proteinu 3 (*RBP3*) nebo exon 28 von Willebrandova faktoru (*VWF*) (Ranwez et al., 2007). Kódující nukleární geny se nejčastěji využívají k inferenci hlubokých fylogenetických vztahů díky vysoké míře konzervace a často také dostupné anotaci, která usnadňuje identifikaci paralogů (genů vzniklých duplikací) (Thomson et al., 2010).

1.2 Selekce

Selekce je spoluzodpovědná za obrovskou komplexitu živých systémů. Přirozený výběr se projevuje nenáhodnou změnou v evoluční zdatnosti (reprodukční úspěšnosti) a často je nepřímo spojena se schopností přežít v určitém prostředí. To vede ke zvýšení frekvence výskytu výhodných resp. snížení frekvence škodlivých dědičných charakteristik v populaci z generace na generaci. Všechny živé struktury jsou, mimo jiné, výsledkem opakované fixace nebo eliminace genetických variant v populacích (Salemi et al., 2009; Gregory, 2009).

Selekci lze rozdělit na několik druhů dle jejího vlivu na četnost různých fenotypových rysů v populaci. Usměrňující selekce upřednostňuje jeden extrémní fenotyp oproti tomu průměrnému a nebo extrému z druhé strany fenotypového spektra. Upřednostňováním extrémních vlastností z obou stran fenotypové škály se vyznačuje disruptivní, nebo také diversifikující selekce, která zvyšuje fenotypovou variability v populaci. Stabilizující selekce naopak variabilitě v populaci brání a upřednostňuje průměrný fenotyp (Sanjak et al., 2018).

Selekci také můžeme na molekulární úrovni rozlišovat dle toho, jak daná mutace ovlivňuje způsobilost organismu. Je-li určitá genetická varianta výhodná a zvyšuje šance organismu přežít a reprodukovat se, dojde k její postupné dominanci a fixaci v populaci. Tento proces se nazývá pozitivní selekce. Naopak škodlivé mutace snižující způsobilost organismu přežít budou z populace časem eliminovány a v takovém případě mluvíme o selekci negativní (Salemi et al., 2009). Speciálním případem pak může být uvolněná selekce, při které dochází k uvolnění selekčního tlaku na dříve důležité geny, což se projeví zvýšenou akumulací změn, které by dříve byly odstraněny (Lahti et al., 2009).

1.3 Metody detekce selekce

Metody detekce selekce na základě sekvencí DNA porovnávají substituční rychlosti nesynonymních (dN) a synonymních substitucí (dS).

Rychlosti nesynonymních a synonymních substitucí představují počet nesynonymních a synonymních substitucí na jednu pozici alignmentu. Nukleotidová substituce v kontextu protein kódujících genetických sekvencí (CDS; coding sequence) se nazývá synonymní, nebo také tichá, jestliže nevede ke změně aminokyseliny, na kterou je příslušný kodon přeložen. V opačném případě se jedná o nesynonymní substituci, tzv. substituci měnící smysl kodonu, v důsledků které dochází ke změně aminokyseliny v peptidovém řetězci (Yang a Bielawski, 2000).

Poměr substitučních rychlostí dN a dS se označuje jako ω a je významným ukazatelem selekce. Hodnota $\omega > 1$ indikuje pozitivní selekci, kdy v sekvencích pozorujeme zvýšenou míru rychlosti nesynonymních substitucí, které vysvětlujeme tím, že docházelo k fixaci výhodných mutací v populaci. Naopak hodnota $\omega < 1$ poukazuje na negativní selekci, kdy nesynonymní substituce měly negativní vliv na životaschopnost organismu a důsledkem toho se akumulují pomaleji a pozorujeme jich v sekvencích méně. Hodnota $\omega = 1$ značí neutrální evoluci, bez vlivu selekčního tlaku (Yang a Bielawski, 2000; Salemi et al., 2009).

Nesynonymní substituce jsou ukazatelem selekce, protože mohou mít přímý vliv na funkci proteinu a ovlivnit tak evoluční zdatnost organismu. Synonymní substituce nepředstavují žádnou funkční změnu a vyvíjí se většinou neutrálně, stejně tak je tomu v případě nesynonymních substitucí, které neovlivní funkci proteinu (Salemi et al., 2009).

1.3.1 Aproximační metody

Aproximační metody zahrnují mnoho intuitivních způsobů odhadů dN a dS (Yang a Bielawski, 2000).

Při porovnání dvou zarovnaných sekvencí můžeme spočítat počet nukleotidových substitucí mezi nimi, následně CDS přeložit do aminokyselinové sekvence a spočítat kolikrát došlo vlivem nukleotidových substitucí ke změně na úrovni proteinu. Dokážeme tak říct, kolik nukleotidových substitucí mezi sekvencemi je nesynonymních (N_{sub}) a kolik synonymních (S_{sub}).

Protože pravděpodobnost vzniku synonymní a nesynonymní substituce se po každou nukleotidovou pozici alignmentu může lišit, je potřeba ji při výpočtu zohlednit. Například máme-li DNA kodon CTG , kódující leucin, můžeme uvažovat všechny možné substituce na všech třech pozicích a sledovat, zda vedou ke změně aminokyseliny či nikoliv. Substituce na třetí pozici kodonu mohou vézt ke vzniku kodonů CTA, CTC a CTT, z nichž všechny tři stále kódují leucin, tudíž třetí pozice kodonu je zcela synonymní. Substitucemi na druhé pozici kodonu mohou vzniknout kodony CAG, CGG a CCG, kódující glutamin, arginin a prolin. Jakákoliv substituce tedy vede ke změně původní aminokyseliny a záměna na druhé pozici kodonu je tedy zcela nesynonymní. Při substituci prvního nukleotidu kodonu mohou vzniknout ATG, TTG a GTG, kódující methionin, leucin a valin. Jedna z možných substitucí zachovává původní aminokyselinu, zatímco druhé dvě vedly ke změně a první pozice kodonu je tedy z $1/3$ synonymní a ze $2/3$ nesynonymní (Miyata a Yasunaga, 1980). Uvažujeme-li sekvenci délky n pak pro počet nesynonymních pozicí N_{poz} a počet synonymních pozicí S_{poz} platí $N_{poz} + S_{poz} = n$. Nyní můžeme spočítat rychlosti nesynonymních substitucí dN a synonymních substitucí dS a porovnat je. Získáme tak poměr ω představující poměr rychlostí nesynonymních a synonymních substitucí

$$\omega = \frac{dN}{dS} = \frac{N_{sub}/N_{poz}}{S_{sub}/S_{poz}}. \tag{1.1}$$

Jedná se pouze o zjednodušený princip výpočtu a v praxi se běžně zohledňují také situace, kdy dojde k více substitucím v rámci jednoho kodonu (Miyata a Yasunaga, 1980; Nei a Gojobori, 1986).

Metody fungující na tomto principu odhadu substitučních rychlostí dN a dS mají zjednodušené předpoklady o procesu nukleotidových substitucí a nezohledňují například

vliv četnosti výskytů tranzicí a transverzí ani preferenční využívání kodonů, které v obou případech mohou mít zásadní dopad na odhad poměru ω (podrobněji v kapitole 2.6.1) (Yang a Bielawski, 2000).

1.3.2 Metoda maximální věrohodnosti

Dalším možným přístupem odhadů ω je metoda maximální věrohodnosti (maximum likelihood; ML). ML metoda obecně spočívá ve snaze maximalizovat věrohodnostní funkci, která představuje pravděpodobnost pozorování dat za předpokladu platnosti určité hypotézy. ML metoda využívá modely kodonových substitucí s několika parametry jako je například divergence sekvencí t, poměr tranzicí a trensverzí κ a především poměr ω. Všechny tyto parametry jsou odhadovány současně maximalizací věrohodností funkce. Model je formulován na úrovni okamžitých změn a nedovoluje tedy více substitucí současně (Yang a Bielawski, 2000). Velikou výhodou ML metody je, že modely dokáží zohlednit vliv poměru rychlostí tranzicí a transverzí a také preferenční využívání kodonů (podrobněji v kapitole 2.6.1).

1.4 Vliv selekce na fylogenetickou rekonstrukci

Nukleotidové substituční modely nezohledňují, že se sekvence mohou vyvíjet pod selekčním tlakem a implicitně předpokládají neutrální evoluci. Selekce ale porušuje předpoklady modelů (např. nezávislost pozic) a potenciálně tak může vézt k inferenci chybných fylogenetických vztahů oproti stromům vytvořených na základě neutrálních markerů. I přes to se detekce selekce před fylogenetickou inferencí běžně neprovádí (Roje, 2014).

Selekce na proteinové úrovni může ovlivnit nukleotidové substituční rychlosti a tím i jejich využitelnost při fylogenetické rekonstrukci (Jansa et al., 2006). Například substituční rychlost kodonu i na j může být jiná než rychlost v opačném směru v závislosti na tom, zda je některý z kodonů výhodný nebo škodlivý, což porušuje předpoklad reversibility substitučních modelů. Rychlosti se mohou také měnit v čase, je-li nutná adaptace na nové prostředí, což může vézt ke zrychlené fixaci výhodných alel.

Změny v působení selekce mohou mít širokou škálu efektů na zastoupení bází v genomech (Ko et al., 2006), čímž selekce porušuje předpoklad stacionarity. Jedním z efektů může být translační selekce, ke které dochází při preferenčním využívání synonymních kodonů odpovídajících nejhojněji dostupné tRNA (Grantham et al., 1981; Plotkin a Kudla, 2011). Vliv na zastoupení bází může mít také náročnost syntézy nukleotidů (Rocha a Danchin, 2002) nebo vliv prostředí a životního stylu (Tarallo et al., 2016).

Pozice v CDS nejsou nezávislé, protože důležitou roli hraje pozice v trojici nukleotidů kodonu, nebo také strukturální omezení proteinu (Tavaré, 1986; Massey et al., 2008). V případě, že se pozice vyvíjí pod silným selekčním tlakem, se dá očekávat, že jejich evoluce bude korelovaná (Roje, 2014).

Porušení předpokladů modelů může vést k systematickým chybám, které na rozdíl od těch stochastických nelze jednoduše eliminovat zvětšením množství sekvencí a jejich délek. Vliv systematických chyb se naopak pro velké fylogenetické datasety ještě umocňuje (Naser-Khdour et al., 2019).

U kódujících nukleotidových sekvencí lze ve většině případů očekávat vysokou míru negativní selekce. Je totiž nutné, aby byla zachována správná funkce proteinu a jakékoli škodlivé substituce budou z populace postupně eliminovány (Popadin et al., 2013; Hughes et al., 2003; Hurst, 2002). Projevem negativní selekce jsou tedy snížené substituční rychlosti, díky čemuž nemusí vzniknout dostatečné množství fylogeneticky informativních pozic (Townsend, 2007). Pozitivní selekce naopak urychluje substituční rychlosti sekvencí. To se může projevit přitažlivostí dlouhých větví (Philippe et al., 2000), což je fylogenetický artefakt, při kterém dochází k chybnému určení blízké příbuznosti dvou nebo více dlouhých vývojových větví, které ale neodráží jejich skutečnou evoluční historii (Bergsten, 2005). Dalším možným projevem je homoplazie, nazývaná také konvergentní molekulární evoluce. Konvergence představuje zisk podobných genotypových změn v nepříbuzných vývojových liniích, u kterých se většinou předpokládá, že jsou poháněny právě selekcí. Důsledkem je, že v současnosti pozorujeme v různých vývojových liniích stejné znaky, které ale nepocházejí od společného předka, což může zmást fylogenetický signál (Edwards, 2009; Roje, 2014; Castoe et al., 2009). K homoplazii může vézt ale také negativní selekce. Vzhledem k tomu, že jsou na substituce kladeny funkční a strukturální omezení proteinu, je očekávané, že se škodlivé změny budou mít tendenci vracet zpět do původního stavu, nebo bude docházet k jiným kompenzujícím změnám. Způsobů, jak změny kompenzovat nebude pravděpodobně mnoho, což může produkovat systematický signál odlišný od signálu skutečného původu (Massey et al., 2008).

Jaký vliv mohou mít porušení předpokladů substitučních modelů na výsledek fylogenetické rekonstrukce studovali autoři studie Naser-Khdour et al. (2019). Zaměřili se na tři předpoklady modelů – stacionaritu, reversibilitu a homogenitu. Testovali širokou škálu nukleotidových sekvencí (nukleární, mitochondriální, kódující, introny, tRNA, rRNA a další) u živočichů, rostlin i virů. Zda jednotlivé datasety splňují předpoklady modelů určovali s pomocí několika párových testů homogenity. Zjistili, že nejčastěji porušují předpoklady modelů třetí kodonové pozice, ale porušení je časté u všech typů sekvencí. Poté provedli fylogenetickou rekonstrukci na základě dat, které předpoklady splňují a zvlášť těch, co je porušují. Výsledné stromy byly často statisticky významně odlišné, z čehož plyne, že porušení předpokladů modelů má na fylogenetickou analýzu nezanedbatelný vliv.

Edwards (2009) ve svém článku hodnotí vliv různých typů selekce na fylogenetickou rekonstrukci. Stabilizující selekce dle něj pouze sníží celkovou rychlost evoluce. Usměrňující selekce může porušit předpoklady modelů, ale pouze mírně. Největší problém tak vidí v molekulární konvergenci způsobené selekcí, protože neexistují spolehlivé způsoby, jak se s ní při fylogenezi vypořádat. Jako příklad uvádí neneutrální konvergentní evoluci zdokumentovanou ve studii Castoe et al. (2009). Vlivem konvergence kódujících mitochondriálních genů získali autoři na základě těchto genů strom, který se markantně lišil od robustnějšího stromu na základě nukleárních markerů. Autoři také upozorňují na možnost, že by konvergenci mohly podléhat také nukleární geny. Pravděpodobnost, že by ke konvergenci docházelo v takové míře u tak velkého množství genů, aby dokázala zmást fylogenetický signál, je ale malá. K tomuto jevu by ale mohlo dojít, pokud došlo v celém genu k tak velkému posunu zastoupení bází oproti příbuzným liniím, že daná vývojová linie působí jako nepříbuzná. Takový případ byl zdokumentován u genu *RAG1* u vačic (Gruber et al., 2007).

1.5 Cíle práce

Cílem této práce je vyhodnotit, zda jsou často využívané fylogenetické markery skutečně vhodné pro fylogenetickou rekonstrukci. Vykazují-li tyto genetické markery selekci, porušují tím předpoklady nukleotidových substitučních modelů, což může představovat překážku pro fylogenetickou inferenci a ovlivnit její výsledek.

Na základě výsledků detekce selekčního signálu v rámci markeru lze rozhodnout, zda je genetický marker vhodný pro fylogenetickou rekonstrukci. Detekcí selekce v rámci jednotlivých taxonů lze prokázat, že marker není vhodný pouze pro některé taxony, které oproti ostatním vykazují zvýšenou míru selekce.

K dosažení cílů této práce jsem postupovala dle následujících kroků:

- Identifikace běžně využívaných fylogenetických markerů obratlovců.
- Výběr analyzovaných druhů (taxonomických jednotek).
- Získání kódující sekvence markerů z veřejných nukleotidových databází.
- Alignment sekvencí fylogenetických markerů.
- Fylogenetická rekonstrukce analyzovaných taxonů.
- Výpočet selekčního signálu v sekvencích genetických markerů na úrovni celého fylogenetického stromu a u vybraných taxonů.

2 Metodika

2.1 Výběr taxonů obratlovců

Při výběru taxonů obratlovců, kterými bych se chtěla v této práci zabývat, jsem došla k závěru, že analýza více tříd obratlovců najednou by mohla být problematická, a to hned z několika důvodů. Obratlovci zahrnují velice různorodou a geneticky divergentní skupinu organismů. Jedním z kroků analýzy selekčního signálu je alignment, který se stává obtížnějším s rostoucí diverzitou sekvencí. Homologické sekvence dvou geneticky vzdálených organismů si nemusí být dostatečně podobné a mohou obsahovat značné množství indelů, což představuje překážku pro správný alignment (Kumar a Filipski, 2007). Chyby v alignmentu mohou mít negativní dopad na následnou analýzu selekčního signálu ve formě falešně pozitivních i falešně negativních výsledků (Jordan a Goldman, 2012). Dalším rozhodujícím faktorem se ukázal být také samotný výběr fylogenetických markerů blíže popsaný v následující podkapitole 2.2. Markery totiž bývají často třídně specifické a jen málo z nich se používá pro fylogenetickou rekonstrukci u všech, nebo alespoň u většiny tříd obratlovců. Všechny tyto charakteristiky by značně ztěžovaly analýzu obratlovců jako celku.

Z výše uvedených důvodů jsem se rozhodla zaměřit pouze na třídu savců, jakožto zajímavou skupinu obratlovců. Při výběru konkrétních zástupců jsem se zaměřila na savce s dostupným referenčním genomem v databázi NCBI (The National Center for Biotechnology Information) (Sayers et al., 2024) k datu 29. 11. 2023. I přes to, že samotné referenční genomy nebyly pro moji práci potřebné, skutečnost, že má savec dostupný referenční genom částečně zaručuje, že se bude jednat o modelový druh s dostupnými nukleotidovými sekvencemi.

Není-li v této práci přímo uveden poddruh savce (viz Přílohu 1), pak jsem jednotlivé poddruhy nerozlišovala, tedy sekvence markerů pro daný druh mohly patřit také některému z jeho poddruhů.

Analýza selekčního signálu jsem provedla většinou na úrovni savčích řádů. V případě hojně zastoupených řádů byla možná také analýza menších podskupin, které jsem případně vybrala na základě zrekonstruovaného fylogenetického stromu (viz podkapitolu 3.5). Podmínkou pro zahrnutí řádu do analýzy bylo, aby obsahoval minimálně čtyři zástupce. Bylo to z důvodu redukce počtu analyzovaných taxonů, ale také lze očekávat, že hojně zastoupený taxon bude poskytovat více informací pro stabilní odhady parametrů při analýze selekce, včetně poměru ω.

2.2 Identifikace fylogenetických markerů

Výběr fylogenetických markerů probíhal formou literární rešerše. Vyhledávala jsem články zabývající se fylogenetickými markery obecně a konkrétní fylogenetické studie savců a v nich použité markery. Pro tuto práci jsem vybrala pouze markery, které byly použity alespoň

ve třech různých fylogenetických studiích savců či alespoň savčích podskupin.

Při výběru markerů jsem se zaměřila pouze na protein kódující geny, což je předpokladem pro pro výpočet selekčního signálu (viz podkapitola 2.6).

Mezi analyzované geny jsem zařadila také gen *TP53*, který je známým nádorovým supresorem. Vzhledem k důležitosti jeho funkce se jedná o gen pod selekčním tlakem (Khan et al., 2011). Tento gen se běžně nevyužívá jako fylogenetický marker a v práci sloužil pouze jako kontrolní gen, u kterého jsem očekávala detekci selekčního signálu. Pro tento gen jsem neprováděla kompletní analýzu, pouze detekci selekce v rámci genu bez dělení na jednotlivé taxony.

2.3 Získání dat z NCBI

Data jsem získala z nukleotidové databáze NCBI (Sayers et al., 2024) s využitím vlastního skriptu v jazyce Python v. 3.8.10. Skript využíval knihovnu Biopython v. 1.81 (Cock et al., 2009). Skrze Entrez modul jsem získala přístup k datům z online databáze NCBI, a SeqIO modul mi umožnil pracovat se sekvenčními daty. Vstupními informacemi skriptu byl pouze seznam genů a seznam druhů. Skript ze vstupních informací vytvořil dotazy pro vyhledávání v online databázi, vybral nejlepší výsledek vyhledávání, stáhl potřebná data v podobě kódujících sekvencí a uložil je ve FASTA formátu. FASTA je textový soubor, který může obsahovat jednu či více sekvencí. Záznam odpovídající jedné sekvenci začíná vždy znakem „>“ na začátku řádku a hned po něm následuje název sekvence. Následující řádek/řádky obsahují samotnou sekvenci. Případný další výskyt „>“ na začátku řádku značí začátek záznamu další sekvence. V mém případě byl název sekvence tvořený přístupovým číslem sekvence, verzí a názvu sekvence z NCBI databáze. Název sekvence u nukleárních genů běžně obsahoval druh savce, označení a celý název genu, případně o jakou transkripční variantu se jednalo. V případě mitochondriálních genů pak šlo o název mitochondriálního genomu. Za přístupovým číslem a verzí sekvence byla navíc intervalem specifikována část genomu odpovídající sekvenci příslušného genu. Název genu jsem s pomocí skriptu doplnila za název genomu.

Nukleární geny byly vyhledávány v databázi RefSeq (O'Leary et al., 2015). Ta obsahuje neredundantní reprezentativní sekvence genů i genomů a je pečlivě spravována. V první řadě proběhl pokus nalézt ověřenou sekvenci genu (a nemá tedy slovo „PREDICTED“ v názvu). Sekvence označené jako „PREDICTED“ (přístupové číslo začínající prefixem XM_ pro mRNA) jsou predikované pomocí automatizovaných bioinformatických nástrojů. Tyto sekvence jsou odvozeny z genomové sekvence a mohou mít různou úroveň podpory na základě homologních transkriptů nebo proteinů. Nemusí však být potvrzené experimentálními daty a mohou se lišit od kurátorsky ověřených sekvencí. Naopak nepredikované sekvence (prefixy NM_ pro mRNA) prošly manuální kurací a jsou podložené experimentálními důkazy. Tyto sekvence jsou považovány za přesnější reprezentaci skutečných biologických molekul (Pruitt et al., 2010). Vyhledávala jsem také podle klíčového slova „mRNA“ v názvu. Pokud nebyl nalezen žádný vyhovující výsledek, vyhledávání se zopakovalo podruhé, tentokrát včetně predikovaných genů. Pokud byly výsledky vyhledávání úspěšné, zkontrolovala jsem, zda existují různé transkripční varianty genu. Pokud existovaly, byla stažena transkripční varianta 1. V případě, že transkripční varianta 1 nebyla nalezena, stáhla se jiná dostupná transkripční varianta. Vybraná sekvence byla stažena

ve formátu GenBank with parts. Ten obsahuje všechny atributy sekvence, jako je název sekvence, druhu i genu, kódující sekvence, exony atd. Kódující sekvence jsem následně uložila do FASTA souborů. Nebyl-li nalezen žádný výsledek, skript označil sekvenci genu pro daný organismus za nedostupnou.

Při stahování mitochondriálních genů se proces stahování dat lišil. Referenční sekvence nebyly v databázi samostatně, ale jako součásti mitochondriálních genomů. Skript se tedy nejprve pokusil vyhledat v databázi RefSeq referenční mitochondriální genom za použití dotazu „mitochondrion, complete genome“ v názvu sekvence. Pokud takový genom nebyl nalezen, vyhledal pak jakékoliv další nereferenční sekvence kompletního mitochondriálního genomu, kromě těch neověřených (unverified). Pokud bylo nalezeno více výsledků, vybral se ten s nejkratším názvem, čímž se přednostně vybral genom bez nadbytečných informací v názvu, jako např. označení konkrétního vzorku, ze kterého byla sekvence získána. V případě více dostupných genomů jsem poté vybraný genom ještě manuálně kontrolovala. Byl-li úspěšně nalezen nějaký mitochondriální genom, byla celá jeho sekvence stažena ve formátu GenBank obsahující informace o genech a jejich umístění v genomu. Skript tedy vyhledal atribut kódující sekvence zájmových genů a dle jejich genomových souřadnic extrahoval příslušnou část genomu.

Na závěr jsem manuálně kontrolovala názvy sekvencí stažených genů v souborech ve formátu FASTA, případně genomů, ze kterých byly geny extrahovány. Všechny chybějící sekvence, které se mi nepodařilo stáhnout automaticky s pomocí skriptu, jsem se pokusila najít a stáhnout manuálně.

Kontrolovala jsem také délky sekvencí, které by u kódujících nukleotidových sekvencí měly být dělitelné třemi, aby bylo možné je přeložit do aminokyselinové sekvence.

2.4 Alignment

Cílem alignmentu, nebo také zarovnání sekvencí, je sekvence zarovnat pod sebe tak, aby si odpovídaly pozice se společnou evoluční historií. Takové pozice se nazývají homologické. Alignment zahrnoval několik kroků:

2.4.1 Příprava dat

V první řadě bylo nutné pro každý marker vložit sekvence všech savců do jediného FASTA souboru. V rámci tohoto kroku jsem zároveň změnila názvy sekvencí jednotlivých savců, které do teď začínaly přístupovým číslem NCBI následovaným názvem sekvence z databáze. Takto podrobné informace už nebyly dál potřebné a názvy sekvencí jsem nahradila pouze druhem savce.

Poté jsem použila software MACSE v. 2.07 (Ranwez et al., 2018). MACSE se specializuje na alignment kódujících nukleotidových sekvencí. Bere v úvahu jejich transkripci do aminokyselinové sekvence, a proto dokáže pracovat i s posuny čtecího rámce a stop kodony. Program TrimNonHomologousFragments jsem spustila ještě před samotným alignmentem. Dokáže nalézt a vystřihnout nehomologické části sekvencí vyskytující se pouze u několika málo sekvencí. Je tedy vhodný k detekci a odstranění inzercí, které mohou vzniknout neúmyslným zahrnutím intronů v kódující sekvenci důsledkem chyb v anotaci. Další

nehomologické úseky mohou vznikat kvůli alternativního sestřihu (různé transkripční varianty genů). Nehomologické fragmenty činí alignment obtížnějším a časově náročnějším a často navíc bývají zarovnány chybně, je proto vhodné je před alignmentem odstranit. Program TrimNonHomologousFragments jsem použila pouze pro nukleární geny, protože délky jejich sekvencí mohou být značně variabilní právě kvůli různým transkripčním variantám, zatímco sekvence mitochondriálních genů mívají podobnou délku. Použila jsem nastavení parametrů `-min_trim_ext 12` a `-min_trim_in 30`, které specifikují minimální délky detekovaných nehomologických fragmentů na koních resp. uvnitř sekvence. Nalezený nehomologický fragment na jednom z konců sekvence tedy musel být dlouhý minimálně 12 nukleotidů, aby ho program odstranil. Dle výchozího nastavení program odstranil nehomologické fragmenty pouze pokud se vyskytovaly u méně než tří sekvencí (GE²pop, 2025).

2.4.2 Prvotní alignment

Následoval samotný alignment s pomocí MACSE programu alignSequences. Program jsem spustila s nastavením parametru `-gap_op 15`. Ten zvyšuje výchozí penaltu za otevření mezery uvnitř sekvence, která je pouze 7. Důležitým faktorem bylo, aby byla tato penalta větší než penalta za otevření mezery na konci sekvence, kterou upravuje parametr `-gap_op_term` s výchozím nastavením 6,3, které jsem ponechala. Důvodem bylo, že se mezi mými sekvencemi nacházelo i několik neúplných a mezery na koncích těchto sekvencí se tedy daly v alignmentu očekávat. Větší hodnota parametru `-gap_op` také částečně zaručuje otvírání méně mezer, které budou delší (oproti většímu množství krátkých mezer) a to v kombinaci s výchozí penaltou za prodloužení již otevřené mezery, která je 1 pro mezery uvnitř sekvence a 0,9 pro mezery na koncích sekvencí. To bylo vhodné pro případ, že by sekvence stále obsahovaly delší inzerce, které se nepodařilo odstranit v předchozím kroku a nevznikalo příliš velké množství mezer. Testovala jsem několik různých hodnot nastavení parametrů penalizujících otevírání mezer a tyto jsem zvolila na základě nejlepších výsledků.

V případě mitochodriálních genů jsem použila také parametr `-gc_def 2`, který specifikuje genetický kód pro mitochondriální geny obratlovců. Pro nukleární geny jsem použila výchozí standardní genetický kód. Genetický kód udává, jak jsou kodony překládány do aminokyselin. Mitochondriální kód obratlovců se od standardního mírně liší, a to například překladem kodonů AGA a AGG, které slouží jako stop kodony, zatímco standardně kódují arginin. Naopak standardní stop kodon TGA v mitochondriální DNA kóduje tryptofan (Osawa et al., 1989). Výstupními soubory programu jsou dva alignmenty ve FASTA formátu, jeden obsahuje nukleotidové sekvence a druhý jejich aminokyselinový překlad.

2.4.3 Filtrace

K filtraci alignmentu jsem použila nástroj HmmCleaner v. 0.180750 (Baurain, 2025a). Software se zaměřuje na filtraci chybně zarovnaných segmentů individuálních sekvencí, které jsou zdrojem signálu bez evolučního opodstatnění a mohou tak vést k chybným závěrům v navazujících analýzách (Di Franco et al., 2019). Filtrace probíhala na aminokyselinovém alignmentu. Nejdůležitější parametry programu definují detekci heterogen-

ních segmentů. Použila jsem parametr `--large`, který načítá předem definované hodnoty těchto parametrů optimalizované pro datasety s více než 50 sekvencemi. Aminokyseliny, které byly programem vyhodnoceny jako nehomologické, byly nahrazeny mezerou. Protože filtrace proběhla na aminokyselinových sekvencích, bylo potřeba výsledky filtrace ještě přenést zpět na sekvence nukleotidové, což HmmCleaner umožňuje prostřednictvím skriptu transferCleaner.pl. Ten pro všechny odstraněné aminokyseliny nahradí mezerami také příslušné kodony v nukleotidovém alignmentu. Jelikož výstupem byl alignment v ali formátu, následně jsem jej ještě převedla zpět do FASTA formátu s pomocí ali2fasta.pl v. 0.250380 (Baurain, 2025b).

Následovala filtrace na úrovni celých pozic (sloupců) alignmentu. Odstranila jsem všechny pozice s chybějícími daty (mezerami) u více než 70 % sekvencí. K tomu jsem použila nástroj trimAl v. 1.4.rev15 (Capella-Gutiérrez et al., 2009) s nastavením parametru `-gt 0,3`. Pozice s velkým podílem chybějících dat mohou být chybně zarovnané a jejich odstranění často vede ke zlepšení výsledků navazujících analýz (Capella-Gutiérrez et al., 2009). V softwaru PAML, který jsem použila pro výpočet selekčního signálu (viz kapitolu 2.6) je doporučeno odstranit pozice tvořené převážně mezerami (Álvarez-Carretero et al., 2023). V tomto bodě filtrace probíhala již na nukleotidových sekvencích. Vzhledem k předchozí filtraci na aminokyselinové úrovni a následnému přenesení zpět na nukleotidovou proběhla veškerá filtrace na úrovni kodonů. To znamená, že došlo buď k odstranění celého kodonu (jeho nahrazení trojicí mezer), nebo celý kodon zůstal zachován a neobsahuje žádné mezery (v případě posunů čtecího rámce byly kodony doplněny znakem „N“, označující neznámý nukleotid). Díky tomu trimAl vždy případně odstranil všechny tři sloupce tvořící jeden kodon, a nedocházelo tak k nechtěnému posunu čtecího rámce.

2.4.4 Rafinace, finální úpravy a kontrola kvality alignmentu

V dalším kroku jsem použila MACSE program refineAlignment, který slouží k vylepšení již existujícího alignmentu. Parametry programu jsou stejné, jako v případě již použitého alignSequences programu. Hodnotu parametru `-gap_op` jsem nyní zvolila pouze 10, pro větší flexibilitu při vytváření nových mezer v alignmentu, který by již nyní neměl obsahovat žádné dlouhé inzerce, přítomné jen u několika sekvencí. Nastavení penalty za otevírání nových mezer jsem opět zvolila na základě testování několika různých hodnot. V případě mitochondriálních markerů jsem specifikovala mitochondriální genetický kód obratlovců.

Pro výpočet selekčního signálu bylo nutné, aby alignment neobsahoval posuny čtecího rámce (značeny „!“) ani stop kodony (viz kapitolu 2.6). To umožňuje další z MACSE programů – exportAlignment. Použila jsem parametry: `-codonForFinalStop --- -codonForInternalStop NNN -codonForExternalFS --- -codonForInternalFS ---`. Stop kodony na konci sekvence tedy byly nahrazeny mezerami, zatímco stop kodony uvnitř sekvence, které mohly vzniknout např. chybou v sekvenování byly nahrazeny „NNN“, tedy neznámými nukelotidy. Kodony, ve kterých se vyskytoval znak posunu čtecího rámce („!“), byly celé nahrazeny mezerami.

Na závěr jsem naposledy odstranila sloupce s mezerami u více jak 70 % sekvencí opět s využitím nástroje trimAl. Nové sloupce s velkým podílem mezer se mohly objevit při vylepšování alignmentu v refineAlignment a také při nahrazování koncových stop kodonů či posunů čtecích rámců za mezery v programu exportAlignment.

Kontrolu kvality alignmentu jsem provedla vizuálně v programu AliView (Larsson, 2014).

Data v této podobě jsem použila jako vstupní data pro fylogenetickou rekonstrukci (viz kapitolu 2.5).

2.4.5 Extrakce markerů z alignmentů celých genů

V několika případech se u nukleárních genů nepoužíval jako fylogenetický marker celý gen, ale jen jeho část (dále jen marker). Protože cílem této práce byl výpočet selekčního signálu v rámci fylogenetických markerů a ne celých genů, musela jsem část odpovídající markerům z alignmentů extrahovat.

Pokud marker zahrnoval i introny či nekódující části exonů, tyto části jsem ignorovala, protože výpočet selekčního signálu jsem mohla provést pouze pro kódující sekvence (viz kapitolu 2.6). U každého genu jsem postupovala individuálně.

Alignment jsem ponechala beze změn v případě, kdy marker pokrýval celou nebo alespoň většinu kódující sekvence genu.

Pokud marker odpovídal celé délce určitého exonu či exonů, stáhla jsem jejich sekvence pro *Homo sapiens* z databáze NCBI.

V případě, kdy marker tvořila specifická část genu běžně získávaná sekvenováním, vyhledala jsem v příslušných fylogenetických studiích používané primery. Sekvence primerů jsem poté zarovnala opět na sekvenci příslušného markeru pro *Homo sapiens* za použití modulů z knihovny Biopython (Cock et al., 2009). Jednalo se o modul Seq, umožňující snadnou manipulaci se sekvencemi a PairwiseAligner z modulu Align, který umožňuje párový alignment sekvencí. Následně jsem extrahovala primery ohraničený úsek sekvence.

Takto získané sekvence markerů pro *Homo sapiens* jsem následně zarovnala s mými výslednými alignmenty celých genů v programu AliView (Larsson, 2014), který rovnou umožňoval také manuální odstranění nepotřebných částí alignmentů na kodonové úrovni. Po ořezání alignmentu jsem přidanou sekvenci markeru opět odstranila. Při extrakci jsem dbala na zachování čtecího rámce a také na to, aby nedošlo ke vložení nových prázdných sloupců. Ty mohly vznikat v případě, kdy přidaná sekvence markeru obsahovala navíc segment, který byl v mém alignmentu již v některém kroku odstraněn.

Následně jsem ze všech alignmentů odstranila sekvence s více jak 50 % mezer. Pokud měl tedy savec známou méně než polovinu sekvence markeru, byl z alignmentu odstraněn. Takto jsem získala finální data, která jsem použila jako vstupní data pro výpočet selekčního signálu (viz kapitolu 2.6).

2.5 Fylogenetická rekonstrukce

Pro výpočet selekčního signálu je kromě alignmentu potřeba také fylogenetický strom.

Jako vstupní data pro fylogenetickou rekonstrukci jsem použila alignmenty celých genů ještě před samotnou extrakcí markerů, tedy tak, jak jsou popsány na konci podkapitoly 2.4.4.

Fylogenetická rekonstrukce na základě více genů se běžně provádí na jednom alignmentu tvořeném spojením alignmentů jednotlivých genů za sebe. To obvykle vede k přesnějšímu výsledku v porovnání se stromy tvořenými na základě alignmentů individuálních genů, tzv. genové stromy, nebo konsenzem těchto stromů (Gadagkar et al., 2005). Ke

spojení alignmentů jsem použila Perl skript catfasta2phyml (Nylander, 2025) s multifunkčním parametrem `-cfs`. Výsledný spojený alignment vygeneroval ve FASTA formátu (`-f; --fasta`), sekvence téhož savce spojené za sebou na jednom řádku na základě přesné shody v názvu sekvencí (`-s; --sequential`). Sekvence se spojily i v případě, že se počet sekvencí v jednotlivých genových alignmentech lišil a chybějící geny byly nahrazeny mezerami (`-c; --concatenate`).

Skript rovněž vygeneroval informace o spojení alignmentů, které zahrnovaly názvy jednotlivých genových alignmentů a jejich odpovídající intervaly v nově vytvořeném spojeném alignmentu. Tyto údaje jsem následně využila při fylogenetické rekonstrukci. Pro jejich uchování jsem vytvořila textový soubor, kde každý řádek souboru odpovídal jednomu genovému alignmentu a obsahoval informace ve formátu: „DNA, BRCA1 = 13720-19308“. Nejprve bylo specifikováno, že se jedná o DNA sekvenci (a ne např. o protein), následoval název genu a interval vymezující jeho pozici ve spojeném alignmentu.

Fylogenetickou rekonstrukci jsem provedla s pomocí programu IQ-TREE v. 2.3.4-mpi (Minh et al., 2020), který k inferenci fylogenetických stromů využívá metodu maximální věrohodnosti.

Výběr substitučního modelu jsem provedla pomocí funkce ModelFinder (Kalyaanamoorthy et al., 2017), která je součástí programu IQ-TREE. Tato funkce testuje všechny dostupné substituční modely na fixním počátečním stromě vytvořeném metodou maximální parsimonie a vybírá model s nejnižší hodnotou BIC (Bayesian Information Criterion). BIC vychází z věrohodnostní funkce, ale zároveň penalizuje modely s velkým počtem parametrů, čímž pomáhá zabránit přetrénování modelů (Chakrabarti a Ghosh, 2011).

Díky specifikaci segmentů (partitions) (Chernomor et al., 2016) mohl ModelFinder hledat optimální substituční model pro každý gen zvlášť, místo aby aplikoval jednotný model na celý dataset. To umožnilo zohlednit rozdílnou evoluční dynamiku jednotlivých genů. Pokud však některé geny vykazovaly podobný vývoj a sdílely stejný nejlepší substituční model, ModelFinder je mohl sloučit dohromady.

Podpory jednotlivých uzlů byly odhadnuty pomocí ultra-rychlého bootstrapu (Hoang et al., 2018) s 2000 replikacemi. Hodnota bootstrapové podpory vyjadřuje, kolikrát byla daná větev pozorována při opakovaném generování stromu na základě převzorkovaných dat. Za dostatečně podpořený je považován uzel s hodnotou alespoň 95, zatímco uzly s nižšími podporami je nutné interpretovat s rezervou (Ojha et al., 2022).

Výsledkem fylogenetické inference byl nezakořeněný strom v Newick formátu. Strom jsem následně ještě restrukturovala tak, aby měl u základny vhodnou monofyletickou skupinu savců. Automaticky se u základny stromu zobrazuje první savec dle abecedy, což způsobovalo problémy při označování podstromů při výpočtu selekce v rámci taxonů (viz podkapitolu 2.6.3). Ke změně struktury Newick formátu jsem použila program FigTree v. 1.4.4 (Rambaut, 2018), který jsem využila také při vizualizaci stromu. K odstranění kořene stromu vzniklého při úpravě ve FigTree jsem použila funkci unroot() z balíku ape v softwaru R (Paradis a Schliep, 2019).

2.6 Výpočet selekčního signálu

Softwarový balík PAML (Phylogenetic Analysis by Maximum Likelihood; Yang (2007)) je považován za zlatý standard pro detekci selekce na molekulární úrovni a je jedním

z nejčastěji používaných nástrojů v této oblasti (Salemi et al., 2009). Detekci selekce s využitím kodonových substitučních modelů umožňuje jeho program codeml.

2.6.1 Modely

Codeml nabízí několik typů kodonových substitučních modelů určených k analýze selekčního tlaku. Modely pozicí (site models) umožňují odlišné hodnoty ω pro různé pozice alignmentu, což umožňuje detekci konkrétních pozicí–kodonů pod vlivem pozitivní selekce. Modely větví (branch models) umožňují, aby se ω lišila mezi větvemi stromu. Tím umožňují detekci pozitivní selekce působící pouze na vybrané vývojové větve. Modely větví-pozicí, (branch-site models), pak kombinují oba přístupy – umožňují detekci pozitivní selekce specifické pro vybranou vývojovou větev a zároveň určí, o které kodony se jedná. Posledním druhem jsou kladové modely (clade models). Ty umožňují srovnání působení selekce mezi různými skupinami organismů (klady). Klad zahrnuje vždy všechny potomky pocházející od jednoho společného předka (Yang, 2020).

Srovnání modelů testem poměru věrohodností

Test poměru věrohodností (Log-likelihood ratio test; LRT) je statistický test sloužící ke srovnání dvou matematických modelů. Pro modely musí platit, že jsou vnořené, tzn. jednodušší (nulový) model je speciálním případem modelu složitějšího (alternativního), vzniklý např. fixací hodnot některých jeho parametrů na konkrétní hodnotu. Testovou statistiku LRT získáme porovnáním maximalizovaných věrohodností obou modelů vztahem

$$2\Delta\ell = 2\ln\left(\frac{L_1}{L_0}\right) = 2\left(\ln L_1 - \ln L_0\right), \tag{2.1}$$

kde L_0 je maximální věrohodnost jednoduššího modelu a L_1 modelu složitějšího. Statistika $2\Delta\ell$ vyjadřuje, do jaké míry popisuje složitější model data lépe než model nulový (Whelan et al., 2001).

Testová statistika $2\Delta\ell$ se porovnává s $\chi^2(df)$ rozdělením pravděpodobnosti, kterým se řídí za předpokladu platnosti nulového modelu. Počet stupňů volnosti df rozdělení je roven rozdílu počtu volných parametrů obou modelů (Yang et al., 2000). Je-li však hodnota $2\Delta\ell$ příliš extrémní na to, aby pocházela z tohoto rozdělení, zamítáme platnost nulového modelu ve prospěch složitějšího modelu. Můžeme tak rozhodnout na základě p-hodnoty představující pravděpodobnost získání alespoň stejně extrémního výsledku za předpokladu platnosti nulového modelu (Wasserstein a Lazar, 2016), tedy platí

$$p = P(\chi^2(df) \geq 2\Delta\ell) = 1 - F^{df}_{\chi^2}(2\Delta\ell), \tag{2.2}$$

kde $F^{df}_{\chi^2}$ je distribuční funkce χ^2 s df stupni volnosti. V této práci jsem uvažovala hladinu významnosti $\alpha = 0,05$, se kterou se výsledná p-hodnota porovnává. V případě, že je p-hodnota menší než stanovená hladina významnosti α, zamítla jsem platnost nulového modelu. Výpočty p-hodnot jsem prováděla v jazyce Python v. 2.7.6-gcc s využitím funkcí modulu scipy.stats (Virtanen et al., 2020).

M0

Základní, nejjednodušší model M0, který uvažuje jednotný poměr ω pro všechny pozice sekvence, byl představen ve studii Yang a Nielsen (1998). Model s pomocí Markovského procesu se spojitým časem popisuje substituce mezi 61 kodony z 64 možných s výjimkou tří stop kodonů (v případě standardního genetického kódu). Stop kodony nekódují aminokyselinu a nejsou proto v proteinu povoleny. Pro libovolné dva kodony model popisuje okamžitou substituční rychlost mezi kodony i a j ($i \neq j$) jako:

$$q_{ij} = \begin{cases} 0, & \text{pokud se kodony liší ve více než jedné pozici,} \\ \pi_j, & \text{pro synonymní transverze,} \\ \kappa\pi_j, & \text{pro synonymní tranzice,} \\ \omega\pi_j, & \text{pro nesynonymní transverze,} \\ \omega\kappa\pi_j, & \text{pro nesynonymní tranzice,} \end{cases} \tag{2.3}$$

kde π_j je rovnovážná frekvence kodonu j (očekávaná frekvence výskytu kodonu za předpokladu genetické rovnováhy v populaci), která může být považována za volný parametr, nebo vypočítána z nukleotidových frekvencí. V této práci rovnovážné frekvence kodonů neuvažuji jako volné parametry, zejména kvůli výpočetní náročnosti. κ je poměr rychlostí tranzicí a transverzí a ω je poměr rychlostí nesynonymních a synonymních substitucí představující míru selekce působící na kódující sekvence (Yang a Nielsen, 1998; Bielawski a Yang, 2003; Yang, 2020).

Pravděpodobnost změny kodonu i na kodon j za čas t je $p_{ij}(t)$ a pro matici pravděpodobností jednotlivých změn platí

$$P(t) = \{p_{ij}(t)\} = e^{\mathbf{Q}t}, \tag{2.4}$$

kde $\mathbf{Q} = \{q_{ij}\}$ a čas nebo délka větve t vyjadřuje očekávaný počet nukleotidových substitucí na jeden kodon, zprůměrovaný pro všechny pozice (Lió a Goldman, 1998; Yang et al., 2000; Bielawski a Yang, 2003).

Parametr π_j slouží k zohlednění vlivu preferenčního využívaní kodonu j (codon usage bias), tedy frekvenci výskytu kodonu v sekvenci. Model nedokáže rozlišit, zda jsou substituce poháněny mutací či selekcí, ale nepředpokládá, že jsou synonymní substituce neutrální. Výrazné preferenční využívání kodonů může být způsobeno jak vlivem mutací, tak selekce (např. kvůli zvýšení efektivity translace (Bulmer, 1991)) a může značně ovlivnit rychlosti synonymních substitucí. Nicméně vynásobíme-li substituční rychlost kodonu i na j parametrem π_j udávajícím frekvenci kodonu j, budou odhady substitučních rychlostí zcela zohledňovat preferenční využívání kodonu bez ohledu na jeho původ (Yang a Bielawski, 2000).

Při transverzi dochází ke změně pyrimidinového nukleotidu na purinový nebo naopak (A, G $\leftrightarrow$ T, C), zatímco u tranzice ke změně nedochází (A $\leftrightarrow$ G nebo C $\leftrightarrow$ T). Protože tranzice na třetí kodonové pozici bývají synonymní s větší pravděpodobností než transverze, nezohlednění vlivu poměru κ (poměr rychlostí tranzicí a transverzí) vede k podhodnocení poměru ω (Yang a Bielawski, 2000). Je také známo, že k tranzicím dochází častěji než k transverzím a z toho důvodu je substituční rychlost násobena parametrem κ v případě, představuje-li změna tranzici (Bielawski a Yang, 2003).

Vliv selekce se projevuje u substitucí způsobujících změnu aminokyseliny. Proto se substituční rychlost při nesynonymních změnách násobí parametrem ω (Bielawski a Yang, 2003).

Poměr ω je stejný pro všechny pozice alignmentu v celém fylogenetickém stromě. Tento model je samozřejmě pro naprostou většinu sekvencí příliš zjednodušující a očekává se jeho zamítnutí ve prospěch složitějšího modelu, umožňujícího variabilitu ω mezi pozicemi alignmentu, větvemi stromu nebo obojím. Je mnohem pravděpodobnější, že pozitivní selekce působí pouze na některé pozice, zatímco nesynonymním změnám na ostatních pozicích je spíše zabraňováno, aby nedocházelo k narušení funkce proteinu a zachovala se jeho sekundární a terciální struktura. Je tedy možné, že negativní selekce na většině pozic a větví převáží signál pozitivní selekce, kterou poté nejsme schopni detekovat (Felsenstein, 2004).

Zjednodušeně se dá říct, že $\omega < 1$ značí selekci negativní, $\omega > 1$ selekci pozitivní a $\omega = 1$ neutrální evoluci. Nicméně průměrná hodnota ω pro všechny pozice alignmentu a větve stromu téměř nikdy nevyjde větší než jedna. Z toho důvodu byly vyvinuty modely zaměřující se na detekci pozitivní selekce působící pouze na určité vývojové větve a/nebo konkrétní kodonové pozice alignmentu (Yang, 2020).

Pozicové modely

Pozicové modely rozšiřují základní model M0 a dovolují, aby se poměr ω lišil mezi jednotlivými kodonovými pozicemi alignmentu (Yang et al., 2000). Nechť h je pozice (kodon) alignmentu a n je celkový počet kodonů v alignmentu v sekvenci $h = 1, 2, \ldots, n$. Dále $\omega^{(h)}$ je poměr platný pro pozici h a $\mathbf{x}_h$ je vektor kodonů všech sekvencí alignmentu na této kodonové pozici. Pak můžeme definovat relativní okamžitou substituční rychlost z kodonu i na kodon j ($i \neq j$) na pozici h jako

$$q_{ij}^{(h)} = \begin{cases} 0, & \text{pokud se kodony liší ve více než jedné pozici,} \\ \pi_j, & \text{pro synonymní transverze,} \\ \kappa\pi_j, & \text{pro synonymní tranzice,} \\ \omega^{(h)}\pi_j, & \text{pro nesynonymní transverze,} \\ \omega^{(h)}\kappa\pi_j, & \text{pro nesynonymní tranzice,} \end{cases} \tag{2.5}$$

kde parametry π_j a κ mají stejný význam a funkci jako v rovnici 2.3. Pravděpodobnost změny kodonu i na kodon j za čas t je $p_{ij}(t)$ a matice pravděpodobností změn

$$P(t) = \{p_{ij}(t)\} = e^{\mathbf{Q}t}, \tag{2.6}$$

kde $\mathbf{Q} = \{q_{ij}^{(h)}\}$ a čas nebo délka větve t vyjadřuje očekávaný počet nukleotidových substitucí na jeden kodon, zprůměrovaný pro všechny pozice (Yang et al., 2000; Anisimova et al., 2001; Bielawski a Yang, 2003).

Použití specifického parametru ω pro každou pozici zvlášť by představovalo příliš velké množství parametrů v modelu. Z toho důvodu se za účelem umožnění heterogenního poměru ω mezi pozicemi běžně používají rozdělení pravděpodobnosti. Předpokládejme K tříd pozic v sekvenci alignmentu s proporcemi $p_0, p_1, \ldots, p_{K-1}$ a poměry $\omega_0, \omega_1, \ldots, \omega_{K-1}$

Chceme spočítat pravděpodobnost pozorování dat $\mathbf{x}_h$ na pozici h. Marginální pravděpodobnost je průměr podmíněných pravděpodobností pro toto rozdělení poměrů ω pro jednotlivé třídy pozic

$$P(\mathbf{x}_h) = \sum_{k=0}^{K-1} p_k P(\mathbf{x}_h|\omega_k). \tag{2.7}$$

Předpokládáme, že je substituční proces mezi kodony nezávislý, a proto můžeme definovat logaritmus věrohodnosti jako sumu logaritmů marginálních pravděpodobností všech pozic v sekvenci

$$\ell = \sum_{h=1}^{n} \log P(\mathbf{x}_h). \tag{2.8}$$

Poté co maximalizací věrohodnostní funkce získáme odhady parametrů, můžeme empirickým Bayesovským přístupem určit, ze které třídy pozice nejpravděpodobněji pochází. Posteriorní pravděpodobnost, že pozice h s daty $\mathbf{x}_h$ patří do třídy k je

$$P(\omega_k|\mathbf{x}_h) = \frac{p_k P(\mathbf{x}_h|\omega_k)}{P(\mathbf{x}_h)} = \frac{p_k P(\mathbf{x}_h|\omega_k)}{\sum_j p_j P(\mathbf{x}_h|\omega_j)}. \tag{2.9}$$

Teoreticky může mít poměr ω také spojité rozdělení $f(\omega)$. V takovém případě bude suma v rovnici 2.7 nahrazena integrálem. Protože ale nebyl nalezen uskutečnitelný způsob provedení výpočtu, využívá se namísto toho aproximace diskrétním rozdělením. Diskrétní rozdělení má K kategorií, každou s pravděpodobností $1/K$ a spojité rozdělení poměru ω v rámci jednotlivých kategorií je reprezentováno mediánem poměru ω uvnitř každé kategorie. Pravděpodobnost pozorování dat $\mathbf{x}_h$ pak můžeme aproximovat jako

$$P(\mathbf{x}_h) = \int_0^\infty f(\omega) P(\mathbf{x}_h|\omega) d\omega \approx \frac{1}{K} \sum_{k=0}^{K-1} P(\mathbf{x}_h|\omega_k) \tag{2.10}$$

(Yang et al., 2000).

M1a a M2a

Model M1a se nazývá téměř neutrální model. Oproti modelu M0 dělí pozice alignmentu do dvou kategorií. Proporce pozic příslušných jednotlivým kategoriím udávají parametry p_0 a p_1. Je zřejmé, že parametry proporcí se v součtu rovnají jedné, a proto je volným parametrem pouze p_0 a p_1 můžeme dopočítat s využitím vztahu $p_1 = 1 - p_0$. Pro část pozic odpovídající proporci p_1 platí ω s fixní hodnotou jedna ($\omega_1 = 1$) a jedná se tedy o pozice s neutrální evolucí. Druhá část pozic proporce p_0 má $\omega_0 < 1$ a představují tak pozice s negativní selekcí. Model má tedy celkem 2 parametry – ω_0 a p_0.

Model M2a, nebo také model pozitivní selekce, rozšiřuje model M1a o jednu další kategorii pozic alignmentu. Pro třetí skupinu pozic platí, že $\omega_2 > 1$, čímž model předpokládá, že se některé pozice vyvíjejí pod vlivem pozitivní selekce. V tomto modelu proporce p_1 představuje další parametr zatímco p_2 se opět dopočítá vztahem $p_2 = 1 - p_0 - p_1$. Model má tedy celkem 4 parametry – ω_0, ω_2, p_0 a p_1.

Porovnáním logaritmů věrohodností modelů M2a (alternativní model) a M1a (nulový model), tak lze detekovat výskyt pozitivní selekce v některých kodonech. Pokud je v datech přítomný signál pozitivní selekce, model M2a by měl data popisovat výrazně lépe než nulový model, který neuvažuje žádné pozice s pozitivní selekcí. Logaritmus věrohodnosti modelu M2a by tak měl být větší než logaritmus věrohodnosti M1a a o tom, jestli je rozdíl statisticky významný, rozhodujeme s pomocí LRT. Počet volných parametrů je 2. Porovnáním těchto modelů jsem tedy byla schopná detekovat pozitivní selekci v rámci jednotlivých markerů napříč všemi savci.

Samotný významný výsledek LRT ale ještě není důkazem pozitivní selekce a je třeba se dívat také na příslušný výsledný poměř ω. Pouze v případě, že je jeho hodnota větší než jedna, lze říct, že detekujeme pozitivní selekci (Anisimova et al., 2002).

M7 a M8

Model M7 (Beta model) umožňuje heterogenní ω mezi kodonovými pozicemi a je založený na rozdělení Beta(p,q). Tvar rozdělení udávají dva kladné parametry p a q. Model M7 tedy odhaduje ω jako spojitou náhodnou veličinu $\omega \sim$ Beta(p,q) s hustotou:

$$f(\omega) = \frac{\omega^{p-1}(1-\omega)^{q-1}}{B(p,q)}, \quad 0 < \omega < 1, \tag{2.11}$$

kde $B(p,q)$ je beta funkce. Hustota beta rozdělení může na intervalu $(0,1)$ nabývat různých tvarů (např. „L“, „U“ nebo obrácené „U“). Vzhledem k tomu, že ω nemůže nabývat hodnot větších než jedna, tento model nedovoluje výskyt pozicí pod pozitivní selekcí.

Model M8 (Beta+ω model) rozšiřuje model M7 a přidává jednu novou třídu pozic představující právě pozice pod pozitivní selekcí. Proporce pozic odpovídající nově přidanému parametru p_0 má poměry $\omega \sim$ Beta(p,q) stejně jako v modelu M7. Zbylé pozice odpovídající $p_1 = 1 - p_0$ mají $\omega_s > 1$ a představují pozitivní pozice. Model M8 má tedy celkem 4 parametry, z toho dva jsou stejné jako v M7 – p a q a dva nové parametry – p_0 a ω_s.

Modely M8 a M7 lze mezi sebou porovnat ekvivalentně jako v případě modelů M1a a M2a a odpovědět na stejnou otázku, zda se v alignmentu vyskytují pozice pod vlivem pozitivní selekce. Test modelů M1a a M2a je však striktnější a může se tedy stát, že příliš slabý signál pozitivní selekce neodhalí (Yang, 2020). Při srovnání modelů M7 a M8 máme opět dva volné parametry.

M2a_rel a kladový model C

Model M2a_rel je téměř identický s již popisovaným modelem M2a s jediným rozdílem – ω pro třetí skupinu pozic nemusí být nutně větší než 1 a podmínka se upravuje na $\omega_2 > 0$. Pro první třídu pozic p_0 platí stále $\omega_0 < 1$ a pro druhou p_1 je poměr $\omega_1 = 1$ Model byl vytvořen primárně jako nulový model pro cmC. Ten se původně porovnával s modelem M1a, ale následné studie prokázaly, že M2a_rel jako nulový model při LRT představuje silnější test (Weadick a Chang, 2012).

Kladový model C (clade model C; cmC) (Bielawski a Yang, 2004b) rozšiřuje model M2a_rel tak, aby bylo možné modelovat selekci specifickou pro určitou zájmovou pod-

skupinu ve fylogenetickém stromě. Jedná se tak o specifický typ modelu větví-pozic, ale bývá řazen do speciální kategorie kladových modelů. Tuto podskupinu, neboli klad všech organismů pocházejících od jednoho společného předka, je třeba ve fylogenetickém stromě označit. To je možné pomocí anotace „#1“, kterou se musí označit všechny jednotlivé větve patřící do zájmového kladu. Anotace větví v pozadí „#0“ není nutná a automaticky sem patří všechny ostatní neoznačené větve. Alternativně lze využít také anotaci „$1“, kterou je možné označit pouze větev společného předka pro výběr celého kladu. Lze označit i více než jeden klad, ale výpočetní náročnost modelu s každou další skupinou rychle narůstá. Proto jsem v této práci označovala vždy pouze jediný klad.

Model cmC má většinu parametrů shodných s modelem M2a_rel. Třída pozic p_0 s $\omega_0 < 1$ i třída p_1 s poměrem $\omega_1 = 1$ jsou pro označený klad i zbytek stromu totožné. Tyto dvě třídy představují pozice, které prochází stejnou evolucí v celém stromě. Podstatná je poslední, třetí třída pozic, kde se od sebe poměry ω_3 kladu a ω_2 zbytku stromu mohou lišit, stále však platí $\omega_2, \omega_3 > 0$. Tyto pozice tak představují pozice, které podléhají jinému selekčnímu tlaku v označeném kladu oproti zbytku stromu. Model cmC má tedy oproti M2a_rel jeden parametr navíc a tím je právě poměr ω_3 specifický pro označený klad.

Významnost rozdílného působení selekce mezi vybraným kladem a větvemi v pozadí se hodnotí pomocí LRT (s jedním stupněm volnosti) a následně také dle velikosti rozdílu $\Delta\omega = \omega_3 - \omega_2$.

Srovnáním parametrů ω_2 a ω_3 lze určit, jak odlišná selekce působí na klad. Případ, kdy $\omega_2 << \omega_3$ značí uvolněnou selekci u kladu. Větší poměr ω_3 u kladu může, ale nemusí být zapříčiněný působením pozitivní selekce a dostatečným důkazem její přítomnosti u kladu je pouze $\omega_3 > 1$ (případně $\omega_2 > 1$ pro větve v pozadí). Naopak $\omega_2 >> \omega_3$ může u kladu indikovat silnější působení negativní selekce a podobné hodnoty ω_2 a ω_3 naznačují podobnou evoluci kladu i zbytku stromu.

Modely větví-pozicí

Modely větví-pozicí při modelování selekčního signálu poskytují největší flexibilitu a umožňují, aby se poměr ω lišil mezi větvemi stromu i pozicemi alignmentu.

Modely větví-pozicí A a A1

Model větví-pozicí A (branch-site model A; bsA) umožňuje, aby na některých pozicích vybraných zájmových větvích stromu docházelo k pozitivní selekci, zatímco na ostatních větvích v pozadí jsou podrobeny negativní nebo neutrální selekci. BsA rozděluje pozice do čtyř tříd. První třída odpovídající proporci p_0 zahrnuje konzervované pozice ($0 < \omega_0 < 1$) a druhá třída s proporcí p_1 obsahuje pozice neutrální ($\omega_1 = 1$). Obě tyto třídy jsou společné pro celý strom, tedy jak pro zájmové větve, tak pro větve v pozadí. Další dvě třídy umožňují výskyt pozitivní selekce pouze u zájmových větví. Třetí třída obsahuje pozice pod negativní selekcí ($0 < \omega_0 < 1$) u větví v pozadí, které jsou ale pod pozitivní selekcí ($\omega_2 \geq 1$) u zájmových větví. Čtvrtá třída zahrnuje pozice pozitivní ($\omega_2 \geq 1$) v zájmových větvích, ale neutrální ve zbytku stromu. Zbytek pozic ($1 - p_0 - p_1$) je rozdělen mezi třídy p_{2a} a p_{2b} podle váhy tříd p_0 (pro třetí třídu s negativními pozicemi ve větvích v pozadí) a p_1 (pro čtvrtou třídu s neutrálními pozicemi u větví v pozadí), viz Tab. 1.

Model větví-pozicí A1 je nulovým modelem pro bsA a liší se od něj pouze tím, že neumožňuje ve třetí a čtvrté třídě pozic u zájmových větví pozitivní selekci, ale pouze neutralitu s fixním poměrem $\omega_2 = 1$.

LRT modelů bsA a bsA1 se nazývá také test pozitivní selekce, protože představuje přímý test pozitivní selekce na zájmových vývojových větvích (Zhang, 2005).

Nulové rozdělení je směsí v poměru 50:50 mezi bodovým rozdělením v nule a rozdělením χ^2_1. *P*-hodnota na základě tohoto směsného rozdělení se počítá jako *p*-hodnota s pomocí rozdělení χ^2_1 dělená dvěma. Doporučuje se však použít pouze rozdělení χ^2_1 namísto směsi, čímž se pomáhá předcházet problémům způsobeným porušením předpokladů modelu. Mezi předpoklady modelu, které mohou být porušeny je například omezení, které klade na větve v pozadí tím, že u nich nedovoluje pozitivní selekci (Yang, 2020; Zhang, 2005).

Konvergence modelů

Modely implementované v codeml mohou skončit v lokálním maximu nebo vykazovat problémy s konvergencí. Proto se doporučuje program spouštět vždy několikrát a vybrat výsledek s největším logaritmem věrohodnosti. Všechny modely jsem počítala vždy alespoň dvakrát s výchozím nastavením parametrů uvedeným v Tab. 1.

Model M8 jsem navíc počítala také s výchozím nastavením `omega = 0,5`, protože se doporučuje model spouštět s různým počátečním nastavením `omega`, nejlépe < 1 i > 1 (Bielawski a Yang, 2003). Stejně jsem postupovala také u bsA modelu.

Model cmC, který je nejproblematičtější, jsem počítala celkem pětkrát, ve třech případech jsem zvolila také odlišné počáteční hodnoty κ a ω, a to `kappa = 6; omega = 1`, dále `kappa = 1,5; omega = 0,2` a nakonec `kappa = 1,5; omega = 2` (Bielawski a Yang, 2004a; Yang, 2005). Nulový model M2a_rel jsem spouštěla navíc také s počátečními hodnotami `kappa = 1,5; omega = 0,2`.

2.6.2 Vstupní data

Alignment

Vstupní alignment by měl být ve PHYLIP formátu. Hlavička souboru obsahuje dvě čísla oddělená mezerou, která představují počet sekvencí a délku alignmentu. Na následujících řádcích jsou jednotlivé sekvence začínající názvem a následované samotnými zarovnanými sekvencemi.

Důležitý je také správný alignment sekvencí, který nesmí obsahovat introny, nekódující úseky ani stop kodony. Stejně tak je vhodné z alignmentu odstranit pozice s příliš velkým zastoupením mezer (viz kapitolu 2.4) (Álvarez-Carretero et al., 2023).

Fylogenetický strom

Vstupní fylogenetický strom v Newick formátu by neměl obsahovat bootstrapové podpory uzlů a ani délky větví, které většinou nejsou pro analýzu potřebné a v některých modelech mohou narušit označování uzlů (Álvarez-Carretero et al., 2023).

Názvy druhů (listů) ve stromu se musí přesně shodovat s názvy sekvencí v alignmentu a strom nesmí obsahovat oproti alignmentu žádné druhy navíc (Yang, 2020).

Použití zakořeněného či nezakořeněného stromu se může lišit v závislosti na modelu. Naprostá většina modelů pracuje s nezakořeněnými stromy s výjimkou ve speciálních případech modelů větví a větví-pozicí. Pokud se očekává, že dvě větve okolo základny stromu prochází jiným evolučním procesem (s jiným poměrem ω), jinými slovy jedna větev je označena jako zájmová, zatímco druhá je v pozadí, je třeba použít zakořeněný strom. V případě, že jsou větve okolo základny obě označené jako zájmové nebo jsou obě větve v pozadí, používá se nezakořeněný strom jako ve všech ostatních případech (Álvarez-Carretero et al., 2023). PAML považuje strom s bifurkací u kořene za zakořeněný, a v případě, že má strom u kořene alespoň trifurkaci, považuje ho za nezakořeněný (Yang, 2020).

V mém výsledném nezakořeněném stromě se u základny nacházela monofyletická skupina hlodavců a primátů, řád hmyzožravců a zbytek řádů savců. Z toho důvodu jsem pro řád Eulipotyphla při výpočtu modelů bsA, bsA1 a cmC musela použít zakořeněný strom (viz Obr. 1).

Kontrolní soubor

Kontrolní soubor obsahuje informace a nastavení potřebná ke specifikaci konkrétní analýzy v codeml. Jedná se textový soubor obsahující jednotlivá nastavení řádek po řádku. Nejdůležitějším nastavením je volba modelu, která probíhá kombinací parametrů `model` a `NSsites` (viz Tab. 1). Nastavení následujících parametrů bylo totožné pro všechny použité modely (šedě jsou značeny některá další možná nastavení) (Yang, 2020):

`Cleandata`	= 0	Ponechává i pozice alignmentu, které obsahují nějaké mezery či nejednoznačné nukelotidy.
	= 1	Odstraní z alignmentu všechny takové pozice.
`seqtype`	= 1	Alignment obsahuje kódující nukleotidové sekvence.
	= 2	Alignment obsahuje aminokyselinové sekvence.
`CodonFreq`	= 2	Parametr π_j je spočítán z průměrných nukleotidových frekvencí zvlášť pro tři kodonové pozice.
	= 0	π_j stejný po všechny kodony ($1/61$ pro standardní kód).
	= 1	π_j je spočítán z průměrných nukleotidových frekvencí.
	= 3	π_j jsou volné parametry (celkem 60).
`fix_blength`	= 0	Délky větví stromu jsou ignorovány a budou odhadnuty.
	= 1	Délky větví stromu jsou použity jako iniciální při odhadování.
	= 2	Délky větví stromu jsou fixní.
`icode`	= 0	Standardní genetický kód.
	= 1	Mitochondriální genetický kód obratlovců.
`fix_kappa`	= 0	Parametr κ bude odhadován, iniciální hodnota je nastavena parametrem `kappa`.
	= 1	Fixuje κ na hodnotu parametru `kappa`.
`kappa`	= 2	Výchozí (nebo fixní) hodnota parametru κ.

`fix_omega`	= 0	Parametr ω bude odhadován, iniciální hodnota je nastavena parametrem `omega`.
	= 1	Fixuje ω na hodnotu parametru `omega`.

Tab. 1: Parametry použitých kodonových substitučních modelů. Volné parametry jsou uvedeny bez závorky a jejich počet je uveden ve sloupci „#p“. Tabulka obsahuje také specifická nastavení jednotlivých modelů v kontrolním souboru codeml.

			Nastavení codeml		
Model	**#p**	**Parametry modelu**	`model`	`NSsites`	`omega`
M1a (téměř neutrální)	2	p_0; $(p_1 = 1 - p_0)$ ω_0; $(\omega_1 = 1)$	0	1	0,7
M2a (pozitivní selekce)	4	p_0; p_1; $(p_2 = 1 - p_0 - p_1)$ $\omega_0 < 1$; $(\omega_1 = 1)$; $\omega_2 > 1$	0	2	1,7
M7 (beta)	2	p; q	0	7	0,7
M8 (beta+ω)	4	p_0; $(p_1 = 1 - p_0)$ p; q; $\omega_s > 1$	0	8	1,7
M2a_rel (nulový pro cmC)	4	p_0; p_1; $(p_2 = 1 - p_0 - p_1)$ $\omega_0 < 1$; $(\omega_1 = 1)$; $\omega_2 > 0$	0	22	1,7
cmC (klad)	5	p_0; p_1; $(p_2 = 1 - p_0 - p_1)$ $\omega_0 < 1$; $(\omega_1 = 1)$; $\omega_2 > 0$; $\omega_3 > 0$	3	2	0,7
bsA1 (nulový pro bsA)	3	p_0; p_1; $(p_2 = (1 - p_0 - p_1)p_0/(p_0 + p_1))$ $(p_3 = (1 - p_0 - p_1)p_1/(p_0 + p_1))$ $0 < \omega_0 < 1$; $(\omega_1 = 1)$; $(\omega_2 = 1)$	2	2	1 (fixní)
bsA (branch-site A)	4	p_0; p_1; $(p_2 = (1 - p_0 - p_1)p_0/(p_0 + p_1))$; $(p_3 = (1 - p_0 - p_1)p_1/(p_0 + p_1))$ $0 < \omega_0 < 1$; $(\omega_1 = 1)$; $\omega_2 \geq 1$	2	2	1,7

2.6.3 Automatizace práce s codeml

ETE (Environment for Tree Exploration) nabízí souhrn mnoha nástrojů pro rekonstrukci, analýzu a vizualizaci fylogenetických stromů i alignmentů (Huerta-Cepas et al., 2016).

Jedním z těchto nástrojů je ete-evol zaměřený na usnadnění a automatizaci práce s codeml. Automaticky generuje kontrolní soubory specifické pro jednotlivé codeml modely (tak, jak jsou nastavení uvedená v podkapitole 2.6.2), ale stejně tak umožňuje i volbu vlastního nastavení. Codeml vyžaduje alignment ve formátu PHYLIP, ale ete-evol umožňuje použít také FASTA formát. Z fylogenetického stromu automaticky odstraňuje délky větví a podpory uzlů, zůstávají tedy pouze názvy listů. Odstranit délky větví a bootstrapové podpory uzlů je doporučeno přímo v PAML manuálu kvůli možným konfliktům se značením uzlů u některých modelů (Álvarez-Carretero et al., 2023). Dále je s využitím ete-evol možný také paralelní výpočet několika modelů najednou. Pomocí ete-evol lze snadno označovat podstromy pro výpočty cmC jednotlivých taxonů. Stačí zadat dva takové zástupce dané podskupiny, aby jejich první společný předek byl zároveň prvním společným předkem celého taxonu, který chceme označit. Ete-evol pak označí všechny listy i uzly pocházející od společného předka těchto dvou savců, včetně něj. Ete-evol tak zjednodušuje celý proces analýzy a zajišťuje lepší reprodukovatelnost a flexibilitu při práci s codeml.

Dalším ETE programem, který jsem využila byl ete-mod, který slouží k modifikaci fylogenetických stromů. Vstupní strom pro codeml totiž nesmí obsahovat žádné listy, které

by zároveň neměly svoji sekvenci v alignmentu. Vzhledem k tomu, že v alignmentech často některé sekvence chyběly, musely být druhy s chybějící sekvencí daného genu ze stromu odstraněny.

3 Výsledky

3.1 Výběr taxonů obratlovců

Dostupný referenční genom v databázi NCBI mělo celkem 222 savců. Z toho jsem jich celkem 5 vyřadila. Jedním z nich byl hybrid *Bos indicus* x *Bos taurus*. Oba rodičovké druhy měly také svůj vlastní referenční genom a zahrnula jsem je do analýzy samostatně. Dalšími čtyřmi savci byli *Ceratotherium simum*, *Enhydra lutris*, *Pantholops hodgsonii* a *Sorex cinereus*, kteří neměli v databázi dostupné sekvence žádných nukleárních genů, které představovaly většinu dat potřebných pro tuto práci.

Zbylých 217 savců zastupovalo celkem 23 řádů, z nichž jsem analyzovala pouze řády obsahující minimálně čtyři zástupce. Podmínku splňovalo pouze 7 řádů, dohromady obsahujících 191 savců. Řád Carnivora jsem neanalyzovala jako celek, protože bylo možné jej na základě výsledného stromu rozdělit na dva rovnoměrně zastoupené podřády – Caniformia a Feliformia.

Posledním analyzovaným taxonem byli kytovci (Cetacea). Infrařád kytovců zahrnuje vodní savce, kteří se vyvinuli před zhruba 50 miliony lety od suchozemského předka (Thewissen et al., 2009). Přechod na vodní prostředí vyžadoval velkou míru anatomické a fyziologické adaptace, ale nutná byla také adaptace na nové patogeny. To muselo vést ke zvýšení selekčního tlaku na genomy kolonizujících druhů a to především na geny spojené s imunitním systémem (Ishengoma a Agaba, 2017). U imunitního genu *RAG1*, který je jedním z analyzovaných genů v této práci, byla dříve detekována uvolněná selekce oproti ostatním savcům (Dias a Nery, 2020).

Analyzovaných taxonů tedy bylo celkem 9 (viz Tab. 2). Seznam všech druhů savců a jejich taxonomii obsahuje Příloha 1.

3.2 Identifikace fylogenetických markerů

Studií, které by se zaměřovaly na fylogenetické markery obecně existuje více, ale jen málo z nich se zaměřuje na kódující markery živočichů. Jednou z nich je přehledová studie Patwardhan et al. (2014), která jako kódující mitochondriální markery uvádí geny pro podjednotky cytochrom *c* oxidázy I a II (*COX1* a *COX2*) a cytochrom *b* (*CYTB*), které jsem v práci použila. Všechny tři markery byly skutečně často využívány ve fylogenetických studiích (Aguilar-Tomasini et al., 2021; Tobe et al., 2010; Sena et al., 2003; Ashley et al., 1996). Nukleární kódující markery však studie neobsahovala. Z toho důvodu jsem se musela spolehnout na fylogenetické studie.

Z často využívaných nukleárních markerů ve fylogenetických studiích savců jsem našla a vybrala celkem 11 a zahrnula jsem také již zmíněný kontrolní gen *TP53* V případě některých nukleárních genů se pro fylogenezi používá pouze jeden či více exonů, nebo jen jejich části, což je vyznačeno v Tab. 3, která shrnuje všechny analyzované geny.

Tab. 2: Seznam všech devíti analyzovaných taxonů a jejich zastoupení, včetně šedě vyznačeného řádu Carnivora, který nebyl analyzován jako celek, ale rozdělený na dva podřády.

Řád	***n***	**Podtaxon**	***n***
Artiodactyla	38	Cetacea	17
Carnivora	43	Caniformia	27
		Feliformia	16
Eulipotyphla	6	–	–
Chiroptera	22	–	–
Perissodactyla	5	–	–
Primates	34	–	–
Rodentia	43	–	–
	∑191		

3.3 Získání dat z NCBI

Úspěšnost automatického stahování referenčních sekvencí jaderných markerů byla poměrně vysoká. Z celkového množství 2292 (12 genů od 191 savců) sekvencí se jich skriptu nepodařilo stáhnout pouze 27. Ověřila jsem, že se sekvence v databázi opravdu nenacházejí, což se až na jednu sekvenci potvrdilo. Sekvence genu *TP53* druhu *Mus musculus* měla v databázi jiné označení genu v názvu sekvence, proto ji skript nenašel a stáhla jsem ji manuálně. Podmínka, že se název genu musí nacházet i v názvu sekvence byla ve skriptu nutná, bez ní se v několika případech stáhly jiné geny z toho důvodu, že měly jako jeden ze synonymních názvů genu uvedený i název mnou hledaného genu. Celkem mi tak chybělo pouze 26 nukleárních sekvencí (1,18 %) (viz Tab. 4). Pouze ve dvou případech se nepodařilo stáhnout transkripční variantu 1 a sekvence genu *TP53* druhů *Mustela erminea* a *Mustela putorius* byla dostupná pouze v transkripční variantě 3. Deset ze stažených sekvencí nebylo kompletních, ale částečných, což ale pro samotnou analýzu nepředstavovalo problém. Při alignmentu jsem tuto skutečnost zohlednila a chybějící části sekvencí byly doplněny mezerami. Naprostá většina stažených sekvencích byla predikovaná. Data byla stažena 5. 12. 2024.

U mitochondriálních genů se již úspěšnost automatického stahování snížila a nepodařilo se stáhnout 66 z celkového počtu 573 genů (3 geny od 191 savců). Z toho jsem 45 stáhla manuálně. Několik genomů skript nenašel, protože v názvu neměly přesnou frázi „mitochondrion, complete genome“, což bylo kritériem vyhledávání. U jiných genomů selhalo extrahování genů, pokud neměly definovanou sekvenci a pouze se odkazovaly na jinou sekvenci v databázi jejím přístupovým číslem a rozsahem. Většina (164) savců měla referenční mitochondriální genom, ze kterého jsem sekvence extrahovala, 10 savců mělo nereferenční mitochondriální genom a 17 savců nemělo dostupný žádný mitochondriální genom. Jejich sekvence jsem hledala v databázi samostatně a tyto sekvence bez výjimky

Tab. 3: Analyzované fylogenetické markery savců a šedě vyznačený kontrolní gen *TP53*. Používá-li se jako marker pouze konkrétní exon nebo exony, je u genu uvedeno, o které se jedná. Typy markerů „mt“ a „nc“ označují mitochondriální resp. nukleární geny.

Typ	Gen	Exon(y)	Protein	Fylogenetické studie
mt	*COX1*	–	Cytochrom c oxidáza podjednotka I	Tobe et al. (2010) Aguilar-Tomasini et al. (2021) Santodomingo et al. (2022)
mt	*COX2*	–	Cytochrom c oxidáza podjednotka II	Adkins a Honeycutt (1991) Ashley et al. (1996) Sena et al. (2003)
mt	*CYTB*	–	Cytochrom b	Martin et al. (2000) Tobe et al. (2010) Aguilar-Tomasini et al. (2021)
nc	*APOB*	26 (část)	Apolipoprotein B	Amrine-Madsen et al. (2003) Meredith et al. (2008) Koepfli et al. (2008)
nc	*BRCA1*	11	Protein náchylnosti k rakovině prsu typu 1	Adkins et al. (2001) Steppan et al. (2004a) Meredith et al. (2008)
nc	*CSN3*	2-4	Kappa-kasein	Chikuni et al. (1995) Gatesy et al. (1996) Cronin et al. (1996)
nc	*DMP1*	6 (část)	Acidický fosfoprotein dentinové matrix 1	Van Den Bussche et al. (2003) Reeder a Bradley (2004) Jansa et al. (2006)
nc	*GHR*	10 (část)	Receptor růstového hormonu	Adkins et al. (2001) Steppan et al. (2004a) Galewski et al. (2006)
nc	*LCAT*	2-6	Lecitin-cholesterol acyltransferáza	Robinson et al. (1997) Robinson-Rechavi et al. (2000) Michaux et al. (2001)
nc	*MYC*	2-3	Protein protoonkogenu MYC	Miyamoto et al. (2000) Steppan et al. (2004a) Steppan et al. (2004b)
nc	*RAG1*	–	Rekombinaci aktivující protein 1	Steppan et al. (2004a) Baker et al. (2004) Meredith et al. (2008)
nc	*RAG2*	–	Rekombinaci aktivující protein 2	Hoofer et al. (2003) Stadelmann et al. (2007) Gaubert et al. (2018)
nc	*RBP3*	1 (část)	Retinol vázající protein 3	Madsen et al. (2001) Jansa a Weksler (2004) Meredith et al. (2008)
nc	*VWF*	28	Von Willebrandův faktor	Porter et al. (1996) Michaux et al. (2001) Meredith et al. (2008)
nc	*TP53*	–	Nádorový protein p53	–

nebyly referenční. Jednalo se opět o CDS z nichž 13 bylo nekompletních. Savec *Piliocolobus tephrosceles* jako jediný neměl v databázi dostupnou žádnou mitochondriální sekvenci. Celkem mi tedy chybělo 21 (3,66 %) mitochondriálních sekvencí (viz Tab. 4). Data byla stažena 10. 12. 2024.

Při kontrole hlaviček FASTA souborů obsahujících názvy sekvencí případně mitochondriálních genomů jsem neodhalila žádné problémy. Přístupová čísla stažených sekvencí genů a mitochondriálních genomů, včetně rozsahu stažené sekvence, obsahuje Příloha 2.

Celkem 62 sekvencí nebylo dělitelných třemi. Pouze ve čtyřech případech se jednalo o sekvence nukleárních genů, z nichž všechny byly zároveň nekompletní, což mohlo vysvětlovat jejich délku. V naprosté většině se ale jednalo o sekvence genů mitochondriálních. U nich je však známým jevem, že v mnoha případech nemají kompletní stop kodon, který bývá doplněn poly(A) koncem až post-transkripční modifikací mRNA zvanou polyadenylace (Taanman, 1999).

Tab. 4: Přehled stažených sekvencí a chybějících dat.

Typ genů	Počet zájmových sekvencí	Staženo automaticky (skript)	Staženo manuálně	Počet chybějících sekvencí	Podíl chybějících sekvencí (%)
Nukleární	2292	2264	1	27	1,18
Mitochondriální	573	507	45	21	3,66

3.4 Alignment

Při alignmentu sekvencí nedělitelných třemi dokázal program MACSE alignSequences úspěšně identifikovat správný čtecí rámec a neúplné kodony doplnit a označit znakem posunu čtecího rámce.

Alignmenty byly po prvotním alignmentu i ve všech následujících fázích analýzy dělitelné třemi, což indikovalo, že nedošlo k žádnému nechtěnému narušení čtecích rámců.

V alignmentech se po jejich rafinaci v refineAlignment už nevyskytovaly žádné posuny čtecího rámce. Co se týče mitochondriálních genů, rafinace nevedla k žádným změnám a alignmenty zůstaly stejné. V případě nukleárních genů však rafinace vedla k lepšímu alignmentu delecí v okolí sloupců odstraněných trimAlem a dalším drobným změnám.

Většina alignmentů neobsahovala žádné stop kodony uvnitř sekvence. Výjimkou byly alignmenty genů *APOB* a *RBP3*, které obsahovaly 11 a 23 stop kodonů. Sekvence, kde se stop kodony vyskytovaly, jsem vizuálně zkontrolovala a vyloučila jsem možnost, že by se mohlo jednat o chybu v alignmentu.

Délky alignmentů v párových bázích (bp; base pair) v jednotlivých krocích alignmentu a filtrace jsou shrnuty v Tab. 5.

Tab. 5: Délky alignmentů (bp) v jednotlivých krocích alignmentu. Předposlední sloupec představuje finální délku alignmentů celých genů a poslední sloupec část těchto alignmentů, představující fylogenetické markery.

Gen	Prvotní alignment	trimAl 1	Rafinace alignmentu	trimAl 2	Fylogenetický marker
APOB	15 117	13 722	13 722	13 719	1 338
BRCA1	7 131	5 592	5 625	5 589	3 423
COX1	1 578	1 545	1 545	1 542	1 542
COX2	690	684	684	681	681
CSN3	1 431	552	558	549	549
CYTB	1 146	1 140	1 140	1 137	1 137
DMP1	2 040	1 566	1 587	1 563	1 194
GHR	2 631	1 917	1 917	1 914	969
LCAT	1 590	1 323	1 323	1 320	1 164
MYC	1 575	1 320	1 320	1 317	1 317
RAG1	3 297	3 132	3 132	3 129	3 129
RAG2	1 599	1 584	1 584	1 581	1 581
RBP3	4 281	3 729	3 729	3 729	1 275
TP53	1 731	1 164	1 185	1 161	1 161
VWF	9 102	8 442	8 442	8 439	1 368

3.4.1 Extrakce markerů z alignmentů celých genů

Tato kapitola popisuje, jaké části genů se běžně ve fylogenetických studiích používají jako markery a jaké části genů jsem tedy použila pro výpočet selekčního signálu. Mitochondriální geny se používaly vždy v celé jejich délce. V případě genu *TP53* jsem také použila celý alignment kódujících sekvencí vzhledem k tomu, že se nejedná a fylogenetický marker, ale pouze o kontrolní gen. Délky alignmentů po extrakci markerů jsou uvedeny v Tab. 5.

Po extrakci markerů jsem z alignmentů odstranila celkem 33 sekvencí, patřící v naprosté většině markerům *BRCA1* (21 sekvencí) a *COX1* (8 sekvencí). Sekvence měly ve výsledném alignmentu více než 50 % chybějících dat. Tyto sekvence jsou v Příloze 2 vyznačeny šedě. Odstraněné sekvence markeru *BRCA1* patřily zástupcům různých savčích řádů, ovšem 4 z nich patřily řádu Eulipotyphla, pro který mi u tohoto markeru zůstaly pouze sekvence dvou zástupců.

APOB

Část exonu 26 genu *APOB* bývá ve studiích uvedených v Tab. 3 amplifikována s využitím primerů uvedených ve studii Amrine-Madsen et al. (2003). Délka výsledného fragmentu byla 1338 bp (NM_000384.3:9421-10758).

BRCA1

Fylogenetickým markerem je část exonu 11, který je celý kódující. Použité primery se lišily v závislosti na fylogenetické studii. Ve všech případech ale amplifikovaly okolo 3000 bp ze začátku exonu 11 (Adkins et al., 2001; Voss a Jansa, 2009; Teeling et al., 2000). Ta tvoří většinu celkové délky tohoto exonu, která je u *Homo sapiens* 3426 bp. Proto jsem použila celý jedenáctý exon (NM_007294.4:784-4209).

CSN3

Ve studii Chikuni et al. (1995) využívají exony 2, 3 a 4, které pokrývají celou délku kódující

sekvence genu. Studie Gatesy et al. (1996); Cronin et al. (1996) pak využívají pouze exon 4, který ale stále pokrývá většinu (cca 80 %) CDS. Využila jsem tedy celou kódující sekvenci.

DMP1
Všechny studie uvedené v Tab. 3 využívají k amplifikaci části exonu 6 primery představené ve studii Toyosawa et al. (1999). Výsledná část šestého exonu má délku 1171 bp (NM_004407.4:468-1638).

GHR
Všechny studie uvedené v Tab. 3 použily pro amplifikaci části exonu 10 odlišné primery. Studie Adkins et al. (2001); Galewski et al. (2006) v obou případech amplifikují více než 900 bp ze začátku exonu 10, která pokrývá celou kódující část desátého exonu. Pro extrakci potřebné části alignmentu jsem tedy mohla použít celý exon 10 (NM_000163.5:1474-4899).

LCAT
Sekvenování markeru obvykle probíhá pro dva fragmenty genu, jak bylo představeno ve studii Robinson et al. (1997). První fragment pokrývá exony 2-5 až na malé okrajové části exonů 2 a 5. Druhý fragment pokrývá většinu exonu 6 až na jeho konec, který ale není kódující. V této práci jsem použila celé exony 2-6. (NM_000229.2:174-1496).

MYC
Ve fylogenetických studiích uvedených v Tab. 3 se využívají exony 2 a 3, které opět pokrývají naprostou většinu kódující části genu a použila jsem tedy celou staženou CDS.

RAG1* a *RAG2
Geny *RAG1* a *RAG2* oba nemají introny. K fylogenetické inferenci se využívají části jejich kódujících sekvencí (Van Den Bussche et al., 2003). Pro *RAG1* byla použita část, pokrývající celou CDS ve studii Steppan et al. (2004a). Pro *RAG2* byla použita 1,4 kb část, která tvoří většinu celkové délky CDS (cca 1,6 kb) (Hoofer et al., 2003). V obou případech jsem tedy použila celé kódující sekvence genů.

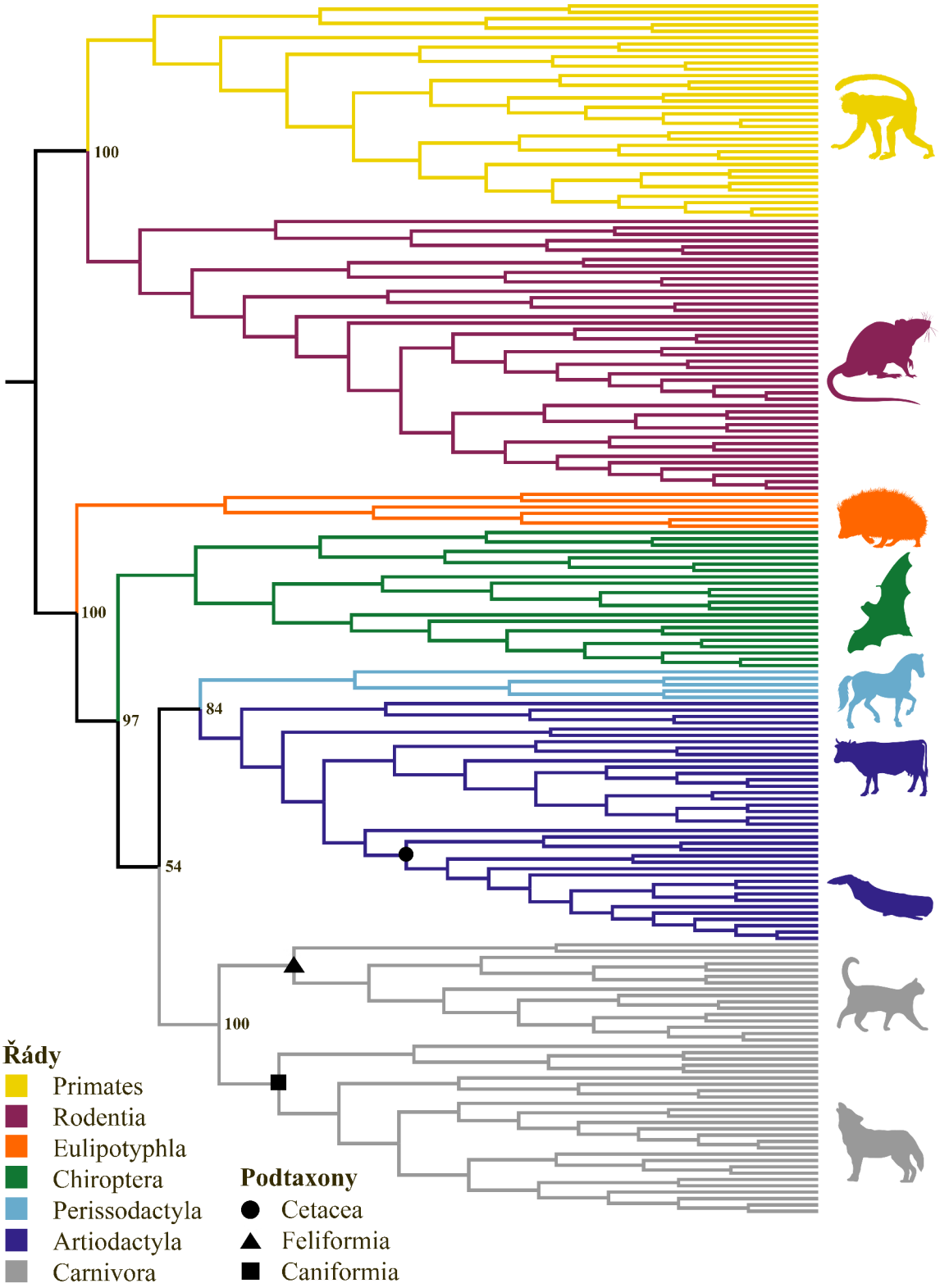
RBP3
Před sekvenováním části exonu 1 se využívají primery představené ve studii Stanhope et al. (1992). Délka výsledného fragmentu byla 1282 bp (NM_002900.3:245-1526).

VWF
Části exonu 28 sekvenované ve studiích Porter et al. (1996); Michaux et al. (2001) pokrývají téměř celou jeho délku. Použila jsem tedy celý exon 28 (NM_000552.5:3925-5303).

3.5 Fylogenetická rekonstrukce

Výsledkem fylogenetické inference v programu IQ-TREE byl nezakořeněný strom. Následně jsem upravila uspořádání taxonů tak, aby se u základny stromu (kde se běžně zobrazuje kořen) nacházela monofyletická skupina hlodavců a primátů. Tyto skupiny se podle fylogenetických studií savců oddělily jako první (Foley et al., 2023; Springer et al., 2004). Strom je v této podobě zobrazen na Obr. 1. Strom má zobrazený také kořen pro lepší vizualizaci – strom je ale nezakořeněný a všechny druhy jsou zobrazeny ve stejné vzdálenosti od kořene stromu.

Podpory hlubokých vztahů byly nedostatečné mezi řády Perissodactyla, Artiodactyla a Carnivora, kde byly hodnoty bootstrapové podpory pouze 84 a 54 (viz Obr. 1). Většina ostatních uzlů vykazovala vysokou úroveň podpory (>95), s výjimkou několika uzlů v těsné blízkosti listů stromu.



Obr. 1: Topologie výsledného fylogenetického stromu s barevně odlišenými řády savců. Geometrickými symboly jsou vyznačeny uzly společného předka dalších tří analyzovaných taxonů. Čísla představují podpory některých hlubokých uzlů stromu.

3.6 Výpočet selekčního signálu

Kompletní výsledky pozicových modelů, detekujících signál pozitivní selekce v rámci markeru (pro celý fylogenetický strom) obsahuje Tab. 6 (modely M1a a M2a) a Tab. 7 (modely M7 a M8).

Výsledky modelu cmC, který detekuje odlišné působení selekce u specifikovaných taxonů, obsahuje Tab. 8. V tabulce jsou zahrnuty pouze signifikantní výsledky s p-hodnotou $\leq 0,05$ a výsledky, kde alespoň jeden z parametrů ω_2 a ω_3 je větší než jedna, nebo je jejich rozdíl $|\Delta\omega| = |\omega_3 - \omega_2| \geq 0,1$. Kompletní výsledky modelu cmC obsahuje Příloha 3. Výsledky tohoto modelu pro kontrolní gen *TP53* nejsou ve výsledcích uvedeny, protože jsem model pro tento gen nepočítala.

Výsledky modelu bsA, který detekuje signál pozitivní selekce u jednotlivých taxonů, obsahuje Tab. 9. Tabulka zahrnuje pouze signifikantní výsledky s p-hodnotou $\leq 0,05$. Kompletní výsledky modelu bsA obsahuje Příloha 4. Výsledky tohoto modelu pro kontrolní gen *TP53* nejsou uvedeny, protože jsem model pro tento gen nepočítala. Výpočet modelu pro řád Eulipotyphla u markeru *BRCA1* nedokonvergoval k výsledku ani po více než 420 hodinách a výpočet trval výrazně déle ve srovnání s většinou ostatních taxonů, proto tento výsledek není zahrnut. Problém konvergence tohoto konkrétního modelu může být způsoben tím, že pro tento marker důsledkem odstranění sekvencí s velkým podílem chybějících dat zbyly pouze sekvence dvou zástupců řádu Eulipotyphla.

S pomocí pozicových modelů M1a a M2a i M7 a M8 se mi podařilo detekovat pozitivní selekci u kontrolního genu *TP53*, čímž jsem si ověřila, že je možné použitou metodou selekci úspěšně detekovat.

Ani u jednoho ze tří mitochondriálních markerů jsem nedetekovala pozitivní selekci, a to jak na úrovni celého stromu tak u jednotlivých taxonů. Působení selekce na tyto geny se na základě výsledků modelu cmC mezi jednotlivými taxony významně nelišilo. Více než 99 % všech jejich pozic je ale pod vlivem silné negativní selekce s poměrem $\omega < 0,03$.

LRT pozicových modelů M1a a M2a odhalil signál pozitivní selekce u 7 z 11 nukleárních fylogenetických markerů. Pouze u markerů *MYC*, *RAG1*, *RBP3* a *VWF* pozitivní selekce detekována nebyla. Tyto závěry byly konzistentní s výsledky modelů M7 a M8, kde sice vyšel statisticky významný výsledek LRT u všech 11 nukleárních markerů, ovšem u těch, u kterých modely M1a a M2a nedetekovaly pozitivní selekci byl odhad parametru ω_s roven jedné, což indikuje neutrální evoluci, nikoliv pozitivní selekci. Pro markery *MYC*, *RBP3* a *VWF* jsem však s pomocí modelu bsA detekovala pozitivní selekci alespoň u jednoho z taxonů a jediným markerem, u kterého se neprokázala pozitivní selekce na žádné úrovni tak byl pouze *RAG1*.

Dle výsledků modelu M2a proporce pozic připadající pozicím pod pozitivní selekcí byla většinou malá (okolo 1 %) s výjimkou markeru *BRCA1*, kde tvořily pozitivní pozice více než 6 % a *CSN3*, kde tvořily 11,5 %. Model M8 detekoval oproti modelu M2a o něco vyšší podíly pozitivních pozic ale s menším poměrem ω.

Dle výsledků M2a modelu je nukleární marker *MYC* pod největším vlivem silné negativní selekce ($p_0 = 95,5\%$ a $\omega_0 = 0,036$) a nejmenším podílem neutrálních pozic (4,5 %). Markery *LCAT*, *RAG1*, *RAG2*, *RBP3* a *VWF* mají více než 80 % pozic pod silnou negativní selekcí ($\omega < 0,13$) a větší podíl neutrálních pozic (9-17 %). Markery *APOB*, *DMP1* a *GHR*, u kterých je negativní selekce slabší, a vykazují také větší proporce neutrálních

pozic (30-36 %) na úkor těch negativních. Největší posun v rozdělení poměrů ω pozic pozorujeme u markerů *BRCA1* a *CSN3*, které mají největší podíl neutrálních pozic (44 % resp. 55 %).

Marker *CSN3* má nejvyšší proporci pozic pod pozitivní selekcí 11,5 % podle modelu M2a a 13,5 % podle modelu M8 s poměry ω okolo 2. Pozitivní selekci u tohoto markeru jsem detekovala také u většiny taxonů až na Perissodactyla a Feliformia. Také dle výsledku cmC modelu mají Feliformia 11,4 % pozic pod mírnou negativní selekcí, zatímco u ostatních savců jsou pozice pod pozitivní selekcí. Ovšem u řádu Perrisodactyla model cmC detekoval ještě silnější působení pozitivní selekce oproti zbytku stromu, i přes to, že bsA žádnou pozitivní selekci neprokázal, závěr pro tento řád proto není zcela jednoznačný. Marker *CSN3* má na druhou stranu ze všech analyzovaných markerů největší podíl neutrálních pozic (přes 55 %) a na zbývající třetinu pozic působí nejmírnější detekovaná negativní selekce ($\omega = 0,42$). Druhým genem s největší proporcí pozitivních pozic byl marker *BRCA1* s proporcí 6,2 % u modelu M2a a 12,2 % pro model M8. Pozitivní selekci jsem pro něj detekovala také u 6 z 9 analyzovaných taxonů. Pro ostatní markery jsem na úrovni taxonů detekovala pozitivní selekci vždy pouze u jednoho či dvou taxonů.

Tab. 6: Výsledky pozicových modelů M1a a M2a. Počet stupňů volnosti LRT je 2.

Gen	Model	LL	ω	Odhady parametrů modelu			LRT	p-hodnota
COX1	M1a	-78457,64	0,01	$p_0 = 0,994$	$p_1 = 0,006$			
				$\omega_0 = 0,006$	$\omega_1 = 1$			
	M2a	-78457,64	0,01	$p_0 = 0,994$	$p_1 = 0,001$	$p_2 = 0,005$	0,00	1,000
				$\omega_0 = 0,006$	$\omega_1 = 1$	$\omega_2 = 1,000$		
COX2	M1a	-36475,69	0,02	$p_0 = 0,991$	$p_1 = 0,009$			
				$\omega_0 = 0,014$	$\omega_1 = 1$			
	M2a	-36475,69	0,02	$p_0 = 0,991$	$p_1 = 0,002$	$p_2 = 0,006$	0,00	1,000
				$\omega_0 = 0,014$	$\omega_1 = 1$	$\omega_2 = 1,000$		
CYTB	M1a	-72651,56	0,03	$p_0 = 0,995$	$p_1 = 0,005$			
				$\omega_0 = 0,027$	$\omega_1 = 1$			
	M2a	-72651,56	0,03	$p_0 = 0,995$	$p_1 = 0,001$	$p_2 = 0,004$	0,00	1,000
				$\omega_0 = 0,027$	$\omega_1 = 1$	$\omega_2 = 1,000$		
APOB	M1a	-33989,26	0,49	$p_0 = 0,673$	$p_1 = 0,327$			
				$\omega_0 = 0,248$	$\omega_1 = 1$			
	M2a	-33969,54	0,52	$p_0 = 0,664$	$p_1 = 0,323$	$p_2 = 0,013$	39,44	0,000**
				$\omega_0 = 0,250$	$\omega_1 = 1$	$\omega_2 = 2,375$		
BRCA1	M1a	-93512,65	0,65	$p_0 = 0,532$	$p_1 = 0,468$			
				$\omega_0 = 0,343$	$\omega_1 = 1$			
	M2a	-93432,29	0,73	$p_0 = 0,501$	$p_1 = 0,437$	$p_2 = 0,063$	160,73	0,000**
				$\omega_0 = 0,354$	$\omega_1 = 1$	$\omega_2 = 1,926$		
CSN3	M1a	-18979,32	0,77	$p_0 = 0,392$	$p_1 = 0,608$			
				$\omega_0 = 0,407$	$\omega_1 = 1$			
	M2a	-18935,94	0,93	$p_0 = 0,333$	$p_1 = 0,552$	$p_2 = 0,115$	86,76	0,000**
				$\omega_0 = 0,420$	$\omega_1 = 1$	$\omega_2 = 2,054$		
DMP1	M1a	-40974,23	0,48	$p_0 = 0,640$	$p_1 = 0,360$			
				$\omega_0 = 0,194$	$\omega_1 = 1$			
	M2a	-40970,50	0,50	$p_0 = 0,638$	$p_1 = 0,351$	$p_2 = 0,011$	7,45	0,024
				$\omega_0 = 0,196$	$\omega_1 = 1$	$\omega_2 = 1,912$		
GHR	M1a	-22044,39	0,39	$p_0 = 0,702$	$p_1 = 0,298$			
				$\omega_0 = 0,126$	$\omega_1 = 1$			
	M2a	-22028,07	0,41	$p_0 = 0,694$	$p_1 = 0,295$	$p_2 = 0,011$	32,64	0,000**
				$\omega_0 = 0,126$	$\omega_1 = 1$	$\omega_2 = 2,396$		
LCAT	M1a	-21680,62	0,21	$p_0 = 0,833$	$p_1 = 0,167$			
				$\omega_0 = 0,052$	$\omega_1 = 1$			
	M2a	-21673,01	0,22	$p_0 = 0,830$	$p_1 = 0,162$	$p_2 = 0,007$	15,22	0,000**
				$\omega_0 = 0,052$	$\omega_1 = 1$	$\omega_2 = 2,355$		
MYC	M1a	-20950,22	0,08	$p_0 = 0,955$	$p_1 = 0,045$			
				$\omega_0 = 0,036$	$\omega_1 = 1$			
	M2a	-20950,22	0,08	$p_0 = 0,955$	$p_1 = 0,045$	$p_2 = 0,000$	-0,00	1,000
				$\omega_0 = 0,036$	$\omega_1 = 1$	$\omega_2 = 23,711$		
RAG1	M1a	-65297,48	0,21	$p_0 = 0,826$	$p_1 = 0,174$			
				$\omega_0 = 0,046$	$\omega_1 = 1$			
	M2a	-65297,48	0,21	$p_0 = 0,826$	$p_1 = 0,174$	$p_2 = 0,000$	0,00	1,000
				$\omega_0 = 0,046$	$\omega_1 = 1$	$\omega_2 = 13,821$		
RAG2	M1a	-27057,37	0,21	$p_0 = 0,859$	$p_1 = 0,141$			
				$\omega_0 = 0,077$	$\omega_1 = 1$			
	M2a	-27053,22	0,22	$p_0 = 0,858$	$p_1 = 0,136$	$p_2 = 0,006$	8,29	0,016
				$\omega_0 = 0,077$	$\omega_1 = 1$	$\omega_2 = 2,097$		
RBP3	M1a	-35671,71	0,23	$p_0 = 0,886$	$p_1 = 0,114$			
				$\omega_0 = 0,127$	$\omega_1 = 1$			
	M2a	-35671,40	0,23	$p_0 = 0,886$	$p_1 = 0,112$	$p_2 = 0,003$	0,62	0,734
				$\omega_0 = 0,127$	$\omega_1 = 1$	$\omega_2 = 1,863$		
VWF	M1a	-37834,41	0,21	$p_0 = 0,876$	$p_1 = 0,124$			
				$\omega_0 = 0,101$	$\omega_1 = 1$			
	M2a	-37834,40	0,21	$p_0 = 0,876$	$p_1 = 0,091$	$p_2 = 0,033$	0,00	0,998
				$\omega_0 = 0,101$	$\omega_1 = 1$	$\omega_2 = 1,000$		
TP53	M1a	-30383,08	0,35	$p_0 = 0,720$	$p_1 = 0,280$			
				$\omega_0 = 0,100$	$\omega_1 = 1$			
	M2a	-30380,08	0,36	$p_0 = 0,718$	$p_1 = 0,272$	$p_2 = 0,010$	6,00	0,050
				$\omega_0 = 0,100$	$\omega_1 = 1$	$\omega_2 = 1,842$		

Pozn.: $*p \leq 0,01$; $**p \leq 0,001$; LL – logaritmus věrohodnosti; ω – průměrný poměr ω pro všechny pozice.

Tab. 7: Výsledky pozicových modelů M7 a M8. Počet stupňů volnosti LRT je 2.

Gen	Model	LL	ω	Odhady parametrů modelu			LRT	p-hodnota
COX1	M7	-77169,81	0,02	$p = 0,156$	$q = 7,669$			
	M8	-77169,81	0,02	$p = 0,156$	$q = 7,669$	$p_0 = 1,000$	-0,01	1,000
				$p_1 = 0,000$	$\omega_s = 1,000$			
COX2	M7	-35653,94	0,03	$p = 0,241$	$q = 6,827$			
	M8	-35653,95	0,03	$p = 0,241$	$q = 6,824$	$p_0 = 1,000$	-0,00	1,000
				$p_1 = 0,000$	$\omega_s = 1,011$			
CYTB	M7	-69853,44	0,03	$p = 0,284$	$q = 8,308$			
	M8	-69853,45	0,03	$p = 0,284$	$q = 8,308$	$p_0 = 1,000$	-0,01	1,000
				$p_1 = 0,000$	$\omega_s = 1,000$			
APOB	M7	-33892,78	0,42	$p = 0,809$	$q = 1,099$			
	M8	-33864,80	0,43	$p = 0,940$	$q = 1,443$	$p_0 = 0,968$	55,96	0,000**
				$p_1 = 0,032$	$\omega_s = 1,675$			
BRCA1	M7	-93432,00	0,60	$p = 1,125$	$q = 0,755$			
	M8	-93312,42	0,67	$p = 1,440$	$q = 1,182$	$p_0 = 0,878$	239,15	0,000**
				$p_1 = 0,122$	$\omega_s = 1,562$			
CSN3	M7	-18977,70	0,73	$p = 1,097$	$q = 0,414$			
	M8	-18928,93	0,88	$p = 1,300$	$q = 0,510$	$p_0 = 0,865$	97,53	0,000**
				$p_1 = 0,135$	$\omega_s = 1,902$			
DMP1	M7	-40635,17	0,36	$p = 0,745$	$q = 1,320$			
	M8	-40625,82	0,36	$p = 0,879$	$q = 1,996$	$p_0 = 0,927$	18,70	0,000**
				$p_1 = 0,073$	$\omega_s = 1,085$			
GHR	M7	-21920,91	0,32	$p = 0,409$	$q = 0,877$			
	M8	-21901,44	0,32	$p = 0,450$	$q = 1,086$	$p_0 = 0,985$	38,93	0,000**
				$p_1 = 0,015$	$\omega_s = 1,935$			
LCAT	M7	-21440,75	0,17	$p = 0,169$	$q = 0,821$			
	M8	-21422,42	0,16	$p = 0,222$	$q = 1,554$	$p_0 = 0,967$	36,65	0,000**
				$p_1 = 0,033$	$\omega_s = 1,289$			
MYC	M7	-20633,64	0,06	$p = 0,231$	$q = 3,182$			
	M8	-20614,56	0,06	$p = 0,247$	$q = 4,134$	$p_0 = 0,994$	38,16	0,000**
				$p_1 = 0,006$	$\omega_s = 1,000$			
RAG1	M7	-64345,96	0,14	$p = 0,196$	$q = 1,157$			
	M8	-64324,20	0,11	$p = 0,226$	$q = 2,220$	$p_0 = 0,970$	43,53	0,000**
				$p_1 = 0,030$	$\omega_s = 1,000$			
RAG2	M7	-26841,73	0,17	$p = 0,255$	$q = 1,178$			
	M8	-26821,25	0,16	$p = 0,287$	$q = 1,604$	$p_0 = 0,988$	40,95	0,000**
				$p_1 = 0,012$	$\omega_s = 1,589$			
RBP3	M7	-35402,67	0,20	$p = 0,741$	$q = 2,929$			
	M8	-35375,95	0,19	$p = 0,951$	$q = 4,958$	$p_0 = 0,963$	53,45	0,000**
				$p_1 = 0,037$	$\omega_s = 1,000$			
VWF	M7	-37199,36	0,16	$p = 0,436$	$q = 2,177$			
	M8	-37191,77	0,16	$p = 0,468$	$q = 2,662$	$p_0 = 0,986$	15,18	0,001**
				$p_1 = 0,014$	$\omega_s = 1,000$			
TP53	M7	-30044,24	0,27	$p = 0,314$	$q = 0,862$			
	M8	-30035,10	0,26	$p = 0,353$	$q = 1,219$	$p_0 = 0,968$	18,28	0,000**
				$p_1 = 0,032$	$\omega_s = 1,315$			

Pozn.: *$p \leq 0,01$; **$p \leq 0,001$; LL – logaritmus věrohodnosti; ω – průměrný poměr ω pro všechny pozice.

Tab. 8: Signifikantní výsledky modelu cmC s rozdílem poměrů $|\Delta\omega| \geq 0,1$ pro jednotlivé taxony. Nulové modely nejsou uvedeny, pouze výsledky LRT testu s jedním stupněm volnosti.

Gen	Taxon	p_2	ω_2	ω_3	$\Delta\omega$	LRT	*p*-hodnota
APOB	Chiroptera	0,463	0,428	0,586	0,158	10,02	0,002*
	Perissodactyla	0,458	0,450	0,809	0,359	6,15	0,013
	Rodentia	0,466	0,525	0,355	−0,170	23,88	0,000**
BRCA1	Caniformia	0,473	0,475	0,364	−0,111	8,82	0,003*
	Cetacea	0,474	0,463	0,678	0,216	7,47	0,006*
	Rodentia	0,471	0,414	0,522	0,109	23,05	0,000**
CSN3	Artiodactyla	0,139	1,837	2,997	1,160	6,19	0,013
	Caniformia	0,104	2,013	3,668	1,655	6,90	0,009*
	Feliformia	0,114	2,099	0,784	−1,315	5,38	0,020
	Perissodactyla	0,113	2,024	4,059	2,036	4,02	0,045
	Primates	0,102	1,952	3,051	1,099	6,42	0,011
	Rodentia	0,099	2,676	1,355	−1,321	25,14	0,000**
DMP1	Cetacea	0,465	0,356	0,535	0,179	14,59	0,000**
	Rodentia	0,457	0,418	0,302	−0,115	27,76	0,000**
GHR	Artiodactyla	0,356	0,374	0,258	−0,115	4,70	0,030
	Chiroptera	0,350	0,388	0,233	−0,155	16,06	0,000**
	Primates	0,356	0,339	0,472	0,133	4,89	0,027
	Rodentia	0,359	0,309	0,451	0,142	16,40	0,000**
LCAT	Caniformia	0,290	0,236	0,120	−0,116	6,75	0,009*
	Chiroptera	0,286	0,200	0,471	0,271	48,76	0,000**
	Primates	0,292	0,250	0,143	−0,107	13,31	0,000**
MYC	Cetacea	0,226	0,165	0,342	0,178	6,13	0,013
RAG1	Cetacea	0,261	0,273	0,591	0,318	33,24	0,000**
RAG2	Feliformia	0,303	0,270	0,143	−0,127	5,44	0,020
RBP3	Eulipotyphla	0,419	0,246	0,417	0,171	15,09	0,000**
	Chiroptera	0,610	0,137	0,408	0,271	161,74	0,000**
VWF	Feliformia	0,390	0,240	0,357	0,118	7,45	0,006*

Pozn.: *$p \leq 0{,}01$; **$p \leq 0{,}001$; $\Delta\omega = \omega_3 - \omega_2$.

Tab. 9: Signifikantní výsledky modelu bsA pro jednotlivé taxony. Nulové modely nejsou uvedeny, pouze výsledky LRT testu s jedním stupněm volnosti.

Gen	Taxon	p_{2a}	p_{2b}	ω_2	LRT	*p*-hodnota
APOB	Caniformia	0,014	0,006	4,838	10,13	0,001*
	Chiroptera	0,077	0,034	1,588	4,91	0,027
BRCA1	Artiodactyla	0,021	0,018	3,034	18,70	0,000**
	Caniformia	0,016	0,014	4,508	39,58	0,000**
	Cetacea	0,028	0,025	5,818	30,16	0,000**
	Feliformia	0,018	0,016	4,755	22,25	0,000**
	Chiroptera	0,014	0,012	2,264	3,91	0,048
	Primates	0,010	0,009	3,233	13,48	0,000**
CSN3	Artiodactyla	0,049	0,077	3,073	20,47	0,000**
	Caniformia	0,025	0,038	4,079	6,11	0,013
	Cetacea	0,038	0,059	5,387	9,52	0,002*
	Eulipotyphla	0,053	0,081	2,571	4,65	0,031
	Chiroptera	0,056	0,085	2,948	24,25	0,000**
	Primates	0,014	0,022	5,270	26,08	0,000**
	Rodentia	0,165	0,155	1,405	9,81	0,002*
DMP1	Cetacea	0,083	0,045	1,669	5,75	0,017
	Chiroptera	0,038	0,020	1,884	4,73	0,030
LCAT	Feliformia	0,003	0,001	7,096	3,96	0,047
MYC	Cetacea	0,013	0,001	3,781	6,49	0,011
RBP3	Cetacea	0,007	0,001	3,083	4,64	0,031
VWF	Feliformia	0,027	0,004	2,239	5,20	0,023

Pozn.: *$p \leq 0{,}01$; **$p \leq 0{,}001$.

4 Diskuze

4.1 Fylogenetická rekonstrukce

V této diplomové práci jsem využila fylogenetický strom vytvořený na základě celých CDS, přestože by se dalo argumentovat, že pro účely rekonstrukce stromu by bylo vhodnější použít pouze extrahované části alignmentů odpovídající fylogenetickým markerům. Tento strom se mi ale kvůli technickým problémům podařilo zrekonstruovat později, ve chvíli kdy už jsem měla část výsledků na základě původního stromu. Proto jsem tyto dva stromy mezi sebou porovnala s pomocí funkce „compare“ z balíku ETE. Stromy sdílely 98 % větví a jejich normalizovaná Robinsonova-Fouldsova vzdálenost byla 0,04. Stromy se tak mezi sebou lišily minimálně a výsledky testů pozitivní selekce jsou obecně robustní k malým rozdílům v topologii stromu (Pie, 2006). Z tohoto důvodu jsem se rozhodla pokračovat v analýze s využitím původního stromu na základě celých CDS.

Podpory hlubokých vztahů byly nedostatečné mezi řády Perissodactyla, Artiodactyla a Carnivora, kde byly hodnoty bootstrapové podpory pouze 84 a 54 (viz Obr. 1). Ani odborná literatura se však v těchto vztazích často neshoduje. Například podle studie Murphy et al. (2001) tvoří sesterské skupiny Carnivora a Perissodactyla, zatímco výsledky studie Liu et al. (2001) podporují bližší příbuznost mezi Perissodactyla a Artiodactyla. Na základě nedávné studie (Foley et al., 2023) vycházející z celogenomových sekvencí savců se ale ukazuje, že Perissodactyla jsou pravděpodobně příbuznější řádu Carnivora. Ve studii zkoumající vliv porušení předpokladů modelů na výsledek fylogenetické inference získali dva stromy, první na základě sekvencí splňující předpoklady, kdy získali stejné vztahy, které jsem získala v této práci s použitím alignmentů celých kódujících sekvencí. Na základě sekvencí porušujících předpoklady získali sesterské skupiny Carnivora a Perissodactyla, shodující se s výsledkem původní studie, což poukazuje na to, že vliv porušení předpokladů na výsledek fylogeneze je značný (Naser-Khdour et al., 2019).

Další otázkou je, zda je pro detekci selekce v codeml vhodnější druhový strom, reprezentující vztahy analyzovaných taxonů či genový strom, odrážející vývoj daného genu (strom tvořený na základě alignmentu jediného genu), který se může od etablovaného druhového stromu lišit. Odpověď není zcela jednoznačná. Platí, že strom by měl reprezentovat evoluční vztahy analyzovaných sekvencí. Je třeba zvážit, který strom bude s větší pravděpodobností správný. V případě příliš krátkého genu nemusí jeho alignment obsahovat dostatek fylogeneticky informativních pozic a je tedy vhodnější zvolit druhový strom (Álvarez-Carretero et al., 2023). Pro potřeby této práce jsem zvolila jednotný strom pro všechny analyzované geny, díky kterému jsem mohla použít jednotné označování taxonů pro modely bsA a cmC.

4.2 Detekce selekce

Model M8 detekoval oproti modelu M2a o něco vyšší podíly pozitivních pozic ale s menším poměrem ω. To naznačuje, že model dokáže detekovat i pozice pod působením

slabší pozitivní selekce, které následně sníží celkový průměrný poměr ω pozitivních pozic. Modely M7 a M8 měly konzistentně větší logaritmus věrohodnosti oproti jednodušším modelům M1a a M2a, což poukazuje na to, že s pomocí spojitého beta rozdělení lze poměry ω jednotlivých pozic modelovat lépe než pomocí tří diskrétních tříd. Zajímavé také je, že se pozitivní selekce na úrovni taxonu prokázala u tří ze čtyř markerů, u kterých ale na úrovni celého stromu detekována nebyla, a to u markerů *MYC*, *RBP3* a *VWF*. Naopak u markerů *GHR* a *RAG2* jsem detekovala signál pozitivní selekce na úrovni celého stromu, ale nikoliv u konkrétního taxonu. Z toho se dá usuzovat, že pozicové modely mohou přehlédnout slabý signál pozitivní selekce, pokud k ní dochází pouze na malé části větví stromu. Celkově je ale model větví-pozicí A při detekci pozitivní selekce více konzervativní v porovnání s pozicovými modely.

Mitochondriální markery vykazovaly největší míru negativní selekce, která je u mitochondriálních genů častá, vzhledem k tomu, že kódují proteiny zapojené v oxidativní fosforylaci s přímým vlivem na metabolismus (Bondareva et al., 2021). Jediným nukleárním markerem který se mírou negativní selekce přiblížil k mitochondriálním markerům byl *MYC*, který se podílí na proliferaci a apoptóze buněk (Bretones et al., 2015; Prendergast, 1999). Silná negativní selekce poukazuje na důležitost jeho funkce.

Výsledky studie Massey et al. (2008) ukazují, že pozice pod silnou negativní selekcí ($\omega < 0,2$) ztrácejí skutečný signál původu, a to dokonce ve větší míře než pozice pod silnou pozitivní selekcí ($\omega > 2$). Skutečný signál dle autorů nesou více neutrální pozice s $0,2 < \omega < 2$. Z tohoto důvodu autoři doporučují před fylogenetickou rekonstrukcí odstranění pozicí pod silnou pozitivní i negativní selekcí. Pokud bychom se rozhodli doporučením řídit, mitochondriální markery analyzované v této práci ani nukleární marker *MYC* by nebylo možné použít, protože podmínku by nesplňovaly téměř žádné pozice. Také u většiny ostatních nukleárních markerů bychom tímto ztratili většinu pozic. V tomto případě by tedy bylo vhodnější využít navrhovaná alternativní řešení autorů, a použít modely uvolňující předpoklady kladené na sekvence markerů (především nezávislost pozic) (Stern a Pupko, 2006), nebo modely zohledňující terciální strukturu proteinů (Parisi a Echave, 2001). Autoři hodnotili signál původu s pomocí nefylogenetických metod, aby co možná nejlépe zabránili zkreslení výsledků. Bylo by proto zajímavé podívat se také na přímý vliv negativní selekce na podobu a topologii fylogenetických stromů.

Silná stránka markerů pod negativní selekcí může ale spočívat například v tom, že selekce snižuje jejich substituční rychlosti a geny jsou proto konzervované. To může být nutné pro rekonstrukci hlubokých fylogenetických vztahů mezi vzdálenými taxony.

I když vezmeme v úvahu působení silné negativní selekce na mitochondriální markery a s tím související konzervaci sekvencí, nejsou tyto geny vhodné pro rekonstrukci hlubokých fylogenetických vztahů a používají se většinou pouze na druhové nebo populační úrovni (Patwardhan et al., 2014). Především rozlišovací schopnost genu *CYTB* se výrazně snižuje s vyššími úrovněmi taxonomických vztahů (Bradley et al., 2007). Mitochondriální geny obecně oproti těm nukleárním rychleji mutují, což způsobuje častou saturaci třetích kodonových pozic, jejíž pravděpodobnost roste se vzdáleností taxonů. To je pravděpodobně jedním z faktorů přispívající k nižší informativnosti mitochondriálních markerů u hlubších taxonomických vztahů (Springer et al., 2001). Oproti tomu nukleární markery mají často optimální substituční rychlosti dostatečně velké na to, aby umožňovaly vysokou informativnost třetích kodonových pozic, ale zároveň dostatečně malé, aby nebyly ovlivněny

výskytem saturace (Galewski et al., 2006).

Pozice s mírnou pozitivní selekcí ($1 < \omega < 2$) dle výsledků studie Massey et al. (2008) neztrácí skutečný signál, ale dokonce ani u pozic s $\omega > 2$ nebyla ztráta signálu zásadní. K podobným závěrům dospěla také studie Roje (2014), kde autor dokázal, že samotná přítomnost selekce v markeru nevypovídala o kvalitě výsledného stromu a případné odstranění pozitivních pozic nevedlo k napravení inkongruence mezi stromy. Další studie Jansa et al. (2006) porovnávala fylogenetickou informativnost markerů *DMP1* a *RBP3* pro rod vačic. Autoři jistili, že *DMP1* má výrazně menší proporci pozic pod silnou negativní selekcí, vyvíjí se až o 30 % rychleji a má asi $2,5\times$ více pozic pod vlivem pozitivní selekce než marker *RBP3*. Tyto závěry se shodují s výsledky této práce, kde jsem u markeru *DMP1* detekovala přibližně o čtvrtinu méně negativních pozic a $2\times$ více pozitivních pozic (dle výsledků modelu M8). I přes značně odlišné působení selekce oba markery vedly k překvapivě kongruentním stromům. *DMP1* důsledkem uvolněné selekce vykazoval více fylogeneticky informativních pozic a dokázal rozlišit více uzlů stromu s většími podporami. Zdá se tedy, že uvolněná selekce dokázala zlepšit efektivitu markeru *DMP1* bez toho, aniž by snížila jeho přesnost. Je ale nutné vzít v úvahu, že analýza probíhala pouze na rodové úrovni a pro vyšší taxonomické jednotky by efekt mohl být odlišný. Pro blíže příbuzné taxony tedy může pozitivní a uvolněná selekce přinášet užitečný fylogenetický signál. Na základě výsledků modelu cmC (Tab. 8) by pak bylo možné při detekci uvolněné selekce v rámci taxonu zvolit fylogenetické markery, které by potenciálně mohly být v rámci taxonu informativnější.

Dalším výsledkem studie Roje (2014) bylo, že strom na základě neneutrálního markeru, který se od ostatních stromů nejvíce odlišoval, byl také nejvíce vzdálený stromu na základě spojeného alignmentu. To poukazuje na skutečnost, že spojení alignmentů více markerů dokázalo problém inkongruence tohoto neneutrálního markeru eliminovat a společná informace více genů konvergovala k jedné hypotéze. Používání několika různých fylogenetických markerů je běžnou praxí ve většině fylogenetických studií uvedených v Tab. 3 a mohlo by pomáhat předcházet možným negativním vlivům selekce na výsledek rekonstrukce.

Autoři studie Massey et al. (2008) také s pomocí simulovaného proteinu dokázali, že pozitivní selekce může vést k paralelní konvergenci sekvencí, které následně vykazují funkční signál namísto signálu skutečného původu. Tato skutečnost představuje další možnou komplikaci působení pozitivní selekce.

Při odstraňování pozicí pod silnou pozitivní selekcí by bylo vhodné zvážit také možné dopady nerovnovážné vazby s okolními pozicemi. Při rychlé fixaci výhodné mutace dochází k lokálnímu snížení genetické variability, zvanému selektivní zametání (selective sweep) (Kim a Nielsen, 2004). Z toho důvodu by odstranění samotné pozice pod pozitivní selekcí nemuselo být dostačující a bylo by potřeba odstranit také pozice v jejím okolí. Velikost ovlivněného okolí závisí na rekombinační rychlosti, obecně se ale předpokládá, že selektivní zametání může mít vliv i na více než 100 kb v okolí výhodné mutace (McVean, 2007). Důsledkem toho zůstává otázkou, zda je vůbec vhodné využít marker, u kterého detekujeme silnou pozitivní selekci, vzhledem k tomu, že délky sekvencí markerů bývají relativně krátké (v jednotkách kb).

Na základě těchto poznatků se tedy jako nejvhodnější fylogenetický marker analyzovaný v této práci jeví *CSN3*, a to i přes to, že se jedná o marker s největší proporcí pozitivních

pozic s ω okolo hodnoty 2, u kterých by se dalo zvážit jejich odstranění. U taxonů Cetacea, Eulipotyphla a Chiroptera však model bsA pozitivní selekci nedetekoval a pro analýzu těchto skupin by tak selekce nemusela představovat žádnou překážku. Marker obsahuje největší podíl neutrálních pozic (55 %) a zbylé pozice podléhají pouze mírné negativní selekci ($\omega = 0,42$). Většina pozic markeru by tak měla poskytovat spolehlivý signál pravého původu. Nízký tlak negativní selekce na sekvence genu se ale projevil jejich vysokou variabilitou mezi taxony. Ačkoliv bylo stále možné provést alignment napříč sedmi analyzovanými savčími řády, vyšší variabilita sekvencí byla v alignmentu na první pohled patrná ve srovnání s alignmenty ostatních markerů. Po zahrnutí dalších vzdálenějších řádů savců by mohl být správný alignment problematický. V takovém případě by bylo vhodnější volit konzervovanější nukleární markery. Dřívější studie prokázaly relativně nízkou podobnost sekvencí genu mezi zástupci řádu hlodavců a sudokopytníků (Adenaike et al., 2020).

CSN3 kóduje protein savčího mléka a je úzce spojován s laktací, která je nezbytná k úspěšnému dokončení reprodukčního procesu savců (Shekar et al., 2006). I přes to na něj působí nejmírnější negativní selekce, což naznačuje, že nesynonymní substituce v tomto genu často nenarušují jeho správnou funkci. Naopak to vypadá, že je jeho variabilita často výhodná a bylo již dříve dokázáno, že geny savčího mléka jsou vysoce polymorfní (Caroli et al., 2009) a varianty genu *CSN3* mohou ovlivnit srážení a množství mléka i množství proteinů, které mléko obsahuje (Comin et al., 2008).

Druhým nejvhodnějším markerem by pak byl *BRCA1* se 44 % neutrálních pozic a poměrem $\omega = 0,35$ u negativních pozic a marker *APOB*. Ostatní nukleární markery mají malou proporci neutrálních pozic a větší podíl pozicí pod silnou negativní selekcí ($\omega < 0,2$).

Na základě literární rešerše se zdá, že se gen *CSN3* jako fylogenetický marker v současnosti nevyužívá. Dostupné studie pocházejí převážně z 90. let 20. století (Tab. 3). *BRCA1* se dle dostupných studií používá převážně pro fylogenezi hlodavců a vačnatců (Tab. 3).

Z celkem osmi nukleárních markerů, u kterých jsem detekovala pozitivní selekci alespoň u jednoho z taxonů, byl nejčastěji detekován infrařád kytovců a to v pěti případech. Jednalo se také o jeden z nejčastějších taxonů s odlišným působením selekce na základě výsledku modelu cmC. Potvrzuje se tak možné působení odlišných selekčních tlaků na tento infrařád nutných pro adaptaci na vodní prostředí (Thewissen et al., 2009; Ishengoma a Agaba, 2017).

5 Závěr

V této práci jsem testovala signál selekce u 11 nukleárních a 3 mitochondriálních fylogenetických markerů běžně využívaných ve studiích savců. Výsledky ukazují, že žádný z nich není evolučně neutrální, přičemž téměř všechny nukleární markery vykazují známky pozitivní selekce. Mitochondriální i většina nukleárních markerů také podléhají silné negativní selekci. Působení selekce se navíc liší mezi jednotlivými taxony, což může dále ovlivnit výsledky fylogenetické rekonstrukce. Tato zjištění potvrzují, že selekce je rozšířeným jevem i u běžně využívaných markerů, a měla by být při jejich výběru zohledněna.

Tato práce zároveň poskytuje ucelený, byť ne zcela vyčerpávající přehled fylogenetických markerů používaných u savců. Na základě zjištěného působení selekce na jednotlivé markery i konkrétní taxony lze zvolit nejvhodnější marker v závislosti na povaze a cílech fylogenetické analýzy.

I přesto, že selekce, ať už pozitivní nebo negativní, porušuje předpoklady modelů a metod fylogenetické rekonstrukce, nelze pouze na základě detekce selekce přímo rozhodnout o tom, zda je gen vhodným fylogenetickým markerem. Určitou míru negativní selekce najdeme u protein kódujících genů ve většině případů. Několik studií také ukazuje, že samotná přítomnost pozitivní selekce v markerech nemusí představovat překážku při fylogenetické analýze (Jansa et al., 2006; Roje, 2014). Možné působení selekce je však vysoce komplexní, zejména v souvislosti s konvergentní evolucí sekvencí a nerovnovážnou vazbou mezi pozicemi. Tato skutečnost poukazuje na potřebu vývoje nových modelů, které uvolňují přísné předpoklady kladené na používané fylogenetické markery, případně modelů schopných lépe zohlednit možný vliv selekce.

Dalším výzkumem navazujícím na výsledky této práce by mohlo být srovnání fylogenetické informativnosti jednotlivých markerů rozdělených na základě přítomnosti pozitivní nebo silné negativní selekce. Případně by bylo možné rozdělit pozice markerů do několika skupin dle poměru ω a porovnat fylogenetický signál pozic pod silnou selekcí s těmi neutrálními. To by mohlo pomoci lépe pochopit komplexní vztah selekce a fylogenetické rekonstrukce.

Příloha 1

Taxonomie analyzovaných savců

Následující tabulka obsahuje seznam všech 191 analyzovaných savců, jejich řád a případně také podtaxon v rámci kterého proběhla analýza selekčního signálu.

Druh/poddruh	Řád	Podtaxon
Bison bison	Artiodactyla	–
Bos indicus	Artiodactyla	–
Bos mutus	Artiodactyla	–
Bos taurus	Artiodactyla	–
Bubalus bubalis	Artiodactyla	–
Bubalus carabanensis	Artiodactyla	–
Budorcas taxicolor	Artiodactyla	–
Camelus bactrianus	Artiodactyla	–
Camelus dromedarius	Artiodactyla	–
Camelus ferus	Artiodactyla	–
Capra hircus	Artiodactyla	–
Cervus canadensis	Artiodactyla	–
Cervus elaphus	Artiodactyla	–
Hippopotamus amphibius	Artiodactyla	–
Moschus berezovskii	Artiodactyla	–
Odocoileus virginianus	Artiodactyla	–
Oryx dammah	Artiodactyla	–
Ovis aries	Artiodactyla	–
Phacochoerus africanus	Artiodactyla	–
Sus scrofa	Artiodactyla	–
Vicugna pacos	Artiodactyla	–
Balaenoptera acutorostrata	Artiodactyla	Cetacea
Balaenoptera musculus	Artiodactyla	Cetacea
Balaenoptera ricei	Artiodactyla	Cetacea
Delphinapterus leucas	Artiodactyla	Cetacea
Delphinus delphis	Artiodactyla	Cetacea
Globicephala melas	Artiodactyla	Cetacea
Kogia breviceps	Artiodactyla	Cetacea
Lagenorhynchus albirostris	Artiodactyla	Cetacea
Lagenorhynchus obliquidens	Artiodactyla	Cetacea
Lipotes vexillifer	Artiodactyla	Cetacea
Mesoplodon densirostris	Artiodactyla	Cetacea
Monodon monoceros	Artiodactyla	Cetacea
Neophocaena asiaeorientalis	Artiodactyla	Cetacea
Orcinus orca	Artiodactyla	Cetacea
Phocoena sinus	Artiodactyla	Cetacea
Physeter catodon	Artiodactyla	Cetacea
Tursiops truncatus	Artiodactyla	Cetacea
Ailuropoda melanoleuca	Carnivora	Caniformia
Callorhinus ursinus	Carnivora	Caniformia
Canis lupus dingo	Carnivora	Caniformia

Druh/poddruh	Řád	Podtaxon
Canis lupus familiaris	Carnivora	Caniformia
Eumetopias jubatus	Carnivora	Caniformia
Halichoerus grypus	Carnivora	Caniformia
Leptonychotes weddellii	Carnivora	Caniformia
Lontra canadensis	Carnivora	Caniformia
Lutra lutra	Carnivora	Caniformia
Meles meles	Carnivora	Caniformia
Mirounga angustirostris	Carnivora	Caniformia
Mirounga leonina	Carnivora	Caniformia
Mustela erminea	Carnivora	Caniformia
Mustela lutreola	Carnivora	Caniformia
Mustela nigripes	Carnivora	Caniformia
Mustela putorius	Carnivora	Caniformia
Neomonachus schauinslandi	Carnivora	Caniformia
Neovison vison	Carnivora	Caniformia
Nyctereutes procyonoides	Carnivora	Caniformia
Odobenus rosmarus	Carnivora	Caniformia
Phoca vitulina	Carnivora	Caniformia
Ursus americanus	Carnivora	Caniformia
Ursus arctos	Carnivora	Caniformia
Ursus maritimus	Carnivora	Caniformia
Vulpes lagopus	Carnivora	Caniformia
Vulpes vulpes	Carnivora	Caniformia
Zalophus californianus	Carnivora	Caniformia
Acinonyx jubatus	Carnivora	Feliformia
Felis catus	Carnivora	Feliformia
Hyaena hyaena	Carnivora	Feliformia
Leopardus geoffroyi	Carnivora	Feliformia
Lynx canadensis	Carnivora	Feliformia
Lynx rufus	Carnivora	Feliformia
Neofelis nebulosa	Carnivora	Feliformia
Panthera leo	Carnivora	Feliformia
Panthera pardus	Carnivora	Feliformia
Panthera tigris	Carnivora	Feliformia
Panthera uncia	Carnivora	Feliformia
Prionailurus bengalensis	Carnivora	Feliformia
Prionailurus viverrinus	Carnivora	Feliformia
Puma concolor	Carnivora	Feliformia
Puma yagouaroundi	Carnivora	Feliformia
Suricata suricatta	Carnivora	Feliformia
Condylura cristata	Eulipotyphla	–
Erinaceus europaeus	Eulipotyphla	–
Sorex araneus	Eulipotyphla	–
Sorex fumeus	Eulipotyphla	–
Suncus etruscus	Eulipotyphla	–
Talpa occidentalis	Eulipotyphla	–
Artibeus jamaicensis	Chiroptera	–
Desmodus rotundus	Chiroptera	–
Eptesicus fuscus	Chiroptera	–
Hipposideros armiger	Chiroptera	–
Miniopterus natalensis	Chiroptera	–
Molossus molossus	Chiroptera	–
Myotis brandtii	Chiroptera	–

Druh/poddruh	Řád	Podtaxon
Myotis daubentonii	Chiroptera	–
Myotis davidii	Chiroptera	–
Myotis lucifugus	Chiroptera	–
Myotis myotis	Chiroptera	–
Phyllostomus discolor	Chiroptera	–
Phyllostomus hastatus	Chiroptera	–
Pipistrellus kuhlii	Chiroptera	–
Pteronotus mesoamericanus	Chiroptera	–
Pteropus alecto	Chiroptera	–
Pteropus giganteus	Chiroptera	–
Pteropus vampyrus	Chiroptera	–
Rhinolophus ferrumequinum	Chiroptera	–
Rhinolophus sinicus	Chiroptera	–
Rousettus aegyptiacus	Chiroptera	–
Sturnira hondurensis	Chiroptera	–
Diceros bicornis	Perissodactyla	–
Equus asinus	Perissodactyla	–
Equus caballus	Perissodactyla	–
Equus przewalskii	Perissodactyla	–
Equus quagga	Perissodactyla	–
Aotus nancymaae	Primates	–
Callithrix jacchus	Primates	–
Carlito syrichta	Primates	–
Cebus imitator	Primates	–
Cercocebus atys	Primates	–
Colobus angolensis	Primates	–
Gorilla gorilla	Primates	–
Homo sapiens	Primates	–
Hylobates moloch	Primates	–
Chlorocebus sabaeus	Primates	–
Lemur catta	Primates	–
Macaca fascicularis	Primates	–
Macaca mulatta	Primates	–
Macaca nemestrina	Primates	–
Macaca thibetana	Primates	–
Mandrillus leucophaeus	Primates	–
Microcebus murinus	Primates	–
Nomascus leucogenys	Primates	–
Nycticebus coucang	Primates	–
Otolemur garnettii	Primates	–
Pan paniscus	Primates	–
Pan troglodytes	Primates	–
Papio anubis	Primates	–
Piliocolobus tephrosceles	Primates	–
Pongo abelii	Primates	–
Pongo pygmaeus	Primates	–
Propithecus coquereli	Primates	–
Rhinopithecus bieti	Primates	–
Rhinopithecus roxellana	Primates	–
Saimiri boliviensis	Primates	–
Sapajus apella	Primates	–
Symphalangus syndactylus	Primates	–
Theropithecus gelada	Primates	–

Druh/poddruh	Řád	Podtaxon
Trachypithecus francoisi	Primates	–
Acomys russatus	Rodentia	–
Apodemus sylvaticus	Rodentia	–
Arvicanthis niloticus	Rodentia	–
Arvicola amphibius	Rodentia	–
Castor canadensis	Rodentia	–
Cavia porcellus	Rodentia	–
Cricetulus griseus	Rodentia	–
Dipodomys ordii	Rodentia	–
Dipodomys spectabilis	Rodentia	–
Fukomys damarensis	Rodentia	–
Grammomys surdaster	Rodentia	–
Heterocephalus glaber	Rodentia	–
Chinchilla lanigera	Rodentia	–
Chionomys nivalis	Rodentia	–
Ictidomys tridecemlineatus	Rodentia	–
Jaculus jaculus	Rodentia	–
Marmota flaviventris	Rodentia	–
Marmota marmota	Rodentia	–
Marmota monax	Rodentia	–
Mastomys coucha	Rodentia	–
Meriones unguiculatus	Rodentia	–
Mesocricetus auratus	Rodentia	–
Microtus fortis	Rodentia	–
Microtus ochrogaster	Rodentia	–
Microtus oregoni	Rodentia	–
Mus caroli	Rodentia	–
Mus musculus	Rodentia	–
Mus pahari	Rodentia	–
Myodes glareolus	Rodentia	–
Nannospalax galili	Rodentia	–
Octodon degus	Rodentia	–
Onychomys torridus	Rodentia	–
Perognathus longimembris	Rodentia	–
Peromyscus californicus	Rodentia	–
Peromyscus eremicus	Rodentia	–
Peromyscus leucopus	Rodentia	–
Peromyscus maniculatus	Rodentia	–
Phodopus roborovskii	Rodentia	–
Psammomys obesus	Rodentia	–
Rattus norvegicus	Rodentia	–
Rattus rattus	Rodentia	–
Sciurus carolinensis	Rodentia	–
Urocitellus parryii	Rodentia	–

Příloha 2

Přístupová čísla NCBI

Šedě vyznačená přístupová čísla označují sekvence, které jsem při výpočtu selekčního signálu nevyužila z důvodu chybějících dat.

Následující tabulka obsahuje přístupová čísla stažených sekvencí mitochondriálních genů z databáze NCBI.

Druh/puddruh	***COX1***	***COX2***	***CYTB***
Acinonyx jubatus	NC_005212.1:6240-7784	NC_005212.1:7925-8608	NC_005212.1:15061-16200
Acomys russatus	–	–	MH044905.1
Ailuropoda melanoleuca	NC_009492.1:6708-8252	NC_009492.1:8392-9075	NC_009492.1:15529-16668
Aotus nancymaae	NC_018116.1:5324-6880	NC_018116.1:7009-7696	NC_018116.1:14143-15282
Apodemus sylvaticus	NC_049122.1:5325-6869	NC_049122.1:7008-7691	NC_049122.1:14135-15278
Artibeus jamaicensis	NC_002009.1:5330-6874	NC_002009.1:7013-7696	NC_002009.1:14150-15289
Arvicanthis niloticus	MN807587.1:5303-6847	MN807587.1:6986-7669	MN807587.1:14113-15256
Arvicola amphibius	MT381921.1:5306-6850	MT381921.1:6989-7672	MT381921.1:14125-15267
Balaenoptera acutorostrata	NC_005271.1:5372-6922	NC_005271.1:7062-7745	NC_005271.1:14204-15343
Balaenoptera musculus	NC_001601.1:5779-7329	NC_001601.1:7469-8152	NC_001601.1:14610-15749
Balaenoptera ricei	EU496283.1	–	KJ586849.1
Bison bison	NC_012346.1:5686-7230	NC_012346.1:7375-8058	NC_012346.1:14514-15653
Bos indicus	NC_005971.1:5690-7234	NC_005971.1:7377-8060	NC_005971.1:14517-15656
Bos mutus	NC_025563.1:5689-7233	NC_025563.1:7376-8059	NC_025563.1:14515-15654
Bos taurus	NC_006853.1:5687-7231	NC_006853.1:7374-8057	NC_006853.1:14514-15653
Bubalus bubalis	NC_049568.1:5700-7244	NC_049568.1:7388-8071	NC_049568.1:14529-15668
Bubalus carabanensis	NC_006295.1:5701-7245	NC_006295.1:7389-8072	NC_006295.1:14530-15669
Budorcas taxicolor	NC_013069.1:5329-6873	NC_013069.1:7016-7699	NC_013069.1:14155-15294
Callithrix jacchus	NC_025586.1:5319-6858	NC_025586.1:7004-7691	NC_025586.1:14141-15280
Callorhinus ursinus	NC_008415.3:5352-6896	NC_008415.3:7037-7721	NC_008415.3:14181-15320
Camelus bactrianus	NC_009628.2:5335-6879	NC_009628.2:7023-7706	NC_009628.2:14161-15300
Camelus dromedarius	NC_009849.1:5333-6877	NC_009849.1:7021-7704	NC_009849.1:14156-15295
Camelus ferus	NC_009629.2:5334-6878	NC_009629.2:7022-7705	NC_009629.2:14160-15299
Canis lupus dingo	MH035676.1:5350-6894	MH035676.1:7035-7718	MH035676.1:14186-15325
Canis lupus familiaris	NC_002008.4:5349-6893	NC_002008.4:7034-7717	NC_002008.4:14183-15322
Capra hircus	NC_005044.2:5329-6873	NC_005044.2:7016-7699	NC_005044.2:14153-15292
Carlito syrichta	NC_012774.1:5367-6908	NC_012774.1:7061-7744	NC_012774.1:14190-15329
Castor canadensis	NC_033912.1:5344-6888	NC_033912.1:7031-7714	NC_033912.1:14163-15302
Cavia porcellus	NC_000884.1:5329-6870	NC_000884.1:7018-7701	NC_000884.1:14164-15303
Cebus imitator	NC_087897.1:5329-6885	NC_087897.1:7014-7701	NC_087897.1:14170-15306
Cercocebus atys	NC_028592.1:5340-6908	NC_028592.1:7022-7705	NC_028592.1:14203-15343
Cervus canadensis	NC_014703.1:5319-6863	NC_014703.1:7006-7689	NC_014703.1:14146-15285
Cervus elaphus	NC_007704.2:5336-6880	NC_007704.2:7023-7706	NC_007704.2:14162-15301
Colobus angolensis	MW049052.1	–	–
Condylura cristata	NC_029762.1:5359-6903	NC_029762.1:7052-7735	NC_029762.1:14195-15334
Cricetulus griseus	NC_007936.1:5310-6854	NC_007936.1:6994-7677	NC_007936.1:14123-15265
Delphinapterus leucas	NC_034236.1:5365-6915	NC_034236.1:7055-7738	NC_034236.1:14193-15332
Delphinus delphis	NC_036415.1:5362-6912	NC_036415.1:7052-7735	NC_036415.1:14194-15333
Desmodus rotundus	NC_022423.1:5334-6878	NC_022423.1:7019-7702	NC_022423.1:14140-15279
Diceros bicornis	NC_012682.1:5351-6895	NC_012682.1:7042-7725	NC_012682.1:14180-15319
Dipodomys ordii	–	–	AF173501.1
Dipodomys spectabilis	NC_068808.1:6168-7715	NC_068808.1:7857-8540	NC_068808.1:14996-16135
Eptesicus fuscus	BK063052.1:6264-7808	BK063052.1:7960-8643	BK063052.1:15096-16235
Equus asinus	NC_001788.1:5357-6901	NC_001788.1:7043-7726	NC_001788.1:14185-15324
Equus caballus	NC_091244.1:5359-6903	NC_091244.1:7045-7728	NC_091244.1:14185-15324
Equus przewalskii	NC_024030.1:5358-6902	NC_024030.1:7044-7727	NC_024030.1:14184-15323
Equus quagga	NC_044858.1:5356-6900	NC_044858.1:7042-7725	NC_044858.1:14182-15321
Erinaceus europaeus	NC_002080.2:5346-6890	NC_002080.2:7034-7717	NC_002080.2:14182-15321

Druh/puddruh	*COX1*	*COX2*	*CYTB*
Eumetopias jubatus	NC_004030.2:5355-6899	NC_004030.2:7039-7722	NC_004030.2:14184-15323
Felis catus	NC_001700.1:6216-7760	NC_001700.1:7902-8585	NC_001700.1:15038-16177
Fukomys damarensis	NC_027742.1:5322-6863	NC_027742.1:7009-7692	NC_027742.1:14150-15289
Globicephala melas	NC_083222.1:5364-6914	NC_083222.1:7054-7737	NC_083222.1:14196-15335
Gorilla gorilla	NC_001645.1:5325-6875	NC_001645.1:7007-7690	NC_001645.1:14169-15309
Grammomys surdaster	NC_053796.1:5304-6848	NC_053796.1:6987-7670	NC_053796.1:14106-15249
Halichoerus grypus	NC_001602.1:6249-7793	NC_001602.1:7933-8616	NC_001602.1:15067-16206
Heterocephalus glaber	NC_015112.1:5320-6861	NC_015112.1:7007-7690	NC_015112.1:14149-15283
Hippopotamus amphibius	NC_000889.1:5339-6884	NC_000889.1:7032-7715	NC_000889.1:14172-15311
Hipposideros armiger	NC_018540.1:5342-6886	NC_018540.1:7027-7710	NC_018540.1:14161-15300
Homo sapiens	NC_012920.1:5904-7445	NC_012920.1:7586-8269	NC_012920.1:14747-15887
Hyaena hyaena	NC_020669.1:6130-7674	NC_020669.1:7816-8499	NC_020669.1:14952-16091
Hylobates moloch	HQ622782.1:5326-6867	HQ622782.1:7009-7692	HQ622782.1:14176-15316
Chinchilla lanigera	NC_021386.1:5331-6875	NC_021386.1:7018-7701	NC_021386.1:14163-15302
Chionomys nivalis	NC_065750.1:5302-6846	NC_065750.1:6985-7668	NC_065750.1:14120-15262
Chlorocebus sabaeus	NC_008066.1:5899-7437	NC_008066.1:7582-8265	NC_008066.1:14732-15872
Ictidomys tridecemlineatus	NC_027278.1:5347-6888	NC_027278.1:7032-7715	NC_027278.1:14162-15301
Jaculus jaculus	NC_005314.1:5350-6894	NC_005314.1:7034-7717	NC_005314.1:14170-15309
Kogia breviceps	NC_005272.1:5362-6912	NC_005272.1:7052-7735	NC_005272.1:14191-15330
Lagenorhynchus albirostris	NC_083094.1:5366-6916	NC_083094.1:7056-7739	NC_083094.1:14199-15338
Lagenorhynchus obliquidens	NC_035426.1:5365-6915	NC_035426.1:7055-7738	NC_035426.1:14198-15337
Lemur catta	NC_059325.1:5341-6882	NC_059325.1:7028-7711	NC_059325.1:14176-15315
Leopardus geoffroyi	NC_028320.1:6123-7667	NC_028320.1:7809-8492	NC_028320.1:14945-16084
Leptonychotes weddellii	NC_008424.1:5351-6895	NC_008424.1:7035-7718	NC_008424.1:14171-15310
Lipotes vexillifer	NC_007629.1:5366-6916	NC_007629.1:7056-7739	NC_007629.1:14196-15335
Lontra canadensis	NC_071788.1:5340-6884	NC_071788.1:7024-7707	NC_071788.1:14170-15309
Lutra lutra	NC_062277.1:5344-6888	NC_062277.1:7028-7711	NC_062277.1:14173-15312
Lynx canadensis	NC_028313.1:6019-7563	NC_028313.1:7705-8388	NC_028313.1:14841-15980
Lynx rufus	NC_014456.1:6241-7785	NC_014456.1:7927-8610	NC_014456.1:15064-16203
Macaca fascicularis	NC_012670.1:5858-7396	NC_012670.1:7540-8223	NC_012670.1:14749-15889
Macaca mulatta	NC_005943.1:5850-7418	NC_005943.1:7532-8215	NC_005943.1:14741-15881
Macaca nemestrina	NC_026976.1:5848-7416	NC_026976.1:7530-8213	NC_026976.1:14739-15876
Macaca thibetana	NC_011519.1:5839-7377	NC_011519.1:7521-8204	NC_011519.1:14718-15858
Mandrillus leucophaeus	NC_028442.1:5332-6900	NC_028442.1:7014-7697	NC_028442.1:14195-15335
Marmota flaviventris	NC_042243.1:5347-6888	NC_042243.1:7032-7715	NC_042243.1:14162-15301
Marmota marmota	MN935776.1:5345-6886	MN935776.1:7030-7713	MN935776.1:14160-15299
Marmota monax	NC_080739.1:5348-6889	NC_080739.1:7033-7716	NC_080739.1:14162-15301
Mastomys coucha	NC_036018.1:5303-6847	NC_036018.1:6985-7668	NC_036018.1:14114-15257
Meles meles	NC_011125.1:5342-6886	NC_011125.1:7026-7709	NC_011125.1:14170-15309
Meriones unguiculatus	NC_023263.1:5336-6880	NC_023263.1:7019-7702	NC_023263.1:14147-15286
Mesocricetus auratus	NC_013276.1:5310-6854	NC_013276.1:6997-7680	NC_013276.1:14125-15265
Mesoplodon densirostris	NC_021974.2:5367-6917	NC_021974.2:7057-7740	NC_021974.2:14194-15333
Microcebus murinus	NC_028718.1:5385-6926	NC_028718.1:7071-7754	NC_028718.1:14222-15361
Microtus fortis	NC_081580.1:5464-7008	NC_081580.1:7147-7830	NC_081580.1:14280-15422
Microtus ochrogaster	NC_027945.1:5303-6847	NC_027945.1:6986-7669	NC_027945.1:14118-15260
Microtus oregoni	–	–	GU954308.1
Miniopterus natalensis	OP137089.1	–	OP157142.1
Mirounga angustirostris	AY377139.1	AY377162.1	AY377325.1
Mirounga leonina	NC_008422.1:5353-6897	NC_008422.1:7037-7720	NC_008422.1:14172-15311
Molossus molossus	NC_065689.1:5351-6895	NC_065689.1:7036-7719	NC_065689.1:14175-15314
Monodon monoceros	NC_005279.1:5364-6914	NC_005279.1:7054-7737	NC_005279.1:14193-15332
Moschus berezovskii	NC_012694.1:5326-6870	NC_012694.1:7013-7696	NC_012694.1:14153-15292
Mus caroli	NC_025268.1:5338-6882	NC_025268.1:7022-7705	NC_025268.1:14153-15296
Mus musculus	NC_005089.1:5328-6872	NC_005089.1:7013-7696	NC_005089.1:14145-15288
Mus pahari	NC_036680.1:5325-6869	NC_036680.1:7009-7692	NC_036680.1:14136-15279
Mustela erminea	NC_025516.1:5645-7189	NC_025516.1:7328-8011	NC_025516.1:14472-15611
Mustela lutreola	NC_056132.1:5340-6884	NC_056132.1:7023-7706	NC_056132.1:14167-15306
Mustela nigripes	NC_024942.1:5341-6885	NC_024942.1:7024-7707	NC_024942.1:14168-15307
Mustela putorius	NC_020638.1:5343-6887	NC_020638.1:7026-7709	NC_020638.1:14170-15309
Myodes glareolus	NC_024538.1:5309-6853	NC_024538.1:6992-7675	NC_024538.1:14128-15270
Myotis brandtii	NC_025308.1:5349-6893	NC_025308.1:7050-7733	NC_025308.1:14176-15315
Myotis daubentonii	MN122860.1:5347-6891	MN122860.1:7049-7732	MN122860.1:14188-15327
Myotis davidii	NC_025568.1:5343-6887	NC_025568.1:7044-7727	NC_025568.1:14186-15325
Myotis lucifugus	NC_029849.1:5344-6888	NC_029849.1:7046-7729	NC_029849.1:14171-15310
Myotis myotis	NC_029346.1:5345-6889	NC_029346.1:7047-7730	NC_029346.1:14186-15325
Nannospalax galili	NC_020754.1:5331-6875	NC_020754.1:7012-7695	NC_020754.1:14150-15289
Neofelis nebulosa	NC_008450.1:6244-7788	NC_008450.1:7930-8613	NC_008450.1:15066-16205

Druh/puddruh	*COX1*	*COX2*	*CYTB*
Neomonachus schauinslandi	NC_008421.1:5357-6901	NC_008421.1:7041-7724	NC_008421.1:14177-15316
Neophocaena asiaeorientalis	NC_026456.1:5366-6919	NC_026456.1:7055-7738	NC_026456.1:14196-15335
Neovison vison	NC_020641.1:5339-6883	NC_020641.1:7026-7709	NC_020641.1:14170-15309
Nomascus leucogenys	NC_021957.1:5325-6875	NC_021957.1:7007-7690	NC_021957.1:14169-15309
Nyctereutes procyonoides	NC_013700.1:5347-6891	NC_013700.1:7042-7725	NC_013700.1:14194-15333
Nycticebus coucang	NC_002765.1:5379-6920	NC_002765.1:7064-7747	NC_002765.1:14202-15341
Octodon degus	NC_020661.1:5344-6888	NC_020661.1:7036-7719	NC_020661.1:14191-15330
Odobenus rosmarus	NC_004029.2:5350-6894	NC_004029.2:7034-7718	NC_004029.2:14173-15312
Odocoileus virginianus	NC_015247.1:380-1924	NC_015247.1:2067-2750	NC_015247.1:9207-10346
Onychomys torridus	HQ980049.1	–	KY754082.1
Orcinus orca	NC_064558.1:5365-6915	NC_064558.1:7055-7738	NC_064558.1:14197-15336
Oryx dammah	NC_058019.1:5325-6869	NC_058019.1:7012-7695	NC_058019.1:14153-15292
Otolemur garnettii	AY671787.1	–	KP410640.1
Ovis aries	NC_001941.1:5332-6876	NC_001941.1:7019-7702	NC_001941.1:14159-15298
Pan paniscus	NC_001644.1:5322-6863	NC_001644.1:7004-7687	NC_001644.1:14166-15306
Pan troglodytes	NC_001643.1:5321-6862	NC_001643.1:7003-7686	NC_001643.1:14165-15305
Panthera leo	NC_018053.1:6222-7775	NC_018053.1:7916-8599	NC_018053.1:15052-16191
Panthera pardus	NC_010641.1:6305-7849	NC_010641.1:7991-8674	NC_010641.1:15127-16266
Panthera tigris	NC_010642.1:6280-7824	NC_010642.1:7966-8649	NC_010642.1:15103-16242
Panthera uncia	NC_010638.1:6099-7643	NC_010638.1:7785-8468	NC_010638.1:14921-16060
Papio anubis	NC_020006.2:5332-6900	NC_020006.2:7014-7697	NC_020006.2:14168-15308
Perognathus longimembris	EF156865.1	–	U65302.1
Peromyscus californicus	OP524493.1:5316-6860	OP524493.1:6999-7682	OP524493.1:14145-15288
Peromyscus eremicus	NC_047188.1:5320-6864	NC_047188.1:7003-7686	NC_047188.1:14148-15291
Peromyscus leucopus	NC_037180.1:5326-6870	NC_037180.1:7009-7692	NC_037180.1:14153-15296
Peromyscus maniculatus	NC_039921.1:5320-6864	NC_039921.1:7003-7686	NC_039921.1:14147-15290
Phacochoerus africanus	NC_008830.1:6618-8162	NC_008830.1:8310-8997	NC_008830.1:15447-16586
Phoca vitulina	NC_001325.1:6275-7819	NC_001325.1:7959-8642	NC_001325.1:15094-16233
Phocoena sinus	NC_053751.1:5388-6938	NC_053751.1:7078-7761	NC_053751.1:14237-15376
Phodopus roborovskii	NC_031809.1:5314-6858	NC_031809.1:6997-7680	NC_031809.1:14124-15266
Phyllostomus discolor	NC_065690.1:5347-6891	NC_065690.1:7032-7715	NC_065690.1:14172-15311
Phyllostomus hastatus	EU096840.1	M80906.1	FJ155479.1
Physeter catodon	NC_002503.2:5369-6919	NC_002503.2:7059-7742	NC_002503.2:14199-15338
Piliocolobus tephrosceles	–	–	–
Pipistrellus kuhlii	KU058655.1:5341-6891	KU058655.1:7032-7715	KU058655.1:14160-15299
Pongo abelii	NC_002083.1:5342-6884	NC_002083.1:7023-7706	NC_002083.1:14201-15341
Pongo pygmaeus	NC_001646.1:5332-6871	NC_001646.1:7014-7697	NC_001646.1:14191-15331
Prionailurus bengalensis	NC_016189.1:6196-7740	NC_016189.1:7882-8565	NC_016189.1:15019-16158
Prionailurus viverrinus	NC_028305.1:6096-7640	NC_028305.1:7782-8465	NC_028305.1:14919-16058
Propithecus coquereli	NC_011053.1:5350-6894	NC_011053.1:7035-7718	NC_011053.1:14182-15321
Psammomys obesus	NC_037509.1:5327-6871	NC_037509.1:7011-7694	NC_037509.1:14151-15294
Pteronotus mesoamericanus	KX590231.1	–	KX589922.1
Pteropus alecto	NC_023122.1:5349-6893	NC_023122.1:7034-7717	NC_023122.1:14172-15311
Pteropus giganteus	MZ573763.1	–	JN398211.1
Pteropus vampyrus	NC_026542.1:5344-6888	NC_026542.1:7029-7712	NC_026542.1:14178-15317
Puma concolor	NC_016470.1:6263-7807	NC_016470.1:7949-8632	NC_016470.1:15085-16224
Puma yagouaroundi	NC_028311.1:6049-7593	NC_028311.1:7735-8418	NC_028311.1:14871-16010
Rattus norvegicus	NC_001665.2:5323-6867	NC_001665.2:7006-7689	NC_001665.2:14136-15278
Rattus rattus	NC_012374.1:5316-6860	NC_012374.1:6999-7682	NC_012374.1:14127-15269
Rhinolophus ferrumequinum	NC_016191.1:5342-6886	NC_016191.1:7027-7710	NC_016191.1:14166-15305
Rhinolophus sinicus	KP257597.1:5343-6887	KP257597.1:7028-7711	KP257597.1:14164-15303
Rhinopithecus bieti	NC_015486.1:5325-6869	NC_015486.1:7008-7691	NC_015486.1:14191-15325
Rhinopithecus roxellana	NC_008218.1:5323-6867	NC_008218.1:7006-7689	NC_008218.1:14190-15324
Rousettus aegyptiacus	NC_007393.1:5347-6891	NC_007393.1:7032-7715	NC_007393.1:14185-15324
Saimiri boliviensis	NC_021966.1:5313-6854	NC_021966.1:6997-7670	NC_021966.1:14131-15270
Sapajus apella	NC_064184.1:5326-6882	NC_064184.1:7010-7697	NC_064184.1:14169-15305
Sciurus carolinensis	NC_050012.1:5352-6893	NC_050012.1:7038-7721	NC_050012.1:14168-15307
Sorex araneus	NC_027963.1:5350-6894	NC_027963.1:7039-7722	NC_027963.1:14382-15521
Sorex fumeus	OR530518.1	–	AB175116.1
Sturnira hondurensis	–	–	KC753793.1
Suncus etruscus	OP855700.1	OP846408.1	LC126597.1
Suricata suricatta	NC_045900.1:5356-6900	NC_045900.1:7057-7740	NC_045900.1:14194-15333
Sus scrofa	NC_000845.1:6511-8055	NC_000845.1:8203-8890	NC_000845.1:15342-16481
Symphalangus syndactylus	NC_014047.1:5329-6870	NC_014047.1:7012-7695	NC_014047.1:14175-15315
Talpa occidentalis	NC_039630.1:5352-6896	NC_039630.1:7043-7726	NC_039630.1:14181-15320
Theropithecus gelada	NC_019802.1:5342-6883	NC_019802.1:7024-7707	NC_019802.1:14177-15311
Trachypithecus francoisi	NC_023970.1:5327-6871	NC_023970.1:7010-7693	NC_023970.1:14186-15326

Druh/puddruh	*COX1*	*COX2*	*CYTB*
Tursiops truncatus	NC_012059.1:5363-6913	NC_012059.1:7053-7736	NC_012059.1:14195-15334
Urocitellus parryii	NC_059785.1:5352-6893	NC_059785.1:7037-7720	NC_059785.1:14167-15306
Ursus americanus	NC_003426.1:6291-7835	NC_003426.1:7975-8658	NC_003426.1:15113-16252
Ursus arctos	NC_003427.1:6484-8028	NC_003427.1:8168-8851	NC_003427.1:15306-16445
Ursus maritimus	NC_003428.1:6479-8023	NC_003428.1:8163-8846	NC_003428.1:15301-16440
Vicugna pacos	NC_002504.1:5326-6870	NC_002504.1:7014-7697	NC_002504.1:14154-15293
Vulpes lagopus	NC_026529.1:5351-6895	NC_026529.1:7036-7719	NC_026529.1:14186-15325
Vulpes vulpes	NC_008434.1:5350-6894	NC_008434.1:7035-7718	NC_008434.1:14185-15324
Zalophus californianus	NC_008416.1:5358-6902	NC_008416.1:7042-7723	NC_008416.1:14185-15324

První ze tří tabulek obsahujících přístupová čísla stažených sekvencí nukleárních genů z databáze NCBI.

Druh/puddruh	*APOB*	*BRCA1*	*CSN3*	*DMP1*
Acinonyx jubatus	XM_027033268.2	XM_027048935.2	XM_015076911.2	XM_015067468.3
Acomys russatus	XM_051143365.1	XM_051158702.1	XM_051170355.1	XM_051170631.1
Ailuropoda melanoleuca	XM_034659850.1	XM_034641122.1	XM_019799956.1	XM_011232259.2
Aotus nancymaae	XM_021677055.2	XM_021673672.2	XM_064364365.1	XM_012437662.2
Apodemus sylvaticus	XM_052184767.1	XM_052194670.1	–	XM_052198604.1
Artibeus jamaicensis	XM_037151750.2	XM_053661375.1	XM_053655669.1	XM_037150023.2
Arvicanthis niloticus	XM_034514286.1	XM_034507638.1	XM_034508204.1	XM_034491815.1
Arvicola amphibius	XM_038318370.2	XM_038328437.1	XM_038317770.1	XM_038315807.1
Balaenoptera acutorostrata	XM_028166784.2	XM_057535628.1	XM_007193562.2	XM_007184122.2
Balaenoptera musculus	XM_036872347.1	XM_036837937.1	XM_036853293.1	XM_036852046.1
Balaenoptera ricei	XM_059943047.1	XM_059907262.1	XM_059924062.1	XM_059923885.1
Bison bison	XM_010840537.1	XM_010842466.1	XM_010839113.1	XM_010847789.1
Bos indicus	XM_019969506.1	XM_019981053.1	XM_019962873.1	XM_019962170.1
Bos mutus	XM_005895759.2	XM_005909180.2	XM_005897042.2	XM_005911009.2
Bos taurus	XM_024999521.2	NM_178573.1	NM_174294.2	NM_174038.3
Bubalus bubalis	XM_044926208.2	XM_025280062.3	NM_001290972.1	XM_006048654.4
Bubalus carabanensis	XM_055541755.1	XM_055575774.1	XM_055588734.1	XM_055587519.1
Budorcas taxicolor	XM_052649316.1	XM_052658434.1	XM_052642252.1	XM_052642304.1
Callithrix jacchus	XM_035273095.2	XM_008996686.4	XM_017970503.4	XM_002745615.5
Callorhinus ursinus	XM_025867581.1	XM_025861663.1	XM_025886930.1	XM_025873390.1
Camelus bactrianus	XM_010947387.2	XM_010968937.2	NM_001303518.1	XM_010966146.2
Camelus dromedarius	XM_031466641.2	XM_010981838.3	NM_001303560.1	XM_031433662.2
Camelus ferus	XM_032497103.1	XM_006175151.3	XM_006188672.2	XM_014551291.2
Canis lupus dingo	XM_025454549.3	XM_025440430.3	XM_035711301.2	XM_025425305.3
Canis lupus familiaris	XM_038690816.1	NM_001013416.1	XM_038684617.1	XM_038444085.1
Capra hircus	XM_018055590.1	XM_018065164.1	NM_001285587.1	XM_005681883.3
Carlito syrichta	XM_021716809.1	XM_021716224.1	XM_008069910.1	XM_021718826.1
Castor canadensis	XM_020187545.1	XM_020165150.1	XM_020151672.1	XM_020153734.1
Cavia porcellus	XM_003473159.3	XM_013149696.3	–	XM_003469364.3
Cebus imitator	XM_017526598.1	XM_017509326.2	XM_037728916.1	XM_017545123.2
Cercocebus atys	XM_012044649.1	XM_012046789.1	XM_012082924.1	XM_012040136.1
Cervus canadensis	XM_043469277.1	XM_043467647.1	XM_043447866.1	XM_043448363.1
Cervus elaphus	XM_043917599.1	XM_043904730.1	XM_043906554.1	XM_043906229.1
Colobus angolensis	XM_011950671.1	XM_011929928.1	XM_011927095.1	XM_011951818.1
Condylura cristata	XM_004686059.1	XM_012726106.1	XM_004681237.1	XM_004681135.2
Cricetulus griseus	XM_035442518.1	XM_027428048.2	XM_035452493.1	XM_035452154.1
Delphinapterus leucas	XM_022574358.1	XM_022557763.2	XM_022561707.1	XM_022582285.2
Delphinus delphis	XM_060027042.1	XM_059997408.1	XM_060011720.1	XM_060011597.1
Desmodus rotundus	XM_045189621.2	XM_053930516.1	XM_024578113.2	XM_024574965.3
Diceros bicornis	XM_058551724.1	XM_058561051.1	XM_058546427.1	XM_058546459.1
Dipodomys ordii	XM_013013544.1	XM_013031946.1	XM_013020079.1	XM_013029827.1
Dipodomys spectabilis	XM_042680732.1	XM_042674019.1	XM_042678381.1	XM_042668439.1
Eptesicus fuscus	XM_054728236.1	XM_054709480.1	XM_054722155.1	XM_054723290.1
Equus asinus	XM_044772007.1	XM_014833644.2	XM_014847264.2	XM_044766644.1
Equus caballus	XM_023619446.1	XM_023652469.1	NM_001081884.1	XM_005608642.3
Equus przewalskii	XM_008539694.1	XM_008543412.1	XM_008512353.1	XM_008541323.1
Equus quagga	XM_046663340.1	XM_046676719.1	XM_046655958.1	XM_046656928.1
Erinaceus europaeus	XM_007516778.3	XM_060203427.1	XM_007523557.2	XM_060188601.1
Eumetopias jubatus	XM_028099515.1	XM_028093921.1	XM_028100524.1	XM_028117743.1

Druh/puddruh	*APOB*	*BRCA1*	*CSN3*	*DMP1*
Felis catus	XM_003984468.6	XM_019817933.3	XM_003985304.2	XM_019829256.2
Fukomys damarensis	XM_033762857.1	XM_010642987.3	XM_019204813.1	XM_010606738.2
Globicephala melas	XM_030858310.2	XM_030840103.2	XM_030880738.2	XM_030865778.2
Gorilla gorilla	XM_004028908.4	XM_031011552.3	XM_063705126.1	XM_063705130.1
Grammomys surdaster	XM_028779394.1	XM_028780663.1	XM_028752574.1	XM_028773716.1
Halichoerus grypus	XM_036112735.1	XM_036095499.1	XM_036068742.1	XM_036096724.1
Heterocephalus glaber	XM_004839070.3	XM_021240178.1	XM_004871751.2	XM_021248804.1
Hippopotamus amphibius	XM_057743096.1	XM_057714392.1	XM_057728576.1	XM_057727720.1
Hipposideros armiger	XM_019651837.1	XM_019667034.1	XM_019663200.1	XM_019668077.1
Homo sapiens	NM_000384.3	NM_007294.4	NM_005212.3	NM_004407.4
Hyaena hyaena	XM_039222451.1	XM_039224283.1	XM_039251779.1	XM_039225116.1
Hylobates moloch	XM_032172346.2	XM_032163745.2	XM_058429953.1	XM_032163799.1
Chinchilla lanigera	XM_005400358.2	XM_013519228.1	XM_013518493.1	XM_005401334.2
Chionomys nivalis	XM_057780956.1	XM_057775846.1	XM_057773219.1	XM_057772442.1
Chlorocebus sabaeus	XM_038006617.1	XM_037993842.1	XM_037994723.1	XM_007999196.2
Ictidomys tridecemlineatus	XM_005322645.4	XM_040268959.1	XM_005333303.2	XM_021721170.2
Jaculus jaculus	XM_004663762.2	XM_045159814.1	XM_045161216.1	XM_012949489.2
Kogia breviceps	XM_059079137.2	XM_067020706.1	XM_059066193.2	XM_067035466.1
Lagenorhynchus albirostris	XM_060168948.1	XM_060135992.1	XM_060147248.1	XM_060147091.1
Lagenorhynchus obliquidens	XM_027113739.1	XM_027094654.1	XM_027089653.1	XM_027116946.1
Lemur catta	XM_045549123.1	XM_045526646.1	XM_045532472.1	XM_045536796.1
Leopardus geoffroyi	XM_045447603.1	XM_045489007.1	XM_045450396.1	XM_045473327.1
Leptonychotes weddellii	XM_031039224.1	–	XM_031023736.1	XM_006727042.2
Lipotes vexillifer	XM_007463760.1	XM_007449887.1	XM_007450170.1	XM_007450688.1
Lontra canadensis	XM_032866920.1	XM_032871740.1	XM_032863403.1	XM_032873928.1
Lutra lutra	XM_047745761.1	XM_047706070.1	XM_047719463.1	XM_047719913.1
Lynx canadensis	XM_030311575.1	XM_032591074.1	XM_030311988.1	XM_030313467.1
Lynx rufus	XM_047067570.1	XM_047093517.1	XM_047091253.1	XM_047090444.1
Macaca fascicularis	XM_005576465.4	XM_065532409.1	XM_005555144.4	XM_065545236.1
Macaca mulatta	XM_001097500.4	NM_001114949.1	XM_028848535.1	XM_015138768.2
Macaca nemestrina	XM_011738185.2	XM_011725001.2	XM_011733810.1	XM_011737742.2
Macaca thibetana	XM_050754588.1	XM_050762519.1	XM_050791434.1	XM_050790472.1
Mandrillus leucophaeus	XM_011971397.1	XM_011999017.1	XM_011968291.1	XM_011990741.1
Marmota flaviventris	XM_027937775.1	XM_034635200.1	XM_027950891.1	XM_027939866.1
Marmota marmota	XM_015479940.2	XM_048790680.1	XM_048785743.1	XM_048784850.1
Marmota monax	XM_046441334.2	XM_046446634.2	XM_046428733.2	XM_046457644.2
Mastomys coucha	XM_031355760.1	XM_031354074.1	XM_031342043.1	XM_031342199.1
Meles meles	XM_045979562.1	XM_045986449.1	XM_045988511.1	XM_045997595.1
Meriones unguiculatus	XM_021654304.2	XM_021641180.2	XM_060378814.1	XM_060381062.1
Mesocricetus auratus	XM_005079084.4	XM_040749756.1	XM_040747354.1	XM_005076370.4
Mesoplodon densirostris	XM_060117213.1	XM_060080410.1	XM_060090460.1	XM_060114964.1
Microcebus murinus	XM_012761208.2	XM_020281092.1	XM_020284230.1	XM_020284563.1
Microtus fortis	XM_050141959.1	XM_050156467.1	XM_050149612.1	XM_050155801.1
Microtus ochrogaster	XM_005360576.3	XM_026777053.1	XM_005359601.1	XM_013350875.2
Microtus oregoni	XM_041639366.1	XM_041644027.1	XM_041671064.1	XM_041642122.1
Miniopterus natalensis	XM_016224750.1	XM_016202128.1	XM_016207082.1	XM_016213041.1
Mirounga angustirostris	XM_045883941.2	XM_064582502.1	XM_045881060.1	XM_064582746.1
Mirounga leonina	XM_035000523.1	XM_035027209.1	XM_035004866.1	XM_035004553.1
Molossus molossus	XM_036263297.1	XM_036269563.1	–	XM_036244687.1
Monodon monoceros	XM_029232060.1	XM_029212505.1	XM_029235780.1	XM_029235989.1
Moschus berezovskii	XM_055400866.1	XM_055407952.1	XM_055434592.1	XM_055434751.1
Mus caroli	XM_021178578.2	XM_021176032.1	XM_021163837.2	XM_021163789.1
Mus musculus	NM_009693.2	NM_009764.3	NM_007786.5	NM_001359013.1
Mus pahari	XM_021202293.2	XM_021211578.1	XM_021211343.1	XM_021211272.2
Mustela erminea	XM_032353203.1	XM_032322956.1	XM_032334502.1	XM_032332837.1
Mustela lutreola	XM_059132540.1	XM_059149814.1	XM_059153813.1	XM_059154998.1
Mustela nigripes	XM_059405208.1	XM_059379007.1	XM_059396689.1	XM_059394465.1
Mustela putorius	XM_004745883.2	XM_004772608.3	XM_045079928.1	XM_004756109.3
Myodes glareolus	XM_048439666.1	XM_048445737.1	XM_048427667.1	XM_048428173.1
Myotis brandtii	XM_005874933.2	XM_014547201.1	XM_014529479.1	XM_014545959.1
Myotis daubentonii	XM_059660452.1	XM_059670278.1	XM_059706547.1	XM_059681911.1
Myotis davidii	XM_015566436.1	XM_015570568.1	XM_015570058.1	XM_015561527.1
Myotis lucifugus	XM_023756429.1	XM_014458810.2	XM_014468332.1	XM_023750292.1
Myotis myotis	XM_036335239.1	XM_036339652.1	XM_036329413.1	XM_036305861.1
Nannospalax galili	XM_029553976.1	XM_029557415.1	XM_008829590.1	XM_008851733.2
Neofelis nebulosa	XM_058682637.1	XM_058702611.1	XM_058722355.1	XM_058721953.1
Neomonachus schauinslandi	XM_021697339.2	XM_021681657.1	XM_044913099.1	XM_021702341.1

Druh/puddruh	*APOB*	*BRCA1*	*CSN3*	*DMP1*
Neophocaena asiaeorientalis	XM_024747803.1	XM_024742367.1	XM_024741611.1	XM_024736205.1
Neovison vison	XM_044262005.1	XM_044247830.1	XM_044225286.1	XM_044224556.1
Nomascus leucogenys	XM_003270568.4	NM_001308668.1	XM_003265702.2	XM_003265878.3
Nyctereutes procyonoides	XM_055327418.1	XM_055322905.1	XM_055315308.1	XM_055313216.1
Nycticebus coucang	XM_053587405.1	XM_053568910.1	XM_053594417.1	XM_053594938.1
Octodon degus	XM_023710934.1	XM_023718048.1	XM_004647923.1	XM_004629159.2
Odobenus rosmarus	XM_004405522.1	XM_012565923.1	XM_004392891.1	XM_004391535.1
Odocoileus virginianus	XM_020898266.1	XM_020886527.1	XM_020873526.1	XM_020904115.1
Onychomys torridus	XM_036171327.1	XM_036194496.1	XM_036200467.1	XM_036200211.1
Orcinus orca	XM_004274849.3	XM_033410957.2	XM_012532551.2	XM_033421505.2
Oryx dammah	XM_040225679.1	XM_040245226.1	XM_040237013.1	XM_040237296.1
Otolemur garnettii	XM_023511812.1	XM_023511242.1	XM_003800978.1	XM_023513418.1
Ovis aries	XM_027966544.2	XM_042256004.1	NM_001009378.1	XM_012180327.4
Pan paniscus	XM_034952688.3	NM_001301758.1	XM_003808999.5	XM_055111604.2
Pan troglodytes	XM_515323.7	NM_001045493.1	XM_054683040.1	XM_054683177.2
Panthera leo	XM_042933585.1	XM_042917145.1	XM_042936239.1	XM_042933970.1
Panthera pardus	XM_053891121.1	XM_019435327.2	XM_019463579.2	XM_019464822.2
Panthera tigris	XM_042982424.1	XM_042967579.1	XM_007086916.2	XM_042982711.1
Panthera uncia	XM_049652009.1	XM_049636158.1	XM_049630585.1	XM_049630855.1
Papio anubis	XM_021924617.2	NM_001302103.1	XM_003898807.5	XM_021938437.2
Perognathus longimembris	XM_048353386.1	XM_048367069.1	XM_048364863.1	XM_048364496.1
Peromyscus californicus	XM_052745659.1	XM_052756257.1	–	XM_052736005.1
Peromyscus eremicus	XM_059248518.1	XM_059271850.1	–	XM_059275678.1
Peromyscus leucopus	XM_028861764.2	XM_037201135.1	XM_028863965.2	XM_028867454.2
Peromyscus maniculatus	XM_006981434.3	XM_042282853.1	XM_006991643.3	XM_042285435.1
Phacochoerus africanus	XM_047782560.1	XM_047759194.1	XM_047799179.1	XM_047797958.1
Phoca vitulina	XM_032391593.1	XM_032396441.1	XM_032400501.1	XM_032400768.1
Phocoena sinus	XM_032653146.1	XM_032615410.1	XM_032632956.1	XM_032633391.1
Phodopus roborovskii	XM_051181687.1	XM_051177549.1	XM_051198594.1	XM_051198525.1
Phyllostomus discolor	XM_036027694.1	XM_036033957.1	XM_028505606.2	XM_036026754.1
Phyllostomus hastatus	XM_045844323.1	XM_045837007.1	XM_045857681.1	XM_045820396.1
Physeter catodon	XM_028496582.2	XM_028498119.2	XM_024132014.1	XM_028491585.2
Piliocolobus tephrosceles	XM_023230042.1	XM_023204001.1	XM_023190216.1	XM_023198126.2
Pipistrellus kuhlii	XM_045580498.1	XM_045586213.1	XM_036421639.2	XM_045570873.1
Pongo abelii	XM_024242617.2	XM_009251700.4	XM_054554034.1	XM_002814953.4
Pongo pygmaeus	XM_054475522.1	XM_054457700.2	XM_063663227.1	XM_054485500.1
Prionailurus bengalensis	XM_043604452.1	XM_043584600.1	XM_043554761.1	XM_043572215.1
Prionailurus viverrinus	XM_047852694.1	XM_047845844.1	XM_047857478.1	XM_047857545.1
Propithecus coquereli	XM_012653063.1	XM_012656643.1	XM_012659821.1	XM_012647195.1
Psammomys obesus	XM_055625998.1	XM_055619709.1	XM_055595496.1	XM_055594612.1
Pteronotus mesoamericanus	XM_054569196.1	XM_054593639.1	XM_054589662.1	XM_054591578.1
Pteropus alecto	XM_006910423.2	XM_006924563.3	XM_006917309.1	XM_006910709.1
Pteropus giganteus	XM_039887734.1	XM_039873350.1	XM_039845375.1	XM_039844972.1
Pteropus vampyrus	XM_011356625.2	XM_011381155.2	XM_011360611.1	XM_011356316.1
Puma concolor	XM_025927155.1	XM_025917666.1	XM_025922436.1	XM_025922001.1
Puma yagouaroundi	XM_040496157.1	XM_040497674.1	XM_040476458.1	XM_040478680.1
Rattus norvegicus	NM_019287.2	NM_012514.2	NM_031562.2	NM_001433761.1
Rattus rattus	XM_032909188.1	XM_032914456.1	XM_032917112.1	XM_032916531.1
Rhinolophus ferrumequinum	XM_033125573.1	XM_033091588.1	XM_033105748.1	XM_033105647.1
Rhinolophus sinicus	XM_019756811.1	XM_019747814.1	XM_019757077.1	XM_019731249.1
Rhinopithecus bieti	XM_017876834.1	XM_017867636.1	XM_017878968.1	XM_017866846.1
Rhinopithecus roxellana	XM_010383001.2	XM_030923110.1	XM_010383853.2	XM_010383120.2
Rousettus aegyptiacus	XM_016144749.2	XM_016150952.2	XM_016123358.1	XM_016161020.2
Saimiri boliviensis	XM_039480113.1	XM_039475995.1	XM_039468481.1	XM_003923991.1
Sapajus apella	XM_032293999.1	XM_032279508.1	XM_032259328.1	XM_032239663.1
Sciurus carolinensis	XM_047522404.1	XM_047543550.1	XM_047566032.1	XM_047565632.1
Sorex araneus	XM_055121351.1	XM_055133209.1	XM_012935729.2	XM_055140180.1
Sorex fumeus	XM_056126573.1	XM_056127887.1	–	XM_056111137.1
Sturnira hondurensis	XM_037035491.1	XM_037033501.1	–	XM_037046143.1
Suncus etruscus	XM_049784666.1	XM_049768089.1	–	XM_049789887.1
Suricata suricatta	XM_029938566.1	XM_029926858.1	XM_029951855.1	XM_029924702.1
Sus scrofa	NM_001375388.1	XM_005656956.3	NM_001004026.1	NM_001129953.2
Symphalangus syndactylus	XM_055252433.2	XM_055256857.2	XM_055297407.2	XM_055296200.2
Talpa occidentalis	XM_037528454.2	XM_037491518.2	XM_037530059.1	XM_037530345.1
Theropithecus gelada	XM_025354212.1	XM_025363311.1	XM_025386464.1	XM_025386295.1
Trachypithecus francoisi	XM_033176393.1	XM_033182851.1	XM_033215982.1	XM_033215971.1
Tursiops truncatus	XM_019943351.2	XM_019939445.2	XM_019928028.1	XM_019926721.2

Druh/puddruh	*APOB*	*BRCA1*	*CSN3*	*DMP1*
Urocitellus parryii	XM_026390830.1	XM_026386835.1	–	XM_026379640.1
Ursus americanus	XM_045813461.1	XM_045781762.1	XM_045789150.1	XM_045791574.1
Ursus arctos	XM_026519577.3	XM_057317881.1	XM_026510026.2	XM_026509605.3
Ursus maritimus	XM_040623722.1	XM_040640558.1	XM_008693175.1	XM_008706189.2
Vicugna pacos	XM_006197288.3	XM_006199189.3	XM_006212504.2	XM_031688384.1
Vulpes lagopus	XM_041754595.1	XM_041724533.1	XM_041724458.1	XM_041759926.1
Vulpes vulpes	XM_026014965.1	XM_025987104.1	XM_026003478.1	XM_025998064.1
Zalophus californianus	XM_027621822.2	XM_027624563.2	XM_027600186.1	XM_027597360.2

Druhá ze tří tabulek obsahujících přístupová čísla stažených sekvencí nukleárních genů z databáze NCBI.

Druh/puddruh	*GHR*	*LCAT*	*MYC*	*RAG1*
Acinonyx jubatus	XM_015067573.3	XM_015068818.3	XM_027063915.2	XM_015074391.3
Acomys russatus	XM_051172293.1	XM_051169304.1	XM_051159605.1	XM_051145002.1
Ailuropoda melanoleuca	NM_001304890.1	XM_011223955.3	XM_034668612.1	XM_002926498.4
Aotus nancymaae	XM_012433844.3	XM_012468171.2	XM_012451635.3	XM_012452362.2
Apodemus sylvaticus	XM_052159774.1	XM_052167139.1	XM_052161161.1	XM_052183073.1
Artibeus jamaicensis	XM_053657234.1	XM_037162578.2	XM_037166711.2	XM_037128294.2
Arvicanthis niloticus	XM_034523342.1	XM_034522655.1	XM_034517270.1	XM_034496296.1
Arvicola amphibius	XM_038321115.1	XM_038312436.2	XM_038342611.2	XM_038331945.1
Balaenoptera acutorostrata	XM_007195530.2	XM_007198151.3	XM_057532313.1	XM_007181092.2
Balaenoptera musculus	XM_036849227.1	XM_036832631.1	XM_036830572.1	XM_036862211.1
Balaenoptera ricei	XM_059916313.1	XM_059905094.1	XM_059901887.1	XM_059931402.1
Bison bison	XM_010832813.1	XM_010851007.1	XM_010844332.1	XM_010831868.1
Bos indicus	XM_019982774.1	XM_019978976.1	XM_019973984.1	XM_019974785.1
Bos mutus	XM_005889721.2	XM_005896529.1	XM_014479490.1	XM_005890914.1
Bos taurus	NM_176608.1	NM_001046069.2	NM_001046074.2	XM_024975902.2
Bubalus bubalis	NM_001290971.1	XM_006040996.4	NM_001290900.1	XM_006062340.3
Bubalus carabanensis	XM_055554857.1	XM_055554177.1	XM_055545790.1	XM_055549322.1
Budorcas taxicolor	XM_052659048.1	XM_052656671.1	XM_052651515.1	XM_052653310.1
Callithrix jacchus	XM_035291616.2	XM_002761078.6	XM_035276808.2	XM_009007944.3
Callorhinus ursinus	XM_025867338.1	XM_025888793.1	XM_025871112.1	XM_025878230.1
Camelus bactrianus	XM_010957409.2	XM_010962501.2	XM_010972639.2	XM_010964664.1
Camelus dromedarius	XM_064480357.1	XM_010979260.3	XM_010988785.3	XM_011000597.3
Camelus ferus	XM_032471550.1	XM_006180600.3	XM_032467638.1	XM_006195031.2
Canis lupus dingo	XM_035714888.2	XM_035715725.2	XM_025450851.3	XM_035702001.2
Canis lupus familiaris	NM_001003123.1	XM_038667269.1	NM_001003246.2	XM_038423502.1
Capra hircus	NM_001285648.1	XM_005692166.1	XM_018058563.1	XM_005690092.2
Carlito syrichta	XM_021709815.1	XM_021716387.1	XM_008061708.2	XM_008061343.1
Castor canadensis	XM_020181292.1	XM_020171127.1	XM_020184875.1	XM_020174920.1
Cavia porcellus	NM_001172706.1	XM_003472142.5	XM_003467354.5	XM_003463833.2
Cebus imitator	XM_017533142.2	XM_017539443.1	XM_017500491.2	XM_017530489.2
Cercocebus atys	XM_012055294.1	XM_012058934.1	XM_012065634.1	XM_012062276.1
Cervus canadensis	XM_043488954.1	XM_043437210.1	XM_043483835.1	XM_043480967.1
Cervus elaphus	XM_043887643.1	XM_043898864.1	XM_043880136.1	XM_043921764.1
Colobus angolensis	XM_011954465.1	XM_011932029.1	XM_011938183.1	XM_011940730.1
Condylura cristata	XM_004678346.2	XM_004690452.2	XM_004679594.2	XM_004682904.1
Cricetulus griseus	XM_016980790.2	XM_007644850.4	XM_007608000.4	XM_027421033.2
Delphinapterus leucas	XM_022598260.2	XM_022583416.1	XM_022554026.1	XM_022583604.1
Delphinus delphis	XM_060009529.2	XM_059999177.1	XM_059995191.1	XM_060017380.1
Desmodus rotundus	XM_045186428.2	XM_024556187.3	XM_024572335.3	XM_024558641.3
Diceros bicornis	XM_058564158.1	XM_058528076.1	XM_058564931.1	XM_058527657.1
Dipodomys ordii	XM_013018643.1	XM_013012870.1	XM_013025209.1	XM_013019614.1
Dipodomys spectabilis	XM_042690566.1	XM_042700764.1	XM_042673099.1	XM_042691791.1
Eptesicus fuscus	XM_054715288.1	XM_008139930.3	XM_054708494.1	XM_008146798.3
Equus asinus	XM_044779368.1	XM_014827680.2	XM_044781215.1	XM_014851022.2
Equus caballus	XM_023625698.1	XM_005608349.3	XM_001497991.5	NM_001256901.1
Equus przewalskii	XM_008524570.1	XM_008518213.1	XM_008544605.1	XM_008534235.1
Equus quagga	XM_046671901.1	XM_046682276.1	XM_046642603.1	XM_046648995.1
Erinaceus europaeus	XM_007517617.3	XM_016185844.2	XM_007530299.3	XM_007521831.1
Eumetopias jubatus	XM_028097994.1	XM_028092080.1	XM_028105649.1	XM_028111099.1
Felis catus	XM_045037560.1	XM_003998147.5	NM_001173446.1	XM_019812286.1

Druh/puddruh	*GHR*	*LCAT*	*MYC*	*RAG1*
Fukomys damarensis	XM_010633879.3	XM_010612286.3	XM_010642883.2	XM_010612551.1
Globicephala melas	XM_060295739.2	XM_030863596.2	XM_030875790.3	XM_030864942.2
Gorilla gorilla	XM_055367529.2	XM_019012012.3	XM_004047526.5	XM_055355540.2
Grammomys surdaster	XM_028784683.1	XM_028768495.1	XM_028753844.1	XM_028767126.1
Halichoerus grypus	XM_036067250.1	XM_036114995.1	XM_036084500.1	XM_036112420.1
Heterocephalus glaber	XM_021261444.1	XM_004842998.3	XM_004837959.3	XM_004852015.2
Hippopotamus amphibius	XM_057708790.1	XM_057711615.1	XM_057735449.1	XM_057695951.1
Hipposideros armiger	XM_019630525.1	XM_019655269.1	XM_019643424.1	XM_019639446.1
Homo sapiens	NM_000163.5	NM_000229.2	NM_002467.6	NM_000448.3
Hyaena hyaena	XM_039234142.1	XM_039223541.1	XM_039245733.1	XM_039233330.1
Hylobates moloch	XM_032135754.2	XM_032157693.2	XM_032760417.2	XM_058446441.1
Chinchilla lanigera	XM_013514982.1	XM_005403602.2	XM_005397989.2	XM_013514280.1
Chionomys nivalis	XM_057789429.1	XM_057753466.1	XM_057792039.1	XM_057781857.1
Chlorocebus sabaeus	XM_007961486.2	XM_007993901.2	XM_008001524.2	XM_037999512.1
Ictidomys tridecemlineatus	XM_005325612.4	XM_005318340.2	XM_013355904.3	XM_005327394.2
Jaculus jaculus	XM_045132800.1	XM_004661670.3	XM_004656161.3	XM_004651251.2
Kogia breviceps	XM_059060882.2	XM_059044888.2	XM_059043078.2	–
Lagenorhynchus albirostris	XM_060142903.1	XM_060130620.1	XM_060127612.1	XM_060157728.1
Lagenorhynchus obliquidens	XM_027101185.1	XM_027119521.1	XM_027132448.1	XM_027115691.1
Lemur catta	XM_045565741.1	XM_045533565.1	XM_045560838.1	XM_045557829.1
Leopardus geoffroyi	XM_045439095.1	XM_045443221.1	XM_045453689.1	XM_045485001.1
Leptonychotes weddellii	XM_031018413.1	XM_006730091.2	XM_031024489.1	XM_006736580.1
Lipotes vexillifer	XM_007447633.1	XM_007457258.1	XM_007456064.1	XM_007466847.1
Lontra canadensis	XM_032851402.1	XM_032854143.1	XM_032873540.1	XM_032841029.1
Lutra lutra	XM_047731501.1	XM_047710650.1	XM_047727577.1	XM_047693236.1
Lynx canadensis	XM_030330509.1	XM_030299658.1	XM_030304073.1	XM_030332658.1
Lynx rufus	XM_047073776.1	XM_047078615.1	XM_047074186.1	XM_047072365.1
Macaca fascicularis	XM_005556812.4	XM_005592313.3	XM_005564062.4	XM_065528314.1
Macaca mulatta	NM_001042667.1	XM_028841537.1	NM_001142873.1	NM_001266772.1
Macaca nemestrina	XM_011750908.2	XM_011757768.1	XM_011760043.2	XM_011724121.2
Macaca thibetana	XM_050793438.1	XM_050773294.1	XM_050800291.1	XM_050759433.1
Mandrillus leucophaeus	XM_011966443.1	XM_011996738.1	XM_011977280.1	XM_011973461.1
Marmota flaviventris	XM_027936309.1	XM_027953907.2	XM_027950100.1	XM_027936383.1
Marmota marmota	XM_048814019.1	XM_015496180.2	XM_048802407.1	XM_015476605.2
Marmota monax	XM_046468258.2	XM_046434398.2	XM_046464943.1	XM_046455203.1
Mastomys coucha	XM_031360110.1	XM_031341256.1	XM_031359117.1	XM_031373590.1
Meles meles	XM_045999829.1	XM_045987460.1	XM_046026180.1	XM_046014232.1
Meriones unguiculatus	XM_060380532.1	XM_021632266.2	XM_021655779.2	XM_021651897.2
Mesocricetus auratus	XM_040757557.1	XM_040736533.1	XM_005078685.4	XM_013120728.3
Mesoplodon densirostris	XM_060091665.1	XM_060083868.1	XM_060116080.1	XM_060104389.1
Microcebus murinus	XM_012777772.2	XM_012742453.2	XM_012785382.2	XM_012738734.1
Microtus fortis	XM_050139023.1	XM_050143781.1	XM_050145311.1	XM_050147173.1
Microtus ochrogaster	XM_005369809.2	XM_005345441.2	XM_005354556.3	XM_026787862.1
Microtus oregoni	–	XM_041650155.1	XM_041672632.1	XM_041679276.1
Miniopterus natalensis	XM_016223746.1	XM_016200752.1	XM_016215994.1	XM_016206146.1
Mirounga angustirostris	XM_045904014.3	XM_045890880.3	XM_045895887.3	XM_045892636.2
Mirounga leonina	–	XM_035024937.1	XM_035011930.1	XM_035012308.1
Molossus molossus	XM_036280505.1	XM_036274135.1	XM_036264004.1	XM_036260178.1
Monodon monoceros	XM_029216310.1	XM_029240192.1	XM_029223188.1	XM_029209720.1
Moschus berezovskii	XM_055422110.1	XM_055412492.1	XM_055432429.1	XM_055422704.1
Mus caroli	XM_021182903.2	XM_021169940.2	XM_021183649.1	XM_029474583.1
Mus musculus	NM_010284.3	NM_008490.2	NM_001177352.1	NM_009019.2
Mus pahari	XM_021208284.2	XM_021220722.2	XM_021216459.1	XM_021194803.1
Mustela erminea	XM_032337495.1	XM_032326354.1	XM_032316455.1	XM_032359447.1
Mustela lutreola	XM_059173896.1	XM_059153282.1	XM_059165626.1	XM_059140049.1
Mustela nigripes	XM_059417066.1	XM_059382710.1	XM_059395158.1	XM_059388381.1
Mustela putorius	XM_045079517.1	XM_004744484.3	XM_004743554.3	XM_004756002.3
Myodes glareolus	XM_048428399.1	XM_048437055.1	–	XM_048443556.1
Myotis brandtii	XM_014547516.1	XM_005876427.2	XM_005885505.2	XM_005872642.2
Myotis daubentonii	XM_059694845.1	XM_059667703.1	XM_059672284.1	XM_059711222.1
Myotis davidii	XM_006774555.2	XM_006769752.2	XM_006761209.2	XM_006756100.2
Myotis lucifugus	–	XM_006099455.3	XM_014459668.2	XM_023763518.1
Myotis myotis	XM_036309526.1	XM_036345785.1	XM_036337663.1	XM_036326159.1
Nannospalax galili	XM_029567818.1	XM_008839051.2	XM_008845471.3	XM_008852387.1
Neofelis nebulosa	XM_058718719.1	XM_058708193.1	XM_058699060.1	XM_058688682.1
Neomonachus schauinslandi	XM_021700587.2	XM_021705746.2	XM_021688451.1	XM_021684748.2
Neophocaena asiaeorientalis	XM_024740332.1	XM_024732430.1	XM_024741267.1	XM_024752616.1

Druh/puddruh	*GHR*	*LCAT*	*MYC*	*RAG1*
Neovison vison	XM_044232377.1	XM_044255897.1	XM_044244949.1	XM_044260941.1
Nomascus leucogenys	XM_004088209.3	XM_012504716.2	XM_003256203.3	XM_030829431.1
Nyctereutes procyonoides	XM_055325044.1	XM_055305910.1	XM_055314420.1	XM_055300734.1
Nycticebus coucang	XM_053591462.1	XM_053604499.1	XM_053559296.1	XM_053560451.1
Octodon degus	XM_023701876.1	XM_004625843.2	XM_004634419.2	XM_004627038.1
Odobenus rosmarus	XM_004409145.2	XM_004393361.2	XM_012566191.1	XM_004409327.1
Odocoileus virginianus	XM_020888240.1	XM_020885063.1	XM_020875260.1	XM_020904965.1
Onychomys torridus	XM_036207314.1	XM_036187084.1	XM_036208372.1	XM_036183199.1
Orcinus orca	XM_033421189.2	XM_004273184.4	XM_004265321.4	XM_004264017.3
Oryx dammah	XM_040262428.1	XM_040253332.1	XM_040261783.1	XM_040259019.1
Otolemur garnettii	XM_023514997.1	XM_003799375.3	XM_012810707.2	XM_023509304.1
Ovis aries	NM_001009323.2	NM_001162568.1	NM_001009426.1	XM_004016411.6
Pan paniscus	XM_063604482.1	XM_008964568.3	XM_003820455.6	XM_008955833.7
Pan troglodytes	XM_024356621.3	XM_063797282.1	NM_001142794.1	XM_001154240.8
Panthera leo	XM_042950820.1	XM_042920939.1	XM_042924185.1	XM_042907087.1
Panthera pardus	XM_053906960.1	XM_019455520.2	XM_019459856.2	XM_019470137.2
Panthera tigris	XM_007095429.2	XM_007090320.3	XM_042973484.1	XM_007082060.2
Panthera uncia	XM_049648698.1	XM_049622408.1	XM_049632908.1	XM_049647885.1
Papio anubis	NM_001112616.1	NM_001112613.1	XM_021942613.2	XM_003910017.5
Perognathus longimembris	XM_048368191.1	XM_048355285.1	XM_048358921.1	XM_048360909.1
Peromyscus californicus	XM_052743766.1	XM_052729808.1	XM_052747887.1	XM_052724256.1
Peromyscus eremicus	XM_059276751.1	XM_059264096.1	XM_059247356.1	XM_059258907.1
Peromyscus leucopus	XM_037209416.1	XM_028854295.1	XM_028882207.2	XM_028895036.2
Peromyscus maniculatus	XM_042261499.1	XM_006986833.3	XM_006989007.2	–
Phacochoerus africanus	XM_047766341.1	XM_047788708.1	XM_047783256.1	XM_047774265.1
Phoca vitulina	XM_032415196.1	XM_032424130.1	XM_032416053.1	XM_032414458.1
Phocoena sinus	XM_032628299.1	XM_032613851.1	XM_032610261.1	XM_032641375.1
Phodopus roborovskii	XM_051182481.1	XM_051201656.1	XM_051202448.1	XM_051184108.1
Phyllostomus discolor	XM_028529502.2	XM_036012042.1	XM_036031420.1	XM_028517512.2
Phyllostomus hastatus	XM_045847874.1	XM_045818536.1	XM_045861509.1	XM_045825446.1
Physeter catodon	XM_028492884.2	XM_028478057.2	XM_007103917.4	XM_007106867.2
Piliocolobus tephrosceles	XM_023216560.1	XM_023210222.1	XM_023224047.3	XM_023185545.2
Pipistrellus kuhlii	XM_036409450.2	XM_036441443.2	XM_036437564.2	XM_036427844.2
Pongo abelii	XM_054555527.2	XM_009250870.4	XM_009244088.4	XM_054524487.2
Pongo pygmaeus	XM_054488654.2	XM_054454020.2	XM_054499106.2	XM_063670983.1
Prionailurus bengalensis	XM_043599136.1	XM_043600882.1	XM_043601753.1	XM_043580977.1
Prionailurus viverrinus	XM_047866660.1	XM_047836513.1	XM_047842871.1	XM_047879247.1
Propithecus coquereli	XM_012664727.1	XM_012660562.1	XM_012645505.1	XM_012649576.1
Psammomys obesus	XM_055601047.1	XM_055620771.1	XM_055626783.1	XM_055604957.1
Pteronotus mesoamericanus	XM_054583980.1	XM_054579300.1	XM_054590093.1	XM_054586223.1
Pteropus alecto	XM_025052160.1	XM_006909787.3	XM_006919843.3	XM_006908040.1
Pteropus giganteus	XM_039842558.1	XM_039851835.1	XM_039853540.1	XM_039864987.1
Pteropus vampyrus	XM_011370701.2	XM_011380523.2	XM_011360819.2	XM_011356837.1
Puma concolor	XM_025924136.1	XM_025920495.1	XM_025922761.1	XM_025928993.1
Puma yagouaroundi	XM_040496262.1	XM_040498425.1	XM_040494416.1	XM_040495424.1
Rattus norvegicus	NM_017094.2	NM_017024.3	NM_012603.2	NM_053468.2
Rattus rattus	XM_032898830.1	XM_032887689.1	XM_032915922.1	XM_032902243.1
Rhinolophus ferrumequinum	XM_033110150.1	XM_033128112.1	XM_033126728.1	XM_033118988.1
Rhinolophus sinicus	XM_019748545.1	XM_019756359.1	XM_019723689.1	XM_019738494.1
Rhinopithecus bieti	XM_017867112.1	XM_017852258.1	XM_017870746.1	XM_017881640.1
Rhinopithecus roxellana	XM_030927287.1	XM_030924629.1	XM_010379223.2	XM_010358405.2
Rousettus aegyptiacus	XM_036226086.1	XM_016164180.2	XM_016136818.2	XM_016123252.2
Saimiri boliviensis	NM_001302931.1	XM_010346808.2	XM_010344242.2	XM_010333497.2
Sapajus apella	XM_032278845.1	XM_032283790.1	XM_032253356.1	XM_032282308.1
Sciurus carolinensis	XM_047556984.1	XM_047529613.1	XM_047526957.1	XM_047515809.1
Sorex araneus	XM_055118179.1	XM_055146121.1	XM_055127392.1	XM_004607911.2
Sorex fumeus	XM_056112075.1	XM_056136354.1	XM_056111994.1	XM_056106567.1
Sturnira hondurensis	XM_037063239.1	XM_037041459.1	XM_037048537.1	XM_037040547.1
Suncus etruscus	XM_049768241.1	XM_049786410.1	XM_049765620.1	–
Suricata suricatta	XM_029940942.1	XM_029924506.1	XM_029923831.1	XM_029956008.1
Sus scrofa	NM_214254.2	NM_001164856.1	NM_001005154.1	NM_001123184.1
Symphalangus syndactylus	XM_063619526.1	XM_063611774.1	XM_055287240.2	XM_063640976.1
Talpa occidentalis	XM_037530378.2	XM_037515944.2	XM_037501810.2	XM_037521628.1
Theropithecus gelada	XM_025387793.1	XM_025370247.1	XM_025394466.1	XM_025357173.1
Trachypithecus francoisi	XM_033179464.1	XM_033219749.1	XM_033202355.1	XM_033205039.1
Tursiops truncatus	XM_033852830.1	XM_033845310.1	XM_004327312.3	XM_004329819.3
Urocitellus parryii	XM_026415485.1	XM_026413312.1	XM_026407393.1	XM_026405176.1

Druh/puddruh	*GHR*	*LCAT*	*MYC*	*RAG1*
Ursus americanus	XM_045787355.1	XM_045814904.1	XM_045812234.1	XM_045808676.1
Ursus arctos	XM_026506893.4	XM_026495049.4	XM_026495412.4	XM_026511863.4
Ursus maritimus	XM_008692585.1	XM_040633472.1	XM_040625987.1	XM_008686790.2
Vicugna pacos	XM_015242106.2	XM_006203718.3	XM_006208010.3	XM_006216856.1
Vulpes lagopus	XM_041747919.1	XM_041769867.1	XM_041769617.1	XM_041773539.1
Vulpes vulpes	XM_026016684.1	XM_026011536.1	XM_025993448.1	XM_026012934.1
Zalophus californianus	XM_027604919.2	XM_027619221.1	XM_027574291.2	XM_027581138.1

Třetí ze tří tabulek obsahujících přístupová čísla stažených sekvencí nukleárních genů z databáze NCBI.

Druh/puddruh	*RAG2*	*RBP3*	*TP53*	*VWF*
Acinonyx jubatus	XM_015074392.3	XM_015088234.3	XM_027047532.2	XM_027074097.2
Acomys russatus	XM_051145001.1	XM_051141867.1	XM_051167465.1	XM_051155405.1
Ailuropoda melanoleuca	XM_034645409.1	XM_002930073.3	XM_002924437.4	XM_034645977.1
Aotus nancymaae	XM_012452363.2	XM_012472634.2	XM_064378361.1	XM_064377521.1
Apodemus sylvaticus	XM_052183071.1	XM_052190665.1	XM_052195723.1	XM_052174915.1
Artibeus jamaicensis	XM_037128300.2	XM_037158456.2	XM_037151170.2	XM_037160260.2
Arvicanthis niloticus	XM_034495241.1	XM_034497486.1	XM_034505539.1	XM_034512425.1
Arvicola amphibius	XM_038332149.1	XM_038349866.1	XM_038328731.1	XM_038320478.1
Balaenoptera acutorostrata	XM_007181094.2	XM_007178750.2	XM_007166385.2	XM_028163508.2
Balaenoptera musculus	XM_036859204.1	XM_036829569.1	XM_036836670.1	XM_036865018.1
Balaenoptera ricei	XM_059929648.1	XM_059900505.1	XM_059909155.1	XM_059935147.1
Bison bison	XM_010831876.1	XM_010854512.1	XM_010843072.1	XM_010845201.1
Bos indicus	XM_019975675.1	XM_019953782.1	XM_019982111.1	XM_019959953.1
Bos mutus	XM_005890916.1	XM_005888669.1	XM_005894802.2	XM_014478162.1
Bos taurus	NM_001082469.2	NM_174164.1	NM_174201.2	NM_001205308.2
Bubalus bubalis	XM_006062335.4	XM_025284888.2	NM_001290844.2	XM_006065776.4
Bubalus carabanensis	XM_055549324.1	XM_055566447.1	XM_055576635.1	XM_055568893.1
Budorcas taxicolor	XM_052653374.1	XM_052641527.1	XM_052657549.1	XM_052641007.1
Callithrix jacchus	XM_035262383.1	XM_002756138.4	XM_002747948.5	XM_002752243.6
Callorhinus ursinus	XM_025878267.1	XM_025880483.1	XM_025864014.1	XM_025851854.1
Camelus bactrianus	XM_045523815.1	XM_010963476.1	XM_010965924.2	XM_010957747.2
Camelus dromedarius	XM_031459781.2	XM_010998716.3	XM_010996514.3	XM_031444271.2
Camelus ferus	XM_032488481.1	XM_006193816.2	XM_006175816.3	XM_032473470.1
Canis lupus dingo	XM_025452502.3	XM_025435444.2	XM_025428522.3	XM_035707072.2
Canis lupus familiaris	XM_038424077.1	XM_038662368.1	NM_001389218.1	NM_001002932.1
Capra hircus	XM_005690095.2	XM_005699333.2	XM_005693530.3	XM_018048607.1
Carlito syrichta	XM_008061345.2	XM_008059145.1	XM_008062341.2	XM_021717812.1
Castor canadensis	XM_020174919.1	XM_020153863.1	XM_020151705.1	XM_020158236.1
Cavia porcellus	XM_023562442.2	XM_003466069.3	NM_001172740.1	NM_001278774.1
Cebus imitator	XM_017530482.2	XM_017504716.2	XM_017531594.2	XM_017546164.1
Cercocebus atys	XM_012062292.1	XM_012072157.1	XM_012052423.1	XM_012054264.1
Cervus canadensis	XM_043481734.1	XM_043475035.1	XM_043475143.1	XM_043440090.1
Cervus elaphus	XM_043909547.1	XM_043924788.1	XM_043903734.1	XM_043880888.1
Colobus angolensis	XM_011940740.1	XM_011950325.1	XM_011958558.1	XM_011955306.1
Condylura cristata	XM_004682903.1	–	XM_004684639.2	–
Cricetulus griseus	XM_027421034.2	XM_027389336.1	NM_001243976.1	XM_027428330.2
Delphinapterus leucas	XM_022583600.2	XM_022564527.1	XM_022595620.2	XM_022577358.2
Delphinus delphis	XM_060017382.1	XM_060033692.1	XM_059998561.1	XM_060024146.1
Desmodus rotundus	XM_045194182.2	XM_024561056.3	–	XM_024579257.3
Diceros bicornis	XM_058527656.1	XM_058542753.1	XM_058559811.1	XM_058559008.1
Dipodomys ordii	XM_013019612.1	XM_013036670.1	XM_013013490.1	XM_013023596.1
Dipodomys spectabilis	XM_042691748.1	XM_042690365.1	XM_042670866.1	XM_042673279.1
Eptesicus fuscus	XM_028153614.2	XM_054729327.1	XM_028128361.2	XM_008144811.3
Equus asinus	XM_014851016.2	XM_014841499.2	XM_014861557.1	XM_044755819.1
Equus caballus	XM_014729427.2	XM_001494823.2	NM_001202405.1	NM_001242566.1
Equus przewalskii	XM_008534232.1	XM_008530266.1	XM_008530347.1	XM_008513920.1
Equus quagga	XM_046648996.1	XM_046652801.1	XM_046675439.1	XM_046661799.1
Erinaceus europaeus	XM_007521832.1	XM_007534910.2	XM_016188237.2	XM_060190227.1
Eumetopias jubatus	XM_028111122.1	XM_028098670.1	XM_028124935.1	XM_028121141.1
Felis catus	XM_004001424.6	XM_003994154.5	NM_001009294.1	NM_001246279.1
Fukomys damarensis	XM_010612550.1	XM_010622627.1	XM_010614891.3	XM_033765953.1
Globicephala melas	XM_030864652.2	XM_030847670.2	XM_060290532.2	XM_060306367.2

Druh/puddruh	*RAG2*	*RBP3*	*TP53*	*VWF*
Gorilla gorilla	XM_004050960.5	XM_004049370.5	XM_063700589.1	XM_031016658.2
Grammomys surdaster	XM_028767018.1	XM_028789145.1	XM_028780049.1	XM_028754414.1
Halichoerus grypus	XM_036112385.1	XM_036103178.1	XM_036077514.1	XM_036117632.1
Heterocephalus glaber	XM_004852014.1	–	NM_001310270.1	XM_021248148.1
Hippopotamus amphibius	XM_057695384.1	XM_057736957.1	XM_057716067.1	XM_057702122.1
Hipposideros armiger	XM_019639435.1	XM_019660466.1	XM_019664961.1	XM_019627083.1
Homo sapiens	NM_000536.4	NM_002900.3	NM_001276760.3	NM_000552.5
Hyaena hyaena	XM_039233329.1	XM_039233648.1	XM_039231497.1	XM_039218403.1
Hylobates moloch	XM_032166243.1	XM_032141830.2	XM_032139337.3	XM_032159366.2
Chinchilla lanigera	XM_005384100.1	XM_005412613.2	XM_005399463.2	XM_005378754.1
Chionomys nivalis	XM_057781920.1	XM_057770009.1	XM_057776262.1	XM_057772108.1
Chlorocebus sabaeus	XM_008001615.2	XM_007962654.2	XM_008010192.2	XM_037988386.1
Ictidomys tridecemlineatus	XM_005327392.3	XM_005329243.3	XM_005332819.4	XM_005338364.3
Jaculus jaculus	–	XM_004658032.2	XM_045131539.1	XM_045139643.1
Kogia breviceps	XM_059068211.1	XM_059053083.1	XM_059048966.2	XM_059082366.2
Lagenorhynchus albirostris	XM_060157727.1	XM_060126763.1	XM_060132840.1	XM_060167140.1
Lagenorhynchus obliquidens	XM_027114962.1	XM_027098528.1	XM_027130927.1	XM_027093340.1
Lemur catta	XM_045557830.1	XM_045567963.1	XM_045570131.1	XM_045554544.1
Leopardus geoffroyi	XM_045485002.1	XM_045437949.1	XM_045490279.1	XM_045466410.1
Leptonychotes weddellii	XM_006736581.2	XM_006749024.1	XM_006738887.2	XM_006735969.2
Lipotes vexillifer	XM_007466849.1	XM_007453444.1	XM_007457790.1	XM_007469511.1
Lontra canadensis	XM_032840994.1	XM_032842973.1	XM_032840028.1	XM_032874808.1
Lutra lutra	XM_047693199.1	XM_047701618.1	XM_047706893.1	XM_047739853.1
Lynx canadensis	XM_030330747.1	XM_030335579.1	XM_030295516.1	XM_030321487.1
Lynx rufus	XM_047070189.1	XM_047072624.1	XM_047089898.1	XM_047066718.1
Macaca fascicularis	XM_005578160.4	XM_005565203.4	XM_005582787.4	XM_005569887.4
Macaca mulatta	XM_001114748.4	XM_015146957.2	NM_001047151.2	NM_001243086.1
Macaca nemestrina	XM_011724124.2	XM_011749690.2	XM_011727613.2	XM_024793545.1
Macaca thibetana	XM_050759434.1	XM_050803183.1	XM_050765346.1	XM_050747383.1
Mandrillus leucophaeus	XM_011973449.1	XM_011990910.1	XM_011994540.1	XM_011982932.1
Marmota flaviventris	XM_027936606.1	XM_027947858.2	XM_027946781.2	XM_027951103.2
Marmota marmota	XM_015477789.2	XM_015498919.2	XM_015504433.2	XM_015500732.2
Marmota monax	XM_046455190.1	XM_046446777.2	XM_046422140.2	XM_046424111.2
Mastomys coucha	XM_031371132.1	XM_031362471.1	XM_031351448.1	XM_031383471.1
Meles meles	XM_046014145.1	XM_046026592.1	XM_045985156.1	XM_046011417.1
Meriones unguiculatus	XM_021651899.2	XM_021628200.2	XM_021648173.2	XM_021636960.2
Mesocricetus auratus	XM_021231573.2	XM_005077637.4	NM_001281661.1	XM_040737568.1
Mesoplodon densirostris	XM_060105316.1	XM_060079495.1	XM_060081214.1	XM_060111711.1
Microcebus murinus	XM_012738727.1	XM_020280499.1	XM_012776058.2	XM_012760580.1
Microtus fortis	XM_050146886.1	XM_050140662.1	XM_050162534.1	XM_050164057.1
Microtus ochrogaster	XM_026787879.1	XM_005348626.2	XM_005349777.1	XM_005365251.2
Microtus oregoni	XM_041678681.1	XM_041662926.1	XM_041664596.1	XM_041641912.1
Miniopterus natalensis	XM_016206233.1	XM_016200835.1	XM_016222486.1	XM_016221858.1
Mirounga angustirostris	XM_045870992.1	XM_045879199.2	XM_064583202.1	XM_064591672.1
Mirounga leonina	XM_035012309.1	XM_034987114.1	XM_034989514.1	XM_035028418.1
Molossus molossus	XM_036260069.1	XM_036265478.1	XM_036269156.1	XM_036250841.1
Monodon monoceros	XM_029209310.1	XM_029245877.1	XM_029212409.1	XM_029204363.1
Moschus berezovskii	XM_055422705.1	XM_055424284.1	XM_055406049.1	XM_055394198.1
Mus caroli	XM_021186290.2	XM_021182453.1	XM_021177210.2	XM_029478271.1
Mus musculus	NM_009020.4	NM_015745.3	NM_011640.4	NM_011708.4
Mus pahari	XM_021194637.1	XM_021204213.1	XM_021212574.2	XM_029534982.1
Mustela erminea	XM_032358597.1	XM_032312199.1	XM_032322544.1	XM_032349032.1
Mustela lutreola	XM_059158517.1	XM_059172492.1	XM_059147788.1	XM_059184573.1
Mustela nigripes	XM_059388379.1	XM_059397691.1	XM_059379539.1	XM_059403765.1
Mustela putorius	XM_004756007.3	XM_004765887.1	XM_004760256.3	XM_004766734.3
Myodes glareolus	XM_048443601.1	XM_048421458.1	XM_048448387.1	XM_048444099.1
Myotis brandtii	XM_014544903.1	XM_014535671.1	XM_014549841.1	XM_005873725.2
Myotis daubentonii	XM_059711220.1	XM_059661966.1	XM_059668742.1	XM_059681979.1
Myotis davidii	XM_006756099.2	XM_015559657.1	XM_015562573.1	XM_006757606.2
Myotis lucifugus	XM_006099888.3	XM_023746257.1	XM_023743933.1	XM_006084318.3
Myotis myotis	XM_036326157.1	XM_036355902.1	XM_036341447.1	XM_036298336.1
Nannospalax galili	XM_008852390.2	XM_008834127.1	NM_001319288.1	XM_008825438.2
Neofelis nebulosa	XM_058688683.1	XM_058696943.1	XM_058702775.1	XM_058743945.1
Neomonachus schauinslandi	XM_021684749.1	XM_021699962.1	XM_021694492.1	XM_044915225.1
Neophocaena asiaeorientalis	XM_024752527.1	XM_024739109.1	XM_024748420.1	XM_024733215.1
Neovison vison	XM_044260942.1	XM_044236829.1	XM_044249002.1	XM_044228748.1
Nomascus leucogenys	XM_003254438.2	XM_003278171.2	XM_030800485.1	XM_030804480.1

Druh/puddruh	*RAG2*	*RBP3*	*TP53*	*VWF*
Nyctereutes procyonoides	XM_055301358.1	XM_055324854.1	XM_055305539.1	XM_055315447.1
Nycticebus coucang	XM_053560454.1	XM_053583486.1	XM_053569948.1	XM_053554967.1
Octodon degus	XM_004627037.1	XM_004641242.1	XM_023723301.1	XM_004643560.2
Odobenus rosmarus	XM_004409325.1	XM_004416717.1	XM_004398491.1	XM_004413880.2
Odocoileus virginianus	XM_020904966.1	XM_020895520.1	XM_020869027.1	XM_020912316.1
Onychomys torridus	XM_036183337.1	XM_036200065.1	XM_036197270.1	XM_036181447.1
Orcinus orca	XM_004264018.3	XM_004286499.3	XM_033422971.2	XM_004279025.3
Oryx dammah	XM_040259021.1	XM_040242143.1	XM_040243464.1	XM_040234946.1
Otolemur garnettii	XM_023509313.1	XM_003801643.3	XM_012806041.2	XM_003796416.3
Ovis aries	XM_060400061.1	XM_004021549.5	NM_001009403.1	XM_027967983.2
Pan paniscus	XM_003830353.4	XM_003809049.5	XM_057300362.2	XM_003820327.5
Pan troglodytes	XM_054662211.2	XM_507774.7	XM_016931470.3	XM_001160508.8
Panthera leo	XM_042906368.1	XM_042909008.1	XM_042915570.1	XM_042945594.1
Panthera pardus	XM_019470138.2	XM_019444852.2	XM_019413568.2	XM_053895673.1
Panthera tigris	XM_007082061.2	XM_007093164.3	XM_042966154.1	XM_007089645.3
Panthera uncia	XM_049647899.1	XM_049646270.1	XM_049637036.1	XM_049624950.1
Papio anubis	XM_003910018.5	XM_003903645.4	XM_003912272.5	XM_021921932.2
Perognathus longimembris	XM_048360255.1	XM_048336003.1	XM_048366174.1	–
Peromyscus californicus	XM_052722445.1	XM_052733940.1	XM_052756670.1	XM_052720566.1
Peromyscus eremicus	–	XM_059274578.1	XM_059270972.1	XM_059258637.1
Peromyscus leucopus	XM_028895024.2	XM_028878306.2	XM_028869449.2	XM_028892286.2
Peromyscus maniculatus	XM_006983930.2	XM_006976708.2	XM_006973448.3	XM_042274144.1
Phacochoerus africanus	XM_047775956.1	XM_047760709.1	XM_047756901.1	XM_047787770.1
Phoca vitulina	XM_032413813.1	XM_032399063.1	XM_032431962.1	XM_032390769.1
Phocoena sinus	XM_032641667.1	XM_032608931.1	XM_032616755.1	XM_032646289.1
Phodopus roborovskii	XM_051184109.1	XM_051202498.1	XM_051180252.1	XM_051173265.1
Phyllostomus discolor	XM_028517065.2	XM_028512741.2	XM_028534686.2	XM_028532222.2
Phyllostomus hastatus	XM_045825819.1	XM_045841799.1	XM_045824646.1	XM_045826343.1
Physeter catodon	XM_007106862.2	XM_007118754.2	XM_028499408.1	XM_028490937.2
Piliocolobus tephrosceles	XM_023185546.2	XM_023230192.3	XM_023193314.2	XM_031936259.1
Pipistrellus kuhlii	XM_036427509.2	XM_045570640.1	XM_045586558.1	XM_036455712.2
Pongo abelii	XM_002821867.3	XM_002820699.6	XM_002826974.6	NM_001246277.1
Pongo pygmaeus	XM_054440959.2	XM_054436368.1	XM_054457336.2	XM_054442726.1
Prionailurus bengalensis	XM_043580978.1	XM_043596802.1	XM_043582969.1	XM_043563566.1
Prionailurus viverrinus	XM_047824195.1	XM_047826277.1	XM_047832341.1	XM_047866328.1
Propithecus coquereli	XM_012649573.1	XM_012651268.1	XM_012655616.1	XM_012637894.1
Psammomys obesus	XM_055605184.1	XM_055612209.1	XM_055594450.1	XM_055609420.1
Pteronotus mesoamericanus	XM_054586222.1	XM_054572396.1	XM_054575722.1	XM_054574700.1
Pteropus alecto	XM_006908043.2	XM_006916945.2	XM_006922335.3	XM_025048649.1
Pteropus giganteus	XM_039867247.1	XM_039870918.1	XM_039847238.1	XM_039874111.1
Pteropus vampyrus	–	XM_011355459.1	XM_011358864.2	XM_023533934.1
Puma concolor	XM_025929022.1	XM_025918239.1	XM_025920288.1	XM_025930120.1
Puma yagouaroundi	XM_040495416.1	XM_040485664.1	XM_040453335.1	XM_040462716.1
Rattus norvegicus	NM_001100528.2	NM_001191832.1	NM_030989.4	NM_053889.1
Rattus rattus	XM_032902242.1	XM_032919064.1	XM_032912336.1	XM_032906594.1
Rhinolophus ferrumequinum	XM_033121417.1	XM_033129982.1	XM_033091194.1	XM_033116665.1
Rhinolophus sinicus	XM_019738609.1	XM_019718074.1	XM_019745191.1	XM_019754370.1
Rhinopithecus bieti	XM_017881641.1	XM_017890984.1	XM_017850544.1	XM_017883319.1
Rhinopithecus roxellana	XM_010358412.2	XM_030941098.1	XM_010362387.1	XM_030939152.1
Rousettus aegyptiacus	XM_016123246.2	XM_016142777.2	XM_016154772.2	XM_016162761.2
Saimiri boliviensis	XM_003920003.2	XM_003929968.2	XM_039476036.1	XM_010330022.2
Sapajus apella	XM_032282299.1	XM_032266806.1	XM_032280430.1	XM_032254536.1
Sciurus carolinensis	XM_047515670.1	XM_047554526.1	XM_047545682.1	XM_047549520.1
Sorex araneus	XM_004607912.2	XM_004607548.2	XM_004604858.2	XM_055143591.1
Sorex fumeus	XM_056135751.1	XM_056130515.1	XM_056114885.1	XM_056108652.1
Sturnira hondurensis	XM_037040548.1	XM_037056658.1	XM_037028382.1	XM_037067832.1
Suncus etruscus	XM_049779661.1	–	XM_049766230.1	XM_049771928.1
Suricata suricatta	XM_029915804.1	XM_029915810.1	XM_029927154.1	XM_029953391.1
Sus scrofa	NM_001128481.1	XM_021072759.1	NM_213824.3	NM_001246221.1
Symphalangus syndactylus	XM_055281887.2	XM_055275424.1	XM_055256596.2	XM_055280967.2
Talpa occidentalis	XM_037522441.1	–	XM_037492036.2	XM_037514506.2
Theropithecus gelada	XM_025358178.1	XM_025396981.1	XM_025363515.1	XM_025401750.1
Trachypithecus francoisi	XM_033205037.1	XM_033209108.1	XM_033184717.1	XM_033223122.1
Tursiops truncatus	XM_019948364.2	XM_004320010.3	XM_033848569.1	XM_033867155.1
Urocitellus parryii	XM_026404849.1	XM_026411988.1	XM_026390537.1	XM_026403302.1
Ursus americanus	XM_045808683.1	XM_045809980.1	XM_045797340.1	XM_045781154.1
Ursus arctos	XM_026511862.2	XM_026481508.3	XM_026520889.4	XM_057303331.1

Druh/puddruh	***RAG2***	***RBP3***	***TP53***	***VWF***
Ursus maritimus	XM_008686789.1	XM_008696329.2	XM_008687972.2	XM_040642060.1
Vicugna pacos	XM_031681026.1	XM_006216104.2	XM_006218777.3	XM_031670534.1
Vulpes lagopus	XM_041773423.1	XM_041748483.1	XM_041771191.1	XM_041735719.1
Vulpes vulpes	XM_026012750.1	XM_026009362.1	XM_026005407.1	XM_025995567.1
Zalophus californianus	XM_027581310.1	XM_027590675.2	XM_027625914.2	XM_027595203.2

Příloha 3

Výsledky kladového modelu C a modelu M2a_rel

Následující tabulka obsahuje kompletní výsledky modelů cmC a M2a_rel. Nulový model M2a_rel nemá uvedený taxon („–“).

Gen	Taxon	LL	p_0	ω_0	p_1	ω_1	p_2	ω_2	ω_3	$\Delta\omega$	LRT	p-hodnota
APOB	–	-33885,784	0,363	0,110	0,179	1	0,458	0,448				
APOB	Artiodactyla	-33885,482	0,363	0,110	0,179	1	0,458	0,444	0,484	0,039	0,603	0,437
APOB	Caniformia	-33885,747	0,362	0,110	0,180	1	0,458	0,449	0,423	-0,026	0,073	0,787
APOB	Cetacea	-33885,684	0,361	0,110	0,180	1	0,459	0,446	0,494	0,048	0,199	0,656
APOB	Eulipotyphla	-33885,230	0,358	0,109	0,182	1	0,460	0,439	0,493	0,054	1,108	0,293
APOB	Feliformia	-33885,654	0,363	0,110	0,179	1	0,458	0,447	0,503	0,056	0,258	0,611
APOB	Chiroptera	-33880,775	0,365	0,111	0,171	1	0,463	0,428	0,586	0,158	10,018	0,002
APOB	Perissodactyla	-33882,709	0,370	0,112	0,172	1	0,458	0,450	0,809	0,359	6,148	0,013
APOB	Primates	-33885,744	0,361	0,110	0,180	1	0,459	0,446	0,460	0,014	0,078	0,780
APOB	Rodentia	-33873,844	0,362	0,110	0,172	1	0,466	0,525	0,355	-0,170	23,880	0,000
BRCA1	–	-93376,976	0,165	0,149	0,362	1	0,473	0,467				
BRCA1	Artiodactyla	-93376,869	0,165	0,149	0,362	1	0,473	0,466	0,482	0,016	0,215	0,643
BRCA1	Caniformia	-93372,566	0,163	0,148	0,363	1	0,473	0,475	0,364	-0,111	8,819	0,003
BRCA1	Cetacea	-93373,241	0,164	0,148	0,362	1	0,474	0,463	0,678	0,216	7,470	0,006
BRCA1	Eulipotyphla	-93375,578	0,165	0,149	0,362	1	0,473	0,471	0,392	-0,079	2,795	0,095
BRCA1	Feliformia	-93375,421	0,165	0,149	0,361	1	0,474	0,471	0,385	-0,087	3,110	0,078
BRCA1	Chiroptera	-93376,813	0,165	0,149	0,362	1	0,473	0,470	0,453	-0,017	0,326	0,568
BRCA1	Perissodactyla	-93375,468	0,162	0,147	0,364	1	0,473	0,468	0,354	-0,113	3,016	0,082
BRCA1	Primates	-93376,703	0,165	0,148	0,363	1	0,473	0,469	0,444	-0,025	0,546	0,460
BRCA1	Rodentia	-93365,450	0,162	0,148	0,367	1	0,471	0,414	0,522	0,109	23,052	0,000
COX1	–	-77400,123	0,871	0,002	0,002	1	0,127	0,041				
COX1	Artiodactyla	-77397,387	0,869	0,002	0,002	1	0,129	0,038	0,049	0,011	5,471	0,019
COX1	Caniformia	-77398,761	0,871	0,002	0,002	1	0,127	0,039	0,048	0,008	2,723	0,099
COX1	Cetacea	-77381,506	0,861	0,002	0,002	1	0,137	0,035	0,082	0,047	37,233	0,000
COX1	Eulipotyphla	-77391,299	0,873	0,002	0,002	1	0,125	0,042	0,020	-0,022	17,647	0,000
COX1	Feliformia	-77400,122	0,871	0,002	0,002	1	0,127	0,041	0,041	0,000	0,001	0,977
COX1	Chiroptera	-77380,705	0,847	0,001	0,004	1	0,149	0,037	0,016	-0,021	38,835	0,000
COX1	Perissodactyla	-77398,290	0,125	0,041	0,002	1	0,873	0,002	0,000	-0,001	3,665	0,056
COX1	Primates	-77340,712	0,838	0,001	0,004	1	0,158	0,026	0,072	0,046	118,821	0,000
COX1	Rodentia	-77388,871	0,861	0,002	0,002	1	0,137	0,043	0,029	-0,014	22,504	0,000
COX2	–	-35823,622	0,751	0,003	0,000	1	0,249	0,045				
COX2	Artiodactyla	-35818,550	0,749	0,003	0,000	1	0,251	0,042	0,060	0,018	10,143	0,001
COX2	Caniformia	-35823,356	0,752	0,003	0,000	1	0,248	0,045	0,049	0,005	0,531	0,466
COX2	Cetacea	-35802,199	0,739	0,003	0,000	1	0,261	0,039	0,109	0,070	42,846	0,000
COX2	Eulipotyphla	-35818,905	0,752	0,003	0,000	1	0,248	0,047	0,027	-0,020	9,434	0,002
COX2	Feliformia	-35823,116	0,751	0,003	0,000	1	0,249	0,045	0,054	0,009	1,010	0,315
COX2	Chiroptera	-35823,545	0,751	0,003	0,000	1	0,249	0,045	0,048	0,002	0,153	0,696
COX2	Perissodactyla	-35820,541	0,248	0,046	0,000	1	0,752	0,003	0,000	-0,003	6,160	0,013
COX2	Primates	-35823,255	0,751	0,003	0,000	1	0,249	0,045	0,049	0,005	0,733	0,392
COX2	Rodentia	-35818,983	0,749	0,003	0,000	1	0,251	0,049	0,036	-0,012	9,278	0,002
CSN3	–	-18935,942	0,333	0,420	0,552	1	0,115	2,054				
CSN3	Artiodactyla	-18932,846	0,331	0,423	0,530	1	0,139	1,837	2,997	1,160	6,191	0,013
CSN3	Caniformia	-18932,492	0,335	0,418	0,562	1	0,104	2,013	3,668	1,655	6,900	0,009
CSN3	Cetacea	-18935,920	0,333	0,420	0,551	1	0,117	2,044	2,231	0,187	0,044	0,833
CSN3	Eulipotyphla	-18934,942	0,332	0,420	0,551	1	0,117	2,090	1,438	-0,652	1,999	0,157
CSN3	Feliformia	-18933,249	0,333	0,420	0,553	1	0,114	2,099	0,784	-1,315	5,385	0,020
CSN3	Chiroptera	-18934,647	0,333	0,420	0,552	1	0,115	1,969	2,608	0,639	2,590	0,108
CSN3	Perissodactyla	-18933,932	0,333	0,419	0,554	1	0,113	2,024	4,059	2,036	4,018	0,045
CSN3	Primates	-18932,733	0,337	0,418	0,561	1	0,102	1,952	3,051	1,099	6,417	0,011
CSN3	Rodentia	-18923,369	0,341	0,417	0,560	1	0,099	2,676	1,355	-1,321	25,145	0,000

Gen	**Taxon**	*LL*	p_0	ω_0	p_1	ω_1	p_2	ω_2	ω_3	$\Delta\omega$	**LRT**	***p*-hodnota**
CYTB	–	-70253,953	0,742	0,006	0,003	1	0,255	0,089				
CYTB	Artiodactyla	-70251,059	0,743	0,006	0,003	1	0,255	0,086	0,101	0,014	5,789	0,016
CYTB	Caniformia	-70251,683	0,742	0,006	0,003	1	0,255	0,087	0,102	0,015	4,539	0,033
CYTB	Cetacea	-70248,770	0,742	0,006	0,003	1	0,255	0,086	0,115	0,029	10,366	0,001
CYTB	Eulipotyphla	-70248,654	0,741	0,006	0,003	1	0,256	0,090	0,061	-0,029	10,598	0,001
CYTB	Feliformia	-70251,356	0,743	0,006	0,003	1	0,255	0,087	0,109	0,021	5,193	0,023
CYTB	Chiroptera	-70242,567	0,745	0,006	0,003	1	0,253	0,093	0,065	-0,028	22,773	0,000
CYTB	Perissodactyla	-70249,338	0,258	0,088	0,003	1	0,739	0,006	0,001	-0,005	9,231	0,002
CYTB	Primates	-70227,177	0,740	0,006	0,003	1	0,258	0,080	0,126	0,046	53,553	0,000
CYTB	Rodentia	-70236,855	0,741	0,006	0,003	1	0,257	0,096	0,069	-0,026	34,196	0,000
DMP1	–	-40680,219	0,349	0,076	0,186	1	0,465	0,371				
DMP1	Artiodactyla	-40675,597	0,348	0,076	0,188	1	0,464	0,355	0,449	0,094	9,244	0,002
DMP1	Caniformia	-40679,414	0,350	0,077	0,186	1	0,465	0,374	0,311	-0,063	1,610	0,205
DMP1	Cetacea	-40672,925	0,344	0,075	0,190	1	0,465	0,356	0,535	0,179	14,589	0,000
DMP1	Eulipotyphla	-40678,385	0,346	0,076	0,191	1	0,463	0,362	0,435	0,073	3,669	0,055
DMP1	Feliformia	-40679,798	0,349	0,076	0,186	1	0,465	0,372	0,312	-0,060	0,842	0,359
DMP1	Chiroptera	-40678,496	0,351	0,077	0,185	1	0,464	0,365	0,434	0,068	3,448	0,063
DMP1	Perissodactyla	-40680,217	0,349	0,076	0,186	1	0,465	0,371	0,364	-0,007	0,005	0,941
DMP1	Primates	-40679,404	0,350	0,077	0,186	1	0,464	0,366	0,405	0,039	1,631	0,202
DMP1	Rodentia	-40666,338	0,352	0,078	0,191	1	0,457	0,418	0,302	-0,115	27,762	0,000
GHR	–	-21924,471	0,476	0,047	0,172	1	0,352	0,357				
GHR	Artiodactyla	-21922,123	0,479	0,047	0,165	1	0,356	0,374	0,258	-0,115	4,698	0,030
GHR	Caniformia	-21924,337	0,477	0,047	0,171	1	0,353	0,360	0,326	-0,034	0,270	0,603
GHR	Cetacea	-21924,113	0,476	0,047	0,172	1	0,352	0,359	0,271	-0,087	0,718	0,397
GHR	Eulipotyphla	-21924,137	0,350	0,360	0,170	1	0,480	0,049	0,037	-0,012	0,669	0,413
GHR	Feliformia	-21922,713	0,481	0,047	0,166	1	0,353	0,367	0,197	-0,170	3,516	0,061
GHR	Chiroptera	-21916,443	0,476	0,047	0,174	1	0,350	0,388	0,233	-0,155	16,056	0,000
GHR	Perissodactyla	-21924,446	0,477	0,047	0,170	1	0,352	0,359	0,311	-0,048	0,050	0,823
GHR	Primates	-21922,027	0,470	0,045	0,174	1	0,356	0,339	0,472	0,133	4,889	0,027
GHR	Rodentia	-21916,271	0,478	0,047	0,164	1	0,359	0,309	0,451	0,142	16,400	0,000
LCAT	–	-21446,004	0,289	0,234	0,074	1	0,637	0,011				
LCAT	Artiodactyla	-21443,699	0,636	0,011	0,075	1	0,289	0,242	0,169	-0,073	4,612	0,032
LCAT	Caniformia	-21442,632	0,631	0,010	0,079	1	0,290	0,236	0,120	-0,116	6,746	0,009
LCAT	Cetacea	-21445,140	0,636	0,011	0,074	1	0,290	0,237	0,146	-0,091	1,728	0,189
LCAT	Eulipotyphla	-21444,975	0,291	0,232	0,075	1	0,634	0,009	0,017	0,007	2,060	0,151
LCAT	Feliformia	-21444,143	0,284	0,238	0,074	1	0,642	0,012	0,000	-0,012	3,723	0,054
LCAT	Chiroptera	-21421,623	0,638	0,011	0,075	1	0,286	0,200	0,471	0,271	48,763	0,000
LCAT	Perissodactyla	-21442,372	0,289	0,234	0,074	1	0,636	0,010	0,056	0,046	7,265	0,007
LCAT	Primates	-21439,347	0,635	0,010	0,073	1	0,292	0,250	0,143	-0,107	13,315	0,000
LCAT	Rodentia	-21445,601	0,638	0,011	0,074	1	0,287	0,243	0,222	-0,021	0,806	0,369
MYC	–	-20647,126	0,226	0,168	0,007	1	0,767	0,010				
MYC	Artiodactyla	-20646,364	0,765	0,010	0,007	1	0,227	0,164	0,202	0,038	1,523	0,217
MYC	Caniformia	-20640,331	0,212	0,175	0,007	1	0,781	0,012	0,000	-0,012	13,588	0,000
MYC	Cetacea	-20644,060	0,767	0,010	0,007	1	0,226	0,165	0,342	0,178	6,131	0,013
MYC	Eulipotyphla	-20645,252	0,229	0,166	0,007	1	0,764	0,011	0,005	-0,006	3,748	0,053
MYC	Feliformia	-20646,665	0,767	0,010	0,007	1	0,226	0,169	0,129	-0,040	0,920	0,337
MYC	Chiroptera	-20644,446	0,224	0,168	0,007	1	0,768	0,012	0,005	-0,006	5,359	0,021
MYC	Perissodactyla	-20642,402	0,229	0,166	0,007	1	0,764	0,010	0,035	0,026	9,447	0,002
MYC	Primates	-20643,648	0,227	0,167	0,007	1	0,766	0,009	0,018	0,009	6,955	0,008
MYC	Rodentia	-20646,827	0,225	0,168	0,007	1	0,767	0,010	0,012	0,002	0,598	0,439
RAG1	–	-64427,513	0,257	0,284	0,039	1	0,704	0,015				
RAG1	Artiodactyla	-64427,188	0,703	0,015	0,039	1	0,258	0,281	0,300	0,018	0,651	0,420
RAG1	Caniformia	-64427,506	0,704	0,015	0,039	1	0,257	0,284	0,281	-0,003	0,013	0,909
RAG1	Cetacea	-64410,892	0,700	0,015	0,039	1	0,261	0,273	0,591	0,318	33,241	0,000
RAG1	Eulipotyphla	-64427,296	0,704	0,015	0,039	1	0,257	0,282	0,299	0,017	0,435	0,510
RAG1	Feliformia	-64426,955	0,704	0,015	0,039	1	0,257	0,282	0,327	0,045	1,116	0,291
RAG1	Chiroptera	-64425,842	0,703	0,015	0,039	1	0,257	0,289	0,254	-0,035	3,342	0,068
RAG1	Perissodactyla	-64427,502	0,704	0,015	0,039	1	0,257	0,284	0,291	0,008	0,022	0,883
RAG1	Primates	-64423,649	0,702	0,015	0,039	1	0,258	0,275	0,340	0,064	7,728	0,005
RAG1	Rodentia	-64426,077	0,704	0,015	0,039	1	0,258	0,293	0,267	-0,026	2,871	0,090
RAG2	–	-26843,539	0,300	0,267	0,064	1	0,636	0,021				
RAG2	Artiodactyla	-26840,798	0,637	0,022	0,064	1	0,299	0,259	0,352	0,093	5,481	0,019
RAG2	Caniformia	-26843,505	0,636	0,021	0,064	1	0,300	0,268	0,257	-0,011	0,067	0,796
RAG2	Cetacea	-26841,793	0,630	0,021	0,065	1	0,305	0,260	0,409	0,149	3,491	0,062
RAG2	Eulipotyphla	-26843,499	0,636	0,021	0,064	1	0,300	0,268	0,257	-0,011	0,080	0,777
RAG2	Feliformia	-26840,817	0,633	0,021	0,064	1	0,303	0,270	0,143	-0,127	5,443	0,020
RAG2	Chiroptera	-26843,042	0,636	0,021	0,064	1	0,300	0,262	0,292	0,030	0,993	0,319

Gen	**Taxon**	*LL*	p_0	ω_0	p_1	ω_1	p_2	ω_2	ω_3	$\Delta\omega$	**LRT**	***p*-hodnota**
RAG2	Perissodactyla	-26841,816	0,635	0,021	0,063	1	0,302	0,263	0,436	0,172	3,444	0,063
RAG2	Primates	-26843,486	0,636	0,021	0,064	1	0,300	0,266	0,277	0,012	0,104	0,747
RAG2	Rodentia	-26841,617	0,638	0,022	0,062	1	0,300	0,286	0,241	-0,046	3,843	0,050
RBP3	–	-35425,235	0,414	0,256	0,056	1	0,530	0,055				
RBP3	Artiodactyla	-35425,040	0,529	0,055	0,056	1	0,415	0,257	0,241	-0,016	0,392	0,531
RBP3	Caniformia	-35419,703	0,513	0,053	0,058	1	0,429	0,257	0,162	-0,095	11,065	0,001
RBP3	Cetacea	-35422,957	0,530	0,055	0,056	1	0,414	0,259	0,168	-0,091	4,557	0,033
RBP3	Eulipotyphla	-35417,690	0,524	0,054	0,056	1	0,419	0,246	0,417	0,171	15,090	0,000
RBP3	Feliformia	-35424,609	0,524	0,054	0,057	1	0,419	0,255	0,206	-0,049	1,253	0,263
RBP3	Chiroptera	-35344,367	0,293	0,027	0,097	1	0,610	0,137	0,408	0,271	161,737	0,000
RBP3	Perissodactyla	-35425,233	0,530	0,055	0,056	1	0,414	0,256	0,259	0,003	0,004	0,949
RBP3	Primates	-35420,981	0,525	0,055	0,057	1	0,418	0,263	0,195	-0,068	8,508	0,004
RBP3	Rodentia	-35411,430	0,432	0,043	0,070	1	0,498	0,256	0,164	-0,092	27,611	0,000
VWF	–	-37298,001	0,389	0,244	0,057	1	0,554	0,029				
VWF	Artiodactyla	-37297,707	0,554	0,029	0,057	1	0,389	0,242	0,257	0,016	0,587	0,444
VWF	Caniformia	-37297,664	0,553	0,029	0,057	1	0,389	0,242	0,265	0,023	0,673	0,412
VWF	Cetacea	-37295,266	0,553	0,029	0,056	1	0,390	0,240	0,338	0,097	5,469	0,019
VWF	Eulipotyphla	-37294,282	0,554	0,029	0,056	1	0,390	0,250	0,186	-0,064	7,437	0,006
VWF	Feliformia	-37294,275	0,553	0,029	0,057	1	0,390	0,240	0,357	0,118	7,453	0,006
VWF	Chiroptera	-37297,445	0,554	0,029	0,058	1	0,389	0,240	0,263	0,022	1,112	0,292
VWF	Perissodactyla	-37297,036	0,554	0,029	0,057	1	0,389	0,243	0,317	0,075	1,931	0,165
VWF	Primates	-37294,240	0,557	0,030	0,056	1	0,386	0,253	0,197	-0,056	7,522	0,006
VWF	Rodentia	-37298,001	0,554	0,029	0,057	1	0,389	0,244	0,244	-0,000	0,000	0,992

Příloha 4

Výsledky modelů větví-pozicí bsA a bsA1

Následující tabulka obsahuje kompletní výsledky modelů bsA a bsA1.

Gen	Model	Taxon	*LL*	p_0	ω_0	p_1	ω_1	p_{2a}	p_{2b}	ω_2	LRT	*p*-hodnota
APOB	bsA1	Artiodactyla	-33984,190	0,610	0,241	0,294	1	0,065	0,031	1,000		
APOB	bsA	Artiodactyla	-33984,190	0,610	0,241	0,294	1	0,065	0,031	1,000	-0,000	1,000
APOB	bsA1	Caniformia	-33988,524	0,642	0,247	0,309	1	0,033	0,016	1,000		
APOB	bsA	Caniformia	-33983,461	0,668	0,250	0,312	1	0,014	0,006	4,838	10,126	0,001
APOB	bsA1	Cetacea	-33987,261	0,576	0,245	0,280	1	0,097	0,047	1,000		
APOB	bsA	Cetacea	-33987,261	0,576	0,245	0,280	1	0,097	0,047	1,000	-0,000	1,000
APOB	bsA1	Eulipotyphla	-33984,958	0,605	0,241	0,289	1	0,072	0,034	1,000		
APOB	bsA	Eulipotyphla	-33984,958	0,605	0,241	0,289	1	0,071	0,034	1,000	-0,000	1,000
APOB	bsA1	Feliformia	-33989,089	0,649	0,248	0,313	1	0,025	0,012	1,000		
APOB	bsA	Feliformia	-33988,125	0,669	0,249	0,318	1	0,009	0,004	5,370	1,929	0,165
APOB	bsA1	Chiroptera	-33969,328	0,585	0,236	0,259	1	0,108	0,048	1,000		
APOB	bsA	Chiroptera	-33966,871	0,620	0,241	0,269	1	0,077	0,034	1,588	4,914	0,027
APOB	bsA1	Perissodactyla	-33989,107	0,632	0,247	0,306	1	0,042	0,020	1,000		
APOB	bsA	Perissodactyla	-33989,098	0,635	0,247	0,307	1	0,039	0,019	1,143	0,018	0,893
APOB	bsA1	Primates	-33983,581	0,602	0,240	0,297	1	0,068	0,033	1,000		
APOB	bsA	Primates	-33983,581	0,602	0,240	0,297	1	0,068	0,033	1,000	0,000	1,000
APOB	bsA1	Rodentia	-33977,740	0,650	0,240	0,286	1	0,045	0,020	1,000		
APOB	bsA	Rodentia	-33977,740	0,650	0,240	0,286	1	0,045	0,020	1,000	0,000	0,999
BRCA1	bsA1	Artiodactyla	-93511,302	0,503	0,340	0,442	1	0,030	0,026	1,000		
BRCA1	bsA	Artiodactyla	-93501,954	0,512	0,343	0,449	1	0,021	0,018	3,034	18,695	0,000
BRCA1	bsA1	Caniformia	-93511,871	0,516	0,342	0,450	1	0,018	0,016	1,000		
BRCA1	bsA	Caniformia	-93492,083	0,520	0,343	0,450	1	0,016	0,014	4,508	39,577	0,000
BRCA1	bsA1	Cetacea	-93504,174	0,388	0,338	0,343	1	0,143	0,126	1,000		
BRCA1	bsA	Cetacea	-93489,093	0,502	0,340	0,445	1	0,028	0,025	5,818	30,163	0,000
BRCA1	bsA1	Feliformia	-93511,836	0,500	0,341	0,438	1	0,033	0,029	1,000		
BRCA1	bsA	Feliformia	-93500,711	0,517	0,342	0,449	1	0,018	0,016	4,755	22,250	0,000
BRCA1	bsA1	Chiroptera	-93510,626	0,503	0,339	0,440	1	0,030	0,027	1,000		
BRCA1	bsA	Chiroptera	-93508,669	0,521	0,342	0,454	1	0,014	0,012	2,264	3,913	0,048
BRCA1	bsA1	Perissodactyla	-93512,656	0,532	0,343	0,468	1	0,000	0,000	1,000		
BRCA1	bsA	Perissodactyla	-93512,655	0,532	0,343	0,468	1	0,000	0,000	1,000	0,002	0,965
BRCA1	bsA1	Primates	-93511,022	0,501	0,340	0,439	1	0,032	0,028	1,000		
BRCA1	bsA	Primates	-93504,282	0,524	0,342	0,458	1	0,010	0,009	3,233	13,480	0,000
BRCA1	bsA1	Rodentia	-93449,185	0,434	0,310	0,328	1	0,136	0,102	1,000		
BRCA1	bsA	Rodentia	-93449,185	0,434	0,310	0,328	1	0,136	0,102	1,000	-0,000	1,000
COX1	bsA1	Artiodactyla	-78441,995	0,990	0,006	0,006	1	0,004	0,000	1,000		
COX1	bsA	Artiodactyla	-78441,995	0,990	0,006	0,006	1	0,004	0,000	1,000	0,000	0,998
COX1	bsA1	Caniformia	-78448,801	0,990	0,006	0,006	1	0,004	0,000	1,000		
COX1	bsA	Caniformia	-78448,801	0,990	0,006	0,006	1	0,004	0,000	1,000	0,000	0,998
COX1	bsA1	Cetacea	-78394,918	0,979	0,006	0,006	1	0,015	0,000	1,000		
COX1	bsA	Cetacea	-78394,918	0,979	0,006	0,006	1	0,015	0,000	1,000	0,000	0,998
COX1	bsA1	Eulipotyphla	-78457,639	0,994	0,006	0,006	1	0,000	0,000	1,000		
COX1	bsA	Eulipotyphla	-78457,639	0,994	0,006	0,006	1	0,000	0,000	1,000	0,000	0,996
COX1	bsA1	Feliformia	-78447,676	0,986	0,006	0,006	1	0,008	0,000	1,000		
COX1	bsA	Feliformia	-78447,676	0,986	0,006	0,006	1	0,008	0,000	1,000	0,000	1,000
COX1	bsA1	Chiroptera	-78457,639	0,994	0,006	0,006	1	0,000	0,000	1,000		
COX1	bsA	Chiroptera	-78457,639	0,994	0,006	0,006	1	0,000	0,000	1,000	-0,000	1,000
COX1	bsA1	Perissodactyla	-78453,864	0,992	0,006	0,006	1	0,002	0,000	1,000		
COX1	bsA	Perissodactyla	-78453,864	0,992	0,006	0,006	1	0,002	0,000	1,000	0,000	1,000
COX1	bsA1	Primates	-78385,869	0,978	0,006	0,006	1	0,016	0,000	1,000		
COX1	bsA	Primates	-78385,869	0,978	0,006	0,006	1	0,016	0,000	1,000	0,000	1,000
COX1	bsA1	Rodentia	-78449,868	0,990	0,006	0,006	1	0,004	0,000	1,000		
COX1	bsA	Rodentia	-78449,868	0,990	0,006	0,006	1	0,004	0,000	1,000	0,000	0,998

Gen	Model	Taxon	*LL*	p_0	ω_0	p_1	ω_1	p_{2a}	p_{2b}	ω_2	LRT	*p*-hodnota
COX2	bsA1	Artiodactyla	-36472,742	0,981	0,014	0,009	1	0,011	0,000	1,000		
COX2	bsA	Artiodactyla	-36472,742	0,981	0,014	0,009	1	0,011	0,000	1,000	0,000	1,000
COX2	bsA1	Caniformia	-36475,688	0,991	0,014	0,009	1	0,000	0,000	1,000		
COX2	bsA	Caniformia	-36475,688	0,991	0,014	0,009	1	0,000	0,000	4,796	0,000	1,000
COX2	bsA1	Cetacea	-36438,294	0,954	0,013	0,008	1	0,038	0,000	1,000		
COX2	bsA	Cetacea	-36438,294	0,954	0,013	0,008	1	0,038	0,000	1,000	0,000	1,000
COX2	bsA1	Eulipotyphla	-36475,615	0,989	0,014	0,009	1	0,002	0,000	1,000		
COX2	bsA	Eulipotyphla	-36475,615	0,989	0,014	0,009	1	0,002	0,000	1,000	0,000	1,000
COX2	bsA1	Feliformia	-36446,709	0,974	0,013	0,009	1	0,018	0,000	1,000		
COX2	bsA	Feliformia	-36446,709	0,974	0,013	0,009	1	0,018	0,000	1,000	0,000	0,999
COX2	bsA1	Chiroptera	-36451,306	0,976	0,014	0,009	1	0,016	0,000	1,000		
COX2	bsA	Chiroptera	-36451,306	0,976	0,014	0,009	1	0,016	0,000	1,000	0,000	1,000
COX2	bsA1	Perissodactyla	-36475,688	0,991	0,014	0,009	1	0,000	0,000	1,000		
COX2	bsA	Perissodactyla	-36475,688	0,991	0,014	0,009	1	0,000	0,000	2,375	0,000	1,000
COX2	bsA1	Primates	-36469,273	0,981	0,014	0,009	1	0,010	0,000	1,000		
COX2	bsA	Primates	-36469,273	0,981	0,014	0,009	1	0,010	0,000	1,000	0,000	1,000
COX2	bsA1	Rodentia	-36473,644	0,987	0,014	0,009	1	0,004	0,000	1,000		
COX2	bsA	Rodentia	-36473,644	0,987	0,014	0,009	1	0,004	0,000	1,000	0,000	1,000
CSN3	bsA1	Artiodactyla	-18979,039	0,364	0,402	0,561	1	0,030	0,046	1,000		
CSN3	bsA	Artiodactyla	-18968,804	0,342	0,403	0,531	1	0,049	0,077	3,073	20,468	0,000
CSN3	bsA1	Caniformia	-18978,840	0,326	0,402	0,504	1	0,067	0,103	1,000		
CSN3	bsA	Caniformia	-18975,783	0,370	0,407	0,567	1	0,025	0,038	4,079	6,113	0,013
CSN3	bsA1	Cetacea	-18978,805	0,329	0,404	0,508	1	0,064	0,099	1,000		
CSN3	bsA	Cetacea	-18974,045	0,354	0,404	0,549	1	0,038	0,059	5,387	9,521	0,002
CSN3	bsA1	Eulipotyphla	-18977,306	0,291	0,396	0,454	1	0,100	0,156	1,000		
CSN3	bsA	Eulipotyphla	-18974,980	0,341	0,399	0,525	1	0,053	0,081	2,571	4,652	0,031
CSN3	bsA1	Feliformia	-18978,116	0,261	0,401	0,406	1	0,130	0,203	1,000		
CSN3	bsA	Feliformia	-18978,116	0,261	0,401	0,406	1	0,130	0,203	1,000	0,000	1,000
CSN3	bsA1	Chiroptera	-18974,014	0,257	0,384	0,393	1	0,138	0,211	1,000		
CSN3	bsA	Chiroptera	-18961,887	0,343	0,398	0,516	1	0,056	0,085	2,948	24,255	0,000
CSN3	bsA1	Perissodactyla	-18979,323	0,392	0,407	0,608	1	0,000	0,000	1,000		
CSN3	bsA	Perissodactyla	-18978,039	0,380	0,409	0,575	1	0,018	0,027	6,519	2,570	0,109
CSN3	bsA1	Primates	-18979,244	0,374	0,405	0,574	1	0,021	0,031	1,000		
CSN3	bsA	Primates	-18966,204	0,369	0,408	0,595	1	0,014	0,022	5,270	26,079	0,000
CSN3	bsA1	Rodentia	-18963,680	0,324	0,378	0,314	1	0,184	0,178	1,000		
CSN3	bsA	Rodentia	-18958,774	0,350	0,400	0,330	1	0,165	0,155	1,405	9,811	0,002
CYTB	bsA1	Artiodactyla	-72583,071	0,975	0,026	0,005	1	0,020	0,000	1,000		
CYTB	bsA	Artiodactyla	-72583,071	0,975	0,026	0,005	1	0,020	0,000	1,000	0,000	1,000
CYTB	bsA1	Caniformia	-72644,990	0,987	0,027	0,005	1	0,008	0,000	1,000		
CYTB	bsA	Caniformia	-72644,990	0,987	0,027	0,005	1	0,008	0,000	1,000	0,000	1,000
CYTB	bsA1	Cetacea	-72578,317	0,948	0,026	0,005	1	0,047	0,000	1,000		
CYTB	bsA	Cetacea	-72578,317	0,948	0,026	0,005	1	0,047	0,000	1,000	0,000	1,000
CYTB	bsA1	Eulipotyphla	-72634,686	0,966	0,027	0,005	1	0,029	0,000	1,000		
CYTB	bsA	Eulipotyphla	-72634,686	0,966	0,027	0,005	1	0,029	0,000	1,000	0,000	0,999
CYTB	bsA1	Feliformia	-72630,690	0,972	0,026	0,005	1	0,022	0,000	1,000		
CYTB	bsA	Feliformia	-72630,690	0,972	0,026	0,005	1	0,022	0,000	1,000	0,000	1,000
CYTB	bsA1	Chiroptera	-72639,386	0,978	0,027	0,005	1	0,016	0,000	1,000		
CYTB	bsA	Chiroptera	-72639,386	0,978	0,027	0,005	1	0,016	0,000	1,000	0,000	1,000
CYTB	bsA1	Perissodactyla	-72651,561	0,995	0,027	0,005	1	0,000	0,000	1,000		
CYTB	bsA	Perissodactyla	-72651,561	0,995	0,027	0,005	1	0,000	0,000	1,455	0,000	1,000
CYTB	bsA1	Primates	-72570,710	0,934	0,026	0,002	1	0,063	0,000	1,000		
CYTB	bsA	Primates	-72570,710	0,934	0,026	0,002	1	0,063	0,000	1,000	0,000	1,000
CYTB	bsA1	Rodentia	-72647,738	0,988	0,027	0,005	1	0,007	0,000	1,000		
CYTB	bsA	Rodentia	-72647,738	0,988	0,027	0,005	1	0,007	0,000	1,000	-0,000	1,000
DMP1	bsA1	Artiodactyla	-40932,919	0,554	0,184	0,278	1	0,112	0,056	1,000		
DMP1	bsA	Artiodactyla	-40932,919	0,554	0,184	0,278	1	0,112	0,056	1,000	0,000	1,000
DMP1	bsA1	Caniformia	-40973,991	0,623	0,194	0,350	1	0,017	0,010	1,000		
DMP1	bsA	Caniformia	-40973,991	0,623	0,194	0,350	1	0,017	0,010	1,000	-0,000	1,000
DMP1	bsA1	Cetacea	-40946,766	0,531	0,188	0,285	1	0,120	0,064	1,000		
DMP1	bsA	Cetacea	-40943,893	0,567	0,189	0,305	1	0,083	0,045	1,669	5,747	0,017
DMP1	bsA1	Eulipotyphla	-40946,594	0,529	0,184	0,298	1	0,110	0,062	1,000		
DMP1	bsA	Eulipotyphla	-40946,594	0,529	0,184	0,298	1	0,110	0,062	1,000	0,000	0,998
DMP1	bsA1	Feliformia	-40974,181	0,632	0,194	0,355	1	0,009	0,005	1,000		
DMP1	bsA	Feliformia	-40974,181	0,632	0,194	0,355	1	0,009	0,005	1,000	0,000	1,000
DMP1	bsA1	Chiroptera	-40953,752	0,586	0,189	0,317	1	0,063	0,034	1,000		
DMP1	bsA	Chiroptera	-40951,389	0,610	0,191	0,332	1	0,038	0,020	1,884	4,727	0,030
DMP1	bsA1	Perissodactyla	-40974,226	0,640	0,194	0,360	1	0,000	0,000	1,000		

Gen	Model	Taxon	*LL*	p_0	ω_0	p_1	ω_1	p_{2a}	p_{2b}	ω_2	LRT	*p*-hodnota
DMP1	bsA	Perissodactyla	-40974,041	0,635	0,194	0,357	1	0,005	0,003	3,293	0,371	0,542
DMP1	bsA1	Primates	-40962,027	0,593	0,190	0,311	1	0,063	0,033	1,000		
DMP1	bsA	Primates	-40962,027	0,593	0,190	0,311	1	0,063	0,033	1,000	-0,000	1,000
DMP1	bsA1	Rodentia	-40963,704	0,629	0,192	0,331	1	0,026	0,014	1,000		
DMP1	bsA	Rodentia	-40963,704	0,629	0,192	0,331	1	0,026	0,014	1,000	-0,000	1,000
GHR	bsA1	Artiodactyla	-22044,030	0,694	0,125	0,296	1	0,007	0,003	1,000		
GHR	bsA	Artiodactyla	-22044,030	0,694	0,125	0,296	1	0,007	0,003	1,000	0,000	1,000
GHR	bsA1	Caniformia	-22044,391	0,702	0,126	0,298	1	0,000	0,000	1,000		
GHR	bsA	Caniformia	-22044,391	0,702	0,126	0,298	1	0,000	0,000	1,000	-0,000	1,000
GHR	bsA1	Cetacea	-22043,998	0,683	0,125	0,289	1	0,020	0,009	1,000		
GHR	bsA	Cetacea	-22043,401	0,697	0,126	0,294	1	0,007	0,003	4,373	1,194	0,275
GHR	bsA1	Eulipotyphla	-22043,472	0,688	0,125	0,288	1	0,017	0,007	1,000		
GHR	bsA	Eulipotyphla	-22042,700	0,693	0,126	0,287	1	0,014	0,006	2,690	1,544	0,214
GHR	bsA1	Feliformia	-22044,391	0,702	0,126	0,298	1	0,000	0,000	1,000		
GHR	bsA	Feliformia	-22044,391	0,702	0,126	0,298	1	0,000	0,000	1,000	0,000	1,000
GHR	bsA1	Chiroptera	-22044,391	0,702	0,126	0,298	1	0,000	0,000	1,000		
GHR	bsA	Chiroptera	-22044,391	0,702	0,126	0,298	1	0,000	0,000	34,444	0,000	0,999
GHR	bsA1	Perissodactyla	-22044,391	0,702	0,126	0,298	1	0,000	0,000	1,000		
GHR	bsA	Perissodactyla	-22044,391	0,702	0,126	0,298	1	0,000	0,000	1,000	0,000	1,000
GHR	bsA1	Primates	-22034,163	0,621	0,118	0,265	1	0,080	0,034	1,000		
GHR	bsA	Primates	-22034,163	0,621	0,118	0,265	1	0,080	0,034	1,000	-0,000	1,000
GHR	bsA1	Rodentia	-22025,990	0,650	0,115	0,242	1	0,079	0,029	1,000		
GHR	bsA	Rodentia	-22025,990	0,650	0,115	0,242	1	0,079	0,029	1,000	0,000	1,000
LCAT	bsA1	Artiodactyla	-21680,622	0,833	0,052	0,167	1	0,000	0,000	1,000		
LCAT	bsA	Artiodactyla	-21680,622	0,833	0,052	0,167	1	0,000	0,000	1,000	0,000	0,998
LCAT	bsA1	Caniformia	-21678,973	0,826	0,051	0,167	1	0,006	0,001	1,000		
LCAT	bsA	Caniformia	-21678,816	0,828	0,051	0,167	1	0,004	0,001	1,696	0,315	0,575
LCAT	bsA1	Cetacea	-21680,622	0,833	0,052	0,167	1	0,000	0,000	1,000		
LCAT	bsA	Cetacea	-21680,622	0,833	0,052	0,167	1	0,000	0,000	1,000	-0,000	1,000
LCAT	bsA1	Eulipotyphla	-21675,873	0,812	0,051	0,157	1	0,026	0,005	1,000		
LCAT	bsA	Eulipotyphla	-21675,873	0,812	0,051	0,157	1	0,026	0,005	1,000	0,000	0,998
LCAT	bsA1	Feliformia	-21677,854	0,824	0,051	0,166	1	0,008	0,002	1,000		
LCAT	bsA	Feliformia	-21675,873	0,829	0,051	0,168	1	0,003	0,001	7,096	3,961	0,047
LCAT	bsA1	Chiroptera	-21618,675	0,756	0,043	0,148	1	0,080	0,016	1,000		
LCAT	bsA	Chiroptera	-21618,675	0,756	0,043	0,148	1	0,080	0,016	1,000	0,000	0,999
LCAT	bsA1	Perissodactyla	-21680,588	0,825	0,051	0,166	1	0,007	0,001	1,000		
LCAT	bsA	Perissodactyla	-21680,588	0,825	0,051	0,166	1	0,007	0,001	1,000	-0,000	1,000
LCAT	bsA1	Primates	-21680,592	0,831	0,051	0,167	1	0,001	0,000	1,000		
LCAT	bsA	Primates	-21680,592	0,831	0,051	0,167	1	0,001	0,000	1,000	0,000	1,000
LCAT	bsA1	Rodentia	-21671,373	0,824	0,050	0,161	1	0,013	0,003	1,000		
LCAT	bsA	Rodentia	-21671,373	0,824	0,050	0,161	1	0,013	0,003	1,000	-0,000	1,000
MYC	bsA1	Artiodactyla	-20933,998	0,929	0,035	0,042	1	0,027	0,001	1,000		
MYC	bsA	Artiodactyla	-20933,098	0,929	0,035	0,042	1	0,027	0,001	1,000	1,800	0,180
MYC	bsA1	Caniformia	-20950,215	0,955	0,036	0,045	1	0,000	0,000	1,000		
MYC	bsA	Caniformia	-20950,215	0,955	0,036	0,045	1	0,000	0,000	16,982	0,000	0,997
MYC	bsA1	Cetacea	-20934,841	0,923	0,035	0,044	1	0,031	0,002	1,000		
MYC	bsA	Cetacea	-20931,598	0,942	0,035	0,045	1	0,013	0,001	3,781	6,486	0,011
MYC	bsA1	Eulipotyphla	-20943,825	0,936	0,035	0,045	1	0,017	0,001	1,000		
MYC	bsA	Eulipotyphla	-20943,825	0,936	0,035	0,045	1	0,017	0,001	1,000	0,000	1,000
MYC	bsA1	Feliformia	-20950,208	0,954	0,036	0,045	1	0,001	0,000	1,000		
MYC	bsA	Feliformia	-20950,208	0,954	0,036	0,045	1	0,001	0,000	1,000	0,000	0,999
MYC	bsA1	Chiroptera	-20931,280	0,938	0,034	0,044	1	0,017	0,001	1,000		
MYC	bsA	Chiroptera	-20931,280	0,938	0,034	0,044	1	0,017	0,001	1,000	0,000	0,999
MYC	bsA1	Perissodactyla	-20950,200	0,951	0,036	0,045	1	0,004	0,000	1,000		
MYC	bsA	Perissodactyla	-20950,200	0,951	0,036	0,045	1	0,004	0,000	1,000	0,000	1,000
MYC	bsA1	Primates	-20943,296	0,943	0,036	0,038	1	0,018	0,001	1,000		
MYC	bsA	Primates	-20943,296	0,943	0,036	0,038	1	0,018	0,001	1,000	0,000	0,997
MYC	bsA1	Rodentia	-20944,878	0,945	0,035	0,040	1	0,015	0,001	1,000		
MYC	bsA	Rodentia	-20944,878	0,945	0,035	0,040	1	0,015	0,001	1,000	-0,000	1,000
RAG1	bsA1	Artiodactyla	-65275,412	0,810	0,045	0,169	1	0,017	0,004	1,000		
RAG1	bsA	Artiodactyla	-65275,412	0,810	0,045	0,169	1	0,017	0,004	1,000	0,000	0,991
RAG1	bsA1	Caniformia	-65275,294	0,803	0,044	0,170	1	0,022	0,005	1,000		
RAG1	bsA	Caniformia	-65275,294	0,803	0,044	0,170	1	0,022	0,005	1,000	-0,000	1,000
RAG1	bsA1	Cetacea	-65275,439	0,768	0,045	0,160	1	0,059	0,012	1,000		
RAG1	bsA	Cetacea	-65275,255	0,779	0,045	0,162	1	0,048	0,010	1,283	0,368	0,544
RAG1	bsA1	Eulipotyphla	-65294,294	0,815	0,046	0,170	1	0,012	0,003	1,000		
RAG1	bsA	Eulipotyphla	-65294,294	0,815	0,046	0,170	1	0,012	0,003	1,000	0,000	0,991

Gen	Model	Taxon	*LL*	p_0	ω_0	p_1	ω_1	p_{2a}	p_{2b}	ω_2	LRT	*p*-hodnota
RAG1	bsA1	Feliformia	-65293,610	0,815	0,046	0,172	1	0,010	0,002	1,000		
RAG1	bsA	Feliformia	-65293,535	0,821	0,046	0,174	1	0,005	0,001	1,779	0,150	0,698
RAG1	bsA1	Chiroptera	-65292,805	0,819	0,046	0,171	1	0,008	0,002	1,000		
RAG1	bsA	Chiroptera	-65292,805	0,819	0,046	0,171	1	0,008	0,002	1,000	-0,000	1,000
RAG1	bsA1	Perissodactyla	-65297,478	0,826	0,046	0,174	1	0,000	0,000	1,000		
RAG1	bsA	Perissodactyla	-65297,478	0,826	0,046	0,174	1	0,000	0,000	1,000	0,000	0,990
RAG1	bsA1	Primates	-65262,134	0,799	0,044	0,164	1	0,031	0,006	1,000		
RAG1	bsA	Primates	-65262,134	0,799	0,044	0,164	1	0,031	0,006	1,000	0,000	0,993
RAG1	bsA1	Rodentia	-65276,862	0,817	0,046	0,162	1	0,017	0,003	1,000		
RAG1	bsA	Rodentia	-65276,861	0,817	0,046	0,162	1	0,017	0,003	1,000	0,000	0,992
RAG2	bsA1	Artiodactyla	-27045,190	0,830	0,075	0,132	1	0,032	0,005	1,000		
RAG2	bsA	Artiodactyla	-27045,190	0,830	0,075	0,132	1	0,032	0,005	1,000	0,000	0,997
RAG2	bsA1	Caniformia	-27055,511	0,845	0,076	0,137	1	0,016	0,003	1,000		
RAG2	bsA	Caniformia	-27055,511	0,845	0,076	0,137	1	0,016	0,003	1,004	0,000	0,994
RAG2	bsA1	Cetacea	-27048,496	0,773	0,075	0,127	1	0,086	0,014	1,000		
RAG2	bsA	Cetacea	-27048,463	0,788	0,075	0,129	1	0,071	0,012	1,201	0,066	0,798
RAG2	bsA1	Eulipotyphla	-27057,366	0,859	0,077	0,141	1	0,000	0,000	1,000		
RAG2	bsA	Eulipotyphla	-27057,366	0,859	0,077	0,141	1	0,000	0,000	1,000	0,000	0,996
RAG2	bsA1	Feliformia	-27057,366	0,859	0,077	0,141	1	0,000	0,000	1,000		
RAG2	bsA	Feliformia	-27057,366	0,859	0,077	0,141	1	0,000	0,000	1,000	0,000	0,996
RAG2	bsA1	Chiroptera	-27052,753	0,846	0,076	0,134	1	0,018	0,003	1,000		
RAG2	bsA	Chiroptera	-27052,753	0,846	0,076	0,134	1	0,018	0,003	1,000	0,000	0,997
RAG2	bsA1	Perissodactyla	-27053,929	0,806	0,076	0,132	1	0,053	0,009	1,000		
RAG2	bsA	Perissodactyla	-27053,323	0,839	0,076	0,136	1	0,022	0,004	2,673	1,210	0,271
RAG2	bsA1	Primates	-27052,012	0,844	0,077	0,131	1	0,022	0,003	1,000		
RAG2	bsA	Primates	-27051,958	0,848	0,077	0,131	1	0,018	0,003	1,186	0,108	0,742
RAG2	bsA1	Rodentia	-27046,082	0,834	0,072	0,134	1	0,028	0,004	1,000		
RAG2	bsA	Rodentia	-27046,082	0,834	0,072	0,134	1	0,028	0,004	1,000	0,000	0,997
RBP3	bsA1	Artiodactyla	-35652,067	0,861	0,124	0,107	1	0,029	0,004	1,000		
RBP3	bsA	Artiodactyla	-35651,794	0,866	0,124	0,107	1	0,024	0,003	1,200	0,546	0,460
RBP3	bsA1	Caniformia	-35670,933	0,879	0,126	0,113	1	0,006	0,001	1,000		
RBP3	bsA	Caniformia	-35670,851	0,881	0,126	0,113	1	0,005	0,001	1,394	0,165	0,684
RBP3	bsA1	Cetacea	-35667,999	0,869	0,126	0,110	1	0,019	0,002	1,000		
RBP3	bsA	Cetacea	-35665,680	0,881	0,126	0,111	1	0,007	0,001	3,083	4,638	0,031
RBP3	bsA1	Eulipotyphla	-35664,919	0,810	0,124	0,102	1	0,078	0,010	1,000		
RBP3	bsA	Eulipotyphla	-35664,919	0,810	0,124	0,102	1	0,078	0,010	1,000	0,000	1,000
RBP3	bsA1	Feliformia	-35671,112	0,874	0,126	0,113	1	0,011	0,001	1,000		
RBP3	bsA	Feliformia	-35670,883	0,878	0,126	0,113	1	0,008	0,001	1,715	0,459	0,498
RBP3	bsA1	Chiroptera	-35515,636	0,536	0,099	0,075	1	0,342	0,048	1,000		
RBP3	bsA	Chiroptera	-35515,636	0,536	0,099	0,075	1	0,342	0,048	1,000	0,000	0,998
RBP3	bsA1	Perissodactyla	-35671,708	0,885	0,127	0,114	1	0,000	0,000	1,000		
RBP3	bsA	Perissodactyla	-35671,677	0,884	0,127	0,114	1	0,002	0,000	2,844	0,063	0,802
RBP3	bsA1	Primates	-35667,018	0,874	0,126	0,112	1	0,012	0,002	1,000		
RBP3	bsA	Primates	-35666,976	0,876	0,126	0,112	1	0,010	0,001	1,166	0,084	0,772
RBP3	bsA1	Rodentia	-35642,537	0,857	0,122	0,094	1	0,044	0,005	1,000		
RBP3	bsA	Rodentia	-35642,536	0,857	0,122	0,094	1	0,044	0,005	1,000	0,001	0,976
VWF	bsA1	Artiodactyla	-37809,282	0,823	0,097	0,116	1	0,053	0,007	1,000		
VWF	bsA	Artiodactyla	-37809,282	0,823	0,097	0,117	1	0,053	0,007	1,000	0,000	0,984
VWF	bsA1	Caniformia	-37827,368	0,853	0,100	0,119	1	0,025	0,004	1,000		
VWF	bsA	Caniformia	-37827,368	0,853	0,100	0,119	1	0,025	0,004	1,000	0,000	1,000
VWF	bsA1	Cetacea	-37815,460	0,811	0,099	0,114	1	0,065	0,009	1,000		
VWF	bsA	Cetacea	-37814,895	0,826	0,099	0,116	1	0,051	0,007	1,361	1,129	0,288
VWF	bsA1	Eulipotyphla	-37828,025	0,847	0,100	0,120	1	0,029	0,004	1,000		
VWF	bsA	Eulipotyphla	-37828,025	0,847	0,100	0,120	1	0,029	0,004	1,000	-0,000	1,000
VWF	bsA1	Feliformia	-37819,199	0,828	0,100	0,114	1	0,050	0,007	1,000		
VWF	bsA	Feliformia	-37816,599	0,852	0,100	0,117	1	0,027	0,004	2,239	5,201	0,023
VWF	bsA1	Chiroptera	-37792,381	0,821	0,097	0,111	1	0,059	0,008	1,000		
VWF	bsA	Chiroptera	-37792,381	0,821	0,097	0,111	1	0,059	0,008	1,000	-0,000	1,000
VWF	bsA1	Perissodactyla	-37830,602	0,835	0,101	0,118	1	0,041	0,006	1,000		
VWF	bsA	Perissodactyla	-37829,903	0,852	0,101	0,121	1	0,024	0,003	2,100	1,399	0,237
VWF	bsA1	Primates	-37829,687	0,856	0,100	0,121	1	0,020	0,003	1,000		
VWF	bsA	Primates	-37829,687	0,856	0,100	0,121	1	0,020	0,003	1,000	0,000	0,998
VWF	bsA1	Rodentia	-37809,620	0,842	0,098	0,118	1	0,035	0,005	1,000		
VWF	bsA	Rodentia	-37809,620	0,842	0,098	0,118	1	0,035	0,005	1,000	-0,000	1,000

Seznam použitých zdrojů

ADENAIKE A. S., PETERS S. O., FAFIOLU A. O., LAWAL R. A., WHETO M., IKEOBI C. O. N., Bioinformatic analyses of kappa casein gene in mammalian livestock species. *Nigerian Journal of Animal Production*. December 2020, 40, 2, s. 7–12. ISSN 0331-2062. doi: 10.51791/njap.v40i2.1126.

ADKINS R. M., HONEYCUTT R. L., Molecular phylogeny of the superorder Archonta. *Proceedings of the National Academy of Sciences*. 1991, 88, 22, s. 10317–10321. doi: 10.1073/pnas.88.22.10317.

ADKINS R. M., GELKE E. L., ROWE D., HONEYCUTT R. L., Molecular Phylogeny and Divergence Time Estimates for Major Rodent Groups: Evidence from Multiple Genes. *Molecular Biology and Evolution*. May 2001, 18, 5, s. 777–791. ISSN 0737-4038. doi: 10.1093/oxfordjournals.molbev.a003860.

AGUILAR-TOMASINI M. A., MARTIN M. D., SPEED J. D., Assessing spatial patterns of phylogenetic diversity of Mexican mammals for biodiversity conservation. *Global Ecology and Conservation*. 2021, 31, s. e01834. ISSN 2351-9894. doi: https://doi.org/10.1016/j.gecco.2021.e01834.

ÁLVAREZ-CARRETERO S., KAPLI P., YANG Z., Beginner's Guide on the Use of PAML to Detect Positive Selection. *Molecular Biology and Evolution*. April 2023, 40, 4, s. msad041. ISSN 1537-1719. doi: 10.1093/molbev/msad041.

AMRINE-MADSEN H., KOEPFLI K.-P., WAYNE R. K., SPRINGER M. S., A new phylogenetic marker, apolipoprotein B, provides compelling evidence for eutherian relationships. *Molecular Phylogenetics and Evolution*. August 2003, 28, 2, s. 225–240. ISSN 1055-7903. doi: 10.1016/S1055-7903(03)00118-0.

ANISIMOVA M., BIELAWSKI J. P., YANG Z., Accuracy and Power of the Likelihood Ratio Test in Detecting Adaptive Molecular Evolution. *Molecular Biology and Evolution*. August 2001, 18, 8, s. 1585–1592. ISSN 1537-1719, 0737-4038. doi: 10.1093/oxfordjournals.molbev.a003945.

ANISIMOVA M., BIELAWSKI J. P., YANG Z., Accuracy and Power of Bayes Prediction of Amino Acid Sites Under Positive Selection. *Molecular Biology and Evolution*. June 2002, 19, 6, s. 950–958. ISSN 0737-4038. doi: 10.1093/oxfordjournals.molbev.a004152.

ASHLEY M. V., NORMAN J. E., STROSS L., Phylogenetic analysis of the perissodactylan family Tapiridae using mitochondrial cytochromec oxidase (COII) sequences. *Journal of Mammalian Evolution*. December 1996, 3, 4, s. 315–326. ISSN 1573-7055. doi: 10.1007/BF02077448.

BAKER M. L., WARES J. P., HARRISON G. A., MILLER R. D., Relationships Among the Families and Orders of Marsupials and the Major Mammalian Lineages Based on

Recombination Activating Gene-1. *Journal of Mammalian Evolution*. March 2004, 11, 1, s. 1–16. ISSN 1573-7055. doi: 10.1023/B:JOMM.0000029143.39776.ec.

BAURAIN D., *Bio-MUST-Apps-HmmCleaner* [online]. 2025a. Navštíveno: 22. března 2025. Dostupné z: https://metacpan.org/dist/Bio-MUST-Apps-HmmCleaner.

BAURAIN D., *ali2fasta.pl - Convert ALI files to FASTA files* [online]. 2025b. Navštíveno: 22. března 2025. Dostupné z: https://metacpan.org/dist/Bio-MUST-Core/view/bin/ali2fasta.pl.

BERGSTEN J., A review of long-branch attraction. *Cladistics*. 2005, 21, 2, s. 163–193. ISSN 1096-0031. doi: 10.1111/j.1096-0031.2005.00059.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1096-0031.2005.00059.x.

BIELAWSKI J., YANG Z. Notes for codon-based Clade models, 2004a. Dostupné z: https://github.com/abacus-gene/paml/blob/master/examples/CladeModelCD/PAMLCladeModelCDnotes.pdf. Navštíveno: 23. března 2025.

BIELAWSKI J. P., YANG Z., Maximum likelihood methods for detecting adaptive evolution after gene duplication. *Journal of Structural and Functional Genomics*. March 2003, 3, 1, s. 201–212. ISSN 1570-0267. doi: 10.1023/A:1022642807731.

BIELAWSKI J. P., YANG Z., A Maximum Likelihood Method for Detecting Functional Divergence at Individual Codon Sites, with Application to Gene Family Evolution. *Journal of Molecular Evolution*. July 2004b, 59, 1, s. 121–132. ISSN 1432-1432. doi: 10.1007/s00239-004-2597-8.

BONDAREVA O. V., POTAPOVA N. A., KONOVALOV K. A., PETROVA T. V., ABRAMSON N. I., Searching for signatures of positive selection in cytochrome b gene associated with subterranean lifestyle in fast-evolving arvicolines (Arvicolinae, Cricetidae, Rodentia). *BMC Ecology and Evolution*. May 2021, 21, 1, s. 92. ISSN 2730-7182. doi: 10.1186/s12862-021-01819-4.

BRADLEY R. D., DURISH N. D., ROGERS D. S., MILLER J. R., ENGSTROM M. D., KILPATRICK C. W., Toward a Molecular Phylogeny for Peromyscus: Evidence from Mitochondrial Cytochrome-b Sequences. *Journal of Mammalogy*. October 2007, 88, 5, s. 1146–1159. ISSN 0022-2372. doi: 10.1644/06-MAMM-A-342R.1.

BRETONES G., DELGADO M. D., LEóN J., Myc and cell cycle control. *Biochimica Et Biophysica Acta*. May 2015, 1849, 5, s. 506–516. ISSN 0006-3002. doi: 10.1016/j.bbagrm.2014.03.013.

BULMER M., The selection-mutation-drift theory of synonymous codon usage. *Genetics*. November 1991, 129, 3, s. 897–907. ISSN 1943-2631. doi: 10.1093/genetics/129.3.897.

CAPELLA-GUTIéRREZ S., SILLA-MARTíNEZ J. M., GABALDóN T., trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*. August 2009, 25, 15, s. 1972–1973. ISSN 1367-4803. doi: 10.1093/bioinformatics/btp348.

CAROLI A. M., CHESSA S., ERHARDT G. J., Invited review: Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. *Journal of Dairy Science*. November 2009, 92, 11, s. 5335–5352. ISSN 0022-0302. doi: 10.3168/jds.2009-2461. Publisher: Elsevier.

CASTOE T. A., KONING A. P. J., KIM H.-M., GU W., NOONAN B. P., NAYLOR G., JIANG Z. J., PARKINSON C. L., POLLOCK D. D., Evidence for an ancient adaptive episode of convergent molecular evolution. *Proceedings of the National Academy of Sciences*. June 2009, 106, 22, s. 8986–8991. doi: 10.1073/pnas.0900233106. Publisher: Proceedings of the National Academy of Sciences.

CHAKRABARTI A., GHOSH J. K. AIC, BIC and Recent Advances in Model Selection. In BANDYOPADHYAY P. S., FORSTER M. R. (Ed.) *Philosophy of Statistics*, 7 / *Handbook of the Philosophy of Science*. Amsterdam: North-Holland, January 2011. s. 583–605. doi: 10.1016/B978-0-444-51862-0.50018-6.

CHERNOMOR O., HAESELER A., MINH B. Q., Terrace Aware Data Structure for Phylogenomic Inference from Supermatrices. *Systematic Biology*. November 2016, 65, 6, s. 997–1008. ISSN 1063-5157. doi: 10.1093/sysbio/syw037.

CHIKUNI K., MORI Y., TABATA T., SAITO M., MONMA M., KOSUGIYAMA M., Molecular phylogeny based on the k-casein and cytochrome b sequences in the mammalian suborder Ruminantia. *Journal of Molecular Evolution*. December 1995, 41, 6, s. 859–866. ISSN 1432-1432. doi: 10.1007/BF00173165.

COCK P. J. A., ANTAO T., CHANG J. T., CHAPMAN B. A., COX C. J., DALKE A., FRIEDBERG I., HAMELRYCK T., KAUFF F., WILCZYNSKI B., HOON M. J. L., Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*. March 2009, 25, 11, s. 1422–1423. ISSN 1367-4803. doi: 10.1093/bioinformatics/btp163. _eprint: https://academic.oup.com/bioinformatics/article-pdf/25/11/1422/48989335/bioinformatics_25_11_1422.pdf.

COMIN A., CASSANDRO M., CHESSA S., OJALA M., DAL ZOTTO R., DE MARCHI M., CARNIER P., GALLO L., PAGNACCO G., BITTANTE G., Effects of Composite β- and κ-Casein Genotypes on Milk Coagulation, Quality, and Yield Traits in Italian Holstein Cows. *Journal of Dairy Science*. October 2008, 91, 10, s. 4022–4027. ISSN 0022-0302. doi: 10.3168/jds.2007-0546.

CRONIN M. A., STUART R., PIERSON B. J., PATTON J. C., K-Casein Gene Phylogeny of Higher Ruminants (Pecora, Artiodactyla). *Molecular Phylogenetics and Evolution*. October 1996, 6, 2, s. 295–311. ISSN 1055-7903. doi: 10.1006/mpev.1996.0078.

CRUICKSHANK R. H., Molecular markers for the phylogenetics of mites and ticks. *Systematic and Applied Acarology*. July 2002, s. 3–14. ISSN 2056-6069. doi: 10.11158/saa.7.1.1.

DI FRANCO A., POUJOL R., BAURAIN D., PHILIPPE H., Evaluating the usefulness of alignment filtering methods to reduce the impact of errors on evolutionary inferences.

BMC Evolutionary Biology. January 2019, 19, 1, s. 21. ISSN 1471-2148. doi: 10.1186/s12862-019-1350-2.

DIAS B. C., NERY M. F., Analyses of RAG1 and RAG2 genes suggest different evolutionary rates in the Cetacea lineage. *Molecular Immunology*. January 2020, 117, s. 131–138. ISSN 0161-5890. doi: 10.1016/j.molimm.2019.10.014.

EDWARDS S. V., Natural selection and phylogenetic analysis. *Proceedings of the National Academy of Sciences*. June 2009, 106, 22, s. 8799–8800. doi: 10.1073/pnas.0904103106. Publisher: Proceedings of the National Academy of Sciences.

FELSENSTEIN J., *Inferring phylogenies*. Sunderland, Mass : Sinauer Associates, 2004. ISBN 978-0-87893-177-4.

FOLEY N. M., MASON V. C., HARRIS A. J., BREDEMEYER K. R., DAMAS J., LEWIN H. A., EIZIRIK E., GATESY J., KARLSSON E. K., LINDBLAD-TOH K., Zoonomia Consortium, SPRINGER M. S., MURPHY W. J., A genomic timescale for placental mammal evolution. *Science*. April 2023, 380, 6643, s. eabl8189. doi: 10.1126/science.abl8189. Publisher: American Association for the Advancement of Science.

GADAGKAR S. R., ROSENBERG M. S., KUMAR S., Inferring species phylogenies from multiple genes: Concatenated sequence tree versus consensus gene tree. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*. 2005, 304B, 1, s. 64–74. ISSN 1552-5015. doi: 10.1002/jez.b.21026. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jez.b.21026.

GALEWSKI T., TILAK M.-k., SANCHEZ S., CHEVRET P., PARADIS E., DOUZERY E. J., The evolutionary radiation of Arvicolinae rodents (voles and lemmings): relative contribution of nuclear and mitochondrial DNA phylogenies. *BMC Evolutionary Biology*. October 2006, 6, 1, s. 80. ISSN 1471-2148. doi: 10.1186/1471-2148-6-80.

GATESY J., HAYASHI C., CRONIN M. A., ARCTANDER P., Evidence from milk casein genes that cetaceans are close relatives of hippopotamid artiodactyls. *Molecular Biology and Evolution*. September 1996, 13, 7, s. 954–963. ISSN 0737-4038. doi: 10.1093/oxfordjournals.molbev.a025663.

GAUBERT P., ANTUNES A., MENG H., MIAO L., PEIGNé S., JUSTY F., NJIOKOU F., DUFOUR S., DANQUAH E., ALAHAKOON J., VERHEYEN E., STANLEY W. T., O'BRIEN S. J., JOHNSON W. E., LUO S.-J., The Complete Phylogeny of Pangolins: Scaling Up Resources for the Molecular Tracing of the Most Trafficked Mammals on Earth. *Journal of Heredity*. May 2018, 109, 4, s. 347–359. ISSN 0022-1503. doi: 10.1093/jhered/esx097.

GE²POP, *MACSE documentation* [online]. 2025. Navštíveno: 25. března 2025. Dostupné z: https://www.agap-ge2pop.org/macse/.

GRANTHAM R., GAUTIER C., GOUY M., JACOBZONE M., MERCIER R., Codon catalog usage is a genome strategy modulated for gene expressivity. *Nucleic Acids Research*. January 1981, 9, 1, s. r43–74. ISSN 0305-1048. doi: 10.1093/nar/9.1.213-b.

GREGORY T. R., Understanding Natural Selection: Essential Concepts and Common Misconceptions. *Evolution: Education and Outreach*. June 2009, 2, 2, s. 156–175. ISSN 1936-6434. doi: 10.1007/s12052-009-0128-1. Number: 2 Publisher: BioMed Central.

GRUBER K. F., VOSS R. S., JANSA S. A., Base-compositional heterogeneity in the RAG1 locus among didelphid marsupials: implications for phylogenetic inference and the evolution of GC content. *Systematic Biology*. February 2007, 56, 1, s. 83–96. ISSN 1063-5157. doi: 10.1080/10635150601182939.

HOANG D. T., CHERNOMOR O., HAESELER A., MINH B. Q., VINH L. S., UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*. February 2018, 35, 2, s. 518–522. ISSN 0737-4038. doi: 10.1093/molbev/msx281.

HOOFER S. R., REEDER S. A., HANSEN E. W., VAN DEN BUSSCHE R. A., MOLECULAR PHYLOGENETICS AND TAXONOMIC REVIEW OF NOCTILIONOID AND VESPERTILIONOID BATS (CHIROPTERA: YANGOCHIROPTERA). *Journal of Mammalogy*. August 2003, 84, 3, s. 809–821. ISSN 0022-2372, 1545-1542. doi: 10.1644/BWG-034.

HUERTA-CEPAS J., SERRA F., BORK P., ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Molecular Biology and Evolution*. June 2016, 33, 6, s. 1635–1638. ISSN 0737-4038. doi: 10.1093/molbev/msw046.

HUGHES A. L., PACKER B., WELCH R., BERGEN A. W., CHANOCK S. J., YEAGER M., Widespread purifying selection at polymorphic sites in human protein-coding loci. *Proceedings of the National Academy of Sciences*. December 2003, 100, 26, s. 15754–15757. doi: 10.1073/pnas.2536718100. Publisher: Proceedings of the National Academy of Sciences.

HURST L. D., The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends in Genetics*. September 2002, 18, 9, s. 486–487. ISSN 0168-9525. doi: 10.1016/S0168-9525(02)02722-1. Publisher: Elsevier.

ISHENGOMA E., AGABA M., Evolution of toll-like receptors in the context of terrestrial ungulates and cetaceans diversification. *BMC Evolutionary Biology*. February 2017, 17, 1, s. 54. ISSN 1471-2148. doi: 10.1186/s12862-017-0901-7.

JANSA S. A., WEKSLER M., Phylogeny of muroid rodents: relationships within and among major lineages as determined by IRBP gene sequences. *Molecular Phylogenetics and Evolution*. April 2004, 31, 1, s. 256–276. ISSN 1055-7903. doi: 10.1016/j.ympev.2003.07.002.

JANSA S. A., FORSMAN J. F., VOSS R. S., Different patterns of selection on the nuclear genes IRBP and DMP-1 affect the efficiency but not the outcome of phylogeny estimation for didelphid marsupials. *Molecular Phylogenetics and Evolution*. February 2006, 38, 2, s. 363–380. ISSN 1055-7903. doi: 10.1016/j.ympev.2005.06.007.

JORDAN G., GOLDMAN N., The Effects of Alignment Error and Alignment Filtering on the Sitewise Detection of Positive Selection. *Molecular Biology and Evolution*. April 2012, 29, 4, s. 1125–1139. ISSN 0737-4038. doi: 10.1093/molbev/msr272.

KALYAANAMOORTHY S., MINH B. Q., WONG T. K. F., HAESELER A., JERMIIN L. S., ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*. June 2017, 14, 6, s. 587–589. ISSN 1548-7105. doi: 10.1038/nmeth.4285. Publisher: Nature Publishing Group.

KAPLI P., YANG Z., TELFORD M. J., Phylogenetic tree building in the genomic age. *Nature Reviews Genetics*. July 2020, 21, 7, s. 428–444. ISSN 1471-0064. doi: 10.1038/s41576-020-0233-0. Publisher: Nature Publishing Group.

KHAN M. M. G., RYDéN A.-M., CHOWDHURY M. S., HASAN M. A., KAZI J. U., Maximum likelihood analysis of mammalian p53 indicates the presence of positively selected sites and higher tumorigenic mutations in purifying sites. *Gene*. September 2011, 483, 1, s. 29–35. ISSN 0378-1119. doi: 10.1016/j.gene.2011.05.011.

KIM Y., NIELSEN R., Linkage Disequilibrium as a Signature of Selective Sweeps. *Genetics*. July 2004, 167, 3, s. 1513–1524. ISSN 1943-2631. doi: 10.1534/genetics.103.025387.

KO W.-Y., PIAO S., AKASHI H., Strong regional heterogeneity in base composition evolution on the Drosophila X chromosome. *Genetics*. September 2006, 174, 1, s. 349–362. ISSN 0016-6731. doi: 10.1534/genetics.105.054346.

KOEPFLI K.-P., DEERE K. A., SLATER G. J., BEGG C., BEGG K., GRASSMAN L., LUCHERINI M., VERON G., WAYNE R. K., Multigene phylogeny of the Mustelidae: Resolving relationships, tempo and biogeographic history of a mammalian adaptive radiation. *BMC Biology*. December 2008, 6, 1, s. 10. ISSN 1741-7007. doi: 10.1186/1741-7007-6-10.

KUMAR S., FILIPSKI A., Multiple sequence alignment: In pursuit of homologous DNA positions. *Genome Research*. February 2007, 17, 2, s. 127–135. ISSN 1088-9051, 1549-5469. doi: 10.1101/gr.5232407. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

LAHTI D. C., JOHNSON N. A., AJIE B. C., OTTO S. P., HENDRY A. P., BLUMSTEIN D. T., COSS R. G., DONOHUE K., FOSTER S. A., Relaxed selection in the wild. *Trends in Ecology & Evolution*. September 2009, 24, 9, s. 487–496. ISSN 0169-5347. doi: 10.1016/j.tree.2009.03.010. Publisher: Elsevier.

LARSSON A., AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*. November 2014, 30, 22, s. 3276–3278. ISSN 1367-4803. doi: 10.1093/bioinformatics/btu531.

LIÓ P., GOLDMAN N., Models of Molecular Evolution and Phylogeny. *Genome Research*. December 1998, 8, 12, s. 1233–1244. ISSN 1088-9051, 1549-5469. doi: 10.1101/gr.8.12.1233. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

LIU F.-G. R., MIYAMOTO M. M., FREIRE N. P., ONG P. Q., TENNANT M. R., YOUNG T. S., GUGEL K. F., Molecular and Morphological Supertrees for Eutherian (Placental) Mammals. *Science*. March 2001. doi: 10.1126/science.1056346. Publisher: American Association for the Advancement of Science.

MADSEN O., SCALLY M., DOUADY C. J., KAO D. J., DEBRY R. W., ADKINS R., AMRINE H. M., STANHOPE M. J., DE JONG W. W., SPRINGER M. S., Parallel adaptive radiations in two major clades of placental mammals. *Nature*. February 2001, 409, 6820, s. 610–614. ISSN 0028-0836, 1476-4687. doi: 10.1038/35054544.

MARTIN Y., GERLACH G., SCHLöTTERER C., MEYER A., Molecular Phylogeny of European Muroid Rodents Based on Complete Cytochrome b Sequences. *Molecular Phylogenetics and Evolution*. 2000, 16, 1, s. 37–47. ISSN 1055-7903. doi: https://doi.org/10.1006/mpev.1999.0760.

MASSEY S. E., CHURBANOV A., RASTOGI S., LIBERLES D. A., Characterizing positive and negative selection and their phylogenetic effects. *Gene*. July 2008, 418, 1, s. 22–26. ISSN 0378-1119. doi: 10.1016/j.gene.2008.03.017.

MCVEAN G., The Structure of Linkage Disequilibrium Around a Selective Sweep. *Genetics*. March 2007, 175, 3, s. 1395–1406. ISSN 0016-6731. doi: 10.1534/genetics.106.062828.

MEREDITH R. W., WESTERMAN M., CASE J. A., SPRINGER M. S., A Phylogeny and Timescale for Marsupial Evolution Based on Sequences for Five Nuclear Genes. *Journal of Mammalian Evolution*. March 2008, 15, 1, s. 1–36. ISSN 1573-7055. doi: 10.1007/s10914-007-9062-6.

MICHAUX J., REYES A., CATZEFLIS F., Evolutionary History of the Most Speciose Mammals: Molecular Phylogeny of Muroid Rodents. *Molecular Biology and Evolution*. November 2001, 18, 11, s. 2017–2031. ISSN 0737-4038. doi: 10.1093/oxfordjournals.molbev.a003743.

MINH B. Q., SCHMIDT H. A., CHERNOMOR O., SCHREMPF D., WOODHAMS M. D., HAESELER A., LANFEAR R., IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*. May 2020, 37, 5, s. 1530–1534. ISSN 0737-4038. doi: 10.1093/molbev/msaa015.

MIYAMOTO M. M., PORTER C. A., GOODMAN M., c-Myc Gene Sequences and the Phylogeny of Bats and Other Eutherian Mammals. *Systematic Biology*. September 2000, 49, 3, s. 501–514. ISSN 1076-836X, 1063-5157. doi: 10.1080/10635159950127367.

MIYATA T., YASUNAGA T., Molecular evolution of mRNA: A method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application. *Journal of Molecular Evolution*. March 1980, 16, 1, s. 23–36. ISSN 0022-2844, 1432-1432. doi: 10.1007/BF01732067.

MURPHY W. J., EIZIRIK E., O'BRIEN S. J., MADSEN O., SCALLY M., DOUADY C. J., TEELING E., RYDER O. A., STANHOPE M. J., JONG W. W., SPRINGER M. S., Resolution of the Early Placental Mammal Radiation Using Bayesian Phylogenetics. *Science*. December 2001, 294, 5550, s. 2348–2351. doi: 10.1126/science.1067179. Publisher: American Association for the Advancement of Science.

NASER-KHDOUR S., MINH B. Q., ZHANG W., STONE E. A., LANFEAR R., The Prevalence and Impact of Model Violations in Phylogenetic Analysis. *Genome Biology and Evolution*. December 2019, 11, 12, s. 3341–3352. ISSN 1759-6653. doi: 10.1093/gbe/evz193.

NEI M., GOJOBORI T., Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*. September 1986, 3, 5, s. 418–426. ISSN 0737-4038. doi: 10.1093/oxfordjournals.molbev.a040410.

NYLANDER J. A. A., *catfasta2phyml – Concatenate FASTA alignments to PHYML, PHYLIP, or FASTA format* [online]. 2025. Navštíveno: 26. března 2025. Dostupné z: https://github.com/nylander/catfasta2phyml.

OJHA K. K., MISHRA S., SINGH V. K. *Chapter 5 - Computational molecular phylogeny: concepts and applications*, s. 67–89. Academic Press, January 2022. doi: 10.1016/B978-0-323-89775-4.00025-0. ISBN 978-0-323-89775-4.

O'LEARY N. A., WRIGHT M. W., BRISTER J. R., CIUFO S., HADDAD D., MCVEIGH R., RAJPUT B., ROBBERTSE B., SMITH-WHITE B., AKO-ADJEI D., ASTASHYN A., BADRETDIN A., BAO Y., BLINKOVA O., BROVER V., CHETVERNIN V., CHOI J., COX E., ERMOLAEVA O., FARRELL C. M., GOLDFARB T., GUPTA T., HAFT D., HATCHER E., HLAVINA W., JOARDAR V. S., KODALI V. K., LI W., MAGLOTT D., MASTERSON P., MCGARVEY K. M., MURPHY M. R., O'NEILL K., PUJAR S., RANGWALA S. H., RAUSCH D., RIDDICK L. D., SCHOCH C., SHKEDA A., STORZ S. S., SUN H., THIBAUD-NISSEN F., TOLSTOY I., TULLY R. E., VATSAN A. R., WALLIN C., WEBB D., WU W., LANDRUM M. J., KIMCHI A., TATUSOVA T., DICUCCIO M., KITTS P., MURPHY T. D., PRUITT K. D., Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*. November 2015, 44, D1, s. D733–D745. ISSN 0305-1048. doi: 10.1093/nar/gkv1189. _eprint: https://academic.oup.com/nar/article-pdf/44/D1/D733/9482930/gkv1189.pdf.

OSAWA S., OHAMA T., JUKES T. H., WATANABE K., Evolution of the mitochondrial genetic code I. Origin of AGR serine and stop codons in metazoan mitochondria. *Journal of Molecular Evolution*. September 1989, 29, 3, s. 202–207. ISSN 0022-2844, 1432-1432. doi: 10.1007/BF02100203.

PARADIS E., SCHLIEP K., ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*. 2019, 35, s. 526–528. doi: 10.1093/bioinformatics/bty633.

PARISI G., ECHAVE J., Structural constraints and emergence of sequence patterns in protein evolution. *Molecular Biology and Evolution*. May 2001, 18, 5, s. 750–756. ISSN 0737-4038. doi: 10.1093/oxfordjournals.molbev.a003857.

PATWARDHAN A., RAY S., ROY A., Molecular Markers in Phylogenetic Studies-A Review. *Journal of Phylogenetics & Evolutionary Biology*. 2014, 02, 02. ISSN 23299002. doi: 10.4172/2329-9002.1000131.

PHILIPPE H., LOPEZ P., BRINKMANN H., BUDIN K., GERMOT A., LAURENT J., MOREIRA D., MüLLER M., LE GUYADER H., Early-branching or fast-evolving eukaryotes? An answer based on slowly evolving positions. *Proceedings of the Royal Society B: Biological Sciences*. June 2000, 267, 1449, s. 1213–1221. ISSN 0962-8452.

PIE M. R., The Influence of Phylogenetic Uncertainty on the Detection of Positive Darwinian Selection. *Molecular Biology and Evolution*. December 2006, 23, 12, s. 2274–2278. ISSN 0737-4038. doi: 10.1093/molbev/msl116.

PLOTKIN J. B., KUDLA G., Synonymous but not the same: the causes and consequences of codon bias. *Nature Reviews Genetics*. January 2011, 12, 1, s. 32–42. ISSN 1471-0064. doi: 10.1038/nrg2899. Publisher: Nature Publishing Group.

POPADIN K. Y., NIKOLAEV S. I., JUNIER T., BARANOVA M., ANTONARAKIS S. E., Purifying Selection in Mammalian Mitochondrial Protein-Coding Genes Is Highly Effective and Congruent with Evolution of Nuclear Genes. *Molecular Biology and Evolution*. February 2013, 30, 2, s. 347–355. ISSN 0737-4038. doi: 10.1093/molbev/mss219.

PORTER C. A., GOODMAN M., STANHOPE M. J., Evidence on Mammalian Phylogeny from Sequences of Exon 28 of the von Willebrand Factor Gene. *Molecular Phylogenetics and Evolution*. February 1996, 5, 1, s. 89–101. ISSN 1055-7903. doi: 10.1006/mpev.1996.0008.

PRENDERGAST G. C., Mechanisms of apoptosis by c-Myc. *Oncogene*. May 1999, 18, 19, s. 2967–2987. ISSN 1476-5594. doi: 10.1038/sj.onc.1202727. Publisher: Nature Publishing Group.

PRUITT K., MURPHY T., BROWN G., MURPHY M., *RefSeq Frequently Asked Questions (FAQ)* [online]. 2010. Navštíveno: 26. března 2025. Dostupné z: https://www.ncbi.nlm.nih.gov/books/NBK50679/.

RAMBAUT A. FigTree v1.4.4. Online, 2018. Dostupné z: http://tree.bio.ed.ac.uk/software/figtree/. Edinburgh: Institute of Evolutionary Biology, University of Edinburgh.

RANWEZ V., DELSUC F., RANWEZ S., BELKHIR K., TILAK M.-K., DOUZERY E. J., OrthoMaM: A database of orthologous genomic markers for placental mammal phylogenetics. *BMC Evolutionary Biology*. November 2007, 7, 1, s. 241. ISSN 1471-2148. doi: 10.1186/1471-2148-7-241.

RANWEZ V., DOUZERY E. J. P., CAMBON C., CHANTRET N., DELSUC F., MACSE v2: Toolkit for the Alignment of Coding Sequences Accounting for Frameshifts and Stop Codons. *Molecular Biology and Evolution*. October 2018, 35, 10, s. 2582–2584. ISSN 0737-4038. doi: 10.1093/molbev/msy159.

REEDER S. A., BRADLEY R. D., MOLECULAR SYSTEMATICS OF NEOTOMINE–PEROMYSCINE RODENTS BASED ON THE DENTIN MATRIX PROTEIN 1 GENE. *Journal of Mammalogy*. December 2004, 85, 6, s. 1194–1200. ISSN 0022-2372, 1545-1542. doi: 10.1644/BEL-105.1.

ROBINSON M., CATZEFLIS F., BRIOLAY J., MOUCHIROUD D., Molecular Phylogeny of Rodents, with Special Emphasis on Murids: Evidence from Nuclear Gene LCAT. *Molecular Phylogenetics and Evolution*. December 1997, 8, 3, s. 423–434. ISSN 1055-7903. doi: 10.1006/mpev.1997.0424.

ROBINSON-RECHAVI M., PONGER L., MOUCHIROUD D., Nuclear Gene LCAT Supports Rodent Monophyly. *Molecular Biology and Evolution*. September 2000, 17, 9, s. 1410–1412. ISSN 0737-4038. doi: 10.1093/oxfordjournals.molbev.a026424.

ROCHA E. P. C., DANCHIN A., Base composition bias might result from competition for metabolic resources. *Trends in genetics: TIG*. June 2002, 18, 6, s. 291–294. ISSN 0168-9525. doi: 10.1016/S0168-9525(02)02690-2.

ROJE D. M., Evaluating the Effects of Non-Neutral Molecular Markers on Phylogeny Inference. *PLOS ONE*. 2014, 9, 2, s. e87428. ISSN 1932-6203. doi: 10.1371/journal.pone.0087428. Publisher: Public Library of Science.

SALEMI M., VANDAMME A.-M., LEMEY P. (Ed.), *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge, UK ; New York : Cambridge University Press, 2nd ed edition, 2009. OCLC: ocn295002266. ISBN 978-0-521-87710-7 978-0-521-73071-6.

SANJAK J. S., SIDORENKO J., ROBINSON M. R., THORNTON K. R., VISSCHER P. M., Evidence of directional and stabilizing selection in contemporary humans. *Proceedings of the National Academy of Sciences*. January 2018, 115, 1, s. 151–156. doi: 10.1073/pnas.1707227114. Publisher: Proceedings of the National Academy of Sciences.

SANTODOMINGO A. M., THOMAS R. S., QUINTERO-GALVIS J. F., ECHEVERRY-BERRIO D. M., FUENTE M. C. S.-d., MORENO-SALAS L., MUÑOZ-LEAL S., Apicomplexans in small mammals from Chile, with the first report of the Babesia microti group in South American rodents. *Parasitology Research*. March 2022, 121, 3, s. 1009–1020. ISSN 1432-1955. doi: 10.1007/s00436-022-07452-4.

SAYERS E. W., BECK J., BOLTON E. E., BRISTER J. R., CHAN J., COMEAU D. C., CONNOR R., DICUCCIO M., FARRELL C. M., FELDGARDEN M., FINE A. M., FUNK K., HATCHER E., HOEPPNER M., KANE M., KANNAN S., KATZ K. S., KELLY C., KLIMKE W., KIM S., KIMCHI A., LANDRUM M., LATHROP S., LU Z., MALHEIRO A., MARCHLER-BAUER A., MURPHY T. D., PHAN L., PRASAD A. B., PUJAR S., SAWYER A., SCHMIEDER E., SCHNEIDER V. A., SCHOCH C. L., SHARMA S., THIBAUD-NISSEN F., TRAWICK B. W., VENKATAPATHI T., WANG J., PRUITT K. D., SHERRY S. T., Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*. January 2024, 52, D1, s. D33–D43. ISSN 1362-4962. doi: 10.1093/nar/gkad1044.

SENA L., VALLINOTO M., SAMPAIO I., SCHNEIDER H., FERRARI S. F., CRUZ SCHNEIDER M. P., Mitochondrial COII Gene Sequences Provide New Insights into the Phylogeny of Marmoset Species Groups (Callitrichidae, Primates). *Folia Primatologica*. 01 2003, 73, 5, s. 240–251. doi: 10.1159/000067456.

SHEKAR P. C., GOEL S., RANI S. D. S., SARATHI D. P., ALEX J. L., SINGH S., KUMAR S., κ-Casein-deficient mice fail to lactate. *Proceedings of the National Academy of Sciences*. May 2006, 103, 21, s. 8000–8005. doi: 10.1073/pnas.0601611103. Publisher: Proceedings of the National Academy of Sciences.

SPRINGER M. S., DEBRY R. W., DOUADY C., AMRINE H. M., MADSEN O., JONG W. W., STANHOPE M. J., Mitochondrial Versus Nuclear Gene Sequences in Deep-Level Mammalian Phylogeny Reconstruction. *Molecular Biology and Evolution*. February 2001, 18, 2, s. 132–143. ISSN 0737-4038. doi: 10.1093/oxfordjournals.molbev.a003787.

SPRINGER M. S., STANHOPE M. J., MADSEN O., JONG W. W. d., Molecules consolidate the placental mammal tree. *Trends in Ecology & Evolution*. August 2004, 19, 8, s. 430–438. ISSN 0169-5347. doi: 10.1016/j.tree.2004.05.006. Publisher: Elsevier.

STADELMANN B., LIN L. K., KUNZ T. H., RUEDI M., Molecular phylogeny of New World *Myotis* (Chiroptera, Vespertilionidae) inferred from mitochondrial and nuclear DNA genes. *Molecular Phylogenetics and Evolution*. April 2007, 43, 1, s. 32–48. ISSN 1055-7903. doi: 10.1016/j.ympev.2006.06.019.

STANHOPE M. J., CZELUSNIAK J., SI J.-S., NICKERSON J., GOODMAN M., A molecular perspective on mammalian evolution from the gene encoding interphotoreceptor retinoid binding protein, with convincing evidence for bat monophyly. *Molecular Phylogenetics and Evolution*. 1992, 1, 2, s. 148–160. ISSN 1055-7903. doi: https://doi.org/10.1016/1055-7903(92)90026-D.

STEPPAN S. J., ADKINS R. M., ANDERSON J., Phylogeny and Divergence-Date Estimates of Rapid Radiations in Muroid Rodents Based on Multiple Nuclear Genes. *Systematic Biology*. August 2004a, 53, 4, s. 533–553. ISSN 1063-5157. doi: 10.1080/10635150490468701.

STEPPAN S. J., STORZ B. L., HOFFMANN R. S., Nuclear DNA phylogeny of the squirrels (Mammalia: Rodentia) and the evolution of arboreality from c-myc and RAG1.

Molecular Phylogenetics and Evolution. March 2004b, 30, 3, s. 703–719. ISSN 1055-7903. doi: 10.1016/S1055-7903(03)00204-5.

STERN A., PUPKO T., An evolutionary space-time model with varying among-site dependencies. *Molecular Biology and Evolution*. February 2006, 23, 2, s. 392–400. ISSN 0737-4038. doi: 10.1093/molbev/msj044.

TAANMAN J.-W., The mitochondrial genome: structure, transcription, translation and replication. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*. February 1999, 1410, 2, s. 103–123. ISSN 0005-2728. doi: 10.1016/S0005-2728(98)00161-3.

TARALLO A., ANGELINI C., SANGES R., YAGI M., AGNISOLA C., D'ONOFRIO G., On the genome base composition of teleosts: the effect of environment and lifestyle. *BMC Genomics*. March 2016, 17, 1, s. 173. ISSN 1471-2164. doi: 10.1186/s12864-016-2537-1.

TAVARÉ S., Some Probabilistic and Statistical Problems in the Analysis of DNA Sequences. *Lectures on Mathematics in the Life Sciences*. 1986, 17.

TEELING E. C., SCALLY M., KAO D. J., ROMAGNOLI M. L., SPRINGER M. S., STANHOPE M. J., Molecular evidence regarding the origin of echolocation and flight in bats. *Nature*. January 2000, 403, 6766, s. 188–192. ISSN 0028-0836, 1476-4687. doi: 10.1038/35003188.

THEWISSEN J. G. M., COOPER L. N., GEORGE J. C., BAJPAI S., From Land to Water: the Origin of Whales, Dolphins, and Porpoises. *Evolution: Education and Outreach*. June 2009, 2, 2, s. 272–288. ISSN 1936-6426, 1936-6434. doi: 10.1007/s12052-009-0135-2.

THOMSON R. C., WANG I. J., JOHNSON J. R., Genome-enabled development of DNA markers for ecology, evolution and conservation. *Molecular Ecology*. 2010, 19, 11, s. 2184–2195. ISSN 1365-294X. doi: 10.1111/j.1365-294X.2010.04650.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-294X.2010.04650.x.

TOBE S. S., KITCHENER A. C., LINACRE A. M. T., Reconstructing Mammalian Phylogenies: A Detailed Comparison of the Cytochrome b and Cytochrome Oxidase Subunit I Mitochondrial Genes. *PLOS ONE*. 11 2010, 5, 11, s. 1–14. doi: 10.1371/journal.pone.0014156.

TOWNSEND J. P., Profiling Phylogenetic Informativeness. *Systematic Biology*. April 2007, 56, 2, s. 222–231. ISSN 1063-5157. doi: 10.1080/10635150701311362.

TOYOSAWA S., O'HUIGIN C., KLEIN J., The Dentin Matrix Protein 1 Gene of Prototherian and Metatherian Mammals. *Journal of Molecular Evolution*. February 1999, 48, 2, s. 160–167. ISSN 1432-1432. doi: 10.1007/PL00006454.

VAN DEN BUSSCHE R. A., REEDER S. A., HANSEN E. W., HOOFER S. R., Utility of the dentin matrix protein 1 (*DMP1*) gene for resolving mammalian intraordinal phylogenetic relationships. *Molecular Phylogenetics and Evolution*. January 2003, 26, 1, s. 89–101. ISSN 1055-7903. doi: 10.1016/S1055-7903(02)00297-X.

VIRTANEN P., GOMMERS R., OLIPHANT T. E., HABERLAND M., REDDY T., COURNAPEAU D., BUROVSKI E., PETERSON P., WECKESSER W., BRIGHT J., WALT S. J., BRETT M., WILSON J., MILLMAN K. J., MAYOROV N., NELSON A. R. J., JONES E., KERN R., LARSON E., CAREY C. J., POLAT İ., FENG Y., MOORE E. W., VANDERPLAS J., LAXALDE D., PERKTOLD J., CIMRMAN R., HENRIKSEN I., QUINTERO E. A., HARRIS C. R., ARCHIBALD A. M., RIBEIRO A. H., PEDREGOSA F., MULBREGT P., SciPy 1.0 Contributors, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*. 2020, 17, s. 261–272. doi: 10.1038/s41592-019-0686-2.

VOSS R. S., JANSA S. A., Phylogenetic Relationships and Classification of Didelphid Marsupials, an Extant Radiation of New World Metatherian Mammals. *Bulletin of the American Museum of Natural History*. June 2009, 2009, 322, s. 1–177. ISSN 0003-0090, 1937-3546. doi: 10.1206/322.1. Publisher: American Museum of Natural History.

WASSERSTEIN R. L., LAZAR N. A., The ASA Statement on p-Values: Context, Process, and Purpose. *The American Statistician*. April 2016, 70, 2, s. 129–133. ISSN 0003-1305. doi: 10.1080/00031305.2016.1154108. Publisher: ASA Website _eprint: https://doi.org/10.1080/00031305.2016.1154108.

WEADICK C. J., CHANG B. S., An Improved Likelihood Ratio Test for Detecting Site-Specific Functional Divergence among Clades of Protein-Coding Genes. *Molecular Biology and Evolution*. May 2012, 29, 5, s. 1297–1300. ISSN 0737-4038. doi: 10.1093/molbev/msr311.

WHELAN S., LIò P., GOLDMAN N., Molecular phylogenetics: state-of-the-art methods for looking into the past. *Trends in Genetics*. May 2001, 17, 5, s. 262–272. ISSN 0168-9525. doi: 10.1016/S0168-9525(01)02272-7. Publisher: Elsevier.

YANG Z., PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution*. April 2007, 24, 8, s. 1586–1591. ISSN 0737-4038, 1537-1719. doi: 10.1093/molbev/msm088.

YANG Z., BIELAWSKI J. P., Statistical methods for detecting molecular adaptation. *Trends in Ecology & Evolution*. December 2000, 15, 12, s. 496–503. ISSN 1872-8383. doi: 10.1016/s0169-5347(00)01994-7.

YANG Z., NIELSEN R., GOLDMAN N., PEDERSEN A. M., Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics*. May 2000, 155, 1, s. 431–449. ISSN 0016-6731.

YANG Z. PAML FAQ, 2005. Dostupné z: http://abacus.gene.ucl.ac.uk/software/pamlFAQs.pdf. Navštíveno: 23. března 2025.

YANG Z. PAML: Phylogenetic Analysis by Maximum Likelihood; User Guide v. 4.9j, 2020. Dostupné z: http://abacus.gene.ucl.ac.uk/software/pamlDOC.pdf. Navštíveno: 23. března 2025.

YANG Z., NIELSEN R., Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *Journal of Molecular Evolution*. April 1998, 46, 4, s. 409–418. ISSN 1432-1432. doi: 10.1007/PL00006320.

YOUNG A. D., GILLUNG J. P., Phylogenomics — principles, opportunities and pitfalls of big-data phylogenetics. *Systematic Entomology*. 2020, 45, 2, s. 225–247. ISSN 1365-3113. doi: 10.1111/syen.12406. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/syen.12406.

ZHANG J., Evaluation of an Improved Branch-Site Likelihood Method for Detecting Positive Selection at the Molecular Level. *Molecular Biology and Evolution*. August 2005, 22, 12, s. 2472–2479. ISSN 0737-4038, 1537-1719. doi: 10.1093/molbev/msi237.