

UNIVERZA V LJUBLJANI
BIOTEHNIŠKA FAKULTETA
ODDELEK ZA BIOLOGIJO

Klement STOJANOVSKI

**Zmanjševanje konflikta taksonomskih znakov z uporabo a-priori
filogenetske informacije**

DIPLOMSKO DELO

Univerzitetni študij

Minimizing character conflict using a-priori phylogenetic information

GRADUATION THESIS

University studies

Ljubljana, 2007

Diplomsko delo sem opravil na Katedri za zoologijo na Oddelku za biologijo Biotehniške fakultete v Ljubljani

Študijska komisija Oddelka za biologijo je za mentorja diplomskega dela imenovala prof. dr. Petra Trontlja

Komisija za oceno in zagovor:

Predsednik: prof. dr. Boris SKET
Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za biologijo

Član: prof. dr. Peter TRONTELJ
Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za biologijo

Član: doc. dr. Rudi VEROVNIK
Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za biologijo

Datum zagovora:

Naloga je rezultat lastnega raziskovalnega dela.

Podpisani se strinjam z objavo svoje naloge v polnem tekstu na spletni strani Digitalne knjižnice Biotehniške fakultete. Izjavljam, da je naloga, ki sem jo oddal v elektronski obliki, identična tiskani verziji.

Klement Stojanovski

KLJUČNA DOKUMENTACIJSKA INFORMACIJA (KDI)

- ŠD Dn
- DK 575.8:577.2(043.2)=863
- KG filogenetika/koflikt znakov/obteževanje znakov/apriorna informacija
- AV STOJANOVSKI, Klement
- SA TRONTELJ, Peter
- KZ SI-1000 Ljubljana, Jamnikarjeva 101
- ZA Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za biologijo
- LI 2007
- IN ZMANJŠEVANJE KONFLIKTA TAKSONOMSKIH ZNAKOV Z UPORABO A-PRIORI FILOGENETSKE INFORMACIJE
- TD Diplomsko delo (univerzitetni študij)
- OP VIII, 64 str., 4 pregl., 13 sl., 3 pril., 105 vir.
- IJ sl
- JI sl/en
- AI V diplomskem delu predstavljamo novo metodo za zmanjševanje konflikta taksonomskih znakov. V nasprotju z mnogimi metodami obteževanja znakov, s to metodo poskušamo razjasniti sorodstvene odnose ene same problematične skupine v kladogramu s sicer dobro razrešenimi sorodstvenimi odnosi. Znane sorodstvene odnose na nižjih hierarhičnih ravneh uporabimo kot apriorno filogenetsko informacijo za vrednotenje znakov. Na podlagi teh odnosov določimo znake, ki so verjetno zavajajoči na ravni preučevanega razvejišča. Problematične so spremembe stanj znakov vzdolž dolgih vej, ki so pogosto posledica velike hitrosti substitucij. Poleg tega z uporabo apriorne sorodstvene informacije lahko razlikujemo med filogenetskim in drugimi zgodovinskimi signali v podatkih. Problematične znake določimo s pomočjo kvartetov v katerih sta dva taksona pripadnika iste monofiletske skupine. Z izločanjem znakov v korakih zmanjšujemo konflikt znakov na hierarhični ravni problema. Po vsakem izločanju naredimo sorodstveno analizo in ugotovimo podporo (bootstrap, Bremerjeva) nasprotujočim hipotezam. Celokupno podporo posamezni hipotezi sorodstvenih odnosov definiramo kot maksimalno podporo, ki jo je ta dobila v celotnem postopku izločanja. Metodo smo preizkusili na treh primerih: položaju glavačev (Physeteridae) znotraj kitov (Cetacea), položaju skakačev (Collembola) znotraj členonožcev (Arthropoda) in sestrski skupini hoacina (*Opisthocomus hoazin*, Aves). V vseh treh primerih je postopno izločanje vodilo v povečanje podpore le eni sorodstveni hipotezi. Iz tega lahko sklepamo, da z izločanjem problematičnih znakov razkrijemo filogenetsko informacijo v podatkih.

KEY WORDS DOCUMENTATION (KWD)

- DN Dn
- DC 575.8:577.2(043.2)=863
- CX phylogenetics/character conflict/character weighting/ a-priori information
- AU STOJANOVSKI, Klement
- AA TRONTELJ, Peter
- PP SI-1000 Ljubljana, Jamnikarjeva 101
- PB University of Ljubljana, Biotechnical Faculty, Department of Biology
- PY 2007
- TI MINIMIZING CHARACTER CONFLICT USING A-PRIORI PHYLOGENETIC INFORMATION
- DT Graduation Thesis (University studies)
- NO VIII, 64 p., 4 tab., 13 fig., 3 ann., 105 ref.
- LA sl
- AL sl/en
- AB A new method for minimizing character conflict is proposed. Contrary to most previously proposed methods of character weighting, this method focuses on a single problematic taxon in an otherwise well resolved cladogram. "Well supported" groups at lower hierarchical levels are used as a priori information for identification of phylogenetically uninformative characters. Identified problematic characters are mostly those that change along long branches, usually as a consequence of high substitution rates. Furthermore, by using a priori information we can differentiate between phylogenetic and other historic signals. Characters are analysed in quartets, two taxa being members of a single monophyletic group. Stepwise character exclusion is used to minimize character conflict at the problematic node. Each character exclusion step is followed by phylogenetic analysis, and support (bootstrap, Bremer) for competing hypotheses is assessed. The overall support for each hypothesis is defined as the highest support it receives over the entire exclusion procedure. The proposed method was tested on three examples: phylogenetic position of sperm whales (Physeteridae) within whales (Cetacea), position of springtails (Collembola) within arthropods, and sister group of the enigmatic bird hoatzin (*Opisthocomus hoazin*, Aves). Gradual exclusion of identified problematic characters led to consistently increasing support of a single hypothesis in all three examples. These results support our presumption that the proposed method recovers hidden phylogenetic information in the data.

KAZALO VSEBINE

	Str.
Ključna dokumentacijska informacija	<i>iii</i>
Key words documentation	<i>iv</i>
Kazalo vsebine	<i>v</i>
Kazalo preglednic	<i>vii</i>
Kazalo slik	<i>vii</i>
Kazalo prilog	<i>viii</i>
Okrajšave in simboli	<i>ix</i>
Slovarček	<i>x</i>
1 UVOD	1
1.1 OPREDELITEV PROBLEMA	3
1.2 OBTEŽEVANJE ZNAKOV	5
1.3 IZBRANI PRIMERI	7
1.3.1 Položaj glavačev (<i>Physeteridae</i>) v kladu kiti (<i>Cetacea</i>)	7
1.3.2 Položaj skakačev (<i>Collembola</i>, <i>Arthropoda</i>)	9
1.3.3 Sestrška skupina hoacina (<i>Opisthocomus hoazin</i>, <i>Aves</i>)	13
2 MATERIALI IN METODE	15
2.1 MATERIALI	15
2.1.1 Zaporedja DNA	15
2.1.1.1 Glavači	15
2.1.1.2 Skakači	15
2.1.1.3 Hoacin	15
2.1.2 Uporabljeni taksoni	16
2.1.2.1 Glavači	16
2.1.2.2 Skakači	16
2.1.2.3 Hoacin	16
2.2 METODE	21
2.2.1 Uskladitev zaporedij	21

2.2.1.1	Glavači	21
2.2.1.2	Skakači	21
2.2.1.3	Hoacin	22
2.2.2	Analiza znakov	22
2.2.2.1	Glavači	23
2.2.2.2	Skakači	23
2.2.2.3	Hoacin	24
2.2.2.4	Analiza izločenih znakov po indeksu homoplastičnosti (HI)	24
2.2.3	Filogenetska analiza	24
3	REZULTATI	25
3.1	METODA	25
3.2	PRIMERI	28
3.2.1	Glavači	28
3.2.2	Skakači	31
3.2.3	Hoacin	36
3.2.4	Analiza izločenih znakov po indeksu homoplastičnosti (glavači)	36
4	RAZPRAVA	39
4.1	METODA	39
4.2	PRIMERI	44
4.2.1	Glavači	44
4.2.2	Skakači	45
4.2.3	Hoacin	48
5	SKLEPI	50
6	POVZETEK	51
7	VIRI	53
	ZAHVALA	
	PRILOGE	

KAZALO PREGLEDNIC

Tabela 1:	Glavači. Taksoni uporabljeni v filogenetski analizi.	18
Tabela 2:	Glavači. Monofiletske skupine uporabljene v analizi znakov.	18
Tabela 3:	Skakači. Taksoni uporabljeni v analizi znakov (v privzetih monofilijah) in filogenetski analizi. V analizo znakov pripadnika skupine Collembola nista bila vključena.	19
Tabela 4:	Uporabljeni taksoni v primeru hoacin. Taksonomska razvrstitev po Sorenson et al. (2003).	20

KAZALO SLIK

Slika 1:	Filogram. A, B, C, D, E, F so (terminalni) taksoni. Dolžine vej predstavljajo število substitucij (sprememb stanj znakov). Merilo kaže dolžino veje, ki ustreza eni spremembi. 1., 2., 3. so različne...	1
Slika 2:	Primer privlaka dolgih vej na nekoreninjenem drevesu s štirimi taksoni (kvartet). Levo drevo predstavlja dejansko topologijo, desno pa topologijo, ki jo izbere metoda največjega verjetja...	5
Slika 3	Shema metode. M, T: izbrane monofiletske skupine in taksoni (glej predpostavke); K_i : i-ti kvartet; H_i : konfliktne znaki v i-tem kvartetu; delež homoplastičnosti: nasprotovanje znakov...	26
Slika 4:	Optimalno drevo (filogram) kito-kopitarjev po kriteriju varčnosti, brez izločanja. V nasprotju s apriorno hipotezo je induška pliskavka (<i>Platanista minor</i>) postavljena kot bazalni takson...	29
Slika 5:	Grafikon spreminjanja bootstrap podpore (BS, vrednosti v odstotkih) alternativnima sestrskima skupinama glavačev v odvisnosti od izločanja znakov. Odontoceti – glavači del...	30
Slika 6:	Grafikon spreminjanja umerjene Bremerjeve podpore (Ubp) alternativnima sestrskima skupinama glavačev v odvisnosti od izločanja znakov. Odontoceti - glavači del monofiluma zobati....	30

Slika 7:	Drevo z najvišjo podporo preizkušani hipotezi (glavači kot bazalna skupina zobatih kitov) v postopku izločanja. Izločeni znaki, katerih delež homoplastičnosti je večji od 0,044. Zgornje število...	31
Slika 8:	Najboljše drevo po kriteriju največje varčnosti v primeru skakačev (Collembola), brez izločanja znakov. Podpora razvejišču je napisana nad vejo, ki vodi k njemu. Zgornje število je umerjena...	33
Slika 9:	Grafikon spreminjanja bootstrap podpore (BS, vrednosti v odstotkih) alternativnim položajem skakačev (Collembola) v kladogramu v postopku izločanja. Skakači kot del monofiluma...	34
Slika 10:	Grafikon spreminjanja umerjene Bremerjeve podpore (Ubp) alternativnim položajem skakačev (Collembola) v kladogramu v postopku izločanja. Skakači kot del monofiluma Hexapoda oz...	34
Slika 11:	Najboljše drevo v primeru skakačev (Collembola) po kriteriju največje varčnosti v petem koraku izločanja, izločeni znaki z deležem homoplastičnosti večjim od 0,135. Podpora razvejišču...	35
Slika 12:	Grafikon spreminjanja bootstrap podpore (BS) alternativnim sestrskim skupinam hoacina (<i>Opisthocomus hoazin</i>) v kladogramu v postopku izločanja. Cleu – golob, Tcur – trogon, Falco - sokol...	36
Slika 13a-f:	Razporeditev izločenih znakov po razredih HI. Izločeni znaki so prikazani kot delež vseh znakov v nekem razredu: a) znaki izločeni v prvem koraku. Znaki z največjim deležem homoplastičnosti...	37-38

KAZALO PRILOG

Priloga A:	Primer vhodne datoteke za program za tvorbo kvartetov (primer skakači).
Priloga B:	Primer vhodne datoteke za analizo znakov na podlagi kvartetov (v programu PAUP*). Primer skakači.
Priloga C:	Vhodna datoteka (za program PAUP*) za zaporedno izločanje znakov. Primer skakači.

OKRAJŠAVE IN SIMBOLI

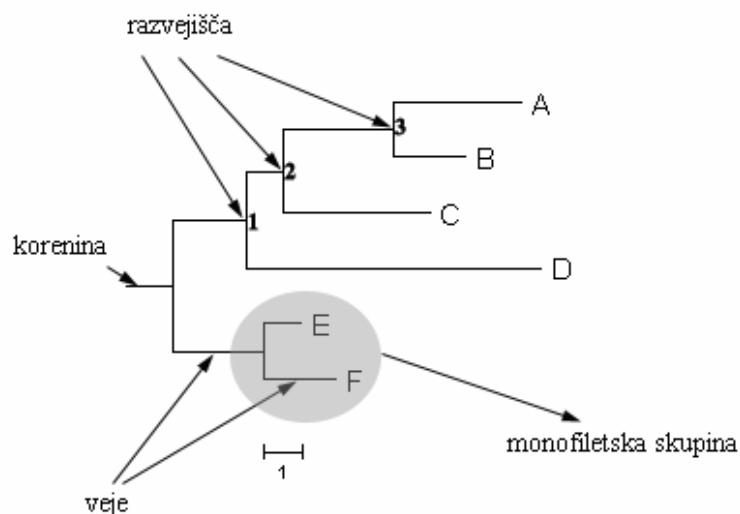
»Myriapoda«	skupina ni monofiletska
BS	(bootstrap support) bootstrap podpora
COI	citokrom oksidaza I
DNA	deoksiribonukleinska kislina
EF-1 α , EF-2	elongacijski faktor 1 α oz. 2
HI	(homoplasy index) indeks homoplastičnosti
INDEL	insercija ali delecija na določenem mestu v nukleotidnem zaporedju
LogDet	LogDet-paralinearne razdalje (LogDet-Paralinear distances)
mtDNA	mitohondrijska DNA. DNA, ki se nahaja v mitohondriju
NADH dehidrogenaza	nikotin amid dinukleotid dehidrogenaza
NADH5 oz. 6	podenota 5 (6) NADH dehidrogenaze
PEPCK	fosfoenolpiruvat karboksikinaza
PolIII	polimeraza II
RNA	ribonukleinska kislina
rRNA	ribosomska RNA
SINE, LINE	(short/long interspersed nuclear elements) kratki/dolgi razpršeni jedrni element
tRNA	transportna RNA
Ubp	umerjena Bremerjeva podpora

SLOVARČEK

Bayesov pristop	metoda v filogenetski analizi, ki vrednoti hipoteze na podlagi posteriorne verjetnosti
indeks homoplastičnosti	nasprotno od indeksa konsistentnosti; $HI=1-CI$
indeks konsistentnosti	(CI, consistency index) mera prileganja znaka določenemu filogenetskemu drevesu. Izračunamo ga po: $c_i = m_i/s_i$, kjer je m_i najmanjše možno število substitucij in s_i najmanjše možno število substitucij na tem drevesu
klad	skupina, ki združuje najbližjega skupnega prednika in vse njegove potomce
metoda največje varčnosti	(Maximum Parsimony) metoda, ki na podlagi vzorcev stanj znakov izbere tisto hipotezo sorodstvenih odnosov, ki predpostavlja najmanj sprememb stanj znakov
metoda največjega verjetja	(Maximum Likelihood) metoda, ki izbere hipotezo sorodstvenih odnosov, ki ima največje verjetje v okviru izbranega evolucijskega modela
monofiletska skupina	glej klad
monofilum	glej klad
takson	(mn.: taksoni) taksonomska skupina kateregakoli ranga; pogosto uporablja v smislu terminalnega taksona (to je najmanjša enota v analizi sorodstvenih odnosov)
zgodovinski signal	vzorec stanj znaka, ki je posledica zgodovinskih procesov ne naključja

1 UVOD

Filogenetika je znanstvena disciplina, ki preučuje evlucijske sorodstvene odnose med taksoni. Cilj je poiskati sorodstvene vezi med vsemi bitji, t.i. drevo življenja. Hipoteze sorodstvenih odnosov prikažemo, v skladu z evlucijsko teorijo (cf. Brower 2000), z drevesu podobnim diagramom, kladogramom (prikaže le sorodstvene odnose) ali filogramom (prikaže tudi mero evlucijske razdalje med taksoni, slika 1).



Slika 1: Filogram. A, B, C, D, E, F so (terminalni) taksoni. Dolžine vej predstavljajo število substitucij (sprememb stanj znakov). Merilo kaže dolžino veje, ki ustreza eni spremembi. 1., 2., 3. so različne hierarhične ravni filograma. Prva hierarhična raven je višja od druge.

Sorodstvene odnose izvajamo iz vzorcev stanj znakov pri različnih osebkih. Znaki so lahko morfološki (npr. krilo pri žuželkah) ali molekularni (zaporedja aminokislin v proteinih, zaporedja DNA in RNA, vrstni red genov, SINE in LINE) ter drugi (fiziološki, vedenjski...). Stanja znaka pa so njegove možne oblike, npr. barva krila. Posamezen znak v zaporedju DNA ima lahko štiri stanja (nukleotidi A, C, G, T). Stanja znakov se spreminjajo zaradi mutacij v procesu evolucije. Na podlagi vzorcev stanj znakov, ki nosijo filogenetsko informacijo, lahko rekonstruiramo sorodstvene odnose.

Najpogosteje uporabljena metoda za ugotavljanje sorodnosti je kladistična analiza. Ta razvršča organizme v skupine na podlagi skupnih izpeljanih stanj znakov (sinapomorfij). To so stanja znakov, ki jih imajo vsi pripadniki neke skupine in jih ni pri najbližjih skupnih

sorodnikih. Pristop, ki se ukvarja le z monofiletskimi skupinami, imenujemo taksistični pristop (Paterson, 1982: 34, op cit Carine in Scotland, 1999). Večina metod v okviru kladistike eksplicitno predpostavlja tudi spreminjanje znakov, t.i. transformacijski pristop. Ta se ukvarja s spremembami, iz česar pa ne »sledi nujno tudi razvrstitev v skupine« (Paterson, 1982: 34, op cit Carine in Scotland, 1999). Torej, za štiri taksone (ABCD), kjer imata A in B enako stanje znaka X_1 in taksona C in D stanje znaka X_2 , taksistični pristop predvideva dve skupini, (AB) in (CD). Iz transformacijske perspektive pa »ni hipotez skupin«, pač pa iz podatkov sledi sprememba $X_1 \leftrightarrow X_2$ (Carine in Scotland, 1999).

Splošno uporabljana transformacijska metoda je metoda največje varčnosti (maximum parsimony), ki deluje po principu Occamove britve (e.g. Hagstrom et al., 2004). Izbere hipotezo sorodstvenih odnosov, ki predpostavlja najmanj sprememb stanj znakov in ima najmanj ad-hoc hipotez. Metoda največje varčnosti za svojo utemeljitev ne potrebuje evolucijske teorije (Brower, 2000) oz. potrebuje le »filogenetsko-nevtralne evolucijske principe« (Lee, 1999; vendar pa Lee, 2002), zato naj bi bila boljša od drugih metod.

Enako stanje nekega znaka pri dveh osebkih je lahko dediščina skupnega prednika, lahko pa je tudi posledica konvergentne evolucije, še posebej pri znakih DNA, ki imajo zelo omejen nabor stanj (e.g. Philippe et al., 1996; Hassanin et al., 1998). Tak znak je homoplastičen in ni posledica evolucijske sorodnosti (e.g. Moreira in Philippe, 2000), zato lahko v kladistični analizi izberemo napačno hipotezo sorodstvenih odnosov.

Homoplastični znaki in znaki, ki nosijo filogenetsko informacijo, podpirajo nasprotujoče sorodstvene odnose v kladogramu. Govorimo o konfliktu znakov (e.g. Lieberman, 1999). Ker metode zgolj izberejo tisto hipotezo, ki najbolje pojasni razporeditev vzorcev stanj znakov, obstaja verjetnost, da prevlada signal homoplazij. V tem primeru dobimo kladogram, ki ne kaže dejanskih sorodstvenih odnosov. Tega neposredno ne moremo vedeti, lahko pa ocenimo podporo hipotetičnim odnosom na podlagi skladnosti znakov. Kajti, podpora neki skupini bo majhna, kadar je veliko znakov v konfliktu, če pa večina znakov podpira enake sorodstvene odnose, bo velika. Ker je verjetnost, da bo veliko naključno spremenjenih znakov nosilo enak vzorec stanj znakov majhna, lahko sklepamo,

da je visoka podpora posledica močnega filogenetskega signala v analiziranih podatkih (e.g. Barome et al., 1998).

V diplomskem delu sem za oceno podpore skupinam uporabil dva indeksa, bootstrap podporo (bootstrap support, BS, Felsenstein, 1985) in Bremerjevo podporo (decay index, Bremer, 1994). Bootstrap (tudi test samovzorčenja) je statistična metoda, ki iz podatkov (vzorca) naredi nove, enako velike vzorce, ki so naključne permutacije izhodiščnega. Pri tem nekaterih znakov ne vključi, nekatere pa v istem vzorcu uporabi večkrat. Za vsak vzorec naredi filogenetsko analizo, podporo skupini v kladogramu pa poda kot delež vzorcev, ki to skupino podpirajo. Bremerjev indeks kot podporo neki skupini oz. veji ki vodi do nje, kot vrednost poda število dodatnih korakov, da ta veja ni več prisotna v konsenzu kladogramov. To pomeni, da ima veja, ki je na najbolj varčnem kladogramu, se pa ne nahaja na konsenzu kladogramov, ki so en korak daljši od tega, Bremerjevo podporo enako 1.

Čeprav konflikt znakov lahko ocenimo, pa lahko le predpostavimo kateri znaki so homoplastični. To so znaki, ki nasprotujejo skupinam v najbolj varčnem kladogramu. Nasprotujejo torej konsenzu vseh znakov. Zato lahko podamo le stopnjo zaupanja v dobljene filogenetske odnose.

1.1 OPREDELITEV PROBLEMA

Znaki se spreminjajo različno hitro. Hitreje spreminjajoči se znaki nosijo informacijo o sorodstvenih odnosih na nižjih, počasneje spreminjajoči se pa na višjih hierarhičnih ravneh kladograma (Moreira in Philippe, 2000). Stanja počasno spreminjajočih se znakov niso informativna na nižjih ravneh, ker se med sabo ne razlikujejo. Uporabnost filogenetskega signala hitro spreminjajočih se znakov na višjih ravneh je zaradi nasičenosti s spremembami vprašljiva, lahko razkrivajo sorodstvene odnose (Björklund, 1999; Källersjö et al., 1999) ali pa so zavajajoči (e.g. Miyamoto et al., 1994).

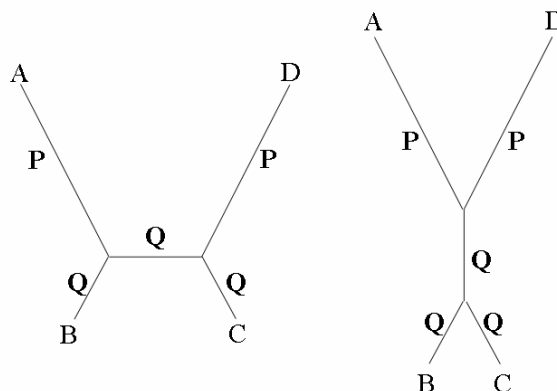
Problem hitro spreminjajočih se znakov predstavljajo večkratne substitucije, saj jih v okviru metode največje varčnosti ne moremo predvideti. Na vsaki veji, ki povezuje dve razvejišči, lahko namreč v okviru te metode predpostavimo le eno ali nič sprememb stanj znaka. Metoda največjega verjetja izbere kladogram, ki najbolje ustreza vzorcu stanj znakov med taksoni in pri tem večkratne substitucije določi ad-hoc, na podlagi konflikta z drugimi znaki. Če je veja kratka, potem bo takih znakov malo ali nič, in signal počasneje spreminjajočih se znakov bo prevladal. Daljša kot je veja, večje bo število znakov z večkratnimi spremembami. Takšni znaki ne nosijo filogenetske informacije o sorodstvenih odnosih na ravneh nad to vejo (Moreira in Philippe, 2000). Pri tem lahko zaradi naključja nastanejo vzorci stanj znakov, ki prekrijejo filogenetski signal v drugih znakih (Wenzel in Siddal, 1999).

Praktični primeri kažejo, da je metoda največje varčnosti precej odporna na šum, ko so homoplazije po vejah enakomerno razporejene (Wenzel in Siddall, 1999). Večji problem predstavljajo nesorazmerja v dolžinah vej, saj lahko pride do pojava privlaka dolgih vej (Felsenstein, 1978; Hendy in Penny, 1989; razširitev problema Kim, 1996). Pri tem je ključno razmerje dolžin vej v filogenetskem drevesu, kjer sta dve dolgi veji ločeni z bistveno krajšo, t.i. Felsensteinova cona. V takih primerih metoda največjega verjetja taksone z dolgimi vejami postavi skupaj, kot sestrške (slika 2). Poleg tega je metoda največjega verjetja v tem območju statistično nekonsistentna¹, saj z večanjem števila znakov konvergira k napačnemu rezultatu (Felsenstein, 1978).

Poleg razlike v hitrosti substitucij v različnih linijah je vezanost velikega števila homoplazij (konflikt) le na določene veje lahko tudi posledica bazne sestave (Collins et al., 1994; Lockhart et al., 1994) in razlik v rabi kodonov (Eyre-Walker, 1991). Tu ne gre za naključne spremembe temveč za zgodovinski signal².

¹ Nekonsistentnost je teoretično izpeljana in jo je v dejanskih podatkih težko predpostaviti (Sanderson et al., 2000), zato uporabljam izraz privlak dolgih vej za stanje, kjer se pojavi pristranskost pri ugotavljanju sorodstvenih odnosov v končnem številu podatkov ali statistična nekonsistentnost zaradi kombinacije dolgih in kratkih vej.

² To je vzorec stanj znaka katerega struktura je posledica zgodovinskega procesa



Slika 2: Primer privlaka dolgih vej na nekoreninjenem drevesu s štirimi taksoni (kvartet). Levo drevo predstavlja dejansko topologijo, desno pa topologijo, ki jo izbere metoda največjega verjetja. A, B, C, D predstavljajo taksoni, P in Q dolžine vej. Metoda je nekonsistentna, če je $P^2 \leq Q(1-Q)$.

1.2 OBTEŽEVANJE ZNAKOV

Dolge veje so pogosto posledica neustreznega nabora taksonov v analizi. Temu se delno lahko izognemo z naborom taksonov, ki skrajšajo problematične veje (Graybeal, 1998; Hillis et al., 2003; Bergsten, 2005; vendar glej Kim, 1996). To ni vedno možno, saj taksoni lahko nimajo bližnjih sorodnih skupin (glej primer hoacina) oz. so dolge veje pač posledica velike hitrosti evolucije. V takih primerih preostaneta le prevrednotenje podatkov ali sprememba metode (primeri občutljivosti drugih metod so povzeti v Bergsten, 2005). V prvem primeru lahko poskusimo identificirati problematične znake (ali spremembe stanj znakov) in jim v filogenetski analizi dati manjšo težo oz. jih izločiti iz analize.

Mnogi raziskovalci so poskušali zmanjšati konflikt znakov z dajanjem večje teže počasneje spreminjajočim se znakom oz. redkejšim substitucijam (Knight in Mindel, 1993). Pogosto uporabljane metode so: obteževanje tretjega mesta v kodonu (e.g. Utevsky in Trontelj, 2004), ki daje manjšo težo spremembam najbolj redundantne pozicije kodona; podobno delujoča metoda, ki DNA zaporedja prevede v aminokislinska (e.g. Barker in Lanyon, 2000); dajanje manjše teže tranzicijam v primerjavi s transverzijami (e.g. Patton et al., 1996); obteževanje na podlagi omejitev substitucij zaradi sekundarne strukture rRNA

(Dixon in Hillis, 1993); zaporedna aproksimacija (successive approximation; Farris, 1969), ki na podlagi indeksa konsistentnosti dobljenega kladograma poda uteži znakom in izračuna nov kladogram. Postopek ponavlja dokler obteženi znaki ne dajo iste topologije; pristop kombinatornih uteži (combinatorial weight approach; Wheeler, 1990) in šest parametrična metoda največje varčnosti (six-parametric parsimony; Williams in Fitch, 1990), ki različno obtežita substitucije glede na ocenjeno pogostost v zaporedjih.

Vendar se te metode osredotočajo na problem hierarhične informativnosti znakov in ne izboljšajo resolucije, če je večina homoplazij vezanih le na določene taksone (e.g. Philippe et al., 1996; Wenzel in Siddall, 1999).

Informativnost (in problematičnost) znakov je torej odvisna od specifičnih taksonov in hierarhične ravni. Zato je ena od možnosti osredotočiti se le na eno problematično vejo, na položaj ene skupine, ter poskušati določiti znake, ki so v konfliktu s filogenetsko informacijo, navzočo v podatkih na tem mestu v drevesu. Te znake lahko določimo le na podlagi hipoteze o sorodstvenih odnosih. Pri tem ne moremo uporabiti hipotez o položaju te skupine, saj to vodi v krožno utemeljevanje.

Lahko pa poznamo filogenetske položaje drugih skupin, katerih monofilija ni sporna, in ki so na hierarhičnih ravneh pod problematičnim razvejiščem³. V tem primeru lahko znake analiziramo na podlagi njihovih topologij. Kajti znaki, ki nasprotujejo tem odnosom, na višjih hierarhičnih ravneh verjetno ne odražajo sorodstvenih odnosov, lahko pa so zavajajoči. Glede na to, da na problematičnem razvejišču obstaja konflikt znakov, je smiselno izločiti te znake.

³ Besedno zvezo »pod hierarhično ravnijo« uporabljam v hierarhičnem in topološkem smislu.

HIPOTEZA

Na podlagi tega razmišljanja smo oblikovali naslednjo hipotezo. Z uporabo predlagane metode identifikacije in izločanja problematičnih znakov na podlagi apriorne filogenetske informacije lahko zmanjšamo konflikt taksonomskih znakov. To dosežemo z izločanjem znakov, ki ne nosijo filogenetske informacije na ravni problema.

NAMEN RAZISKAVE

Namen raziskave predstavljene v diplomskem delu je bil izdelati postopek in preveriti delovanje nove metode za identifikacijo in izločanje problematičnih znakov. Preverjali smo jo na treh izbranih primerih, in sicer na že znanem položaju glavačev (*Physeteridae*) znotraj skupine kitov (*Cetacea*) ter na dveh še nerazjasnjenih primerih, položaju skakačev (*Collembola*) znotraj členonožcev (*Arthropoda*) in sestrski skupini hoacina (*Opisthocomus hoazin*, *Aves*)

V prvem primeru smo ugotavljali ali je rezultat, ki ga poda predstavljena metoda, v skladu s splošno sprejeto in prepričljivo potrjeno hipotezo. V preostalih dveh primerih nas je bolj kot skladnost z »najverjetnejšimi« hipotezami zanimalo obnašanje metode. In sicer, ali je rezultat dobljen s to metodo dosežen v celotnem postopku izločanja znakov.

1.3 IZBRANI PRIMERI

1.3.1 Položaj glavačev (*Physeteridae*) v skupini kiti (*Cetacea*)

V tem poskusu smo preverjali delovanje metode na znanem položaju kitov glavačev znotraj skupine kitov, pri katerem so bili dobljeni nasprotujoči rezultati v analizi mitohondrijskih zaporedij (e.g. Arnason et al., 2004).

Predznanje (Background knowledge):

Zaradi hitre in skrajne morfološke adaptacije so bili sorodstveni odnosi kitov (Cetacea) z drugimi sesalci dolgo nejasni (Arnason et al., 2004). Po nekaterih morfoloških znakih, kot je paraksonija (os simetrije noge je med tretjim in četrtem prstom) (glej Milinkovitch in Thewissen, 1997) so bili umeščeni kot sestrška skupina sodoprstim kopitarjem (Artiodactyla) (Flower, 1883, op cit Milinkovitch et al., 1994; Van Valen, 1966: 93). Kasnejše molekularne raziskave so pokazale, da so kiti pravzaprav ugnezdeni globoko med sodoprstimi kopitarji, kot sestrška skupina povodnim konjem (Hippopotamidae) (Irwin in Arnason, 1994; Ursing in Arnason, 1998; Nikaido et al., 1999), zato so skupino preimenovali v Cetartiodactyla (Montgelard et al, 1997), kito-kopitarji. Čeprav nekateri morfologi na podlagi fosilov še vedno postavljajo kite kot sestrsko skupino sodoprstim kopitarjem (e.g. Thewissen et al., 2001), mnoge novejšje morfološke študije hipotezi pritrjujejo (Geisler in Uhen, 2003; Boisserie et al., 2005).

Tradicionalno delimo živeče kite na dve morfološko dobro podprti skupini, vosate (Mysticeti) in zobate kite (Odontoceti). Prvi »pripadajo monofiletski skupini, ki je diagnosticirana s tremi nedvoumnimi siapomorfijami« (Geisler in Sanders, 2003). V raziskavi sta Messenger in McGuire (1998) našla 22 morfoloških sinapomorfij za zobate kite, vendar pa so molekularne raziskave monofilijo teh postavile pod vprašaj. Milinkovitch et al. (1993, 1994) so, sprva na podlagi analize mitohondrijskih ribosomskih genov 16 vrst kitov, nato pa še delnih zaporedij treh mitohondrijskih (mtDNA) genov za 12S, 16S ribosomsko RNA (rRNA) in citokrom b, pri 21 vrstah kitov ugotovili, da so glavači (Physeteridae) bolj sorodni vosatim kot ostalim zobatim kitom (vendar glej kritiko uskladitve zaporedij, Cerchio in Tucker, 1998). Podoben rezultat so dobili tudi na večjem naboru taksonov in ob uporabi zaporedja gena za citokrom b (Milinkovitch et al., 1996). Po drugi strani pa sta Arnason in Gullberg (1996), prav tako z uporabo zaporedja mitohondrijskega gena za citokrom b dobila (sicer brez statistične podpore) kladogram z monofiletskimi zobatimi kiti. Tudi vključitev aminokislinskega zaporedja mioglobina (poleg mtDNA) v analizi Hasegawa et al. (1997) je dala podporo Milinkovitchevi hipotezi. Po več študijah predvsem mitohondrijske DNA naj bi torej bili zobati kiti parafiletski, ehlokacija pa pleziomorfija, predniški znak, ki so ga vosati kiti izgubili (prevrednotenje

morfoloških znakov v Milinkovitch, 1995). Tega zaključka niso podprle združene analize morfoloških in mtDNA znakov (Messenger in McGuire, 1998) ter morfoloških, jedrnih in mitohondrijskih znakov (Gatesy et al., 1999). Ta hipoteza je bila ovržena tudi s študijo transpozonskih zaporedij (Nikaido et al., 2001), tako da je monofilija zobatih kitov sedaj splošno sprejeta. Po drugi strani pa celo študija celotne mtDNA (Arnason et al., 2004) ni prepričljivo zavrgla Milinkovitcheve hipoteze. Ta zaporedja smo uporabili v raziskavi.

Monofiliji vosatih in zobatih kitov sta torej splošno sprejeti in dobro podprti, zato smo ju privzeli kot monofiletski skupini. Privzeti klad zobatih kitov pa ni vključeval glavačev, saj smo prav njihov položaj znotraj kitov ugotavljali. Predstavnik glavačev sta bila tudi izključena iz analize znakov.

1.3.2 Položaj skakačev (Collembola, Arthropoda)

V tem primeru smo želeli preveriti, ali je nenavaden sorodstveni položaj skupine skakačev, kot ga nakazuje študije mitohondrijskih zaporedij (e.g. Nardi et al., 2003), posledica velike količine homoplazij, ali dejanskih sorodstvenih odnosov.

Predznanje:

Členonožci (Arthropoda)

V klasičnih morfoloških študijah so členonožce večinoma šteli kot monofiletsko skupino (Lankester, 1904, op cit Giribet in Ribera, 2000; Snodgrass, 1938, op cit Giribet in Ribera, 2000). Nekateri avtorji (Tiegs in Manton, 1958, op cit Giribet in Ribera, 2000; Anderson, 1973, op cit Giribet in Ribera, 2000) so sicer spodbijali njihov enkratni izvor, vendar pa danes splošno veljajo za dobro podprto skupino (e.g. Edgecombe et al., 2000; z izjemo Fryer, 1996; glej še Scholtz in Edgecombe, 2006). Skupino so podprle tudi novejšje molekularne študije, npr. Mallatt in Giribet (2006, študija 18S in 28S rRNA genov), ter obsežna študija na podlagi morfoloških znakov, vrstnega reda genov ter mitohondrijskih in jedrnih (ribosomskih in kodirajočih genov) zaporedij (Giribet et al., 2001).

Pipalkarji (Chelicerata)

Znotraj pipalkarjev veljajo Euchelicerata (pravi pipalkarji) za monofiletsko skupino, ki se deli na skupini praskrluparji (Merostomata) s predstavnikom ostvarjem (*Xiphosura*) in terestične pajkovce (Arachnida) (Weygoldt, 1998). Skupino podpirajo nedavne študije (pregled starih hipotez v Wheeler in Hayashi, 1998), tako analize morfoloških znakov (Edgecombe et al., 2000), kot tudi molekularnih (Giribet in Ribera, 2000 ter Mallatt in Giribet, 2006 - 18S in 28S rRNA; Regier et al., 2005a – kodirajoči jedrni geni) ter združenih analiz znakov (Wheeler in Hayashi, 1998; Edgecombe et al., 2000; Giribet et al., 2001).

V analizi smo imeli zaporedja DNA skupin pršic (*Acarina*, Arachnida) in ostvarjev, oboji so pripadniki skupine pravi pipalkarji, katero smo privzeli kot monofiletsko skupino za analizo znakov.

Stonoge (Myriapoda)

V primeru smo imeli dva razreda (po Giribet in Ribera, 2000) stonog, strige (Chilopoda) in kačice (Diplopoda). Oba razreda veljata za monofiletska (Edgecombe in Giribet, 2002) in ju podpirajo morfološke (Edgecombe et al., 2000; Bitsch C. in Bitsch J., 2004; Giribet et al., 2005) ter molekularne (Regier et al., 2005b; Mallatt in Giribet, 2006) študije. Tako sta v svoji študiji morfoloških in molekularnih znakov Edgecombe in Giribet (2004) komentirala monofilijo skupine strig kot »nekontroverzen rezultat«. Podala sta tudi pregled sinapomorfij strig (Edgecombe in Giribet, 2002) in kačic (Edgecombe in Giribet, 2004).

Monofilija stonog je kontroverzna. Na podlagi morfologije so skupino uvrščali kot sestrsko šesteronožnim členonožcem (Hexapoda) ali kot parafiletsko, kjer je del »stonog« sestrski monofilumu sestavljenem iz preostalih »stonog« in šesteronožnih členonožcev (pregled mnogih hipotez delitev stonog v Bitsch C. in Bitsch J., 2004). Dohle (1997, op cit Edgecombe et al., 2000) trdi, da stonoge nimajo nobene pozitivne sinapomorfije, to je stanja znaka, ki ne pomeni njegove odsotnosti ali izgube. Edgecombe et al. (2000) v obsežni morfološki študiji identificirajo le en znak, enoslojni akrosomski kompleks v spermijih, kot nedvoumno sinapomorfijo pri vseh transformacijskih optimizacijah (vendar glej Edgecombe in Giribet, 2002, za drugačno definicijo tega znaka). Nasprotno pa

nekatero molekularno študijo močno podpirajo stonoge kot monofiletsko skupino (Regier et al., 2005b).

Problematika monofilije stonog se je pojavljala v okviru hipotetične skupine vzdušničarji (Tracheata oz. Atelocerata), ki združuje šesteronožne členonožce in stonoge, vzpostavljene na podlagi morfoloških znakov. Regier et al. (2005b) vprašljivost morfoloških znakov komentirajo: »... trdimo, da morfologija ne podaja niti nedvoumnih sinapomorfij stonog niti prepričljivih alternativ njihovi monofiliji...«. Molekularne študije (Giribet et al., 2001; Negrisoló et al., 2004; Cook et al., 2005; Luan et al., 2005; Regier et al., 2005a) pa postavljajo šesteronožne členonožce kot skupino znotraj rakov, t.i. hipotezo Pancrustacea oz. Tetraconata. To hipotezo je nakazala tudi novejša morfološka študija Giribet et al. (2005; vendar glej Bitsch C. in Bitsch J., 2004). Edgecombe in Giribet (2002) sta v luči pankrustacijske hipoteze kot sinapomorfije stonog navedla več znakov.

Nasprotno pa nekatere molekularne študije nakazujejo novo monofiletsko skupino, Paradoxopoda (Mallatt et al., 2004), ki združuje skupini pipalkarji in stonoge (18S in 28S rRNA - Mallatt et al., 2004; mtDNA - Negrisoló et al., 2004; Hox geni – Cook et al., 2001). Slednji so lahko v okviru te hipoteze celo parafiletska skupina (Cook et al., 2001; Negrisoló et al., 2004).

Ker še ni konsenza o statusu stonog ter zaradi pomanjkanja podrobnejših analiz notranjih struktur pri členonožcih, »z ozirom na notranje strukture... smo upoštevali le makroskopsko anatomijo, zaradi pomanjkanja natančnih informacij v mnogih taksonomskih skupinah« (Bitsch C. in Bitsch J., 2004) in v luči novih hipotez o položaju stonog, smo kot monofiletski predpostavki vzeli skupini strige in kačice.

Raki (Crustacea)

V analizi so bili raki zastopani s tremi taksoni, listonožcem (Branchiopoda), ceponožcem (Copepoda) in deseteronožcem (Decapoda, del višjih rakov – Malacostraca). Mitohondrijski genom ceponožca, *Tigriopus japonicus*, je močno preurejen glede na ostale členonožce, zato smo ga iz analize izključili (Cameron et al., 2004; vendar vključen v

Cook et al., 2005). Preostali dve skupini sta tako morfološko kot molekularno dobro podprti (Giribet et al., 2005; glej še združena analiza Giribet et al., 2001; jedrna DNA v Regier et al., 2005a; jedrna rRNA v Mallatt in Giribet, 2006; mitohondrijski genom v Cook et al., 2005).

Novejša hipoteza, ki združuje rake in šesteronožne členonožce v skupino Pancrustacea, postavlja pod vprašaj monofilijo prvih. Šesteronožni členonožci so lahko rakom sestrski (hipoteza je predvsem posledica pomanjkljivega nabora taksonov znotraj rakov v analizah, e.g. Nardi et al., 2003), ali pa, verjetneje, predstavljajo skupino znotraj njih, kot sestrski listonožcem (Regier et al., 2005a), sestrski ceponožcem (Mallatt in Giribet, 2006), sestrski monofiliji (listonožci+višji raki) (Mallatt et al., 2004 in Cook et al., 2005).

Kot apriorni hipotezi smo torej privzeli monofiliji taksonov listonožci in deseteronožci (višji raki).

Šesteronožni členonožci (Hexapoda; Insecta s.l. – po Luan et al., 2005)

Znotraj te skupine, katere monofilijo smo preverjali, so bili v sorodstveno analizo vključeni predstavniki skakačev (Collembola) in skupine žuželke (Insecta s.s.; Ectognatha sensu Hennig). Drugo skupino sestavljajo ščetinorepke (Archaeognatha), srebrne ribice (Zygentoma) in krilate žuželke (Pterygota). Prepričljivo jo podpirajo recentne študije morfoloških znakov (Wheeler et al., 2001; Bitsch C. in Bitsch J., 2004; Giribet et al., 2004, 2005). Wheeler et al. 2001 so v raziskavi našli 13 sinapomorfij. Od molekularnih analiz, jedrni 18S in 28S rRNA geni (Kjer, 2004; Luan et al., 2005; Mallatt in Giribet, 2006) dajejo zelo prepričljivo podporo monofiliji žuželk. Tudi analize jedrnih genov za EF-1 α , EF-2 in PolII (Regier et al., 2005a) monofilijo dobro podpirajo. Študije mtDNA (e.g. Nardi et al., 2003; Lavrov et al. 2004, Cameron et al., 2004) pa nimajo ustreznega nabora taksonov za reševanje tega problema. Analiza Cook et al. (2005) celotnega mitohondrijskega genoma je sicer podprla monofilijo žuželk, vendar v raziskavo ni bila vključena bazalna skupina ščetinorepke.

Podpora monofiliji skakačev je še bolj prepričljiva. Zelo dobro jo podpirajo vse analize (e.g. D'Haese 2002; Cook et al. 2005; Giribet et al., 2005; Luan et al., 2005; Regier et al.,

2005a). Bitsch C. in Bitsch J. (2004) podata kritičen pregled sinapomorfij, vendar monofilije ne spodbijata.

Kot dobro podprti hipotezi smo privzeli monofiliji žuželk in skakačev. Monofiletska skupina skakači je bila iz analize znakov (kvarteti) izključena, saj smo položaj te skupine preverjali.

1.3.3 Sestrsko skupino hoacina (*Opisthocomus hoazin*, Aves)

Že odkar je bil leta 1776 prvič opisan, je *Opisthocomus hoazin* taksonomska uganka (Hedges et al., 1995). »Prvotno opisan kot *Phasianus hoazin*, torej kot pripadnik kur, je bil nato uvrščen v mnoge skupine, kot so turaki (Musophagidae), kljunorožci (Bucerotidae), med golobe, kukavice in žerjavovce (Gruiformes)« (Hedges et al., 1995). Položaj hoacina je problematičen zaradi bizarnega izgleda (Hughes in Baker, 1999) in očitne filogenetske oddaljenosti od ostalih recentnih vrst ptic, tako da so ga celo postavljali v svoj red Opisthocomiformes (e.g. Hedges et al., 1995). Problematičnost potrjuje tudi njegov nestabilen položaj v kladogramih različnih analiz Sorenson et al. (2003).

Predznanje:

Čas in vzorec evlucijske radiacije modernih ptic (Neornithes) sta še v veliki meri nejasna (Dyke in Van Tuinen, 2004). Odnosi višjih skupin so sicer znani, korenina je med skupinama staročeljustnic (Paleognathae) in novočeljustnic (Neognathae), slednjo delimo na monofiletski skupini Galloanserae (kure in plojkokljuni) in Neoaves (Harrison et al., 2004). Odnosi znotraj monofiletske skupine Neoaves, ki obsega večino živečih vrst ptic, pa so zelo nejasni (Edwards et al., 2005). V okviru nabora taksonov analize Sorenson et al. (2003) lahko le malo skupin označimo kot zanesljive monofilume. Dobro podprti redovi so mišaki (Coliiformes) (Dyke in Van Tuinen, 2004), pevci (Passeriiformes) (Livezey in Zusi, 2001; Harrison et al., 2004) in pobrežniki (Charadriiformes) (Erikson et al., 2003; Paton et al., 2003; Van Tuinen et al., 2004; cf. Fain in Houde, 2004). Več zanesljivih monofilumov

je na ravni družin, e.g. štorke (Ciconiidae), kukavice (Cuculidae) in prave papige (Psittacidae) (Slikas 1997; Hughes in Barker, 1999; Dyke in van Tunien, 2004). Navedene skupine so bile v analizi uporabljene kot apriorne monofilije. Čeprav je skupina Galloanserae sestrsko Neoaves, in tako manj relevantna pri ugotavljanju položaja hoacina, je bila zaradi pomanjkanja ustrežnejših skupin prav tako vključena v analizo znakov.

Od hipotetičnih sestrskih skupin iz Sorenson et al. (2003) smo preverjali le trogona (*Trogon curucui*), vijoličnega turaka (*Musophaga violacea*), beloglavega goloba (*Columba leucocephala*) in anija (*Crotophaga ani*). Čeprav sta se v rezultatih nekaterih sorodstvenih analiz študije Sorenson et al. (2003) tudi sokol selec (*Falco peregrinus*) in rdečerepa kanja (*Buteo jamaicensis*) pojavljala kot sestrski skupini hoacina, je to po mnenju avtorjev posledica privlaka dolgih vej, zato v analizo znakov nista bila vključena.

2 MATERIALI IN METODE

2.1 MATERIALI

2.1.1 Zaporedja DNA

Pri analizi sorodstvenih odnosov smo uporabili zaporedja DNA dobljena od avtorjev, kito-kopitarji - Arnason et al. 2004, ptice - Sorenson et al. 2003, členonožci - Cameron et al. 2004.

2.1.1.1 Glavači

Pri kito-kopitarjih zaporedje obsega celotno mitohondrijsko DNA, v analizi smo uporabili 12 kodirajočih genov, gena za NADH6 (glej Arnason et al., 2004) pa nismo vključili. Poleg kodirajočih genov smo dodali še zaporedji genov za 12S in 16S RNA, ki smo jih pridobili iz GeneBank-a. Usklajena zaporedja so vsebovala 13.468 znakov (vključno z INDEL-i).

2.1.1.2 Skakači

Uporabili smo celotno mitohondrijsko kodirajočo DNA, to je 13 genov, Usklajena zaporedja so bila dolga 11.793 znakov (vključno z INDEL-i).

2.1.1.3 Hoacin

V raziskavi so bili uporabljeni geni (navedeni v skupinah, kot so bili sekvencirani) za 12S rRNA in del tRNA; del 16S rRNA ter celotni podenoti I in II gena za NADH dehidrogenazo; 36 baz gena za COI, 9 genov za tRNA; del gena za NADH5, cel gen za

citokrom b in del gena za tRNA-treonin (vse mitohondrijska DNA); intron 9 gena za PEPCCK (jedrna). Zaporedja so skupaj obsegala 6349 znakov (vključno z INDEL-i).

2.1.2 Uporabljeni taksoni

2.1.2.1 Glavači

Uporabili smo iste taksone kot Arnason (2004), vendar brez pripadnikov lihoprstih kopitarjev, ki jih je avtor uporabil kot zunanjo skupino za postavitev korenine kito-kopitarjem. V izbranem problemu so vlogo zunanje skupine imeli taksoni iz »sodoprstih kopitarjev«. Skupine, ki smo jih uporabili v analizi sorodstvenih odnosov so navedene v tabeli 1.

V kvartetih za analizo deleža homoplastičnosti znakov (nova metoda) smo uporabili vse taksone razen glavačev. Predpostavljene monofiletske skupine so navedene v tabeli 2.

2.1.2.2 Skakači

V poskusu so bili uporabljeni vsi predstavniki členonožcev, razen ceponožca (glej uvod). Iz analize deleža homoplastičnosti znakov smo izključili predstavnike monofilije skakačev, katerega položaj znotraj členonožcev smo želeli ugotoviti. V tabeli 3 so navedeni uporabljeni taksoni kot pripadniki privzetih monofilij.

2.1.2.3 Hoacin

V raziskavi smo uporabili enak nabor taksonov kot Sorenson et al. (2003), vendar brez skupine staročeljustnic (tabela 4). Korenina je v skladu s predhodnimi študijami (e.g. Livezey in Zusi, 2001) postavljena med monofiluma Galloanserae in Neoaves.

V analizi deleža homoplastičnosti znakov smo uporabili predpostavljene monofiletske skupine (glej uvod) in štiri hipotetične sestrskse taksone hoacina.

Tabela 1: Glavači. Taksoni uporabljeni v filogenetski analizi.

SKUPINA	LATINSKO IME	DRUŽINA	SLOVENSKO IME
Vosati kiti	<i>Balaena mysticetus</i>	Balaenidae	grenlandski kit
(Mysticeti)	<i>Caperea marginata</i>	Neobalaenidae	mali gladki kit
	<i>Eschrichtius robustus</i>	Eschrichtiidae	sivi kit
	<i>Balaenoptera acutorostrata</i>	Balaenopteridae	ščukasti ali mali kit
	<i>Balaenoptera physalus</i>	Balaenopteridae	brazdasti (hrbtoplavuti) kit
	<i>Balaenoptera musculus</i>	Balaenopteridae	sinji kit
Zobati kiti	<i>Lagenorhynchus albirostris</i>	Delphinidae	belokljuni pisani delfin
(Odontoceti)	<i>Monodon monoceros</i>	Monodontidae	samorog
	<i>Phocoena phocoena</i>	Phocoenidae	pristaniška rjava pliskavka
	<i>Inia geoffrensis</i>	Iniidae	orinoška pliskavka
	<i>Pontoporia blainvillei</i>	Pontoporiidae	laplatska pliskavka
	<i>Hyperoodon ampullatus</i>	Ziphiidae	severni račji kit
	<i>Berardius bairdii</i>	Ziphiidae	črni kit
	<i>Platanista minor</i>	Platanistidae	induška pliskavka
	<i>Kogia breviceps</i>	Kogiidae	mali glavač
	<i>Physeter macrocephalus</i>	Physeteridae	kit glavač
Sodoprsti kopitarji	<i>Sus scrofa</i>	Suidae	evrazijska divja svinja
(»Artiodactyla«)	<i>Bos taurus</i>	Bovidae	domače govedo
	<i>Ovis aries</i>	Bovidae	ovca
	<i>Muntiacus muntjac</i>	Bovidae	indijski muntjak
	<i>Hippopotamus amphibius</i>	Hippopotamidae	veliki povodni konj
	<i>Lama pacos</i>	Camelidae	alpaka

Tabela 2: Glavači. Monofiletske skupine uporabljene v analizi znakov.

»Odontoceti«:	<i>Lagenorhynchus albirostris</i> , <i>Monodon monoceros</i> , <i>Phocoena phocoena</i> , <i>Inia geoffrensis</i> , <i>Pontoporia blainvillei</i> , <i>Hyperoodon ampullatus</i> , <i>Berardius bairdii</i> , <i>Platanista minor</i>
Mysticeti:	<i>Balaena mysticetus</i> , <i>Caperea marginata</i> , <i>Eschrichtius robustus</i> , <i>Balaenoptera acutorostrata</i> , <i>Balaenoptera physalus</i> , <i>Balaenoptera musculus</i>
sus:	<i>Sus scrofa</i>
ovis:	<i>Ovis aries</i>
hippo:	<i>Hippopotamus amphibius</i>
bos:	<i>Bos taurus</i>
alpaca:	<i>Lama pacos</i>
muntjac:	<i>Muntiacus muntjac</i>

Tabela 3: Skakači. Taksoni uporabljeni v analizi znakov (v privzetih monofiletskih skupinah) in filogenetski analizi. V analizo znakov pripadnika skupine Collembola nista bila vključena.

MONOFILIJA	VRSTA	KRATICA
Collembola	<i>Gomphiocephalus hodgsoni</i>	Gomp
	<i>Tetrodontophora bielanensis</i>	Tbie
Insecta	<i>Nesomachilis australica</i>	Arch
	<i>Tricholepidion gertschi</i>	Tric
	<i>Locusta migratoria</i>	Lmig
	<i>Triatoma dimidiata</i>	Tdim
	<i>Tribolium castaneum</i>	Tcas
	<i>Crioceris duodecimpunctata</i>	Cduo
	<i>Pyrocoelia rufa</i>	Pruf
	<i>Antheraea pernyi</i>	Aper
	<i>Bombyx mori</i>	Bmor
	<i>Bombyx mandarina</i>	Bman
	<i>Ostrinia furnacalis</i>	Ofur
	<i>Ostrinia nubilalis</i>	Onub
	<i>Anopheles quadrimaculatus</i>	Aqua
	<i>Anopheles gambiae</i>	Agam
	<i>Drosophila yakuba</i>	Dyak
	<i>Drosophila melanogaster</i>	Dmel
	<i>Crysomya putoria</i>	Cput
<i>Cochliomyia hominivorax</i>	Chom	
<i>Ceratitis capitata</i>	Ccap	
Branchiopoda	<i>Daphnia pulex</i>	Dpul
	<i>Artemia franciscana</i>	Afra
	<i>Triops cancriformis</i>	Tcan
Decapoda	<i>Panulirus japonicus</i>	Pjap
	<i>Portunus trituberculatus</i>	Ptri
	<i>Pagurus longicarpus</i>	Plon
	<i>Penaeus monodon</i>	Pmon
Diplopoda	<i>Thyropygus sp.</i>	Thyr
	<i>Narceus annularis</i>	Nann
Chilopoda	<i>Lithobius forficatus</i>	Lfor
Chelicerata	<i>Varroa destructor</i>	Vdes
	<i>Limulus polyphemus</i>	Lpol
	<i>Ornithodoros moubata</i>	Omou

se nadaljuje

nadaljevanje tabele

<i>Ixodes hexagonus</i>	Ihex
<i>Ixodes persulcatus</i>	Iper
<i>Rhipicephalus sanguineus</i>	Rsan

Tabela 4: Uporabljeni taksoni v primeru hoacin. Taksonomska razvrstitev po Sorenson et al. (2003).

RED	DRUŽINA	VRSTA
Galliformes	Cracidae	<i>Crax rubra</i>
	Megapodiidae	<i>Alectura lathamii</i>
	Megapodiidae	<i>Megapodius eremita</i>
	Phasianidae	<i>Gallus gallus</i>
	Numididae	<i>Acryllium vulturinum</i>
Anseriformes	Anhimidae	<i>Chauna torquata</i>
	Anseranatidae	<i>Anseranas semipalmata</i>
	Dendrocygnidae	<i>Dendrocygna arcuata</i>
	Anatidae	<i>Aythya americana</i>
Bucerotiformes	Bucerotidae	<i>Tockus erythrorhynchus</i>
Trogoniformes	Trogonidae	<i>Trogon curucui</i>
Coraciiformes	Coraciidae	<i>Coracias spatulata</i>
Coliiformes	Coliidae	<i>Colius striatus</i>
	Coliidae	<i>Urocolius macrourus</i>
Cuculiformes	Cuculidae	<i>Cuculus canorus</i>
	Centropodidae	<i>Centropus cupreicaudus</i>
	Coccyzidae	<i>Coccyzus erythrophthalmus</i>
	Opisthocomidae	<i>Opisthocomus hoazin</i>
	Crotophagidae	<i>Crotophaga ani</i>
Neomorphidae	Neomorphidae	<i>Neomorphus geoffroyi</i>
	Psittacidae	<i>Neophema elegans</i>
Psittaciformes	Psittacidae	<i>Nandayus nenday</i>
	Musophagidae	<i>Musophaga violacea</i>
Musophagiformes	Musophagidae	<i>Crinifer piscator</i>
	Strigidae	<i>Asio otus</i>
Strigiformes	Columbidae	<i>Columba leucocephala</i>
	Columbidae	<i>Treron sieboldii</i>
Columbiformes	Scolopacidae	<i>Scolopax minor</i>
	Burhinidae	<i>Burhinus senegalensis</i>

se nadaljuje

		nadaljevanje tabele
	Accipitridae	<i>Buteo jamaicensis</i>
	Falconidae	<i>Falco peregrinus</i>
	Phoenicopteridae	<i>Phoenicopus ruber</i>
	Ciconiidae	<i>Mycteria americana</i>
	Ciconiidae	<i>Ciconia ciconia</i>
Passeriformes	Eurylaimidae	<i>Smithornis sharpei</i>
	Tyrannidae	<i>Sayornis phoebe</i>
	Corvidae	<i>Corvus frugilensis</i>
	Passeridae	<i>Vidua chalybeata</i>

2.2 METODE

2.2.1 Uskladitev zaporedij

2.2.1.1 Glavači

Zaporedja kodirajočih genov smo dobili že usklajena od avtorjev (Arnason et al., 2004). Dodani zaporedji 12S in 16S RNA sta bili usklajeni s pomočjo programa ClustalX (Thompson et al., 1997) pod pogoji: »gap opening penalty 15, gap extension penalty 6,66, DNA transition weight 1,0«. Nadaljnje analize vprašljivo usklajenih znakov nismo naredili, saj smo predvidevali, da jih bo identificirala analiza homoplastičnosti znakov.

2.2.1.2 Skakači

Že usklajena DNA zaporedja smo dobili od avtorjev, Cameron et al. (2004). Ti so v svoji študiji naredili več različnih uskladitev zaporedij, ki so se razlikovale glede na vključene taksone zunanjih skupin. Cameron et al. (2004) so v analizi kot zunanje skupine uporabili gliste (Nematoda) ali kolobarnike (Annelida) in mehkužce (Mollusca), glede na dve hipotezi o položaju členonožcev. Drugi izbor je v skladu s tradicionalno hipotezo o položaju členonožcev, uporaba glist pa odseva novejšo hipotezo o členonožcih kot

ekdisozojih (e.g. Mallat in Giribet, 2006). V obeh primerih ne gre za najbližje sorodnike in boljša izbira bi bili taksoni iz monofiletske skupine Pararthropoda. Uporaba bolj oddaljenih skupin vodi v slabšo usklajenost zaporedij, to je v pretirano število INDEL-ov (Cameron et al., 2004), ta mesta pa so iz analize po kriteriju največje varčnosti izključena. Posledica je manjši nabor potencialno informativnih znakov. Zunanja skupina lahko povzroči tudi privlak dolgih vej (e.g. Bergsten, 2005). Zato smo v končni študiji uporabili uskladitev zaporedij brez zunanjih skupin. V preliminarni analizi smo metodo uporabili tudi na usklajenih zaporedjih členonožcev in glist, da smo določili korenino.

2.2.1.3 Hoacin

Uporabili smo usklajena zaporedja pridobljena od avtorjev Sorenson et al. (2003). Izločili smo le [PEPCK-gap characters] matriko, v kateri so bili kodirani INDEL-i introna 9 gena za PEPCK.

2.2.2 Analiza znakov

Kvarteti

Kvartete uporabljene za ugotavljanje konflikta znakov smo sestavili s pomočjo računalniškega programa (avtor Tori Andraž) (priloga A - vhodna datoteka). Program vrne vse možne kombinacije štirih taksonov (zapisane v tekstovnem formatu; priloga B), pri čemer sta dva pripadnika iste monofiletske skupine, druga dva pa ne.

Analiza znakov

Analiza znakov je bila opravljena v programu PAUP* 4.0b10 (Swofford, 2002). Na podlagi vnaprej znanih topologij kvartetov smo identificirali homoplastične znake, to je znake, ki so topologiji nasprotovali. Zaradi obsežnosti smo uporabili računalniški program (avtor Tori Andraž), ki izpiše znake, katerih indeks homoplastičnosti (HI) ni enak 0 (glej rezultate in diskusijo). Kot rezultat vrne za vsak znak delež kvartetov v katerih ta znak ni podpiral apriornih hipotez, t.i. delež homoplastičnosti.

Izločanje in sorodstvena analiza

Znake smo razvrstili po deležu homoplastičnosti in jih nato v korakih zaporedno izločali, od najbolj do najmanj homoplastičnih (priloga C). Po vsakem koraku smo naredili filogenetsko analizo, podporo vsaki sorodstveni hipotezi pa smo ovrednotili na podlagi bootstrap podpore (bootstrap support) in umerjene Bremerjeve podpore (rescaled Bremer support).

Ker se zaradi izločanja znakov dolžine dreves krajšajo, Bremerjev indeks ne podaja primerljivih vrednosti. Zato smo uporabili umerjeno Bremerjevo podporo, ki je enaka Bremerjevi podpori deljeni z deležem vključenih informativnih znakov (enačba 1):

$$Ubp_t = Bp_t * (N_o / N_t) \quad \dots(1)$$

kjer je Ubp_t umerjena Bremerjeva podpora, Bp_t Bremerjeva podpora razvejišču v koraku t , N_o začetno število informativnih znakov, N_t število informativnih znakov v koraku t . Ker se zaradi izločanja najbolj homoplastičnih znakov dolžina drevesa zmanjšuje hitreje kot upada število informativnih znakov, je Ubp konzervativna ocena spreminjanja podpore. Za primerjavo, po izločitvi 133 (2,5% vseh informativnih) znakov z največjim deležom homoplastičnosti v zaporedjih kito-kopitarjev se dolžina drevesa zmanjša za 979 (3,6%) korakov.

2.2.2.1 Glavači

Sestavili smo 3.213 kvartetov. Kot homoplastične v vsaj enem kvartetu smo določili 4.018 znakov, najvišji delež homoplastičnosti je bil 0,32. Izločali smo po 100-150 znakov na korak, v 24 korakih je bilo skupaj izločenih 3.048 (delež homoplastičnosti večji od 0,044).

2.2.2.2 Skakači

V okviru analize homoplastičnosti znakov smo iz predpostavljenih monofilij naredili 31.403 kvartetov. 7.825 znakov je bilo homoplastičnih v vsaj enem kvartetu, najbolj homoplastični so nasprotovali apriornim hipotezam v 0,28 vseh kvartetov. V 17 zaporednih korakih smo izločali po približno 350 znakov, vse skupaj 6.232.

2.2.2.3 Hoacin

Kvartete smo tvorili tako, da smo v vsakem uporabili dva predstavnika ene predpostavljene monofilije in dva hipotetična sestrskata taksona hoacina. Za vsako dvojico izbrane monofilije smo naredili vse možne kombinacije kvartetov za štiri hipotetične sestrskate taksona. Tako smo v primeru apriornega monofiluma mišaki z dvema predstavnikoma tvorili kvartete: (mišak1, mišak2, golob, turako), (mišak1, mišak2, golob, ani), (mišak1, mišak2, golob, trogon), (mišak1, mišak2, turako, ani), (mišak1, mišak2, turako, trogon), (mišak1, mišak2, ani, trogon).

2.2.2.4 Analiza izločenih znakov po indeksu homoplastičnosti (HI)

V primeru glavačev, katerih sestrski odnosi so že bili razjasnjeni, smo za vse izločene znake ugotovili tudi njihov HI na podlagi prvotne topologije. Razporedili smo jih v razrede od R_1 (od vključno 0 do 0,1) do R_{10} (od vključno 0,9 do 1). Zanimalo nas je katere znake po HI identificira analiza znakov kot problematične in kateri razred najbolj prispeva k spremembi podpore topologiji.

2.2.3 Filogenetska analiza

V vseh filogenetskih analizah smo uporabili program PAUP 4.0b10 (Swofford, 2002), z iskalnimi nastavitvami: »Hsearch nreps=100 addseq=random swap=tbr;« BS podpora: »Bootstrap nreps=1000/nreps=10 addseq=random swap=nni;«.

Bremerjevo podporo smo izračunali na podlagi razlike v dolžinah optimalnih dreves z oz. brez preučevane skupine. To smo naredili z ukazom »constrain« in uporabo istih iskalnih nastavitev kot v filogenetski analizi.

V okviru analize znakov smo prav tako izvajali filogenetsko analizo, vendar z vnaprej določeno topologijo kvarteta. S tem smo dobili HI vrednosti znakov. Uporabili smo algoritem »branch and bound« (glej prilogo 2).

3 REZULTATI

3.1 METODA

Za razliko od metod, s katerimi skušamo izboljšati resolucijo celotnega kladograma, se predlagana metoda osredotoča na določeno problematično skupino in njen položaj v filogenetskem drevesu. Gre za višje taksonomske skupine, katerih filogenetski položaj ni jasen zaradi mnogih homoplastičnih znakov.

Metoda zahteva poznavanje položaja korenine v kladogramu. To nam omogoča, da opredelimo hierarhično raven problema in skupine na nižjih taksonomskih ravneh. Na podlagi tega lahko poiščemo za problem relevantne skupine, katerih monofilijam zaupamo.

Temeljni principi metode:

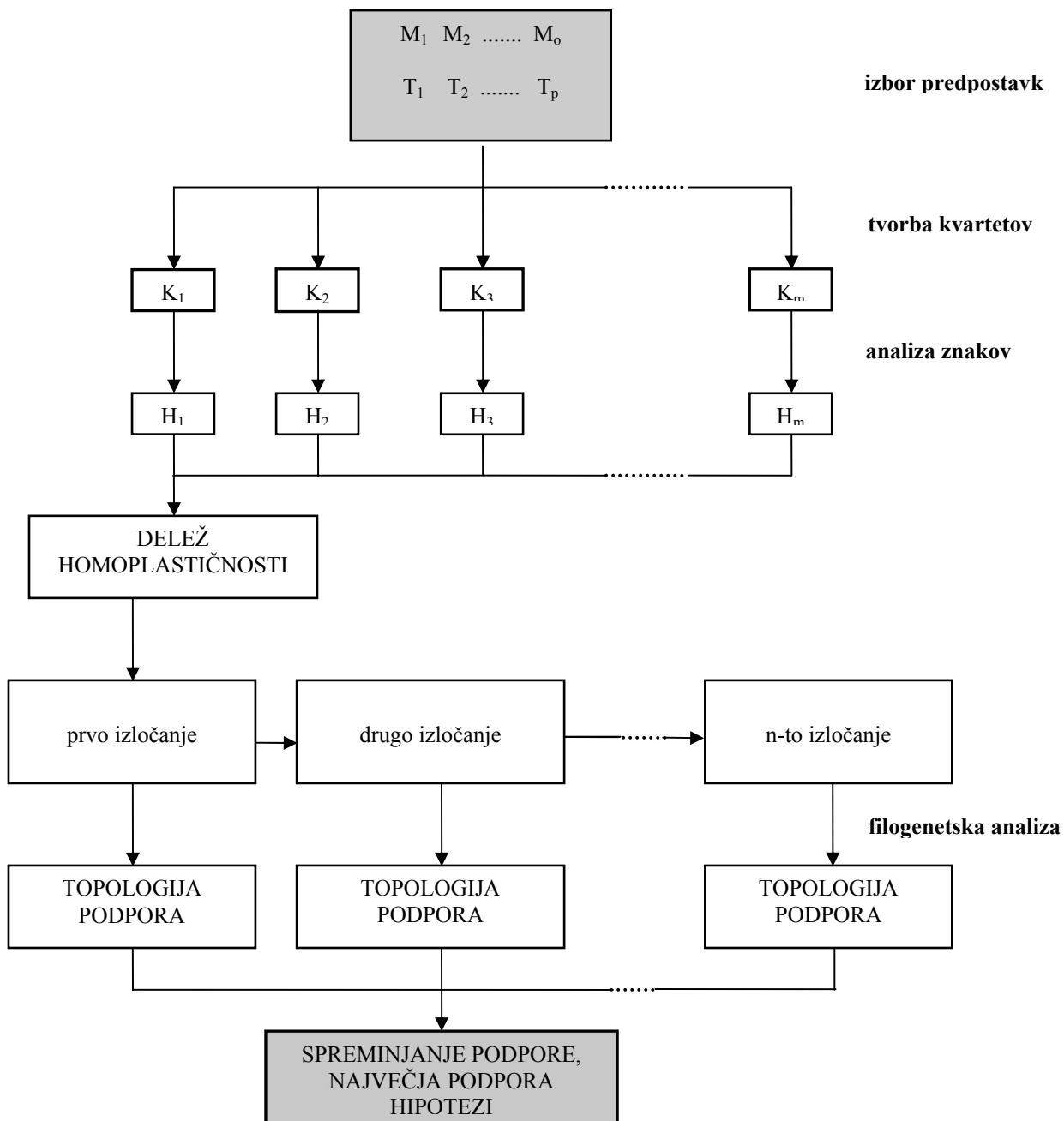
- predznanje; apriorno, od v raziskavi uporabljenih podatkov (znakov) neodvisno znanje;
- znake, ki so v konfliktu z apriornim znanjem izločimo.

Postopek razdelimo v štiri korake (slika 3). Predhodno moramo opredeliti problem, to je skupino, katere položaj v kladogramu ugotavljamo, in hipoteze o njenih sestrskih odnosih.

Postopek:

1. izbor relevantnega apriornega znanja (monofiletskih skupin),

2. tvorba kvartetov,
3. analiza znakov – ugotavljanje deleža homoplastičnosti,
4. kumulativno izločanje znakov na podlagi deleža homoplastičnosti in filogenetska analiza po vsakem izločanju.



Slika 3: Shema metode. M, T: izbrane monofiletske skupine in taksoni (glej predpostavke); K_i : i-ti kvartet; H_i : konfliktni znaki v i-tem kvartetu; delež homoplastičnosti: nasprotovanje znakov privzetim monofilijam.

Predpostavke

V prvem koraku opredelimo predpostavke, monofiletske skupine, ki jih raziskave opravljene na drugih podatkih (znakih) dobro podpirajo, ter ostale taksone, ki jih želimo uporabiti v analizi sorodstvenih odnosov. Idealno je če so privzete monofilije neposredno pod problematičnim razvejiščem, saj tako metoda identificira največji delež problematičnih znakov.

Kvarteti

Iz izbranih taksonov tvorimo kvartete, to je skupine štirih taksonov, pri čemer dva pripadata isti monofiletski skupini, druga dva pa ne. Topologija takega, nekoreninjenega, filogenetskega drevesa je vnaprej znana (po apriornih hipotezah) kar nam omogoča vrednotenje znakov. Kvartet nam omogoča, da preverjamo le, ali znak podpira našo hipotetično skupino, in ne kako je razporejen na celem drevesu. Poleg tega je taka analiza zelo rigorozna, saj morata biti v enem kvartetu dve homoplaziji, da znak nasprotuje predpostavljeni monofiliji.

Analiza znakov

Za vsak v kvartetih informativen znak na podlagi topologije predpostavljamo eno ali dve spremembi stanja znaka. Znaki z dvema spremembama stanja topologiji nasprotujejo, so homoplastični, in zato tudi na ravni problema lahko zavajajoči. V kvartetu ima tak znak HI enak 0,5. Ker je rezultat analize nekega znaka odvisen od konkretne kombinacije štirih taksonov in zato lahko posledica naključja, naredimo analizo znakov na vseh možnih kvartetih. Homoplastične znake razvrstimo po deležu kvartetov v katerih so nasprotovali predpostavljenim monofilijam, po »deležu homoplastičnosti« (enačba 2):

$$DH_j = KV_j / KV_{tot} \quad \dots(2)$$

kjer je DH_j delež homoplastičnosti j-tega znaka, KV_j število kvartetov v katerih je j-ti znak nasprotoval privzeti monofiletski skupini in KV_{tot} število vseh kvartetov. Torej ima znak, ki je nasprotoval predpostavljeni topologiji v dveh od osmih kvartetov delež homoplastičnosti 0,25.

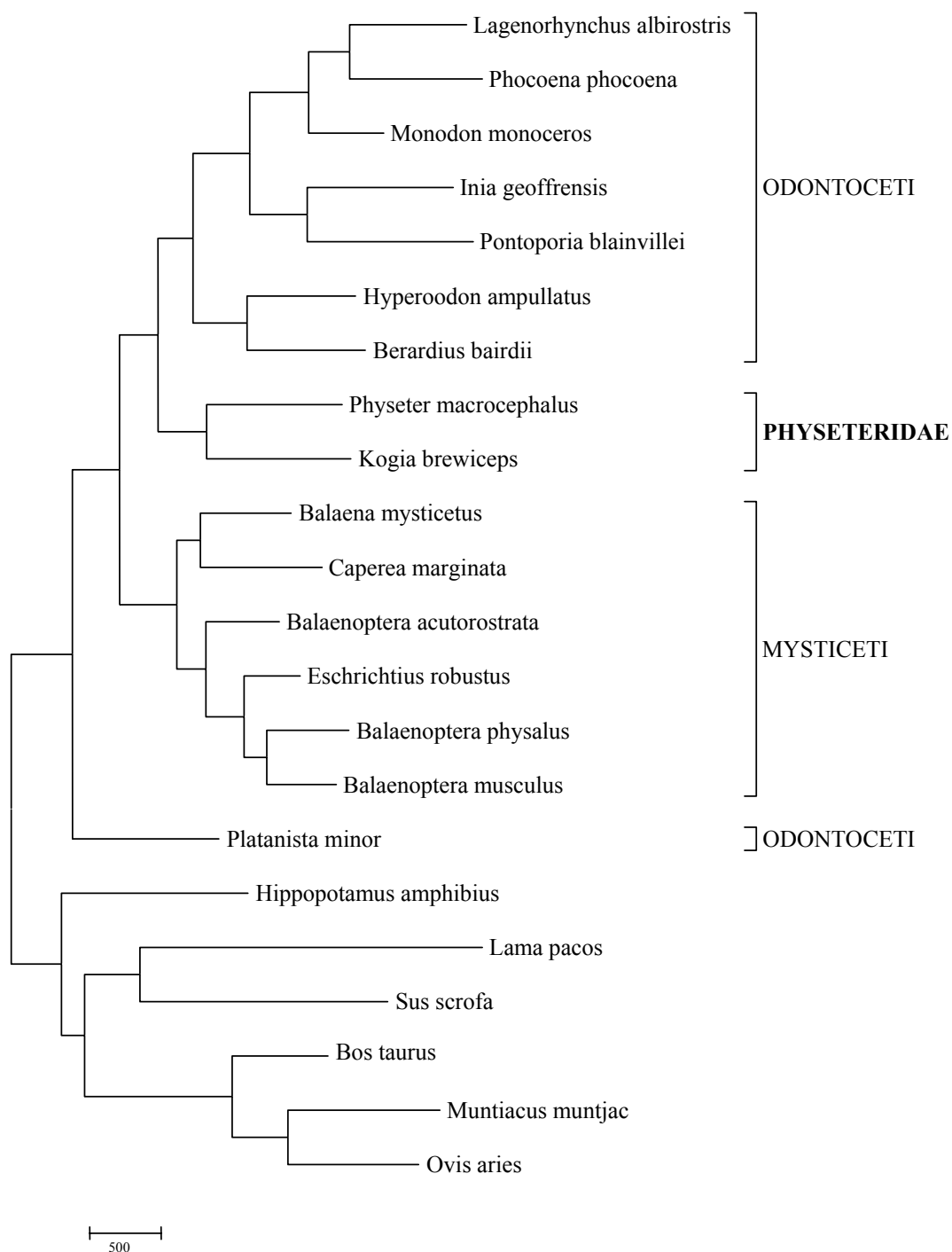
Filogenetska analiza

V analizi sorodstvenih odnosov izbrane skupine problematične znake izločamo kumulativno po korakih, od tistih z najvišjim deležem homoplastičnosti do znakov z najnižjim. V vsakem koraku po izločanju naredimo filogenetsko analizo, ki nam poda kladogram ter podporo za alternativne hipoteze o položaju problematičnega monofiluma. Opazujemo spremembe v položaju preiskovanega monofiluma in statistični podpori različnim hipotezam.

3.2 PRIMERI

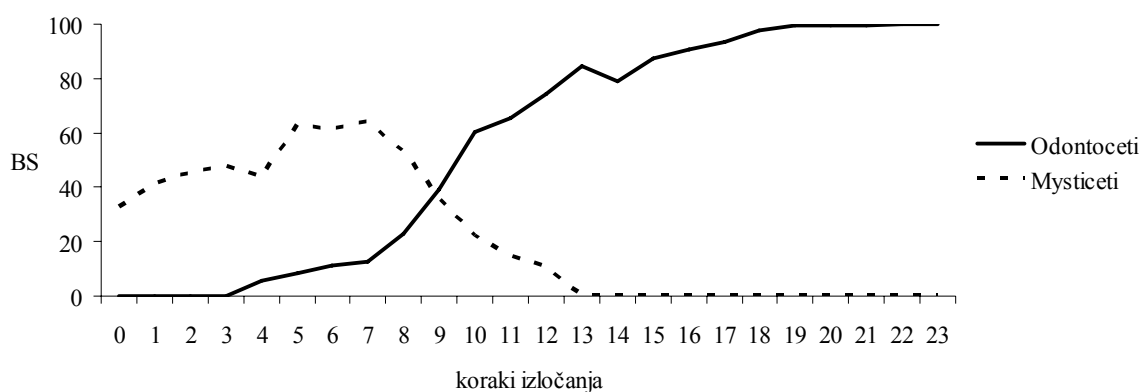
3.2.1 Glavači

Prvotna filogenetska analiza brez izločanja znakov je sicer pokazala glavače kot del zobatih kitov (z zelo nizko podporo BS = 22,3%; slika 4), vendar pa je postavila induško pliskavko kot bazalen takson vsem kitom, kar je v nasprotju z apriorno hipotezo o monofiliji zobatih kitov (glej uvod – glavači). Položaj induške pliskavke kot pripadnice zobatih kitov v vseh ostalih študijah ni vprašljiv in že ob vključitvi štirih taksonov zunanje skupine (glej Arnason et al., 2004) lihoprstih kopitarjev je postavljena v ta klad. Zato smo rezultat prvotne filogenetske analize označili za artefakt, ter upoštevali (bootstrap in Bremerjevo) podporo skupini glavačev in zobatih kitov vključno z induško pliskavko. Bootstrap podpora tej skupini je 0% (program PAUP* bootstrap podpore nižje od 5% ne prikaže) in umerjena Bremerjeva podpora -39 korakov.

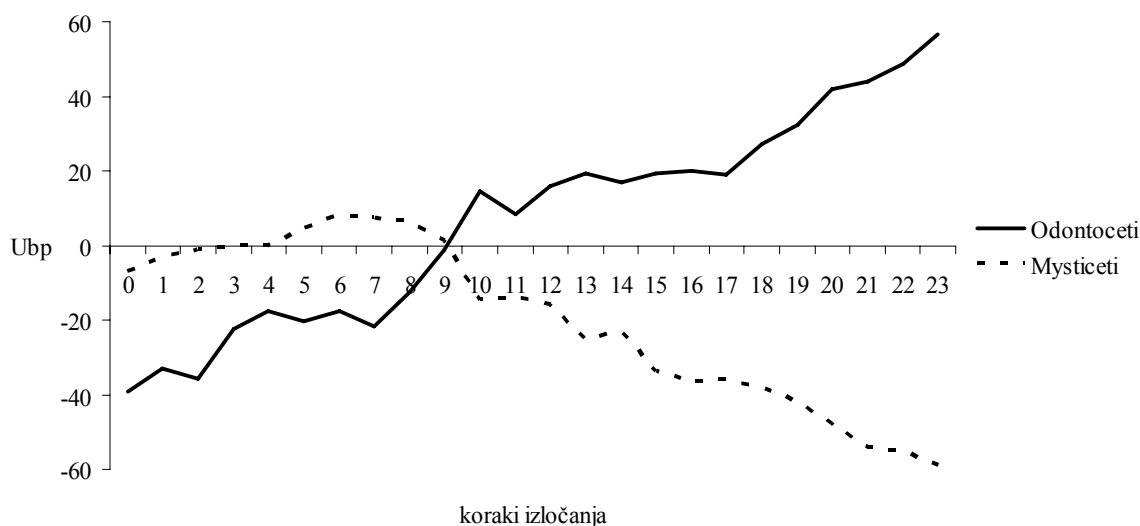


Slika 4: Optimalno drevo (filogram) kito-kopitarjev po kriteriju varčnosti, brez izločanja. V nasprotju s apriorno hipotezo je induška pliskavka (*Platanista minor*) postavljena kot bazalni takson drugim kitom (glej besedilo). Merilo predstavlja 500 korakov.

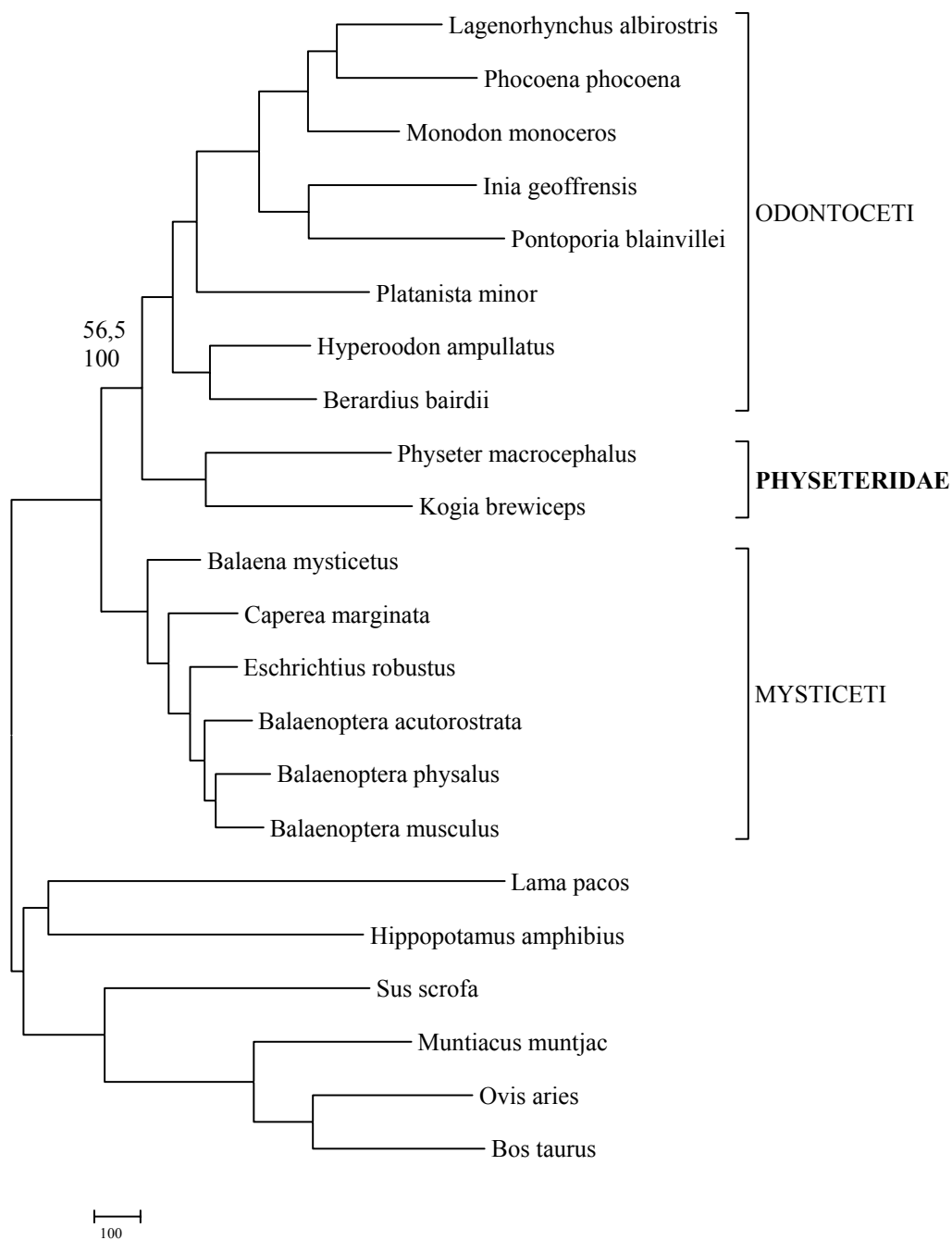
V analizi sorodstvenih odnosov po predstavljeni metodi je prepričljivo podporo dobila tradicionalna hipoteza, na kar kaže naraščanje podpore tej hipotezi ob izločanju (glej sliki 5 in 6). Najvišjo podporo za katerokoli hipotezo je pokazala analiza z izločenimi znaki, katerih delež homoplastičnosti je večji od 0,044. V tej analizi imajo glavači kot del monofiluma zobatih kitov BS = 100% in umerjeno Bremerjevo podporo 56,5 relativnih korakov (slika 7).



Slika 5: Grafikon spreminjanja bootstrap podpore (BS, vrednosti v odstotkih) alternativnima sestrskima skupinama glavačev v odvisnosti od izločanja znakov. Odontoceti – glavači del monofiluma zobati kiti, Mysticeti - glavači kot sestrška skupina vosatim kitom.



Slika 6: Grafikon spreminjanja umerjene Bremerjeve podpore (Ubp) alternativnima sestrskima skupinama glavačev v odvisnosti od izločanja znakov. Odontoceti - glavači del monofiluma zobati kiti, Mysticeti - glavači kot sestrška skupina vosatim kitom.



Slika 7: Drevo z najvišjo podporo preizkušani hipotezi (glavači kot bazalna skupina zobatih kitov) v postopku izločanja. Izločeni znaki, katerih delež homoplastičnosti je večji od 0,044. Zgornje število je umerjena Bremerjeva podpora, spodnje bootstrap podpora (v odstotkih). Merilo predstavlja 100 korakov.

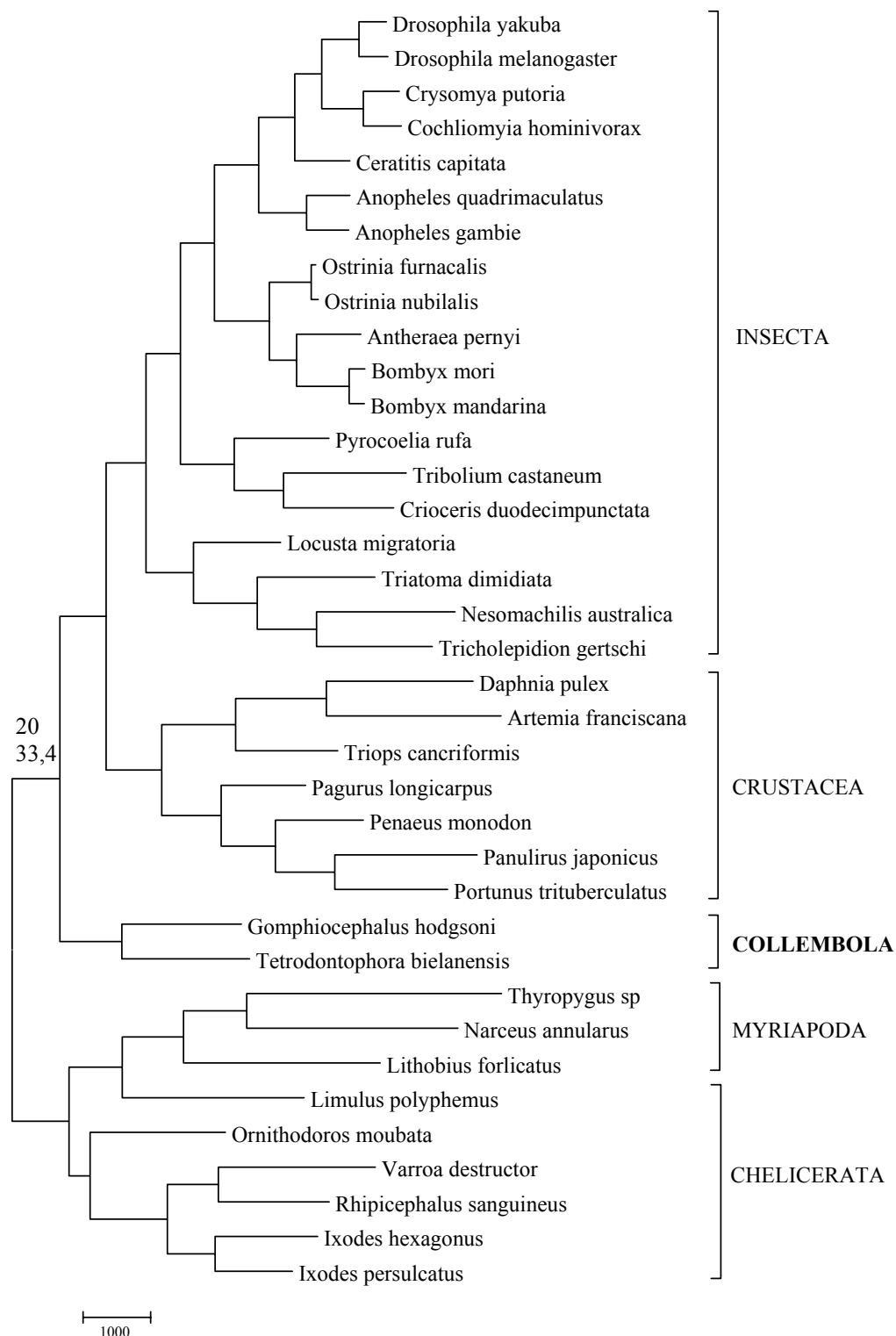
3.2.2 Skakači

Korenina členonožcev, ugotovljena v preliminarnem poskusu, je bila med pravimi pipalkarji (*Euchelicerata*) in čeljustarji (*Mandibulata*).

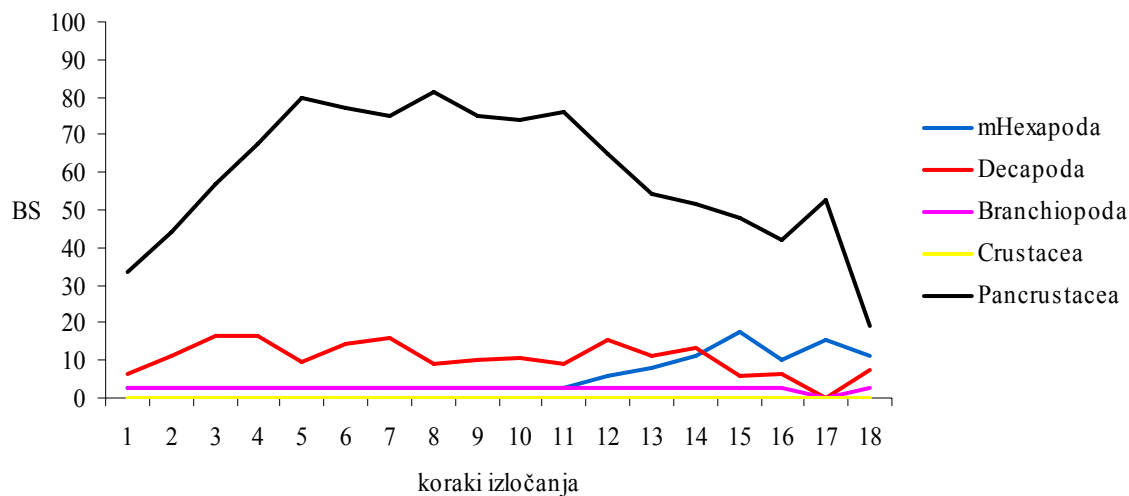
Analiza brez izločanja znakov je podala predstavnika monofiletske skupine skakači kot bazalno vejo skupine Pancrustacea (slika 8). Bootstrap⁴ podpora je nizka (33,4%), Bremerjeva podpora pa je 20 korakov za spremembo te topologije. Klasična hipoteza, ki predvideva, da so skakači del monofiletskih šesteronožnih členonožcev, ima bootstrap podporo pod 5%, drevo s tako topologijo pa je potrebovalo 77 dodatnih korakov.

Največja podpora katerikoli topologiji v procesu izločanja je po izločenih znakih z deležem homoplastičnosti višjim ali enakim 0,135. Takrat ima topologija (slika 9), ki predvideva skakače kot bazalno skupino v skupini Pancrustacea bootstrap podporo 80% (slika 10) in umerjeno Bremerjeveo podporo 53 relativnih korakov (slika 11). Alternativna hipoteza, monofiletski šesteronožni členonožci, ima bootstrap podporo manjšo od 5% in umerjeno Bremerjevo podporo -67. To pomeni, da je drevo, ki podpira tako topologijo, 67 relativnih korakov daljše od optimalnega.

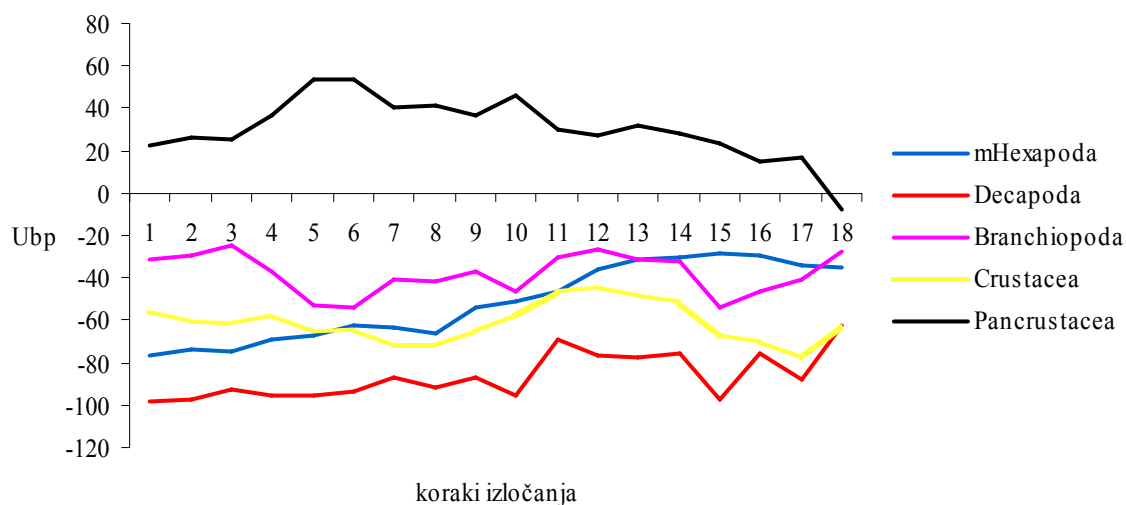
⁴ Bootstrap podpora za monofilum *Pancrustacea* pomeni podporo za (*Collembola+Branchiopoda+Decapoda+Insecta*), pri čemer je lahko položaj skakačev tudi znotraj ene od preostalih skupin, tudi alternativnih hipotez (*Collembola+Branchiopoda*), (*Collembola+Decapoda*) ali (*Collembola+Insecta*). Zato smo od dobljene BS podpore za Pancrustacea odšteli podpore teh hipotez. Ker PAUP* skupinam, katerih podpora je nižja od 5%, pripiše vrednost nič, smo v takem primeru odšteli povprečno morebitno podporo, to je 2.5%.



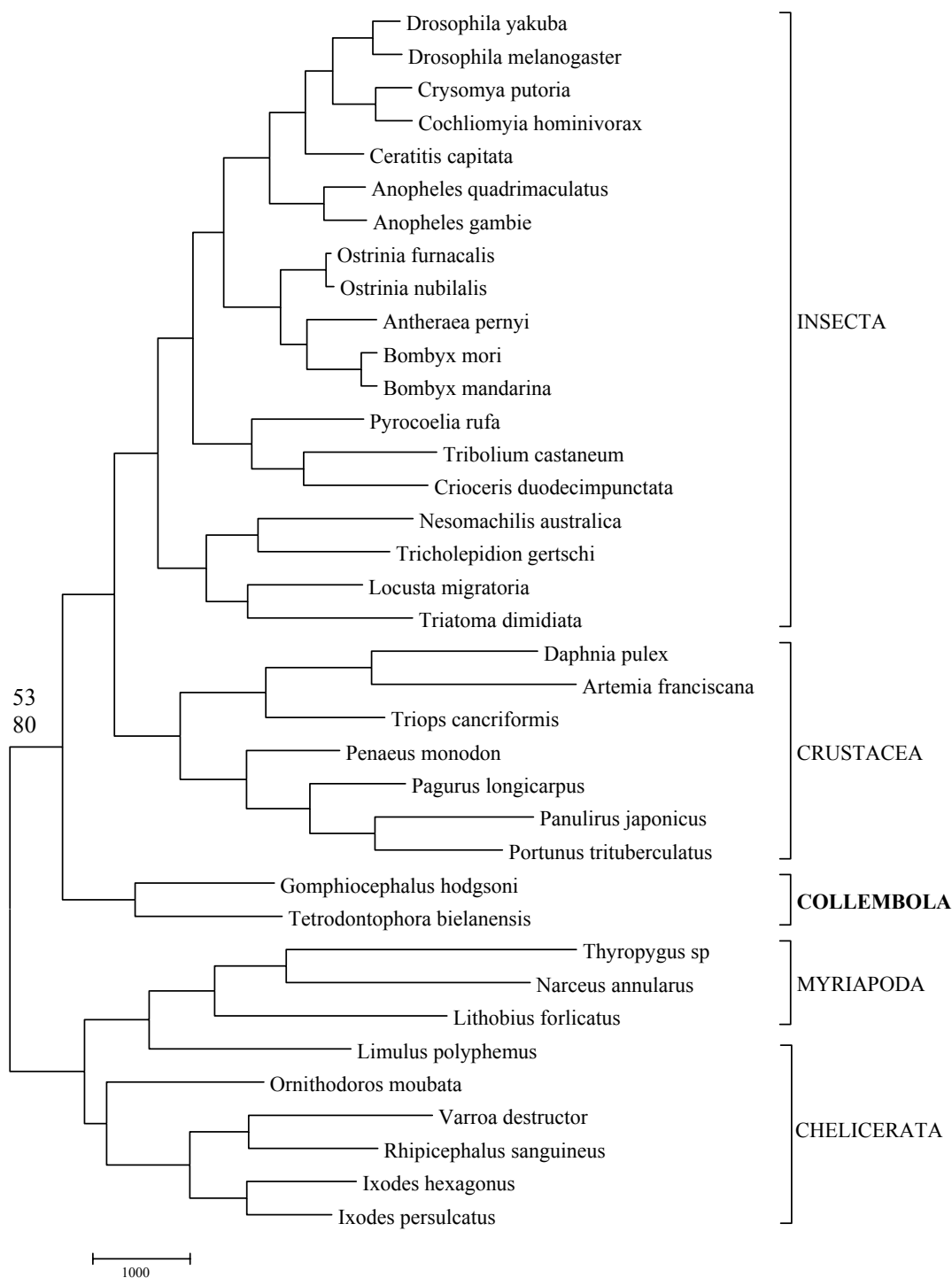
Slika 8: Najboljše drevo po kriteriju največje varčnosti v primeru skakačev (Collembola), brez izločanja znakov. Podpora razvejišču je napisana nad vejo, ki vodi k njemu. Zgornje število je umerjena Bremerjeva podpora, spodnje bootstrap podpora (v odstotkih). Merilo predstavlja 1000 korakov.



Slika 9: Grafikon spreminjanja bootstrap podpore (BS, vrednosti v odstotkih) alternativnim položajem skakačev (Collembola) v kladogramu v postopku izločanja. Skakači kot del monofiluma Hexapoda oz. sestrski skupini Decapoda, Branchiopoda, Crustacea, Pancrustacea, Myriapoda ali Chelicerata.



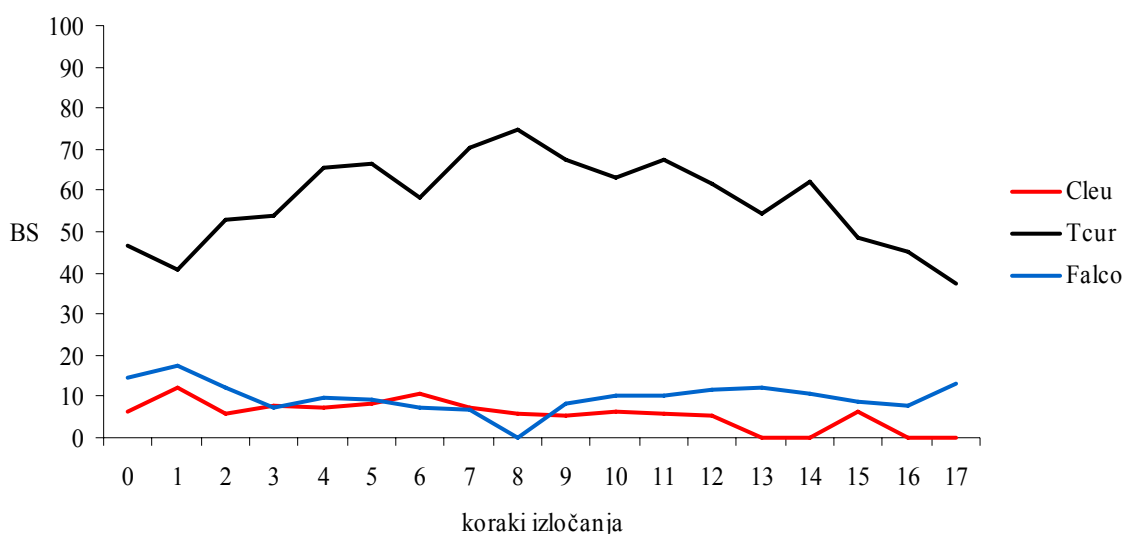
Slika 10: Grafikon spreminjanja umerjene Bremerjeve podpore (Ubp) alternativnim položajem skakačev (Collembola) v kladogramu v postopku izločanja. Skakači kot del monofiluma Hexapoda oz. sestrski skupini Decapoda, Branchiopoda, Crustacea, Pancrustacea, Myriapoda ali Chelicerata.



Slika 11: Najboljše drevo v primeru skakačev (Collembola) po kriteriju največje varčnosti v petem koraku izločanja, izločeni znaki z deležem homoplastičnosti večjim od 0,135. Podpora razvejišču je napisana nad vejo, ki vodi k njemu. Zgornje število je umerjena Bremerjeva podpora, spodnje bootstrap podpora (v odstotkih). Merilo predstavlja 1000 korakov.

3.2.3 Hoacin

V analizi filogenetskih odnosov z izločanjem je največjo podporo dobila monofiletska skupina (trogon+hoacin). Podpora tej skupini narašča iz prvotne bootstrap podpore za to hipotezo ki je 46,7% in Bremerjeve podpore, ki je enaka 0 korakov, do maksimalne (po osmem zaporednem izločanju, izločeni znaki z deležem homoplastičnosti večjim od 0,13), ki je BS = 74,7% (slika 12), umerjena Bremerjeva podpora pa je 22,4 relativnih korakov.



Slika 12: Grafikon spreminjanja bootstrap podpore (BS) alternativnim sestrskim skupinam hoacina (*Opisthocomus hoazin*) v kladogramu v postopku izločanja. Cleu – golob, Tcur – trogon, Falco - sokol. Ostale skupine (bananojed, kukavica in ani) niso bile podprte.

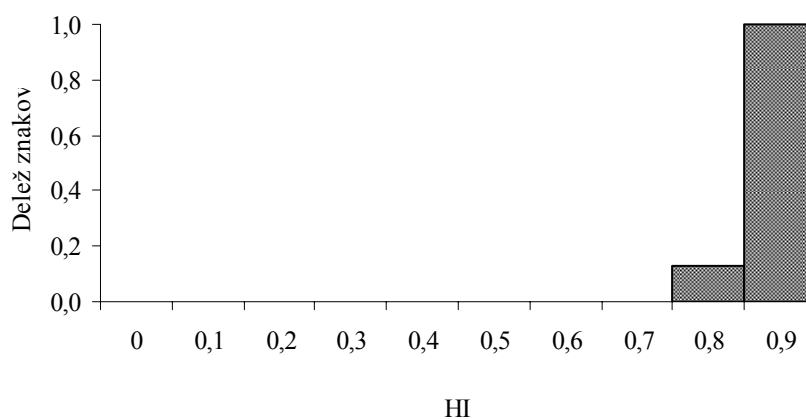
3.2.4 Analiza izločenih znakov po indeksu homoplastičnosti (glavači)

Po pričakovanju imajo znaki z najvišjim deležem homoplastičnosti, ti znaki so bili izločeni v prvih korakih, tudi najvišji indeks HI (slika 13a). Vendar pa na spremembo podpore nimajo bistvenega vpliva (slika 5). Največja sprememba v bootstrap podpori med izločanjem je v 9. in 10. koraku (delež homoplastičnosti 13,9 in 13) (glej sliko 13b in c, ter 13d in e; glej tudi sliko 5), ko se podpora monofiliji zobatih kitov poveča za 16,5 in 20,8

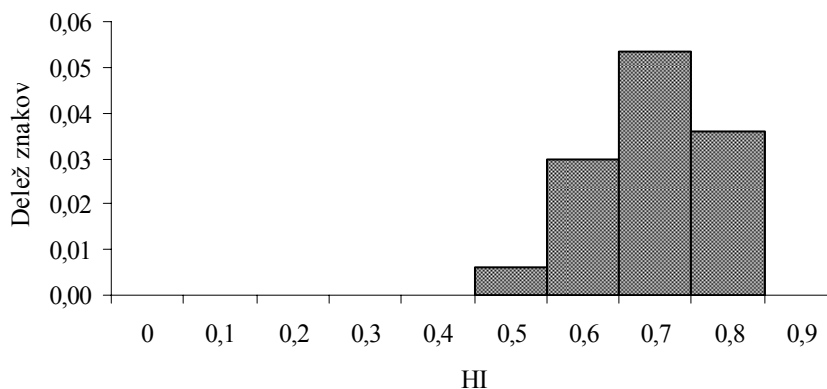
odstotnih točk. Podpora nato raste do BS=100% (slika 13f in slika 5). Znaki, ki so najbolj vplivali na spremembo podpore pripadajo razredom, katerih HI je nižji od 0,8.

Znaki, ki jih je kot problematične določila analiza homoplastičnosti niso enaki tistim, ki jih identificira analiza po HI (slika 13f).

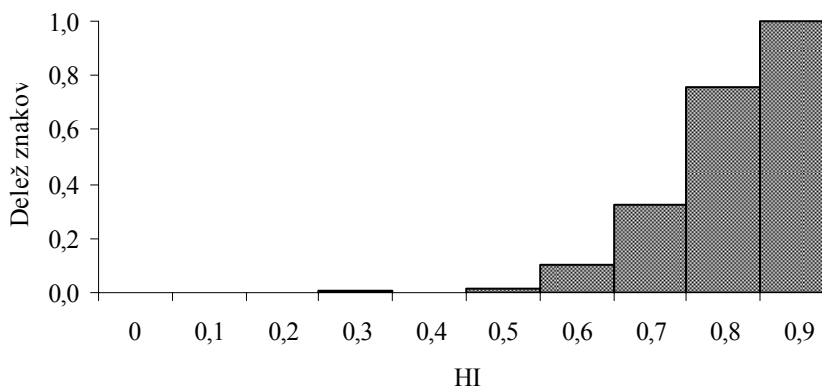
a.) znaki izločeni v prvem koraku



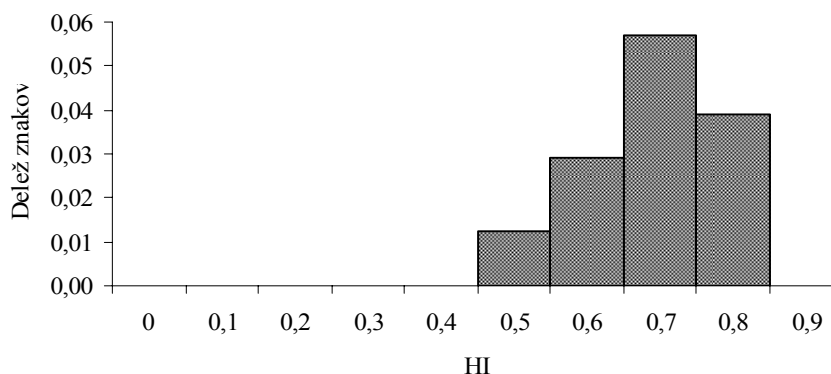
b.) znaki, katerih izločitev je najbolj prispevala k spremembi podpore (9. korak)



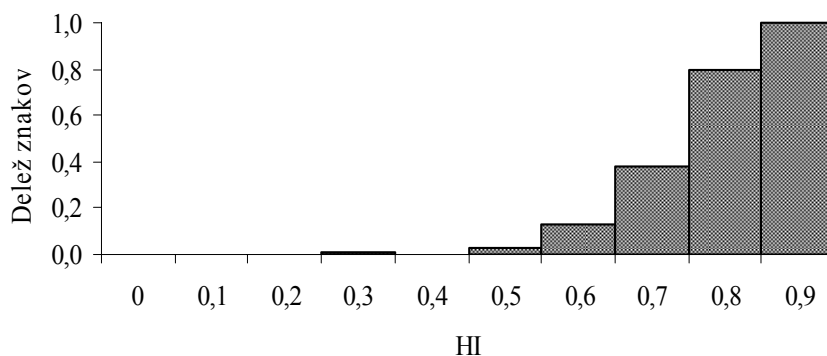
c.) vsi izločeni znaki v 9. koraku



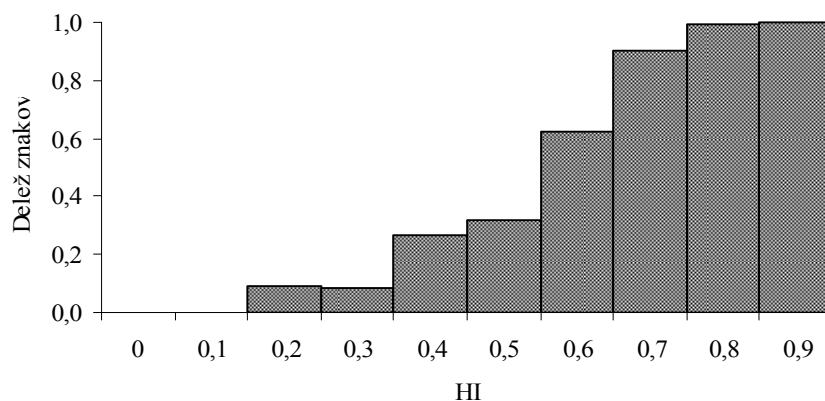
d.) znaki, katerih izločitev je najbolj prispevala k spremembi podpore (10. korak)



e.) vsi izločeni znaki v 10. koraku



f.) najvišja podpora hipotezi (glavači+zobati kiti), vsi izločeni znaki



Slika 13a-f: Razporeditev izločenih znakov po razredih HI. Izločeni znaki so prikazani kot delež vseh znakov v nekem razredu: a) znaki izločeni v prvem koraku. Znaki z največjim deležem homoplatičnosti imajo tudi najvišji HI. b) in d) znaki, katerih izločitev je najbolj prispevala k spremembi podpore (9. in 10. korak). c) in e) vsi izločeni znaki (9. in 10. korak). f) najvišja podpora hipotezi (glavači+zobati kiti), vsi izločeni znaki.

4 RAZPRAVA

4.1. METODA

Predlagamo rešitev problema konflikta znakov z izločanjem tistih, ki nasprotujejo apriorni informaciji na nižjih ravneh kladograma. Izločanje utemeljujemo na logični predpostavki, da znaki, ki nasprotujejo privzetim monofiletskim skupinam v filogenetski hierarhiji, ne morejo pravilno podpirati odnosov skupin nad njimi. Pogoj za delovanje metode je torej dobro poznavanje skupin na nižjih hierarhičnih nivojih. Za preverjanje ali znak nasprotuje monofilijam uporabljamo kvartete. Na ta način ne preštejemo (glej spodaj) preprosto koliko substitucij je potrebnih na kladogramu, pač pa neposredno ugotavljamo nasprotovanje točno določenim skupinam. Predvidevamo, da s tem določimo problematične znake v dolgih vejah. Z izločanjem teh znakov zmanjšamo vpliv privlaka dolgih vej, ki je znan problem metode največje varčnosti (Felsenstein, 1978; glej uvod).

Izločanje znakov, ki zmanjšuje pomen transformacij, in način identifikacije znakov dajeta večji vpliv počasneje spreminjajočim se znakom, še posebej tako imenovanim komplementarnim sinapomorfijam (Carine in Scotland, 1999), zato je metoda blizu taksičnemu pristopu.

Da znak v kvartetu nasprotuje monofiletski skupini sta potrebni dve dolgi veji. Ker v predstavljenih primerih v analizo znakov nismo vključili problematičnega taksona (glavači, skakači oz. hoacin), na to dolgo vejo nismo neposredno vplivali. Za tak pristop smo se odločili, da bi zagotovili čim večjo neodvisnost med analizo znakov in določitvijo položaja problematične skupine v sorodstveni analizi. S skrajšanjem dolgih vej bližnjih skupin smo kljub temu zmanjšali možnost privlaka dolgih vej. Skrajšanje dolge veje problematičnega taksona bi verjetno močno povečalo učinkovitost metode, zato predlagamo, da je tudi ta vključen v analizo znakov.

Predlagana metoda je podobna postopku, s katerim sta Roelants in Bossuyt (2005) ugotavljala robustnost položaja korenine žabam (Anura). Taksone sta razdelila v »filogenetsko dobro podprte« skupine, ki sta jih določila na podlagi prvotne analize. Za

vsak znak sta ocenila število substitucij znotraj vsake od skupin ter jih nato seštela. Na podlagi števila substitucij sta znake razdelila v diskretne kategorije od hitro spreminjajočih se znakov (veliko število substitucij) do nevariabilnih (brez substitucij). Kategorije sta zaporedno izločala iz analize in preverjala spremembo bootstrap podpore nasprotujočima hipotezama. Ta postopek se od v diplomskem delu predstavljene metode razlikuje v dveh pomembnih točkah, izboru monofilij in načinu določanja problematičnih znakov. »Filogenetsko dobro podprte« skupine so določene na podlagi istih znakov kot so uporabljeni v kasnejši analizi. Na ta način ocenjujeta znake na podlagi njih samih, kar vodi v krožno utemeljevanje. Poleg tega je zaradi tako omejenega nabora znakov zaupanje v te skupine majhno. S postopkom določita le znake z mnogo substitucijami in ne neposredno znake, ki nasprotujejo »dobro podprtim« hipotezam.

Filozofsko ozadje

Tako nasprotniki (e.g. Kluge, 1997) kot tudi zagovorniki (Lecointre in Deleporte, 2004) obteževanja znakov se strinjajo, da je v okviru kladistike treba uporabiti vse relevantne podatke oz. sinapomorfije. Razlika je torej v interpretaciji, kjer prvi trdijo, da so to vsi razpoložljivi znaki, drugi pa relevantnost znakov določijo z analizo po nekem kriteriju (e.g. zaporedna aproksimacija; Farris, 1969). V osnovi se strinjamo z zagovorniki obteževanja znakov, vendar menimo, da znaka ne moremo ocenjevati neposredno na podlagi neke skladnosti z drugimi, v analizi uporabljenimi znaki. Tak postopek je krožen, saj na tem principu temelji tudi filogenetska analiza.

Po naši definiciji je znak relevanten, če nam potencialno lahko razkrije filogenetske odnose, če vsebuje »filogenetsko informacijo«. Znaki, ki so zavajajoči na nižji hierarhični ravni, tudi na višjih ravneh ne morejo razkrivati sorodstvenih odnosov.

Glavni argument nasprotnikov obteževanja znakov je edinstvenost zgodovinskih dogodkov. Na podlagi sprememb stanj znaka v enem delu drevesa ne moremo sklepati na večjo verjetnost spremembe tudi v drugih delih drevesa saj so ti dogodki med seboj neodvisni (Kluge, 1997a). Zagovorniki obteževanja pa izpostavljajo dejstvo, da so nekateri nukleotidi (znaki) pač pod manjšim selekcijskim pritiskom in so zato substitucije verjetnejše (e.g. Dixon in Hillis, 1993). V pogosto uporabljani metodi največjega verjetja

eksplicitno modeliramo različno variabilnost znakov. V tukaj predlagani metodi te dileme ni. Ker je problematična skupina na višji hierarhični ravni istega dela kladograma kot apriorne monofilije, ne gre za ekstrapolacijo rezultatov analize znakov na druge taksone.

Zaupanje v rezultat analize je omejeno z zaupanjem v predpostavke. Več predpostavk vključimo v predznanje bolj bo od njih odvisen rezultat analize in s tem zmanjšana moč testa. Manjšo moč analize sorodstvenih odnosov nekateri raziskovalci (e.g. Kluge, 1997b) navajajo kot argument proti obteževanju znakov. Menimo, da to ne izpodbija same metode, temveč le omejuje njeno uporabnost. Uporaba je smiselna v primerih, ko predhodne analize vseh znakov ne razrešijo konflikta znakov.

Delež homoplastičnosti, izločanje v korakih

Nek znak je lahko bolj ali manj problematičen, odvisno od deleža taksonov v katerih je stanje tega znaka homoplastično. Znak torej ni popolnoma informativen ali popolnoma zavajajoč. Analiza znakov na podlagi kvartetov je primerna saj tako ugotavljamo dejansko problematične znake. Po drugi strani pa v posamezno analizo znakov ne vključimo vseh taksonov, ki so hierarhično pod problematičnim razvejiščem. Za razlago odnosov v neki skupini niso pomembna le stanja znaka taksonov te skupine pač pa stanja znaka vseh v analizo vključenih taksonov. Tako ima lahko že sprememba zunanje skupine močan vpliv na odnose v kladogramu. V primeru glavačev je bila v sorodstveni analizi brez izločanja znakov induška pliskavka postavljena kot bazalen kit. Analiza z vključeno zunanjo skupino lihoprstih kopitarjev (kot v študiji Arnason et al., 2004) pa jo postavi med zobate kite.

Homoplastičnosti znaka v kvartetu torej ne moremo posplošiti na celoten nabor uporabljenih taksonov. Zato kot rezultat analize znakov podamo delež kvartetov v katerih je nek znak nasprotoval apriornim sorodstvenim odnosom. Prav tako ni razloga, da bi omejili izločanje znakov na neko vrednost deleža homoplastičnosti. Bolj smiselno je izločati znake v korakih ter ugotavljati tendenco, to je spreminjanje podpore različnim sorodstvenim hipotezam, skozi celoten postopek izločanja znakov.

V diplomskem delu smo podporo hipotezam kvantificirali z maksimalno podporo (bootstrap in Bremerjevo), ki so jo hipoteze dobile v postopku izločanja. Konsistentnosti

metode, to je naraščanja podpore le eni hipotezi skozi postopek izločanja, nismo ocenjevali kvantitativno. Na nedoslednost metode bi sklepali, če bi izmenoma dobivali dobro podprte nasprotujoče hipoteze, česar v poskusih nismo zasledili (glej slike 5, 6, 9, 10 in 12). Tak rezultat bi lahko pomenil tudi odsotnost zgodovinskega signala v podatkih (glej spodaj).

»Šum« in zgodovinski signal

Rezultati analize znakov po indeksu homoplastičnosti kažejo, da najbolj homoplastični znaki niso nujno tudi najbolj problematični pri ugotavljanju sorodstvenih odnosov na višjih hierarhičnih ravneh. To je v skladu z ugotovitvijo, da s substitucijami nasičeni znaki, tako imenovan šum, nimajo velikega vpliva na delovanje metode največje varčnosti (e.g. Wenzel in Siddall, 1999). Še več, izločitev znakov z najvišjim indeksom homoplastičnosti poveča podporo (prepričljivo zavrjeni, glej uvod in rezultate) Milinkovitchevi hipotezi o glavačih kot vosatih kitih, kar kaže, da ti znaki kljub zasičenosti s spremembami podpirajo sorodstvene odnose. Da s substitucijami nasičeni znaki lahko podpirajo »dejanske« sorodstvene odnose so pokazali tudi v drugih študijah (Björklund, 1999; Källersjö et al., 1999). V takih primerih seveda ne moremo govoriti o »šumu«. Obteževanje znakov zgolj na podlagi opaženega ali predpostavljenega števila substitucij je zato vprašljivo.

Znaki, ki imajo indeks HI med 0,6 in 0,8, najbolj vplivajo na spremembo topologije. Ti znaki so homoplastični le v nekaterih taksonih oz. problematične spremembe stanj so le na nekaterih vejah. Ugotovitev je v skladu z našo predpostavko, da je sporen rezultat sorodstvenih analiz brez izločanja znakov posledica privlaka dolgih vej, ter da s predlagano metodo te veje skrajšamo. Ali so te homoplazije dejansko vezane na dolge veje bi lahko preverili s podrobnejšo analizo problematičnih znakov. Za vsak znak bi bilo treba ugotoviti na kateri veji so substitucije problematične. Če bi bila večina substitucij znakov, katerih izločitev je najbolj vplivala na spremembo topologije problematičnih na istih vejah, bi to potrdilo našo predpostavko o vplivu privlaka dolgih vej.

Na podlagi dosledne podpore le eni hipotezi v vseh treh primerih trdimo, da predlagana metoda odkriva zgodovinski signal v podatkih. Poleg sorodnosti organizmov je zgodovinski signal lahko tudi posledica nekaterih drugih procesov (glej uvod; Eyre-Walker, 1991; Collins et al., 1994; Lockhart et al., 1994). To predstavlja problem pri

metodah, ki vrednotijo znake na podlagi skladnosti z drugimi znaki. Če nek signal v podatkih prevladuje, bodo te metode določile kot »slabe« znake, ki mu nasprotujejo. Z vrednotenjem na podlagi sorodstvenih odnosov se temu izognemo. Identificiramo in izločimo tudi znake, ki vsebujejo zgodovinski signal, ki ni posledica sorodstvenih odnosov. Kljub temu priporočamo previdnost pri interpretaciji rezultatov, ki niso tako nedvoumni kot v primeru glavačev.

Vpliv apriornih monofilij na delovanje metode

V nalogi smo preizkusili delovanje metode na treh izbranih primerih. Izločanje znakov je vedno vodilo v povečanje podpore eni hipotezi v večjem delu (hoacin) postopka izločanja oz. v celotnem postopku (glavači).

Moč signala je v različnih primerih močno variirala. Največje in med izločanjem najbolj dosledne podpore so bili deležni glavači, kar kaže na pomembnost izbora predpostavk. Pri glavačih je poznavanje sorodstvenih odnosov kitov omogočalo izbor monofiletskih skupin neposredno pod problematičnim razvejiščem. Zato smo lahko neposredno določili znake, ki so problematični. Pri problemih položaja hoacina in monofilije šesteronožnih členonožcev smo zaradi nepoznavanja odnosov oz. slabega nabora taksonov vzeli problematični skupini najbližje zanesljive monofilume. S tem nismo določili znakov, ki so zavajajoči v problematičnem taksonu ali hipotetičnih sestrskih skupinah, pač pa znake zavajajoče v manj sorodnih skupinah. Ni nujno, da ti znaki zakrivajo filogenetsko informacijo tudi na problematičnem razvejišču.

Predlagamo uporabo metode le, če imamo v naboru zastopane vse večje skupine pod problematičnim razvejiščem, ter če naše poznavanje sorodstvenih odnosov omogoča postavitev apriornih monofilij pod problematičnim razvejiščem.

Predlagani testi

Z vedno obsežnejšimi knjižnicami molekularnih zaporedij postajajo sorodstveni odnosi organizmov jasnejši. Vedno več pa je primerov, ko imamo mnogo znakov in načeloma dobro podprte sorodstvene odnose, pri tem pa položaj neke skupine ni jasen. Še več, različne vrste znakov (mitohondrijska, jedrna DNA, morfološki znaki) prepričljivo

podpirajo nasprotujoče hipoteze. Tak aktualen problem je sestrška skupina vretenčarjev, kjer najnovejše, na molekularnih znakih narejene raziskave (e.g. Delsuc et al, 2006) kažejo, da so to plaščarji (Tunicata oz. Urochordata) in ne brezglavci (Cephalochordata) (morfološko) (Schaeffer, 1987). Predlagamo uporabo metode na tem primeru.

Delovanje metode bi bilo dobro preizkusiti tudi s ponovno študijo položaja skakačev z uporabo istih genov, a z ustrežnejšim naborom taksonov (glej diskusija – skakači). Pričakujem, da bo hipoteza o polifiliji šesteronožnih členonožcev dobila še prepričljivejšo podporo.

Koristen način za testiranje je tudi uporaba simulacij. Te imajo v primerjavi z dejanskimi zaporedji dve pomembni prednosti. (1) Poznamo dejanske »sorodstvene« odnose uporabljenih taksonov. (2) Z njimi lahko simuliramo različne modele evolucije znakov in tako preverimo občutljivost metode in njene omejitve. Pomembno bi bilo preveriti obnašanje metode pri različnih razmerjih dolžin vej ter različnemu številu dolgih vej.

4.2 PRIMERI

4.2.1 Glavači

Rezultati analize so v skladu z dobro podprto in splošno sprejeto hipotezo o sorodstvenih odnosih znotraj skupine kitov. Primer kitov je idealen, saj smo, zaradi jasnega filogenetskega položaja večine skupin lahko postavili apriorne hipoteze monofiletskih skupin le eno hierarhično raven pod problematičnim razvejiščem.

Uspeh metode je razviden tudi v luči najnovejše študije gena za citokrom b, v kateri sta May-Collado in Agnarson (2006) na širokem vzorcu taksonov, 66 kitov in 24 zunanjih skupin ter ob uporabi Bayesovega pristopa sicer dobila skoraj enak rezultat sorodstvenih odnosov kot Nikaido et al. (2001), vendar z nizko podporo za monofilijo zobatih kitov (posteriorna verjetnost je bila 67%).

4.2.2 Skakači

Interpretacija podpore kladu Pancrustacea (bootstrap podpora) je v okviru hipotetične skupine čeljustarji, saj kladogram nima korenine. Po tej hipotezi imajo členonožci korenino med čeljustarji (Mandibulata; stonoge, raki in šesteronožni členonožci) in pipalkarji, kar je v skladu z večino študij (ter preliminarnim poskusom – glej rezultate). Nekatere raziskave živčevja (Dove in Stollewerk, 2003) in predvsem molekularnih znakov (e.g. Cook et al., 2001; Mallatt et al., 2004) pa temu nasprotujejo, ter postavljajo korenino med skupini Pancrustacea in Paradoxopoda (Myriapoda+Chelicerata). Ker nobena hipoteza do sedaj ni predvidela položaja skakačev kot sestrskes kateri od skupin paradoksopodov, menimo, da nekoreninjenost kladograma ne zmanjšuje relevantnosti podpore skupini Pancrustacea.

Povečanje podpore topologiji, ki postavlja skupino skakači kot bazalno vejo skupine Pancrustacea, v procesu izločanja znakov za več kot dvakrat, je vsekakor presenetljiva. Čeprav so že nekatere predhodne analize molekularnih znakov podprle hipotezo o polifiliji šesteronožnih členonožcev, pa so to bile (tako kot tudi ta) predvsem analize mitohondrijskih genov (Nardi et al., 2003; Cameron et al., 2004; Negrisolo et al., 2004; Cook et al., 2005). Le Cook et al. (2005) so v svoji analizi celotne mitohondrijske kodirajoče DNA dobili dobro podporo tej hipotezi. Od študij drugih znakov sta dali podporo polifiliji šesteronožnih členonožcev še raziskavi Giribet et al. (2001, morfološki znaki in 9 lokusov) in Luan et al. (2005, nuklearna rRNA). Pri prvi je šlo očitno za artefakt, saj sta monofilijo razbila en pripadnik dvorepk (Diplura) in pa *Drosophila sp.*, ki sta bila prikazana kot del rakov. Luan et al. (2005) pa so imeli v analizi vključene le tri rake, tako da je bila pozicija solinskega rakca (*Artemia sp.*) znotraj šesteronožnih členonožcev lahko posledica privlaka dolgih vej. Neustrezen nabor taksonov za reševanje tega problema je splošna značilnost vseh molekularnih študij z izjemo Regier et al. (2004, 2005a), Giribet et al. (2005) in Mallatt in Giribet (2006), od morfoloških študij pa Bitsch C. in Bitsch J. (2004). Od analiz vseh znakov, mitohondrijskih in jedrnih kodirajočih in ribosomskih genov ter morfoloških znakov, ni podpirala monofilije še študija Giribet et al. (2004), ki je imela le tri predstavnike rakov, v kladogramu pa so bili očitni artefakti. Po drugi strani pa so tudi raziskave, ki podpirajo šesteronožne členonožce kot monofiletsko skupino (Kjer et al., 2004; Regier et al., 2004, 2005a; Mallatt in Giribet 2006), le redko dale močno podporo

tej hipotezi. Tako sta Mallatt in Giribet (2006) z Bayesovim pristopom in metodo največjega verjetja dobila močno podporo v korist monofilije šesteronožnih členonožcev, analizi po metodah največje varčnosti in razdalj LogDet pa je nista podprli. Dobro podporo z vsemi metodami je dala analiza Regier et al. (2005a), vendar pa niso bili vključeni predstavniki proturov (Protura). Giribet et al. (2005) so v združeni analizi (morfološki znaki in 9 lokusov jedrne in ribosomske DNA) sicer dobili kladogram z monofiletskimi šesteronožnimi členonožci, a brez statistične podpore.

Na podlagi tradicionalnih morfoloških znakov je monofilija šesteronožnih členonožcev na videz dobro utemeljena (Bitsch C. in Bitsch J., 2004; Giribet et al., 2004). Bitsch C. in Bitsch J. ugotavljata, da čeprav večina podatkov morfoloških znakov podpira šesteronožne členonožce kot monofiletsko skupino, ima ta (presenetljivo) malo sinapomorfij. Kot najočitnejši sta navedla redukcijo členov v tričlensko oprsje in tri pare oprsnihih okončin. Ostale potencialne sinapomorfne znake sta kritizirala. Npr. zlitje sekundarnih maksil v labium naj bi bil diagnostični znak šesteronožnih členonožcev, vendar je podobna struktura prisotna pri drobnonogah (Symphyla) (tudi Scholtz in Edgecombe, 2006; za kritiko preostalih znakov glej Bitsch C. in Bitsch J., 2004). Njuna analiza podpira hipotezo vzdušnjičarjev, kar je v neskladju s sedanjim konsenzom (glej uvod). Tudi Giribet et al. (2004) so navedli nedvoumne sinapomorfije iz njihove morfološke analize, ki podpira Pancrustacejsko hipotezo. Kljub temu so dobili razmeroma nizko podporo (BS = 61%) v korist monofilije šesteronožnih členonožcev.

Problematiko ugotavljanja sorodstvenih odnosov členonožcev sta Mallatt in Giribet (2006) opisala: »Postavitev filogenetskih odnosov znotraj členonožcev je težavna, saj so se pojavili pred približno 520 milijoni let in glavne skupine pipalkarjev, rakov, stonog in šesteronožnih členonožcev so bile navzoče že pred okoli 325 milijonov leti, torej so se njihove morfološke in genetske značilnosti zgodaj diverzificirale, čemur je sledilo dolgo obdobje evlucijskega spreminjanja«.

Kljub šibki podpori molekularnih znakov in malo sinapomorfijam v okviru morfologije, pa je monofilija šesteronožnih členonožcev najbolj prepričljiva hipoteza. Regier et al. (2004) menijo, da v večini študij: »ponujeni alternativni filogenetski sistemi (če sploh so kakšni)

niso prepričljivi iz več razlogov. Osredotočijo se na en sistem morfoloških znakov, predvidijo, da so stonoge najbližja zunanja skupina šesteronožnim členonožcem, ali uporabijo majhen ali analitsko kompleksen vzorec genetskih znakov«. V prihodnosti lahko pričakujemo večji doprinos k osvetlitvi problema s strani študij razvoja ter strukture živčnega sistema. Te se trenutno srečujejo s problemom pomanjkanja študij določenih skupin, kot so bazalni šesteronožni členonožci (e.g. Schachtner et al., 2005), ter konsenza o kodiranju znakov (Scholtz in Edgecombe, 2006).

Značilnost večine študij je neustrezna zastopanost taksonov za reševanje problema položaja bazalnih šesteronožnih členonožcev. Večinoma se raziskave osredotočajo na odnose višjih skupin členonožcev in je v analizah malo predstavnikov šesteronožnih členonožcev (e.g. Mallatt et al., 2004), ali pa so ti neustrezno zastopani. V drugem primeru v študijah običajno manjkajo mnogi predstavniki bazalnih skupin, kot so proturi, dvorepke, ščetinorepke, srebrne ribice, skakači, medtem ko je zastopano veliko taksonov iz skupine krilatih žuželk (e.g. Nardi et al., 2003). Tiste raziskave, ki so preučevale odnose med šesteronožnimi členonožci, pa so uporabile le nekaj predstavnikov rakov kot zunanjo skupino (e.g. Wheeler et al., 2001).

Enak problem kot Nardi et al. (2003) je imela tudi naša raziskava. Znotraj šesteronožnih členonožcev so bili, poleg mnogih krilatih žuželk, ena srebrna ribica in ena ščetinorepka, ter dva predstavnika skakačev iz skupine Poduromorpha, kar vsekakor ne predstavlja diverzitete te skupin. Poleg tega so manjkali predstavniki proturov in dvorepk, ki so hipotetične sestrške skupine skakačem. Skupino rakov sta zastopala le dva razreda od šestih (Regier et al., 2004), pri čemer so bili vsi predstavniki višjih rakov iz reda *Decapoda*. Pri stonogah smo imeli le enega predstavnika strig, dva predstavnika kačic, ter nobenega predstavnika skupin drobnonožke (*Symphyla*) in malonožke (*Pauropoda*). Predstavniki pipalkarjev so bili ostvar in taksoni le ene skupine pršic znotraj pajkovcev. Pomanjkljiv nabor taksonov znotraj predpostavljenih monofilij omejuje delovanje naše metode. Poleg tega pa pomanjkanje predstavnikov bazalnih šesteronožnih členonožcev še dodatno povečuje vpliv privlaka dolgih vej.

Rezultat analize je torej v nasprotju z najverjetnejšo hipotezo položaja skakačev. Vendar pa je bila pri izločanju znakov podpora vseskozi v korist pankrustacejske hipoteze (največ 80%), kar težko pojasnimo z naključjem in kaže na zgodovinski signal. Zaradi omejenega nabora apriornih monofilij na hierarhičnih ravneh pod obravnavanim problemom ne moremo z gotovostjo trditi, da je ta signal posledica sorodstvenih odnosov.

Danes je nabor vrst, katerih zaporedje celotne mitohondrijske DNA je dostopno, bistveno večji in poskus je mogoče ponoviti z ustrežnejšim izborom monofiletskih skupin in taksonov znotraj njih.

4.2.3 Hoacin

V tem primeru smo zaradi nejasnih hierarhičnih ravni hipotetičnih razvejišč tvorili kvartete z manj sorodnimi taksoni. Na ta način smo želeli posredno zmanjšati vpliv dolgih vej na hipotetične sestrške skupine.

Odnosi med skupinami (razredi in družinami) klada Neoaves niso jasni, kot kažejo nasprotujoči rezultati filogenetskih analiz Fain in Houde (2004), Harison et al. (2004) in Poe in Chubb (2004). Prav tako ni konsenza o sistematski pripadnosti nekaterih fosilov in posledično o času radiacije modernih ptic ter nastanka posamičnih skupin (Dyke in Van Tuinen, 2004).

Analiza filogenetskih odnosov z izločanjem daje največjo podporo sestorskemu odnosu hoacina in trogona. V svoji študiji so Sorenson et al. (2003) enak rezultat analize, vendar z bistveno nižjo podporo, pripisali privlaku dolgih vej in trdili, da je najbolj podprt sestrski odnos hoacina in golobov. Trditev je vprašljiva, saj je razmeroma visoka podpora temu odnosu lahko tudi posledica zelo podobne bazne sestave. Razmerja baz hoacina so namreč najbolj podobna razmerjem pri golobih, medtem so pri bananojedih in kukavicah drugačna (Sorenson et al., 2003). Poleg tega smo z našo metodo identificirali znake, ki so bili v dolgih vejah nasičeni s spremembami in smo tako zmanjšali vpliv naključja. Kljub temu se je podpora hipotetični monofiletski skupini (hoacin+trogon) še povečala, kar nasprotuje

trditvi, da je podpora skupini posledica vpliva znakov z veliko hitrostjo substitucij, pač pa kaže na signal v mitohondrijskem genomu. Pripadniki reda Trogoniformes v predhodnih analizah niso bili obravnavani kot morebitni najbližji sorodniki hoacina (e.g. glej Hedges et al., 1995). Razlog je, da si morfološko niso podobni. Za to skupino do sedaj ni bilo predlaganih sinapomorfij. Zgolj rezultat te raziskave ni dovolj za postavitev trogonov kot najbližjih živečih sorodnikov hoacinu (glej diskusijo-skakači), kaže pa, da bi bilo treba v bodočih analizah ta klad obravnavati kot možno hipotezo (c.f. Sorenson et al., 2003).

5 SKLEPI

Izločanje znakov na podlagi apriorne informacije zmanjša konflikt znakov.

Kot apriorno informacijo smo uporabljali dobro podprte monofiletske skupine, ki so hierarhično pod problematičnim razvejiščem.

S substitucijami nasičeni znaki, tako imenovan šum, lahko lahko razkrivajo sorodstvene odnose tudi na višjih hierarhičnih ravneh. Z uporabo kvartetov za identifikacijo znakov določimo tiste znake, ki so na hierarhičnem nivoju problema dejansko zavajajoči.

Znaki, katerih izločitev najbolj vpliva na spremembo podpor sorodstvenim hipotezam, imajo HI med 0,6 in 0,8. Predvidevamo, da so homoplazije zbrane na majhnem številu vej, na t.i. dolgih vejah.

Naraščanje podpore eni hipotezi v večini postopka izločanja znakov pomeni, da s predlagano metodo razkrivamo dejanski zgodovinski signal. Ta je poleg sorodnosti organizmov lahko posledica drugih procesov. Z uporabo apriornih monofilij v analizi znakov izločimo tudi znake, katerih vzorci stanj znakov so posledica teh procesov.

Z vključitvijo sorodstvenih hipotez v predznanje zmanjšamo zaupanje v rezultat. Uporaba metode je smiselna, ko kljub povečanju števila znakov analize brez izločanja ne razrešijo konflikta znakov.

Delovanje metode je treba preveriti na simuliranih zaporedjih. S tem bi preverili občutljivost metode na različna razmerja v dolžinah vej.

6 POVZETEK

Vzorci stanj znakov, ki ne odražajo sorodstvenih odnosov taksonov (t.i. homoplazije), predstavljajo problem v filogenetski analizi. Pogosto uporabljena metoda največje varčnosti je posebej občutljiva na veliko število homoplazij, če te niso enakomerno razporejene po vejah. Te imenujemo dolge veje, pojav, ko jih analiza združi, pa privlak dolgih vej. Dodaten problem so nenaključne homoplazije, ko so vzorci stanj znakov posledica različnih zgodovinskih procesov. V takih primerih lahko v filogenetski analizi dobimo visoko podporo za neobstoječe sorodstvene odnose. Na podlagi razmišljanja, da so znaki, ki ne podpirajo »znanih« sorodstvenih odnosov na nižjih hierarhičnih ravneh, verjetno zavajajoči tudi na višjih, smo razvili novo metodo za obteževanje znakov. V okviru metode uporabljamo apriorno znanje za določitev problematičnih znakov. Metodo smo preizkusili na treh primerih: položaju glavačev (*Physeteridae*) znotraj kitov (*Cetacea*), položaju skakačev (*Collembola*) v skupini *Pancrustacea* in sestrski skupini hoacina (*Opisthocomus hoazin*, *Aves*). V poskusih smo uporabili DNA zaporedja kito-kopitarjev (*Cetartiodactyla*), členonožcev (*Arthropoda*) in ptic (*Aves*).

Predlagana metoda je uporabna v primerih, ko imamo načeloma jasne sorodstvene odnose, položaj neke skupine pa ni razrešen. Zanesljive monofiletske skupine pod hierarhično ravno problematičnega razvejišča uporabimo kot apriorno informacijo v analizi znakov. To izvajamo na kvartetih v katerih sta dva taksona pripadnika istega monofiluma. Topologija takih kvartetov je znana in znaki, ki ji nasprotujejo so homoplastični. Analizo naredimo na vseh možnih kombinacijah taksonov. Znake razvrstimo po deležu kvartetov v katerih so nasprotovali apriorni informaciji. Znake nato izločamo v korakih, od najbolj do najmanj homoplastičnih. Po vsakem koraku naredimo filogenetsko analizo in določimo podporo hipotezam o sorodstvenih odnosih na ravni problematičnega razvejišča. Izberemo hipotezo z najvišjo podporo v celotnem postopku izločanja.

V primeru glavačev je podpora monofiliji zobatih kitov (*Odontoceti*) narasla iz prvotne, ki je bila 0 za bootstrap podporo in -39 korakov za Bremerjevo podporo, na najvišjo, ki je bila 100% za bootstrap podporo in 56,5 relativnih korakov za umerjeno Bremerjevo podporo. V primeru skakačev je dobila najvišjo podporo hipoteza, ki jih postavlja kot bazalen takson v

skupini Pancrustacea. Prvotna podpora za to hipotezo je bila 33,4% (bootstrap podpora) in 20 korakov (Bremerjeva podpora). Najvišja bootstrap podpora je bila 80%, najvišja umerjena Bremerjeva podpora pa 53 relativnih korakov. Najbolje podprti hipotetični sestrski takson hoacinu je bil trogon. Iz prvotne bootstrap podpore enake 46,7% in Bremerjeve podpore enake 0 je podpora narasla na 74,7% za bootstrap podporo in 22,4 relativnih korakov za umerjeno Bremerjevo podporo.

V vseh treh primerih je podpora naraščala le eni hipotezi v večjem delu postopka. Sklepamo, da metoda ne izloča znakov naključno, pač pa z izločanjem razkrije vzorec v podatkih. Ker znake vrednotimo na podlagi sorodstvenih odnosov, izločimo tudi zgodovinske signale, ki niso posledica sorodnosti taksonov. Zato metoda izbere hipotezo, ki odraža dejanske sorodstvene odnose.

7 VIRI

Arnason U., Gullberg A. 1996. Cytochrome b nucleotide sequences and the identification of five primary lineages of extant Cetaceans. *Molecular Biology and Evolution*, 13, 2: 407-417

Arnason U., Gullberg A., Janke A. 2004. Mitogenomic analyses provide new insights into cetacean origin and evolution. *Gene*, 333: 27-34

Barker F.K., Lanyon S.M. 2000. The Impact of Parsimony Weighting Schemes on Inferred Relationships among Toucans and Neotropical Barbets (Aves: Piciformes). *Molecular Phylogenetics and Evolution*, 15, 2: 215–234

Barome P.O., Monnerot M., Gautun J.C. 1998. Intrageneric Phylogeny of *Acomys* (Rodentia, Muridae) Using Mitochondrial Gene Cytochrome b. *Molecular Phylogenetics and Evolution*, 9, 3: 560-566

Bergsten J. 2005. A review of long-branch attraction. *Cladistics*, 21: 163-193

Bitsch C., Bitsch J. 2004. Phylogenetic relationships of basal hexapods among the mandibulate arthropods: a cladistic analysis based on comparative morphological characters. *Zoologica Scripta*, 33, 6: 511-550

Björklund M. 1999. Are Third Positions Really That Bad? A Test Using Vertebrate Cytochrome b. *Cladistics*, 15, 2: 191-197

Boisserie J.R., Lihoreau F., Brunet M. 2005. The position of hippopotamidae within cetartiodactyla. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 5: 1537-1541

Bremer, K. 1994. Branch support and tree stability. *Cladistics*, 10, 3: 295-304

Brower A.V.Z. 2000. Evolution is not a necessary assumption of cladistics. *Cladistics*, 16, 1: 143-154

Cameron S.L., Miller K.B., D'Haese C.A., Whiting M.F., Barker S.C. 2004. Mitochondrial genome data alone are not enough to unambiguously resolve the relationships of Entognatha, Insecta and Crustacea sensu lato (Arthropoda). *Cladistics*, 20, 6: 534-557

Carine M.A., Scotland R.W. 1999. Taxic and transformational homology: Different ways of seeing. *Cladistics*, 15, 2: 121-129

Cerchio S., Tucker P. 1998. Influence of alignment on the mtDNA Phylogeny of Cetacea: Questionable Support for a Mysticeti/Physeteroidae Clade. *Systematic Biology*, 47, 2: 336-344

Collins T.M., Wimberger P.H., Naylor G.J.P. 1994. Compositional bias, character-state bias, and character-state reconstruction using parsimony. *Systematic Biology*, 43: 482–496.

Cook C.E., Smith M.L., Telford M.J., Bastianello A., Akam M. 2001. Hox genes and the phylogeny of the arthropods. *Current Biology*, 11, 10: 759-763

Cook C.E., Yue Q.Y., Akam M. 2005. Mitochondrial genomes suggest that hexapods and crustaceans are mutually paraphyletic. *Proceedings of the Royal Society B-Biological Sciences*, 272, 1569: 1295-1304

Delsuc F., Brinkmann H., Chourrout D., Philippe H. 2006. Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature*, 439, 7079: 965-968

D'Haese C.A. 2002. Were the first springtails semi-aquatic? A phylogenetic approach by means of 28S rDNA and optimization alignment. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 269, 1496: 1143-1151

- Dixon M.T., Hillis D.M, 1993. Ribosomal RNA secondary structure: compensatory mutations and implications for phylogenetic analysis. *Molecular Biology and Evolution*, 10, 1: 256-267
- Dove H., Stollewerk A. 2003. Comparative analysis of neurogenesis in the myriapod *Glomeris marginata* (Diplopoda) suggests more similarities to chelicerates than to insects. *Development*, 130, 10: 2161-2171
- Dyke G.J., Van Tuinen M. 2004. The evolutionary radiation of modern birds (Neornithes): reconciling molecules, morphology and the fossil record. *Zoological Journal of the Linnean Society*, 141, 2: 153-177
- Edgecombe, G.D., Giribet, G. 2002. Myriapod phylogeny and the relationships of Chilopoda. V: *Biodiversidad, Taxonomía y Biogeografía de Artrópodos de México: Hacia una Síntesis de su Conocimiento*. Llorente Bousquets J. and Morrone J.J. (eds.). Prensas de Ciencias, Universidad Nacional Autónoma de México: 143-168
- Edgecombe G.D., Giribet G. 2004. Adding mitochondrial sequence data (16S rRNA and cytochrome c oxidase subunit I) to the phylogeny of centipedes (Myriapoda : Chilopoda): an analysis of morphology and four molecular loci. *Journal of Zoological Systematics and Evolutionary Research*, 42, 2: 89-134
- Edgecombe G.D., Wilson G.D.F., Colgan D.J., Gray M.R., Cassis G. 2000. Arthropod cladistics: Combined analysis of histone H3 and U2 snRNA sequences and morphology. *Cladistics*, 16, 2: 155-203
- Edwards S.V., Jennings W.B., Shedlock A.M. 2005. Phylogenetics of modern birds in the era of genomics. *Proceedings of the Royal Society B-Biological Sciences*, 272, 1567: 979-992

- Ericson P.G.P., Envall I., Irestedt M., Norman J.A. 2003. Inter-familial relationships of the shorebirds (Aves: Charadriiformes) based on nuclear DNA sequence data. *BMC Evolutionary Biology*, 3: 16
- Eyre-Walker A.C. 1991. An analysis of codon usage in mammals: selection or mutation bias? *Journal of Molecular Evolution*, 33: 442–453
- Fain M.G., Houde P. 2004. Parallel radiations in the primary clades of birds. *Evolution*, 58, 11: 2558-2573
- Farris J.S. 1969. A successive approximations approach to character weighting. *Systematic Zoology*, 18: 374–385
- Felsenstein J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology*, 27, 4: 401-410
- Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39, 4: 783-791
- Fryer G. 1996. Reflections on arthropod evolution. *Biological Journal of the Linnean Society*, 58, 1: 1-55
- Gatesy J., Milinkovitch M., Waddell V., Stanhope M. 1999. Stability of cladistic relationships between Cetacea and higher-level Artiodactyl taxa. *Systematic Biology*, 48, 1: 6-20
- Geisler J.H., Sanders A.E. 2003. Morphological evidence for the phylogeny of Cetacea. *Journal of Mammalian Evolution*, 10, 1-2: 23-129
- Geisler J.H., Uhen M.D. 2003. Morphological support for a close relationship between hippos and whales. *Journal of Vertebrate Paleontology*, 23, 4: 991-996

- Giribet G., Edgecombe G.D., Carpenter J.M., D'Haese C.A., Wheeler W.C. 2004. Is *Ellipura* monophyletic? A combined analysis of basal hexapod relationships with emphasis on the origin of insects. *Organisms, Diversity & Evolution*, 4, 4: 319-340
- Giribet G., Edgecombe G.D., Wheeler W.C. 2001. Arthropod phylogeny based on eight molecular loci and morphology. *Nature*, 413, 6852: 157-161
- Giribet G., Ribera C.A. 2000. A review of arthropod phylogeny: New data based on ribosomal DNA sequences and direct character optimization. *Cladistics*, 16, 2: 204-231
- Giribet G., Richter S., Edgecombe G.D., Wheeler W.C. 2005. The position of crustaceans within Arthropoda - Evidence from nine molecular loci and morphology. *Crustacean Issues*, 16: 307–352
- Giribet G., Wheeler W.C. 1999. The position of arthropods in the animal kingdom: Ecdysozoa, islands, trees, and the "parsimony ratchet". *Molecular Phylogenetics and Evolution*, 13, 3: 619-623
- Goertzen L.R., Theriot E.C. 2003. Effect of taxon sampling, character weighting, and combined data on the interpretation of relationships among the heterokont algae. *Journal of Phycology*, 39, 2: 423-439
- Graybeal A. 1998. Is It Better to Add Taxa or Characters to a Difficult Phylogenetic Problem? *Systematic Biology*, 47, 1: 9-17
- Hagstrom G.I., Hang D.H., Ofria C., Tornig E. 2004. Using Avida to Test the Effects of Natural Selection on Phylogenetic Reconstruction Methods. *Artificial Life*, 10: 157–166
- Harrison G.L., McLenachan P.A., Phillips M.J., Slack K.E., Cooper A., Penny D. 2004. Four new avian mitochondrial genomes help get to basic evolutionary questions in the late Cretaceous. *Molecular Biology and Evolution*, 21, 6: 974-983

Hasegawa M., Adachi J., Milinkovitch M.C. 1997. Novel phylogeny of whales supported by total molecular evidence. *Journal of Molecular Evolution*, 44: S117-S120

Hassanin A., Lecointre G., Tillier S. 1998. The 'evolutionary signal' of homoplasy in protein-coding gene sequences and its consequences for a priori weighting in phylogeny. *Comptes rendus de l'Académie des sciences*, 321, 7: 611-620

Hedges S.B., Simmons M.D., van Dijk M.A.M., Caspers G., de Jong W.W., Sibley C.G. 1995. Phylogenetic Relationships of the Hoatzin, an Enigmatic South American Bird. *Proceedings of the National Academy of Sciences of the United States of America*, 92: 11662-11665

Hendy M.D., Penny D. 1989. A framework for the quantitative study of evolutionary trees. *Systematic Zoology*, 38, 297–309

Hillis D.M., Pollock D.D., McGuire J.A., Zwickl D.J. 2003. Is Sparse Taxon Sampling a Problem for Phylogenetic Inference? *Systematic Biology*, 52, 1: 124-126

Hughes J.M., Baker A.J. 1999. Phylogenetic relationships of the enigmatic hoatzin (*Opisthocomus hoazin*) resolved using mitochondrial and nuclear gene sequences. *Molecular Biology and Evolution*, 16, 9: 1300-1307

Källersjö M., Albert V.A., Farris J.S. 1999. Homoplasy Increases Phylogenetic Structure. *Cladistics*, 15, 1: 91–93

Kim J. 1996. General Inconsistency Conditions for Maximum Parsimony: Effects of Branch Lengths and Increasing Numbers of Taxa. *Systematic Biology*, 45, 3: 363-374

Kjer K.M. 2004. Aligned 18S and insect phylogeny. *Systematic Biology*, 53, 3: 506-514

- Kluge A.G. 1997a. Sophisticated falsification and research cycles: Consequences for differential character weighting in phylogenetic systematics. *Zoologica Scripta*, 26, 4: 349-360
- Kluge A.G. 1997b. Testability and the refutation and corroboration of cladistic hypotheses. *Cladistics*, 13, 1-2: 81-96
- Lavrov D.V., Brown W.M., Boore J.L. 2004. Phylogenetic position of the Pentastomida and (pan)crustacean relationships. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 271, 1538: 537-544
- Lee M.S.Y. 1999. Circularity, evolution, systematics ... and circularity. *Journal of Evolutionary Biology*, 12, 4: 724-734
- Lee M.S.Y. 2002. Divergent evolution, hierarchy and cladistics. *Zoologica Scripta*, 31, 2: 217-219
- Lieberman D.E. 1999. Homology and hominid phylogeny: Problems and potential solutions. *Evolutionary Anthropology*, 7, 4: 142-151
- Livezey B.C., Zusi R.L. 2001. Higher-order phylogenetics of modern Aves based on comparative anatomy. *Netherlands Journal of Zoology*, 51, 2: 179-205
- Lockhart P.J., Steel M.A., Hendy M.D., Penny D. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. *Molecular Biology and Evolution*, 11: 605-612
- Luan Y.X., Mallatt J.M., Xie R.D., Yang Y.M., Yin W.Y. 2005. The phylogenetic positions of three basal-hexapod groups (Protura, Diplura, and Collembola) based on ribosomal RNA gene sequences. *Molecular Biology and Evolution*, 22, 7: 1579-1592

- Mallatt J., Giribet G. 2006. Further use of nearly complete, 28S and 18S rRNA genes to classify Ecdysozoa: 37 more arthropods and a kinorhynch. *Molecular Phylogenetics and Evolution*, 40, 3: 772-794
- Mallatt J.M., Garey J.R., Shultz J.W. 2004. Ecdysozoan phylogeny and Bayesian inference: first use of nearly complete 28S and 18S rRNA gene sequences to classify the arthropods and their kin. *Molecular Phylogenetics and Evolution*, 31, 1: 178-191
- May-Collado L., Agnarsson I. 2006. Cytochrome b and Bayesian inference of whale phylogeny. *Molecular Phylogenetics and Evolution*, 38, 2: 344-354
- Messenger S. L., McGuire J. A. 1998. Morphology, molecules, and the phylogenetics of cetaceans. *Systematic Biology*, 47, 1: 90-124
- Milinkovitch M.C. 1995. Molecular phylogeny of cetaceans prompts revision of morphological transformations. *Trends in Ecology & Evolution*, 10, 8: 328-334
- Milinkovitch M. C., Meyer A., Powell J. R. 1994. Phylogeny of all major groups of cetaceans based on DNA-sequences from 3 mitochondrial genes. *Molecular Biology and Evolution*, 11, 6: 939-948
- Milinkovitch M.C., Orti G., Meyer A. 1993. Revised phylogeny of whales suggested by mitochondrial ribosomal DNA-sequences. *Nature*, 361, 6410: 346-348
- Milinkovitch M.C., Thewissen J.G.M. 1997. Evolutionary biology - Even-toed fingerprints on whale ancestry. *Nature*, 388, 6643: 622-624
- Miyamoto M.M., Allard M.W., Adkins R.M., Janecek L.L., Honeycutt R.L. 1994. A Congruence Test of Reliability Using Linked Mitochondrial DNA Sequences. *Systematic Biology*, 43, 2: 236-249

Montgelard C., Catzeflis F.M., Douzery E. 1997. Phylogenetic relationships of artiodactyls and cetaceans as deduced from the comparison of cytochrome b and 12S rRNA mitochondrial sequences. *Molecular Biology and Evolution*, 14, 5: 550-559

Moreira D., Philippe H. 2000. Molecular phylogeny: pitfalls and progress. *International Microbiology*, 3: 9–16

Nardi F., Spinsanti G., Boore J.L., Carapelli A., Dallai R., Frati F. 2003. Hexapod origins: Monophyletic or paraphyletic? *Science*, 299, 5614: 1887-1889

Negrisola E., Minelli A., Valle G. 2004. The mitochondrial genome of the house centipede *Scutigera* and the monophyly versus paraphyly of myriapods. *Molecular Biology and Evolution*, 21, 4: 770-780

Nikaido M., Matsuno F., Hamilton H., Brownell R. L., Cao Y., Ding W., Zuoyan Z., Shedlock A. M., Fordyce R. E., Hasegawa M., Okada N. 2001. Retroposon analysis of major cetacean lineages: The monophyly of toothed whales and the paraphyly of river dolphins. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 13: 7384-7389

Nikaido M., Rooney A.P., Okada N. 1999. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspersed elements: Hippopotamuses are the closest extant relatives of whales. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 18: 10261-10266

Paton T.A., Baker A.J., Groth J.G., Barrowclough G.F. 2003. RAG-1 sequences resolve phylogenetic relationships within Charadriiform birds. *Molecular Phylogenetics and Evolution*, 29, 2: 268-278

Patton J.L., dos Reis S.F., da Silva M.N.F. 1996. Relationships among didelphid marsupials based on sequence variation in the mitochondrial cytochrome B gene. *Journal of Mammalian Evolution*, 3, 1: 3-29

- Philippe H., Lecointre G., Hoc Lanh Van Le, Le Guyader H. 1996. A critical study of homoplasy in molecular data with the use of a morphologically based cladogram, and its consequences for character weighting. *Molecular biology and evolution*, 13, 9: 1174-1186
- Poe S., Chubb A.L. 2004. Birds in a bush: five genes indicate explosive evolution of Avian orders. *Evolution*, 58, 2: 404-415
- Regier J.C., Shultz J.W., Kambic R.E. 2004. Phylogeny of basal hexapod lineages and estimates of divergence times. *Annals of the Entomological Society of America*, 97, 3: 411-419
- Regier J.C., Shultz J.W., Kambic R.E. 2005a. Pancrustacean phylogeny: hexapods are terrestrial crustaceans and maxillopods are not monophyletic. *Proceedings of the Royal Society B-Biological Sciences*, 272, 1561: 395-401
- Regier J.C., Wilson H.M., Shultz J.W. 2005b. Phylogenetic analysis of Myriapoda using three nuclear protein-coding genes. *Molecular Phylogenetics and Evolution*, 34, 1: 147-158
- Roelants K., Bossuyt F. 2005. Archaeobatrachian paraphyly and pangaean diversification of crown-group frogs. *Systematic Biology*, 54, 1: 111-126
- Sanderson M.J., Wojciechowski M.F., Hu J.-M., Sher Khan T., Brady S.G. 2000. Error, bias, and long-branch attraction in data for two chloroplast photosystem genes in seed plants. *Molecular Biology and Evolution*, 17, 5: 782-797
- Schachtner J., Schmidt M., Homberg U. 2005. Organization and evolutionary trends of primary olfactory brain centers in Tetraconata (Crustacea plus Hexapoda). *Arthropod Structure & Development*, 34, 3: 257-299

- Schaeffer B. 1987. Deuterostome monophyly and phylogeny. *Evolutionary Biology*, 21: 179–235
- Scholtz G., Edgecombe G.D. 2006. The evolution of arthropod heads: reconciling morphological, developmental and palaeontological evidence. *Development Genes and Evolution*, 216, 7-8: 395-415
- Slack K.E., Janke A., Penny D., Arnason U. 2003. Two new avian mitochondrial genomes (penguin and goose) and a summary of bird and reptile mitogenomic features. *Gene*, 302, 1-2: 43-52
- Slikas B. 1997. Phylogeny of the Avian Family Ciconiidae (Storks) Based on Cytochrome b Sequences and DNA–DNA Hybridization Distances. *Molecular Phylogenetics and Evolution*, 8, 3: 275-300
- Sorenson M.D., Oneal E., García-Moreno J., Mindell D.P. 2003. More Taxa, More Characters: The Hoatzin Problem Is Still Unresolved. *Molecular Biology and Evolution*, 20, 9: 1484-1498
- Swofford, D.L. 1998. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts
- Thewissen J.G.M., Williams E.M., Roe L.J., Hussain S.T. 2001. Skeletons of terrestrial cetaceans and the relationship of whales to artiodactyls. *Nature*, 413, 6853: 277-281
- Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F., Higgins D.G. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, 24: 4876-4882
- Ursing B.M., Arnason U. 1998. Analyses of mitochondrial genomes strongly support a hippopotamus-whale clade. *Proceedings Of The Royal Society Of London Series B-Biological Sciences*, 265, 1412: 2251-2255

- Utevsky S.Y., Trontelj P. 2004. Phylogenetic relationships of fish leeches (Hirudinea, Piscicolidae) based on mitochondrial DNA sequences and morphological data. *Zoologica Scripta*, 33, 4: 375–385
- Van Tuinen M., Waterhouse D., Dyke G.J. 2004. Avian molecular systematics on the rebound: a fresh look at modern shorebird phylogenetic relationships. *Journal of Avian Biology*, 35: 191-194
- Van Valen L. 1966. Deltatheridia, a new order of mammals. *Bulletin of the American Museum of Natural History*, 132, 1: 1-126
- Wenzel J.W., Siddall M.E. 1999. Noise. *Cladistics*, 15, 1: 51-64
- Weygoldt P. 1998. Evolution and systematics of the Chelicerata. *Experimental and Applied Acarology*, 22, 2: 63-79
- Wheeler W.C. 1990. Combinatorial weights in phylogenetic analysis: A statistical parsimony procedure. *Cladistics*, 6: 269–275
- Wheeler W.C., Hayashi C.Y. 1998. The phylogeny of the extant chelicerate orders. *Cladistics*, 14, 2: 173-192
- Wheeler W.C., Whiting M., Wheeler Q.D., Carpenter J.M. 2001. The phylogeny of the extant hexapod orders. *Cladistics*, 17, 2: 113-169
- Williams P.L., Fitch W.M. 1990. Phylogeny determination using the dynamically weighted parsimony method. *Methods in Enzymology*, 183: 615–626

ZAHVALA

Zahvaljujem se mentorju prof. dr. Petru Trontlju za strokovno svetovanje, potrpežljivost in spodbudo pri nastajanju diplomskega dela. Hvala recenzentu doc. dr. Rudiju Verovniku, da mi je nesebično odstopil pisarno in računalnik, ter za hitro posredovanje koristnih popravkov diplomskega dela. Iskrena hvala dragi družini za vso podporo in finančno pomoč pri študiju. Drejkotu za 24 urno družbo in kavice, ki sva jih spila. Nenazadnje, hvala Janji za vse.

Priloga A

Primer vhodne datoteke za program za tvorbo kvartetov (primer skakači).

Insecta:

Arch
Tric
Lmig
Tdim
Tcas
Cduo
Pruf
Aper
Bmor
Bman
Ofur
Onub
Aqua
Agam
Dyak
Dmel
Cput
Chom
Ccap

Chelicerata:

Lpol
Vdes
Omou
Ihex
Iper
Rsan

Decapoda:

Pjap
Ptri
Plon
Pmon

Branchiopoda:

Dpul
Afra
Tcan

Diplopoda:

Thyr
Nann

Chilopoda:

Lfor

Priloga B

Primer vhodne datoteke za analizo znakov na podlagi kvartetov (v programu PAUP*).
Primer skakači.

```
#nexus
Begin paup;
Log file=paupVrne.log;
Set autoclose=yes;
set warntree=no;
set warnredef=no;
End;

[!(Arch Tric Lpol Vdes) st kvarteta: 1 ]
begin paup;
delete all;
undelete Arch Tric Lpol Vdes;
constraints zaostri (monophyly)=[&U] (( Arch, Tric),( Lpol, Vdes));
bandb enforce=yes;
describetrees/diag=yes labelnode=no;
cleartrees;
end;

[!(Arch Tric Lpol Omou) st kvarteta: 2 ]
begin paup;
delete all;
undelete Arch Tric Lpol Omou;
constraints zaostri (monophyly)=[&U] (( Arch, Tric),( Lpol, Omou));
bandb enforce=yes;
describetrees/diag=yes labelnode=no;
cleartrees;
end;

...

[!(Thyr Nann Tcan Lfor) st kvarteta: 31416 ]
begin paup;
delete all;
undelete Thyr Nann Tcan Lfor;
constraints zaostri (monophyly)=[&U] (( Thyr, Nann),( Tcan, Lfor));
bandb enforce=yes;
describetrees/diag=yes labelnode=no;
cleartrees;
end;

begin paup;
log stop;
end;
```

Priloga C

Vhodna datoteka (za program PAUP*) za zaporedno izločanje znakov. Primer skakači.

```
#nexus

Begin paup;
Log file=ColembollaIzloc-BS.log;
End;

Begin paup;
Set dstatus=no status=yes;
Set autoclose=yes warntree=no warnredef=no;
Set MaxTrees=100 increase=auto autoinc=100;
Set criterion=parsimony;
End;

Begin paup;
Delete Tjap;
Outgroup Vdes Omou Ihex Iper Rsan;
[Lpol nisem vkljucil v OG, ker v prvih izlocanjih potegne za sabo
Myriapoda]
End;

[!0 brez izlocanja]
Begin paup;
Hsearch nreps=100 addseq=random swap=tbr;
Describetrees/diag=no labelnode=no plot=phylogram;
savetrees file=zacetnoDrevo.tre brlens=no replace=yes;
Bootstrap nreps=1000/nreps=10 addseq=random swap=nni;
End;

[!1 freq=0,19 , 362 znakov izlocenih]
Begin paup;
Exclude 4776 9342 8550 752 4173 5262 3382 2259 2109 4456 8118 3059 3595
4614 1305 4852 2196 3810 2835 4767 5053 2247 415 1638 8815 9002 4303 8409
3645 2884 3894 2238 3745 1956 1507 4674 5352 107 8327 2208 2547 3295 3924
6598 1231 2784 5622 11163 1249 3843 5814 3897 3727 2880 8352 2326 4071
8768 5090 3972 8112 1812 1252 7130 4534 2973 6797 3016 4331 4930 7491
2448 2017 350 9355 4759 2052 8724 3224 5808 1788 3813 1158 3454 6846 5607
5838 4863 4176 1119 3472 6039 3648 1513 11313 11434 4580 6443 7672 6370
3507 10260 1998 6879 46 1200 472 9561 3775 1560 7050 3771 8028 3792 4233
11451 2090 10842 4332 9801 10302 3850 2974 815 511 514 3844 1543 6420
8044 4131 2025 2941 2421 5672 2274 7938 297 5014 1611 10651 4117 7797
2311 7669 8706 646 1086 7484 5448 2046 809 4965 4692 6014 5268 11716 8802
4845 2209 2244 4120 8610 1071 5628 1729 2193 1765 4794 10554 5556 1491
7650 2958 3495 3510 5060 3954 3138 1899 3840 2098 4677 1569 6991 4791 951
8121 3321 6897 1108 6033 942 4395 11067 1755 1581 990 10674 9372 7440 63
484 8064 9240 5331 6582 2309 1107 4726 9888 10026 10557 3876 4632 12 1039
2059 8551 604 10886 8127 3576 1188 3453 1599 8933 1366 5085 8877 2887
1321 4408 2442 1522 8898 5295 7509 7587 11782 2022 2538 8364 1818 5877
2630 2184 2215 8515 4833 9360 2908 948 4800 289 6717 6261 1957 2262 11459
9814 1479 5648 7051 4962 2500 4056 5328 4680 5451 10989 11737 1722 3669
4873 1312 7161 9687 9625 4095 4671 3228 3273 321 4986 10383 4296 4066
3650 1990 2710 1134 4215 2817 941 8070 3517 10843 6397 10791 5236 10905
```

```
1284 3274 5307 5071 5833 969 8743 10237 1477 9312 1650 2165 11472 8788
4545 3350 3402 1551 2901 2556 7449 9186 1830 4404 8604 7194 7500 3955
9222 489 4730 4476 2386 2004 2429 6048 1419 6074 6423 9534 1821 5100 4891
1703 2684 8066 8805 8539 10710 7649 2632 7413 870 4365 5823 2502 7540
3036 2144 3429;
Hsearch nreps=100 addseq=random swap=tbr;
Describetrees/diag=no labelnode=no plot=phylogram;
Bootstrap nreps=1000/nreps=10 addseq=random swap=nni;
End;
```

...

```
[!17 freq=0,029 , 325 znakov izlocenih]
Begin paup;
Exclude 7556 7973 10483 9298 2592 6980 11257 4099 11129 6608 10996 6938
929 5686 5908 9197 9359 10424 11038 2453 5590 6809 10567 11010 11254
11442 3041 866 5415 8700 10318 7588 8475 11624 1262 3143 3686 8751 11531
142 10477 9185 4735 5521 6625 1074 2169 10638 2823 8351 10387 2079 8736
10666 11690 2985 4416 7288 7767 7945 8755 9158 10262 5067 7188 8011 10784
7855 9328 3900 27 1461 3751 2458 4425 7562 8265 1944 7738 5966 2486 3534
6281 5369 6380 10990 5123 6087 7961 11661 7879 9378 7238 9531 10706 112
1697 3122 3397 3962 5609 6953 1032 3443 10692 8965 9592 3315 5207 11612
817 3889 5944 10309 11762 2459 7124 542 9863 2609 2652 3212 8682 10056
4645 6702 7781 8963 10810 11144 1548 7761 6776 11651 7947 8911 5269 9332
10821 11071 6817 3684 5909 7861 8389 8781 3767 5045 5712 6612 10605 486
1012 5955 8608 11385 4704 10936 9621 8071 5638 11652 7662 10095 1692 4575
5536 7096 2086 4783 10320 4898 7207 7974 8731 6067 6296 10107 8334 11080
6508 4649 8310 220 6118 876 2005 5230 8613 9373 11378 10950 6465 2260
4124 9213 11464 7865 10736 11711 4315 7097 1911 7398 7518 9198 10254 9721
5633 10703 7787 9829 1326 11667 1699 2119 3341 3355 5274 11511 8867 6905
10054 6555 10944 6407 7726 6042 7261 9851 11525 10167 8646 7208 11262
9650 9065 9793 4123 7828 9725 10897 6113 11619 3256 7926 591 536 7270 760
10110 9313 9927 10253 10595 10828 3139 11084 11132 2308 5186 7665 9683
9958 3209 9786 4586 8056 9202 2211 2852 7735 8497 925 5565 6015 4135 5491
7967 3400 1355 1458 903 5139 1701 2355 3289 5915 7946 8146 9530 4069
11089 8158 8885 11327 10001 2318 11259 11756 3967 8328 8576 9929 8208
8258 9643 8459 10690 11567 2407 9244 10735 5752 3657 9001 11665 5005 9142
381 1656 4619 11037 3298 724 9605 90 5122 1089 1354 9263 11308 803 2147
6320 7287 5898 9290 4387 9026 9678;
Hsearch nreps=100 addseq=random swap=tbr;
Describetrees/diag=no labelnode=no plot=phylogram;
Bootstrap nreps=1000/nreps=10 addseq=random swap=nni;
End;

Begin paup;
Log stop;
End;
```