

Jihočeská univerzita v Českých Budějovicích
Přírodovědecká fakulta



Bakalářská práce

**Aplikace barcodingu na rod *Folsomia* (Collembola)
a mitochondriální genom *F. candida***

Martina Slámová

Školitel : PaedDr. Martina Žurovcová, PhD.

Školitel specialista : Prof. RNDr. Josef Rusek DrSc.

České Budějovice

2008

Bakalářská práce

Slámová, M., 2007. Aplikace barcodingu na rod *Folsomia* (Collembola) a mitochondriální genom *F. candida*. [Application of DNA barcoding on genus *Folsomia* (Collembola) and mitochondrial genome of *F. candida*. Bc. Thesis, in Czech.] Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic, 33 pp.

Annotation

Mitochondrial molecular marker COI was tested for use in species identification of selected species of genus *Folsomia* (Collembola). Marker was successfully amplified and sequenced. Dendrogram constructed by Neighbor-Joining method with Kimura-2-Parameter model grouped all individuals into presumed species clusters and high intraspecific variability of *F. quadrioculata* suggests the existence of cryptic species.

Furthermore, 65 % of mitochondrial genome of *F. candida* was obtained with 16 tRNA genes, 9 protein-coding genes and 2 rRNA genes identified. So far the genome characteristics correspond to the one described in *G. hodgsoni*.

Tato práce byla financována ze záměru Entomologického ústavu Z-50070508.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své bakalářské práce, a to v nezkrácené podobě Přírodovědeckou fakultou, elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejich internetových stránkách.

Prohlašuji, že jsem tuto, bakalářskou práci vypracovala samostatně, pouze s použitím uvedené literatury.

V Českých Budějovicích, 11.1.2007

.....

PODĚKOVÁNÍ

Chtěla bych poděkovat mé školitelce Martině Žurovcové za odborné vedení v průběhu práce, Michalu Žurovcovi a kolektivu jeho laboratoře za vše, co mě naučili v mých laboratorních začátcích a profesoru Josefu Ruskovi za poskytnutí analyzovaného materiálu. Dále děkuji Luce Kučerové a Vašku Brožovi, kteří mi vždy ochotně poradili a pomohli, a kolektivu naší laboratoře za vytvoření přátelské atmosféry jak v laboratoři, tak mimo ní. Zvláštní poděkování patří mojí rodině a přátelům, kteří mi byly velkou oporou po celou dobu studia.

OBSAH:

1. ÚVOD	1
2. CÍLE PRÁCE	5
3. MATERIÁL A METODY	6
3.1 Extrakce DNA	6
3.1.1 Chelexová izolace	7
3.1.2 Izolace pomocí ZR Genomic DNA II Kit TM	7
3.1.3 Izolace pomocí DNeasy Tissue Kit (QIAGEN)	7
3.1.4 Izolace extrakčním pufrem	7
3.2 Primery	7
3.3 PCR	10
3.3 Gelová elektroforéza	10
3.4 Purifikace PCR	11
3.4.1 DNA Clean & Concentrator TM -5 (ZYMO RESEARCH)	11
3.4.2 Zymoclean TM Gel DNA Recovery Kit (ZYMO RESEARCH)	11
3.4.3 Čištění pomocí směsi Exosap	11
3.5. Sekvenování DNA	12
3.5.1. Sekvenační směs	12
3.5.2. Sekvenační program	12
3.5.3. Přečištění sekvenační reakce pomocí Sephadexu a aerosolových špiček	12
3.5.4 Sekvenace	12
3.6 Analýza sekvencí	13
4. VÝSLEDKY	14
4.1 Izolace DNA	14
4.2 Barcoding rodu <i>Folsomia</i>	14
4.3 Mitochondriální genom <i>Folsomia candida</i>	16
4.3.1 tRNA	18
4.3.2 Geny kódující proteiny	21
4.3.3 Ribozomální RNA	22
5. DISKUZE	23
5.1 Izolace DNA	23
5.2 Použití molekulárního markeru COI (DNA barcoding) pro identifikaci druhů rodu <i>Folsomia</i>	23
5.3 Mitochondriální genom <i>Folsomia candida</i>	25
5.3.1 Popis genomu	25
5.3.2 tRNA	26
5.3.3 Geny kódující proteiny	26
5.3.4 Ribozomální RNA	27
5.3.4 A+T oblast	27
6. ZÁVĚR	28
7. LITERATURA	29
8. PŘÍLOHA	33

1. ÚVOD

Chvostokoci (Collembola)

Chvostokoci jsou jedni z nejpočetnějších členovců na zemi s dlouhou evoluční historií; (Engel & Grimaldi, 2004) nejstarší fosilní chvostokoci nalezení v jantaru byly více než 400 milionů let staří. Zcela určitě jsou chvostokoci monofyletickou větví, která se brzy oddělila od linie, jež vedla k vyššímu hmyzu. Jestli k tomu došlo před oddělením koryšů nebo po něm však ještě není jednoznačně vyřešeno, ačkoliv nejnovější studie založené na srovnávání celých mitochondriálních genomů tomu nasvědčují (Delsuc *et al.*, 2003; Nardi *et al.*, 2003a; Nardi *et al.*, 2003b; Carapelli *et al.*, 2007).

Chvostokoků je známo si 7000 druhů a většina z nich se živí houbami v půdě nebo opadaným listím. Rozšířili se do mnoha prostředí, od litorální zóny do hor a jsou zvláště početní v epifytech tropických deštných lesů (Hopkin, 1997). Taxonomicky je dělíme do tří hlavních řádů:

Artheopleona (asi 5500 druhů) mají protáhlé tělo, a zástupci žijí jak na povrchu půdy tak i celý život pod zemí,

Symphyleona (asi 1000 druhů) s kulatějším tělem, jsou výrazněji zbarveni a žijí na povrchu půdy,

Neelipleona podobající se předchozím, jsou velice malí (kolem 0,5 mm), žijí v půdě, a patří do nich pouze 25 známých druhů.

Nejvýraznějším znakem většiny chvostokoků (jež vlastně vedl k jejich taxonomickému názvu) je skákací orgán zvaný furka. Vyvinula se splynutím přívěsků čtvrtého zadečkového článku jako únikový mechanismus k obraně před predátory. U druhů žijících pod zemí je redukována, aby nebránila jejich pohybu mezi půdními částicemi, a některé druhy ji ztratili úplně. Maximální počet oček (ocelli) v každém oku je osm, ale jsou často redukovány a mnoho půdních a jeskynních druhů je slepých. Všichni chvostokoci mají trubicovitý orgán „ventrální tubus“, vzniklý splynutím přívěsků prvního zadečkového článku. Tento orgán je velice důležitý pro udržení vodní rovnováhy ale může sloužit i k adhezi chvostokoků na kluzkém povrchu. (Fountain & Hopkin, 2005).

Vzhledem k jejich rozměrům a malým, jen obtížně pozorovatelným morfologickým rozdílům není snadné tyto členovce taxonomicky určovat. Je nutné je fixovat a pozorovat pod mikroskopem. Časté jsou i případy kryptických druhů u celé skupiny Collembola (Stevens & Hogg, 2003)

Protože jsou nedílnou součástí všech půdních ekosystémů a přitom velice citliví na jejich znečištění, je jejich množství a diverzita (tj. druhové spektrum) často používána ke zjištění vlivu širokého spektra polutantů na životní prostředí (Fountain & Hopkin, 2005). Jsou tedy významnými bioindikátory, i když jsou pro svou nepatrnou velikost pouhým okem téměř nepozorovatelní. Nepigmentovaný chvostoskok *Folsomia candida* (Willem, 1902) byl dokonce vybrán jako modelový druh pro testování vlivů různých látek na kvalitu půdy, zřejmě pro snadnost, s jakou se dají jeho kultury uchovávat v laboratorních chovech. Tento druh byl také použit ke sledování kontaminované půdy při navrácení do normálního stavu (Crouau *et al.*; 2002, Diez *et al.*, 2001; Fava *et al.*, 1999) nebo v legislativním procesu schvalování nových chemických látek (Van Straalen, 1997). Mezinárodní standardizační organizace (ISO) nedávno vydala protokol pro použití druhu *F. candida* jako ekotoxikologického testovacího druhu (ISO, 1999). Tento druh je také možno použít ve školní výuce, neboť se na jeho kulturách dají snadno demonstrovat např. vliv teploty nebo herbicidů na jeho životní cyklus (Moore *et al.* 2000).

Vzhledem k tomu, jak důležitým modelovým ekotoxikologickým organismem *F. candida* po desetiletí byla, je překvapivé, že se její genetikou téměř nikdo nezabýval. Prvním velkým projektem se proto stalo sekvenování jaderného genomu neboli “The *Folsomia candida* EST (Expressed Sequence Tag) project” započatý v r. 2004 M.Timmermansem (Vrije Universiteit, Amsterdam). Práce jeho týmu vyústila do zřízení webové databáze “Collembase” (<http://www.collembase.org>) v r. 2007, jejímž cílem je doplnit potřebné informace o druhu *F. candida*, aby se tento druh mohl stát i standardním genomovým modelem.

Větší zájem však byl už dříve o mitochondriální genom jiných druhů chvostoskoků. Zvláště kompletní mtDNA sekvence může být totiž zdrojem informací na hluboké fylogenetické úrovni (Curole & Kocher, 1999). Její využití bylo prokázáno v různých skupin živočichů včetně ryb (Myia *et al.*, 2003), ptáků (Mindel *et al.*, 1999), plazů (Kumazawa & Nishida, 1999), savců (Reyes *et al.*, 1998) a bezobratlých (Boore & Staton, 2002). Mitochondriální DNA je kruhová molekula 15-20 kb dlouhá, která obsahuje 37 genů, z toho 13 pro proteiny, 2 pro rRNA a 22 pro transferovou RNA (tRNA), a nekódující kontrolní oblast, která reguluje transkripci mitochondriálního genomu. (Boore., 1999). MtDNA je často používána pro studium populační struktury, fyto geografie a fylogenetických vztahů na různých taxonomických úrovních (Simon *et al.*, 1994, Boore *et al.*, 1995, Boore *et al.*, 1998) protože se dědí v mateřské linii, má relativně vysokou evoluční rychlost a nedochází v ní ke genetické rekombinaci.

Porovnání genových přestaveb mtDNA živočichů se stává důležitým nástrojem v současné fylogenetice. Právě tyto přestavby byly použity pro objasnění fylogenetických vztahů mezi hlavními liniemi v rámci členovců (Cook *et al.*, 2005), k čemuž přispělo osekvenování kompletní nebo téměř kompletní mtDNA z chvostoskoků *Onychiurus orientalis*, *Gomphiocephalus hodgsoni*, *Podura aquatica* a *Tetrodontophora bielanensis*.

Pro účely populační genetiky, fylogeografie nebo molekulární taxonomie se nepoužívají jen celé mitochondriální genomy, ale hlavně jednotlivé geny nebo jejich části (např. 12S RNA, 16S RNA, COI, COII). Tyto geny se liší evoluční rychlostí a tudíž mají využití pro zkoumání vztahů na různých taxonomických úrovních. Nejčastěji používaným markerem u chvostoskoků je cytochrom oxidáza I (COI), která byla vybrána pro projekt DNA barcoding započatý Hebertem *et al.* (2003a, b). Je to technika využívající krátký úsek genu pro cytochrom C oxidázu I (asi 650 bp poblíž 5' konce) pro přesné určení a pozdější rychlou identifikaci rostlinných a živočišných druhů (Hebert *et al.*, 2003a). Slouží nejen k rozlišení už známých druhů (Hebert *et al.*, 2003b, Ward *et al.*, 2005), ale také k identifikaci nových nebo kryptických druhů (Hebert *et al.*, 2004, Hajibabei *et al.*, 2006; Smith *et al.*, 2006), invazivních druhů (Siddal & Budinoff, 2005) a ke stanovení biodiverzity (Janzen *et al.*, 2005). Pro COI platí, že vnitrodruhová variabilita je pod 1%, maximálně 2%, mezidruhová variabilita je pak větší než 10%, což zobecněno znamená, že variabilita sekvence mezi druhy je 10x větší než variabilita vnitrodruhová (Hebert *et al.*, 2004). Barcoding je užitečný právě u chvostoskoků, kteří se mohou lišit ekologicky, ale jejich morfologické znaky, které byly dosud pro taxonomy hlavním rozlišovacím znakem, jsou stejné. Bylo tak úspěšně určeno 13 rodů a 19 druhů arktických chvostoskoků (Hogg & Hebert, 2004).

Pro naši práci byly vybrán rod *Folsomia* z čeledi *Isotomidae*, jehož druhy byly na molekulární úrovni studovány jen okrajově při zjišťování druhového spektra severoamerických lokalit (Hogg & Hebert, 2004) a v rámci projektu BOLD (Barcoding Of Life). K hlavním morfologickým rozlišovacím znakům tohoto rodu patří dobře vyvinutá furka, absence snovacích bradavek a na zadečku mají spojené poslední tři články. Jedinci jsou v dospělosti 1,5 až 3 mm dlouzí, bílí nebo světle žlutí a nemají oči, hlavním rozlišovacím znakem je přítomnost štětinek (stout setae) na ventrální straně manubria furky. Vzhledem k tomu, v jak odlišných biotopech se tyto druhy nacházejí, je zde právě možnost existence kryptických druhů (J. Rusek, osobní sdělení).

Ve své práci jsem se proto zabývala tím, zda je možné pomocí barcodingu tyto morfologicky neodlišitelné druhy rozpoznat. Pro významnost druhu *Folsomia candida* jsem

se pak také zabývala popsáním struktury jeho mitochondriálního genomu a porovnat ho s ostatními publikovanými genomy jiných chvostoskoků.

2. CÍLE PRÁCE

- 1) Naučit se použití metod molekulární biologie pro molekulární taxonomii.
- 2) Druhová identifikace rodu *Folsomia* pomocí „DNA Barcoding“.
- 3) Stanovení struktury mitochondriálního genomu *Folsomia candida*.

3. MATERIÁL A METODY

3.1 Materiál

V práci bylo použito celkem 8 druhů rodu *Folsomia*, které byly nasbírány a určeny prof. J. Ruskem. *F. candida* byla brána z laboratorního chovu založeného z jedinců, které nám poskytnuli V. Šustr (Biologické centrum AV ČR, ÚPB), J. Hofman (Výzkumné centrum pro chemii životního prostředí a ekotoxikologii - RECETOX, Přírodovědecká fakulta Masarykovy univerzity) a M. Timmermans (Department of Animal Ecology, Vrije Universiteit, Amsterdam). Podrobnosti k jednotlivým druhům jsou uvedeny v Tabulce 1.

Tabulka 1. Seznam druhů použitých pro DNA barcoding

Druh	Lokalita	
<i>F. manolachei</i>	zahrada ÚPB ČB	
	Bílé Karpaty	Vrch výzkum, louka
<i>F. quadriocullata</i>	Stromovka ČB	
	Jeseníky, NPR Praděd	Sjezdovka pod Petrovými kameny vzorek půdy s kapradinami
		Velká hole, alpská (vrcholová) louka s porostem lišejníků
	NP Šumava	D 879
	Budislav u Litomyšle	borový les
<i>F. candida</i>	z laboratorních chovů	České Budějovice
		Amsterdam
		Brno
	Tišnovský kras	Králová jeskyně, Tišnovský dóm
<i>F. tesari</i>	Tatranský NP, Slovensko	Červené vrchy, dolina Rozpadliny, rostlinné spoločenstvo <i>Adinostletum alliariae</i>
<i>F. penicula</i>	Petrovice u Karviné	porost bříz u haldy
<i>F. inoculata</i>	E-595-1	
<i>F. ksenemani</i>	Blanský les	severní svah Kletě, lučina
<i>F. sensibilis</i>	Tatranský NP, Slovensko	

Při všech laboratorních analýzách byly používány špičky s filtrem a sterilní materiál pro zamezení kontaminace mezi vzorky nebo cizorodou DNA.

3.1 Extrakce DNA

Chvostoskoci byli před extrakcí DNA fixováni v 96% etanolu, uchováni živí v destilované vodě nebo odebráni přímo z chovu. Vzorky uchované v etanolu jsme před extrakcí nechali oschnout na filtračním papíře při pokojové teplotě, dokud se všechen etanol neodpařil.

Pro získání DNA jsem testovala čtyři různé metody extrakce. Vždy byli použiti celí jedinci, kromě vzorků *F.candida* použitých pro sekvenování mitochondriálního genomu. V tomto případě jsem použila směs jedinců z téhož chovu. .

3.1.1 Chelexová izolace (Hoelzel, 1998)

Celý jedinec byl přenesen do 300 μ l 5% Chelexu (BioRad), rozdrčen pomocí sterilního homogenizátoru a inkubován 60 minut při teplotě 56° C. Po této době byl vzorek promíchán a inkubován 10 minut při 98°C. Izolovaná DNA byla uskladněna při teplotě -20° .

3.1.2 Izolace pomocí ZR Genomic DNA II Kit™

Vzhledem k malé velikosti jedinců jsem použila protokol výrobce s několika doporučenými úpravami. Celý jedinec byl přenesen do mikrozkušavky se 125 μ l extrakčního pufru „Genomic Lysis Buffer“ (GL). Vzorek byl pak zcentrifugován 6 minut při 14500 g, což usnadnilo následnou homogenizaci vzorku pomocí sterilního homogenizátoru. Pak bylo přidáno dalších 125 μ l GL pufru a znovu zcentrifugováno. Supernatant byl přenesen na sloupeček a centrifugován 90 sekund při 14.5g. Pak bylo přidáno 250 μ l promývacího pufru „Wash Buffer“ a znovu zcentrifugováno 90 sekund při 14500 g. Tento krok byl ještě jednou opakován. Po přidání 25 μ l ddH₂O, která se nechala 3 minuty na sloupku, a po následné centrifugaci došlo k uvolnění DNA z membrány do vody.

3.1.3 Izolace pomocí DNeasy Tissue Kit (QIAGEN)

Izolace byla provedena podle návodu výrobce.

3.1.4 Izolace extrakčním puftrem (Cox & Hebert, 2001, Frati *et al.*, 2001)

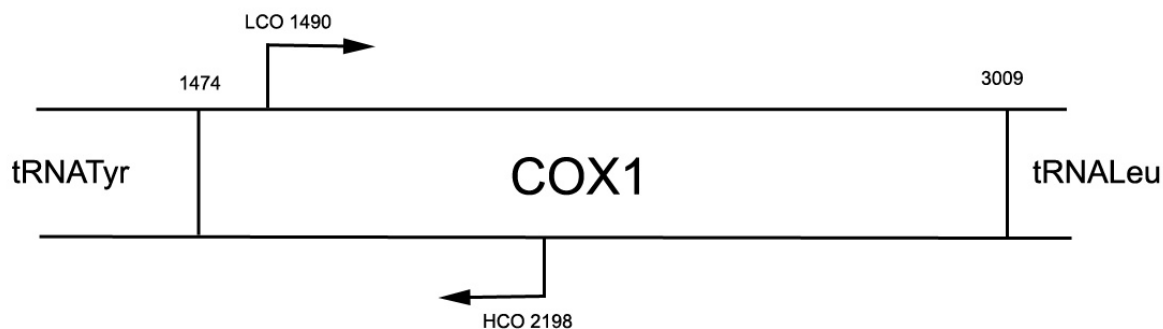
Pro izolaci jsem použila extrakční pufr (EP) o následujícím složení: ddH₂O (885 μ l) 1x PCR pufr s MgCl₂ (100 μ l), 1% Tweenu 20 (10 μ l) a proteináza K o koncentraci 100 μ g/ml (5 μ l). Ke vzorku bylo přidáno 25 μ l EP a na 15 minut byl zmražen na -70°C, pak byl inkubován při 55°C 8 hodin nebo při 65°C 2 hodiny. Nakonec byl 15 minut inkubován při 98°C. Izolovaná DNA byla uskladněna při teplotě -20°C.

3.2 Primery

Úsek genu pro cytochrom oxidázu byl amplifikován pomocí univerzálních primerů (Folmer *et al.*, 1994), které jsou používány standardně pro tzv. DNA Barcoding (viz Tabulka 2). Jejich polohu vzhledem k referenční sekvenci *D. yakuba* znázorňuje Obr. 1.

Tabulka 2. Primery pro amplifikaci COI.

Název primeru	Sekvence primeru 5'→ 3'	Směr
LCO 1490 ³	GGTCAACAAATCATAAAGATATTGG	F
HCO 2198 ³	TAAACTTCAGGGTGACCAAAAAATCA	R

**Obrázek 1.** Schéma umístění primerů pro DNA Barcoding.

Pro amplifikaci jednotlivých úseků mitochondriálního genomu byly použity primery převzaté z práce Simon *et al.* (1994 a 2006) nebo byly navrženy na základě sekvencí chvostoskoků dostupných v databázi OGRE (<http://drake.physics.mcmaster.ca/ogre/>) (Tabulka 3 a 4). Po získání prvních sekvencí *F. candida* jsme pak také navrhovali primery specifické. Pozice jednotlivých primerů je dána číslem, vztaženým na referenční sekvenci *D. yakuba*. Kombinace různých primerů byly zkoušeny pro získání překrývajících se PCR produktů o délce cca 2kbp. Sekvenování pak probíhalo z těch, které byly dostatečně specifické, tj. produktem byl jediný fragment očekávané velikosti). 16 získaných fragmentů bylo přímo sekvenováno, většina z obou řetězců.

Tabulka 3. Primery pro amplifikaci mtDNA (forward).

Název primeru	Sekvence primeru 5'→ 3'
TM-J-210b ¹	AAGCTACTGGGTTTCATACCC
TBf1 ²	AACCTATTCTTCGATTTCACTCAG
LCO1490 ³	GGTCAACAAATCATAAAGATATTGG
LPFf1 ⁴	TGCTTTCCCTTCCAGTTCTTG
FoIC1-J-fa ²	ATGCATTTACTACTTGAAACG
FoIC2-J-fa ²	ACGCTGTGCCAGGACGTCT

FAf1 ⁴	CTGAAAGTAAGCGTCGGTCTCT
N3-J-5747 ¹	CCATTTGAATGTGGRTTTGAYCC
N5-J-7077 ¹	TTAAATCCTTWGARTAAAAYCC
7Rfor ⁴	TACCCATGTTATTAAGAAAGAA
7Ff1 ⁴	AGGAGCCTTAGACTGAAA
N4-J-8944 ¹	GGAGCTTCAACATGAGCTTT
CB-J-10933 ¹	TATGTACTACCATGAGGACAAATATC
LAf1 ⁴	ATTTCAACTTCCCCCTACG
Coll-LR-12542 ²	AGCCAGGTTGGTTTCTATCT

Tabulka 4. Primery pro amplifikaci mtDNA (revers).

Název primeru	Sekvence primeru 5'→3'
TBr1 ⁴	TTGAAGGGCTGTAAAGGATGC
CBr1 ⁴	ACTAAAGAAGACGCAACCAATGTAAT
FolC1-N-ra ²	ATTATAAATTTGGTCGTCTCC
LPRr1 ⁴	CATATTAAGTTGGGGTGAT
L2-N 3104 (Pat) ¹	TCCAATGCACTAATCTGCCATATTA
FolC2-N-ra ²	CCGCAAATTTCTGAGCATTG
NAr ²	GGTCCCCTCTCGTCTGATA
C3-N-5460 ¹	TCAACAAAGTGTCAGTATCA
7Rr1 ⁴	TTTCTACTTTTGCTCTGACTCCTATT
7Rr2 ⁴	CTCTTTGTGGGTTTCCGTTTTT
N5-N-7211 ¹	TTAAGGCTTTAYTATTTATRTGYGC
N5-N-7793 ¹	TTAGGTTGRGATGGNYTAGG
HAr1 ⁴	CATTTACACTGAGCTTATTTAGA
LR-N-13398 ¹	CGCCTGTTTAACAAAAACAT
SR-N-14745b ¹	CCAGCAGYYGCGGTTATAC

¹ Simon et al.(1994, 2006)

² navržený podle OGRE, Žurovcová a Slámová

³ Folmer (1994)

⁴ navržený podle získaných sekvencí *F.candida*, Žurovcová a Slámová

3.3 PCR

Požadované úseky DNA byly amplifikovány metodou polymerázové řetězové reakce, a to na termocyklérech Eppendorf Master Cycler, T3 Termocycler Biometra a T personal Biometra.

Typická PCR reakce měla objem 12 μl a obsahovala tyto složky (převzato z návodu výrobce pro TaKaRa Ex Taq) : 7,25 μl dd H_2O + 1,25 μl 10x Ex Taq pufru (TaKaRa) + 1,0 μl dNTP (2,5mM každý, TaKaRa) + 0,75 μl primer forward (5 μM) + 0,75 μl primer revers (5 μM) + 1,0 μl templátová DNA (100 ng / μl) + 0,05 μl TaKaRa Ex Taq polymerázy (5 jednotek / μl). Podle potřeby bylo přidáno různé koncentrace MgCl_2 . Reakce byly připravovány na ledu, polymeráza přidána jako poslední složka směsi.

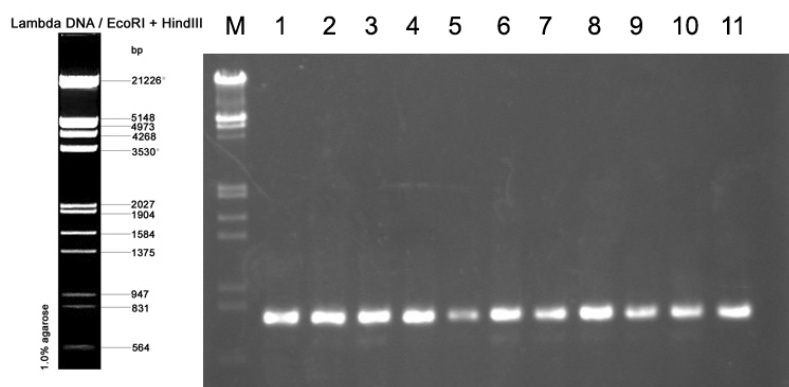
Standardní program PCR probíhal v následujících krocích : predenaturace při 94°C po 60 sekund, následovalo 35 cyklů 94°C 1 minuta, 37-56°C 20-30 sekund, 72°C 1-2 minuty). Reakci ukončila terminální elongace při 72°C po dobu 10 minut.

Pokud se nezdařilo získat požadovaný fragment standardním programem, případně jeho variacemi s jinou teplotou pro nasedání primeru na templát (annealing), použila jsem i metodu tzv. „step-up“ PCR. Amplifikace pomocí Step-up PCR probíhala v následujících krocích : predenaturace při 94°C po 60 sekund, následovalo 20 cyklů A (94°C 1 minuta, 37°C 20-30 sekund, 68°C 1-2 minuty) a 20 cyklů B (94°C 1 minuta, 37-56°C 20-30 sekund, 72°C 1-2 minuty). Reakci ukončila terminální elongace při 72°C po dobu 10 minut.

3.3 Gelová elektroforéza

Kontrola přítomnosti a velikosti DNA fragmentů byla prováděna elektroforézou na 1,5% agarozovém gelu. Gel byl připraven z agarozy (SeaKream LE Agarose) a 1x TAE pufru (50x TAE pufr: 242 g TRIS, 57,1 ml kyseliny octové, 100 ml 0,5M EDTA, 1000 ml H_2O , pH 8,0). a rozvařením v mikrovlnné troubě. Po ochlazení na cca 50°C byl nalit do připravené formy s vloženými hřebínky a ponechán přibližně 40 minut tuhnout při pokojové teplotě. Na připravený gel bylo nanášeno buď 5 μl PCR produktu anebo pouze 1 μl , pokud šlo o PCR produkt přečištěný.

Elektroforéza probíhala při pokojové teplotě a napětí 120 V. DNA byla následně barvena v lázni (100 ml 1x TAE pufru + 5 μl ethidiumbromidu o koncentraci 5 $\mu\text{g}/\text{ml}$) po dobu 25 minut a vizualizována na UV transiluminátoru (UVP Transilluminator). K odhadu velikosti naamplifikovaných fragmentů byl použit velikostní marker Lambda DNA /EcoRI+Hind III (Fermentas). Ukázka fotografie gelu po úspěšné PCR je na Obrázku 2.



Obrázek 2. Elektroforeogram produktů PCR s primery LCO1490 a HCO2198 (M – velikostní marker Lambda DNA /EcoRI+Hind III, 1-11 testované vzorky druhu *F. candida*)

3.4 Purifikace PCR

K přečištění PCR produktu byly použity tři následující metody, a to podle toho, jak kvalitní produkt se podařilo získat.

3.4.1 DNA Clean & ConcentratorTM-5 (ZYMO RESEARCH)

Tuto metodu jsem použila pro vzorky, u kterých jsem získala specifický, ale slabý produkt. Proto jsem tuto reakci opakovala ve větším objemu a zakoncentrovala podle návodu výrobce.

3.4.2 ZymocleanTM Gel DNA Recovery Kit (ZYMO RESEARCH)

Pokud se v produktu PCR vyskytlo více fragmentů, byly rozděleny pomocí elektroforézy, požadovaný fragment vyříznut z gelu a DNA izolována tímto kitem podle návodu výrobce.

3.4.3 Čištění pomocí směsi Exosap (Dugan et al., 2002)

Čištění pomocí enzymatické směsi ExoSAP-IT[®] (USB) je postupem pro specifické a silně koncentrované PCR produkty. Je také doporučeným standardem podle protokolu BOLD. Použila jsem upravený protokol výrobce, ve kterém jsem snížila množství použité směsi Exosap, ale prodloužila doby inkubace: k 10 μ l PCR produktu byl na ledu přidán 1 μ l Exosapu, následně pak byla směs inkubována 30 minut při 37°C a 15 minut při 87°C v termocykléru Perkin Elmer. Poté je produkt přímo sekvenován nebo zamražen po další použití.

3.5. Sekvenování DNA

Kvalitu přečištěných PCR produktů jsem opět zkontrolovala pomocí elektroforetické separace, při níž také byla vizuálně odhadnuta jejich koncentrace porovnáním s velikostním markerem Lambda DNA /EcoRI+Hind III (Fermentas).

3.5.1. Sekvenační směs

Pro sekvenování byl použit BigDye Terminator v 3.1. Cycle Sequencing Kit (Applied Biosystems). Složení reakční směsi pro jeden vzorek (20 μ l) : 2,0 μ l sekvenačního mixu + 3,0 μ l 5x sekvenačního pufru + 1,0 μ l primeru (5 μ M) + 10 - 40 ng PCR produktu; vše doplněno ddH₂O do 20 μ l.

3.5.2. Sekvenační program

Sekvenační reakce proběhla v 25 cyklech při teplotách 96°C 15 s, 50°C 20 s, 60°C 4 min.; s úvodní denaturací 96°C 1 min. (Eppendorf Master cycler, Perkin Elmer).

3.5.3. Přečištění sekvenační reakce pomocí Sephadexu a aerosolových špiček

Hydratace Sephadexu: Pro přípravu 10 sloupečků bylo rozmícháno 0,5 g Sephadex G-50 v 7 ml MilliQ H₂O. Bylo ponecháno minimálně 45 min při 4°C.

Příprava sloupečku: 1 ml aerosolová špička byla uříznuta cca 5 mm pod filtrem a umístěna do 1,5 ml mikrozkuřavky. Roztok Sephadexu byl důkladně promíchán. Bylo nanášeno 400 μ l na každý sloupeček a byla protlačena přebytečná voda. Bylo nanášeno dalších 400 μ l a stočeno 2 min při 1 000 g. Sloupečky byly umístěny do nových 1,5 ml mikrozkuřavek.

Čištění sekvenační reakce: Sekvenační reakce byla nanášena na sloupek sephadexu. Sloupečky byly centrifugovány 2 min. při 1 000 g. Přečištěná sekvenační reakce byla vysušena ve Speed-Vac (cca 20 min) při pokojové teplotě.

3.5.4 Sekvenace

Sekvenace byla provedena na přístroji ABI PRISM 3130xl firmy Applied Biosystems v Laboratoři genomiky, ÚMBR AV ČR.

3.6 Analýza sekvencí

Při práci s molekulárními daty a pro úpravu sekvencí byly použity programy EditSeq a SeqManII z programového balíku DNASTAR ver. 4.0 (DNASTAR, Inc.). Program EditSeq

sloužil jako textový editor sekvencí a také pro vyhledávání čtecích rámců (ORF, open reading frame) u genů kódujících proteiny. V programech SeqManII a MEGA verze 4 (Tamura *et al.* 2007) byly komplementární sekvence spojeny, upraveny a byla z nich vytvořena konsenzuální sekvence. Kontrolní identifikace sekvencí byla provedena pomocí databáze NCBI BlastSearch (<http://www.ncbi.nlm.nih.gov>). Aligmenty byly vytvořeny metodou ClustalW a také v programu MEGA verze 4, případně byly manuálně upraveny.

Statistika sekvencí byla provedena v programu MEGA verze 4, genetická vzdálenost byla spočítána pomocí modelu Kimura-2-parametr.

V programu MEGA verze 4 byla provedena shlukovací (klastrová) analýza metodou Neighbour-Joining (NJ). Pomocí modelu Kimura-2-parametr byly vytvořeny dendrogramy, (použita volba „Pairwise deletion“). Statistická podpora stromu byla otestována pomocí metody bootstrap (1000 opakování).

Analýza haplotypů a jejich frekvencí byla provedena pomocí programu DnaSP ver. 3 (Rozas & Rozas, 1999).

Analýza transferových RNA byla provedena v programu tRNAscan-SE (Lowe & Eddy, 1997) s nastavením „mitochondrial predictors“ a „cove score cut off = 1“. Protein kódující sekvence a rRNA byly identifikovány pomocí programu DOGMA (Wyman *et al.*, 2004). V případě neúspěchu uvedených programů jsme použili alignment (srovnávání) sekvencí s údaji o genomech dalších 4 druhů chvostoskoků, uvedených v databázi mitochondriálních genomů OGRE (<http://drake.physics.mcmaster.ca/ogre>; Jameson *et al.*, 2003)

4. VÝSLEDKY

4.1 Izolace DNA

Izolace DNA byla prováděna pomocí čtyř různých metod. Metoda chelexové izolace používaná v začátcích práce měla velmi nízkou úspěšnost a to zejména při izolaci DNA z jednotlivých jedinců. Jako nejlepší se ukázala metoda Extrakčního pufru, která byla ke konci práce téměř stoprocentně úspěšná.

4.2 Barcoding rodu *Folsomia*

Oblast genu COI používaná pro DNA barcoding byla získána z 28 jedinců z 5 druhů, u zbývajících 3 druhů se nepodařilo osekvenovat. Průměrná délka sekvence COI byla 672,5 bazí. Další statistické informace o použitých sekvencích jsou uvedeny v tabulce níže.

Tabulka 5. Statistika sekvencí COI.

Druh (počet jedinců)	Nukleotidové frekvence (%)				Délka sekvence
	T (U)	C	A	G	
<i>F. candida</i> (12)	34,5	20,9	25,3	19,4	684
<i>F. manolachei</i> (4)	34,8	18,9	28,1	19,5	560
<i>F. tesari</i> (1)	32,6	23,9	25	18,5	696
<i>F. quadriocullata</i> (9)	35,8	19,4	26,8	18,4	697
<i>F. penicula</i> (2)	30,3	23,4	26,3	20	707
Průměr	33,6	21,3	26,3	19,16	668,8

Počet variabilních míst získaných sekvencí COI je 310, což je 43,6 % z celkové délky sekvence.

Pro vyhodnocení vhodnosti uvedeného markeru pro druhovou identifikaci byly spočítány genetické vzdálenosti uvnitř jednotlivých druhů a mezi druhy navzájem (Tab. 6, 7)

Tabulka 6. Průměrné vnitrodruhové genetické vzdálenosti.

Druh (počet jedinců)	Průměr	Směrodatná odchylka
<i>F. candida</i> (12)	0,007	0,002
<i>F. manolachei</i> (4)	0,025	0,005
<i>F. tesari</i> (1)	n/c	n/c
<i>F. quadriocullata</i> (9)	0,147	0,011
<i>F. penicula</i> (2)	0,007	0,003

n/c ... průměr a odchylka nebyla spočítána u druhů s jedním osekvenovaným jedincem

Tabulka 7. Průměrné mezidruhové genetické vzdálenosti.

	1	2	3	4	5
1		0,025	0,018	0,017	0,020
2	0,271		0,025	0,017	0,023
3	0,207	0,265		0,018	0,022
4	0,237	0,191	0,242		0,017
5	0,232	0,232	0,257	0,227	

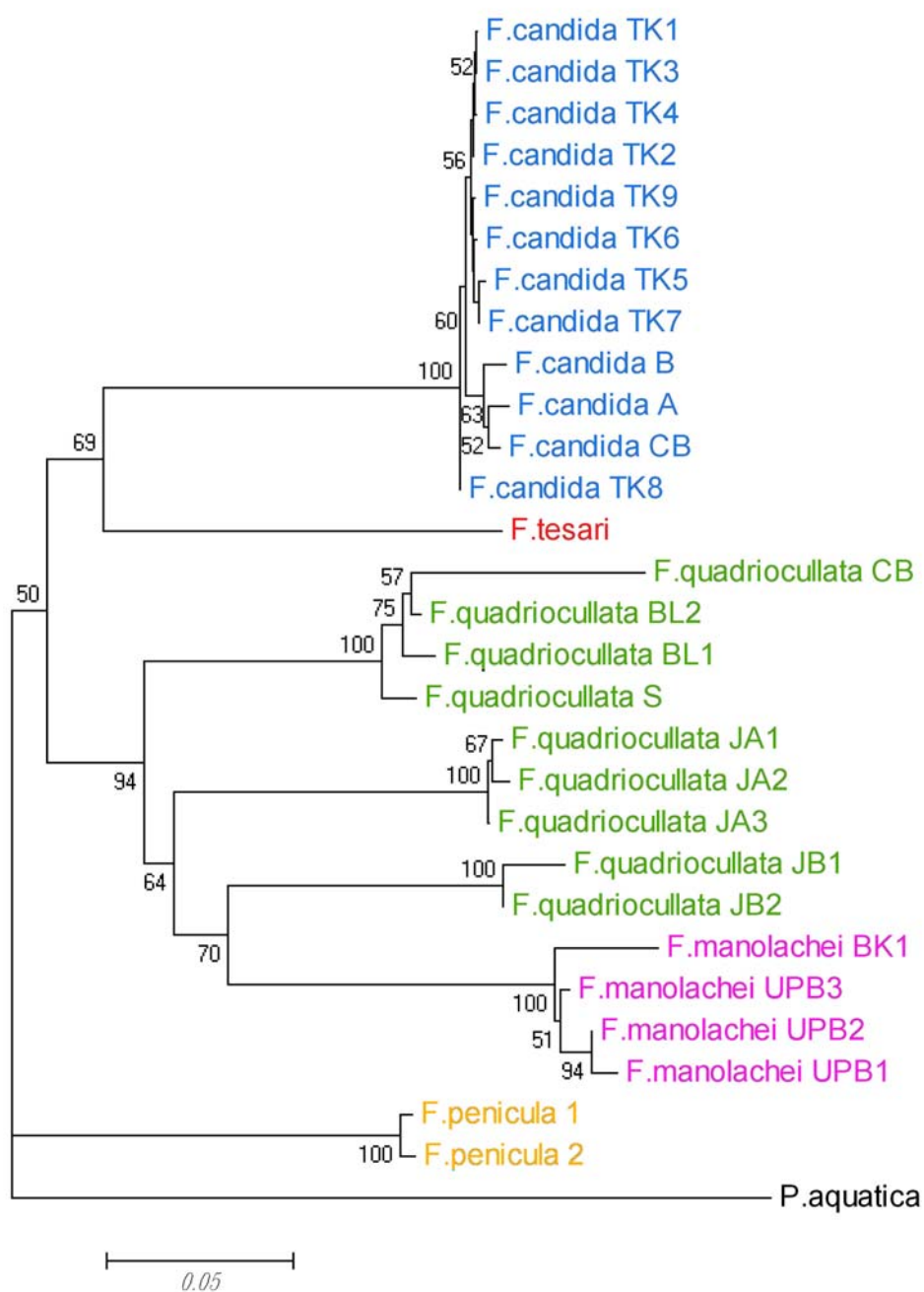
1 – *F. candida*, 2 – *F. manolachei*, 3 – *F. tesari*, 4 – *F. quadriocullata*, 5 – *F. penicula*. Černě jsou psány průměry, modře směrodatné odchylky od průměru.

Ve 28 sekvencích se vyskytovalo 18 různých haplotypů (viz Tab.8), které jsou rovnoměrně rozdělené mezi zkoumanými druhy a rody.

Tabulka 8. Seznam haplotypů COI

druh	haplotyp	frekvence	označení jedince
<i>F. candida</i>	1	9	TK1, 2, 3, 4, 5, 6, 7, 8, 9
	2	1	A
	3	1	CB
	4	1	B
<i>F. manolachei</i>	5	1	BK1
	6	1	UPB2
	7	1	UPB1
	8	1	UPB3
<i>F. tesari</i>	9	1	
<i>F. quadriocullata</i>	10	2	JB1, JB2
	11	1	JA3
	12	2	JA1, JA 2
	13	1	S
	14	1	CB
	15	1	BL1
	16	1	BL2
<i>F. penicula</i>	17	1	1
	18	1	2

Výsledkem shlukovací analýzy je dendrogram (Obrázek 3), který na základě vypočtené genetické vzdálenosti ukazuje, jak dalece jsou si vzorky podobné.



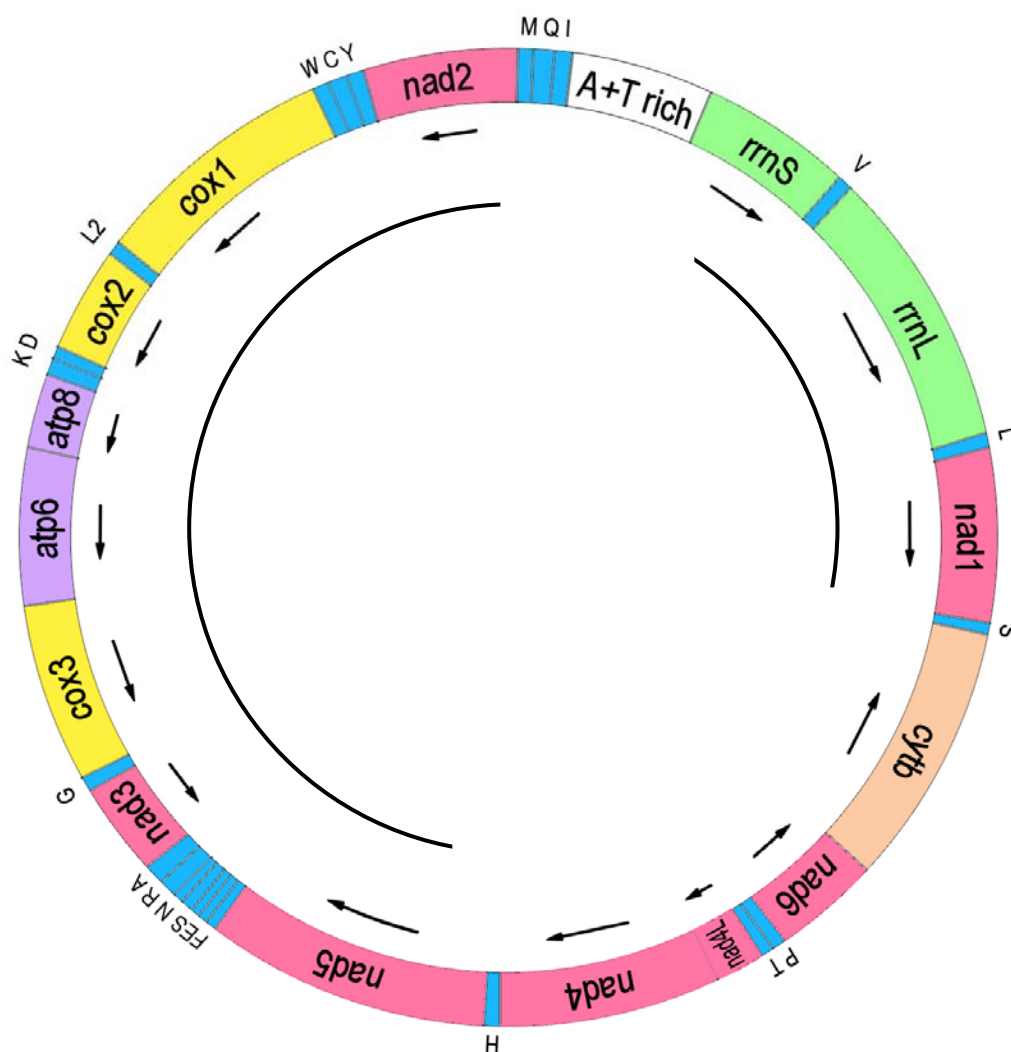
Obrázek 3. Dendrogram COI (Neighbour-Joining, Kimura-2-parametr, Pairwise deletion, Bootstrap 1000x) Zobrazeny jsou jen hodnoty bootstrapu 50 a vyšší. Jako outgroup použita *Podura aquatica* (převzato z databáze GeneBank, AY665319). Jednotlivé druhy jsou odlišeny různými barvami. Čísla označují různé jedince z téže lokality: TK – Těšínský kras, CB – České Budějovice, BL – Budislav u Litomyšle, JA – Jeseníky, Velká Hole, JB – Jeseníky – sjezdovka, BK – Bílé Karpaty, UPB – zahrada UPB v ČB

4.3 Mitochondriální genom *Folsomia candida*

Podarilo se nám osekvenovat celkem 10023 bp mitochondriálního genomu *F. candida*, a to ve dvou fragmentech dlouhých 7766 bp a 2257 bp, poloha těchto fragmentů na genomu *G. hodgsoni* je na Obr. 4. V těchto úsecích jsme identifikovali 16 genů pro transferovou

RNA, 10 genů kódujících proteiny a dva geny pro ribozomální RNA. Genomová mapa je v Příloze 1. Na „hlavním“ řetězci (J-strand) je kódováno 17 genů (*cox1*, *tRNA^{Leu-UUR}*, *cox2*, *tRNA^{Lys}*, *tRNA^{Asp}*, *atp8*, *atp6*, *cox3*, *tRNA^{Gly}*, *nad3*, *tRNA^{Ala}*, *tRNA^{Arg}*, *tRNA^{Asn}*, *tRNA^{Ser-AGN}*, *tRNA^{Glu}*, *nad2*, *tRNA^{Trp}*) a protější řetězec (N-strand) kóduje dalších 9 genů (*tRNA^{Phe}*, *nad5*, *nad1*, *tRNA^{Leu-CUR}*, *lrRNA (16S)*, *tRNA^{Val}*, *srRNA (12S)*). Pořadí genů, jejich délka a pozice na J/N řetězci je v Tabulce 9.

Frekvence nukleotidů A+T v celém genomu je 69,4 % (A = 35,7 %; T = 33,7 %; C = 13,0 %; G = 17,7 %). V genech kódujících protein je obsah A+T 69,1%, 78 % v tRNA a 74,7 % v rRNA. Frekvence nukleotidů pro jednotlivé geny je v Tabulkách 10, 11 a 12.



Obrázek 4. Poloha námi osekvenovaných fragmentů na genomu *G. hodgsoni*.

Tabulka 9. Pořadí genů, délka a pozice na osekvenovaném fragmentu a na J/N řetězci.

gen/tRNA	začátek	konec	délka	J/N (+/-)
Fragment 1				
ND2 nc	1	1049	1048	J
Trp	1048	1113	65	J
Cys	1174	1113	61	N
Tyr	1242	1178	64	N
COX1	1244	2782	1538	J
Leu-UUR	2778	2841	63	J
COX2	2842	3528	686	J
Lys	3528	3605	77	J
Asp	3605	3671	66	J
ATP8	3672	3848	176	J
ATP6	3842	4525	683	J
(Ser-UCN)	4165	4070	95	N
COX3	4533	5321	788	J
Gly	5325	5390	65	J
ND3	5391	5732	341	J
Ala	5754	5815	61	J
Arg	5819	5880	61	J
Asn	5879	5942	63	J
Ser-AGY	5942	6016	74	J
Glu	6012	6078	66	J
Phe	6146	6081	65	N
ND5 nc	6152	7766	1614	N
Fragment 2				
ND1 nc	1	320	319	N
Leu-CUN	324	405	81	N
16S RNA	1636	406	1230	N
Val	1637	1704	67	N
12S RNA nc	1705	2257	552	N

nc ... sekvence není úplná

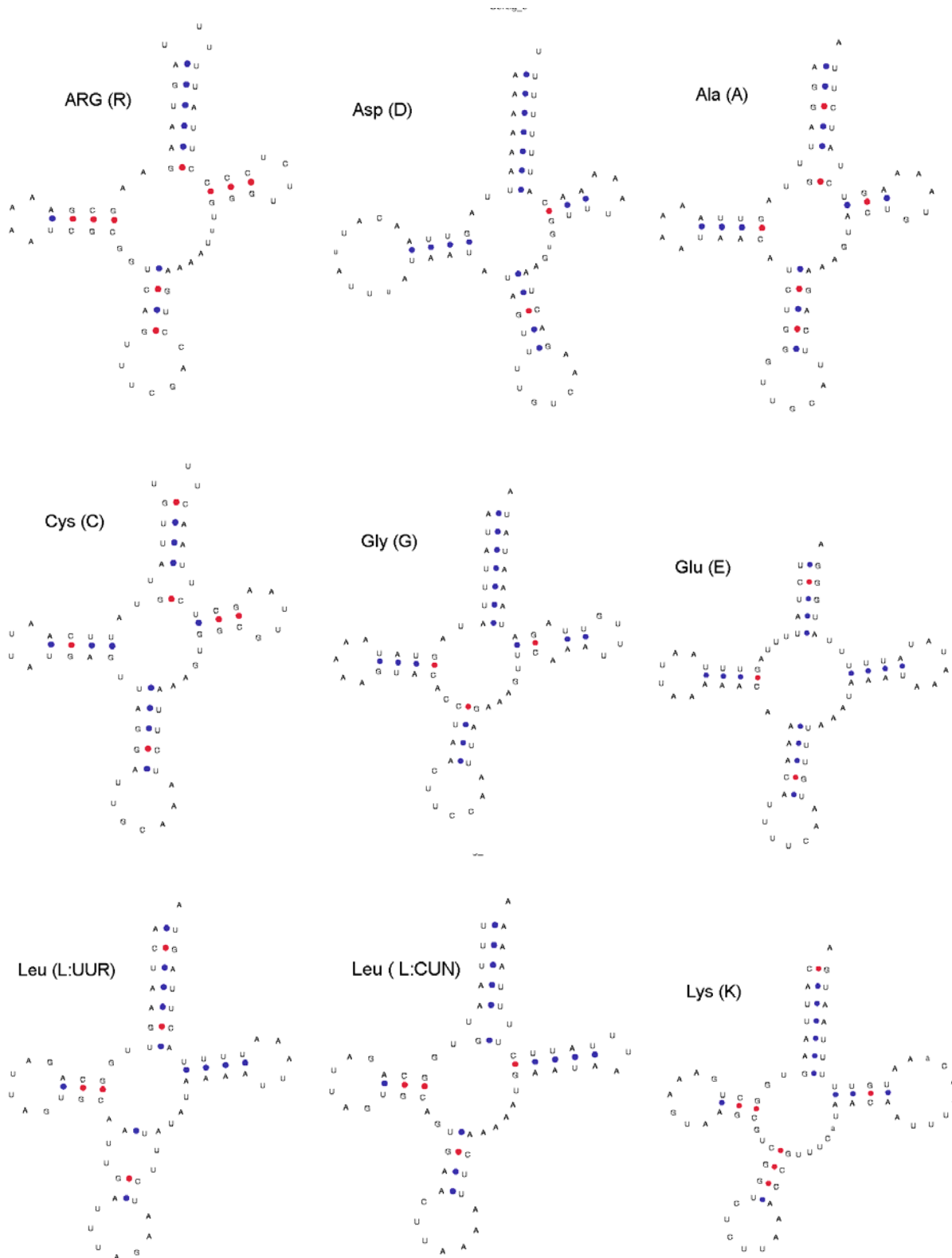
4.3.1 tRNA

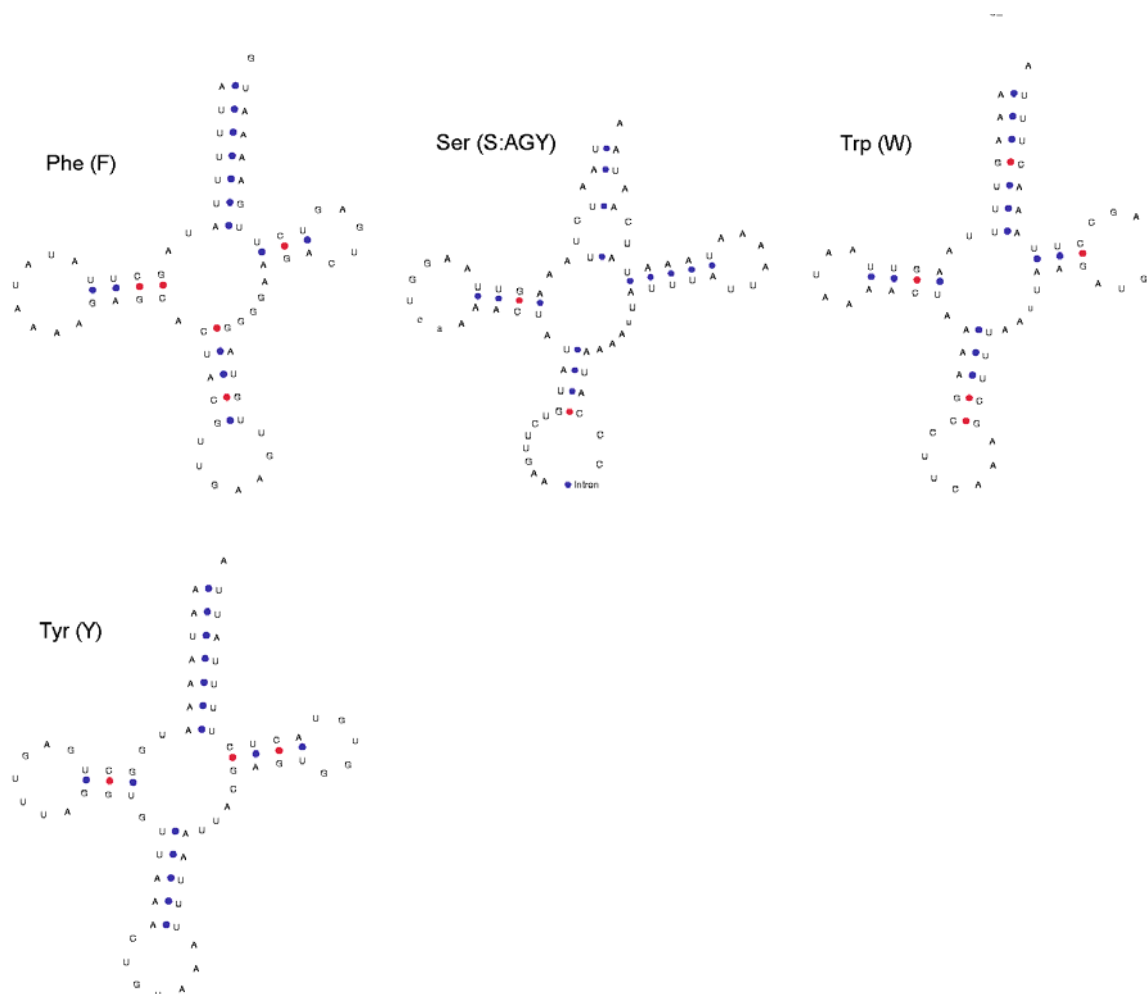
V získané sekvenci jsme za pomoci programu tRNAscan-SE (Lowe & Eddy, 1997) nebo srovnáním s dříve publikovanými sekvencemi identifikovali 16 tRNA. tRNAscan-SE také umožnil predikci jejich sekundární struktury a uspořádání do typických „trojlístků“ (viz Obr. 5), nukleotidové složení a antikodony jednotlivých tRNA jsou v Tab.10.

Tabulka 10. Nukleotidové složení jednotlivých tRNA a jejich antikodony.

tRNA	nukleotidové frekvence (%)				antikodon
	T (U)	C	A	G	
Trp	32,3	12,3	41,5	13,8	tca
Cys	45,2	3,2	45,2	6,5	gca
Tyr	32,3	20	40	7,7	gta
Leu-UUR	37,5	7,8	40,6	14,1	taa
Lys	34,6	15,4	34,6	15,4	ctt
Asp	43,3	6	40,3	10,4	gtc
(Ser-UCN)	47,9	10,4	32,3	9,4	tga
Gly	34,8	10,6	45,5	9,1	tcc
Ala	32,4	11,3	37,1	19,4	tgc
Arg	31,1	19,7	27,9	21,3	tcg
Asn	37,3	16,0	33,3	13,3	ni
Ser-AGY	36,5	16,2	33,8	13,5	ni
Glu	43,3	6	43,3	7,5	ttc
Phe	34,8	24,2	31,8	9,1	gaa
Leu-CUN	34,1	13,4	46,3	9	tag
Val	35,3	13,2	39,7	11,8	tac

ni ... nelze identifikovat





Obrázek 5. Sekundární struktura tRNA určená pomocí tRNAscan-SE (Lowe & Eddy, 1997), antikodony jednotlivých tRNA jsou na „lístku“ orientovaném směrem dolů.

4.3.2 Geny kódující proteiny

Devět genů kódujících protein bylo v sekvencích identifikováno pomocí programu DOGMA (Wyman *et al.*, 2004) a srovnáním s dříve publikovanými sekvencemi ostatních chvostoskoků (*Onychiurus orientalis*, *Gomphiocephalus hodgsoni*, *Podura aquatica* a *Tetrodontophora bielanensis*; OGRE databáze). Průměrný obsah A+T ve všech sekvencích kódujících protein je 69,1 %, přičemž nejvyšší byl v genu *nd1* (73,5 %) a nejnižší v *cox3* (61,9 %). Nukleotidové složení, iniciační a terminační kodony jednotlivých genů jsou v Tabulce 11.

Tabulka 11. Statistika sekvencí genů kódujících proteiny.

Geny	Nukleotidové složení (%)				Iničiační kodon	Terminační kodon
	T (U)	C	A	G		
COX1	35,3	20,6	26,8	17,2	atg	taa
COX2	33,4	19,5	33,8	13,3	att	taa
COX3	33,8	26,7	28,1	17,4	atg	tag
ND1	50,4	10,7	23,1	15,7	ata	nc
ND2	39,1	17,7	32,7	10,5	nc	taa
ND3	36,2	19,5	32,9	11,4	att	taa
ND5	42,6	14,3	25,2	17,8	nc	tag
ATP6	37,3	18,9	32	11,7	atg	taa
ATP8	34,5	20,3	33,9	11,3	att	taa
Průměr	38,3	17,1	30,8	14,3		

nc ... neúplná sekvence

4.3.3 Ribozomální RNA

Pomocí programu DOGMA (Wyman *et al.*, 2004) pro anotaci genomů byly v námi získaných sekvencích identifikovány 2 geny pro ribozomální RNA, jeden pro velkou a jeden pro malou ribozomální podjednotku. Průměrný obsah A+T je 74,7 %. Nukleotidové složení jednotlivých genů je uvedeno v Tabulce 12.

Tabulka 12. Statistika sekvencí získaných genů pro RNA

Geny	Nukleotidová variabilita			
	T (U)	C	A	G
RNL	38,9	10,6	33,9	16,7
RNS nc	39,9	9,3	36,7	14,2
Průměr	39,4	9,95	35,3	15,45

nc ... neúplná sekvence

5. DISKUZE

5.1 Izolace DNA

Byly vyzkoušeny čtyři různé postupy izolace DNA z jedinců i ze směsného vzorku z laboratorních chovů. Jako nejvhodnější se ukázala metoda Extrakčního puftru (Cox & Hebert, 2001; Frati et al. 2001), kterou bylo možné získat DNA i z některých velmi malých druhů (*F. tesari*). Při izolaci DNA pro sekvenování mitochondriálního genomu se osvědčil upravený postup s použitím ZR Genomic DNA II KitTM. Chelexová izolace se dařila jen ze směsných vzorků, ale s mnohem menší úspěšností než pomocí kitu.

Neúspěchy při kterémkoli typu izolace mohly být způsobeny velmi malou velikostí jedinců nebo degradací DNA při uchovávání vzorků v etanolu. Naopak u jedinců živých nebo uchovaných krátkou dobu v destilované vodě byla izolace úspěšná ve většině případů.

5.2 Použití molekulárního markeru COI (DNA barcoding) pro identifikaci druhů rodu *Folsomia*

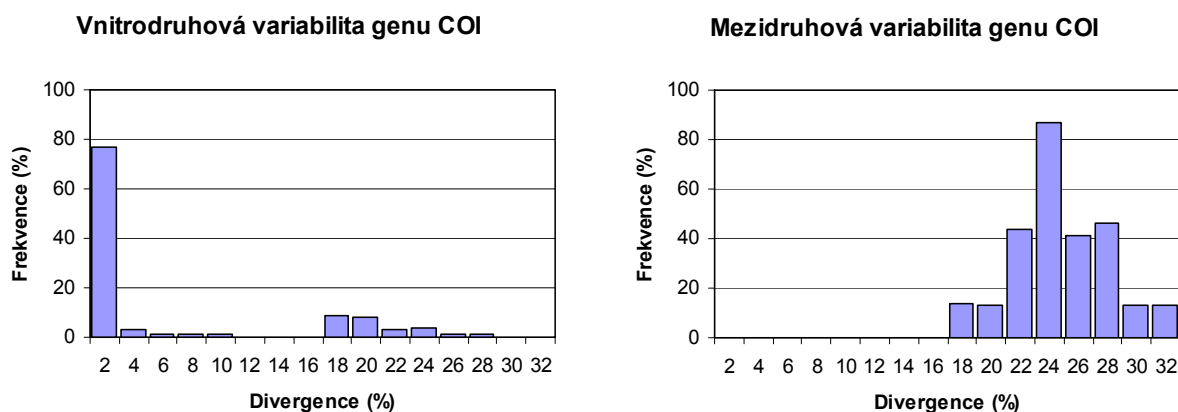
Základním předpokladem úspěšného použití barcodingu pro identifikaci druhů je, že vnitrodruhová variabilita DNA sekvencí tohoto markeru by měla být zhruba 10x vyšší než mezidruhová. Důležité však také je, v jakém rozmezí se tyto variability pohybují, neboť jde především o to, aby se tyto dva parametry významně nepřekrývaly.

V práci Hogga a Heberta z roku 2004, kde určovali známé druhy arktických zástupců třídy Collembola, byla pozorována ve většině případů vnitrodruhová variabilita pod 1 % a mezidruhová vždy přesáhla 8 %. V rámci druhu *F. quadriocullata* ale variabilita dosahovala 13 %, což autoři připisují přítomnosti dosud nepopsaných sesterských druhů. Přítomnost těchto morfologicky nerozlišitelných blízce příbuzných (tzv. kryptických) druhů je ve skupině Collembola dobře známá (Stevens & Hogg, 2003), proto je tento závěr velmi pravděpodobný. Průměrné zastoupení A+T v získaných frekvencích bylo 60 %, což odpovídá námi zjištěné hodnotě 59,9 %

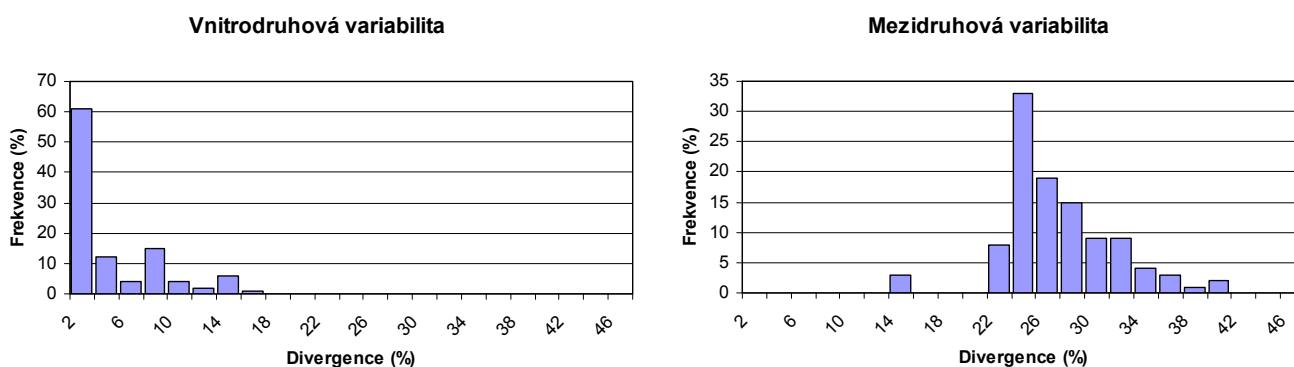
Průměrná mezidruhová variabilita v našem souboru rodu *Folsomia* byla 23,5 % (18 – 32 %), vnitrodruhová variabilita s průměrem 4,5 % se pohybovala mezi 0 – 10 % a 18 – 28 %. Avšak úrovně 18-28 %, kde se hodnoty překrývají, nabývala pouze v rámci druhu *F. quadriocullata* (zjištěná vysoká hodnota vnitrodruhové variability tohoto druhu je v souladu s předchozími výsledky Hogga a Heberta, 2004). Příčinou je pravděpodobně to, že zkoumaní jedinci tohoto druhu pocházeli z pěti různých lokalit i odlišných biotopů. Odpovídá tomu i jejich rozdělení do tří skupin v dendrogramu (Obr.3) Jedinci z Jeseníků byly v dendrogramu

umístění do dvou větví, což odpovídá jejich původu z různých populací žijících v odlišných prostředích. Mohlo by se tedy jednat o kryptické druhy, ovšem k potvrzení této teorie by bylo třeba analyzovat větší počet jedinců z každé populace. Další větev překvapivě spojuje zástupce pocházející z poměrně vzdálených lokalit (Budislav u Litomyšle, Šumava a České Budějovice), jejich zahrnutí do jedné skupiny ale může být zkreslené malým počtem analyzovaných jedinců (2 z Budislavi u Litomyšle a pouze po jednom z ČB a Šumavy).

Když vypustíme takto způsobené vysoké hodnoty vnitrodruhové variability odpovídají získaná data hodnotám zjištěným u ostatních chvostoskoků. Toto je možné pozorovat při srovnání histogramů vytvořených na základě matice genetických vzdáleností (analýza Pairwise distance calculation, Kimura-2-parametr v programu MEGA verze 4) na Obr. 6 a histogramů získaných z databáze BOLD (www.barcodinglife.org) na Obr. 7. V obou případech je vnitrodruhová variabilita nejčastěji v rozmezí 0 – 2 % a mezidruhová 22 – 26 %.



Obrázek 6. Vnitrodruhová a mezidruhová variabilita genu COI u rodu *Folsomia*



Obrázek 7. Vnitrodruhová a mezidruhová variabilita genu COI u skupiny Collembola (převzato z databáze BOLD, www.barcodinglife.org)

Nejnižší variabilita byla zjištěna uvnitř druhů *F. penicula* (0,7%), avšak to nemusí nutně znamenat genetickou homogenitu této populace, protože sekvence byly získány pouze ze dvou jedinců. Stejná míra variability byla zjištěna u druhu *F. candida*, který byl zastoupen celkem 12 jedinci a to jak z přírodní populace (Těšínský kras), tak ze tří laboratorních chovů založených ze vzorků z geograficky vzdálených lokalit. Všichni tito jedinci jsou v dendrogramu zařazeni do jedné skupiny, která je podpořena vysokou hodnotou bootstrapu. To ukazuje na pravděpodobnou homogenitu tohoto druhu, ačkoli pro potvrzení by bylo třeba analyzovat jedince z dalších přírodních populací. Tento druh je však považován za velmi „cestovatelský“, neboť se často dostává na velké vzdálenosti díky lidské činnosti, např. v půdě komerčně šířených hrnkovaných rostlin (Fountain a Hopkin, 2005), takže homogenizace populací může být skutečná. Populace jsou také převážně tvořeny partenogenetickými samicemi, což může přispět ke stejnorodosti i v rámci jediného vzorku.

5.3 Mitochondriální genom *Folsomia candida*

V současnosti jsou známy pouze 2 kompletní mitochondriální genomy chvostoskoků. a to *Gomphiocephalus hodgsoni* a *Tetrodontophora bielanensis*. U dvou dalších druhů (*Onychiurus orientalis* a *Podura aquatica*) je sekvence neúplná (chybí část zahrnující počátek 12S RNA a AT rich oblast), ale tyto jsou blízké příbuzné výše uvedeným druhům. Podle nejnovější práce Carapelliho *et al.* (2007) jsou již osekvenovány další 4 druhy, jejich sekvence však ještě nebyly zveřejněny a uvedená publikace se zmiňuje jenom o fylogeneticky významném pořadí genů. Z těchto důvodů jsme se při srovnávání námi získaných dat opírali především o *Gomphiocephalus hodgsoni*, který je fylogeneticky rodu *Folsomia* nejbližší, a další kompletní genom *T. bielanensis*. Tento však vykazuje několik translokací (Příloha 1), zřejmě specifických pro tuto evoluční linii.

5.3.1 Popis genomu

Získali jsme sekvenci mitochondriálního genomu *F. candida* pokrývající cca 65 % z předpokládané velikosti, a z dané sekvence jsme získali mnoho poznatků o struktuře mitochondriálního genomu. Zjištěné pořadí genů je stejné jako u *G. hodgsoni* s výjimkou *tRNA^{Ser-UGN}*, která byla u *F. candida* nalezena v oblasti genu *atp6* kódovaná na N řetězci, zatímco podle genomu *G. hodgsoni* by měla ležet v zatím neosekvenované části (mezi *cytb* a *nd1*) na J řetězci. U *T. bielanensis* je tento gen mezi *tRNA^{Met}* a *tRNA^{Gln}*, což je právě jedna z přestaveb, které ji odlišují od *G. hodgsoni*. U některých genů bylo zjištěno několik přesahů

(celkem 7), kdy čtecí rámec jednoho genu zasahoval do jiného. Tento jev je u členovců poměrně častý (Boore, 1999), takže *F. candida* není v tomto ohledu žádnou výjimkou.

5.3.2 tRNA

Při porovnání sekvencí jednotlivých tRNA se sekvencemi ostatních čtyř osekvenovaných druhů chvostoskoků byla zjištěna největší variabilita u $tRNA^{Cys}$ (71,9 %) a nejnižší u $tRNA^{Lys}$. Hodnoty pro ostatní tRNA jsou v tabulce níže.

Tabulka 13. Počet variabilních míst jednotlivých tRNA v porovnání s tRNA ostatních osekvenovaných druhů.

tRNA	Trp	Cys	Tyr	Leu2	Lys	Asp	(Ser UCN)	Gly	Ala	Arg	Asn	Ser-AGY	Glu	Phe	Leu CUN	Val
počet variabilních míst (%)	54,4	71,9	41,5	47,0	12,7	52,2	43,1	63,2	46,0	67,2	70,4	68,9	60,3	64,3	32,9	50,0

U dvou tRNA (pro $tRNA^{Asn}$ a $tRNA^{Ser-AGY}$), které se nepodařilo identifikovat pomocí tRNAScan-SE, ale na základě porovnání se sekvencemi ostatních chvostoskoků, nebylo možno spolehlivě predikovat sekundární strukturu. I když jsme se pokusili ji sestavit pomocí programu Sfold (Ding a Lawrence, 2003, <http://sfold.wadsworth.org/>) určeného právě pro práci s RNA, ani jedna tRNA neměla charakteristickou strukturu „trojlístku“. Bude tedy nutné další ověření jejich sekvence, případně parametrů, nezbytných pro správnou strukturální predikci.

Otázkou také zůstává umístění $tRNA^{Ser-UCN}$, která je podle výsledku tRNAScan-SE ve čtecím rámci *atp6*, což by bylo velmi neobvyklé. Vzhledem k tomu, že by byla přesunuta z oblasti, jejíž sekvenci se zatím nepodařilo získat, nelze o tomto umístění spolehlivě rozhodnout. Při práci s uvedeným programem jsme ale zjistili, že k takovému „falešnému“ umístění tRNA může také dojít v důsledku nastavení vyhledávacích parametrů na velmi nízké hodnoty.

5.3.3 Geny kódující proteiny

Ze 13 genů kódujících proteiny, které se v mitochondriích vyskytují se podařilo osekvenovat 6 kompletních genů a 3 částečně. Jejich pořadí v genomu se shoduje s pořadím zjištěným u *G. hodgsoni*.

Byly určeny iniciační a terminační kodony. Tři geny začínají ATG (*cox1*, *atp6* a *cox3*), tři ATT (*cox2*, *atp8* a *nd3*) a jeden ATA (*nd1*). Jsou to triplety, kterými obvykle začínají mitochondriální geny mnohobuněčných živočichů (Boore, 1999), stejně jako

nalezené terminační kodony; TAA v šesti genech (*nd2*, *cox1*, *cox2*, *atp8*, *atp6*, *nd3*) a TAG u genu *nd5*.

5.3.4 Ribozomální RNA

Byly identifikovány obě očekávané rRNA umístěné mezi tRNA v pořadí *tRNA^{Leu-CUR}*, *lrRNA (16S)*, *tRNA^{Val}* a *srRNA (12S)*, což odpovídá pořadí u *G. hodgsoni*.

5.3.4 A+T oblast

I přesto, že se dařilo naamplifikovat tuto oblast několika různými kombinacemi primerů, sekvenování nikdy neposkytlo výsledky použitelné pro analýzu. Obvykle totiž docházelo k amplifikaci několika produktů o různé velikosti, které se ani po izolaci z gelu a přečištění nedařilo sekvenovat. Tento problém však není ojedinělý. Jak je možno vidět v databázi OGRE, u řady nekompletních genomů chybí právě tento úsek (viz např. neúplný genom *O. orientalis* a *P. aquatica*). Je to způsobeno především vysokým obsahem nukleotidů A+T, podle nichž se daný úsek jmenuje, a také jeho sekundární strukturou. Nukleotidy se navíc nacházejí v delších opakovaných úsecích, jež způsobují předčasné ukončení PCR amplifikace či sekvenační reakce. Tento problém se jeví jako řešitelný amplifikací delšího fragmentu, který by danou oblast zahrnoval (Burger *et al.*, 2007) nebo zaklonováním všech PCR produktů a jejich následným sekvenováním.

6. ZÁVĚR

Testovali jsme použití mitochondriálního molekulárního markeru COI pro druhovou identifikaci vybraných zástupců rodu *Folsomia* (Collembola). Marker byl úspěšně amplifikován a osekvenován. Alignment sekvencí byl použit pro vytvoření dendrogramu metodou Neighbor-Joining s použitím modelu Kimura-2-Parameter. Zjistili jsme průměrnou mezidruhovou variabilitu 23,5 % a vnitrodruhovou variabilitu 4,5 %, přičemž u většiny druhů byla do dvou procent, jen u *F. quadriocullata* dosahovala 13 %. Vybrané druhy se podařilo rozdělit podle předpokladu a ukázat pravděpodobnou přítomnost kryptických druhů uvnitř druhu *F. quadriocullata*.

Získali jsme také přibližně 2/3 mitochondriálního genomu *F. candida*, kde jsme identifikovali 16 genů pro tRNA, 9 genů kódujících proteiny a dva geny pro ribozomální RNA. Pozice všech nalezených genů se shodují s pořadím zjištěným u *G. hodgsoni*, a jejich charakteristika se výrazně neodlišuje od ostatních mitochondriálních genomů zástupců řádu Collembola.

7. LITERATURA

- Boore J.L., Brown W.M., (1995) Complete sequence of the mitochondrial DNA of the anellid worm *Lumbricus terrestris*. *Genetics* 141:305-319
- Boore J.L. (1999) Animal mitochondrial genomes. *Nucleic Acids Research* 27: 1767 – 1780.
- Boore J.L., Lavrov D.V., Brown W.M. (1998) Gene translocation links insects and crustaceans. *Nature* 392:667-668
- Boore J.L., Staton J.L. (2000) The mitochondrial genome of the sipunculid *Phascolopsis fouldii* supports its association with Annelida rather than Mollusca. *Mol Biol Evol* 19:127-137
- Burger G, Lavrov D.V, Forget L, Lang B.F. (2007) Sequencing complete mitochondrial and plastid genomes. *Nat Protoc.* 2(3):603-14.
- Carapelli A., Liò P., Nardi F., van der Wath E., Frati F. (2007) Phylogenetic analysis of mitochondrial protein coding genes confirms the reciprocal paraphyly of Hexapoda and Crustacea. *BMC Evol Biol.* 7 Suppl 2:S8
- Cook C.E., Yue Q., Akam M. (2005) Mitochondrial genomes suggest that hexapods and crustaceans are mutually paraphyletic. *Proc Biol Sci.* 272(1569):1295-304.
- Cox A.J., Hebert P.D.N. (2001) Colonization, extinction, and phylogeographic patterning in a freshwater crustacean. *Molecular Ecology* 10:331-386
- Coroau Y., Gisclard C., Perotti P. (2002) The use of *Folsomia candida* (Collembola: Isotomidae) in bioassays of waste. *Appl. Soil Ecol.* 19:65-70
- Curole J.P., Kocher T.D. (1999) Mitogenomics? Digging deeper with complete mitochondrial genomes. *Trends Ecol Evol* 14:394-398
- Delsuc F., Philips M.H., Penny D. (2003). Comment on „Hexapod origins: monophyletic or paraphyletic?“ *Science*:1482d
- Diez J. A., De La Torre A.I., Cartagena M.C., Caravallo M., Vallejo S. (2001) Evaluation of the application of pig slurry to an experimental crop using agronomic and ecotoxicological approaches. *J. Environ. Qual.* 30:2165-75
- Ding, Y. and Lawrence, C.E. (2003) A statistical sampling algorithm for RNA secondary structure prediction. *Nucleic Acids Res.* 31:7280-7301.
- Dugan K.A., Lawrence H.S., Hares D.R., Fisher C.L., Budowle B. (2002) An Improved method for post-PCR purification for mtDNA sequence analysis. *Journal of Forensic Science* 47(4):811-8
- Engel M.S., Grimaldi D.A. (2004) New light shed on the oldest insect. *Nature* 427:627-30

- Fava F., Bertin L. (1999) Use of wxogenous specialised bacteria in the biological detoxification of dump site polychlorobiphenyl-contaminated soil in slurry phase conditions. *Biotechnol. Bioeng* 64:240-49
- Folmer L.J., Black M., Hoeh W., Lutz R., Vrijenhoek R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* 3:294-299
- Fountain M.T., Hopkin S.P. (2005) *Folsomia candida* (Collembola): A "Standard" Soil Arthropod. *Annu. Rev. Entomol.* 50:201-22
- Fрати F., Spinsanti G., Dallai R. (2001) Genetic variation of mt COII gene in the collembolan *Isotoma klovstadi* from Victoria land, Antarctica: evidence for population differentiation. *Polar Biology* 24:934-940
- Hajibabaei M., Janzen D.H., Burns J.M., Hallwachs W., Hebert P.D.N. (2006) DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Science of the United States of America* 103:968-971
- Hebert P.D.N., Cywinska A., Ball S.L., DeWaard JR (2003a) Biological identifications through DNA Barcodes. *Proceedings of the Royal Society of London Series B* 270:313-321
- Hebert P.D.N., Penton E.H., Burns J.M., Janzen D.H., Hallwachs W. (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences, USA* 101:14812-14817
- Hebert P.D.N., Stoeckle M.Y., Zemlak T.S., Francis C.M. (2004) Identification of Birds through DNA barcodes. *PLoS Biol* 2(10):e312
- Hoelzel A.R. (1998) *Molecular Genetics analysis of population* 2nd edition. Oxford University Press, New York.
- Hogg I.D., Hebert P.D.N. (2004) Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. *Can. J. Zool.* 82: 749-754
- Hopkin S.P. (1997) *Biology of the Springtails (Insecta: Collembola)*. Oxford, UK: Oxford Univ. Press. 330pp.
- ISO (1999) Soil duality-inhibition of reproduction of Collembola (*Folsomia candida*) by soil pollutants. Rep.No.ISO11267:1999(E). Geneva: International Standard Organization 16 pp.

- Jameson D., Gibson A.P., Hudelot C. & Higgs P.G. (2003) OGRé: a relational database for comparative analysis of mitochondrial genomes. *Nucl. Acids. Res.* 31, 202-206
- Jensen J., Lokke H., Holmstrup M., Krogh P.H., Rlsgaard L. (2001) Effects and risk assessment of linear alkylbenzene sulfonates in agricultural soil.5. Probabilistic risk assessment of linear lakylbenzene sulfonates in sludge-amended soils. *Environ. Toxicol. Chem.* 20:1690-97
- KumazawaY., Nishida M. (1999) Complete mitochondrial DNA sequence of the green turtle and blue-tailed mole skink: statistical evidence for archosaurian affinity of turtles. *Mol Biol Evol* 16:784-792
- Lowe T.M. & Eddy S.R. (1997) "tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence." *Nucl. Acids Res.* 25: 955-964
- Moore, J.C., B.B. Tripp, R. Simpson, and D.C. Coleman (2000). A springtail in the classroom: *Folsomia candida* as a model for inquiry-based laboratories. *American Biology Teacher* 62:512-519.
- Mindell D.P., Sorenson M.D., Dimcheff D.R, Hasegawa M., Ast J.C., Yuri T. (1999) Interordinal relationships of birds and other reptiles based on whole mitochondrial genomes. *Syst Biol* 48:138-152
- Myia M., Takeshima H., Endo H., Ishiguro N.B., Inoue J.G., Mukai T., Satoh T.P., Yamaguchi M., Kawaguchi A., Mabuchi K., Shirai S.M., Nishida M. (2003) Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences. *Mol Phylogenet Evol* 26:121-138
- Nardi F., Spinsanti G., Boore J.L., Carapelli A., Dallai R., Frati F. (2003) Hexapod origins: monophyletic or paraphyletic? *Science* 299:1887-89
- Nardi F., Spinsanti G., Boore J.L., Carapelli A., Dallai R., Frati F. (2003) Response to comment on „Hexapod origins: monophyletic or paraphyletic?“ *Science* 301:1482e
- Reyes A., Pesole G., Saccone C. (1998) Complete mitochondrial DNA sequence of the fat dormouse, *Glis glis*: further evidence of rodent paraphyly. *Mol Biol Evol* 15:499-505
- Rozas J., Rozas R. (1999) DnaSP version 3: an integrated program for molekular analysis. *Bioinformatics* 15: 174-175
- Siddall M.E, Budinoff R.B. (2005) DNA-barcoding evidence for widespread introductions of a leech from the South American *Helobdella triserialis* complex. *Conservation Genetics* 6:467-472

- Simon C., Frati F., Beckenbach A., Crespi B., Liu H., Flook P. (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.* 87:651-701
- Simon C., Buckley T.R., Frati F., Stewart J.B., Beckenbach A.T. (2006) *Annual Review of Ecology, Evolution, and Systematics* 37:545-79
- Smith M.A., Woodley N.E., Janzen D.H., Hallwachs W., Hebert P.D.N. (2005) DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). *Proceedings of the National Academy of Sciences, USA* 103:3657-3662
- Stevens M.I., Hogg I.D. (2003) Long term isolation and recent range expansion revealed for the endemic springtail *Gomphiocephalus hodgsoni* from southern Victoria Land, Antarctica. *Mol. Ecol.* 12:2357-2369
- Tamura K., Dudley J., Nei M., Kumar S. (2007) MEGA 4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24: 1596-1599
- Van Straalen N.M., Lokke H. (1997) *Ecological risk assessment of contaminants in soil.* London: Chapman & Hall.
- Ward R.D., Zemplak T.S., Innes B.H., Last P.R., Hebert P.D.N. (2005) DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society of London. Series B, Biological Science* 360:1847-1857
- Wyman S.K., Jansen R.K., Boore J.L. (2004) Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* 20(17):3252-3255

Internetové zdroje :

BOLD : <http://www.barcodinglife.org>

Collembase : <http://www.collembase.org>

GeneBank : <http://www.ncbi.nlm.nih.gov>

OGRé : <http://drake.physics.mcmaster.ca/ogre>

Sfold : <http://sfold.wadsworth.org>

tRNAscan-SE : <http://lowelab.ucsc.edu/tRNAscan-SE>

8. PŘÍLOHA

T. bielensis

M ND2 W C- Y- COX L2 COX2 K D ATP8 ATP6 COX3 G ND3 A R N S2 E F- ND5- H- ND4- ND4L T P- ND6 CYTB ND1- L- RNL- V- RNS- Q- I S

D. yakuba

M ND2 W C- Y- COX L2 COX2 K D ATP8 ATP6 COX3 G ND3 A R N S2 E F- ND5- H- ND4- ND4L T P- ND6 CYTB S ND1- L- RNL- V- RNS- I Q-

Gomphiocephalus hodgsoni

M ND2 W C- Y- COX L2 COX2 K D ATP8 ATP6 COX3 G ND3 A R N S2 E F- ND5- H- ND4- ND4L T P- ND6 CYTB S ND1- L- RNL- V- RNS- I Q-

Folsomia candida

M ND2 W C- Y- COX L2 COX2 K D ATP8 ATP6 S- ATP6 COX3 G ND3 A R N S2 E F- ND5- H- ND4- ND4L T P- ND6 CYTB S ND1- L- RNL- V- RNS- I Q-

Příloha 1. Pořadí genů v genomu *F. candida* a srovnání s *G. hodgsoni*, *T. bielensis* a *D. yakuba*. Šipky znázorňují translokace.