

Studium genetické diverzity a struktury populací pomocí mtDNA a jaderných mikrosatelitových markerů u včely medonosné

Přednáška

Modul č. 1: Genetika zvířat

Tomáš Urban, Aleš Knoll

Mendelova univerzita v Brně, Agronomická fakulta



Co-funded by
the European Union

Obsah

- Získávání genetických dat
- Hodnocení genetické variability pomocí sekvencí mtDNA
- Hodnocení genetické variability pomocí jaderných STR markerů (mikrosatelity)



Co-funded by
the European Union

Proč je důležitá genetická diverzita?

- Genetická diverzita pro zdraví v populaci, zásadní pro toleranci a rezistenci k současným i budoucím nemocem, patogenům a predátorům,
- Současný stav lze přisoudit mimo jiné i ke snížení diverzity,
- Diverzita včel byla hodnocena morfometrickými znaky (parametry křídel, pigmentace...)
- Je popsáno na 31 poddruhů (plemen, ras) včely medonosné
- Pomocí analýz DNA (hlavně mitochondriální) byly popsány evoluční linie: Západní mediteránní **M**, severní mediteránní **C**, africká **A**, orientální linie **O**...



Chromosome	GenBank	RefSeq	Size (bp)	GC content (%)
LG1	CM000054.5	NC_007070.3	29 893 408	n/a
LG2	CM000055.5	NC_007071.3	15 549 267	n/a
LG3	CM000059.5	NC_007072.3	13 234 341	n/a
LG4	CM000060.5	NC_007073.3	12 718 334	n/a
LG5	CM000058.5	NC_007074.3	14 363 272	n/a
LG6	CM000056.5	NC_007075.3	18 472 937	n/a
LG7	CM000063.5	NC_007076.3	13 219 345	n/a
LG8	CM000061.5	NC_007077.3	13 546 544	n/a
LG9	CM000062.5	NC_007078.3	11 120 453	n/a
LG10	CM000068.5	NC_007079.3	12 965 953	n/a
LG11	CM000057.5	NC_007080.3	14 726 556	n/a
LG12	CM000067.5	NC_007081.3	11 902 654	n/a
LG13	CM000064.5	NC_007082.3	10 288 499	n/a
LG14	CM000065.5	NC_007083.3	10 253 655	n/a
LG15	CM000066.5	NC_007084.3	10 167 229	n/a
LG16	CM000069.5	NC_007085.3	7 207 165	n/a
MT	na	NC_001566.1	16 343	15

Genů celkem	12 398
Protein-kódujících	9 935
Nekódující geny	2 421

CHROMOZOMY VČELY MEDONOSNÉ



Co-funded by
the European Union

- Mitochondriální DNA (mtDNA) většiny živočišných druhů má 16 až 20 kb,
- Mitochondriální genom včely medonosné západní má přibližně 16,5-17 kb,
- Referenční genom NC_001566 má 16 343 bp, Kompletní genom mtDNA karniolské rasy (MN250878) má 16 358 bp,

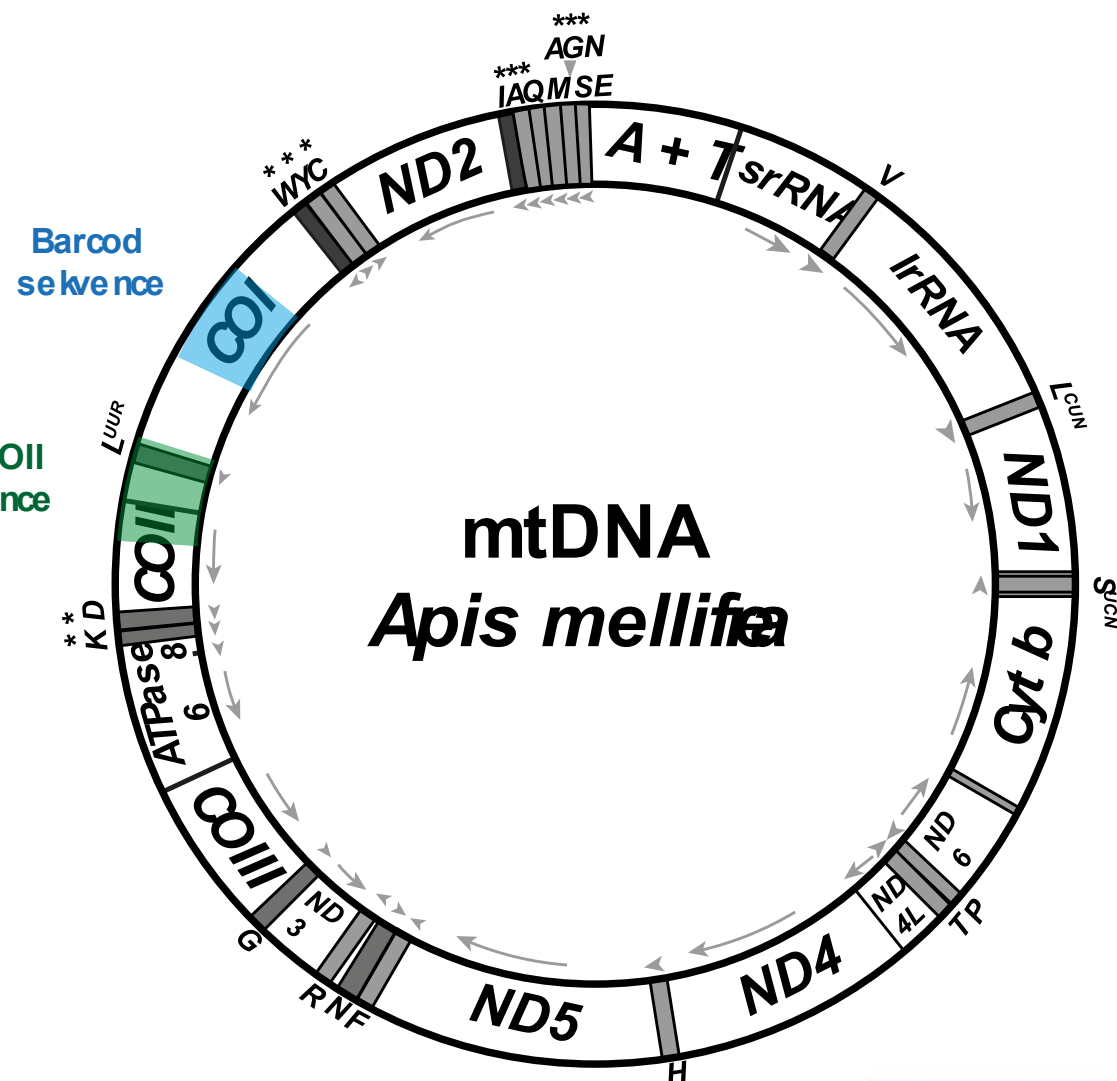


Co-funded by
the European Union

mtDNA

- Kompletní sekvence mitochondriální DNA včely medonosné (*Apis mellifera*) je u sekvenovaného kmene dlouhá **16 343 bp**,
- **13 genů kódující proteiny**
- 22 genů tRNA
- geny rRNA (s a l-rRNA)
- R H Crozier, Y C Crozier, The mitochondrial genome of the honeybee *Apis mellifera*: complete sequence and genome organization,, *Genetics*, Volume 133, Issue 1, 1 January 1993, Pages 97–117, <https://doi.org/10.1093/genetics/133.1.97>

MT DNA se používá pro fylogeografii a genetické studie , nejsou nebo nízké rekombinace, výhodné pro studium evoluce, konzervované struktury a uniparentální dědičnost,



Co-funded by
the European Union

Jakou variabilitu lze pozorovat?

Molekulární genetické markery

- bi-alelický jednonukleotidový polymorfismus (**SNP**) v kódujících nebo častěji v nekódujících intronových či intergenových regionech
- inzerce-delece nukleotidu (**indel**)

...ccccggctgacaagtgtg**C**ggtccacagg...
...ggggccgactgttcacac**G**ccagggtgtcc...

...ccccggctgacaagtgtg**A**ggtccacagg...
...ggggccgactgttcacac**T**ccagggtgtcc...

↑
SNP

delece C v oblasti sekvence ACA

...ccccggctg**aca**agtgtgcggtccacagg...
...ggggccgact**gtt**tcacacgccagggtgtcc...

...ccccggctg**aa**agtgtgcggtccacagg...
...ggggccgact**ttt**tcacacgccagggtgtcc...

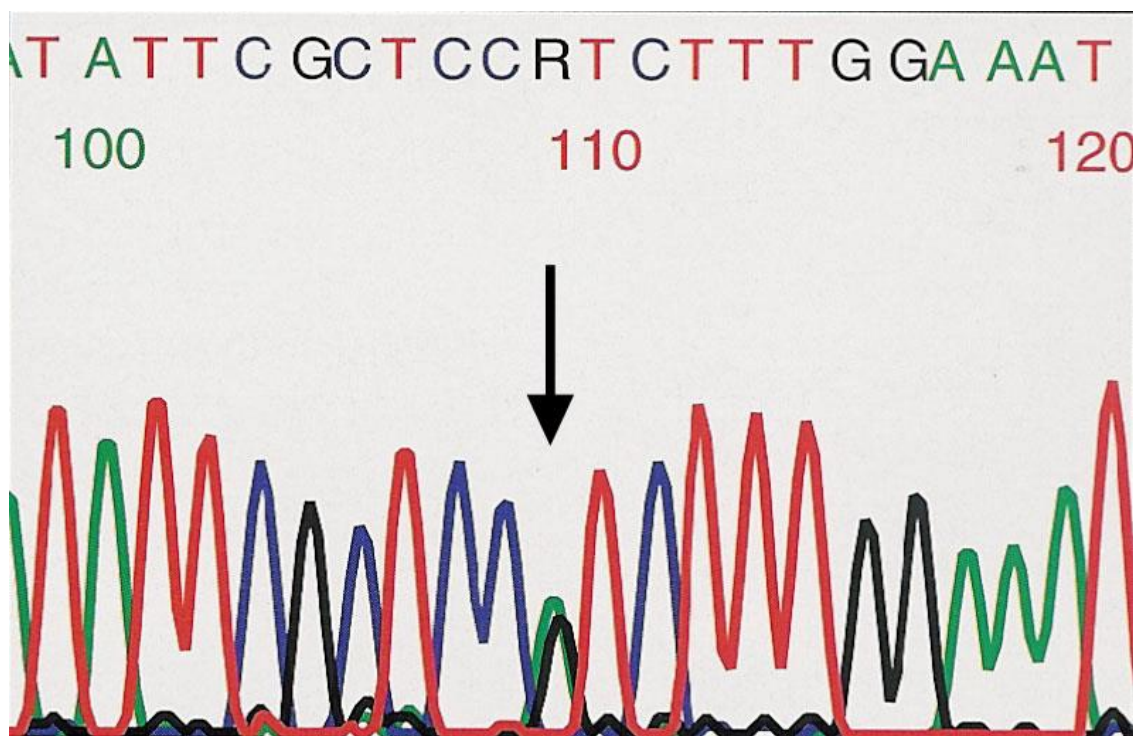


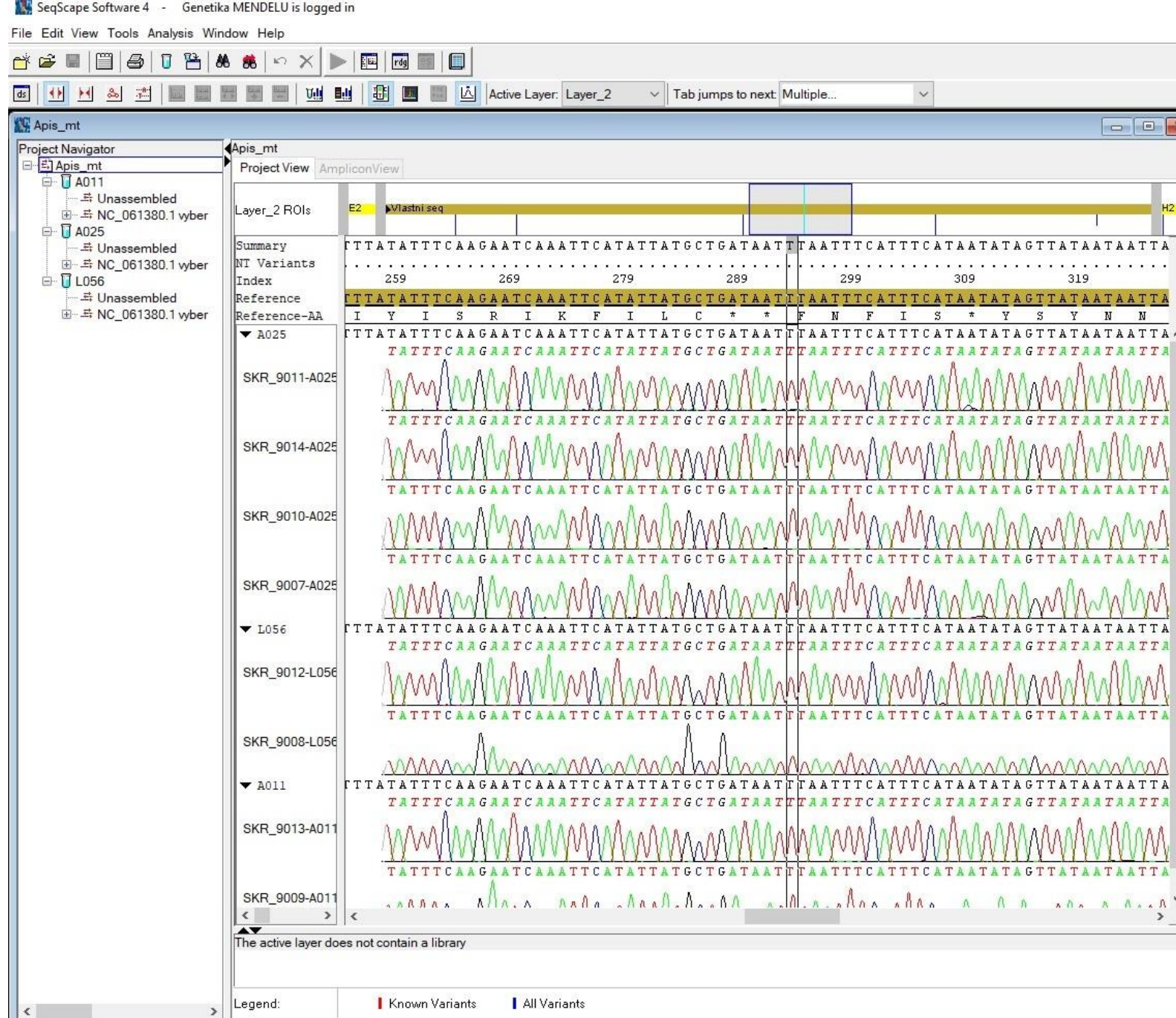
Co-funded by
the European Union

Sekvenátor

Kapilární elektroforézou se jednotlivá vlákna různé délky oddělí podle velikosti a odečte se odpovídající fluorescenční barva, která označuje bázi přítomnou v daném místě DNA.

Např, řetězec dlouhý 100 nukleotidů vyzařuje zelené světlo, tj, v pozici 100 templátu se nacházel adenin.





mt DNA -
sekvenování



Co-funded by
the European Union

Identifikace sublinií (mitotypy) v linii C (carnica)

Využívají se zejména tyto dvě sekvence mtDNA

- Sekvenování intergenové oblasti COI-COII (*tRNA^{leu}-cox2*)
 - se skládá ze dvou mitochondriálních genů: transferové RNA (tRNA)-Leu, a podjednotka cytochromoxidázy II (*cox2*)
 - Tento fragment obsahuje vysoké množství AT a vykazuje významné rozdíly v délce a složení nukleotidů mezi populacemi včely medonosné, Tento ampikon se štěpí restriční endonukleázou *DraI*, která specificky rozpoznává sekvenci TTAAA, aby se určila linie, Sekvenování oblasti barcodingu v genu COI (*cox1*)
 - To se porovnává se sekvencemi uloženými v databázích, jako je BoldSystem a GenBank, Tento fragment DNA je v rámci taxonu vysoce konzervovaný a často se používá k rozlišení taxonů.



Co-funded by
the European Union

Evoluční linie podle oblasti COI-COII u *A*, *m*,

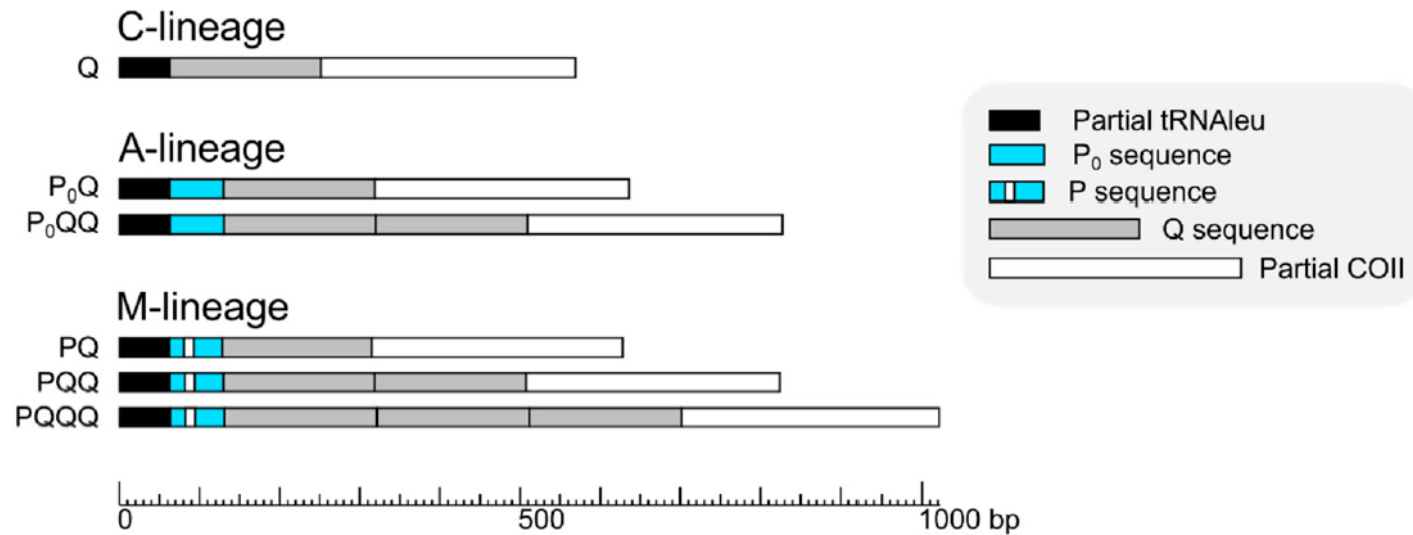
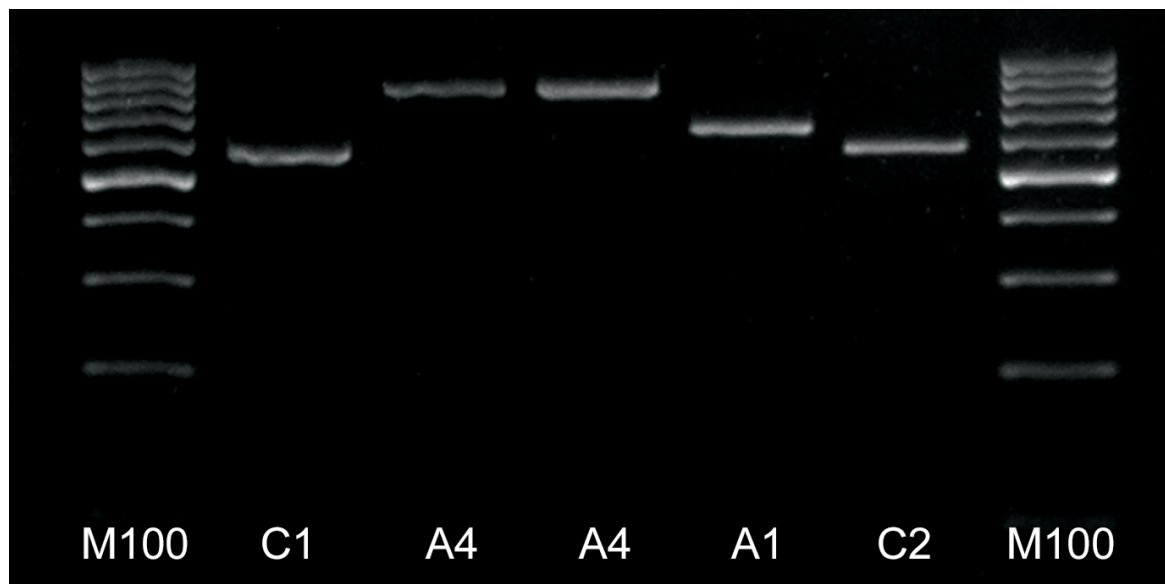


Figure 2. Schematic representation of COI-COII intergenic region of the mtDNA examined in this study. It is composed of partial sequences of tRNA^{leu} and COII genes, and a variable number of non-coding sequences P and Q. Bees of C-lineage have only one Q-sequence and no P-sequences, while A- and M-lineages have up to three Q-sequences and one P-sequence, which occurs in two variants, i.e., P₀ in A-lineage, and P in M-lineage.

Oleksa et al, 2021



Jakoukoliv odchylku (v důsledku mutace) nerozpozná -> místo nebude štěpit

Výsledkem je vznik několika fragmentů, jejichž počet a pozice se využívají pro určení evoluční linie (M, C, A, O...)

- PCR produkt oblasti *tRNA^{leu}-cox2* mitochondriální DNA pomocí elektroforézy v agarosovém gelu, M: 100 bp DNA ladder, C1 a C2 (Q), A4 (P_0 QQ) a A1 (P_0 Q) linie

>A054. 709 nucleotides.

```
-----GATCTGGAaTACTAGGATCATCAATGAGA
CTTATTATTCGAATAGAATTAAGATCCCCAGGATCATGAATTAACAATGATCAAATTTATAATACAATTGTTACTAGTCA
TGCATTCTAATAATTTTTTTTTATAGTTATACCATTTTTTAATTGGAGGATTGGAAATTGGCTTATTCCTTTAATACTAG
GATCACCTGATATAGCATTCCCCGAATAAATAATATTAGATTTTGATTACTTCTCCCTCATTATTTATACTTTTATTA
AGAAATTTATTTTATCCAAGACCAGGAACGGATGAACAGTATATCCACCATTATCAGCATATTTATATCATTCTTCACC
TTCAGTAGATTTTGCAATTTTTCTCTTCATATATCAGGAATTCCTCAATTATAGGATCATTAACTTAATAGTTACAA
TTATAATAATAAAAAATTTTTCTATAAATTATGACCAAATTCATTATTTCCATGATCAGTTTTTATTACAGCAATTTTA
TTAATTATATCATTACCTGTATTAGCTGGAGCAATTACTATACTATTATTTGATCGAAATTTTAATACATCATTTTTTCGA
TCCTATAGGAGGTGGAGATCCAATTCCTTTATCAACATTTATTT-----
```

>A071. 709 nucleotides.

```
-----GATCTTGTATATTATTCTAGCTTTATGATCTGGAATACTAGGATCATCAATGAGA
CTTATTATTCGAATAGAATTAAGATCCCCAGGATCATGAATTAACAATGATCAAATTTATAATACAATTGTTACTAGTCA
TGCATTCTAATAATTTTTTTTTATAGTTATACCATTTTTTAATTGGAGGATTGGAAATTGGCTTATTCCTTTAATACTAG
GATCACCTGATATAGCATTCCCCGAATAAATAATATTAGATTTTGATTACTTCTCCCTCATTATTTATACTTTTATTA
AGAAATTTATTTTATCCAAGACCAGGAACGGATGAACAGTATATCCACCATTATCAGCATATTTATATCATTCTTCACC
TTCAGTAGATTTTGCAATTTTTCTCTTCATATATCAGGAATTCCTCAATTATAGGATCATTAACTTAATAGTTACAA
TTATAATAATAAAAAATTTTTCTATAAATTATGACCAAATTCATTATTTCCATGATCAGTTTTTATTACAGCAATTTTA
TTAATTATATCATTACCTGTATTAGCTGGAGCAATTACTATACTATTATTTGATCGAAATTTTAATACATCATTTTTTCGA
TCCTATAGGAGGTGGAGATCCAATTCCTTTATCAACATTTATTT-----
```

>A1026. 709 nucleotides.

```
-----GATCTTGTATATTATTCTAGCTTTATGATCTGGAATACTAGGATCATCAATGAGA
CTTATTATTCGAATAGAATTAAGATCCCCAGGATCATGAATTAACAATGATCAAATTTATAATACAATTGTTACTAGTCA
TGCATTCTAATAATTTTTTTTTATAGTTATACCATTTTTTAATTGGAGGATTGGAAATTGGCTTATTCCTTTAATACTAG
GATCACCTGATATAGCATTCCCCGAATAAATAATATTAGATTTTGATTACTTCTCCCTCATTATTTATACTTTTATTA
AGAAATTTATTTTATCCAAGACCAGGAACGGATGAACAGTATATCCACCATTATCAGCATATTTATATCATTCTTCACC
TTCAGTAGATTTTGCAATTTTTCTCTTCATATATCAGGAATTCCTCAATTATAGGATCATTAACTTAATAGTTACAA
TTATAATAATAAAAAATTTTTCTATAAATTATGACCAAATTCATTATTTCCATGATCAGTTTTTATTACAGCAATTTTA
TTAATTATATCATTACCTGTATTAGCTGGAGCAATTACTATACTATTATTTGATCGAAATTTTAATACATCATTTTTTCGA
TCCTATAGGAGGTGGAGATCCAATTCCTTTATCAACATTTATTT-----
```

>A1065. 709 nucleotides.

```
-----GATCTTGTATATTATTCTAGCTTTATGATCTGGAATACTAGGATCATCAATGAGA
CTTATTATTCGAATAGAATTAAGATCCCCAGGATCATGAATTAGCAATGATCAAATTTATAATACAATTGTTACTAGTCA
TGCATTCTAATAATTTTTTTTTATAGTTATACCATTTTTTAATTGGAGGATTGGAAATTGGCTTATTCCTTTAATACTAG
GATCACCTGATATAGCATTCCCCGAATAAATAATATTAGATTTTGATTACTTCTCCATCATTATTTATACTTTTATTA
AGAAATTTATTTTATCCAAGACCAGGAACGGATGAACAGTATATCCACCATTATCAGCATATTTATATCATTCTTCACC
TTCAGTAGATTTTGCAATTTTTCTCTTCATATATCAGGAATTCCTCAATTATAGGATCATTAACTTAATAGTTACAA
TTATAATAATAAAAAATTTTTCTATAAATTATGACCAAATTCATTATTTCCATGATCAGTTTTTATTACAGCAATTTTA
TTAATTATATCATTACCTGTATTAGCTGGAGCAATTACTATACTATTATTTGATCGAAATTTTAATACATCATTTTTTCGA
TCCTATAGGAGGTGGAGATCCAATTCCTTTATCAACATTTATTT-----
```

Individuální sekvence *tRNA^{leu}-cox2*

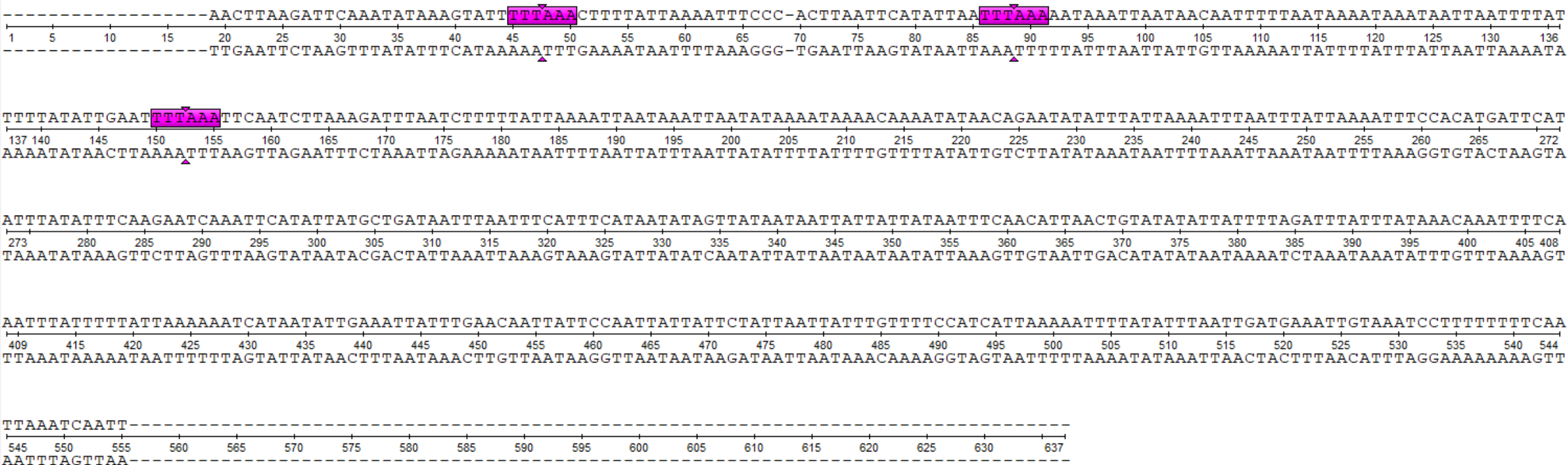


Co-funded by
the European Union

DraI test pro oblast COI-COII

Linie C

The *DraI* restriční spektrum *tRNA^{leu}-cox2* fragmentu byly analyzovány in silico pomocí programu **Unipro UGENE**



■ DraI
■ DraI
■ DraI

Restriction Site 45..50
Restriction Site 86..91
Restriction Site 149..154



Co-funded by
the European Union

Linie A

```

-----AACTTAAGATTCAAATATAAAGTATTTTTAAACCTTTTATTAAAATTAATAAATTAATATAAAATAAAACAAAATATAACAAAATATATTTATTAAAATTTAATTTATTAAAATTCCCC
1      5      10     15     20     25     30     35     40     45     50     55     60     65     70     75     80     85     90     95     100    105    110    115    120    125    130    136
-----TTGAATTCTAAGTTTATATTTTCATAAAAAATTTGAAAATAATTTTAATTATTTAATTATATTTTATTTTGTTTTATATTGTTTTATATAAATAATTTTAAATTAAATAATTTTAAGGGG

ACTTAATTCATATTAATTTTAAATAATAAATTAATAACAATTTTAATAAAATAAATAATTAATTTATTTTTATATTGAATTTTAAATTCAATCTTAAAGATTTAATCTTTTATTAAAATTAATAAATTAATATAAAA
137 140    145    150    155    160    165    170    175    180    185    190    195    200    205    210    215    220    225    230    235    240    245    250    255    260    265    272
TGAATTAAGTATAAATTAATTTTATTTAATTATTGTTAAAATTATTTTATTTATTAATTAATAAAAAATATAACTTAAAAATTAAGTTAGAATTTCTAAATTAGAAAAATAATTTTAATTATTTAATTATATTTT

TAAACAAAATATAATAGAATATATTTATTAAAATTTAATTTATTAAAATTTCCACATGATTCATATTTATATTTCAAGAATCAAATTCATATTATGCTGATAATTTAATTTTCATTTTCATAATATAGTTATAATAA
273    280    285    290    295    300    305    310    315    320    325    330    335    340    345    350    355    360    365    370    375    380    385    390    395    400    405    408
ATTTTGTTTTATATTATCTTATATAAATAATTTTAAATTAATAAATTTTAAAGGTGTACTAAGTATAAATATAAAGTCTTAGTTTAAAGTATAATACGACTATTAAATTAAGTAAAGTATTATATCAATATTATT

TCATTATTATAATTTCAACATTAAGTGTATATATTATTTTAGATTTATTTATAAATAAATTTTCAAATTTATTTTTATTAAAAAATCATAATATTGAAATTATTTGAACAGTTATTCCAATTATTATTCTATTAAT
409    415    420    425    430    435    440    445    450    455    460    465    470    475    480    485    490    495    500    505    510    515    520    525    530    535    540    544
AGTAATAATATTAAAGTTGTAATTGACATATATAATAAAATCTAAATAAATATTTATTTTAAAGTTTAAATAAAAAATAATTTTTTAGTATTATAACTTTAATAAACTTGTCATAAAGGTTAATAATAAGATAATTA

TATTTGTTTTCCATCATTAAAAATTTTATATTTAATTGATGAAATTGTAAATCCTTTTTTTTCAATTAAATCAATT-----
545    550    555    560    565    570    575    580    585    590    595    600    605    610    615    620    625    630    637
ATAAACAAAAGGTAGTAATTTTAAAATATAAATTAAGTACTTTAACATTTAGGAAAAAAAGTTAATTTAGTTAA-----

```

■ Dral

■ Dral

Restriction Site

45..50

Restriction Site

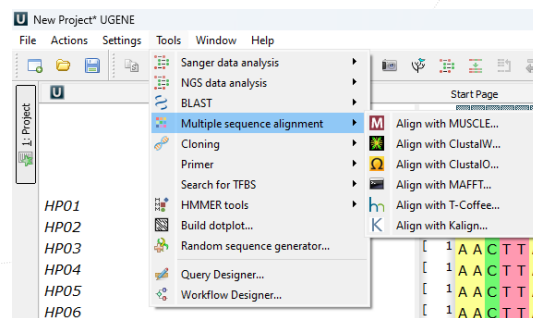
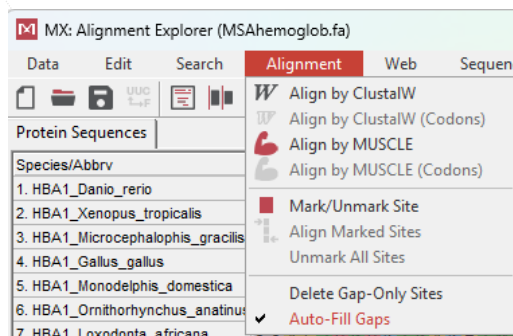
153..158



Co-funded by
the European Union

Nástroje pro multiple sequence alignment - MSA

- Haplotypy identifikované v obou oblastech byly charakterizovány pomocí vícenásobného přirovnání sekvencí s použitím **Kalign**
www.ebi.ac.uk/jdispatcher/msa/kalign
- Další vhodné MSA nástroje vhodné pro středně velké zarovnání :
 - Clustal Omega** (který ke generování zarovnání používá nasazené vodící stromy HMM profile-profile techniky k tvorbě alignmentů)
 - MAFT** (používá Fast Fourier Transforms)
 - www.ebi.ac.uk/jdispatcher/msa
- Programy MEGA, Unipro UGENE



Erasmus+ project 2021-1-SK01-KA220-HED-000032068



Co-funded by
the European Union

Restriction site Dml*tRNA^{Leu}* – gray; P0 – orange; Q – blue (2nd duplication of Q – black); *cox2* – purple

Site position by Franck et al. (2000)

Site position without P₀ and Q insertion

C1a	1	a
C21	25	
C2e		
C2d		
C2c		
C2j		
C2s		
C2d7		
C2i		
C2y		
Alha		
A4s		
A4na		

Site position with P₀ and Q insertion

C1a	45	46
C21		
C2e		
C2d		
C2c		
C2j		
C2s		
C2d7		
C2i		
C2y		
Alha		
A4s		
A4na		

112 – parsimony informative site

C1a	11	12	13	14	j	90
C21						
C2e						
C2d						
C2c						
C2j						
C2s						
C2d7						
C2i						
C2y						
Alha						
A4s						
A4na						

C1a	116	15	16	17	18	19	20	21
C21								
C2e								
C2d								
C2c								
C2j								
C2s								
C2d7								
C2i								
C2y								
Alha								
A4s								
A4na								

201, 202 – parsimony informative site

C1a	187	189	17	205	207	223	230
C21							
C2e							
C2d							
C2c							
C2j							
C2s							
C2d7							
C2i							
C2y							
Alha							
A4s							
A4na							

C1a	TAAA	-----
C21	TAAA	-----
C2e	TAAA	-----
C2d	TAAA	-----
C2c	TAAA	-----
C2j	TAAA	-----
C2s	TAAA	-----
C2d7	TAAA	-----
C2i	TAAA	-----
C2y	TAAA	-----
Alha	TAAA	-----
A4s	TAAA	-----
A4na	TAAA	-----

Site position with P₀ and Q insertion

C1a	45	46
C21		
C2e		
C2d		
C2c		
C2j		
C2s		
C2d7		
C2i		
C2y		
Alha		
A4s		
A4na		

112 – parsimony informative site

C1a	11	12	13	14	j	90
C21						
C2e						
C2d						
C2c						
C2j						
C2s						
C2d7						
C2i						
C2y						
Alha						
A4s						
A4na						

C1a	21	22
C21		
C2e		
C2d		
C2c		
C2j		
C2s		
C2d7		
C2i		
C2y		
Alha		
A4s		
A4na		

511 – parsimony informative site

C1a	22	23
C21		
C2e		
C2d		
C2c		
C2j		
C2s		
C2d7		
C2i		
C2y		
Alha		
A4s		
A4na		

586 – parsimony informative site

C1a	23	24	25
C21			
C2e			
C2d			
C2c			
C2j			
C2s			
C2d7			
C2i			
C2y			
Alha			
A4s			
A4na			

640, 646 – parsimony informative site

C1a	26	27
C21		
C2e		
C2d		
C2c		
C2j		
C2s		
C2d7		
C2i		
C2y		
Alha		
A4s		
A4na		

695 – parsimony informative site

C1a	27	28
C21		
C2e		
C2d		
C2c		
C2j		
C2s		
C2d7		
C2i		
C2y		
Alha		
A4s		
A4na		

712 – parsimony informative site

C1a	28	29
C21		
C2e		
C2d		
C2c		
C2j		
C2s		
C2d7		
C2i		
C2y		
Alha		
A4s		
A4na		

712 – parsimony informative site

Zarovnání vícenásobného sekvenování pomocí Kalign a ručního zpracování

tRNA^{Leu}-cox2



Co-funded by
the European Union

- Pro detekci polymorfismů DNA a haplotypů v obou oblastech mtDNA s využitím všech sekvencí z vícenásobného zarovnání sekvencí (ve formátu FASTA) byl použit program **DnaSP**



- Nukleotidové substituce a inserce/delece pro každý haplotyp byly určeny porovnáním pozic se sekvencemi bez primerů a referenčním genomem NC_001566, (www.ncbi.nlm.nih.gov/nuccore/NC_001566)
- K identifikaci specifických haplotypů *tRNA^{leu}-cox2* linií C a A byly vyhledány referenční sekvence se 100% identitou, (**BLAST**, blast.ncbi.nlm.nih.gov/Blast.cgi) v nukleotidové databázi Národního centra pro biotechnologické informace (NCBI) (GenBank),
- K ověření haplotypů cox1 byl použit BLAST, Sekvence cox1 byla také ověřena pomocí databáze **BOLD**(boldsystems.org/index.php/IDS_IdentificationRequest)



Detekované haplotypy *tRNA^{leu}-cox2*

L	G	Haplotype	NCBI accession	Dral spectrum (bp)	Seq. length (bp)	NCBI reference	
						accession number (% similarity)	designated as
C	Q	C1a	PP350804	47, 41, 64, 420	572	JQ97769.1 (100)	carnica
C	Q	C2l	PP350810	47, 40, 64, 419	570	OR761864.1 (100)	carnica
C	Q	C2e	PP350812	47, 40, 63, 420	570	JQ977702.1 (100)	carnica
C	Q	C2d	PP350805	47, 40, 64, 420	571	JF723977.1 (100)	carnica
C	Q	C2c	PP350806	47, 40, 64, 420	571	JF723976.1 (100)	carnica
C	Q	C2j	PP350807	47, 40, 64, 420	571	KX463941.1 (100)	carnica
C	Q	C2s	PP350808	47, 40, 63, 421	571	JF723979.1 (100)	carnica
C	Q	C2d7*	PP397042	46, 40, 64, 420	570	JF723977.1 (99,81)	carnica
C	Q	C2i	PP350809	47, 40, 64, 420	571	JQ977703.1 (100)	carnica
C	Q	C2y	PP350811	47, 40, 64, 420	571	JQ754650.1 (100)	carnica
A	P ₀ Q	A1ha*	PP430326	47, 108, 482	637	KX463739.1 (99,83)	iberiensis
A	P ₀ QQ	A4s	PP430327	47, 107, 192, 483	829	MW939597.1 (100)	n,a,
A	P ₀ QQ	A4na*	PP430328	47, 108, 190, 482	827	KX463793.1 (99,75)	iberiensis

C haplotypy COI-COI intergenové oblasti

Pozice/ Haplotyp	T/del 29	C/del 50:1 (inz)	T/del 92:1 (inz)	T/A 132	A/T 185	A/del 187	G/A 225	A/del 221	T/C 250	C/T 385
C1a	T	C	T	T	A	A	G	A	T	C
C2l	T	del	T	T	A	del	G	A	T	C
C2e	T	del	del	T	A	A	G	A	C	T
C2d	T	del	T	T	A	A	G	A	C	T
C2c	T	del	T	T	A	A	G	A	T	C
C2j	T	del	T	T	T	A	G	A	C	T
C2s	T	del	T	A	A	A	G	A	C	T
C2d7*	del	del	T	T	A	A	G	A	C	T
C2i	T	del	T	T	A	A	A	A	C	T
C2y	T	del	T	T	A	A	G	del	T	C

IS AGREED

Haplotypy v sekvenci pro barcoding (v COI)

Pozice/ Haplotyp	G/A 1	G/A 52	A/G 99	C/T 142	G/A 196	T/C 259	T/A 271	C/A 274	C/A 421	A/G 448	T/C 571
HpB_1	G	G	A	C	G	T	T	C	C	A	T
HpB_2	G	G	G	C	G	T	T	A	C	A	T
HpB_3	G	G	G	C	G	T	T	C	C	A	T
HpB_4	G	G	G	C	G	T	T	C	C	G	T
HpB_5	G	G	A	C	G	T	T	C	A	A	C
HpB_6	G	G	G	T	G	T	T	C	C	A	T
HpB_7	A	G	A	C	G	T	T	C	C	A	T
HpB_8	G	A	A	C	G	T	T	C	C	A	T
HpB_9	G	G	A	C	G	C	T	C	C	G	T
HpB_10	G	G	G	C	A	T	T	C	C	A	T
HpB_11	G	G	G	C	G	T	A	C	C	A	T
HpB_12	A	G	G	C	G	T	T	C	C	G	T
HpB_13	G	G	A	C	A	T	T	C	C	A	T

ISAGREED

Frekvence zjištěných haplotypů u *A*, *m*, v České republice

tRNA^{leu}-cox2

Haplotype	N	Frequency
C1a	155	0,5032
C2l	25	0,0812
C2e	58	0,1883
C2d	12	0,0390
C2c	47	0,1526
C2j	1	0,0032
C2s	4	0,0130
C2d7	1	0,0032
C2i	1	0,0032
C2y	1	0,0032
A1ha	1	0,0032
A4s	1	0,0032
A4na	1	0,0032
Total	308	1,0000

cox1

Haplotype	N	Frequency
HpB01	64	0,2078
HpB02	134	0,4351
HpB03	65	0,2110
HpB04	24	0,0779
HpB05	4	0,0130
HpB06	4	0,0130
HpB07	4	0,0130
HpB08	1	0,0032
HpB09	1	0,0032
HpB10	1	0,0032
HpB11	1	0,0032
HpB12	1	0,0032
HpB13	1	0,0032
HpB14	1	0,0032
HpB15	1	0,0032
HpB16	1	0,0032
Total	308	1,0000

Parametry diverzity haplotypů *tRNA^{leu}-cox2* a *cox1*

Haplotypová diverzita Hd

Molekulární diverzita π

Tajima's D

Haplotype	N	H	Hd	π	D
<i>tRNA^{leu}-cox2</i> (796 bp)	308	13	0,68186	0,00172 (1,38726-06)	-3,02386 (0,00249)
<i>cox1</i> (658 bp)	308	16	0,71866	0,00203 (1,96124e-06)	-1,60658 (0,10815)

Indexy genetické diverzity, včetně nukleotidové diverzity π , haplotypové diverzity Hd a Tajimova D pro testování hypotézy neutrální mutace, byly vyhodnoceny pomocí balíčku **pegas** v programu R,
Lze použít také **DnaSP**, **MEGA**, **Arlequin**



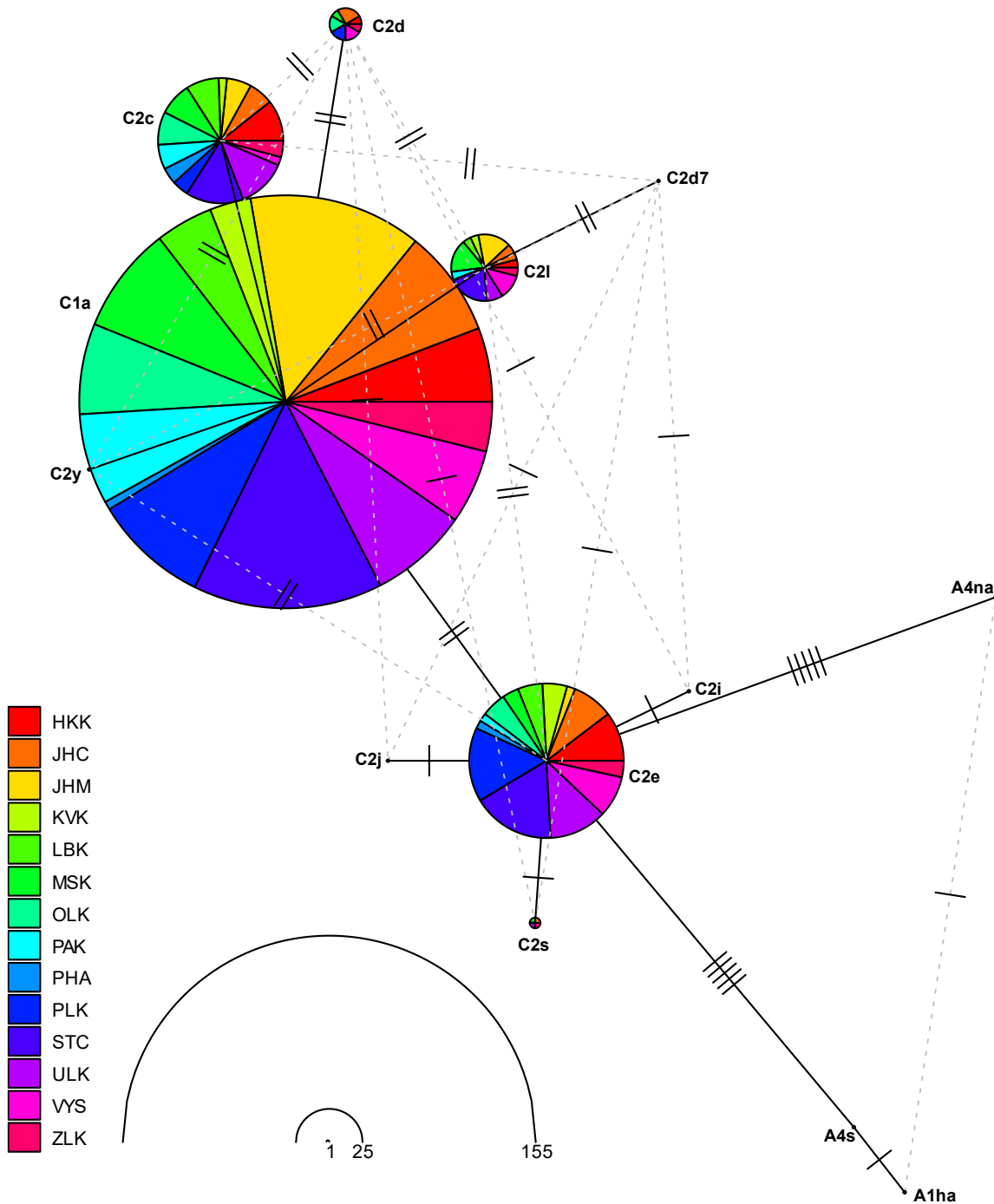
Co-funded by
the European Union

Haplotypové sítě

- **RMST** (the randomized minimum spanning tree) haplotypové sítě znázorňující frekvence a vztahy haplotypů *tRNA^{leu}-cox2* a *cox1* zjištěných v České republice.
- Fylogeografické sítě byly generovány metodou RMST s 1000 iteracemi pomocí funkce `rmst` v balíčku **pegas** v **R**
- Může být použit program:
 - PopART (Leigh, Bryant, 2015)

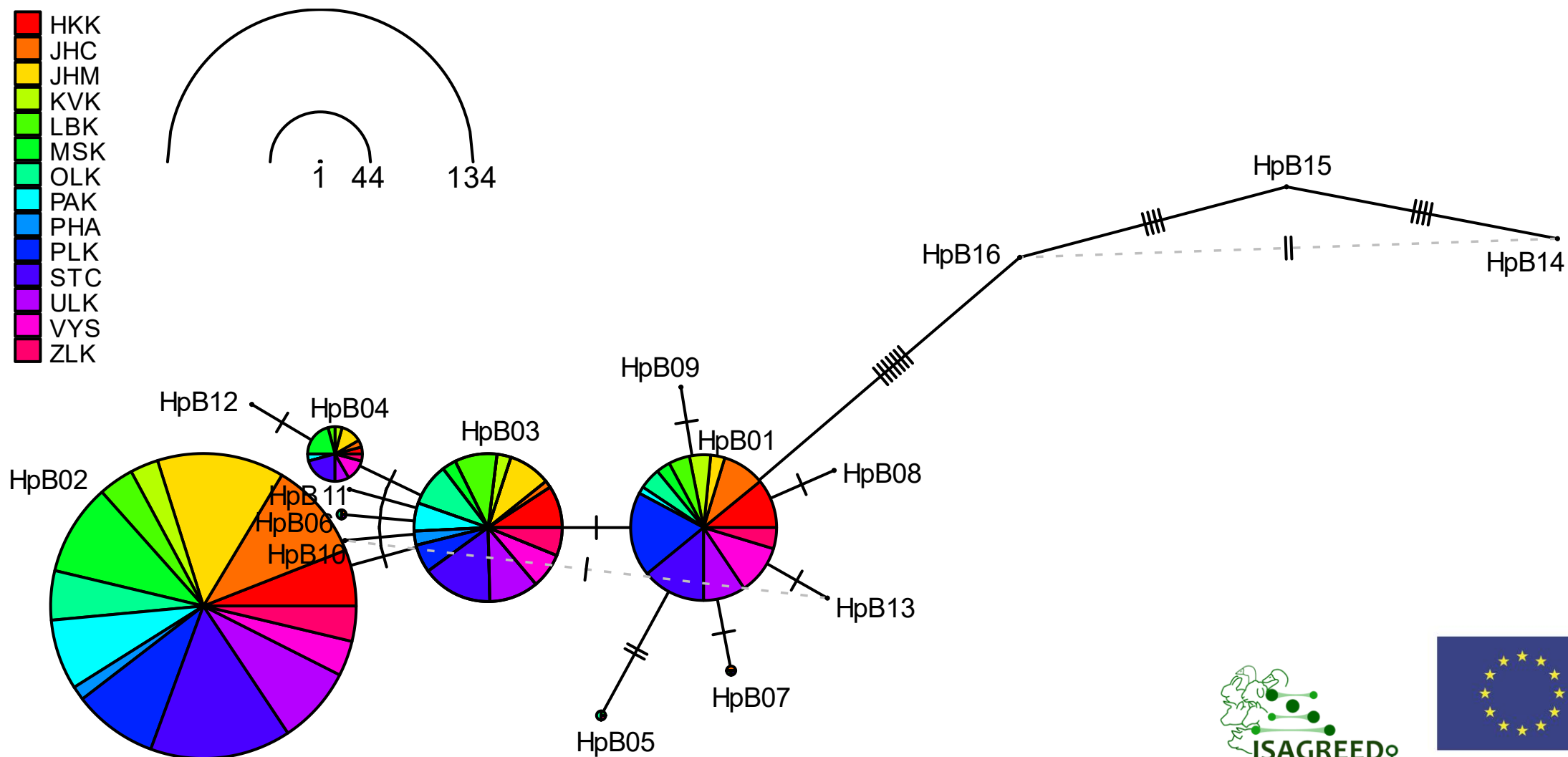


Co-funded by
the European Union



- Analýza haplotypové sítě RMST mezi studovanými 308 sekvencemi *tRNA^{leu}-cox2*
- Velikost kruhů je úměrná počtu jedinců a různé barvy představují různé kraje

- Analýza haplotypové sítě RMST mezi studovanými 308 sekvencemi *cox1*



Fylogenetická analýza na základě sekvencí

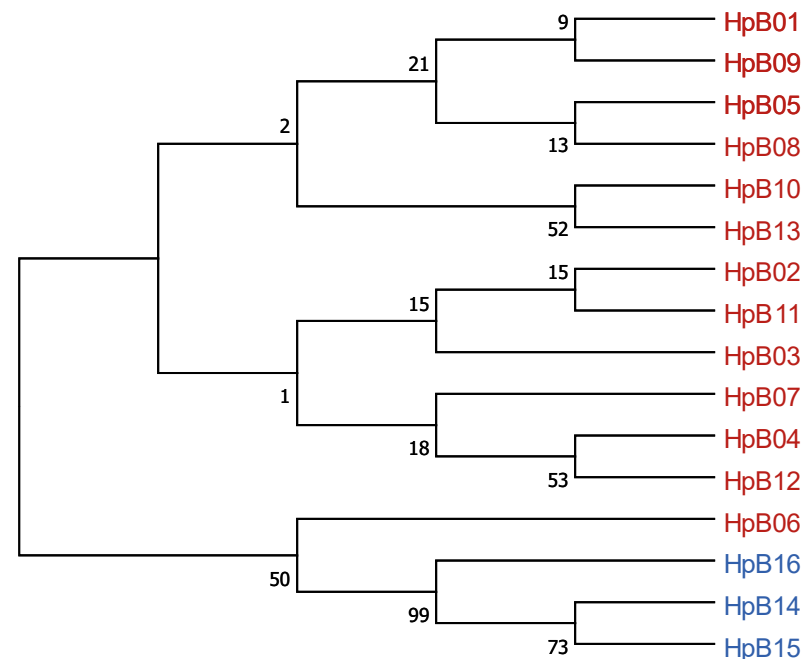
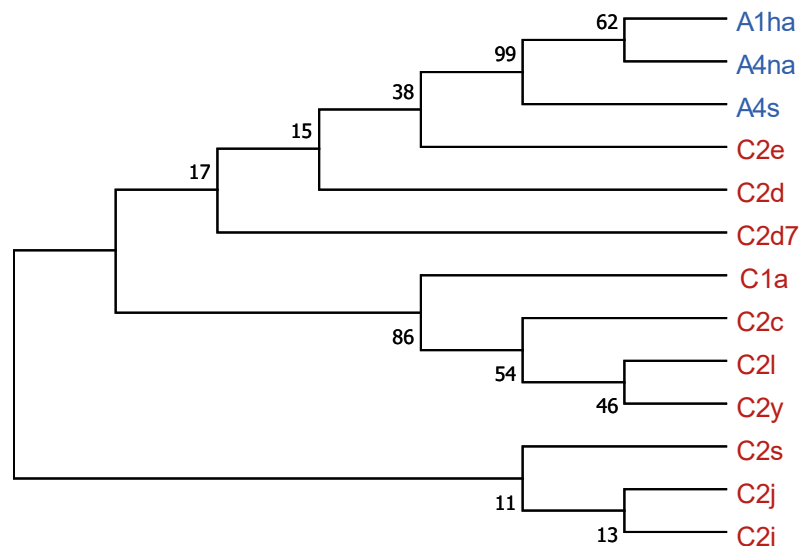
1. MSA

2. Fylogenetická analýza

- Konstrukce fylogenetických stromů pro určení evolučních vztahů mezi haplotypy A, melifera ($tRNA^{leu}$ -cox2 a cox1) byla odvozena pomocí metody Maximum Likelihood a Tamura-Nei modelu v programu Mega X.
- Bootstrapový konsenzuální strom byl odvozen z 10 000 replikátů, Větve odpovídající oddílům reprodukováným v méně než 50 % bootstrapových replikátů byly zhrouceny, Procento replikačních stromů, v nichž se související haplotypy shlukovaly v bootstrapovém testu (10 000 replikátů), je uvedeno vedle větví.
- Počáteční strom pro heuristické vyhledávání byl získán automaticky použitím algoritmů Neighbor-Join a BioNJ na matici párových vzdáleností odhadnutých pomocí modelu Tamura-Nei a následným výběrem topologie s nejvyšší hodnotou logaritmické pravděpodobnosti.



Co-funded by
the European Union



C lineage are in red and A lineage in blue,

- Bootstrap konsenzuální strom haplotypů *tRNA^{leu}-cox2* v České republice, Analýza zahrnovala 13 nukleotidových sekvencí, Do konečného souboru dat bylo zahrnuto celkem 796 pozic,

- The bootstrap consensus tree ***cox1*** haplotypes in the Czech Republic, Analýza zahrnovala 16 nukleotidových sekvencí, A total of 658 positions were included in the final dataset,

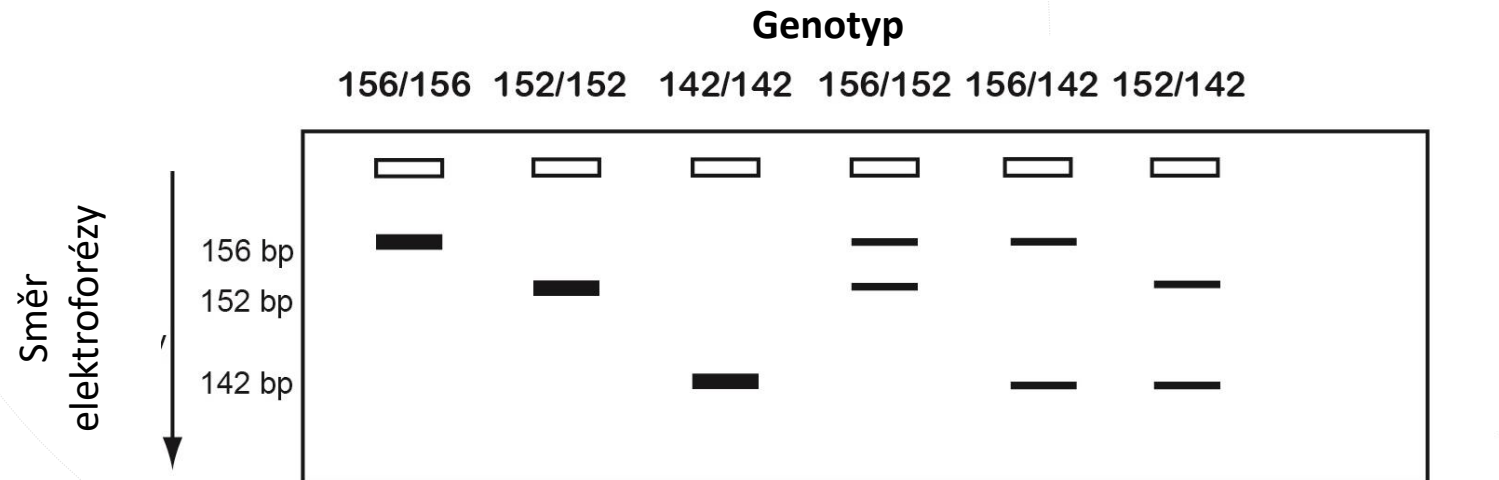
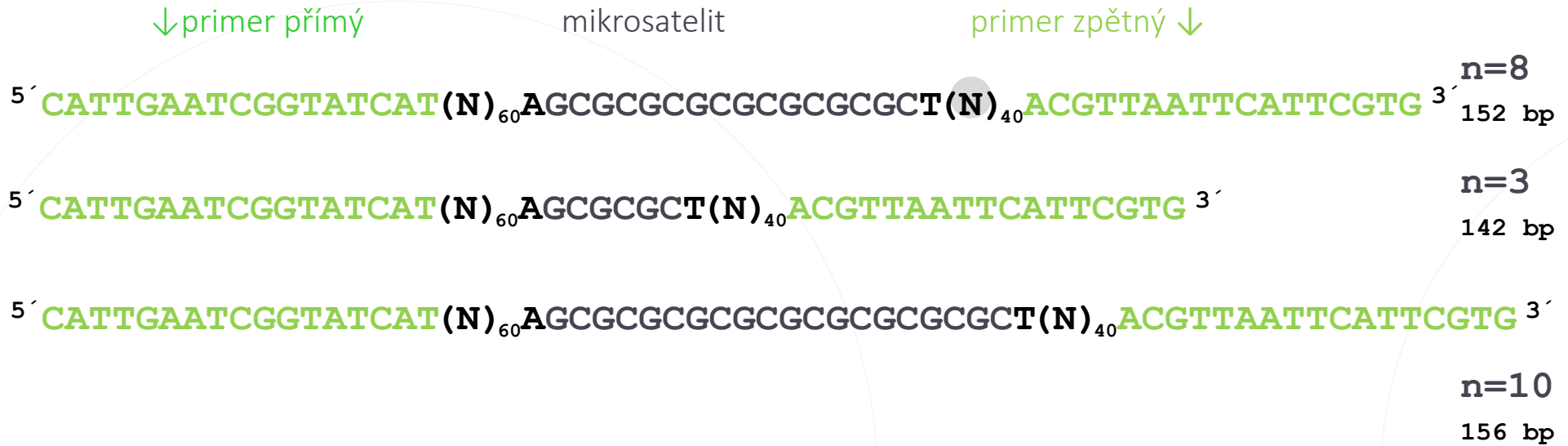
Bootstrapový konsenzuální strom pomocí metody maximální věrohodnosti a modelu Tamura-Nei, vedle větví je procento replikátů, u kterých bylo zjištěno shlukování odpovídajících haplotypů (bootstrapový test - 10 000 replikátů).



Co-funded by
the European Union

Mikrosatelity

(GC)_n
velikost PCR



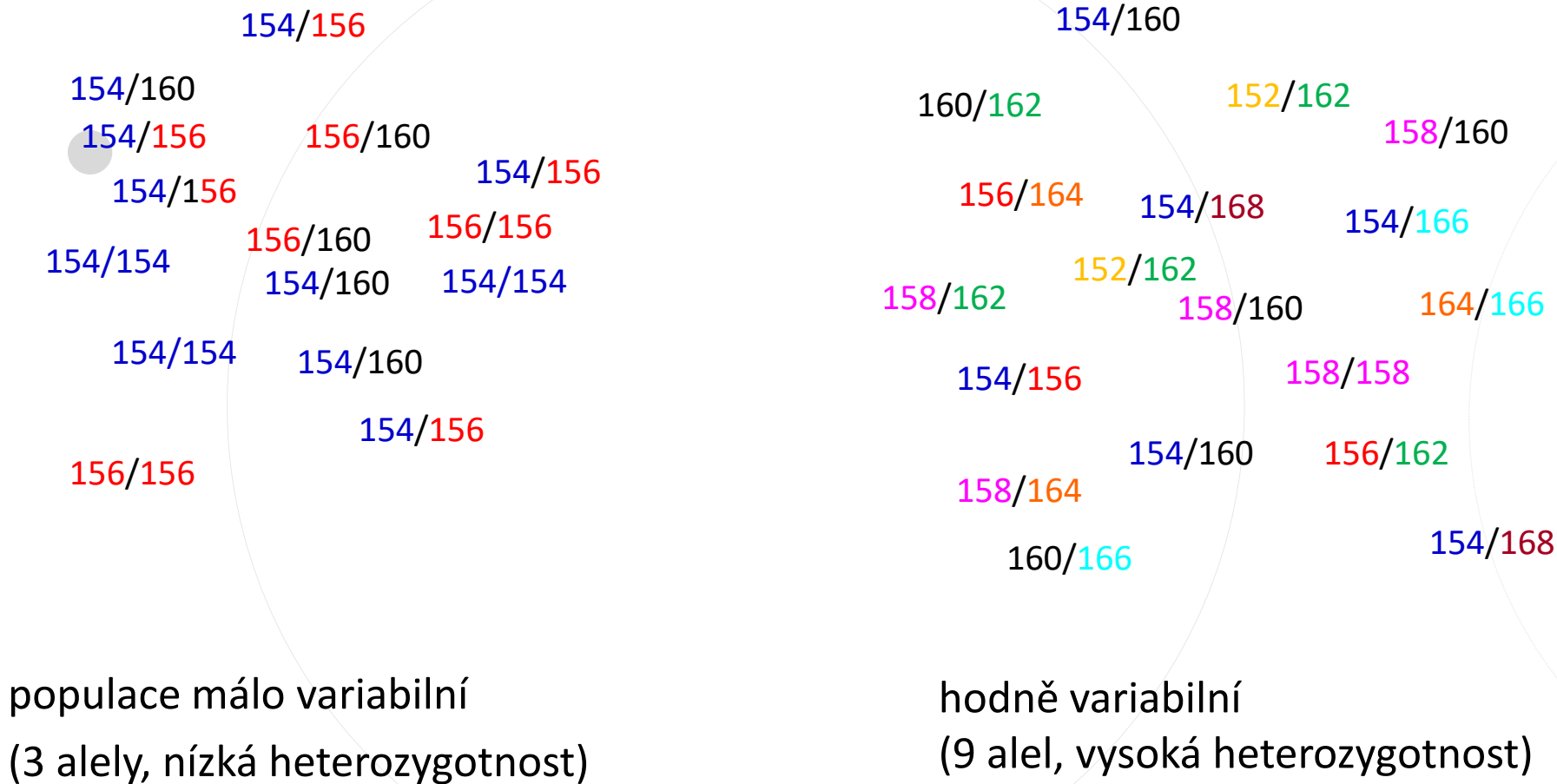
různé kombinace genotypů, které mohou být v populaci s výskytem uvedených 3 alel mikrosatelitu (3 homozygotní, 3 heterozygotní)



Co-funded by
the European Union

Příklad variability ve dvou populacích

ISAGREED



Co-funded by
the European Union

- Mikrosatelity zůstávají cenově dostupným genetickým markerem se schopností zachytit multilokusovou genotypovou informaci, kterou lze následně využít k odhadu genetické diverzity, populační struktury, s využitím v konzervační genetice a šlechtění široké škály organismů.



Co-funded by
the European Union

Fragmentační analýza

Kapilární elektroforéza + fluorescenční detekce (pro přesné určení velikosti fragmentů - alel)

Fragmenty se dělí podle své velikosti v kapiláře, Snímač detekuje průchod molekuly - její barvu a intenzitu signálu v určitém čase - menší fragmenty doputují ke snímači dříve než větší, Zároveň se zkoumanými vzorky jde velikostní standard, software vykreslí kalibrační křivku a podle ní určí velikost i zkoumaných fragmentů,

Přístroj: genetický analyzátor

Analýza fragmentů byla provedena pomocí genetického analyzátoru ABI PRISM 3500 (Applied Biosystems) za standardních podmínek (polymer POP-7, matrice G5), velikost fragmentů byla přesně stanovena pomocí GeneScan™ 600 LIZ™ Size Standard a vyhodnocena pomocí softwaru GeneMapper v 6,0 (Applied Biosystems),



Co-funded by
the European Union

Mikrosatelitní lokusy a skupina multiplexních primerů

Locus	Primer sequence	Length (nt)	Range (bp)	Fluorescent label	Motif	Author
A079	CGAAGGTTGCGGAGTCCTC GTCGTCGGACCGATGCG	19 17	89-124	6-FAM	(CT) ₁₄	Estoup et al, 1994
Ac306	GAATATGCCGCTGCCACC TTTCGTTGCATCCGAGCG	18 18	161-187	6-FAM	(CT) ₁₁	Solignac et al,, 2003
Ap226	AACGGTGTTCGCGAAACG AGCCAACTCGTGCGGTCA	18 18	228-248	6-FAM	(CT) ₈	Solignac et al,, 2003
A007	CCCTTCCTCTTTCATCTTCC GTTAGTGCCCTCCTCTTGC	20 19	94-136	VIC	(CT) ₃ (T) ₇ CTTCG(CT) ₂₄	Estoup et al, 1994
HB-C16-01	AAAATGCGATTCTAATCTGG TTGCCTAAAATGCTTGCTAT	20 20	249-309	VIC	(GA) ₃₅	Shaibi et al,2008
Ap068	TGTCTGCCCTCCTCTCTGTT CACATCGAGCGAGAAGGC	20 18	147-169	NED	(CT) ₁₂ (TA) ₈	Solignac et al,, 2003
A014	GTGTCGCAATCGACGTAACC GTCGATTACCGATCGTGACG	20 20	214-247	NED	(CT) ₁₃ (GGT) ₉	Estoup et al, 1994
Ap223	TCGTACAACGTCGCGCAA GCCGCTCGCCTGTATCTG	18 18	147-182	PET	(T) ₅ (C) ₄ A(T)6(C) ₅	Solignac et al,, 2003

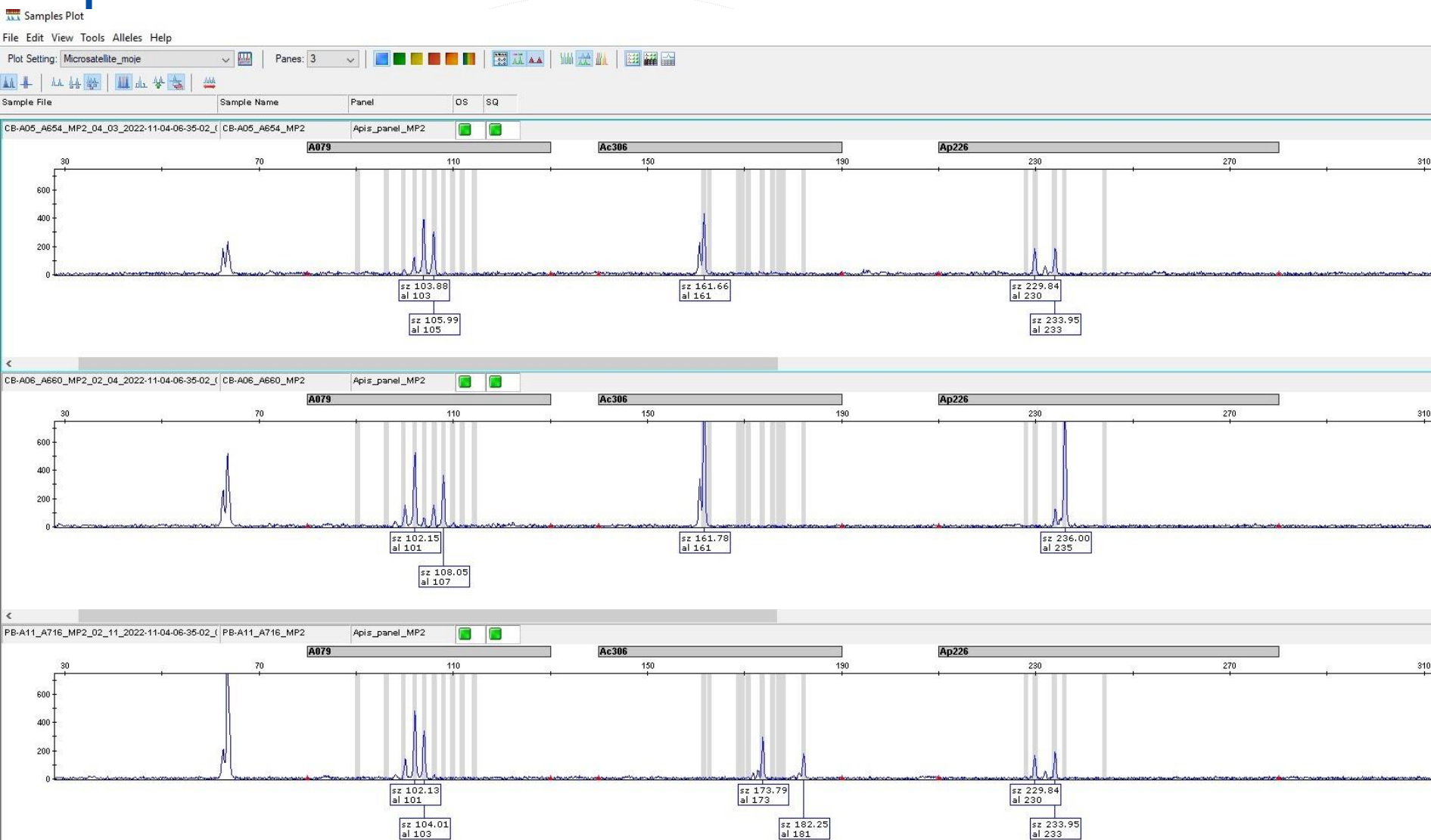
Podmínky cyklování byly následující:
95 ° C (2 min); 30 cyklů: 20 s
denaturace při 95 ° C, 20 s při 57
° C, 30 s elongace při 72 ° C a
závěrečný krok prodloužení při 72
° C po dobu 60 min

Výsledek mikrosatelity

př, MP1



Variabilita - mikrosatelite



Locus	N	Na	Ne	I	Ho	He	uHe	F	HWE
A(B)024	545	4	2,605	1,018	0,591	0,616	0,617	0,041	ns
A088	546	10	2,296	1,174	0,551	0,564	0,565	0,023	ns
Ap043	543	10	3,984	1,593	0,729	0,749	0,750	0,026	ns
Ap113	550	16	1,770	0,934	0,438	0,435	0,435	-0,007	***
Ap218	542	5	2,700	1,060	0,546	0,630	0,630	0,133	***
Ap249	550	8	1,847	0,832	0,467	0,459	0,459	-0,019	ns
A007	548	15	4,374	1,765	0,765	0,771	0,772	0,009	ns
A014	548	22	2,023	1,298	0,500	0,506	0,506	0,011	ns
A079	548	13	6,169	1,992	0,816	0,838	0,839	0,026	ns
Ac306	548	11	1,975	1,076	0,484	0,494	0,494	0,020	ns
Ap068	547	11	3,589	1,514	0,702	0,721	0,722	0,027	ns
Ap223	549	5	1,744	0,796	0,439	0,427	0,427	-0,029	ns
Ap226	549	9	3,093	1,310	0,676	0,677	0,677	0,001	***
HB-C16-01	511	20	10,98 9	2,582	0,738	0,909	0,910	0,188	***
A(B)124	532	17	4,283	1,844	0,774	0,767	0,767	-0,010	ns
AP019	541	7	2,815	1,290	0,647	0,645	0,645	-0,003	ns
Ap273	551	4	1,466	0,624	0,318	0,318	0,318	0,001	ns
Ap289	552	24	1,942	1,317	0,476	0,485	0,485	0,018	ns
HB-C16-05	546	17	2,305	1,379	0,535	0,566	0,567	0,055	ns
A043	541	8	2,250	1,162	0,505	0,556	0,556	0,092	***
Ap049	541	6	1,767	0,926	0,370	0,434	0,434	0,148	***
Ap288	541	4	1,218	0,382	0,174	0,179	0,179	0,031	ns
Mean	544,04	11,18	3,055	1,267	0,556	0,579	0,580	0,036	
SE	1,86	1,29	0,454	0,104	0,035	0,037	0,037	0,012	

- Výpočty byly provedeny v prostředí **GenAlEx** verze 6,5
- Byly vypočteny následující parametry: počet alel (Na), efektivní počet alel (Ne), Shannonův informační index (I), pozorovaná (Ho), očekávaná (He) a nezkreslená očekávaná heterozygotnost (uHe) a index fixace (F)
- Lze využít i program: balíček **diveRsity** v R



Co-funded by
the European Union

- Wrightovy F statistiky (F_{ST} , F_{IS} a F_{IT}) navržené Weirem a Cockerhamem a rozložení genetické diverzity byly analyzovány pomocí analýzy molekulární variance (AMOVA),

Locus	F_{IS}	F_{IT}	F_{ST}
A(B)024	-0,041	0,047	0,085
A088	-0,096	0,000	0,087
AP043	-0,067	0,021	0,083
Ap113	-0,092	-0,005	0,080
Ap218	0,059	0,150	0,097
Ap249	-0,100	-0,011	0,081
A007	-0,067	0,017	0,079
A014	-0,085	0,004	0,082
A079	-0,061	0,030	0,086
Ac306	-0,060	0,024	0,079
Ap068	-0,063	0,028	0,086
Ap223	-0,124	-0,031	0,082
Ap226	-0,110	-0,008	0,092
HB-C16-01	0,082	0,186	0,113
A(B)124	-0,108	-0,018	0,080
AP019	-0,084	0,005	0,083
Ap273	-0,072	-0,007	0,060
Ap289	-0,064	0,019	0,078
HB-C16-05	-0,030	0,058	0,085
A043	0,018	0,110	0,093
Ap049	0,039	0,155	0,121
Ap288	-0,066	0,015	0,076
Mean	-0,054	0,036	0,086
SE	0,012	0,013	0,003

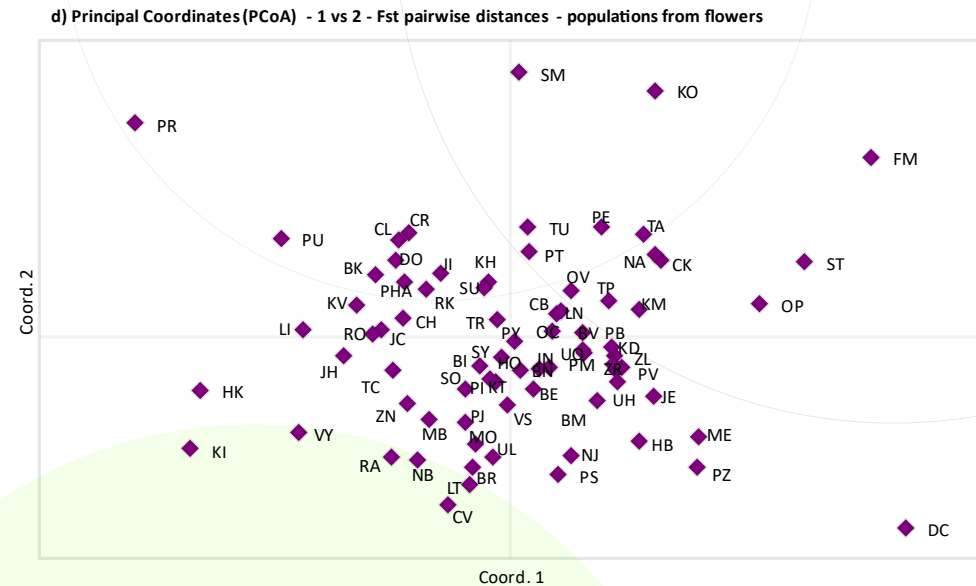
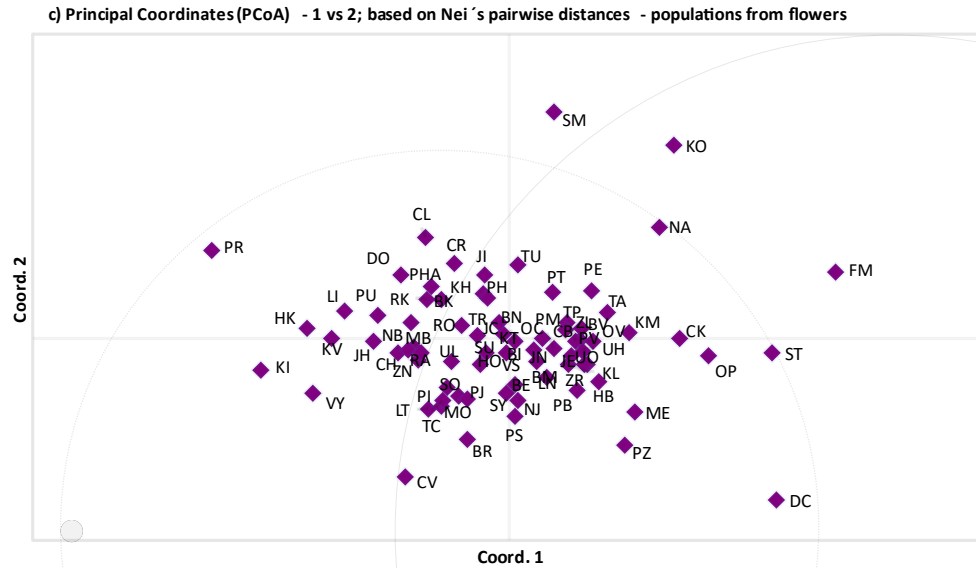
GenAlEx

Source	df	Sum of squares	Mean of squares	Estimated variance	Percentage of variation (%)
Between regional populations	77	619,940	8,157	0,095	1
Among Individuals	476	3231,345	6,789	0,388	6
Within Individuals	553	3325,000	6,013	6,013	93
Total	1105	7176,285		6,496	100



Co-funded by
the European Union

- Byly vypočteny párové Neiovy nestranné a párové F_{ST} genetické distance mezi populacemi a použity pro analýzu hlavních komponent (PCA),

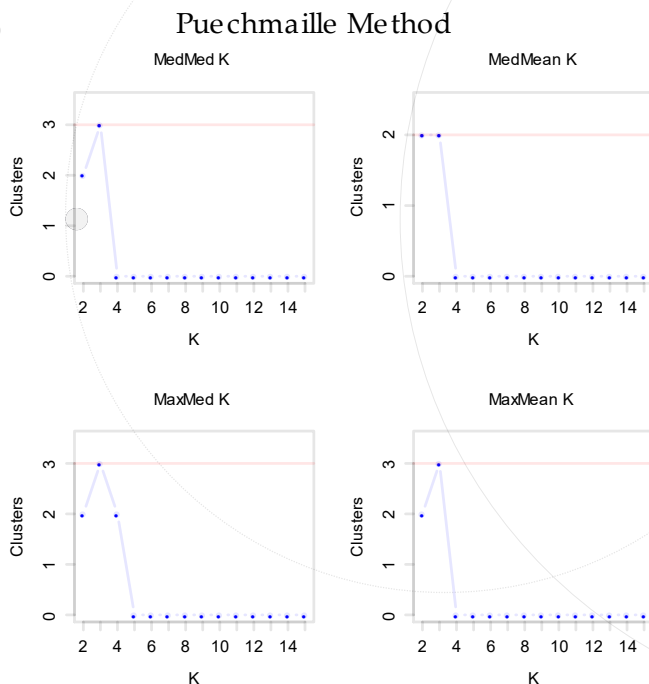


GenAlEx

- K analýze genetické diverzity a míry příměsi populací včely medonosné byla použita bayesovská metoda shlukování **STRUCTURE ver, 2.3.4**. Bylo provedeno deset nezávislých simulací, z nichž každá zahrnovala 10 000 burn-in kroků a následně 100 000 iterací Markovova řetězce Monte Carlo (MCMC). Následně jsme použili programy **Clumpak** a **Structure Selector**, které implementují Evannovu metodu a Puechmailleovu metodu, abychom zjistili **optimální počet shluků (K)**, který nejlépe odpovídá datům (**ΔK , MedMeaK, MaxMeaK, MedMedK a MaxMedK**).

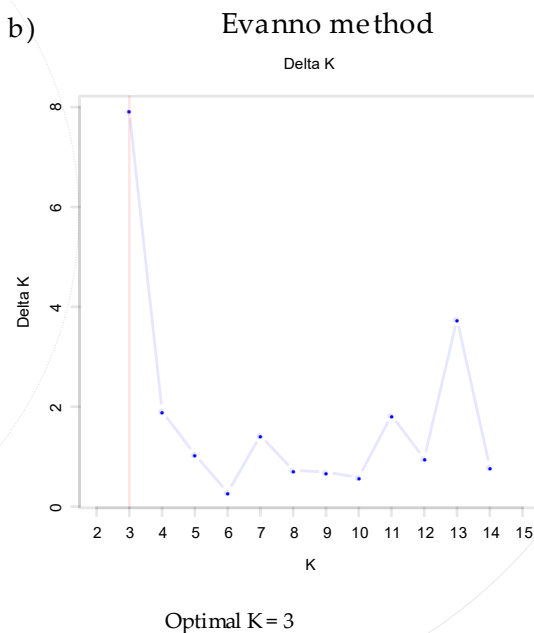
Population from flowers

a)



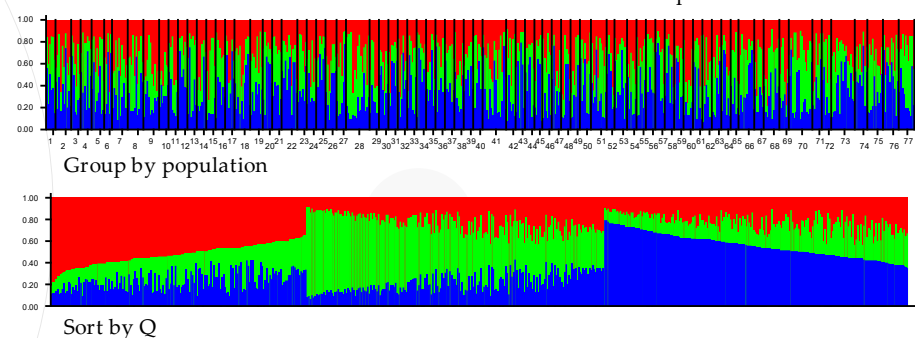
Optimal K after removing spurious clusters (indicated by red lines) are for MedMeaK, MaxMeaK, MedMedK, MaxMedK (Threshold=0.5) = 3

b)



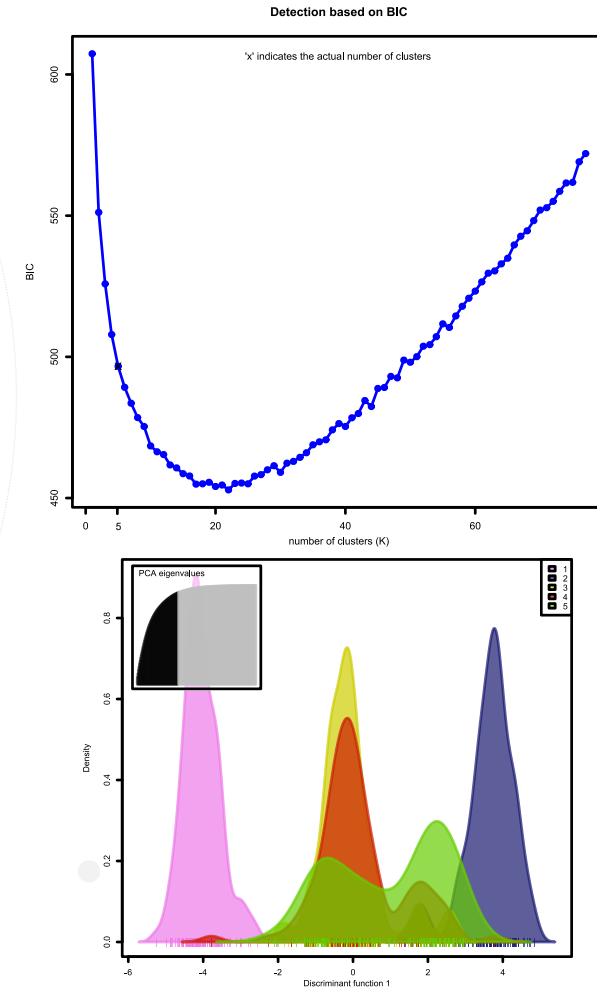
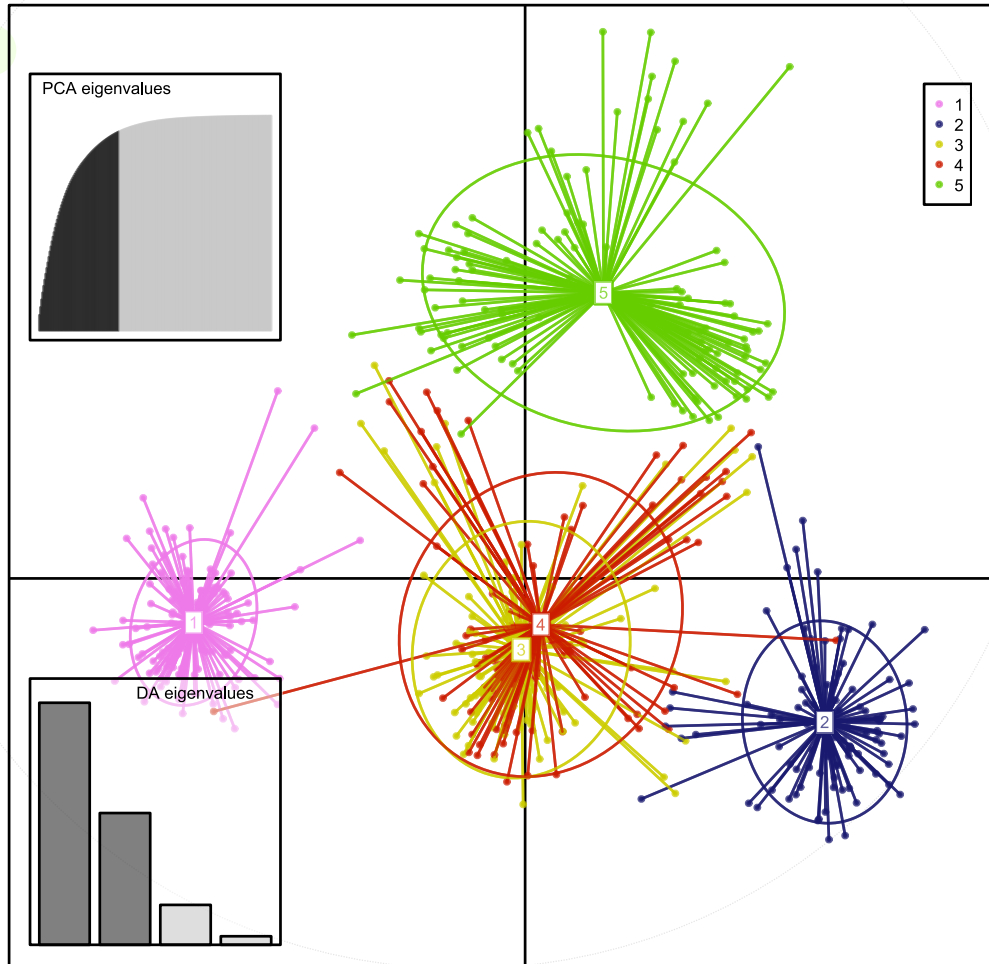
K = 3

Population from flowers



Co-funded by
the European Union

- Pro určení genetické struktury a odvození genetické příměsi byla provedena **diskriminační analýza hlavních komponent (DAPC)** pomocí balíčku **adegenet** R (v rámci programu R).





Siedlce University
of Natural Sciences
and Humanities

Mendel
University
in Brno



Czech University
of Life Sciences Prague

Partners:



Thank you for your attention!

This presentation has been supported by the Erasmus+ KA2 Cooperation Partnerships grant no, 2021-1-SK01-KA220-HED-000032068 "Innovation of the structure and content of study programs in the field of animal genetic and food resources management with the use of digitalisation - Inovácia obsahu a štruktúry študijných programov v oblasti manažmentu živočíšnych genetických a potravinových zdrojov s využitím digitalizácie", The European Commission support for the production of this presentation does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein,



Tomáš Urban



urban@mendelu.cz



Co-funded by
the European Union