

Studie genetické variability populací matizny bahenní – souhrnná zpráva.

RNDr. Miloslav Kitner, Ph.D.
Univerzita Palackého v Olomouci
Přírodovědecká fakulta
Katedra botaniky
Šlechtitelů 27
Olomouc-Holice
78371

Olomouci 30.8.2016

1. Úvod

Tato zpráva shrnuje výstupy molekulárně genetických analýz provedených na základě výsledků veřejné zakázky a její zadávací dokumentace s názvem „Studie genetické variability matizny bahenní a Ověření hybridizace druhů rodu Schoenus“. Popisované analýzy se vztahují k části 1. veřejné zakázky, tj. „Studie genetické variability matizny bahenní“. Projekt je financován za finanční podpory EHP fondů 2009-2014 a Ministerstva životního prostředí na základě „Smlouvy o dílo“ ze dne 26.10.2015 a „Dodatku č. 1 ke smlouvě o dílo ze dne 26. 10. 2015“ ze dne 4.2.2016, které byly uzavřeny mezi sdružením Sagittaria - Sdružení pro ochranu přírody střední Moravy (objednatel), a Univerzitou Palackého v Olomouci (zhotovitel).

Zhotovitel díla se v těchto smlouvách zavazuje provést detailní vyhodnocení genetické variability kultivované moravské populace matizny bahenní v porovnání s populacemi v okolních státech. Jako vstupní materiál pro genetické analýzy měly být použity vzorky rostlin dodané Objednatelem díla, a to alespoň 12 populací z lokalit v ČR, Polska, Maďarska a případně Německa (10 rostlin na populaci) – tj. 120 vzorků. Z metodického hlediska měly být použity dvě metody - metoda AFLP (Vos et al. 1995) a sekvenací vybraných úseků chloroplastové DNA (např. Shaw et al. 2005). Výstupem projektu má být srovnání základních indexů genetické variability analyzovaného souboru položek (polymorfismus, očekávaná heterozygotnost populací, přítomnost privátních markerů, analýza distribuce molekulární variability, vizualizace vzájemných vztahů mezi analyzovanými populacemi a jedinci ve formě fylogenetických stromů /NeighborNetwork diagramů / haplotypových sítí/ Bayesovského klastrování).

Termín zhotovení díla byl stanoven nejpozději do 31.8.2016, ve formě závěrečné zprávy zhotovené ve třech výtiscích (2 pro objednatele, 1 pro zhotovitele díla).

2) Materiál a metody

2a) Komentář k analyzovaným vzorkům

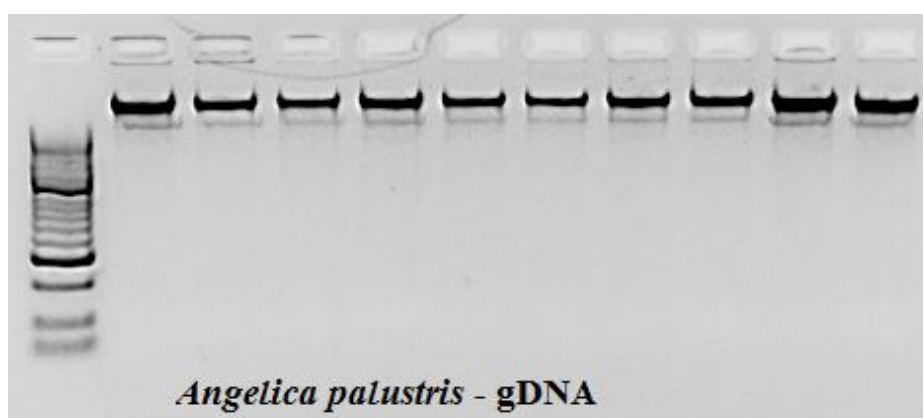
Objednatelem díla dodáno celkem 215 vzorků, ze kterých byly provedeny extrakce DNA (viz. Tabulka č.1). Důvodem extrakce vyššího počtu vzorků oproti smlouvě (max. 120) bylo dodání nekvalitních vzorků z Německa (sušené poškozené listy vykazující vizuální infekci) z lokality Alperstedt a Hassleben na konci vegetační sezóny (21.10.2015). Objednatel díla při dodání těchto vzorků upozornil na nejednoznačnost morfologické determinace rostlin, které vykazovaly znaky matizny bahenní i děhele lesní (*An. sylvestris*). Na jaře 2016 bylo po další konzultaci dosavadních výsledků s Objednatelem díla dohodnuto osekvenování ITS regionu u výběru vzorků pocházejících ze všech lokalit, doplněných o vzorky děhele lesní. Důvodem bylo provedení molekulární identifikace (srovnáním se sekvencemi NCBI, <http://www.ncbi.nlm.nih.gov/>) a vyloučení mezidruhových hybridů matizna x děhel. Označení sedmi referenčních vzorků je doplněno o „x_“, na začátku zkratky (tj. 2 vzorky z ČR, 2 vzorky z německé lokality Rohlbeck, a po jednom vzorku z maďarských lokalit Daru-Rét, Albert-Tó a Falu-Rét). Vzorky z Německých lokalit byly dodány znovu dne 20.5.2016. Jako poslední byly dodány vzorky z Maďarska dne 3.6.2016, přičemž celkový počet dodaných vzorků výrazně převyšoval původně plánovaných 120 jedinců. K tomuto datu již byly zpracovány výsledky sekvenace cpDNA doposud dodaných vzorků. Na základě konzultace postupu dalších analýz bylo mezi Objednatelem a Zhotovitelem dohodnuto, že pro sekvenace cpDNA bude použito po 5 jedincích z Maďarských lokalit, 1-2 vzorky pro sekvenaci ITS a pro AFLP analýzy budou použity pouze vzorky ověřené sekvenováním jako matizna bahenní.

2b) Rostlinný materiál a extrakce genomické DNA

Pro extrakci DNA byly zadavatelem práce dodány listy v relativně čerstvém stavu (cca 1 den po odběru na lokalitě), popřípadě usušené, které byly před dalším zpracováním přeneseny do označených papírových sáčků, ošetřeny tekutým dusíkem a uloženy v hlubokomrazícím boxu (-80°C). Extrakce DNA byla provedena podle standartního CTAB protokolu (Doyle & Doyle 1987) používaného v Laboratoři molekulárních markerů Katedry botaniky PřF UP v Olomouci. Pro posouzení integrity vizolované DNA byly vzorky separovány v 1,5% agarózovém gelu (Obr. 1). Koncentrace DNA byla změřena na UV-spektrofotometru Nanodrop 2000. Po naředění DNA na pracovní koncentraci 50 ng·μl⁻¹ byly vzorky uloženy v hlubokomrazícím boxu (-80°C).

Tabulka č. 1. Přehled použitých vzorků pro extrakce DNA, AFLP a sekvenaci vybraných regionů cpDNA (vzorky děhele lesní jsou označeny „x_“, jejich počty jsou uvedeny v závorce u příslušné analýzy, společně s počtem vzorků matizny).

Stát/lokalita	označení	N_{EX}	N_{cpDNA}	N_{ITS}	N_{AFLP}
Česká republika	CZ, x_CZ	13	13 (11+2)	13 (11+2)	13 (11+2)
Polsko					
Nowosiolki	PL-NO	8	8	5	8
Czesniki	PL-CZ	10	10	8	10
Krakow	PL-KR	6	5	5	5
Kamien	PL-KA	10	10	8	8
Maďarsko					
Júlia Liget	HU-JL	10	5	5	10
Fülöpi	HU-FL	10	5	2	10
Biri Nagy-Rét	HU-BR	10	5	2	10
Daru-Rét	HU-DR, x_HU-DR	11	6 (5+1)	3 (2+1)	11
Bátorliget	HU-BT	10	5	2	10
Teleki Legelo	HU-TL	10	5	2	10
Albert-Tó	HU-AT, x_HU-AT	11	6 (5+1)	3 (2+1)	11
Falu-Rét	HU-FR, x_HU-FR	11	6 (5+1)	3 (2+1)	11
Německo (průměr)					
Alperstedt	DE-AT	31	12 (9+3)	12 (9+3)	10 (9+1)
Hassleben	DE-HA	24	14 (0+14)	19 (0+19)	0
Bad Dürrenberg - Balditz	DE-BD	20	10 (0+10)	1 (0+1)	0
Zwintschöna	DE-ZW	11	10 (0+10)	1 (0+1)	0
Rohlbeck	x_DE-RH	2	2 (0+2)	1 (0+1)	2 (0+2)
Celkem	18	215	137	95	139
plánováno	12	120	120	0	120



Obr. 1. Ukázka výsledků agarózové elektroforézy vzorků vyzolované genomické DNA

2c) sekvenace cpDNA regionů

PCR reakce byly provedeny v 20 µl objemech obsahující 1,5 µl genomické DNA (50ng/µl); po 0,8 µl příslušného primeru (20µM); 2µl pufru dodávaného s polymerázou; 0,38 µl dNTP's (10 mM); 0,08 µl polymerázy KAPA (Kapabiosystems, Boston, USA) a 14,5 µl vody v PCR kvalitě. PCR reakce proběhly v přístroji Eppendorf Mastercycler (Eppendorf, Germany) za použití různých PCR profilů pro každý z analyzovaných regionů, podle původních publikací (Tab. 2). PCR produkty byly po přečištění (GenElute PCR Clean-Up Kit, Sigma-Aldrich, St. Louis, USA) sekvenovány komerční firmou (Macrogen Europe, Amsterdam, the Netherlands).

Tabulka č. 2. Přehled analyzovaných regionů, sekvence použitých primerů a odkaz na příslušný literární zdroj (F, přímý; R, zpětný; I, vnitřní sekvenační primer)

region		přímý	zdroj
<i>trnS-trnG</i>	F	AGATAGGGATTCTGAACCCCTCGGT	Shaw et al., 2005
	R	GTAGCGGGAATCGAACCCGCATC	
	I	TTTTACCACTAACTATAACCCGC	
<i>trnS-trnfM</i>	F	GAGAGAGAGGGATTCTGAACC	Shaw et al., 2005
	R	CATAACCTTGAGGTCACGGG	
<i>rpS16-trnK</i>	F	AAACGATGTGGTARAAAGCAAC	Shaw et al., 2005
	R	AACATCWATTGCAASGATTGATA	
<i>psbD-trnT</i>	F	CTCCGTARCCAGTCATCCATA	Shaw et al. 2007
	R	CCCTTTTAACTCAGTGGTAG	
<i>trnL-trnF</i>	F	ATTTGAACTGGTGACACGAG	Taberlet et al., 1991
	R	GGTTCAAGTCCCTCTATCCC	
<i>ndhJ-trnF-trnL</i>	F	ATGCCYGAAAGTTGGATAGG	Taberlet et al., 1991; Shaw et al., 2007
	R	GGTTCAAGTCCCTCTATCCC	
<i>psbA-trnH</i>	F	GTTATGCATGAACGTAATGCTC	Hamilton, 1999
	R	CGCGCATGGATTACAAATCC	
<i>ITS</i>	F	TCCGTAGGTGAACCTGCGG	White et al. 1990
	R	TCCTCCGCTTATTGATATGC	

2d) AFLP

Na 24 vzorcích bylo provedeno testování funkčnosti metody AFLP na DNA vyizolované pro tuto studii. Zároveň byl proveden výběr primerových kombinací, které byly použity pro analýzu celého datasetu. Finální analýzy byly provedeny pomocí 3 primerových kombinací (Tabulka 2). Pro

provedení metody AFLP byl použit původní protokol podle Vos et al. (1995). Genomická DNA byla štěpena pomocí enzymů *EcoRI* a *MseI* v celkovém objemu 20 µl; následně byla provedena ligace *EcoRI* a *MseI* adapterů k fragmentům DNA získaných v předchozím kroku. Preamplifikace takto získaných konjugátů byla provedena v 20 µl objemu PCR mastermixu. PCR podmínky preamplifikace : 94°C – 3:00; 20 cyklů 94°C (30s), 56°C (60s), 72°C – (1:45); 72°C – 5:00. Preamplifikát DNA byl zředěn v poměru 1:7 a použit jako templát pro selektivní amplifikaci pomocí tří primerových kombinací uvedených v Tabulce 2. Pro fluorescenční značení fragmentů byly použity forward primery s fluorescenčními značkami Well Red (D2, D3, D4) pro sekvenátory Beckman Coulter. PCR podmínky amplifikace byly následující: denaturace 95°C (4:30); 20 cyklů 95°C (30 s), annealing 60,1°C (30 s), extenze 72°C (60 s); následně 1 cyklus s annealingovou teplotou 56°C. Amplifikované fragmenty byly separovány a detekovány na kapilárním sekvenátoru GenomeLab™ GeXP Genetic Analysis System (Beckman Coulter, Inc., Kalifornie, USA).

Tabulka č. 3. Použité primerové kombinace pro preselektivní a selektivní amplifikaci, sekvence použitých primerů

Primerové kombinace pro selektivní amplifikaci	
Set A	D2- <i>EcoRI</i> primer E-ACC / <i>MseI</i> primer M-ACT
Set B	D3- <i>EcoRI</i> primer E-ACT / <i>MseI</i> primer M-AAC
Set C	D4- <i>EcoRI</i> primer E-AGC / <i>MseI</i> primer M-AAT
Sekvence primerů použitých pro preselektivní amplifikaci:	
<i>EcoRI</i> primer: 5'-GACTGCGTACCAATTCA-3'	
<i>MseI</i> primer: 5'-GATGAGTCCTGAGTAAC-3'	

2e) Analýza dat

Elektroforetogramy získané ze sekvenační laboratoře byly zpracovány pomocí programu Geneious 7.1.8 (Biomatters Ltd., Auckland, New Zealand) – t.j. vizuální kontrola, asemblování dílčích čtení do kontigů, tvorba alignmentů a spojování dílčích alignmentů. Základní statistiky polymorfismu (počet SNP, mezer, počet sigletonových a parsimonicky informativních pozic) hodnocených úseků byly vypočítány pomocí programu DnaSP 5.10 (Librado and Rozas 2009). Haplotypová síť byla vytvořena programem TCS (TCS Network, Clement et al. 2002). Pro přípravu vstupního souboru pro TCS a pro identifikaci haplotypů byla použita internetová aplikace FaBox (Villesen 2007).

Detekované AFLP fragmenty o délce 60-550bp byly automaticky vyhodnoceny pomocí programu GenomeLab Genetic Analysis System Software ver.11.0.24 (threshold 100, bin 1) a vyexportovány ve formě binární matice (1 = přítomnost, 0 = nepřítomnost bandu). Tato byla dále použita pro další

analýzy. Pro vizualizaci genetické příbuznosti analyzovaných položek byly použity dva přístupy: a) formou Neighbor-joining dendrogramu (Diceův koeficient podobnosti, 1000 bootstrap replikací) pomocí software FreeTree a FigTree (Pavlíček et al. 1999, FigTree 2015); b) Bayesovským clusterováním pomocí programů STRUCTURE 2.2 (Falush et al. 2007) a GENELAND (<http://www2.imm.dtu.dk/~gigu/Geneland/>).

Pro posouzení genetické diverzity byly vypočítány indexy heterozygotnosti (H_E), polymorfismus (P%) (GENALEX 6 software, Peakall & Smouse 2006), počet privátních bandů (N_{PB} ; markery přítomné pouze u jedinců určité skupiny vzorků, nemusí se nacházet u všech jedinců dané skupiny/populace) a počet privátních fixovaných bandů (N_{FPB} ; markery unikátní pro danou skupinu/populaci, které byly nalezeny u všech jedinců skupiny/populace) (FAMD 1.31, Schlüter and Harris 2006). Zatímco indexy P a H_E vypovídají o celkové variabilitě souboru, indexy N_{PB} a N_{FPB} ukazují na možnou izolovanost hodnocených skupin/populací. Výpočet celkové genetické variability byl proveden v programu AFLP-SURV 1.0. Pro výpočet fixačního indexu (F_{ST}), a analýzy molekulární variance (AMOVA) byl použit program ARLEQUIN 3.5 (Excoffier and Lischer 2010) pro posouzení distribuce celkové genetické variability uvnitř a mezi hodnocenými skupinami vzorků/populací. Hodnoty F_{ST} se pohybují v rozmezí 0-1, přičemž hodnoty blízké 0 ukazují na vysokou podobnost populací danou vysokým počtem sdílených alel. Pro kompletně izolované populace se hodnota F_{ST} blíží 1 (nemají žádné společné alely).

3. Výsledky

3a) Sekvenování

Před provedením sekvenace všech vzorků bylo nutné provést předvýběr vhodných regionů, a to na osmi vzorcích dodaných na podzim 2015 (1 z České republiky, 2 z Německa, 2 z Maďarska, 3 z Polska). Celkem bylo otestováno 5 hypervariabilních regionů cpDNA: *trnS-trnG* intergenic spacer (*trnS-trnG*), *trnS-trnfM*, *rpS16-trnK* (*rpS16*), *psbD-trnT* intergenic spacer (*psbD-trnT*), *ndhJ-trnF-trnL* (*ndhJ-trnL*). Všechny testované regiony obsahovaly velké množství variabilních míst odlišující testované vzorky z Německa od ostatních. Pro analýzu celého sample setu byly vybrány tři nejvíce variabilní regiony, a to *trnS-trnG*, *ndhJ-trnL* a *rpS16*. Tento výběr byl proveden na podzim 2015 a podle dohody s Objednatelem díla byly osekvenovány všechny doposud dodané vzorky. Ze získaných výsledků byla zřejmá výrazná odlišnost části německých vzorků.

Po sloučení dílčích čtení všech tří chloroplastových genů 138 analyzovaných jedinců byl získán 3088 bp (base pairs, nukleotidů) dlouhý úsek (alignment), obsahující 189 variabilních míst (SNP). Sekvenací 84 jedinců byl získán 629 bp dlouhý alignment ITS regionu obsahující 98 SNP. Detailní statistiky jsou pro jednotlivé sekvenované úseky uvedeny v Tabulce č. 4a. Po srovnání získaných ITS sekvencí se

záznamy v nukleotidové databázi NCBI bylo u části vzorků z Německa nalezena shoda s uloženými záznamy odpovídající děhele lesní, stejně jako se sekvencemi referenčních vzorků děhele lesní dodaných zadavatelem. Navíc se ITS sekvence jednoho vzorku z lokality Alperstedt shodovala s deponovanou sekvencí *Aegopodium podagraria* (bršlice kozí noha). Část vzorků *An. sylvestris* a vzorek *Ae. podagraria* byly dále ve studii použity jako srovnávací a jejich variabilita není dále diskutována. Po vyloučení vzorků děhele lesní, měl výsledný alignment cpDNA regionů délku 2927 bp a obsahoval 6 SNP. Zajímavým zjištěním je 18 bp dlouhá delece v regionu *psbD-trnT*, která byla charakteristická pro všechny vzorky z České republiky, a díky které se tyto vzorky výrazně odlišují od ostatních vzorků (Obr. 2). Ve stejném regionu byla nalezena u 4 vzorků záměna guaninu za tymin u čtyř vzorků z Polské lokality Nowosiolki. U tří vzorků ze stejné lokality byla v ITS regionu (624bp) nalezena SNP záměna A/G v pozici 597 alignmentu, která je odlišovala od vzorků z ostatních lokalit a států. Detailní statistiky jsou uvedeny v Tabulce 4b.

Tabulka č. 4. Přehled základních indexů variability sekvenovaných úseků. Část a) udává statistiky celého souboru vzorků, část b) se vztahuje ke statistikám jedinců *A. palustris*

lokus	N	délka (bp)	Nhap	Nsnp	Nss	Npi	gaps (bp)	in/del
a) <i>rpS16-trnK</i>	138	895	9	61	15	46	13 bp	105 bp (4-22-8-6-9-19-4-9-8-11-5)
<i>psbD-trnT</i>	138	1307	8	82	23	59	21 bp	84 bp (10-40-7-7-20bp)
<i>ndhJ-trnL</i>	138	886	6	46	17	29	7 bp	143 bp(5-10-18-25-4-6-47-28)
ITS	95	629	4	98	0	98	12 bp	0 bp
b) <i>rpS16-trnK</i>	94	848	3	4	0	4	0 bp	0 bp
<i>psbD-trnT</i>	94	1233	2	1	0	1	0 bp	18 bp (18)
<i>ndhJ-trnL</i>	94	846	2	1	0	1	0 bp	0 bp
ITS	63	624	2	1	0	1	0 bp	0 bp

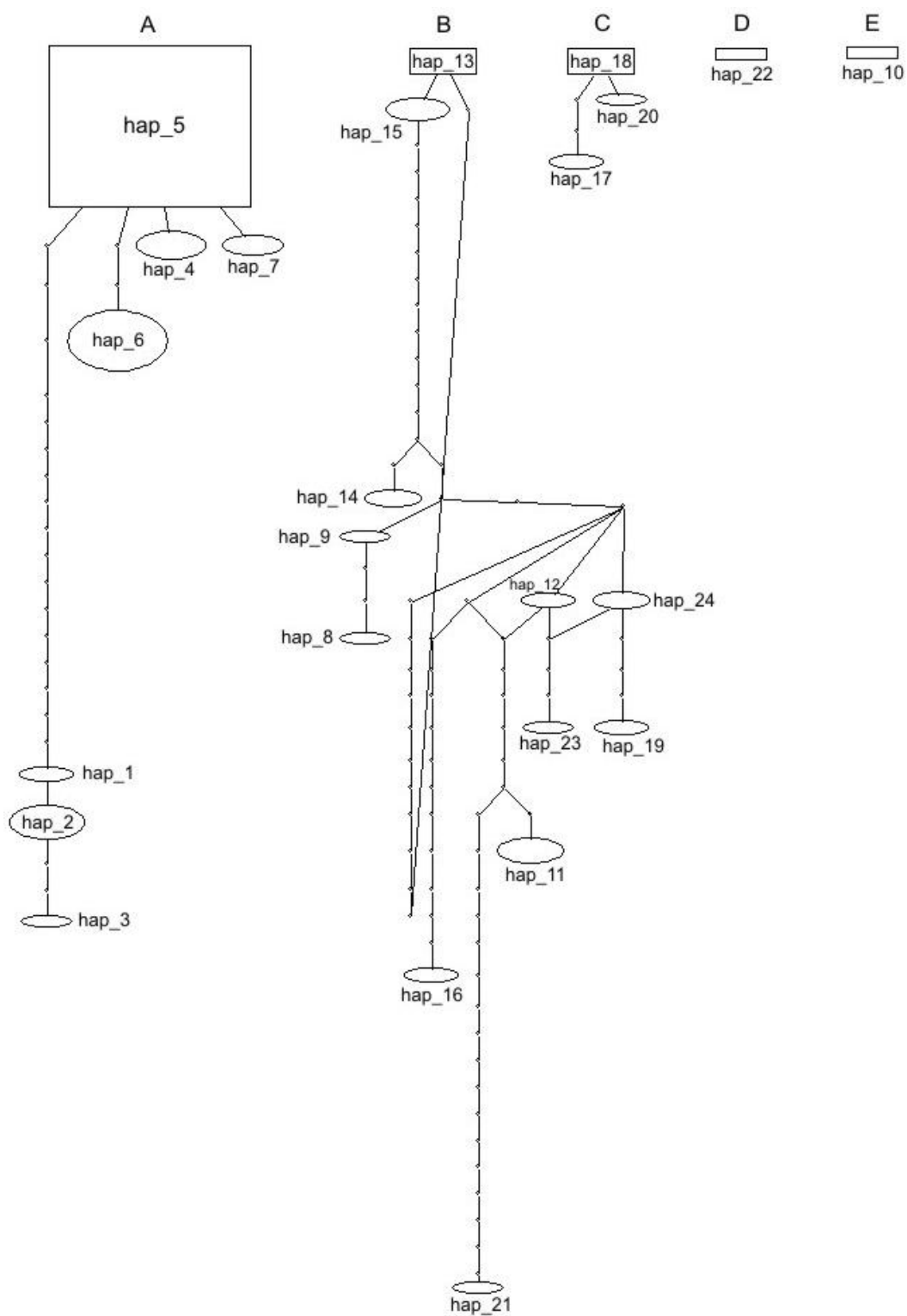
Legenda:
N, počet hodnocených jedinců;
délka, počet bazí daného alignmentu;
Nhap, počet haplotypů;
Nsnp, počet pozic v alignmentu, na kterých se vyskytuje u různých sekvencí více variant různých bazí (tj. 2 a více bazí);
Nss, singleton sites = počet SNP, které se vyskytují pouze u 1 jedince;
Npi, parsimony informative site = počet SNP, které se vyskytují u více než u 2 a více jedinců;
gaps, mezery = počet pozic s chybějící bazí u některé ze sekvencí (zde jsou za mezeru považovány i dvoj- a troj-mezer);
in/del, délka alignmentu s inzercemi nebo delecemi. Udáváme celkový počet chybějících bazí a v závorce délku jednotlivých delecí podle pořadí v alignmentu).

Analýzou spojených chloroplastových regionů bylo identifikováno celkem 24 haplotypů (unikátních sekvencí) (Tabulka 5), které byly programem TCS rozděleny do celkem 5 nezávislých sítí/linií (Obrázek 2). Kromě více či méně komplikovaných sítí B-D tvořené jedinci děhele lesní a linie E (bršlice kozí noha), všech 7 haplotypů matizny bahenní tvořilo haplotypovou síť A. V této síti byl identifikován

haplotyp *hap_5* jako ancestrální (mateřský), ze kterého se postupně odvozovaly evolučně mladší haplotypy tvořící čtyři samostatné, nepropojené linie. Ancestrální haplotyp *hap_5* je představován 53 jedinci původem především z Maďarska (37 jedinců ze všech lokalit), Polska (14 – 4 Nowosiolki, 9 Kamien, 1 Czesniky) a dvěma jedinci z Německa (Alperstedt). Z této skupiny se odštěpuje skupina tří haplotypů charakteristických pouze pro jedince z České republiky (*hap_1*, *hap_2*, *hap_3*); dále linie 18 jedinců haplotypu *hap_6* z Polska (15 – Czesniky 9, Krakow 6) a Maďarska (3, Júlia Liget); linie *hap_4* je charakteristická pro dva jedince z lokality Alperstedt; a konečně linie haplotypu *hap_7* představovaná čtyřmi vzorky z polské lokality Nowosiolki.

Jedinci pocházející z jedné populace mohou nést různé chloroplastové haplotypy. Naopak u některých populací byl nalezen pouze jeden haplotyp – např. *hap_5* u všech lokalit z Maďarska (kromě lokality Júlia Liget) nebo *hap_6* na lokalitě Czesniky (Polsko).

Jako vývojově původní byl identifikován haplotyp *hap_5* charakteristický pro většinu jedinců z Maďarska, polovinu jedinců z Polska (14) a Německa (2). Další nejrozšířenější linií byl haplotyp *hap_6*, detekovaný v druhé polovině vzorků z Polska a 3 vzorků z Maďarska. Samostatnou linii představuje populace *A. palustris* z České republiky, s unikátními haplotypy, které nebyly identifikovány u žádného z jedinců *A. palustris* ze sousedních států. Jako unikátní lze označit i haplotypy *hap_4* (Alperstedt) a *hap_7* (Nowosiolki), avšak jedinci nesoucí tyto haplotypy koexistují na svých lokalitách s jedinci nesoucími haplotypy jiných linií.



Obr. 2: Haplotypové sítě A-E vytvořené na základě sekvenovaných chloroplastových regionů *rpS16-trnK psbD-trnT* a *ndhJ-trnL*. Obdélníky označují ancestrální haplotypy identifikované programem TCS, ze kterých se odvozovaly další mladší linie. Tečky na spojnicích jednotlivých skupin vyjadřují počet rozdílných mutací mezi skupinami. Velikost oválů resp. obdélníků odpovídá počtu jedinců nesoucích daný haplotyp – viz. Tabulka č. 5.

Tabulka č. 5. Detailní přehled rozdělení analyzovaných vzorků podle nalezených haplotypů (taxon, molekulární identifikace sekvenací ITS; MB, matizna bahenní; DL, děhel lesní; bršlice kozí noha).

Haplotyp	Počet jedinců	Délka sekvence (bp)	Jedinci s daným haplotypem	Taxon
<i>hap_1</i>	2	2909	CZ-01,09 (1,16 ndhj)	MB
<i>hap_2</i>	9	2909	CZ-02-08, 10	MB
<i>hap_3</i>	1	2909	CZ-11 (pův 19 - rps16)	MB
<i>hap_4</i>	7	2927	DE-AL-01,02 DE-AL-04-08	MB
<i>hap_5</i>	53	2927	DE-AL-03,38 HU-AT-01-05 HU-BR-01-05 HU-BT-01-05 HU-DR-01-05 HU-FL-01-05 HU-FR-01-05 HU-JL-04,05 HU-TL-01-05 PL-NO-37-39 PL-NO-42 PL-CZ-45 PL-KA-61-66 PL-KA-68-70	MB
<i>hap_6</i>	18	2927	HU-JL-01-03 PL-CZ-46-54 PL-KR-56-60 PL-KA-67	MB
<i>hap_7</i>	4	2927	PL-NO-40,41 PL-NO-43,44	MB
<i>hap_8</i>	1	2909	DE-AL-25	DL
<i>hap_9</i>	1	2909	DE-AL-27	DL
<i>hap_10</i>	1	2845	DE-AL-28	BKN
<i>hap_11</i>	6	2908	DE-BD-01-04 DE-BD-07 DE-BD-10	DL
<i>hap_12</i>	2	2910	DE-BD-05 DE-HA-74	DL
<i>hap_13</i>	3	2917	DE-BD-06 DE-BD-08,09 DE-HA-52,53 DE-HA-56	DL
<i>hap_14</i>	3	2912	DE-HA-46-48 DE-HA-47 DE-HA-48	DL

Tabulka 5. -pokračování

Haplotyp	Počet jedinců	Délka sekvence (bp)	Jedinci s daným haplotypem	Taxon
<i>hap_15</i>	5	2917	DE-HA-49 DE-HA-51 DE-HA-55 DE-HA-61,62	DL
<i>hap_16</i>	2	2917	DE-HA-01,02	DL
<i>hap_17</i>	2	2918	x_DE-RH-01,02	DL
<i>hap_18</i>	6	2920	DE-ZW-01-06	DL
<i>hap_19</i>	2	2910	DE-ZW-07 DE-ZW-10	DL
<i>hap_20</i>	1	2920	DE-ZW-08	DL
<i>hap_21</i>	1	2896	DE-ZW-09	DL
<i>hap_22</i>	1	2867	x_CZ-01	DL
<i>hap_23</i>	1	2911	x_CZ-02	DL
<i>hap_24</i>	3	2909	x_HU-DR x_HU-FR x_HU-AT	DL

3b) Výsledky - AFLP

Výsledky této části studie jsou založeny na hodnocení 132 jedinců *A. palustris*, doplněných o srovnávací skupinu sedmi jedinců *A. sylvestris*, které byly analyzovány pomocí tří AFLP primerových kombinací (Tabulka č. 3). Celkem bylo detekováno 249 bandů, přičemž 235 z nich bylo polymorfních (94,4%).

Základní indexy genetické variability

Zjištěné hodnoty očekávané heterozygotnosti se pohybovaly v rozmezí od 0,126 (Kamien, Polsko) do 0,182 (Nowosiolki, Polsko) (Tabulka č. 6). Nejnižší hodnoty polymorfismu, dalšího ukazatele variability populací, byly zjištěny u polské populace Kamien (36.14%), naopak nevyšší hodnota (58,2%) byla zjištěna u populace z České republiky (s nejvyšším počtem analyzovaných jedinců).

Pro hodnocení rozdělení celkové genetické variability na variabilitu představovanou rozdíly mezi jedinci uvnitř populací a podílu variability na úrovni populací jako takových, byla provedena analýza

molekulární variance (AMOVA). Celková hodnota genetické variability (H_t) je 0,257, která je z 66,74% představována rozdíly mezi jedinci uvnitř populací a z 33,25% vzájemnými rozdíly mezi populacemi jako takovými. Poukazuje to určitý stupeň strukturovanosti populací (odlišnosti), stejně jako hodnota fixačního indexu ($F_{ST} = 0.339$) a přítomnost privátních bandů u několika lokalit (Tabulka č. 6).

Tabulka č. 6. Seznam analyzovaných vzorků s výsledky statistických analýz (n, počet vzorků; n.a., neanalyzováno; N_{polB} počet polymorfních bandů, N_{FB} počet fixovaných bandů N_{PrB} , počet privátních bandů; N_{FPrB} , počet fixovaných privátních bandů; P%, polymorfismus; H_E , očekávaná heterozygotnost; SE, směrodatná odchylka)

Stát	Lokalita	Pop	n	H_E	$He(SE)$	%P	N_{PB}	N_{FB}	N_{PrB}	N_{FPrB}
Česko	Hrdibořice	CZE	11	0,169	0,012	58,23%	145	47	0	0
Německo	Alperstedt	DE-ALP	9	0,129	0,012	37,35%	93	84	1	0
Maďarsko	Júlia Liget	HU-JL	10	0,168	0,012	52,21%	130	48	0	0
	Fülöpi	HU-FL	10	0,168	0,012	52,21%	130	51	1	0
	Biri Nagy Rét	HU-BR	10	0,135	0,012	38,15%	95	84	0	0
	Daru-Rét	HU-DR	10	0,181	0,013	52,21%	130	56	0	0
	Bátorliget	HU-BT	10	0,144	0,012	42,97%	107	81	0	0
	Teleki Legelo	HU-TL	10	0,127	0,012	37,35%	93	70	0	0
	Albert - Tó	HU-AT	10	0,172	0,013	49,00%	122	54	0	0
	Falu-Rét	HU-FR	10	0,160	0,012	48,59%	121	75	0	0
	Nowosiolki	PL-NO	8	0,182	0,013	51,41%	128	77	1	0
Polsko	Czesnik	PL-CZ	10	0,146	0,012	49,00%	122	55	1	0
	Kraków	PL-KR	5	0,133	0,012	37,35%	93	77	0	0
	Kamien	PL-KA	9	0,126	0,012	36,14%	90	60	1	1

Populačně genetická struktura

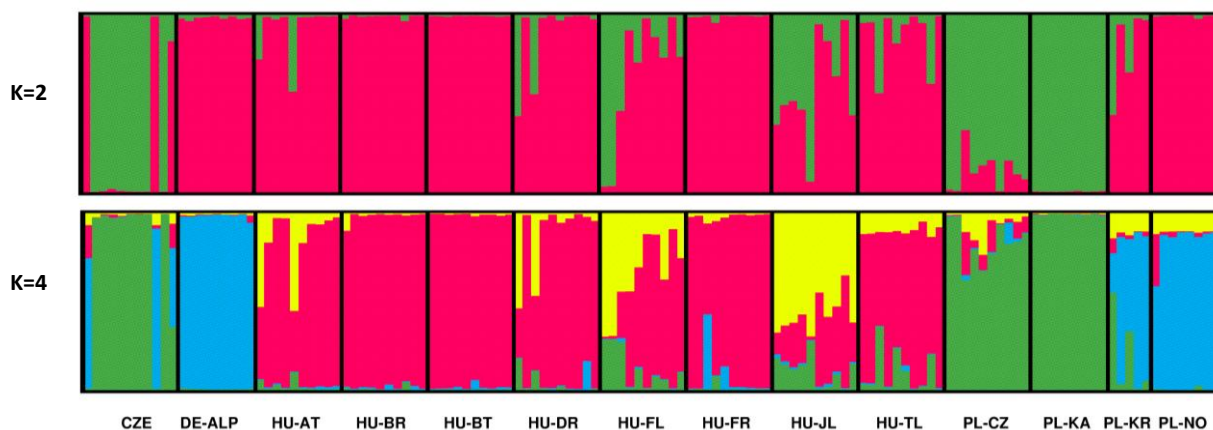
Vizualizace vzájemných vztahů mezi analyzovanými vzorky ve formě Neighbor-Joining (NJ) dendrogramu rozdělila dataset do několika skupin, ale bez statisticky významné podpory (Obr. 3). Kromě samostatné skupiny vzorků děhele (Skupina O) jsou vzorky matizny bahenní rozděleny do tří hlavních skupin A-C. Vzorky původem z Maďarska jsou seskupeny ve skupinách A a B. Součástí skupiny B jsou i vzorky z polské lokality Nowosiolki tvořící samostatnou podskupinu, stejně jako samostatná podskupina jedinců z německé lokality Alperstedt a samostatně stojící tři vzorky z ČR. Konečně poslední skupinu C tvoří vzorky ze zbytku polských lokalit ve více či méně oddělených podskupinách a samostatná větev zbývajících sedmi vzorků z ČR.



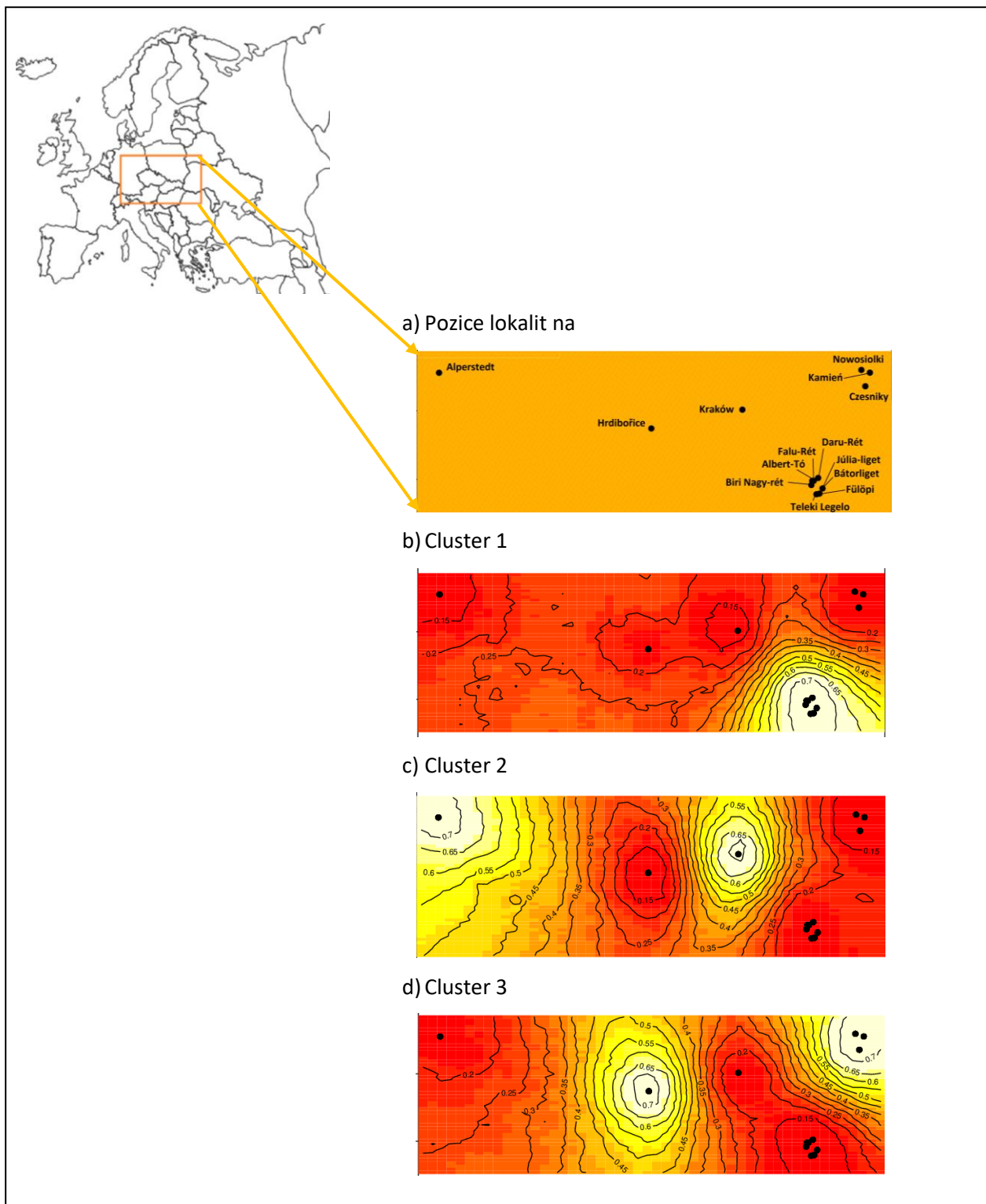
Obr. 4. Neighbour-joining dendrogram vygenerovaný pomocí 249 AFLP markerů na souboru 139 vzorků *A. palustris* a 7 vzorků *A. sylvestris* ze čtyř Evropských států.

Při Bayesovském hodnocení vztahů mezi jednotlivými populacemi programem STRUCTURE byly nalezeny dva způsoby nejvíce pravděpodobného rozdělení vzorků, a to do dvou ($K = 2$; $\Delta K = 14.01$) nebo do čtyř skupin ($K = 4$; $\Delta K = 13.36$) (Obr. 5).

Při rozdělení vzorků do dvou skupin ($K = 2$, Obr. 3) je jednak zřejmá shoda profilů populace z ČR s polskými populacemi Czesniki a Nowosiolki (zelený signál), a dále významný příspěvek tohoto signálu u maďarských lokalit Júlia-Liget, Fülöpi, Teleki Legelo. Při navrženém modelu rozdělení do čtyř skupin ($K=4$, Obr. 3) je zachován společný základ populací ČR a lokalit Czesniki a Nowosiolki (zelený signál), dále se ukazuje podobnost populací Alperstedt (Německo) s populacemi Krakow a Nowosiolki (převaha modrého signálu). U dvou jedinců z ČR byl při tomto modelu identifikován prakticky stejný profil jako u jedinců z populací Krakow a Nowosiolki. U vzorků z maďarských lokalit převažuje červený signál, který je u populací Júlia Liget, Fülöpi a Teleki Legelo doplněn o žlutý signál, přítomný i v profilech polských populací Krakow a Nowosiolki.



Obr. 5. Výsledek Bayesovského clusterování programem STRUCTURE, rozdělující 139 vzorků do dvou ($K=2$) a čtyř ($K=4$) skupin. Každý vertikální sloupec představuje jednoho jedince, přičemž barevný profil jedince odpovídá pravděpodobnému zařazení daného jedince do jedné nebo více skupin. Geografický původ analyzovaných vzorků je uveden pod diagramem ve formě zkratky odpovídající označení v Tabulce č.1.



Obr. 6. Výsledek Bayesovského clusterování programem GENELAND. Pravděpodobnost rozdělení analyzovaných populací do stejných skupin. S přechodem barvy červené do žluté se zvyšuje pravděpodobnost zařazení lokalit do stejné skupiny.

a) **Komentář k výsledkové části a závěr**

Bayesovské rozdělování dat do skupin provádí program GENELAND., přičemž v analýze zohledňuje vzájemnou geografickou polohu studovaných jedinců a populací. Výstup programu zobrazuje mapy „posterior probabilities“ příslušnosti do jednotlivých skupin. Při této simulaci program přiřadil všechny populace z Maďarska do samostatného klastru (Obr. 6a), populace Alperstedt a Krakow do druhého klastru (Obr. 6b) a do samostatného klastru c) zbývající polské populace a populaci z ČR.

Závěr

- Sekvenací vybraných chloroplastových regionů a metodou AFLP byla provedena analýza genetické variability a vzájemných vztahů populací matizny bahenní z České republiky, Maďarska, Polska a Německa.
- Z dodaného rostlinného materiálu byla provedena extrakce genomické DNA, ověřena její integrita a koncentrace.
- Z celkem 215 extrahovaných vzorků bylo analyzováno 137 vzorků sekvenací tří regionů cpDNA a 139 metodou AFLP. Oproti požadavku uvedených ve smlouvě o dílo byla navíc u 95 vzorků provedena sekvenace ITS regionu pro ověření morfologické identifikace dodaných rostlin Objednatelem díla.
- Kombinací výsledků všech molekulárních analýz bylo zjištěno, že se jedná o 134 jedinců *An. palustris*, 43 *An. sylvestris* a 1 jedince *Aegopodium podagraria*. Zbývající vzorky nebyly použity, protože u nich neproběhla PCR reakce.
- Celkový počet analyzovaných populací byl 18, přičemž se jednalo o 14 populací *An. palustris* a 4 populace *An. sylvestris*. Pro výpočet hodnot genetické variability populací (metoda AFLP) byl dodržen požadavek min. 10 analyzovaných jedinců na populaci, s výjimkou dvou populací z Polska (5 a 8 vzorků) a německé populace Alperstedt (9 vzorků), z důvodu malého počtu nalezených rostlin na lokalitě.
- U analyzovaných populací byly provedeny výpočty základních indexů genetické variability a vizualizovány vztahy mezi analyzovanými jedinci ve formě haplotypových sítí, fylogenetického stromu a Bayesovského klastrování. Vizualizace ve formě NeighborNetworku neposkytla další informaci k ostatním použitým metodám a byla ze studie vypuštěna.

- Sekvenační data ukazují na existenci pěti chloroplastových linií *A. palustris*, které nejsou vázány na konkrétní lokalitu nebo stát. Výjimkou je populace z České republiky nesoucí unikátní haplotypy, které nebyly identifikovány u žádného z jedinců *A. palustris* ze sousedních států. V rámci jedné populace byli detekováni jedinci nesoucí buď pouze jeden haplotyp nebo maximálně dva různé haplotypy. Výjimkou je opět populace ze záchranné kultivace v České republice, kde byly nalezeny tři. Z výsledků lze usuzovat na vzájemné propojení (gen-flow) Evropských populací matizny bahenní demonstrované přítomností dvou převažujících haplotypů (*hap_5* a *hap_6*) v různých populacích napříč Evropou, které postupně vytlačují minoritní haplotypy z okrajových populací se kterými jsou v kontaktu – tj. haplotyp *hap_4* na lokalitě Alperstedt nebo *hap_7* převažující na lokalitě Nowosiolki. Tato lokalitě leží blízko hranic Polska s Ukrajinou, takže nelze vyloučit, že *hap_7* představuje linii charakteristickou pro Ukrajinu (a popřípadě další pobaltské státy).
- Na základě AFLP dat byly vypočteny hodnoty očekávané heterozygotnosti, které lze označit za středně vysoké a odráží převahu pohlavního rozmnožování v populacích.
- Vizualizace vzájemných vztahů mezi analyzovanými jedinci pomocí Neighbor-Joining dendrogramu (AFLP data) poskytla nejednoznačné rozdělení vzorků do jasně definovaných skupin podle jejich geografického původu. V zásadě podporuje výsledky sekvenačních dat v identifikaci několika linií, které jsou v analyzovaných lokalitách promíchány. Populace Alperdstadt a Nowosiolki se shlukují ve společném sub-klastru, stejně jako na výstupu Bayesovského shlukování. Nejednoznačně působí postavení populace Krakow, pokud srovnáváme výstupy ze všech použitých programů. Na dendrogramu je populace Krakow umístěna vedle většiny vzorků z ČR a polských populací Kamien a Czesniki, zatímco Structure identifikuje podobný profil s populacemi Nowosiolki a Alperstedt a třemi vzorky z ČR. Naopak program Geneland identifikuje populaci Krakow ve stejném clusteru jako populaci Nowosiolki. Tyto dílčí rozdíly ve shlukování mohou ukazovat na společný původ populací z ČR, Německa a Polska, oproti populacím z Maďarska.

Summary

- In this study, we used sequencing of selected chloroplast regions and AFLP markers to describe genetic variability and relationship among populations of *Angelica palustris* originated from the Czech Republic, Hungary, Poland and Germany.
- Genomic DNA was extracted from plant material and its integrity and concentration was verified.

- From the total of 215 extracted samples, 137 samples were used for sequencing and 139 for AFLP's. Contrary to requirements specified in the contract, sequencing of nuclear ITS region was performed (95 samples) to verify morphological determination of samples provided by the Client.
- The result of both molecular methods have identified 134 *Angelica palustris*, 43 *Angelica sylvestris* and one *Aegopodium podagraria*. The PCR reactions performed on remaining samples were unsuccessful.
- The total number of analysed populations was 18, 14 populations of *An. palustris* and 4 *An. sylvestris* populations. To calculate population genetic variability indices using AFLP method, we used 10 analyzed individuals per population (as required in contract), with the exception of two populations from Poland (5 and 8 samples) and Alperstedt (9 samples) from Germany because of the small number of plants found at particular site.
- We performed calculations of basic genetic variability indices, visualisations of relationships among analysed individuals by means of haplotype network, phylogenetic tree, and Bayesian clustering. Visualisation by NeighborNetwork provided no additional information to other methods and was excluded from the study.
- Sequencing data showed the existence of five chloroplast lines of *A. palustris*, which are not linked to a specific location nor state. The exception is the Czech population with unique haplotypes that were not identified in any of *A. palustris* individuals from neighboring countries. Within one population, we identified individuals carrying either only one haplotype or a maximum of two different haplotypes. There is single exception again, the Czech population, where three haplotypes have been found. Based on acquired results we can deduce on the interconnection (gene-flow) of European populations of *A. palustris* which is demonstrated by the presence of two predominant haplotypes (*hap_5* and *hap_6*) scattered in several populations across Europe, which gradually displace the minor haplotypes from peripheral populations - i.e. haplotype *hap_4* (Alperstedt) or *hap_7* prevailing in the locality Nowosiółki. This locality is situated near the border with Ukraine, so it can not be ruled out that *hap_7* line represents a characteristic for Ukraine (and possibly other Baltic States).
- Based on the AFLP data values of expected heterozygosity were calculated, which can be described as medium-high and reflects the predominance of sexual reproduction in populations.
- Visualization of relationships between analysed individuals on Neighbor-Joining dendrogram (AFLP data) showed ambiguous clustering of samples into clearly defined groups according to their geographic origin. In principle it is in line with the sequencing data identifying several lines,

being mixed in analyzed populations. Individuals from Alperdstadt and Nowosiółki populations fell into single sub-cluster on Neighbor-Joining dendrogram as well as on Bayesian plots. Ambiguous character of the population Krakow is evident when comparing outputs of all software used. On the dendrogram, population Krakow is located next to the majority of samples from the Czech Republic, and Polish populations Kamien and Czesniki, while Structure identifies a similar profile with populations Nowosiółki and Alperstedt and three samples from the Czech Republic. Conversely Geneland program assigns Krakow population in the same cluster as the population Nowosiółki. These minor differences in clustering may indicate a common ancestry of populations from the Czech Republic, Germany and Poland, contrary to populations from Hungary.

Literatura

- Clement M, Snell Q, Walke P, Posada D, Crandall, K (2002). TCS: estimating gene genealogies. *Proc 16th Int Parallel Distrib Process Symp* 2: 184.
- Doyle J.J. & Doyle J.L. (1987): A rapid DNA isolation procedure for small quantities of fresh leaf tissue. – *Phytochem. Bull.* 19: 11–15.
- Excoffier L, Lischer HEL (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*. 10: 564-567.
- Falush D, Stephens M, Pritchard JK (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. – *Mol. Ecol. Notes* 7: 574–578.
- Hamilton M (1999). Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Mol. Ecol.* 8: 521-523.
- Librado P, Rozas J (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.
- Pavlíček BA, Hrda S, Flegr J (1999) FreeTree - Freeware program for construction of phylogenetic trees on the basis of distance data and bootstrap/jackknife analysis of the tree robustness, application in the RAPD analysis of the genus *Frenkelia*. – *Folia Biologica* 45: 97–99.
- Peakall R, Smouse PE (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. – *Mol. Ecol. Notes* 6: 288–295.
- Schlüter PM, Harris SA (2006). Analysis of multilocus fingerprinting data sets containing missing data. *Mol. Ecol. Notes* 6, 569–572.
- Shaw J, Lickey EB, Beck JT, Farmer SB, et al. (2005). The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am. J. Bot.* 92: 142-166.

- Shaw J, Lickey EB, Schilling EE, Small RL (2007). Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *Am. J. Bot.* 94: 275-288.
- Taberlet P, Gielly L, Patou G, Bouvet J (1991). Universal primers for amplification of three noncoding regions of chloroplast DNA. *Pl. Mol. Biol.* 17: 1105-1109.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013). MEGA6. Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*. 30: 2725-2729.
- Villesen P (2007). FaBox: an online toolbox for fasta sequences. *Mol. Ecol. Notes* 7: 965–968.
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucl Acids Res* 23: 4407–4414.
- White TJ, Bruns T, Lee S, Taylor JW (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pp. 315-322 In: PCR Protocols: A Guide to Methods and Applications, eds. Innis, M. A., D. H. Gelfand, J. J. Sninsky, and T. J. White. Academic Press, Inc., New York.