

Využití molekulárních markerů v systematice a populační biologii rostlin

5. AFLP

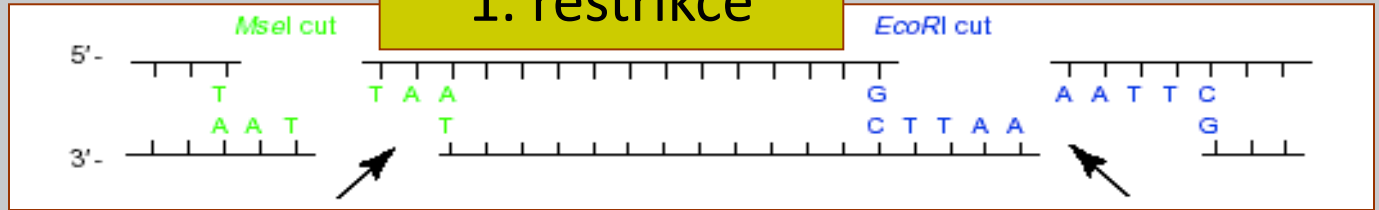
Amplified Fragment Length Polymorphism

- polymorfismus v délce amplifikovaných fragmentů
- základní princip metody
 - restrikce celkové DNA
 - selektivní PCR amplifikace jen některých fragmentů

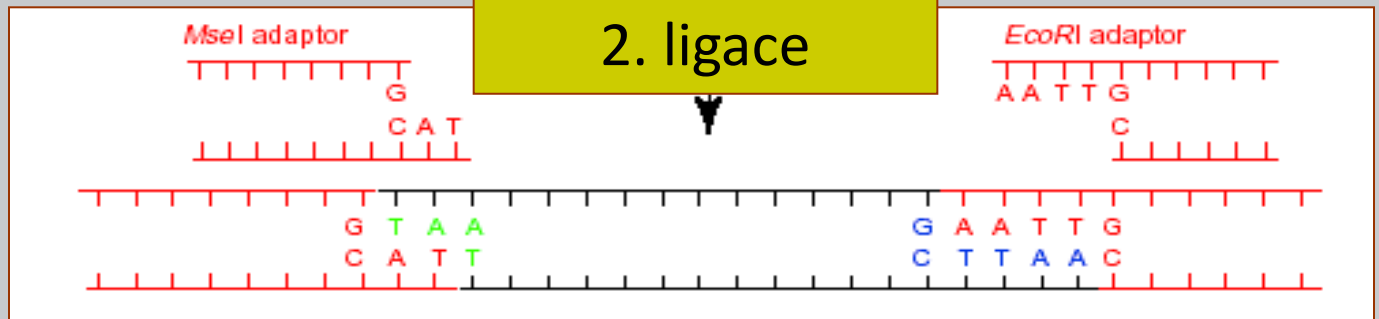


celková DNA

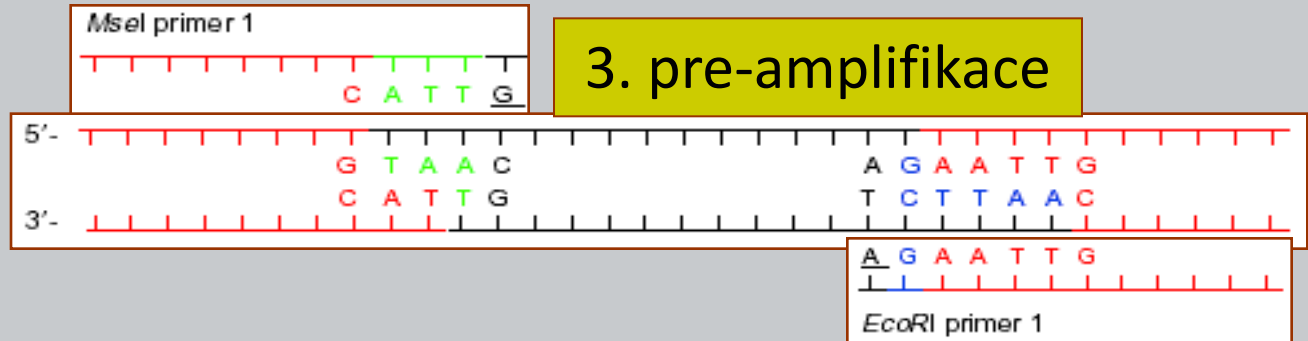
1. restrikce



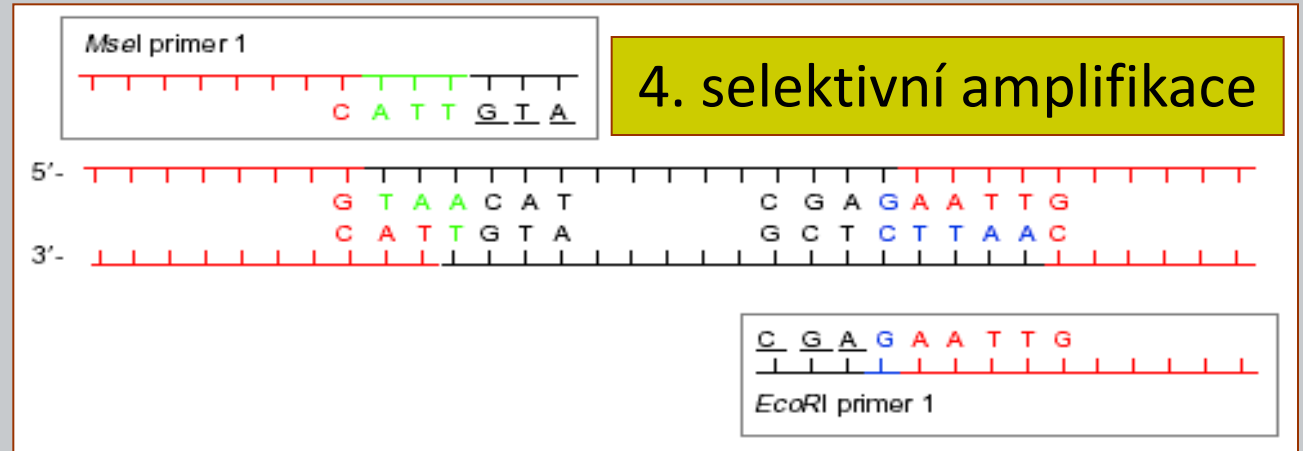
2. ligace



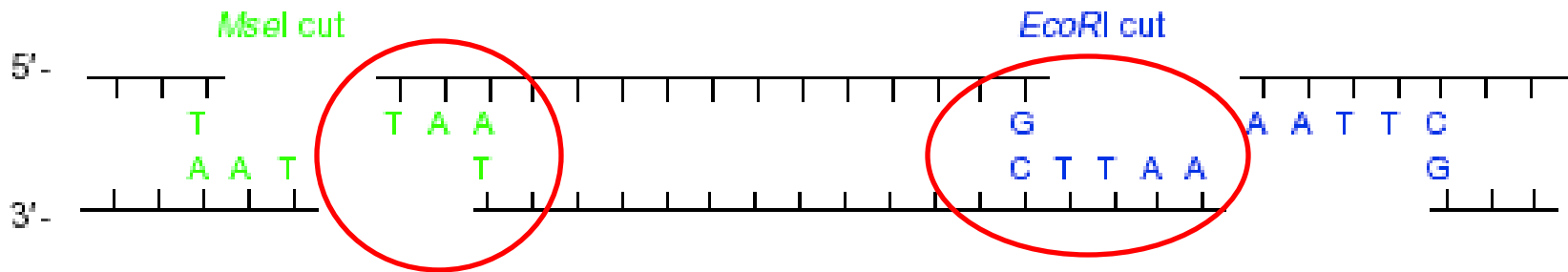
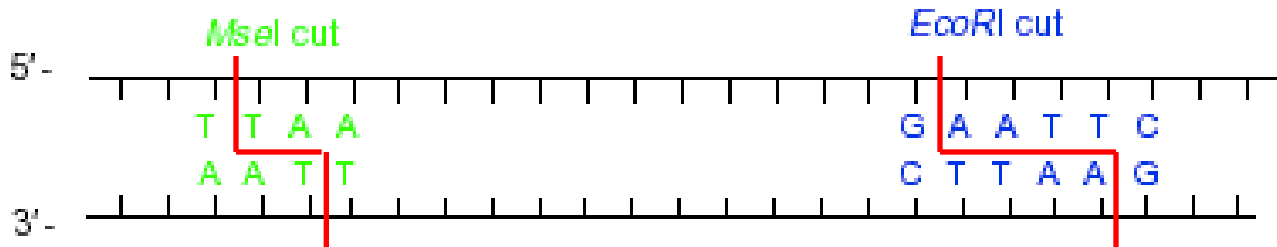
3. pre-amplifikace



4. selektivní amplifikace

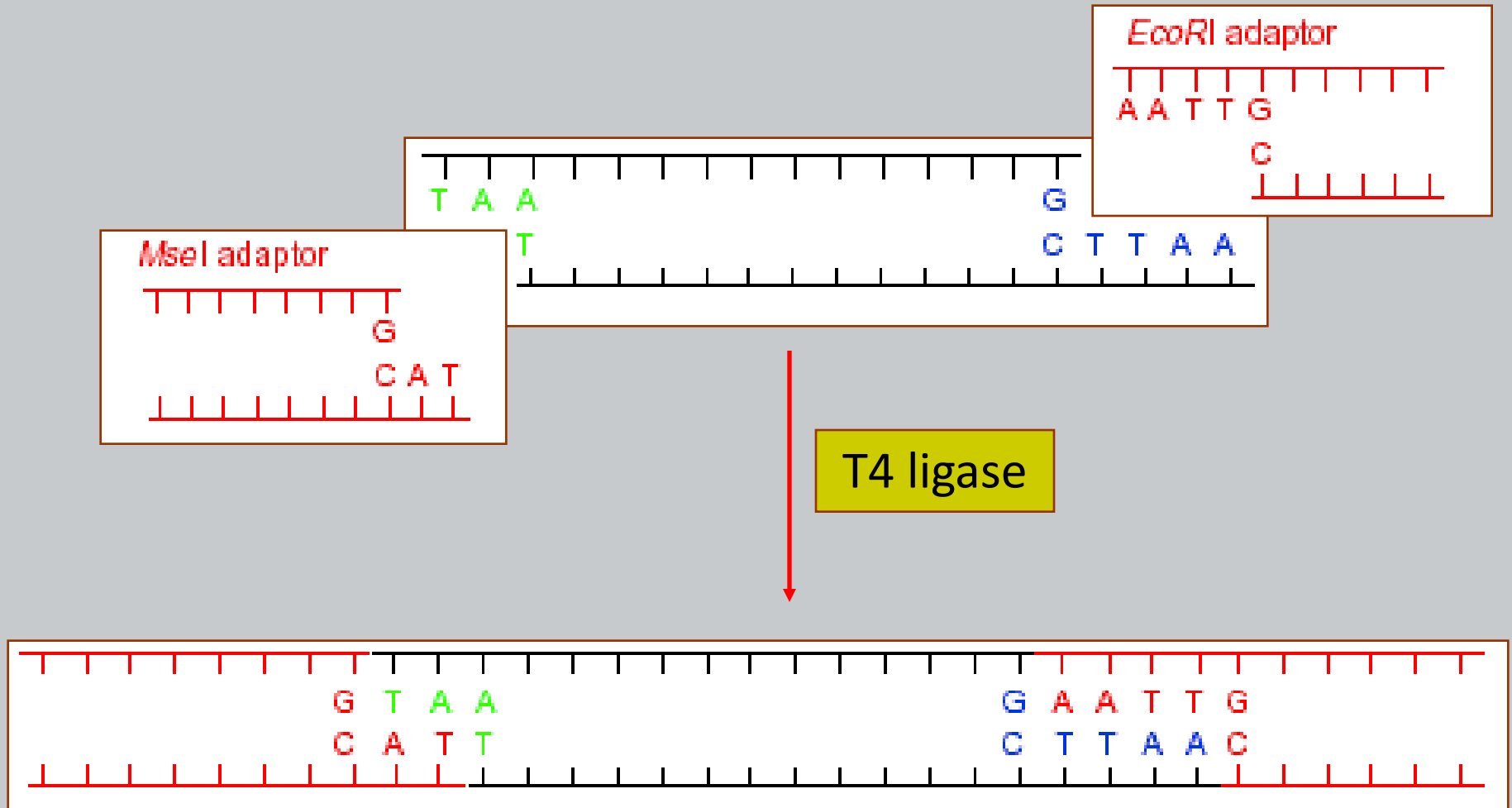


1. Restriktion



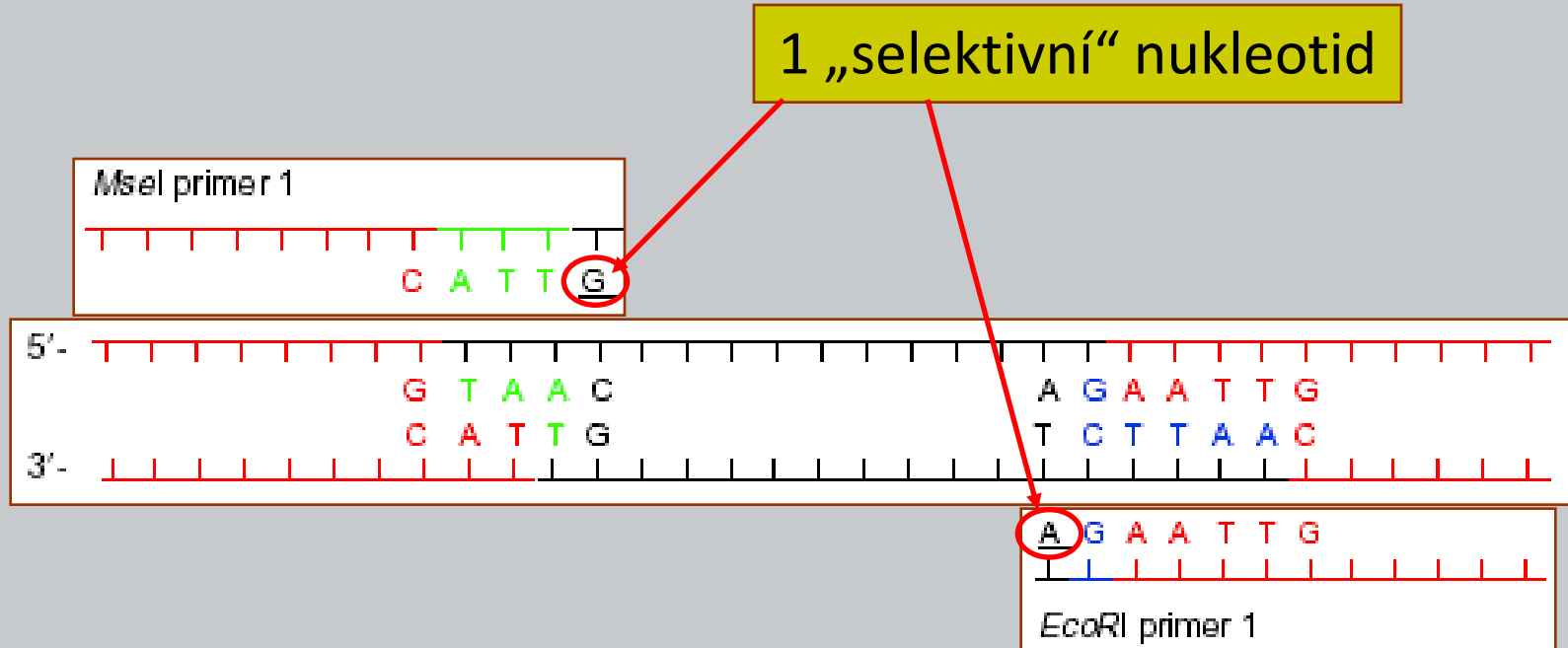
<i>MseI</i>	4 bp	„frequent cutter“
<i>EcoRI</i>	6 bp	„rare cutter“

2. Ligase



3. Pre-amplifikace

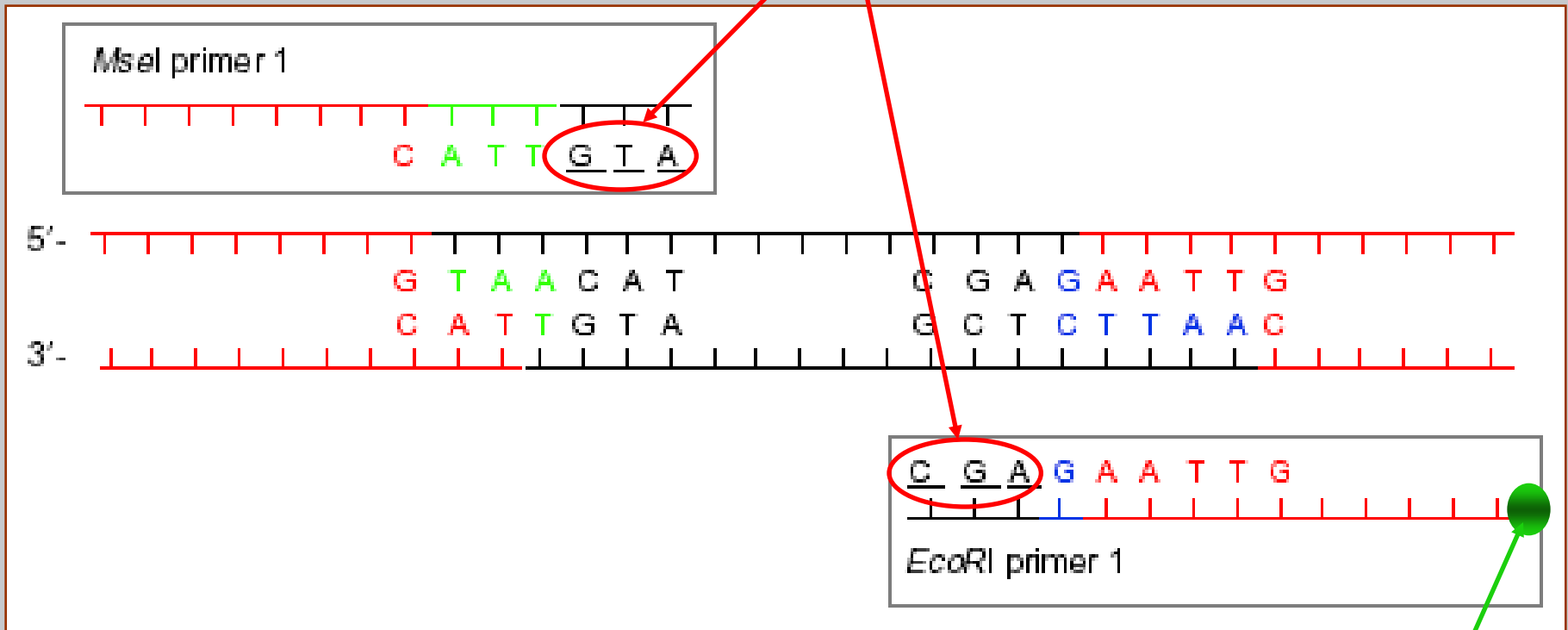
1 „selektivní“ nukleotid



PCR amplifikace jen 1/16 všech proužků

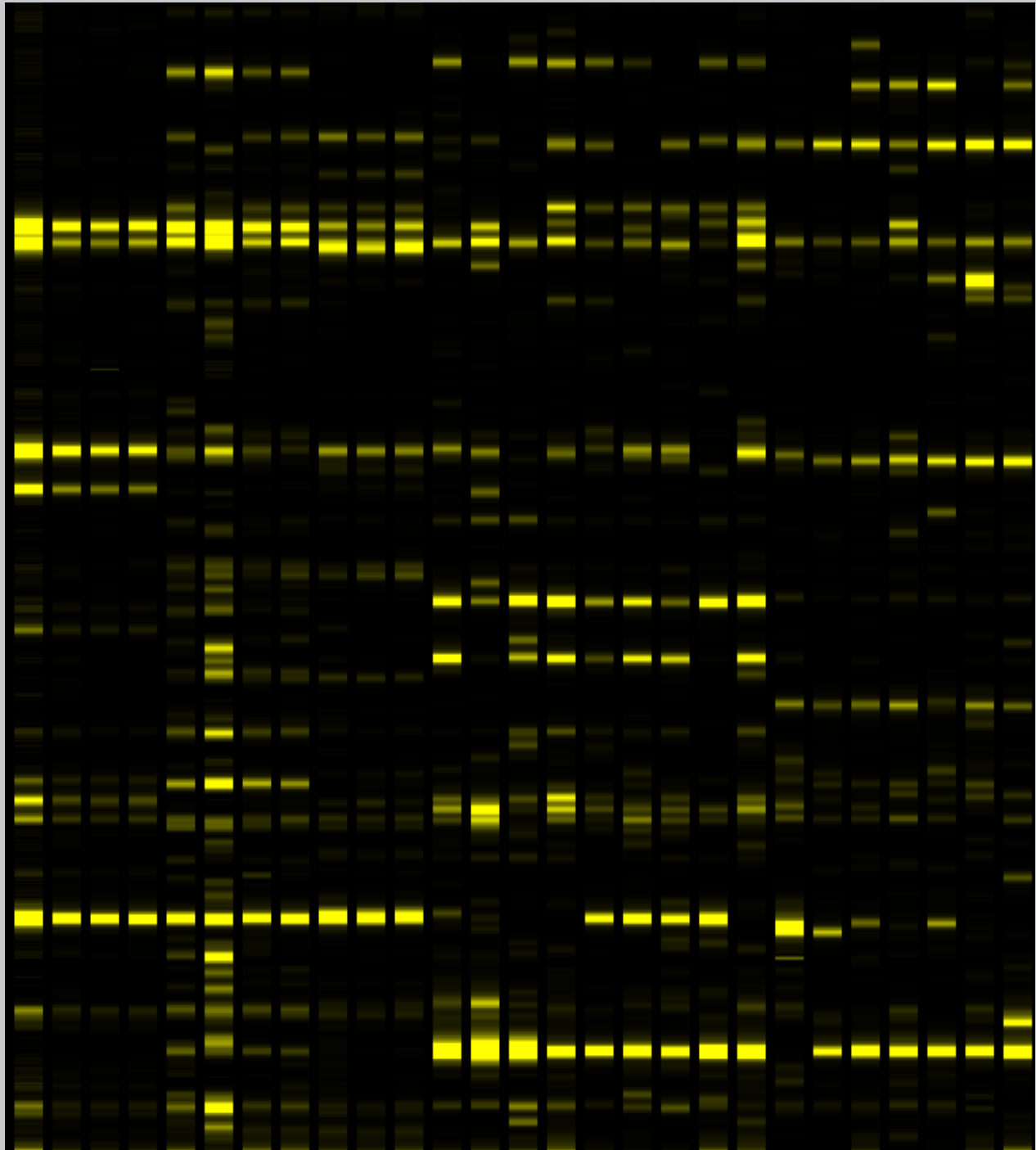
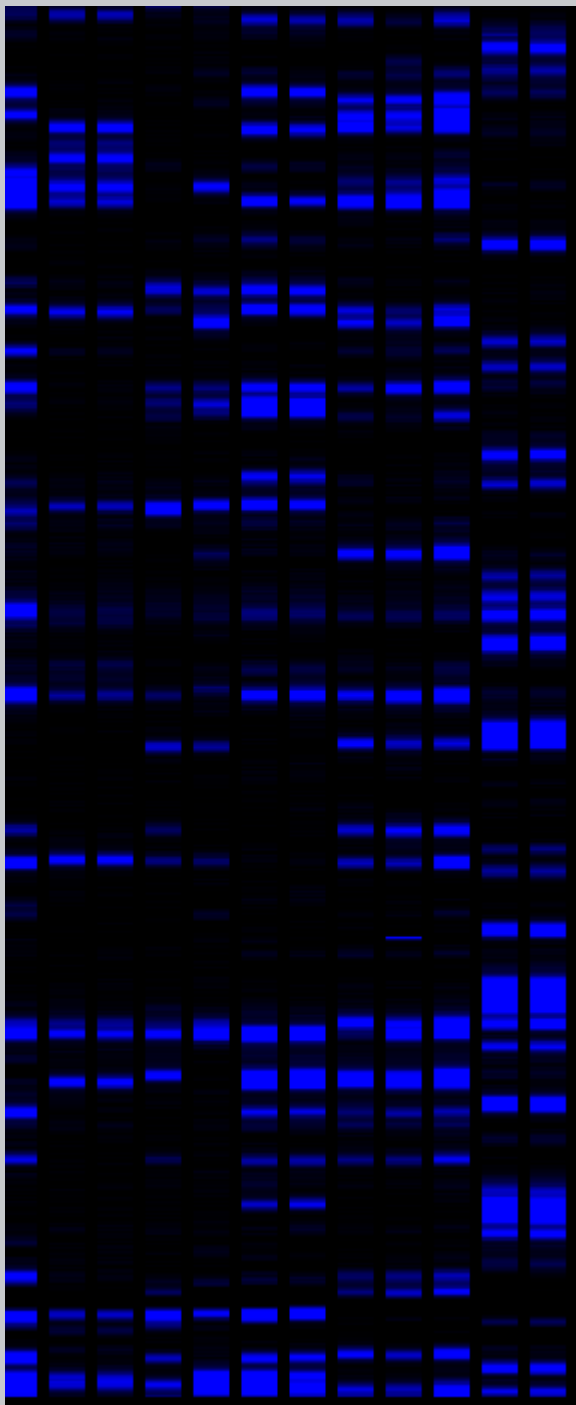
4. Selektivní amplifikace

3 „selektivní“ nukleotidy



fluorescenčně značený primer

PCR amplifikace jen 1/256 všech proužků



Visualizace

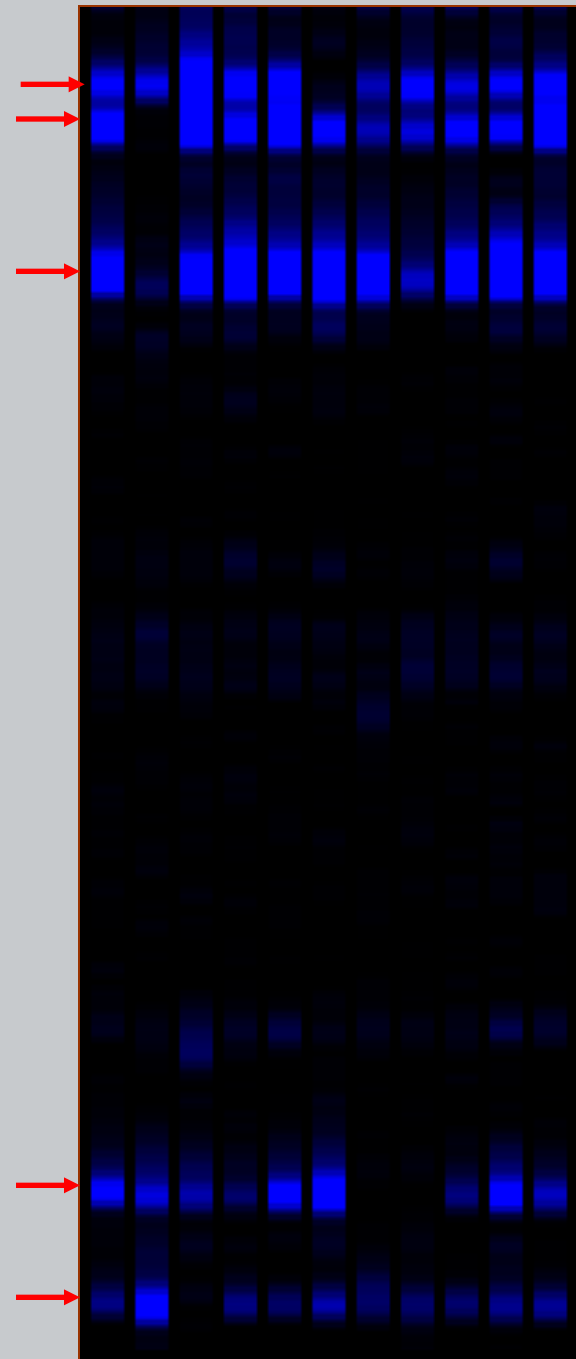
- fluorescenčně značené fragmenty
- využití automatického sekvenátoru
- vyhodnocení pattern v programu
← GENOGRAPHER
- proužky cca 50-500 bp

Variabilita dána

- mutací v restričním místě
- insercí delecí mezi restričními místy

Předpoklady pro interpretaci

- homologie komigrujících fragmentů
- nezávislost fragmentů



AFLP

výhody

- vysoký stupeň polymorfismu
– i 100 fragmentů / 1 kombinace primerů
- není potřeba žádná předchozí znalost o studovaném organismu
- vysoce reprodukovatelné pattern
- variabilita skrz celý genom

nevýhody

- dominantní marker
 - neznámý původ proužků
 - homologie?
 - intenzita?
- vyžaduje optimalizaci (nalezení vhodné kombinace primerů)
- relativně komplikovaná (a drahá) metoda

Standardizace AFLP

- stejné množství vstupní DNA
- replikace cca 10% vzorků – výpočet *error rate* (počet rozdílů/počet srovnávaných lokusů)
- negativní kontroly (test kontaminace)
- vyloučit fragmenty s frekvencí nižší než „*error rate*“

Problematika hodnocení AFLP profilů

- nezávislost fragmentů ?
- problém homologie fragmentů (komigrace nehomologních fragmentů?)
 - různá intenzita
 - drobné rozdíly v mobilitě homologních fragmentů
- asymetrie v pravděpodobnosti získání a ztráty fragmentu (kritika pro použití MP na 0-1 data)
- nemožnost odlišit homozygoty od heterozygotů

Vyhodnocení

- prezenze/absence → 0-1 matice
- počet fragmentů, % polymorfních fragmentů
- počet unikátních a vzácných fragmentů, DW-index
- koeficienty **podobnosti** (*Nei & Li, Jaccard, ...*)
 - dendrogramy (UPGMA, NJ ...)
 - sítě (neighbour-net)
 - PCoA – *principal coordinate analysis*
- *Bayesian clustering* (BAPS, Structure...)
- AMOVA (Arlequin ...) – rozdělení variance, F_{ST} analog...
- vnitropopulační diverzita – Shannonův index diverzity, *average gene diversity...*
- Kinship coefficient (Hardy 2003)

Koeficienty podobnosti

		jedinec A	
		presence 1	absence 0
jedinec B	presence 1	a	b
	absence 0	c	d

	A	B
a	—	—
b		—
c	—	
d		

- Jaccardův koeficient (Jaccard 1908)

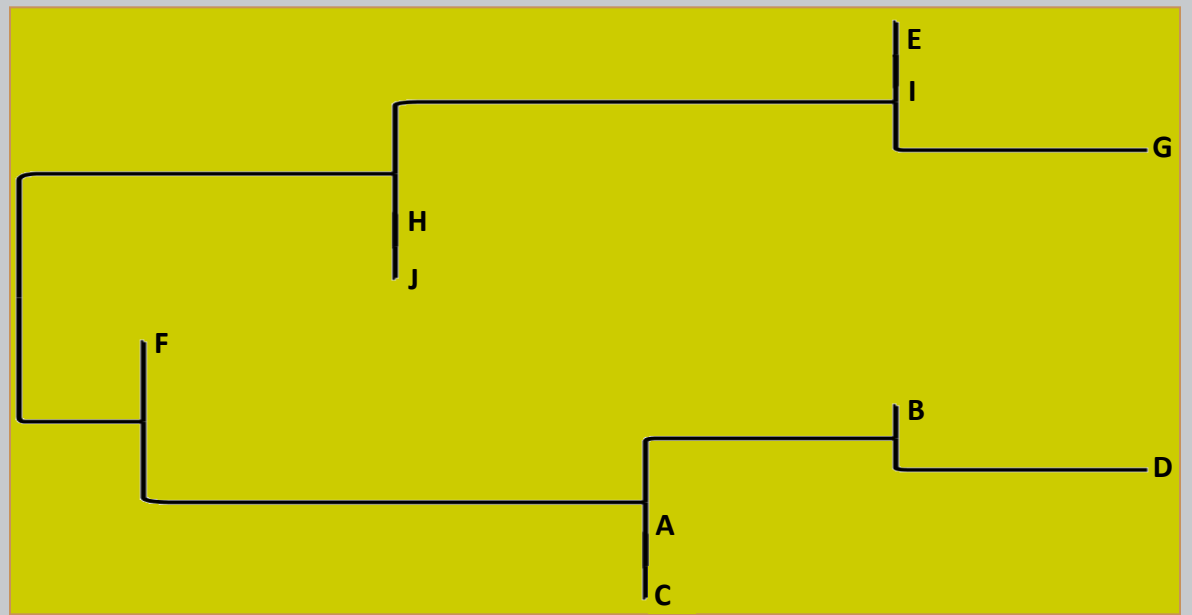
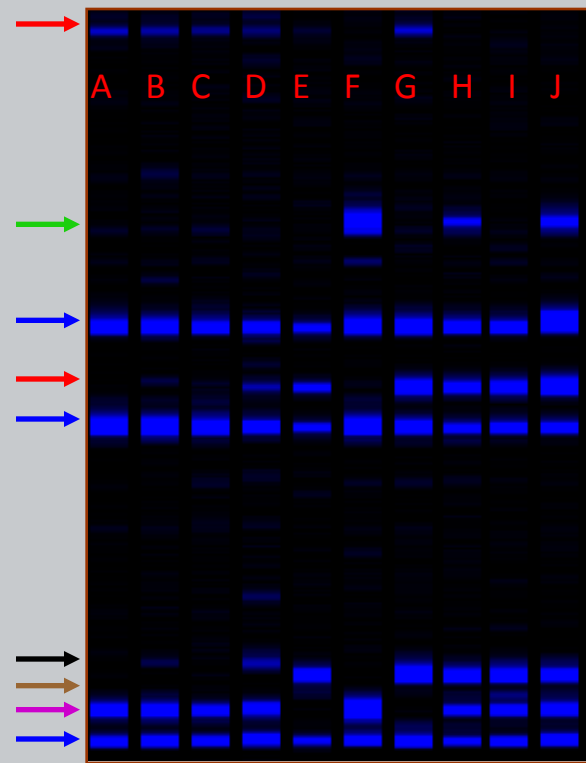
$$\frac{a}{a+b+c}$$

- Dice koeficient (Dice 1945) = Nei & Li 1979, Sørensen 1948

$$\frac{2a}{2a+b+c}$$

- „simple-matching“ koeficient (Sokal & Michener 1958)

$$\frac{a+d}{a+b+c+d}$$

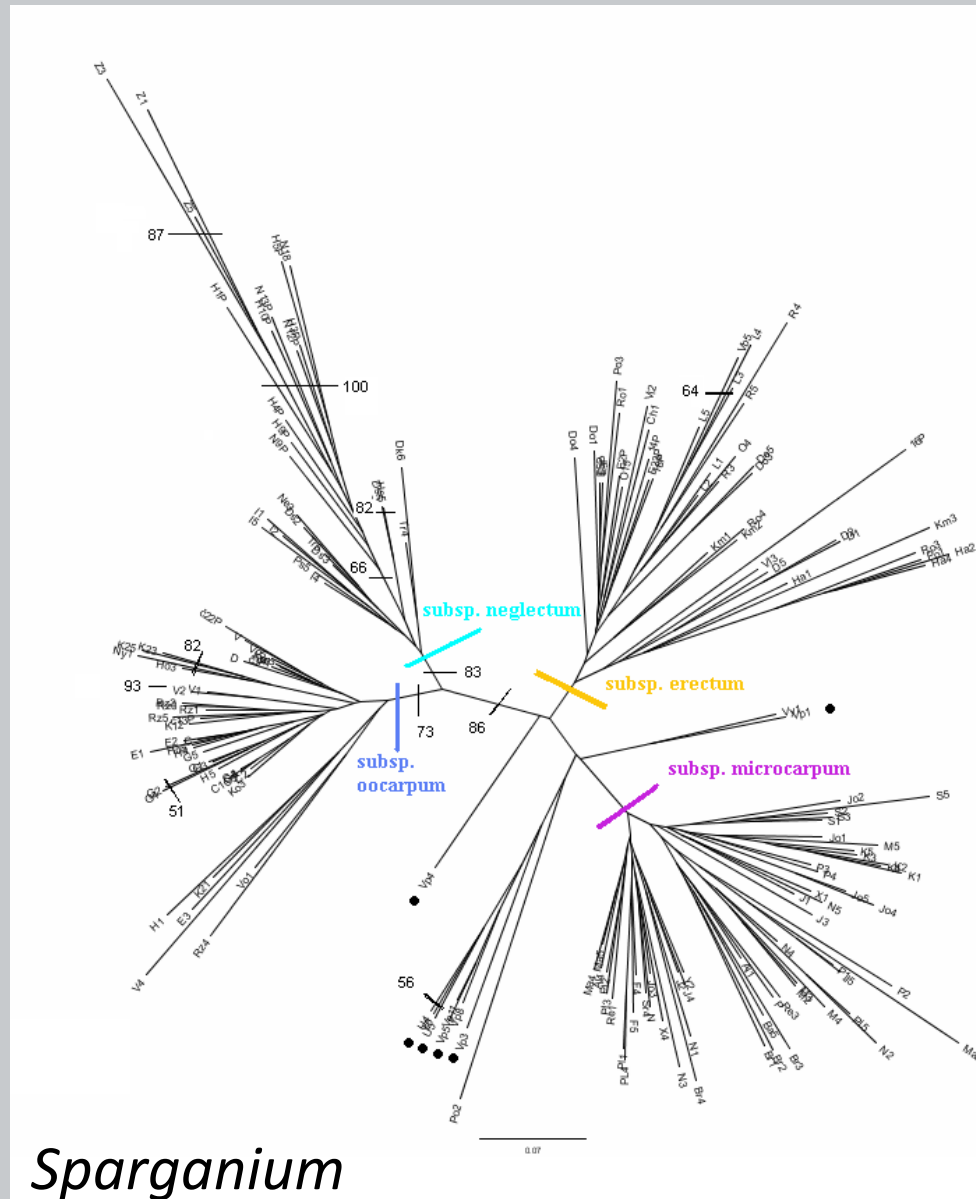


neighbour-joining tree,
midpoint rooting

1	1	1	1			1				
					1		1			1
1	1	1	1	1	1	1	1	1	1	1
			1	1		1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1
	1		1							
				1		1	1	1	1	
1	1	1	1		1		1			1
1	1	1	1	1	1	1	1	1	1	1

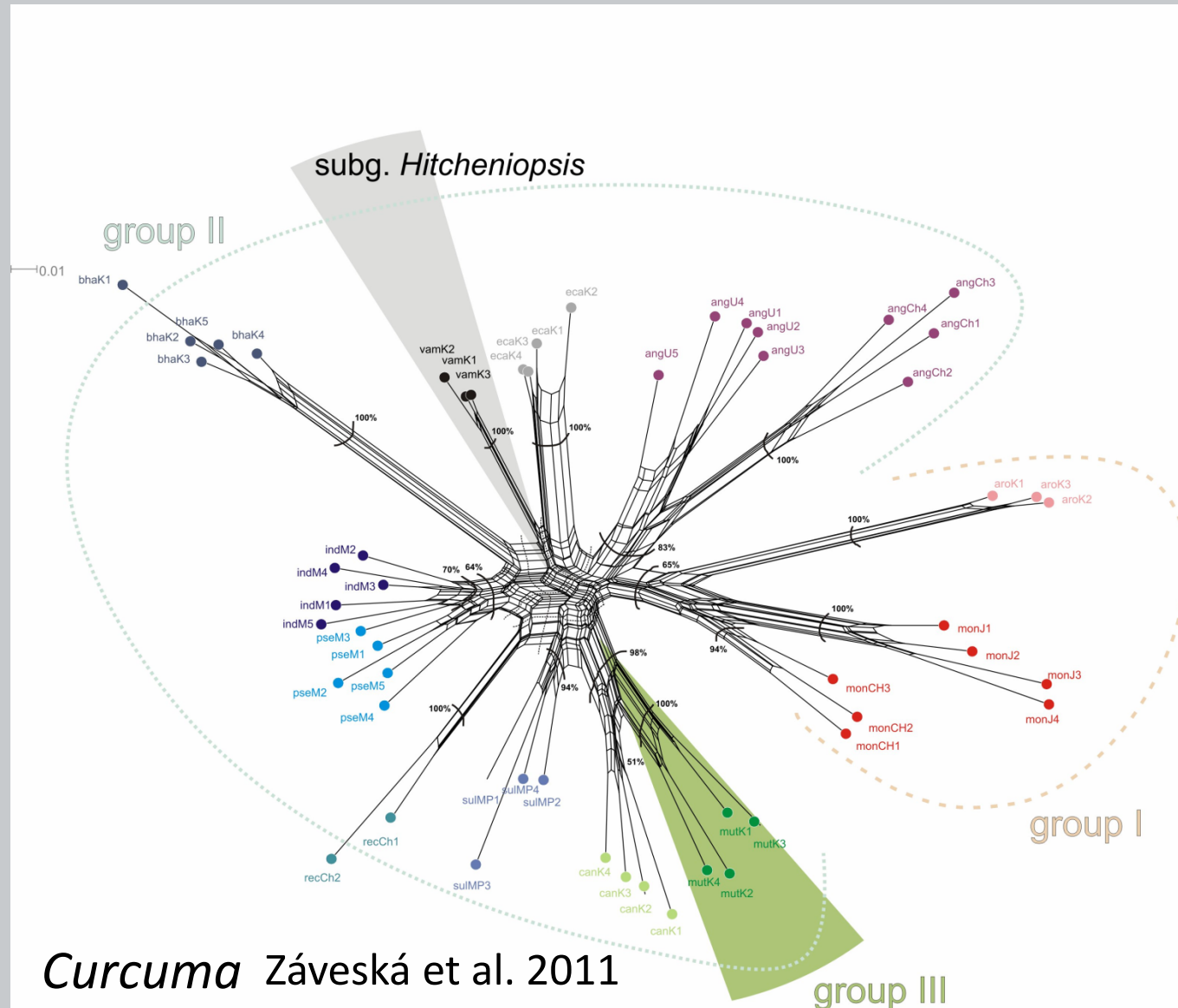
	A	B	C	D	E	F	G	H	I	J
A	0,00000									
B	0,01591	0,00000								
C	0,00000	0,01591	0,00000							
D	0,03046	0,01336	0,03046	0,00000						
E	0,08576	0,10190	0,08576	0,06797	0,00000					
F	0,03731	0,05331	0,03731	0,06797	0,08576	0,00000				
G	0,05331	0,06797	0,05331	0,04389	0,01591	0,10190	0,00000			
H	0,06797	0,08148	0,06797	0,05634	0,03046	0,03046	0,04389	0,00000		
I	0,08576	0,10190	0,08576	0,06797	0,00000	0,08576	0,01591	0,03046	0,00000	
J	0,06797	0,08148	0,06797	0,05634	0,03046	0,03046	0,04389	0,00000	0,03046	0,00000

NJ strom - nezakořeněný



Sparganium

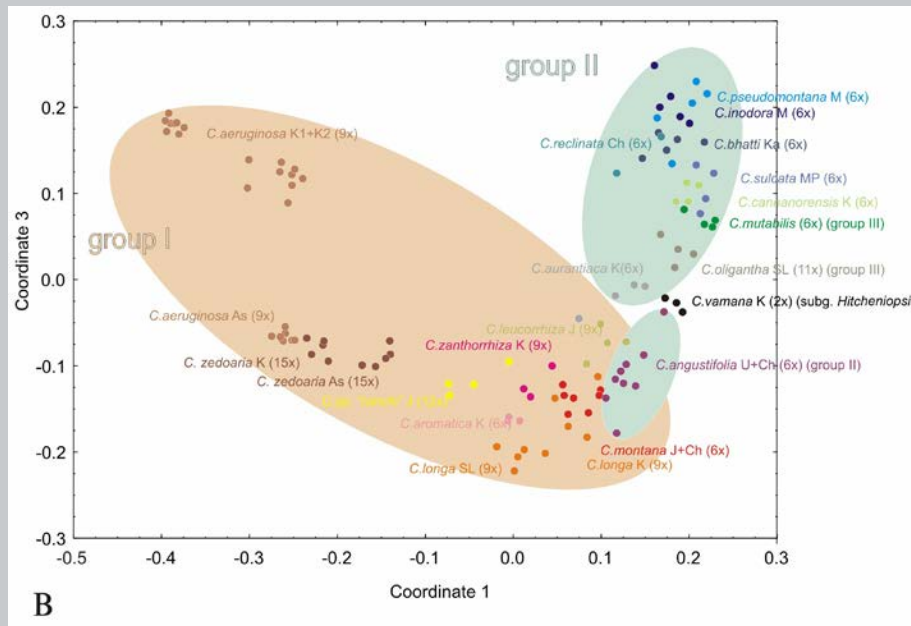
Neighbour-network



Principal Coordinate Analysis

PCoA

- mnohorozměrná analýza
- založená na koeficientech podobnosti
- nalezení směrů nejvyšší variability v souboru
- vizualizace variability v souboru dat



Populační diverzita a divergence

- genetická diverzita
 - Shannonův index diverzity
 - average gene diversity
 - reflektuje reprodukční systém, recentní procesy (genový tok, velikost populace...)
- vzácnost (*rarity*)
 - počet (nebo %) vzácných fragmentů (arbitrární definice vzácnosti)
 - počet (nebo %) unikátních (privátních) fragmentů
 - DW-index (*frequency down-weighted marker values*)
 - reflektuje historické procesy (dlouhodobá izolovanost...)

Shannonův index diverzity

$$H_{Sh} = - \sum_{i=1}^k p_i \ln p_i$$

p_i – frekvence i -tého fragmentu



Saponaria pumila, AFLP

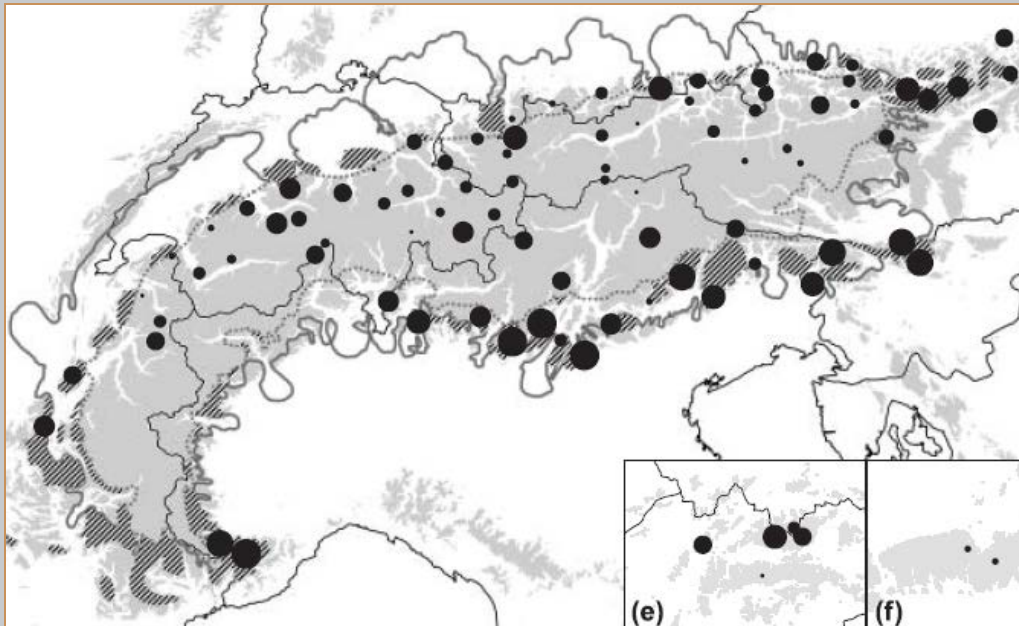
Velikost H_{Sh} ve vztahu k
maximálnímu zalednění
na území Alp

(Tribtsch et al. 2002)

DW-index

frequency down-weighted marker values

- počet výskytů AFLP markeru v populaci/počet výskytů markeru v celém datovém souboru -> součet hodnot pro všechny markery
- vyšší hodnoty v dlouhodobě izolovaných populacích (akumulace díky mutacím)
- nižší hodnoty v nově vzniklých populacích (recentní šíření)



Ranunculus alpestris, AFLP

distribuce DW-indexu ve vztahu k maximálnímu zalednění na území Alp

(Paun et al. 2008)

Mantelův test – prostorová autokorelace

- porovnání dvou matic \rightarrow Mantel R_M

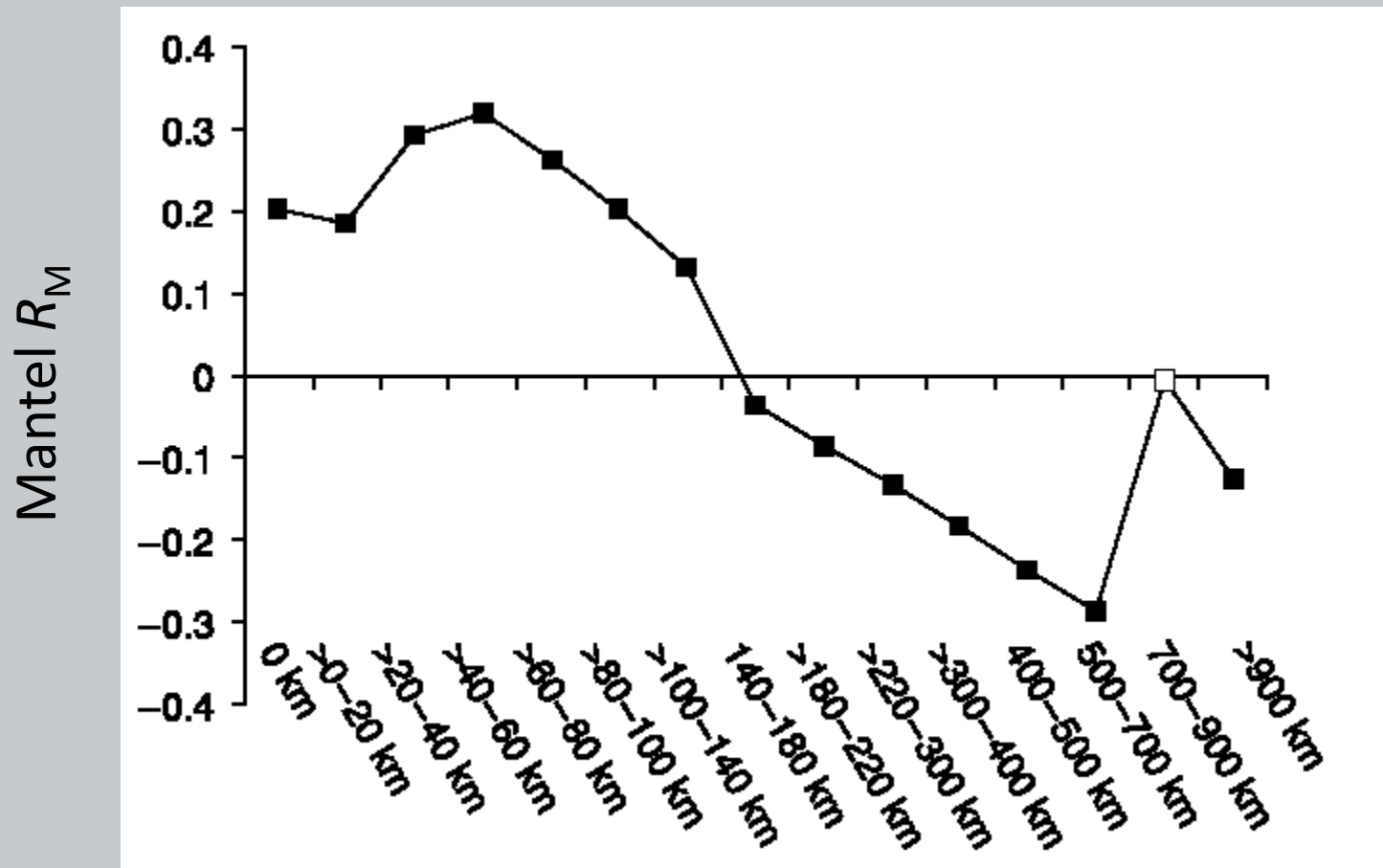
	A	B	C	D	E	F	G	H	I	J
A	0,00000									
B	0,01591	0,00000								
C	0,00000	0,01591	0,00000							
D	0,03046	0,01336	0,03046	0,00000						
E	0,08576	0,10190	0,08576	0,06797	0,00000					
F	0,03731	0,05331	0,03731	0,06797	0,08576	0,00000				
G	0,05331	0,06797	0,05331	0,04389	0,01591	0,10190	0,00000			
H	0,06797	0,08148	0,06797	0,05634	0,03046	0,03046	0,04389	0,00000		
I	0,08576	0,10190	0,08576	0,06797	0,00000	0,08576	0,01591	0,03046	0,00000	
J	0,06797	0,08148	0,06797	0,05634	0,03046	0,03046	0,04389	0,00000	0,03046	0,00000

genetické vzdálenosti

	A	B	C	D	E	F	G	H	I	J
A	0,0									
B	0,2	0,0								
C	0,2	0,1	0,0							
D	0,1	0,6	0,6	0,0						
E	0,3	0,2	0,6	0,6	0,0					
F	0,3	0,8	0,3	0,2	0,6	0,0				
G	0,3	0,1	0,3	0,8	0,8	0,6	0,0			
H	0,3	0,8	0,6	0,2	0,6	0,2	0,1	0,0		
I	0,8	0,3	0,2	0,6	0,0	0,8	0,2	0,1	0,0	
J	0,8	0,1	0,6	0,1	0,6	0,6	0,3	0,0	0,3	0,0

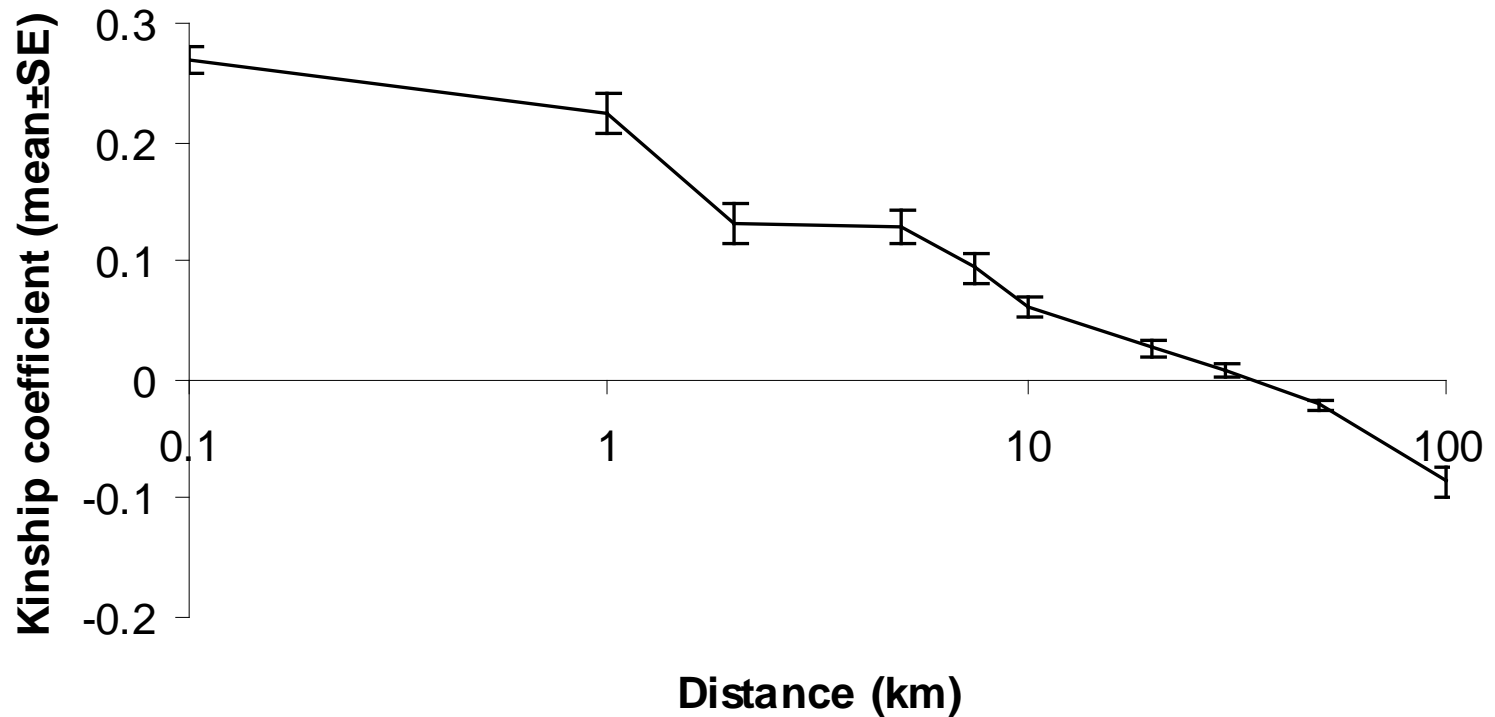
geografické vzdálenosti

Mantel correlogram



Phyteuma globulariifolium v Alpách (Schönswetter et al. 2002)

Spatial autocorrelation



Sparganium erectum v povodí Cidliny (Fér & Pfosser unpubl.)

Bayesian clustering

- hledání takového rozdělení individuů do K clusterů, které je optimální, tj. má maximální záporný logaritmus marginální pravděpodobnosti
- výsledkem je zjištění optimálního počtu clusterů, tj. „reálných populací“ a rozřazení všech individuů
- software
 - BAPS - Bayesian clustering of population structure (Corander *et al.* 2003)
 - *Structure* (Pritchard *et al.* 2000)
 - *Geneland* (Guillot *et al.* 2005)

Výstup z programu BAPS v3.2

RESULTS OF INDIVIDUAL LEVEL MIXTURE ANALYSIS:

Data file: S_preprocessed_BAPS.mat

Number of clustered individuals: 258

Number of groups in optimal partition: 8

Log(marginal likelihood) of optimal partition: -8062.66

Best Partition:

Cluster 1: {10, 11, 33, 40, 41, 58, 59, 60, 61, 101, 114, 123,
124, 126, 131, 132, 137, 159, 161, 162, 164, 166,
167, 168, 169, 243, 244, 245, 246, 251, 252, 253,
254}

Cluster 2: {47, 49, 50, 51, 52, 56, 57, 64, 68, 69, 70, 75,
77, 81, 85, 86, 91, 112, 113, 115, 117, 130, 155,
173}

Cluster 3: {25, 26, 27, 28, 29, 138, 139, 140, 141, 146, 147,
148, 149, 150, 151, 156, 160, 163, 165, 184, 185,
186, 187, 188, 189, 190, 191, 192, 194, 196, 198,
199, 200, 201, 202, 203, 204, 205, 206, 207}

rozdělení jedinců do skupin

Changes in log(marginal likelihood) if individual i is moved to group j:

ind	1	2	3	4	5	6	7	8
1:	-61.0	-20.1	-77.6	-94.0	.0	-67.1	-125.9	-71.3
2:	-50.4	-18.2	-69.4	-85.4	.0	-60.5	-117.8	-70.9
3:	-22.5	-60.5	-29.7	-83.0	-64.8	.0	-70.6	-130.4
4:	-22.9	-58.9	-28.8	-83.2	-61.3	.0	-69.3	-126.6
5:	-22.3	-53.1	-28.7	-78.5	-61.0	.0	-65.7	-123.4

změna likelihood modelu
při přesunu jedince do jiné
skupiny

KL-divergence matrix (Kullback-Leibler):

	1	2	3	4	5	6	7	8
1								
2	0.415							
3	0.507	1.062						
4	0.407	0.637	1.515					
5	0.782	0.244	1.030	1.433				
6	0.327	0.817	0.269	1.353	0.826			
7	0.583	1.069	1.009	0.780	1.603	0.944		
8	1.050	0.613	1.833	0.910	0.953	1.751	0.776	

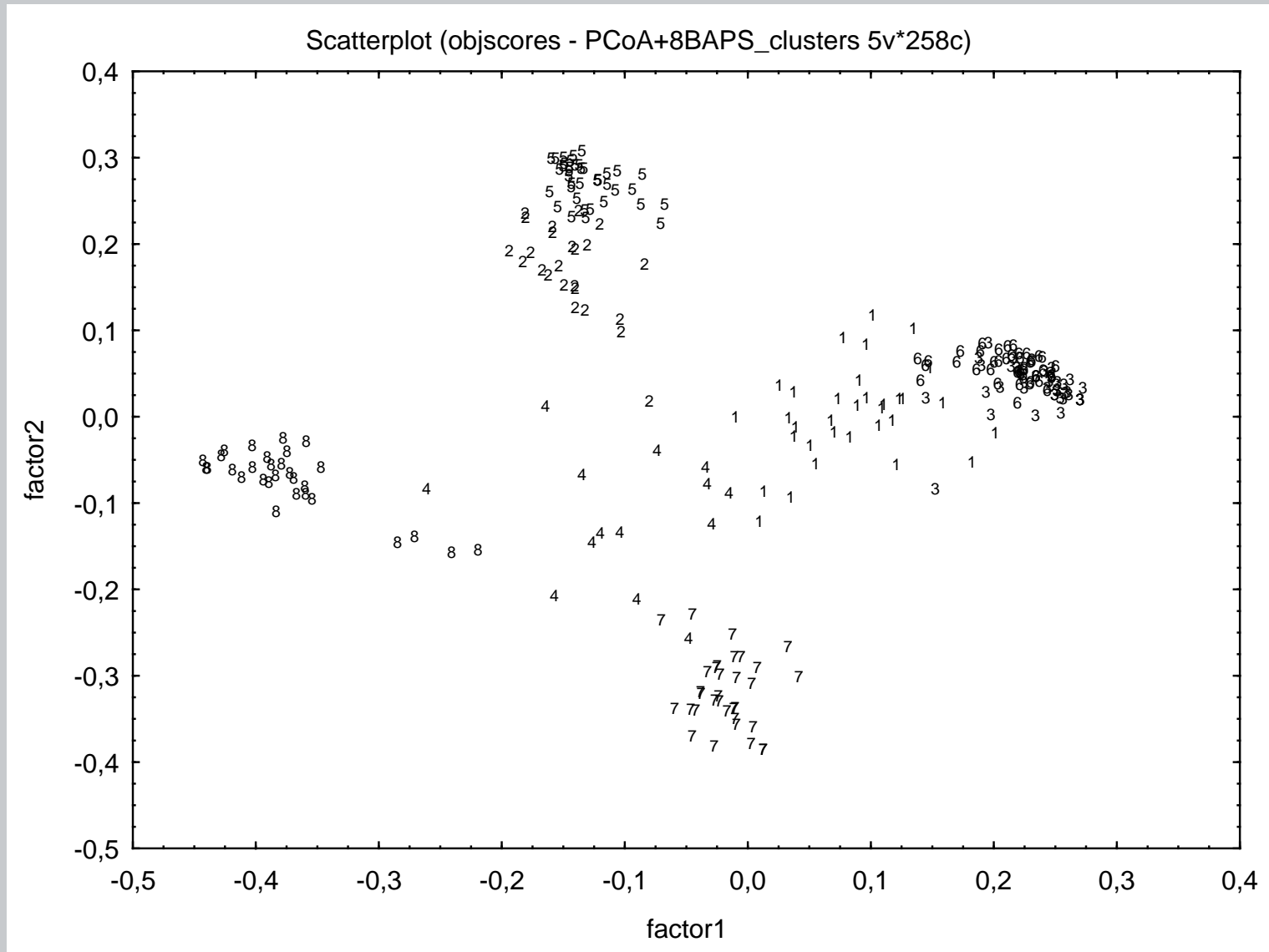
podobnosti mezi skupinami

Probabilities for number of clusters

8 0.9984
9 0.001605

pravděpodobnost modelu

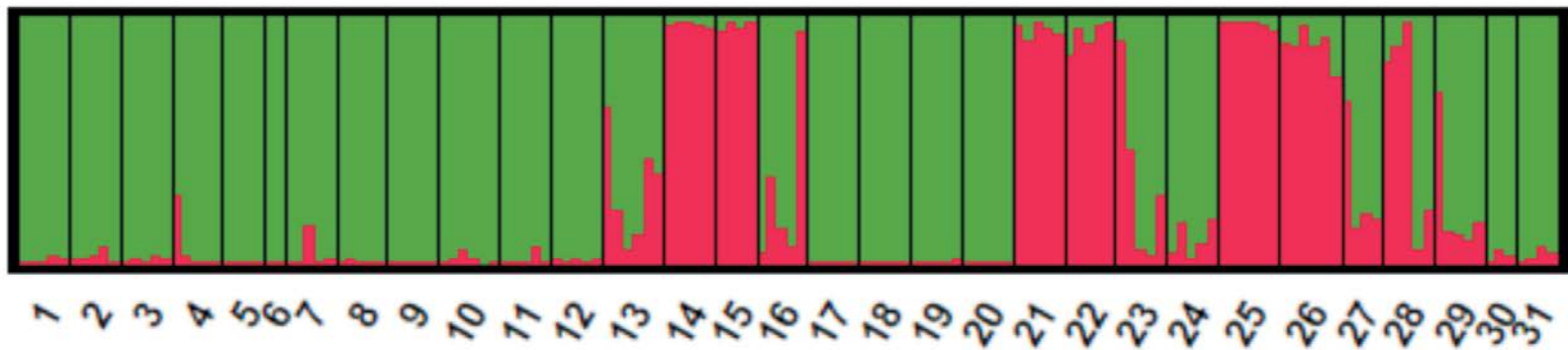
Porovnání výsledků PCoA a bayesovského clusterování



Výsledky programu Structure

Structure (Pritchard *et al.* 2000)

- minimalizuje odchylky od H-W ekvilibria
- minimalizuje *linkage disequilibrium*
- výsledky dále zpracovány programy CLUMPP a Distruct



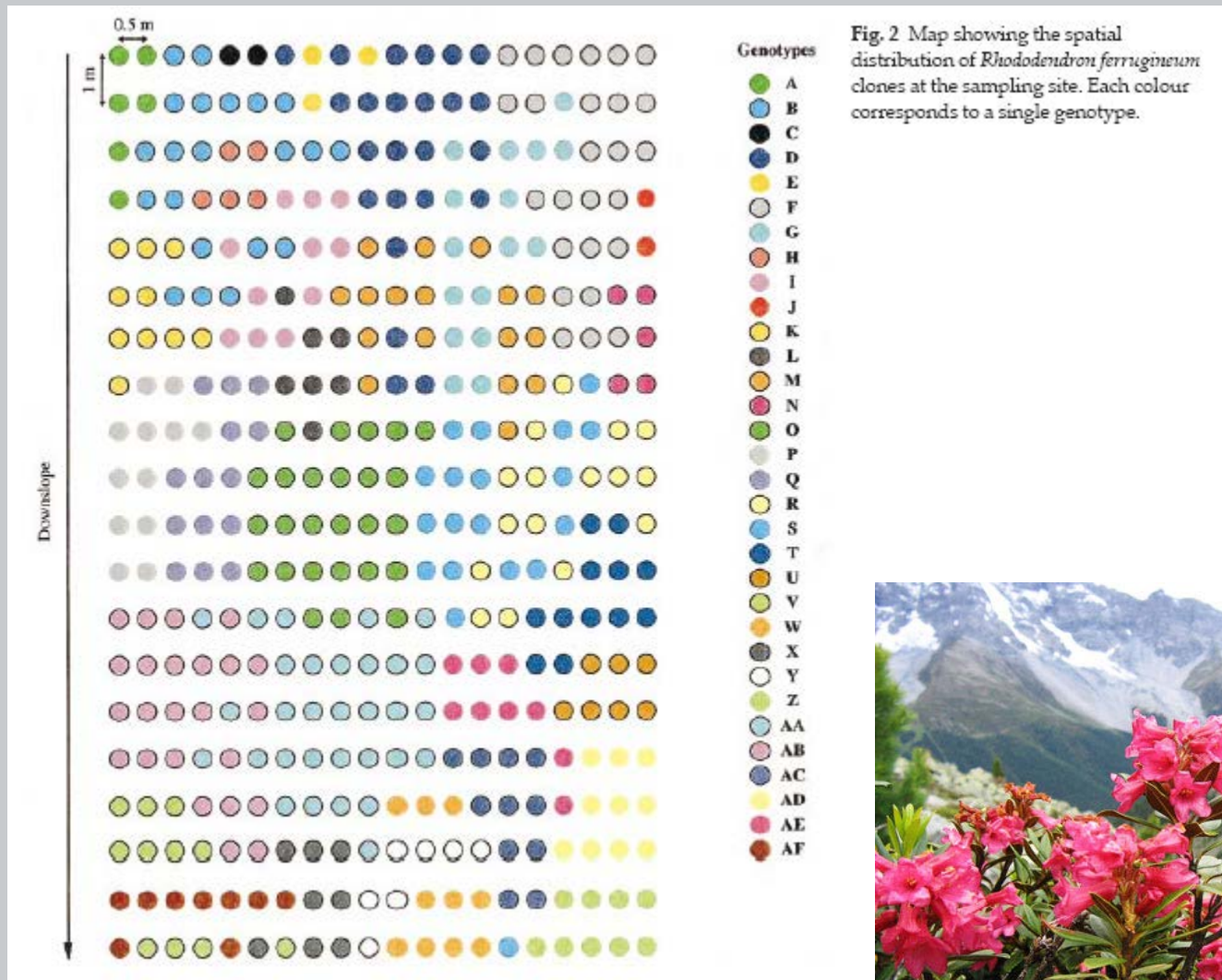
Lonicera nigra – Daneck *et al.* 2011

Další aplikace AFLP

- definice klonů
- identifikace kultivarů
- fylogeografie
- analýza rodičovství (*parentage analysis*)
 - identifikace rodičů jedince
- systematika – fylogeneze, polyploidní speciace...
- hybridizace
- ...

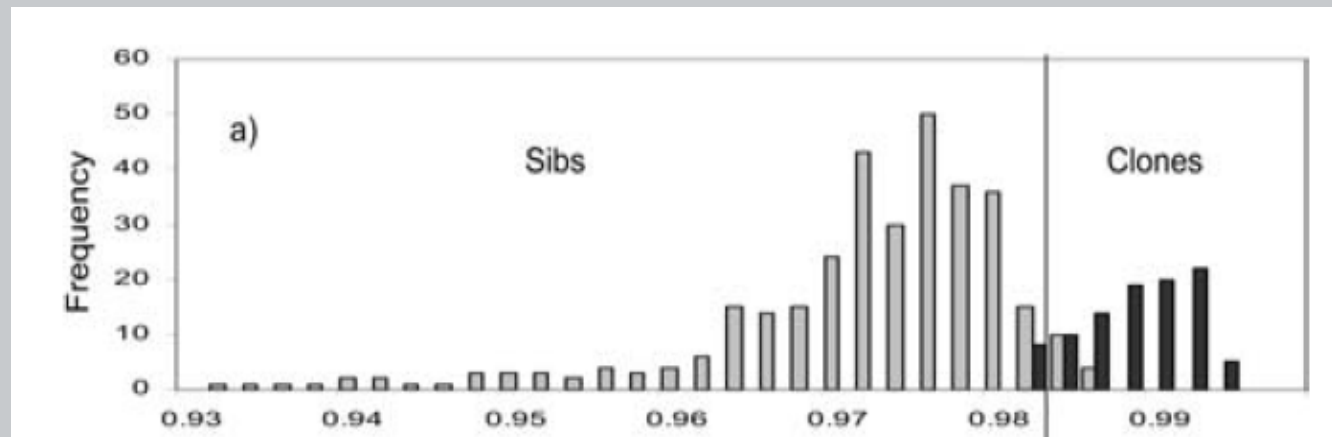
Definice klonů

Escavara N. et al (1998): Clonal diversity in a *Rhododendron ferrugineum* L. (Ericaceae) population inferred from AFLP markers. *Molecular Ecology* 7:975-982.

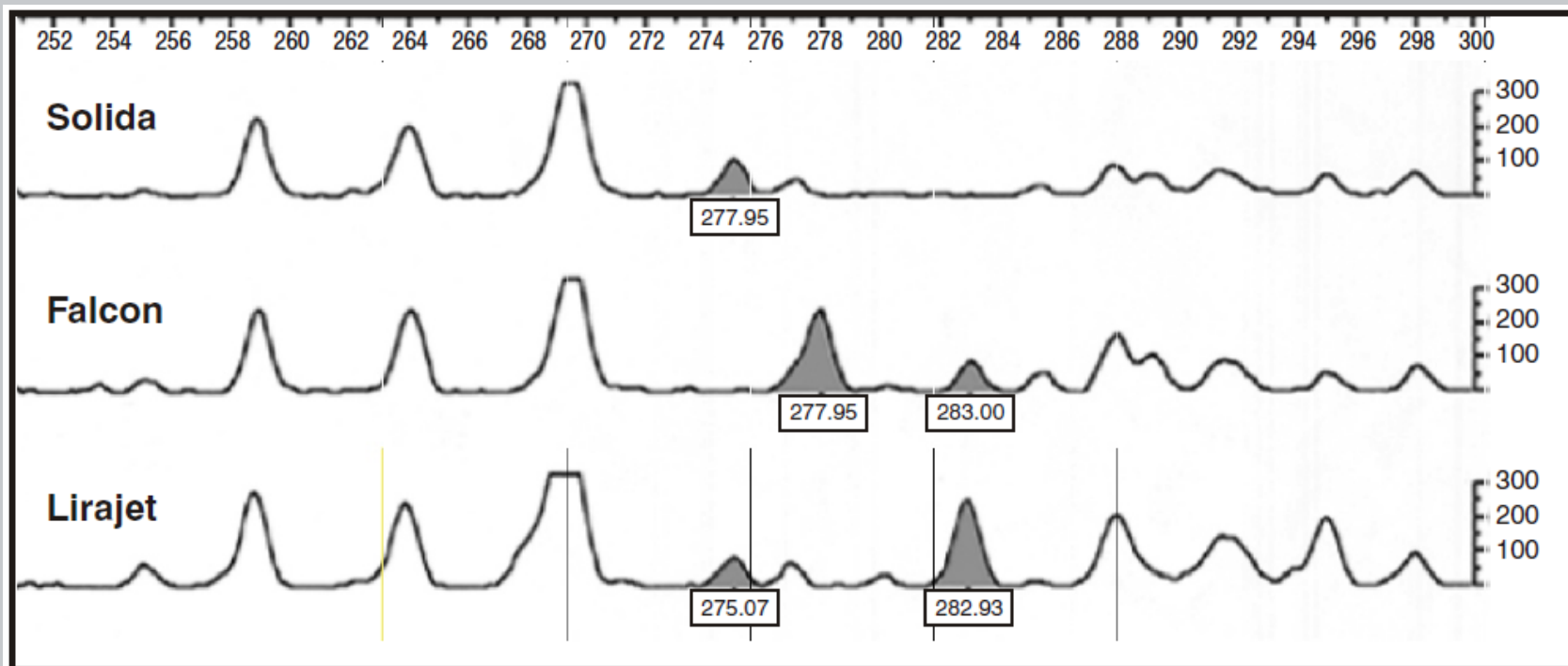


Definice klonů

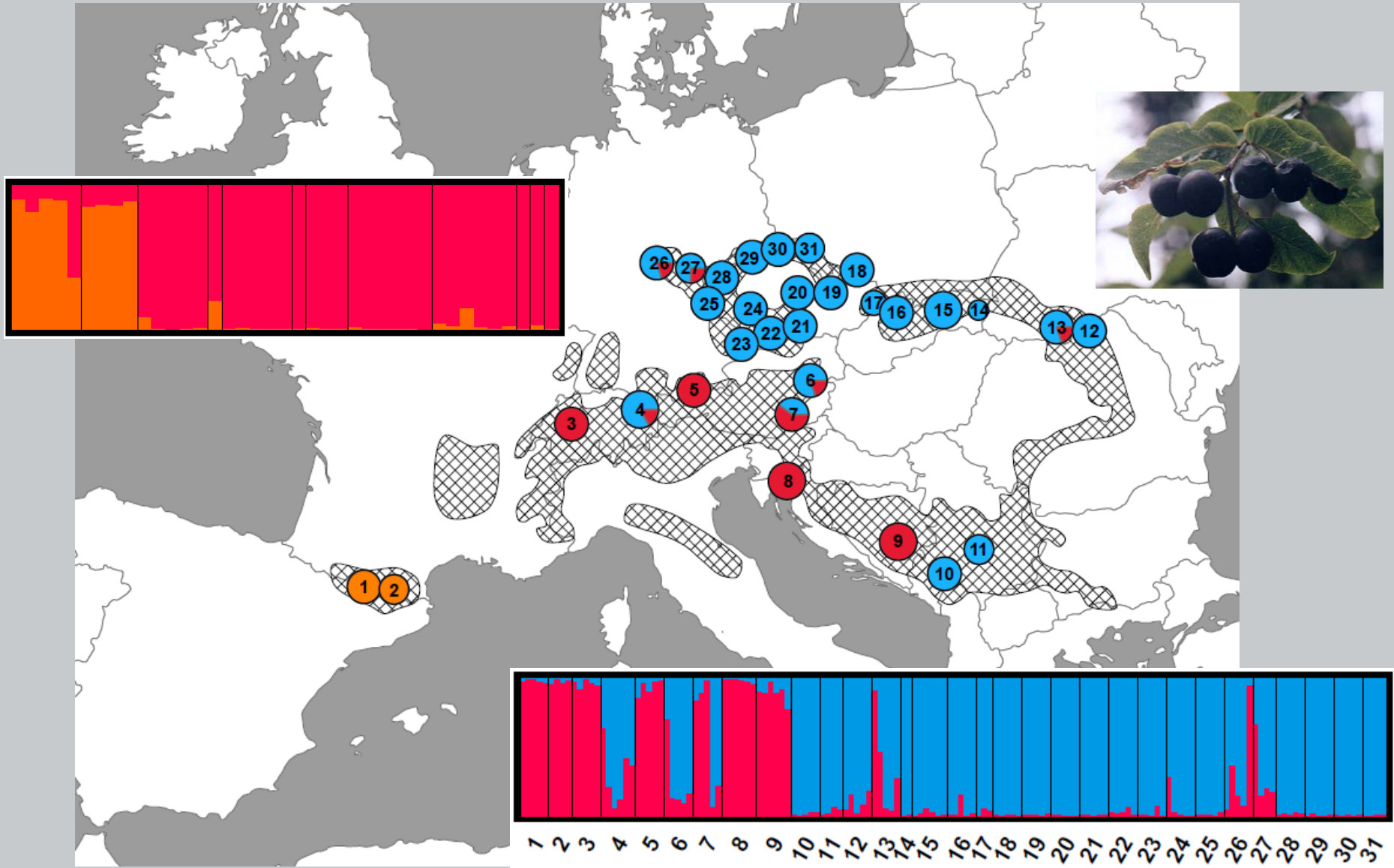
- klon – identický AFLP profil
- *pairwise similarity threshold* – jedinci s určitou podobností (např. 0.98) jsou stále bráni jako klony
- zdroje chyb
 - kontaminace
 - metodologické problémy (nespecif. restrikce, PCR inhibice)
 - chyby při skorování (+ subjektivita při výběru lokusů)
 - somatické mutace



Identifikace kultivarů

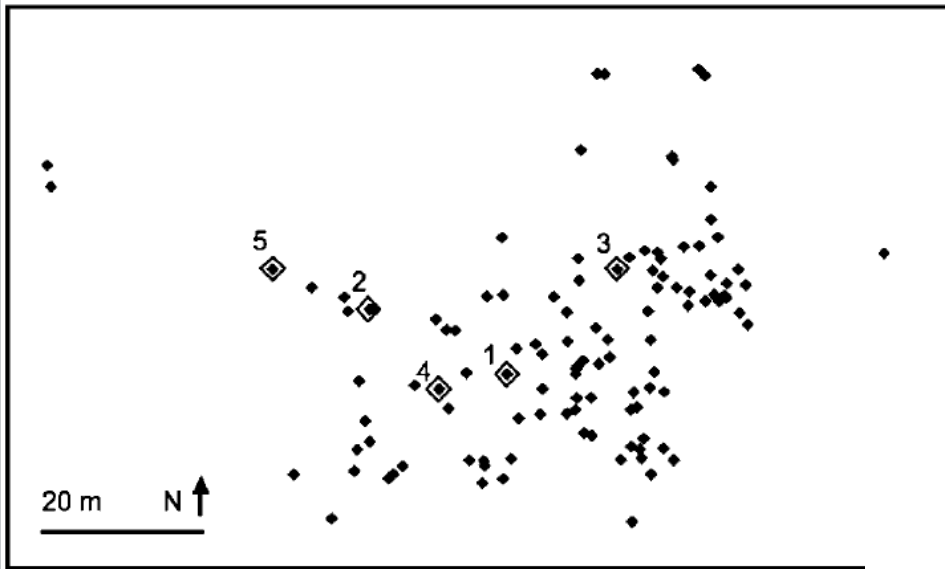


Fylogeografie

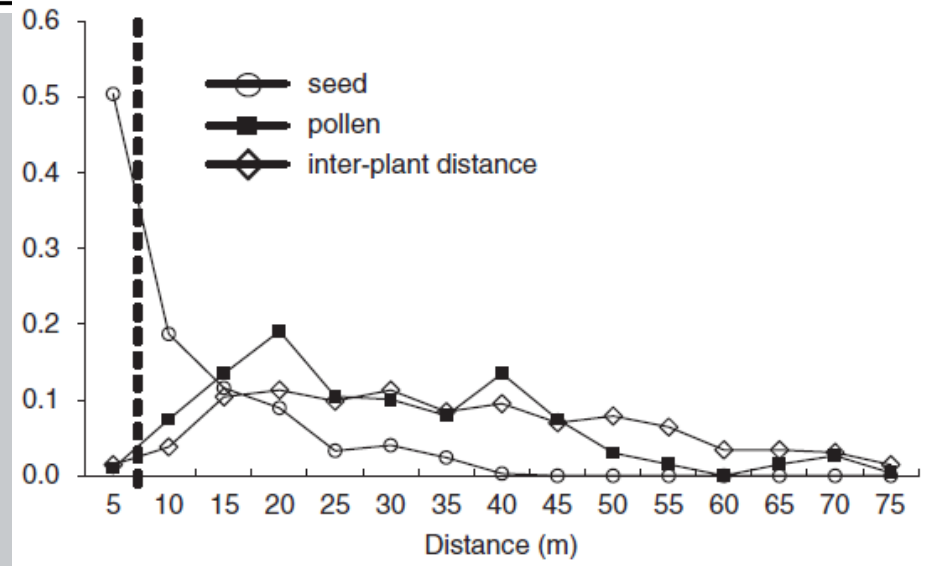


Lonicera nigra – Daneck et al. 2011

Parentage analysis



- určení rodičů semen
- analýza pohybu pylu a semen



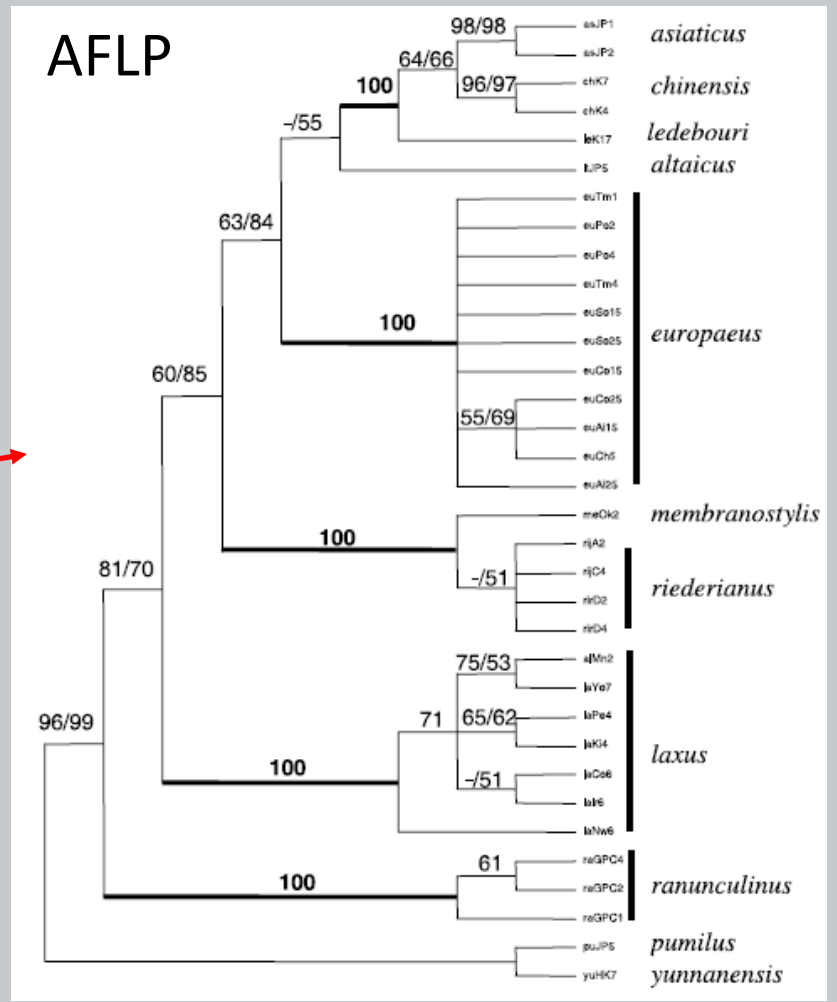
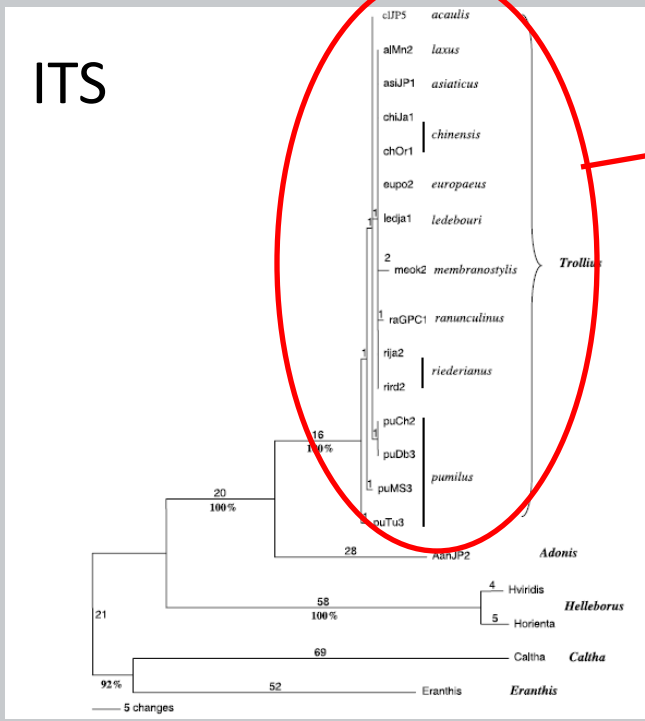
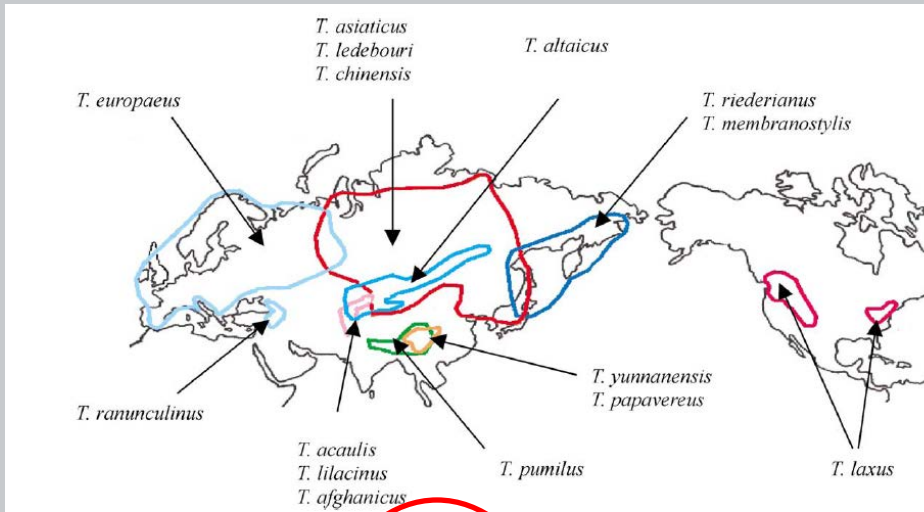
AFLP v systematice

- stejný problém jako u RAPD
 - homologie – čím podobnější taxony, tím pravděpodobnější
 - použitelnost pokud rozdíly v ITS max. 30 bází (Koopman 2005)
- vztahy mezi druhy v rámci rodu
- infraspecifická úroveň
- evoluce polyploidů
- studium hybridizace, introgrese
- úspěšně použito, pokud např. sekvenční informace nebyly dostatečně variabilní

Analýzy pro systematické účely

- fenetické metody
 - koeficienty podobnosti – Nei&Li, Jaccard, Dice ...
 - clusterové analýzy – UPGMA, NJ ...
- kladistické metody
 - rekonstrukce fylogeneze
 - změna znaků ($0 \rightarrow 1 \rightarrow 0$)
 - maximální parsimonie
 - kladogram

Určení fylogeneze



Trollius L. – Després et al., 2003

Studium hybridizace a polyploidní evoluce

Guo Y.-P. et al. (2006): Hybrid origin and differentiation of two tetraploid *Achillea* species in East Asia: molecular, morphological and ecogeographical evidence. *Molecular Ecology* 15:133-144.

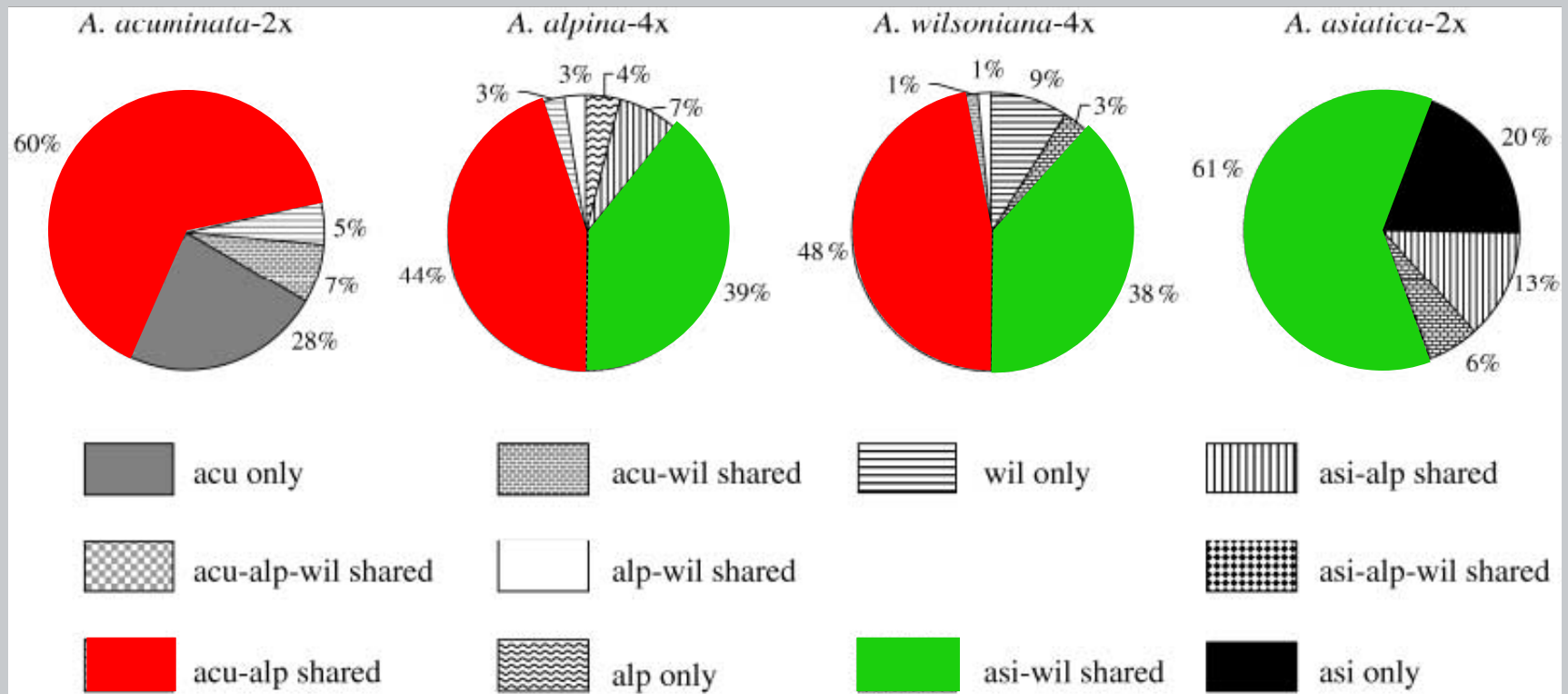
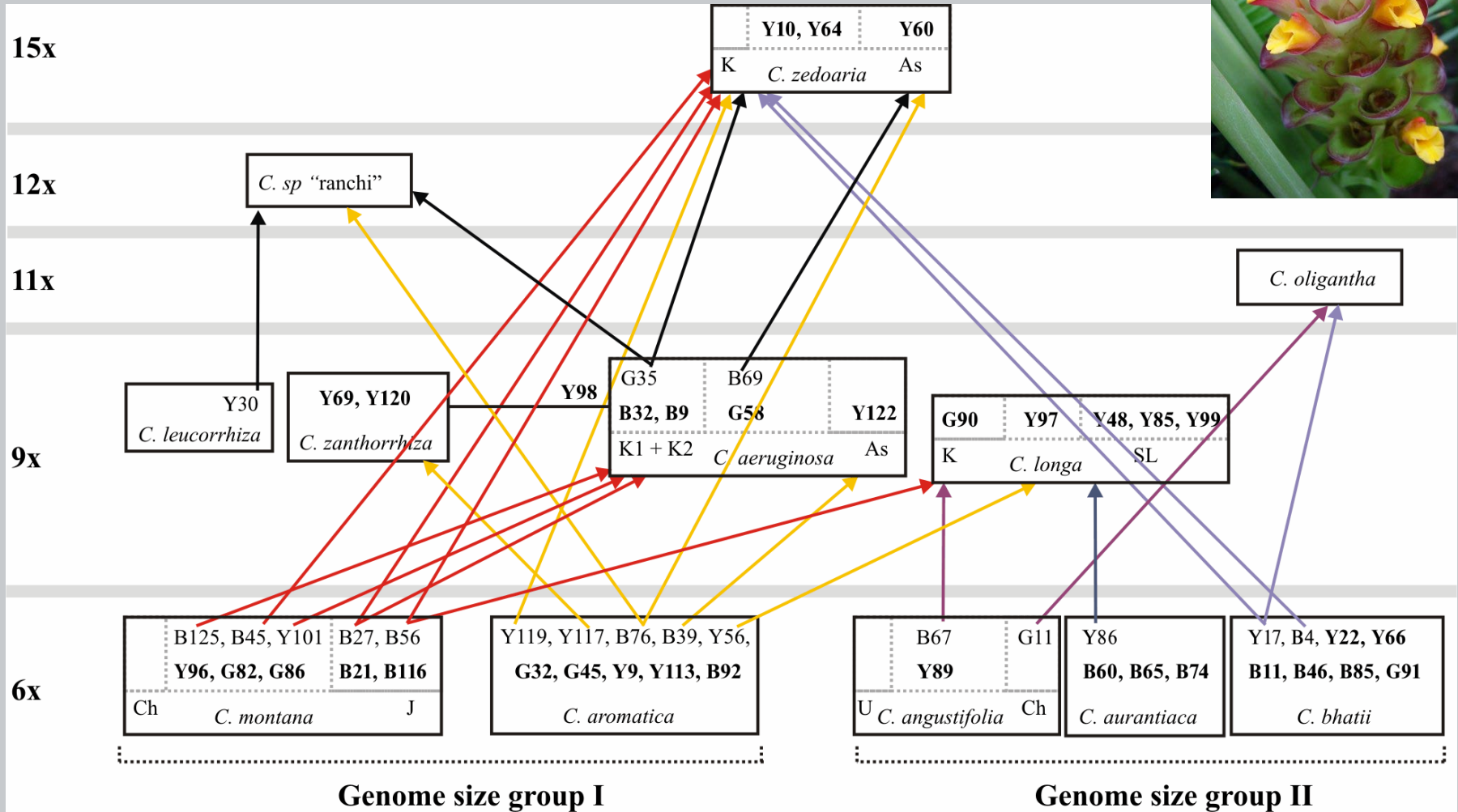
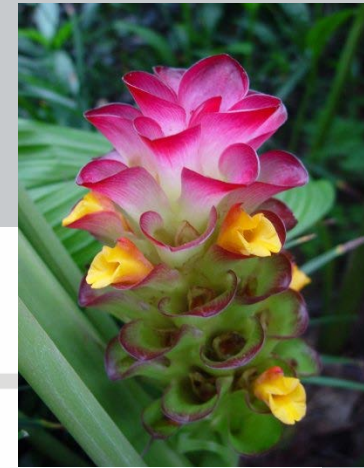


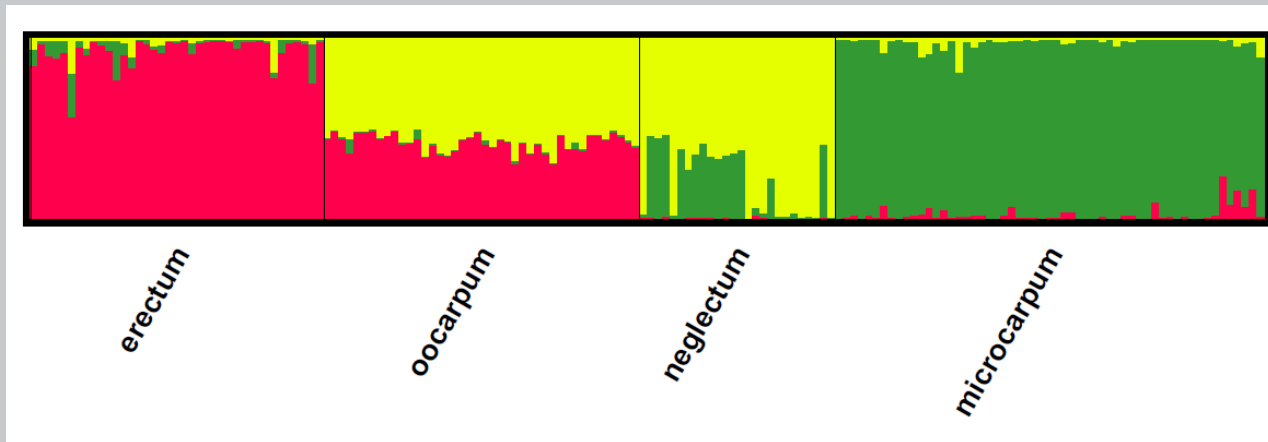
Fig. 2 Frequencies of species-specific and species-shared AFLP bands in populations of *Achillea acuminata-2x*, *Achillea alpina-4x*, *Achillea wilsoniana-4x*, and *Achillea asiatica-2x* in East Asia. The 4x-species share many AFLP bands with the two presumed 2x parental species, but also exhibit a few specific bands.

Polyploidní komplexy



Testování hybridního původu

- hybrid má sdílené proužky od obou rodičů
- pravděpodobnostní (modelová) detekce
 - Structure (Pritchard et al. 2000; Falush et al. 2003)
 - Newhybrids (Anderson & Thompson 2002)



4 subspecie *Sparganium erectum* – Píšová & Fér (unpubl.)

Populační studie

Tribsch A., Schönswetter P. & Stuessy T. (2002):
Saponaria pumila (Caryophyllaceae) and the Ice
Age in the European Alps. *American Journal of
Botany* 89(12): 2024-2033



Systematická studie

Koopman W.J.M et al. (2001): Species relationships in *Lactuca* s.l. (Lactuceae, Asteraceae) inferred from AFLP fingerprints. *American Journal of Botany* 88(10): 1881-1887



Literatura

- Vos P. et al. (1995): *AFLP: a new technique for DNA fingerprinting*. Nucleic Acids Research, 23(21):4407-4414.
- Mueller U.G. & Wolfenbarger L.L. (1999): *AFLP genotyping and fingerprinting*. Trends in Ecology & Evolution 14:389-394.
- Meudt H.M. & Clarke A.C. (2007): *Almost forgotten or latest practice? AFLP applications, analyses and advances*. Trends in Plant Science:106-117.
- Bonin A. et al. (2004): *How to track and assess genotyping errors in population genetics studies*. Molecular Ecology 13:3261-3273.
- Bonin A. et al. (2007): *Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists*. Molecular Ecology 16:3737-3758.
- Robinson J.P. & Harris S.A. (1999): *Amplified Fragment Length Polymorphisms and Microsatellites: A phylogenetic perspective*. In: Gillet E.M.[ed.]: Which DNA marker for which purpose? <http://webdoc.sub.gwdg.de/ebook/y/1999/whichmarker/index.htm>
- Koopman W.M.J. (2005): *Phylogenetic signal in AFLP data sets*. Systematic Biology 54:197-217.
- Fér T. (2011): Dominantní (binární) molekulární markery a jejich využití v populační a systematické biologii. Zprávy České botanické společnosti, Praha, Materiály, 25, 61-78.