

Analýza populační struktury



Hierarchická populační struktura

Druh → populace → subpopulace (demy)

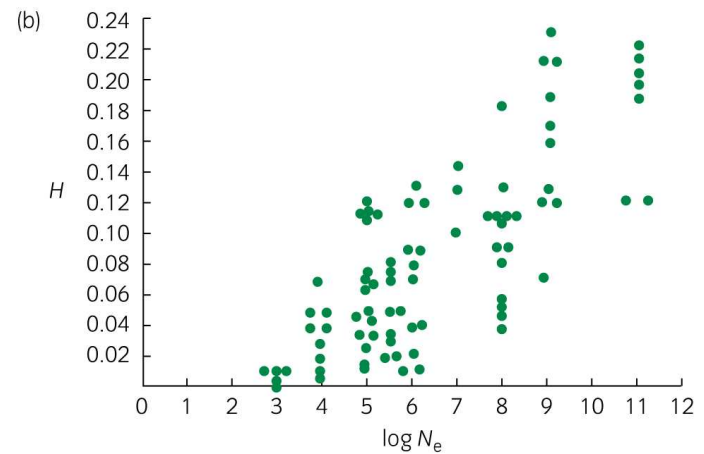
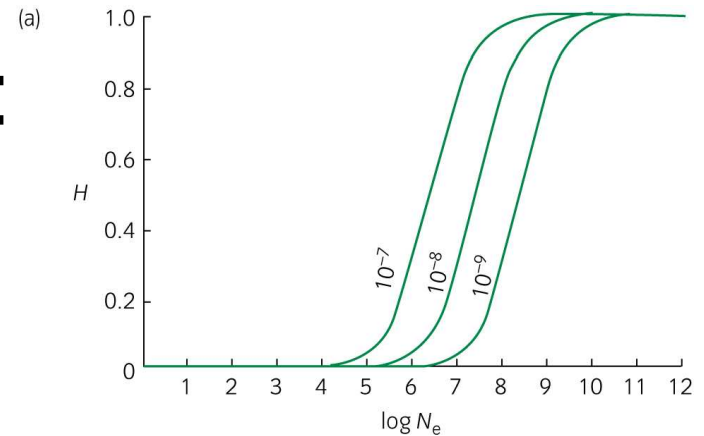
Předpoklady studia populačně-genetické struktury

- lokusy používané pro analýzu populační struktury jsou neutrální vůči selekci
- množství adaptivní variability koreluje s neutrální variabilitou
- popsáno mnoho výjimek (rozdíly v subpopulacích - neutrální vs. adaptivní)

Hierarchická populační struktura

Druh → populace → subpopulace (demy)

- neutrální genetická teorie:
 $H_e = 4N_e\mu / [4N_e\mu + 1]$
- mutation-drift equilibrium



... nemusí to být pravda

- vliv historického vývoje populací („bottlenecks“)
- Northern elephant seals *Mirounga angustirostris* – 120 000 jedinců – 50 allozymových lokusů – žádná variabilita

Indian rhinoceros

- 29 allozyme loci – vyšší variabilita než byla očekávána
- 1960 – 60-80 jedinců v Nepálu
- 1980 – nárůst populace na několik set – max. 2 alely na lokus
- méně intenzivní bottleneck



Hierarchická populační struktura

Druh → populace → subpopulace (demy)

- Fixační indexy
- AMOVA
Analýza genetické variance
- Assignments
- Fylogeneze alel, fylogeografie (někdy příště)

Polymorfismus a heterozygotnost

Polymorfismus

- podíl polymorfních lokusů (znaků) – 95 % nebo 99 % (např. 0,8 = 4 z pěti zkoumaných mikrosatelitů mají v populaci alespoň 2 alely)

Heterozygotnost

- průměrná četnost heterozygotů v jednotlivých lokusech

Očekávaná heterozygotnost: $H_{oček} = 1 - (p^2 + q^2)$ pro 1 lokus se 2 alelami s četností p a q

Hardy-Weinbergova rovnováha

Př. Jeden lokus se 2 alelami

Alela	Četnost alely
A	p
a	q

$$p + q = 1$$

p, q - zjistíme analýzou tkání

Genotyp	Očekávaná četnost genotypu
AA	p^2
Aa	$2pq$
aa	q^2

= Hardy-Weinbergova rovnováha

➤ četnosti genotypů zjistíme analýzou tkání

➤ odchylky od očekávaných četností $\Rightarrow \chi^2$ test

Genetická struktura populací drift, mutace a migrace

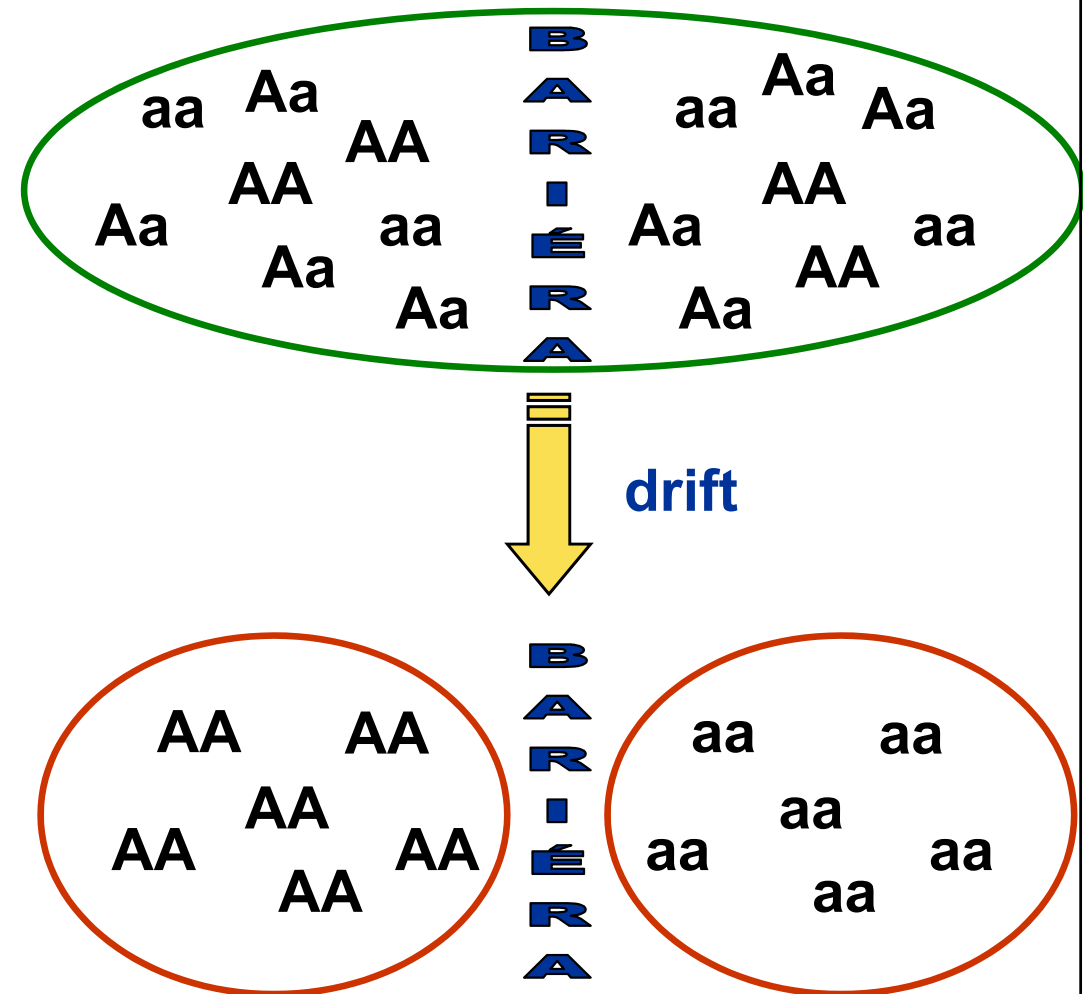
- **Drift**

→ diferenciaci subpopulací díky fixaci alternativních alel

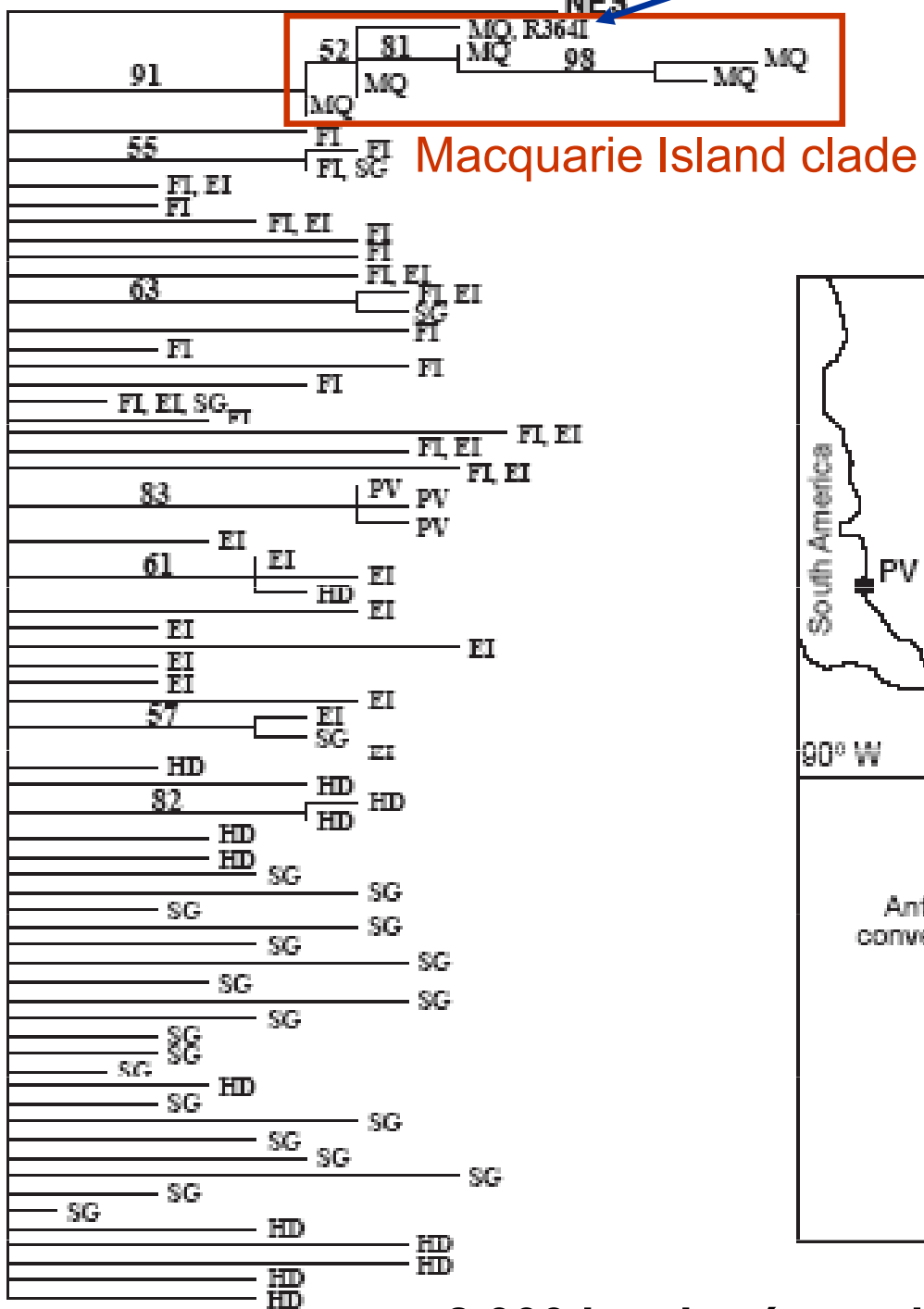
- **Mutace**

mohou zvýšit diferenciaci (odlišit subpopulace) ale riziko homoplázií

- Diferenciaci „pokazí“ **migrace**
1 migrant na generaci může stačit k setření rozdílů!

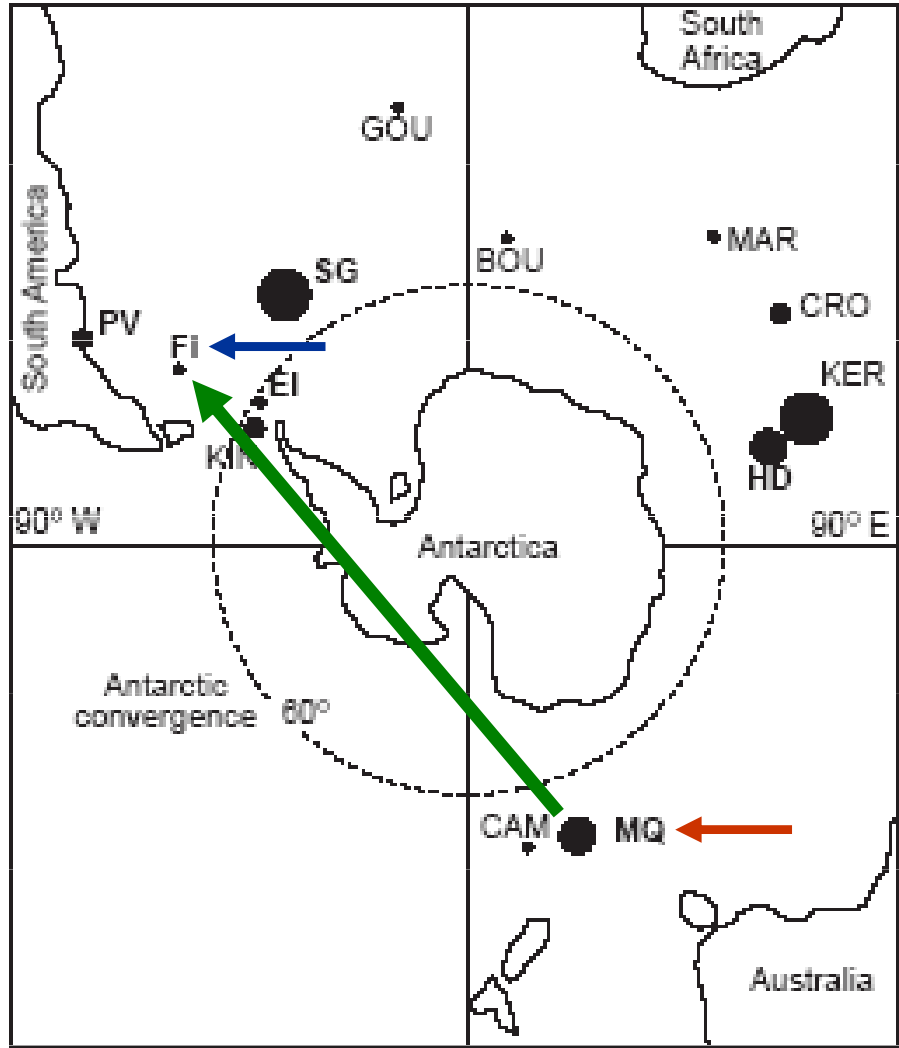


B Samec z Falkland



Macquarie Island clade

Mirounga leonina
Fabiani et al. 2003

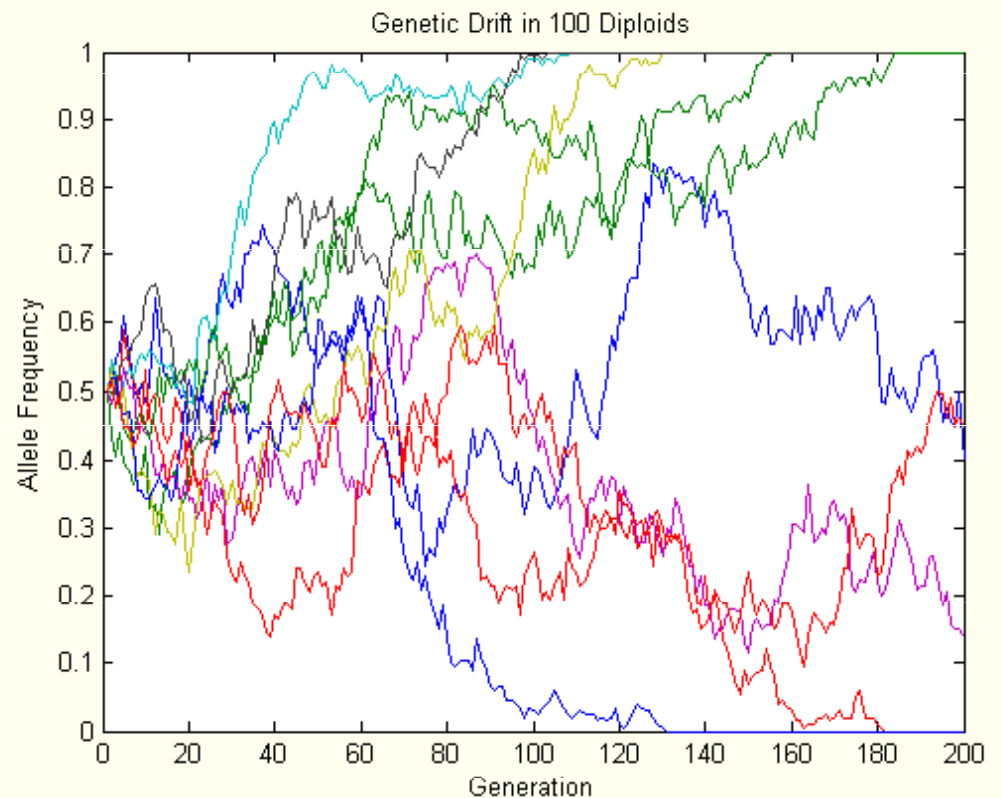
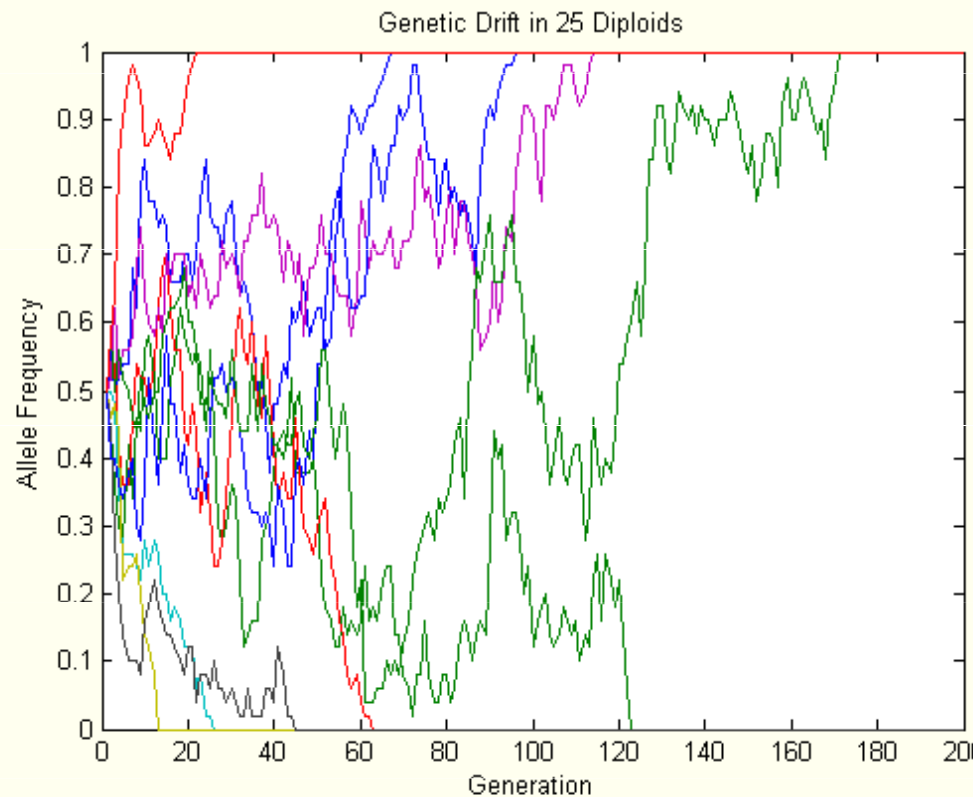


8 000 km, harém, nejméně 19 potomků (testováno mikrosatelity)

5 changes

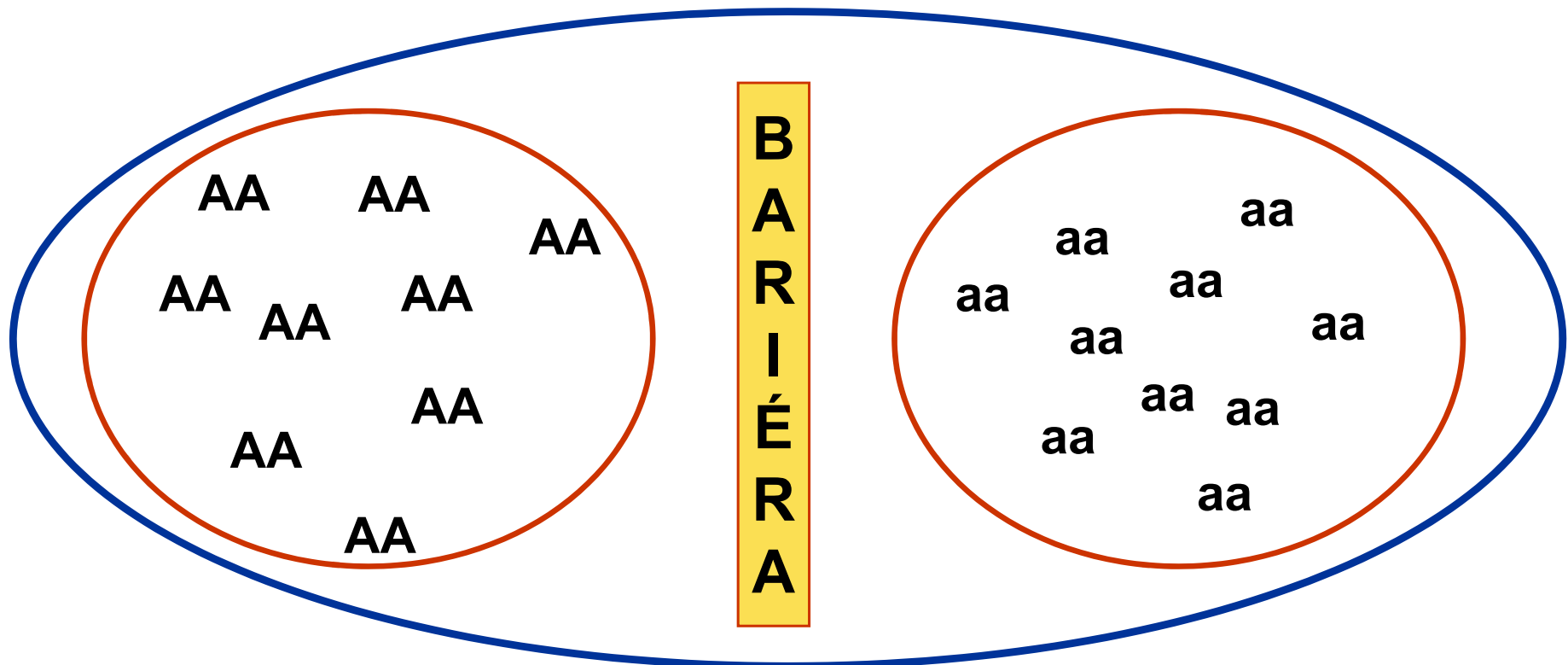
Genetická struktura populací

drift, mutace a migrace



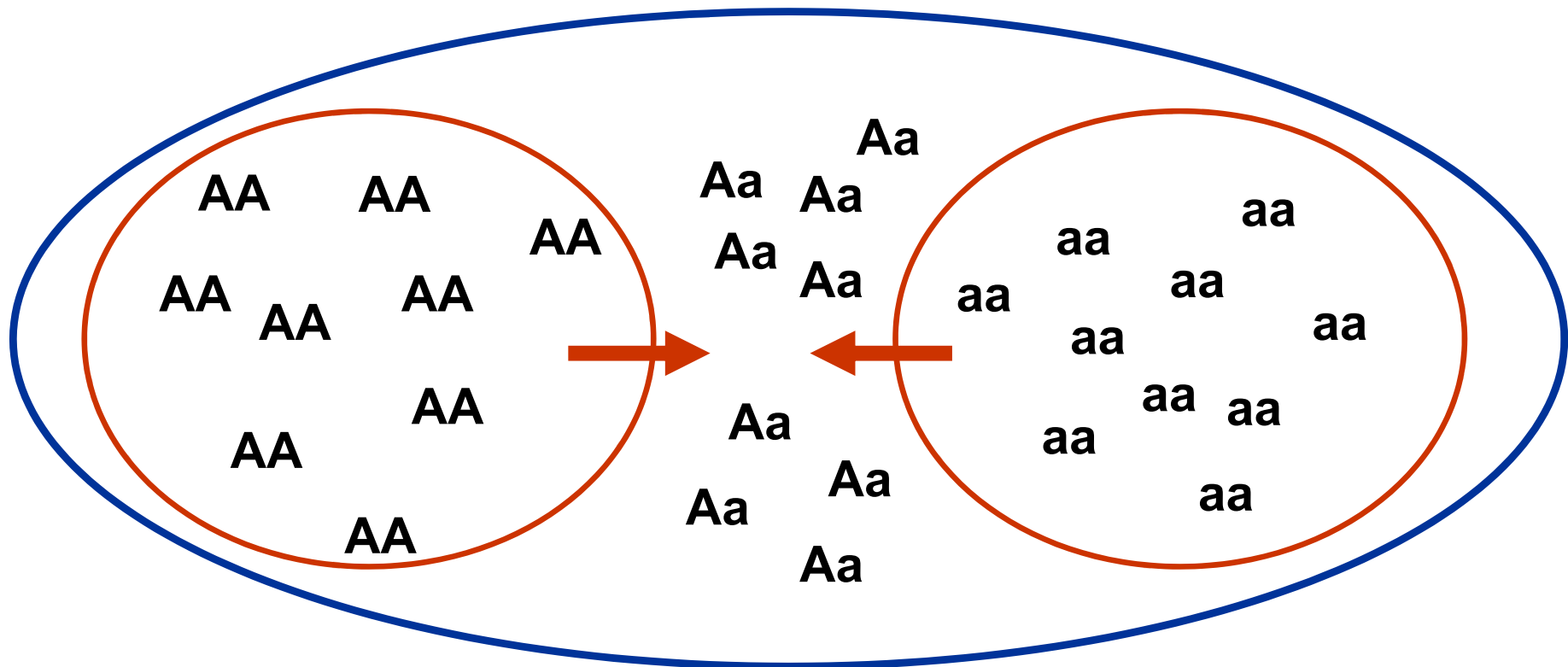
Vliv populační struktury na heterozygotnost

- Extrémní příklad
- Dvě izolované subpopulace s fixovanými alelami
- Subpopulace v HW, celkově v populaci však nedostatek heterozygotů



Wahlundův princip (isolate breaking)

- Pokles homozygotnosti při sloučení subpopulací



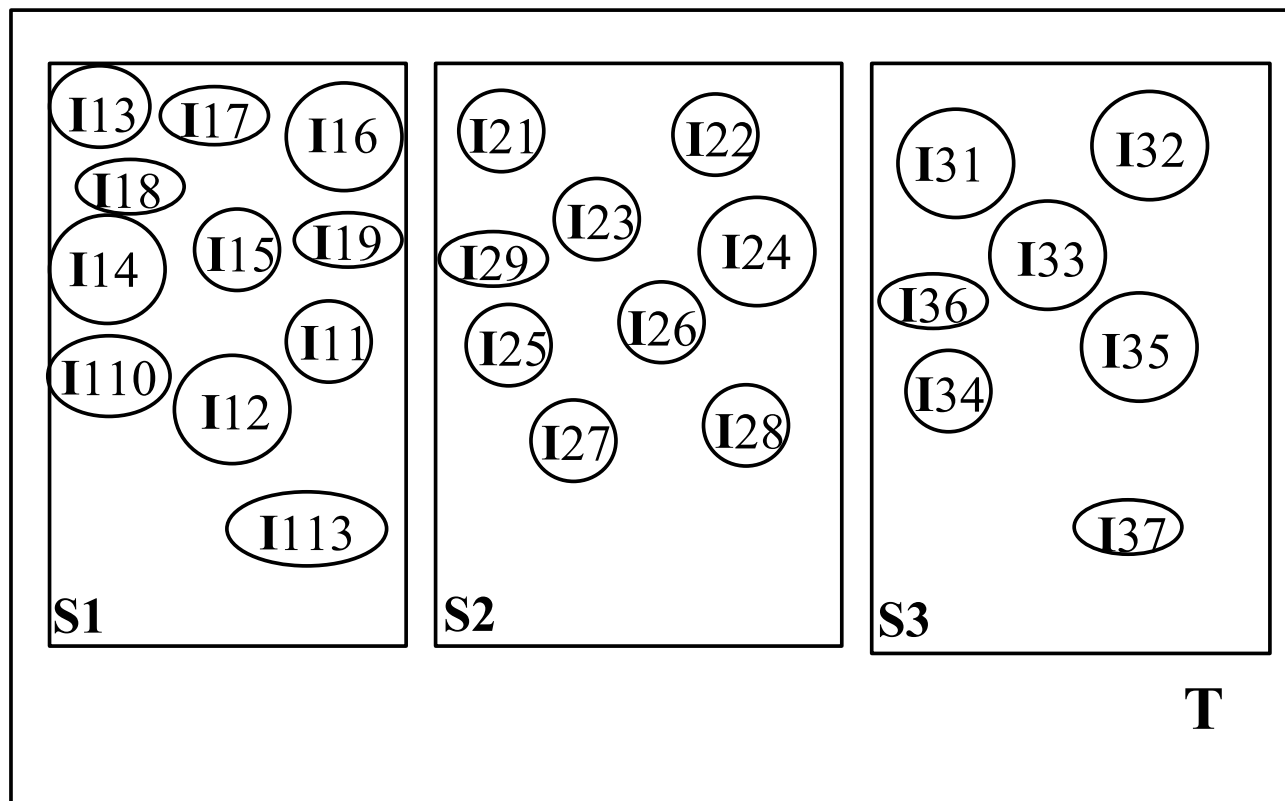
F-statistika



- Wright, Nei F_{IS} , F_{ST} , F_{IT}
- Popisují heterozygotnost (odchylky od HW) na různých měřítkách

Odhad vlivu populační struktury na genetický make-up populace

Celková populace



- 3 úrovně (T, S, I)
- x subpopulací ($x = 1$ až k ; zde $k = 3$)
- každá subpopulace má N_x jedinců
- AA, Aa, aa – odlišný symbol
- př. I1-13 = 13. jedinec z první subpopulace

Koncept heterozygotnosti

H_I – průměrná pozorovaná heterozygotnost jedince v subpopulaci

H_S - očekávaná heterozygotnost jedince v subpopulaci za předpokladu náhodného páření

H_T - očekávaná heterozygotnost jedince v celé populaci za předpokladu náhodného páření

$$H_I = \sum_{x=1}^k H_x / k \quad H_x = \text{pozorovaná heterozygotnost v subpopulaci } x$$

$$H_S = 1 - \sum_{i=1}^j p_{i,x}^2 \quad p_{i,x}^2 = \text{frekvence } i\text{-té alely v subpopulaci } x \quad \bar{H}_S = \sum_{x=1}^k H_S / k \quad \text{průměrná oček. heterozygotnost v populaci}$$

$$H_T = 2p_0q_0$$

- pouze pro dvě alely na jednom lokusu (Wright 1931)
- pro více alel je výpočet složitější (Nei 1987)

F statistiky

$$F_{IS} = \frac{\overline{H}_S - H_I}{\overline{H}_S}$$

Snížení heterozygotnosti jedince kvůli nenáhodnému páření v subpopulaci

$$F_{ST} = \frac{H_T - \overline{H}_S}{H_T}$$

Vliv rozdělení populace na subpopulace (genetický drift)

$$F_{IT} = \frac{H_T - H_I}{H_T}$$

Celkový koeficient inbreedingu F_{IT} - měří redukci heterozygotnosti jedince ve vztahu k celkové populaci

$$(1-F_{IT}) = (1-F_{ST})(1-F_{IS})$$

Weir & Cockerham (1984) **f, θ, F**

Korekce na velikost vzorku a počet subpopulací

Výpočet F statistik - příklad

Průměrná frekvence alely A v celé populaci

Lokus	Subpopulace 1 (N ₁ =40)				Subpopulace 2 (N ₂ =20)				p _{0(i)}	Pozn.
	AA	AB	BB	p _{1(i)}	AA	AB	BB	p _{2(i)}		
A	10	20	10	0.5	5	10	5	0.5	0.5	H.-W. rovnováha
B	16	8	16	0.5	4	4	12	0.3	0.4	deficit heterozygotů
C	12	28	0	0.65	6	12	2	0.6	0.625	přebytek heterozygotů
D	0	0	40	0.0	20	0	0	1.0	0.5	alternativně fixované alely

Výpočet alelových frekvencí

Lokus	Pozorovaná heterozygotnost		Očekávaná heterozygotnost			Wrightova F-statistika		
	H _{1(i)}	H _{2(i)}	H _{I(i)}	H _{S(i)}	H _{T(i)}	F _{IS(i)}	F _{ST(i)}	F _{IT(i)}
A	0.5	0.5	0.5	0.5	0.5	0.0	0.0	0.0
B	0.2	0.2	0.2	0.46	0.48	0.565	0.042	0.583
C	0.7	0.6	0.65	0.4675	0.46875	-0.39	0.0027	-0.387
D	0.0	0.0	0.0	0.0	0.5	---	1.0	1.0
Průměr						0.058	0.261	0.300

Průměrné hodnoty F statistik mohou maskovat odlišnou evoluční historii na různých lokusech

Výpočet

- Software FSTAT, GENEPOP, GENETIX...
- S.E. - jackknifing
- Průkaznost odchylky od nulové hodnoty – permutační (randomizační) testy

F-statistika

- F_{IS} snížení heterozygotnosti v lokální subpopulaci
vysoké hodnoty – inbreeding
- F_{IT} souhrnná hodnota, heterozygotnost v celé populaci
- F_{ST} míra „rozdělenosti“ = snížení toku mezi subpopulacemi
Vliv driftu – fixuje odlišné alely v subpopulacích

Hodnoty F_{ST}

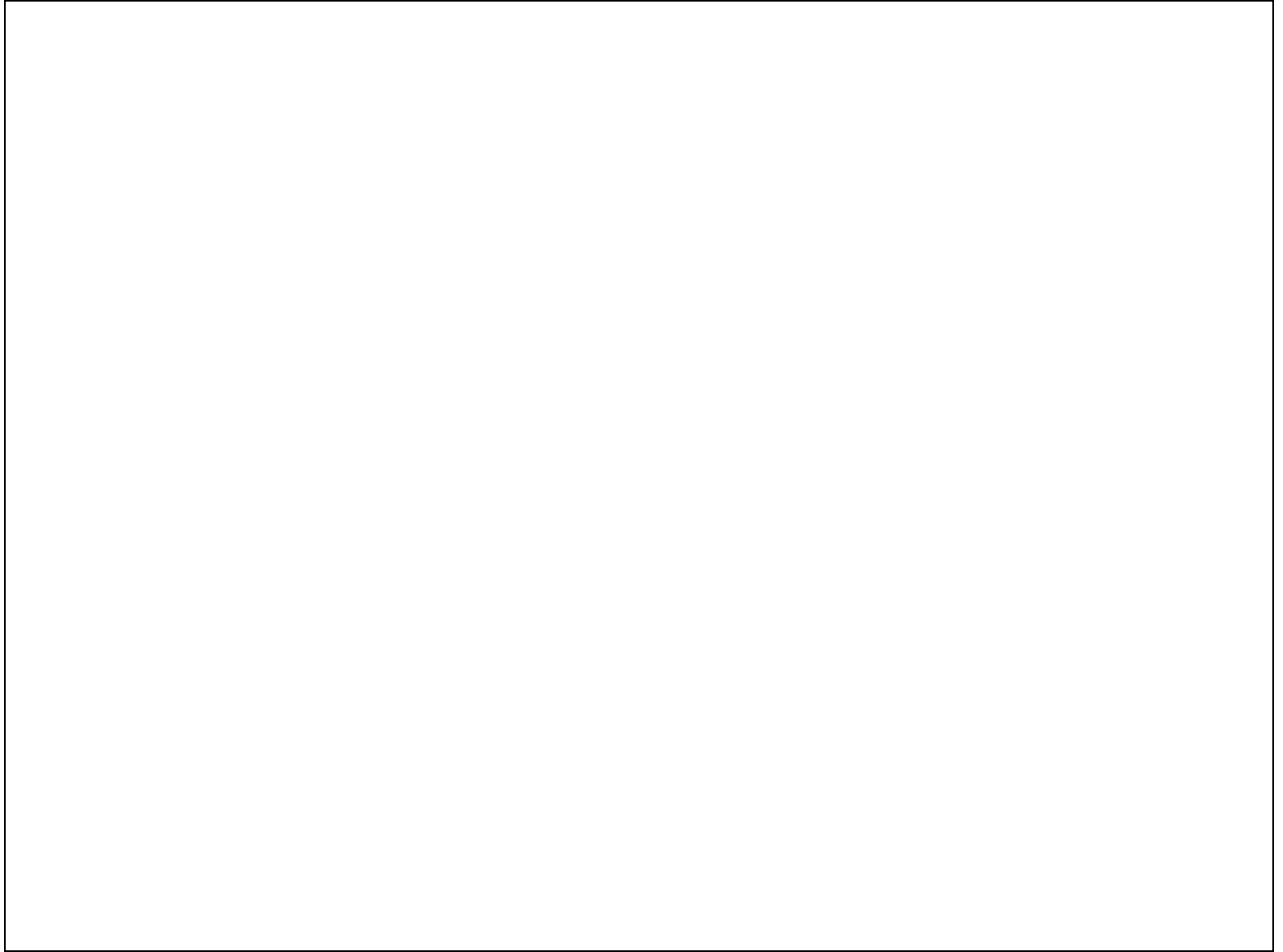
- 0 – 0.05 malá diferenciace (zanedbatelná)
- 0.05 – 0.15 střední
- 0.15 – 0.25 velká
- > 0.25 velmi velká

Příklad: Hořavka duhová (*Rhodeus sericeus*)



téměř 600 jedinců, 21 lokalit, 12 mikrosatelitů

Bryja & Reichard, nepubl. data





Emberiza citrinella

Lee et al. 2001



- 7 mikrosatelitových lokusů, F_{ST}
- Průkazné rozdíly mezi populacemi
- Ve Velké Británii není panmiktická populace

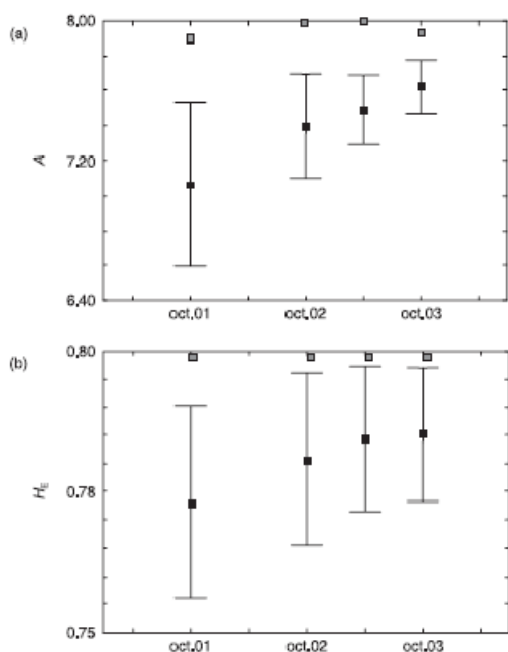
	DEV	GWE	CUM	OXO	YOR	T & W	LEI	SUF
DEV	—	NS	1, 3, 5, 6, 7	4, 5	1, 5	5	1, 5, 6	NS
GWE	0.0055	—	3	NS	NS	3, 5	5	NS
CUM	<u>0.0171*</u>	0.0041	—	5, 6	NS	1, 3, 5	1, 5, 6	3, 5
OXO	0.0111*	-0.0007	0.009*	—	NS	NS	1	3
YOR	0.0085	-0.002	0.0095*	0.0013	—	5	1, 5	5
T & W	0.0087	0.0117*	<u>0.0251**</u>	0	0.004	—	1, 3	5
LEI	<u>0.0238***</u>	0.0061	<u>0.0267***</u>	0.0083	0.0062	<u>0.0174*</u>	—	1, 5, 6
SUF	0.0073	0.0026	0.0188**	0.0047	0.005	0.0107*	<u>0.0222*</u>	—

Populační cykly hrabošovitých hlodavců (Berthier et al. 2006)

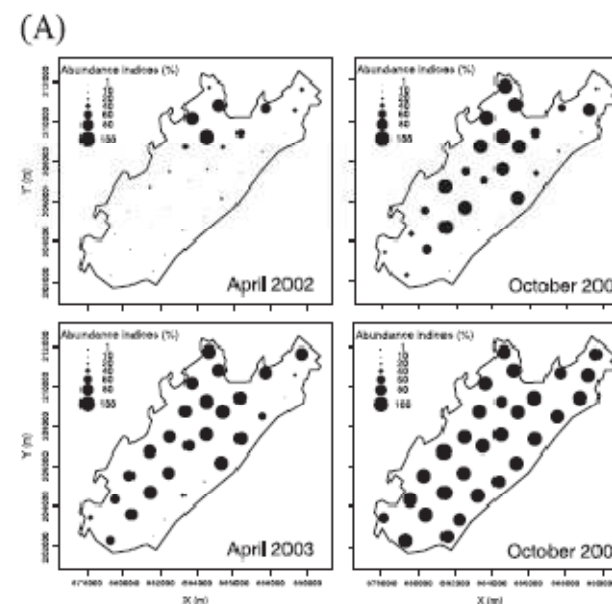


Arvicola terrestris

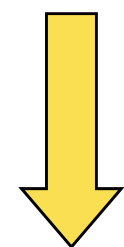
- Výrazné cyklické změny v početnosti populací – střídají se období „izolovanosti“ s „kontinuálním rozšířením“



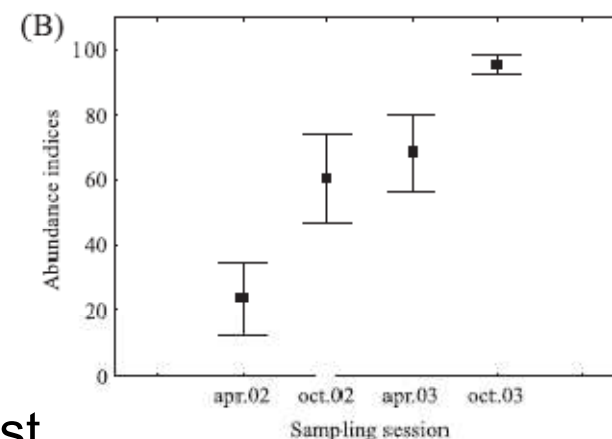
- Nízká hustota – extrémní vliv driftu
- S nárůstem populační hustoty se zvyšuje intenzita toku genů mezi subpopulacemi a narůstá genetická variabilita subpopulací



Sampling session	Min. pairwise F_{ST}	Max. pairwise F_{ST}	Overall F_{ST}
October 2001	0.018	0.058	0.037
October 2002	0.010	0.047	0.026
April 2003	0.002	0.033	0.020
October 2003	0.007	0.035	0.017



Snižování F_{ST}





Myotis bechsteinii

Kerth et al. 2002

Table 2 Genetic differentiation based on nuclear and mitochondrial DNA among 10 maternity colonies of the Bechstein's bat

Loci	No. of alleles	$F_{ST} \pm SE$	Population differentiation
Nuclear DNA			
b15	23	0.023 ± 0.009	$P < 0.0005$
b22	12	0.003 ± 0.005	$P = 0.07$
b23	20	0.008 ± 0.007	$P = 0.02$
MM5	10	0.010 ± 0.006	$P = 0.01$
NNS	10	0.031 ± 0.012	$P = 0.001$
p20	20	0.014 ± 0.007	$P < 0.0005$
paur3	10	0.012 ± 0.008	$P = 0.002$
All		0.015 ± 0.003	$P < 0.0005$
Mitochondrial DNA			
AT-1	10	0.658 ± 0.058	$P < 0.0005$
AT-2	3	0.961 ± 0.050	$P < 0.0005$

Wright's fixation index F_{ST} ($\pm SE$) was calculated for seven nuclear and two mitochondrial microsatellite loci. Significant deviations from zero, indicating population differentiation, were assessed using permutation tests (see Materials and methods for details).

- Letní kolonie samic (15-40 jedinců)
- Vzorky - jen ♀♀
- Nukleární i mitochondriální mikrosatelity
- ♀♀ filopatrie
- ♂♂ disperze

G_{ST} (Nei 1973)

- Analogie F_{ST}
- Haploidní (haplodiploidní) organismy, sekvence mtDNA
- Počítá s „gene diversity“ místo heterozygotnosti
- Pracuje tedy jen s frekvencemi alel, ne s procentem heterozygotů

R_{ST}

- Obdoba F_{ST}
- Pracuje však s velikostí alel (počet repeatů u mikrosatelitů)
- Předpoklad známého mutačního modelu jen při platnosti SMM (*stepwise mutation model*)
- „Paměť“ mutací v minulosti
 - $R_{ST} > F_{ST}$ větší vliv mutací
 - $R_{ST} = F_{ST}$ větší vliv driftu
- Potvrzení významnosti rozdílu randomizačními testy (Hardy et al. 2003, program SPAGeDi 1.1)

AMOVA

Excoffier et al. 1992

Arlequin ver. 2.000

A software for
population genetics
data analysis

Authors:

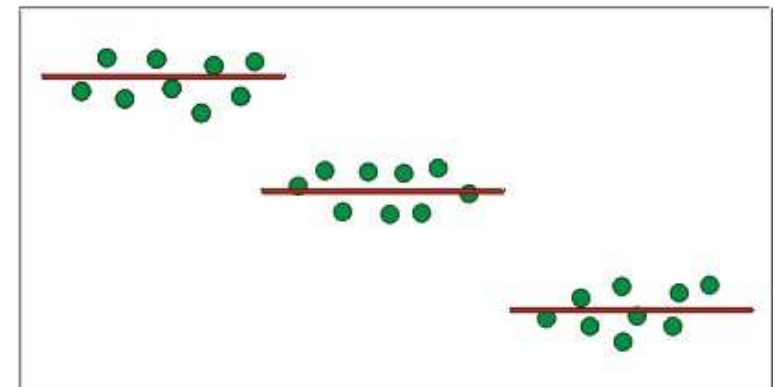
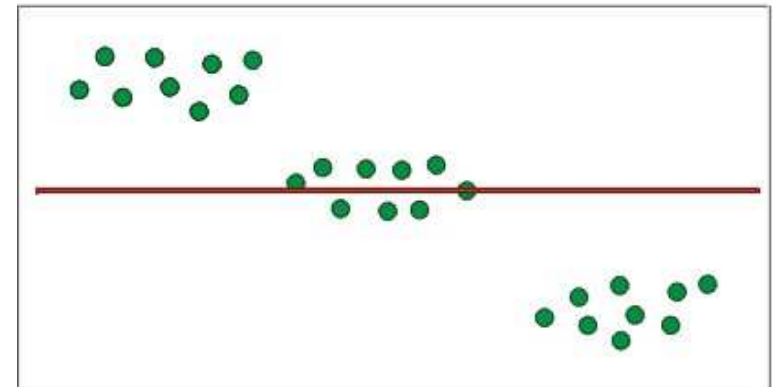
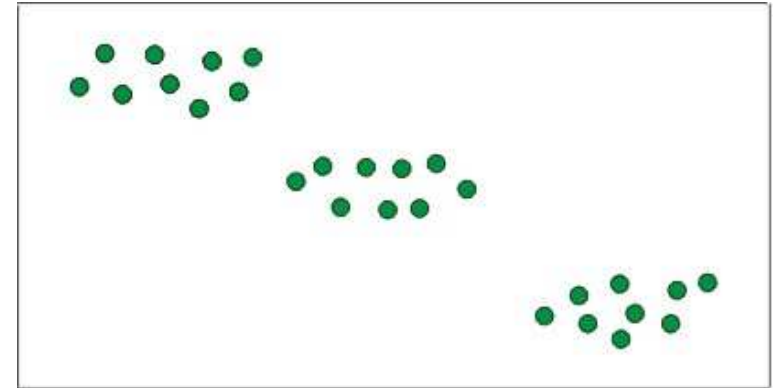
Stefan Schneider
David Roessli
Laurent Excoffier

Contact Arlequin:

Url: <http://anthropologie.unige.ch/arlequin/>
Mail: arlequin@sc2a.unige.ch



- **A**nalýza **M**olekulární **V**ariance
- Analogie R_{ST} a spojení s ANOVA
- Analýza variance alelických frekvencí
(již dříve Cockerham & Weir
1987, 1993)
- Započítává se ale rozdílnost (mutace)
alel
- Φ_{ST} , Program ARLEQUIN (umí i F_{ST})
- Data:
sekvence
mikrosatelity (jen při platnosti SMM
stepwise mutation model)



Lagopus lagopus scoticus



Piertney et al. 1998 a 2000

- Mikrosatelity a mtDNA
- R_{ST} , AMOVA
- NcDNA strukturovanost
- mtDNA není strukturovanost
- **Filopatrie samců, disperze samic, průkazný vztah R_{ST} a vzdálenosti**

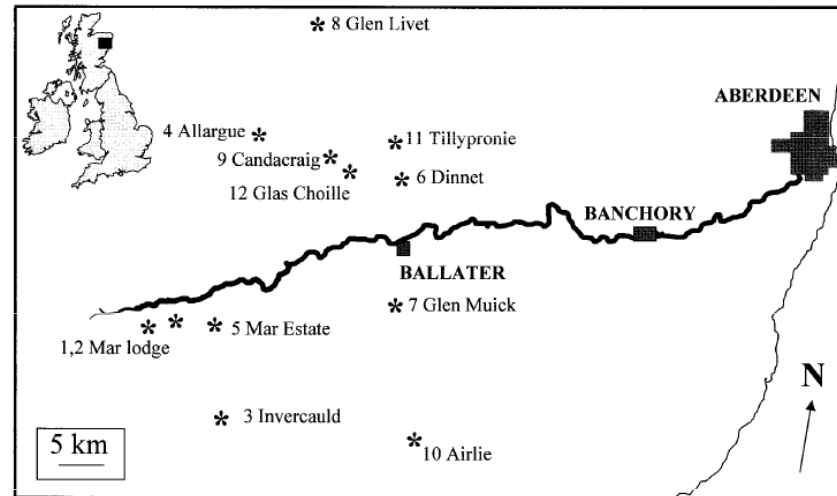
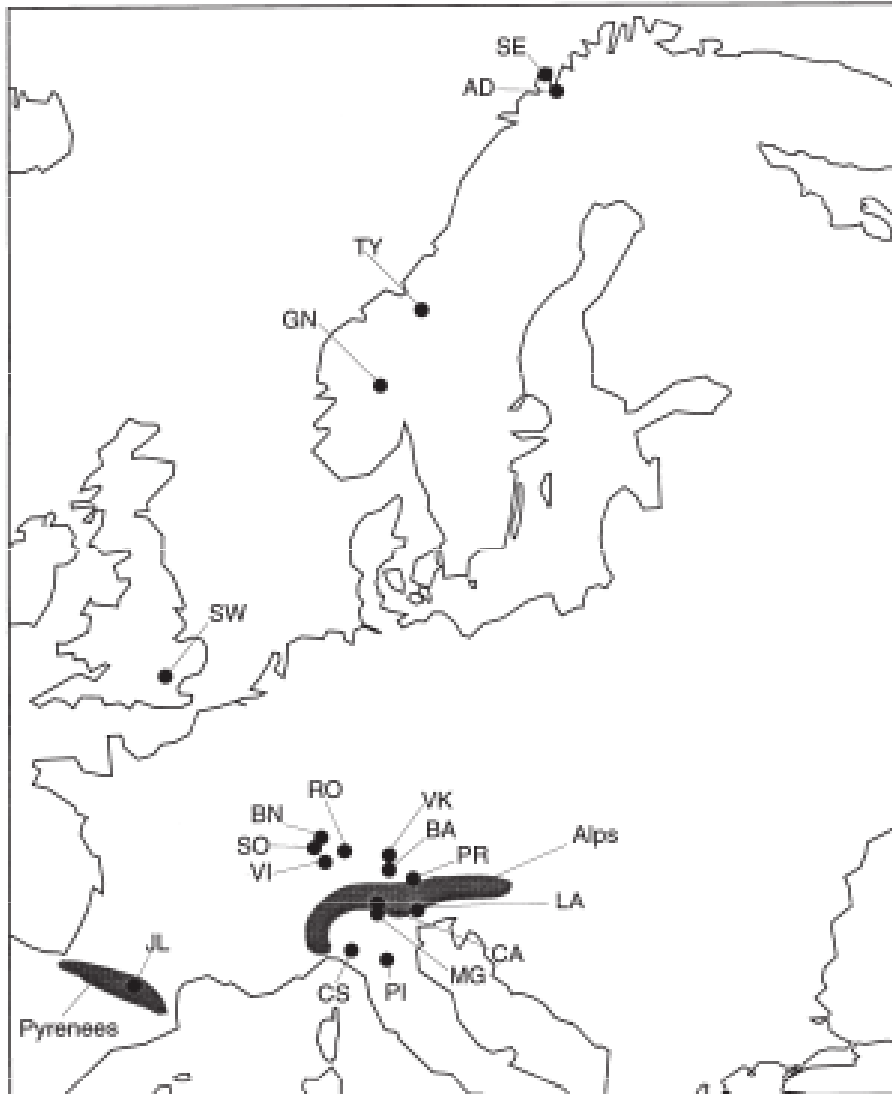


Table 3 Unbiased, multilocus estimates of Slatkin's (1995) R_{ST} between all population pairs. Significant values are underlined

	Mar Lodge 1	Mar Lodge 2	Invercauld 1	Invercauld 2	Allargue	Mar Estate	Dinnet	Glen Muick	Glen Livet	Tillypronie	Candacraig 1	Candacraig 2	Airlie
Mar Lodge 1													
Mar Lodge 2	0.039												
Invercauld 1	<u>0.227</u>	0.060											
Invercauld 2	-0.020	-0.011	<u>0.117</u>										
Allargue	0.048	<u>0.174</u>	<u>0.380</u>	0.075									
Mar Estate	<u>0.152</u>	<u>0.285</u>	<u>0.477</u>	<u>0.177</u>	0.009								
Dinnet	0.017	<u>0.137</u>	<u>0.352</u>	0.040	-0.016	<u>0.049</u>							
Glen Muick	-0.032	0.009	<u>0.153</u>	-0.033	0.024	<u>0.110</u>	-0.002						
Glen Livet	<u>0.181</u>	0.309	<u>0.491</u>	<u>0.204</u>	0.028	-0.012	<u>0.074</u>	<u>0.137</u>					
Tillypronie	0.328	<u>0.448</u>	<u>0.619</u>	<u>0.334</u>	<u>0.117</u>	0.026	<u>0.195</u>	<u>0.252</u>	-0.004				
Candacraig 1	0.332	<u>0.458</u>	<u>0.641</u>	<u>0.333</u>	<u>0.100</u>	0.009	<u>0.185</u>	<u>0.243</u>	-0.017	-0.039			
Candacraig 2	0.236	<u>0.374</u>	<u>0.565</u>	<u>0.254</u>	0.040	-0.019	<u>0.103</u>	<u>0.173</u>	-0.030	-0.025	-0.045		
Airlie	<u>0.249</u>	<u>0.384</u>	<u>0.574</u>	<u>0.263</u>	0.051	-0.008	<u>0.166</u>	<u>0.182</u>	-0.020	-0.012	-0.028	-0.039	
Glas Choille	-0.001	<u>0.040</u>	<u>0.208</u>	-0.013	<u>0.044</u>	<u>0.136</u>	0.017	-0.025	<u>0.163</u>	<u>0.239</u>	0.291	<u>0.209</u>	<u>0.220</u>

Bombus pascuorum

Widmer & Schmid-Hempel 1999



	F/Φ	d.f.	SSD†	Variance component	% Total variance*
Among populations	F	17	77.71	0.07	4.51*
	Φ	17	5198.20	5.02	8.74*
Among regions	F	4	56.15	0.08	5.16*
	Φ	4	3464.94	4.58	7.49*
Among populations within regions	F	11	24.35	0.02	1.11*
	Φ	11	1773.71	2.16	3.53*
Between north and south of Alps	F	1	38.57	0.11	7.12*
	Φ	1	2622.89	7.25	11.74*
Among populations north and south of the Alps, respectively	F	16	39.14	0.02	1.46*
	Φ	16	2575.31	2.18	3.53*

†Sum of squared deviations.

* $P < 0.001$.

**Mikrosatelity, AMOVA
Nejvíc vysvětlují Alpy**



Neighbour-joining

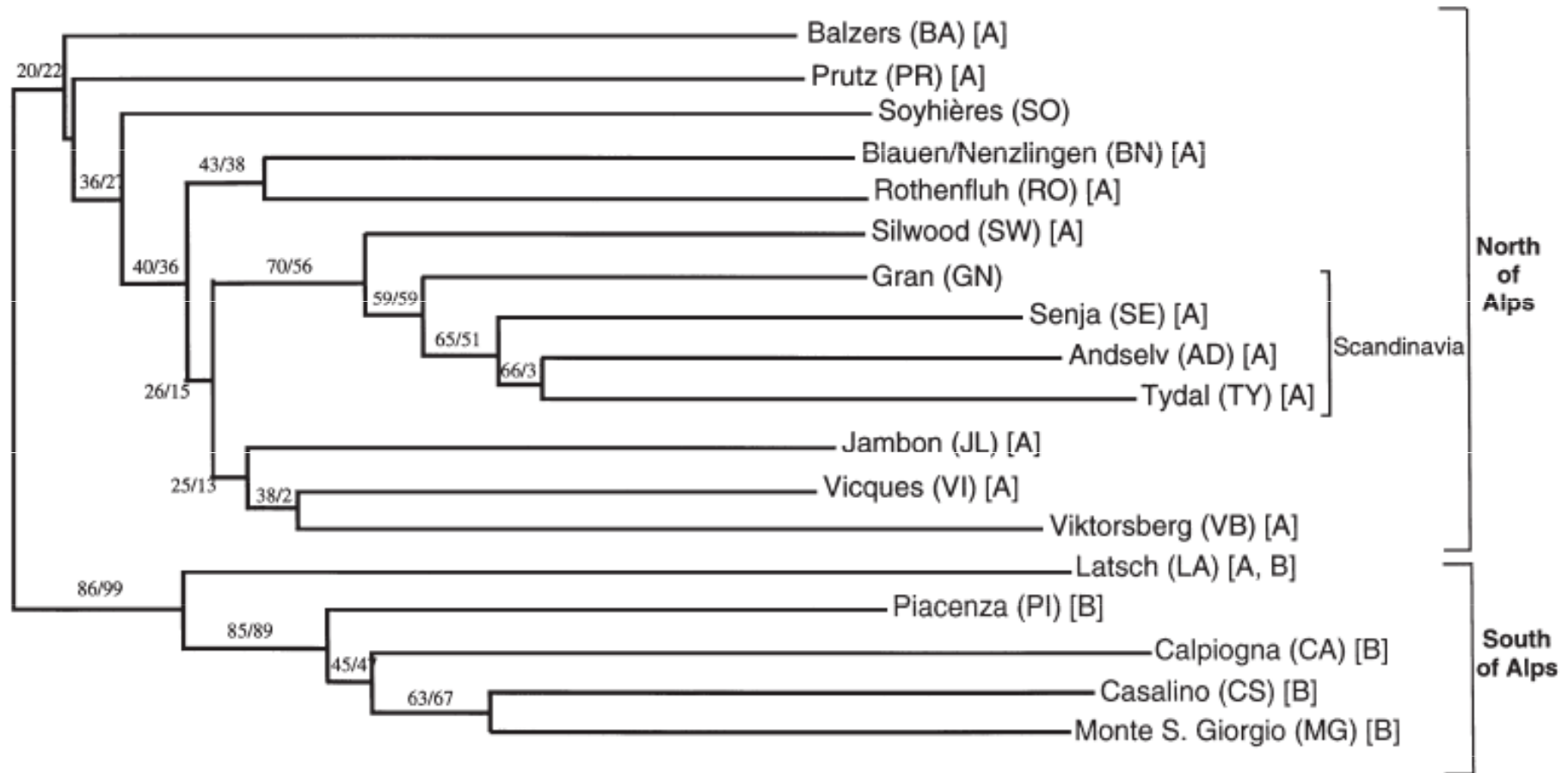


Fig. 2 Neighbour-joining (NJ) dendrogram based on the chord distance of microsatellite data (Cavalli-Sforza & Edwards 1967) connecting 18 populations of *Bombus pascuorum*. Bootstrap values (percentages) were computed over 2000 replications by resampling individuals within populations (left values) and by resampling loci (right values). Letters in parentheses denote mtDNA haplotypes.

PC1 z alelických frekvencí vs. zeměpisná délka

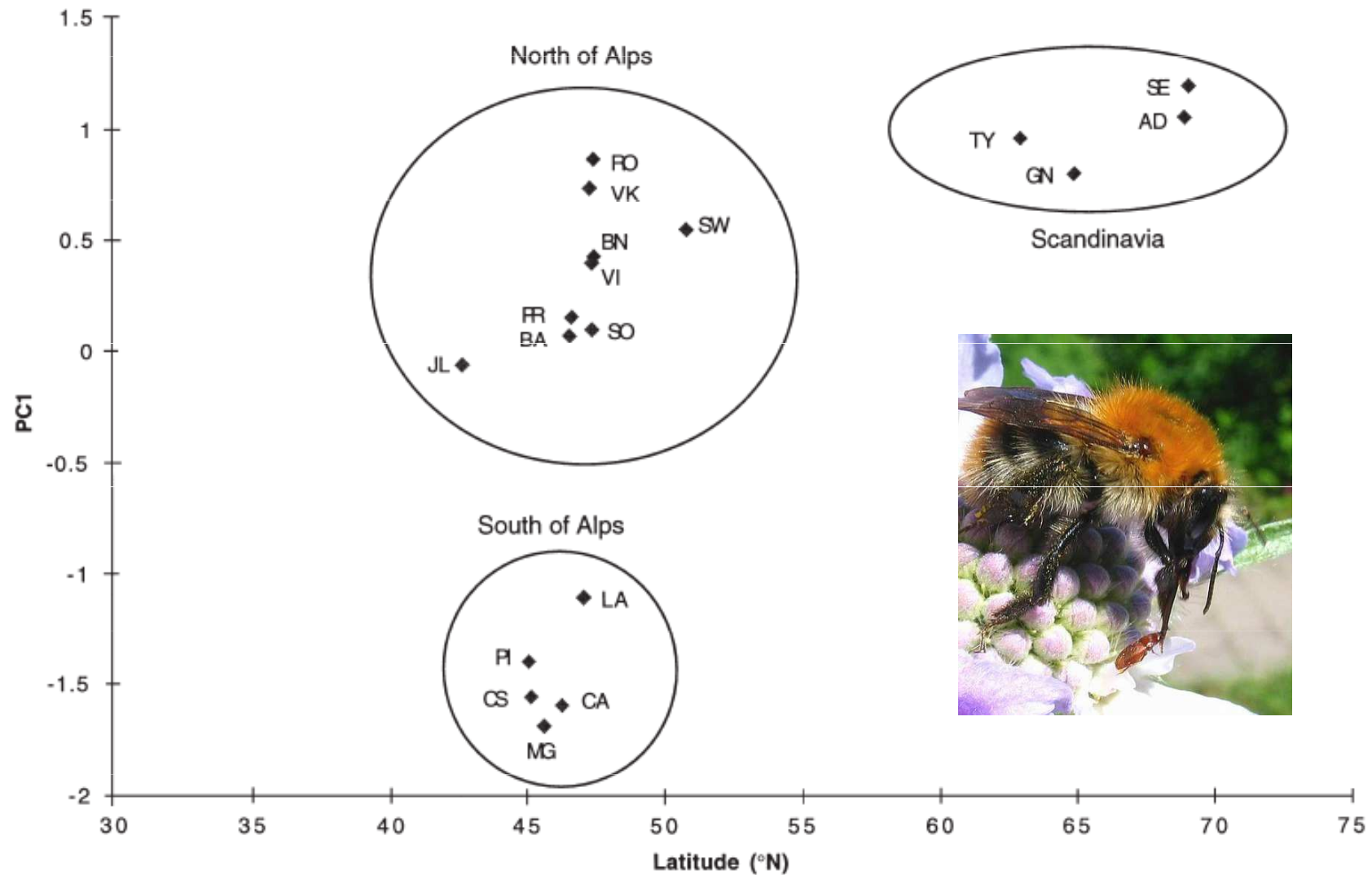
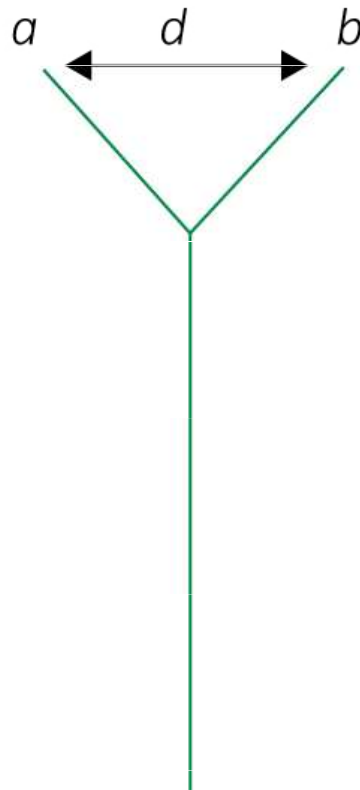


Fig. 3 The first principal components (PC1) for microsatellite allele frequencies vs. latitude. Identification codes refer to populations, see Fig. 1.

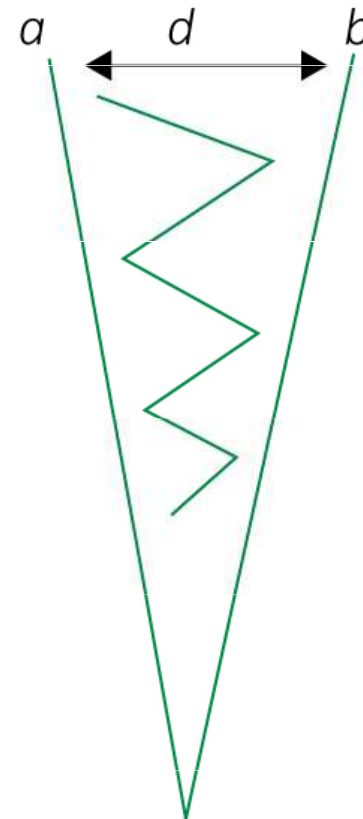
AMOVA a F statistika

popis výsledku nikoliv příčin → možná alternativní vysvětlení

Recent separation,
no gene flow



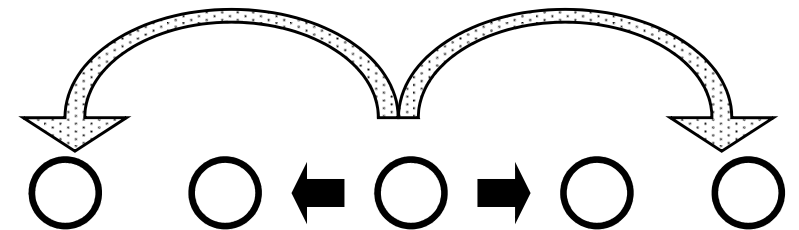
Old separation, but
continuous (low)
gene flow



Time

Modely toku genů (alel)

- **Island model** (Wright)
Stejně velké subpopulace
Symetrický tok genů
Stejná pravděpodobnost výměny
mezi jakýmkoliv subpopulacemi
- **Stepping-stone model**
(Kimura)
Výměna jen mezi sousedními
subpopulacemi
- **Isolation by distance**
Tok slábne se vzrůstající
vzdáleností subpopulací



Isolation by distance

- rozumné geografické měřítko
(závisí na schopnosti disperze)
- musí být ustanovena rovnováha mezi migrací a driftem
- IBD (isolation-by-distance) nebude
 - u velmi recentně izolovaných populací
 - u zcela izolovaných populací
 - při značné migraci



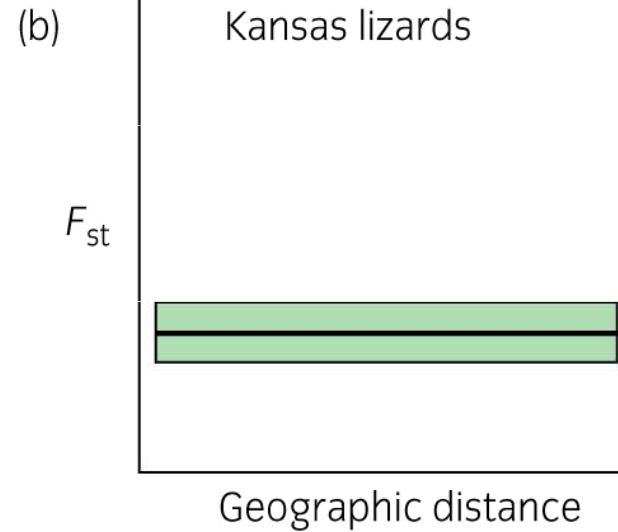
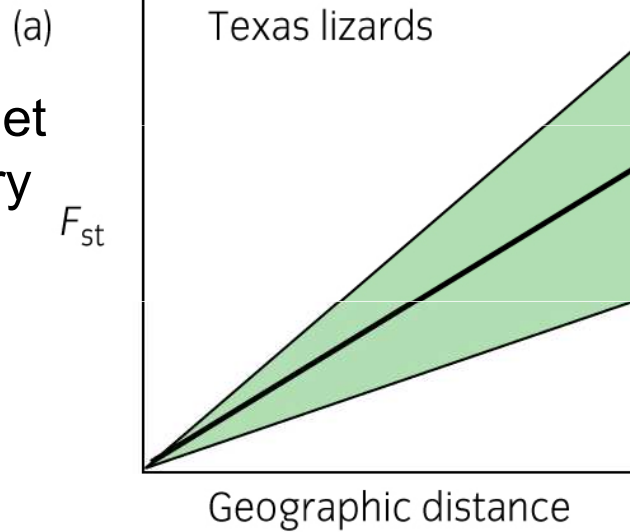
Isolation by distance

Crotaphytus collaris

Hutchinson & Templeton 1999

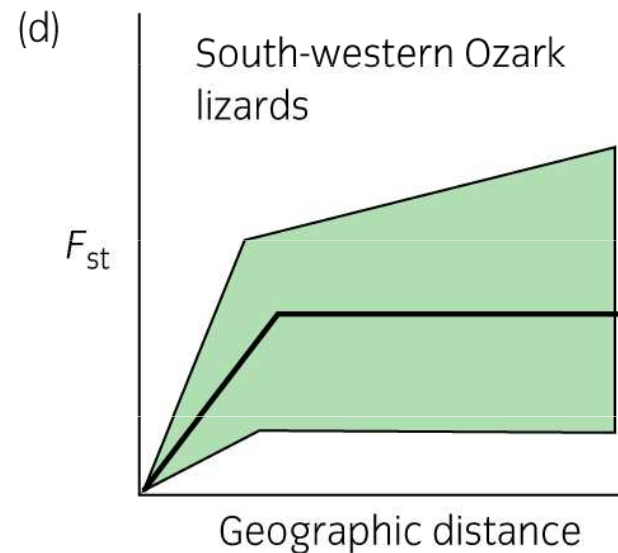
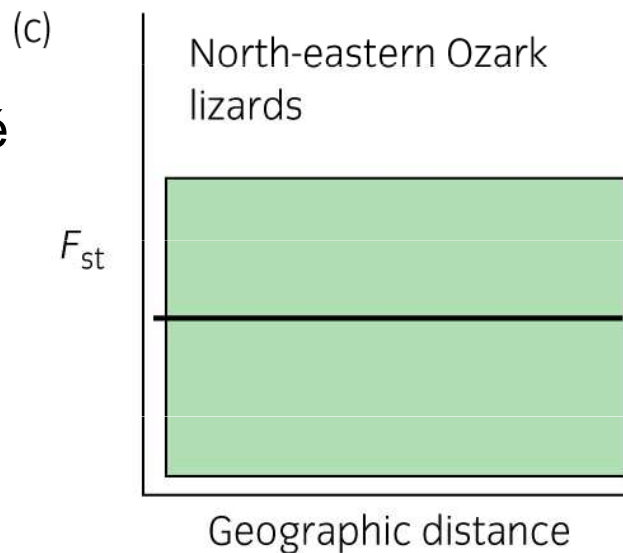


desítky tisíc let
nejsou bariéry
rovnováha
mezi driftem
a migrací



postglaciálně
nejsou bariéry
vliv migrace

postglaciálně
fragmentace
vliv driftu



postglaciálně
vzrůstající
fragmentace
vliv driftu
na velkých
měřítkách
rovnováha
na malých
měřítkách



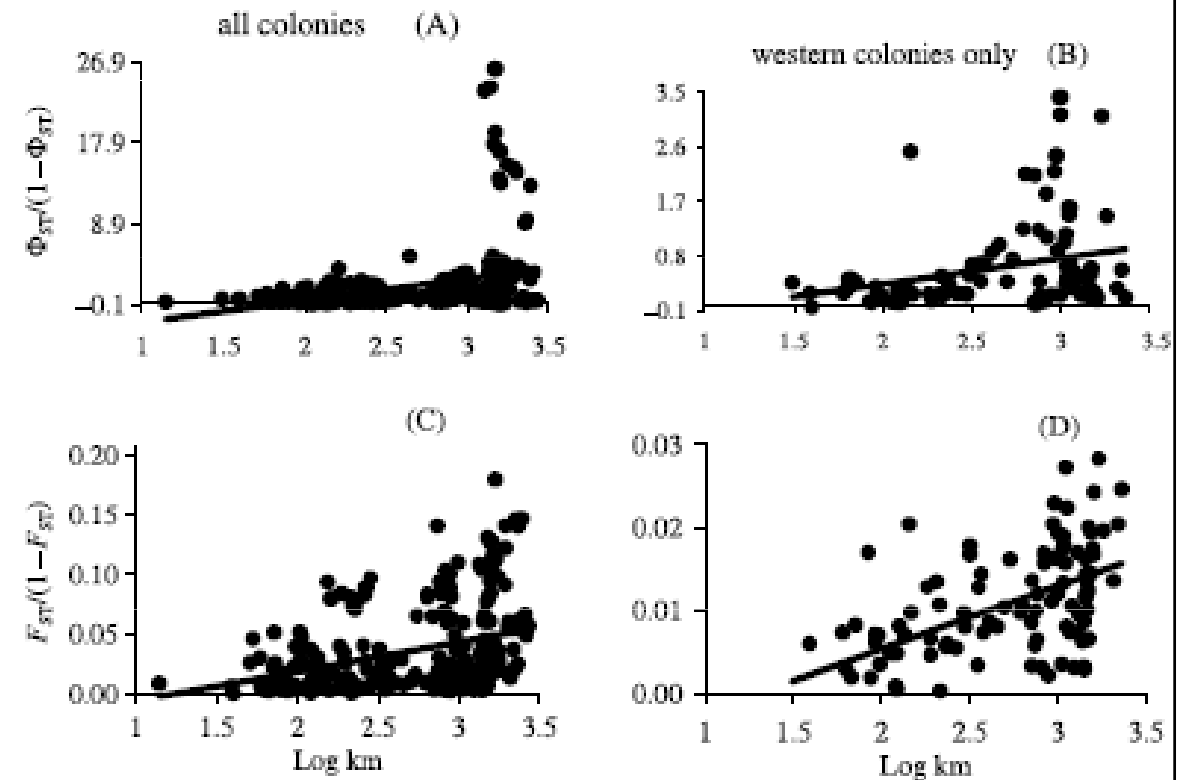
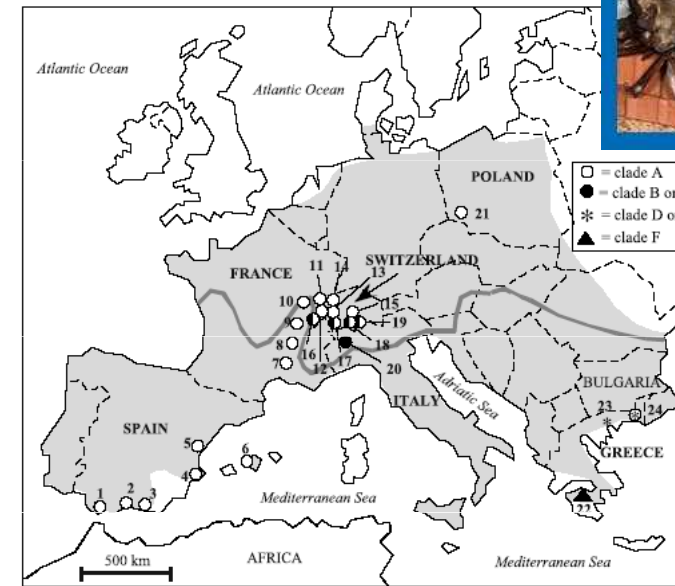
Isolation by distance

Myotis myotis

Ruedi & Castella 2003



- Regrese $\Phi_{ST}/(1-\Phi_{ST})$ nebo $F_{ST}/(1-F_{ST})$ oproti logaritmu geografické vzdálenosti
- Průkaznost testována permutacemi
- Mantelův test (FSTAT)
- Lépe vyšly mikrosatelity než mtDNA



Tok genů – metody

Nm = počet migrantů / subpopulace / generace

Jde o hrubý odhad ve škále: málo, středně, hodně!

- **Privátní alely** (Slatkin)

Alely vyskytující se jen v jedné subpopulaci

$p(1)$ frekvence privátních alel

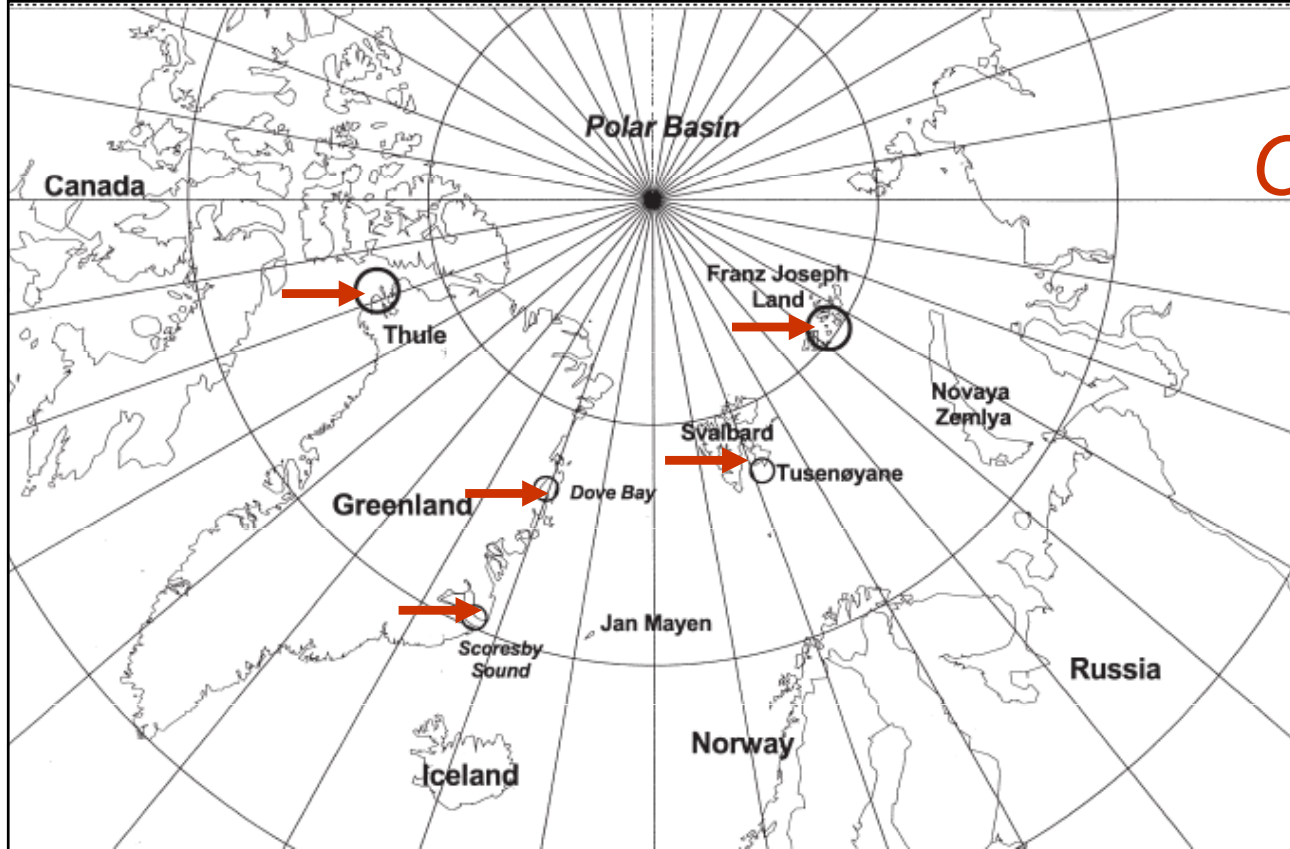
$$\ln p(1) = -0,505 \ln(Nm) - 2.44$$

- **F statistika**

$$F_{ST} = \frac{1}{1 + 4N_m}$$

Co ale rozhodně NE!

- Dvě hodně vzdálené populace
- F_{ST}
- Vyjde nenulové Nm
- → v minulosti došlo k výměně jedinců
- I populace, které si nikdy nevyměnily migranty mohou mít nenulové Nm .

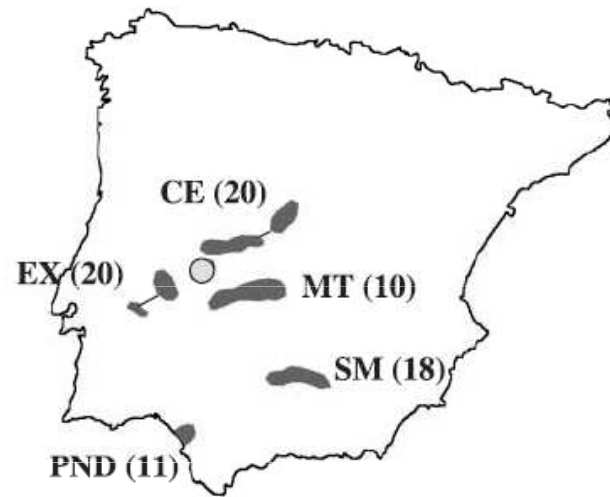


Odobenus rosmarus

Andersen et al. 1998



- Mikrosatelity
- Nejvíce izolované SZ Grónsko
- Nízké hodnoty FST (okolo 0.05)
- Několik migrantů za generaci (asi 4)
- Shodně vyšla i mtDNA



Aquila adalberti

Martínez-Cruz et al. 2004

Table 3 Differentiation indexes among the five breeding nuclei

	CE	EX	MT	PND	SM
CE		0.0227*	0.0078	0.1100***	0.0274**
EX	0.0025*		-0.0001	0.0790**	0.0263**
MT	0.0644	0.2364		0.0557**	0.0018
PND	0.0001**	0.0001**	0.0031*		0.1015***
SM	0.0008**	0.0021*	0.2441	0.0002**	

CE, Centro; EX, Extremadura; MT, Montes de Toledo; PND, Parque Nacional de Doñana; SM, Sierra Morena.

F_{ST} probability values above the diagonal and for the exact test of differentiation under the diagonal.

Levels of significance are * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$.

- Mikrosatelity (a mtDNA), F_{ST} , GENETIX, GENEPOP
- Průkazné rozdíly, hlavně Doñana a zbytek
- Průkazně vyšlo Isolation by distance
- Asi rychlý drift, malé N_e , rozdíly v reprodukčním úspěchu

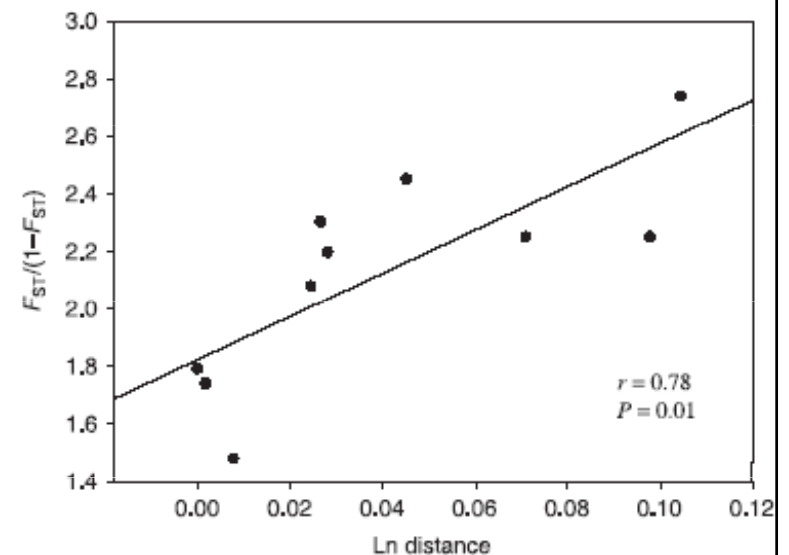


Fig. 3 Mantel test showing the correlation between genetic and geographical distances among the five nuclei ($r = 0.78$, $P = 0.01$).

Detekce bottlenecku

- **Při bottlenecku vymizí nejdříve vzácné alely**
 - Alely rozděleny do kategorií podle frekvence
 - Relativní nedostatek alel s nízkou frekvencí → bottleneck
- **Zvýšená heterozygotnost – program BOTTLENECK**
 - Nutný předpoklad mutačního modelu (SSM, TPM, IAM)
 - Testuje „mutation-drift equilibrium“
- **Bayesiánský přístup** (založen na koalescenční teorii)
 - Detekce bottlenecku i expanze
 - Předpoklad rovnováhy
 - Vhodné při dlouhodobějších změnách
 - Markov chain Monte Carlo simulations
 - Beaumont MA – např. program MSVAR

Bottleneck



- Při bottlenecku
 - redukce počtu alel
 - ovlivnění heterozygotnosti není tak rychlé
 - více heterozygotů než by vyplývalo z populačního modelu (IAM, TPM, SMM)
- Nutno definovat model, nepředpokládám HW rovnováhu
- Program BOTTLENECK

- Strnadi
- Záleží na zvoleném modelu
- Ale SMM asi neplatí stoprocentně

Table 4 Departures from mutation-drift equilibrium in yellowhammer samples

Sampling sites	IAM	TPM	SMM
DEV	0.0039**	0.0078**	0.3438
GWE	0.0039**	0.0391*	0.7656
CUM	0.0039**	0.0391*	0.8125
OXO	0.0039**	0.0117*	0.3438
YOR	0.0039**	0.0117*	0.7656
T & W	0.0078**	0.1875	0.8125
LEI	0.0195*	0.2891	0.6563
SUF	0.0039**	0.0195*	0.6563

Detekce bottlenecku

Bayesiánský přístup (založen na koalescenční teorii)

Detekce bottlenecku i expanze

Vhodné při dlouhodobějších změnách

Markov chain Monte Carlo simulations

Beaumont MA – program **Msva**

Shrnutí

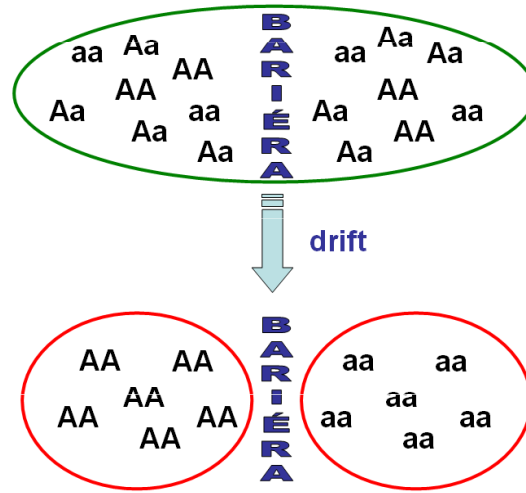
- **Klasické fixační indexy (F_{ST})**
 - Značně robustní
 - Mutačního model není nutný
 - Nutno mít rozumnou škálu (geografickou vzdálenost)
- **AMOVA**
 - Jemnější rozlišení
 - Možno použít i na sekvence (např. mtDNA)
 - Předpoklad modelu (SMM)
 - Platnost modelu většinou není známa

Genetická struktura populací drift, mutace a migrace

- **Drift**
→ diferenciace subpopulací díky fixaci alternativních alel

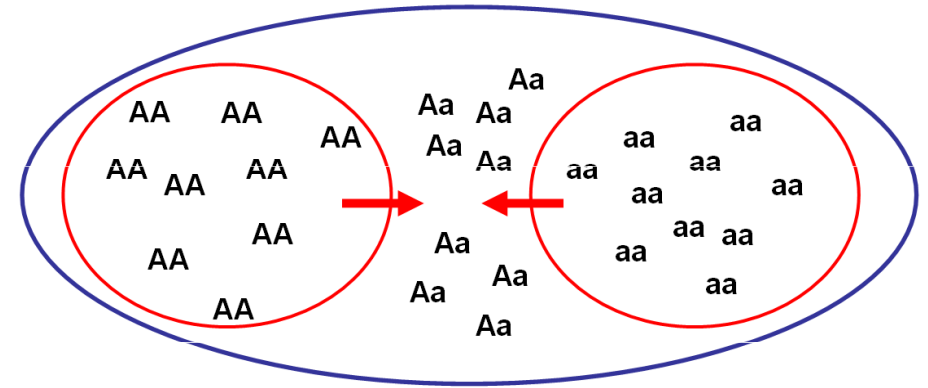
- **Mutace**
mohou zvýšit diferenciaci (odlišit subpopulace) ale riziko homoplázií

- Diferenciaci „pokazí“ **migrace**
1 migrant na generaci může stačit k setření rozdílů!



Wahlundův princip (isolate breaking)

- Pokles homozygotnosti při sloučení subpopulací



F-statistika

- Wright, Nei F_{IS} , F_{ST} , F_{IT}
- Weir & Cockerham (1984) f , θ , F
Korekce na velikost vzorku a počet subpopulací
- Popisují heterozygotnost (odchylky od HW) na různých měřítkách
- F_{IS} snížení heterozygotnosti v lokální subpopulaci vysoké hodnoty – inbreeding
- F_{IT} souhrnná hodnota, heterozygotnost v celé populaci
- F_{ST} míra „rozdělenosti“ = snížení toku mezi subpopulacemi
Vliv driftu – fixuje odlišné alely v subpopulacích



$$F_I = 1 - (h_{obs} / h_{exp})$$

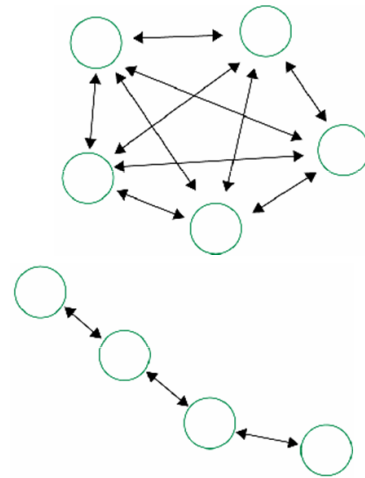
$$(1 - F_{IT}) = (1 - F_{ST})(1 - F_{IS})$$

$$R_{ST}$$

- Obdoba F_{ST}
- Pracuje však s velikostí alel (počet repeatů u mikrosatelitů)
- Předpoklad známého mutačního modelu jen při platnosti SMM (*stepwise mutation model*)
- „Paměť“ mutací v minulosti
 - $R_{ST} > F_{ST}$ větší vliv mutací
 - $R_{ST} = F_{ST}$ větší vliv driftu
- Potvrzení významnosti rozdílů randomizačními testy (Hardy et al. 2003, program SPAGeDi 1.1)

Modely toku genů (alel)

- **Island model (Wright)**
Stejně velké subpopulace
Symetrický tok genů
Stejná pravděpodobnost výměny mezi jakýmkoliv subpopulacemi
- **Stepping-stone model (Kimura)**
Výměna jen mezi sousedními subpopulacemi
- **Isolation by distance**
Tok slábne se vzrůstající vzdáleností subpopulací



Tok genů – metody

Nm = počet migrantů / subpopulace / generace
Jde o hrubý odhad ve škále: málo, středně, hodně!

- **Privátní alely**
Alely vyskytující se jen v jedné subpopulaci

$p(1)$ frekvence privátních alel
 $\ln p(1) = -0,505 \ln(Nm) - 2.44$

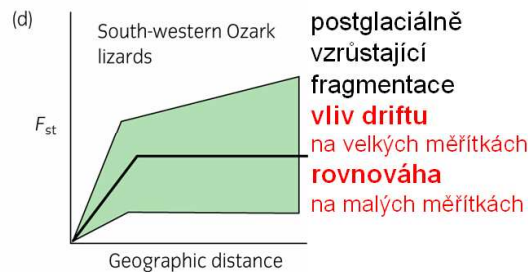
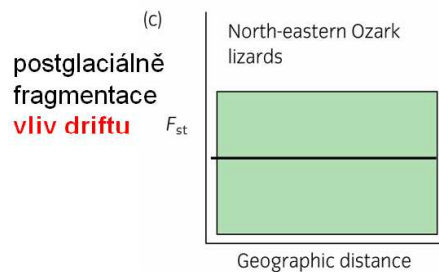
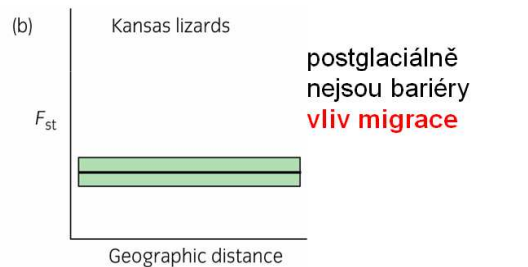
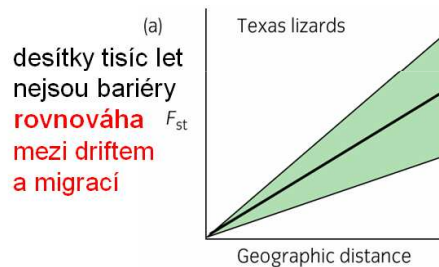
- **F statistika** $F_{ST} = \frac{1}{1 + 4N_m}$
- **Fylogeneze alel**



Isolation by distance

leguánovec obojkový
Crotaphytus collaris

Hutchinson & Templeton 1999



Myotis bechsteinii

Kerth et al. 2002

- Letní kolonie samic (15-40 jedinců)
- Vzorke - jen ♀♀
- Nukleární i mitochondriální mikrosatelity
- ♀♀ filopatrie
- ♂♂ disperze

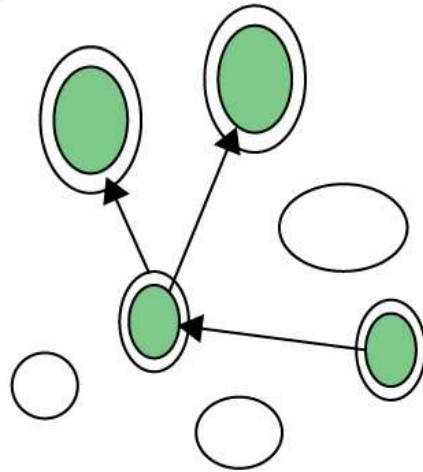
Table 2 Genetic differentiation based on nuclear and mitochondrial DNA among 10 maternity colonies of the Bechstein's bat

Loci	No. of alleles	FST ± SE	Population differentiation
Nuclear DNA			
b15	23	0.023 ± 0.009	$P < 0.0005$
b22	12	0.003 ± 0.005	$P = 0.07$
b23	20	0.008 ± 0.007	$P = 0.02$
MM5	10	0.010 ± 0.006	$P = 0.01$
NN8	10	0.031 ± 0.012	$P = 0.001$
p20	20	0.014 ± 0.007	$P < 0.0005$
paur3	10	0.012 ± 0.008	$P = 0.002$
All		0.015 ± 0.003	$P < 0.0005$
Mitochondrial DNA			
AT-1	10	0.658 ± 0.058	$P < 0.0005$
AT-2	3	0.961 ± 0.050	$P < 0.0005$

Genetika metapopulací

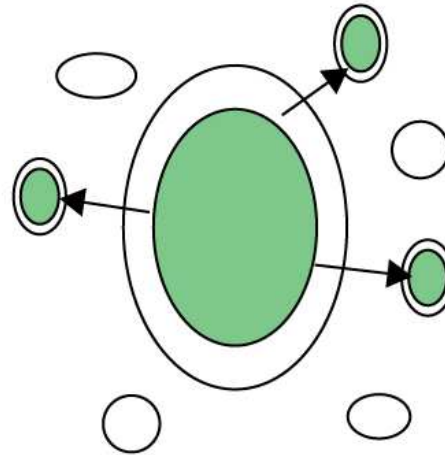
„classical“

(a)



„mainland-island“

(b)



- **Jednoduchá populační struktura vs. metapopulace** → populační genetika
- stable pattern of genetic differentiation vs. pattern vary over time

Atlantic salmon – Sainte Marguerite river (Quebec) (Garant et al. 2000)

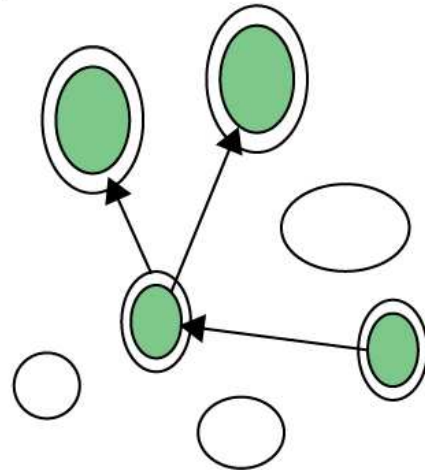
- výrazná populační struktura mezi řekami
- v rámci jedné řeky – různé mikrohabitaty obývané (1) **demy** nebo (2) **metapopulacemi**?
- dvouleté sledování, mikrosatelity, hierarchická AMOVA
- lokálně adaptované demy, ale i metapopulace (např. po povodních je mikrohabitat obsazen náhodně a ne jedinci z původního demu)
- long-term (stable) vs. short-term (metapopulations) substructures



Který typ metapopulační struktury?

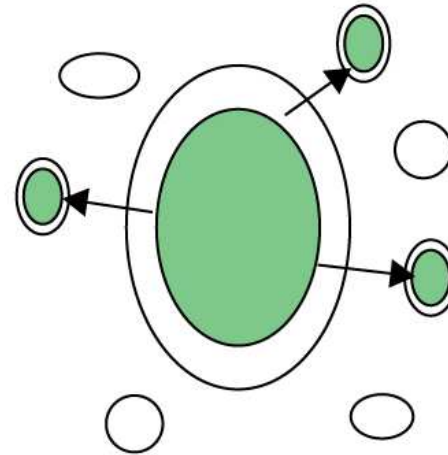
„classical“

(a)



„mainland-island“

(b)

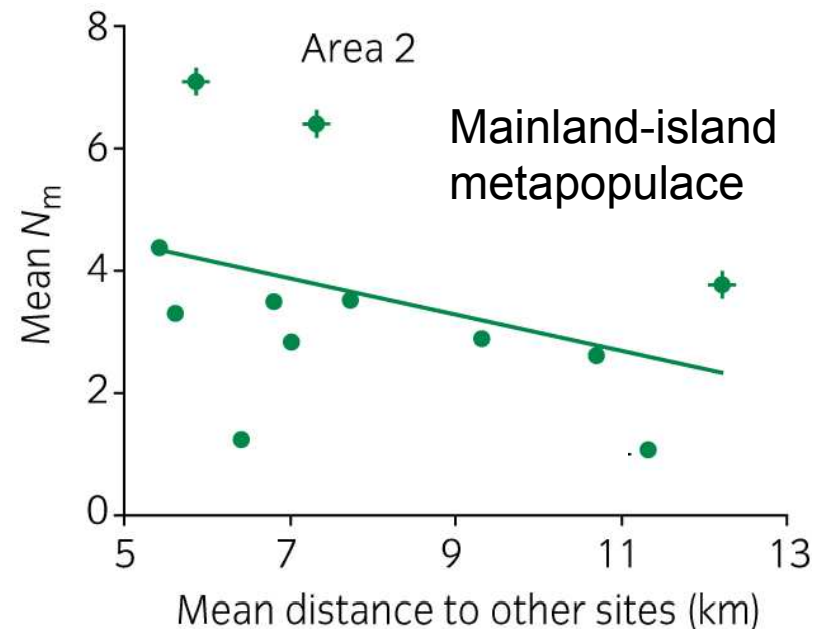
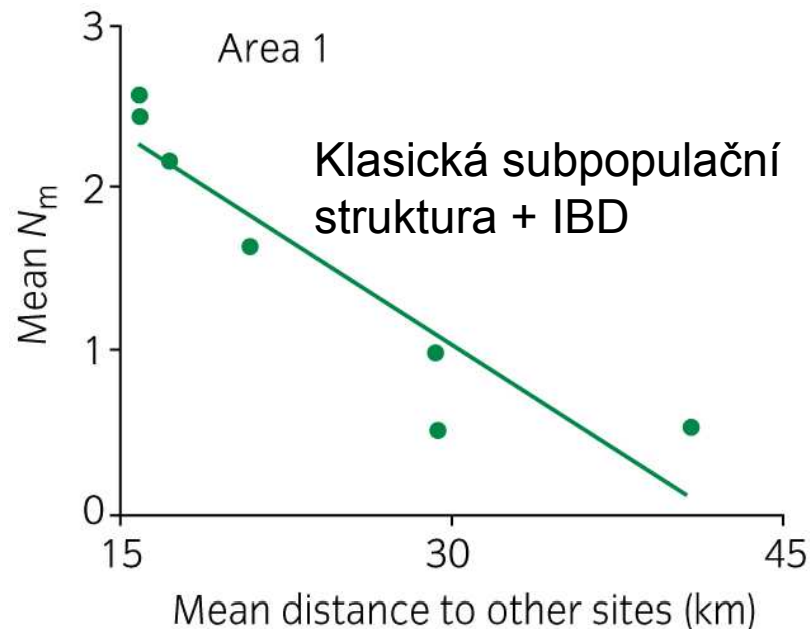


„source and sinks“

- ad a) *Cristatella mucedo* - Bryozoa (Freeland et al. 2000) – long-distance migration (birds), no IBD, 5 ms for clone identification
- black rats in Senegal – transport by traffic

ad b) Mainland-island model

- *Bufo calamita* (Rowe et al. 2000) – different ponds have different breeding success
- klasické subpopulace (Area 1) i metapopulace (Area 2); křížek – vyšší tok genů než by se předpokládalo

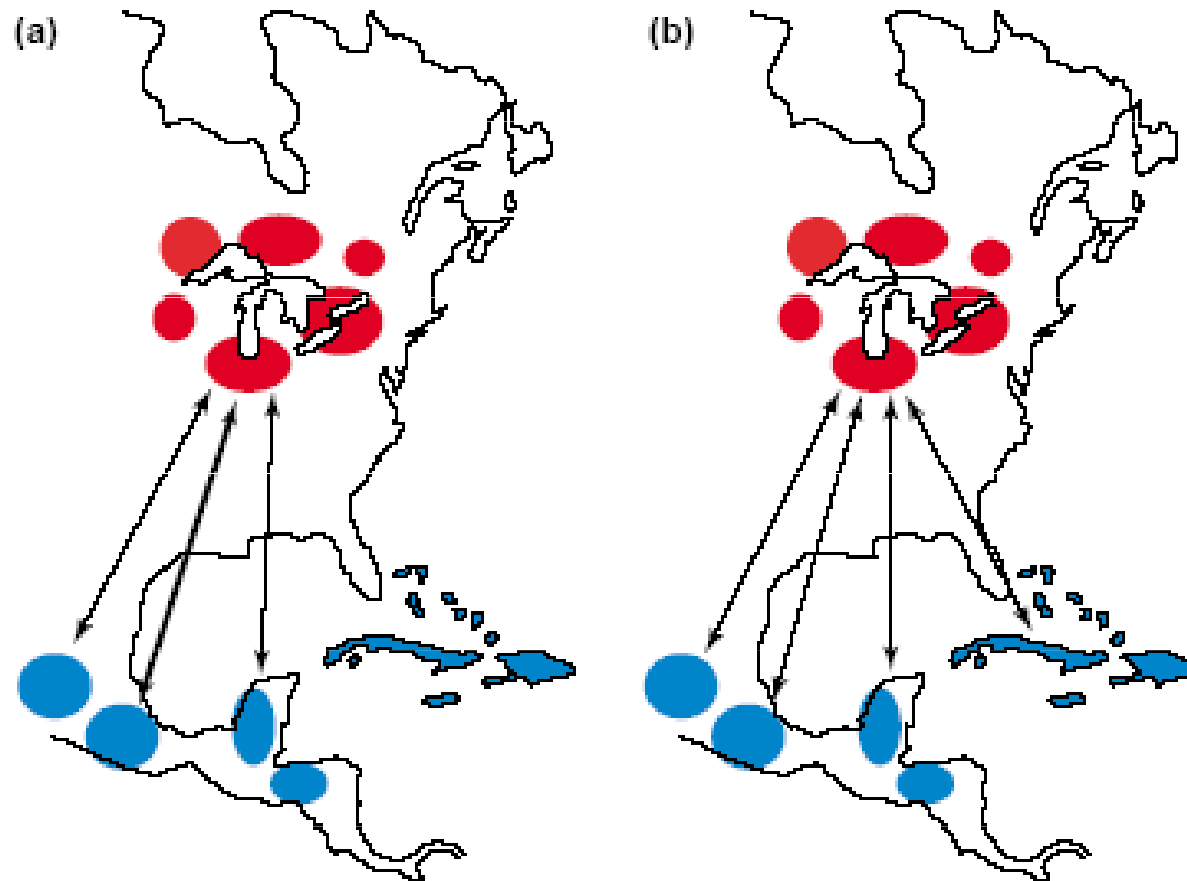


Population assignments

Klasické problémy populační genetiky

- Populace dány, jedinci předem zařazeni do populací, zajímají nás vlastnosti populací
- Populace sice definovány, ale chceme k nim přiřadit jedince neznámého původu
- Kryptická populační struktura = předem není dáno nic → chci zjistit klastry (tj. přirozené populace) a rozřadit individua do klastrů (**population assignments**)

Unraveling migratory connectivity



(a) strong vs. (b) weak connectivity

1. Direct methods

- morphological variation (geographical races)
- leg-bands or similar markers (ex. over one million *Ficedula hypoleuca* have been ringed in UK and Sweden – only six recaptured on wintering grounds in Africa)
- satellite telemetry – expensive

2. Biogeochemical approaches

- ratios of stable isotopes of naturally occurring elements (C, H, N, Sr) vary across the landscape
 - determined by the relative frequency of C3 and C4 plants, climate, and bedrock
- (1) geographical structure of isotopic ratio distributions
 - (2) knowledge about where animals incorporate isotopes
 - (3) tissue samples from individuals at different parts of their annual cycle

3. Genetic approaches

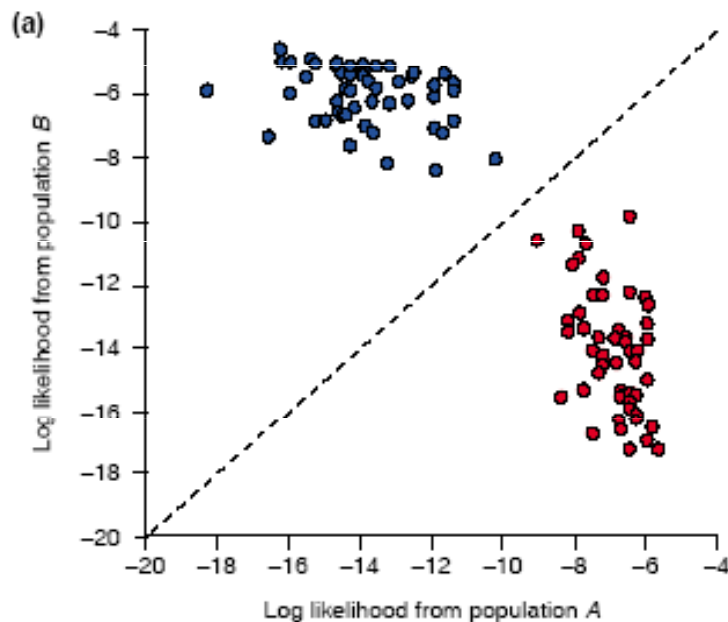
- « very few birds have bands, but all have genotypes »
- genetic data on population structure
- problems: (1) weak genetic differentiation among populations (widespread dispersal), (2) lack of differentiation in northern temperate zone – recent postglaciation expansion
- solution: (1) more genetic markers, (2) study of avian parasites DNA

Useful (and used) genetic markers

- **mtDNA** – smaller effective population size, population specific markers arise quickly, often fixed differences
- **microsatellites** – pronounced polymorphism; rare in the avian genome
- **dominant nuclear markers** (RAPD, AFLP) – almost infinite number of markers

Population assignment tests

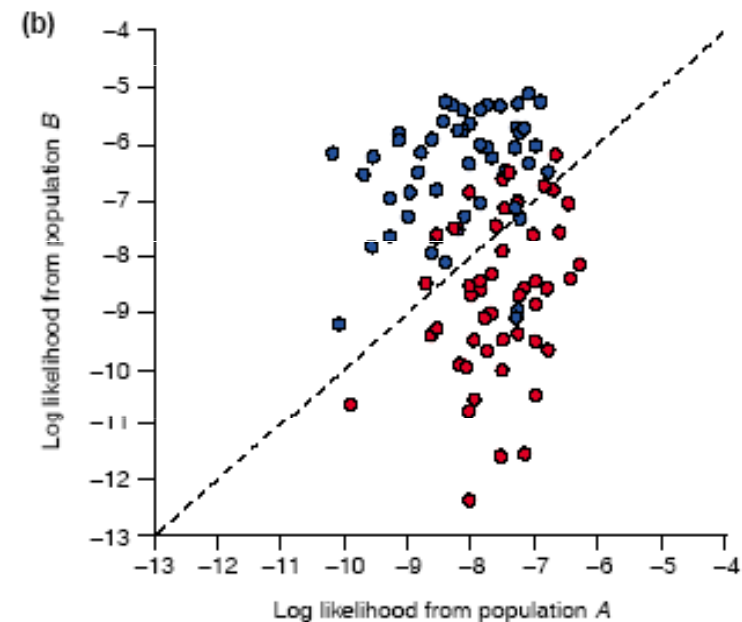
- program GeneClass (Piry et al. 2004)
- calculates the probability that an individual's genotype might exist in a particular population
- can combine data from multiple genetic marker types



5 microsatellite loci

$$F_{st} = 0.14$$

99.9% assigned correctly



5 microsatellite loci

$$F_{st} = 0.04$$

90.2% assigned correctly

Klastrování

- **Distance based methods**

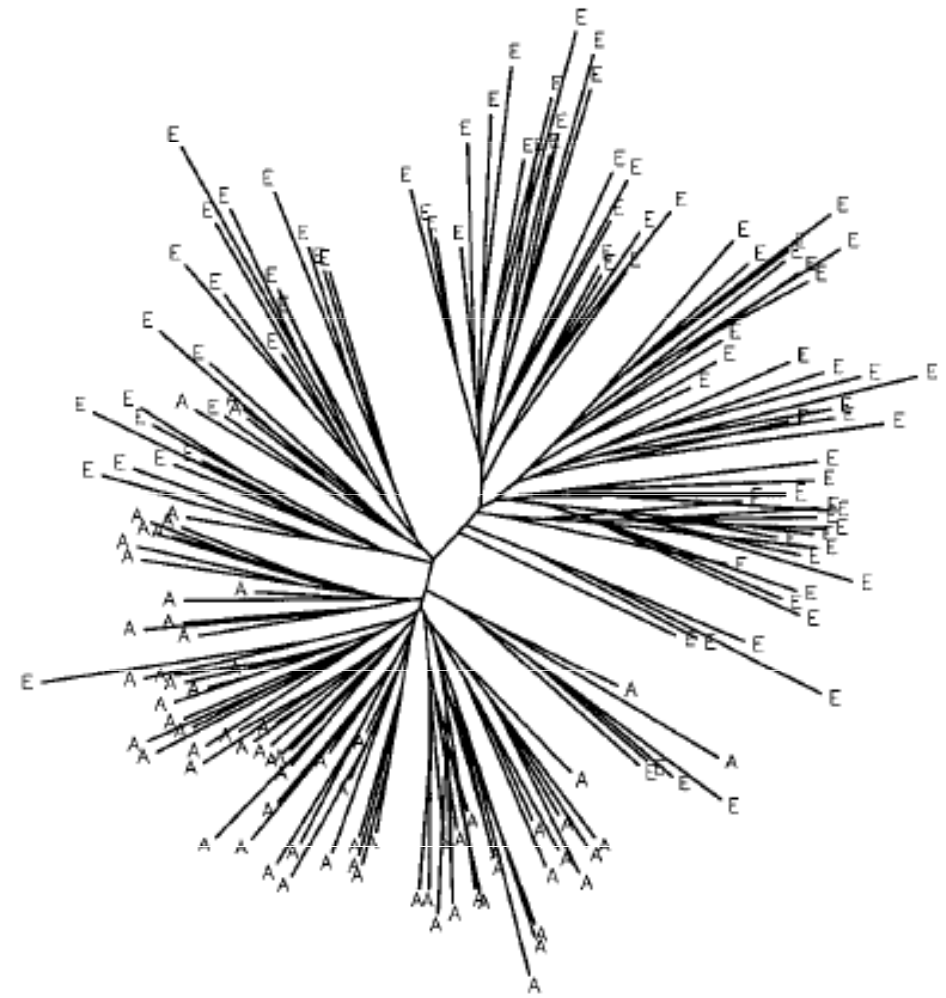
Matice párových vzdáleností
(vzdálenosti mezi každým párem individuí)
Znázornění mnohorozměrným grafem →
klastry

Explorační metoda!
Důvěryhodnost klastrů není známa
Závislost na distanční míře i na grafickém
zobrazení
Např. neighbour-joining

- **Model based methods**

Použijí parametrický model
Současně hledám parametry pro klastry a
určuji členy klastrů
Určím věrohodnost výsledků

(Maximum likelihood, Bayesianá
metoda)

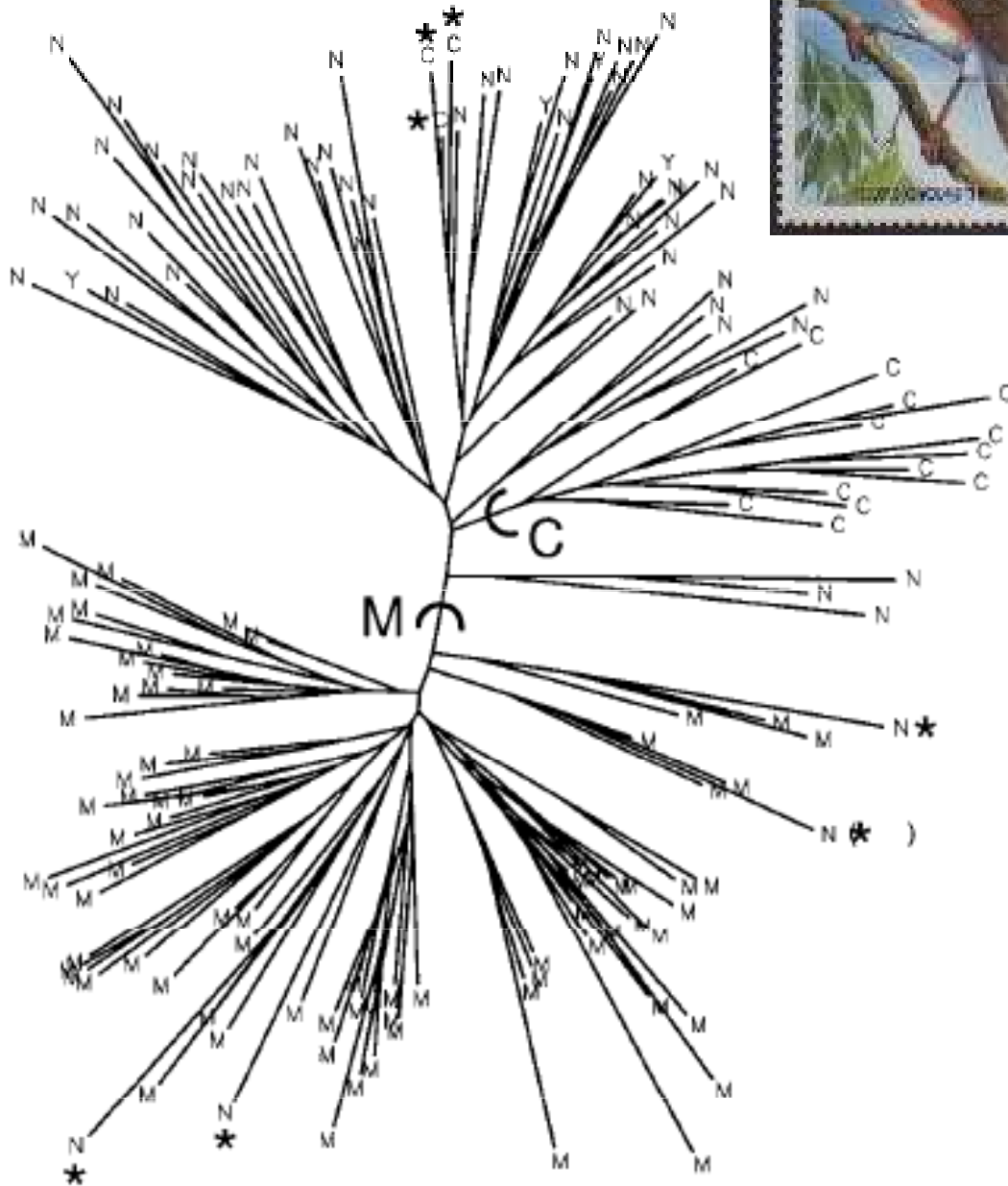


Bayesian clustering approach

STRUCTURE - *Pritchard et al. 2000*

- Neznámý počet populací charakterizovaných různými frekvencemi alel → počet populací a frekvence zjišťují
- Současně přiřazují individua do populací
- Lokusy, které nejsou ve vazbě, HW uvnitř subpopulací (mikrosatelity, RFLPs, SNPs)
- Možno předem zahrnout geografickou polohu individuí
- Model se snaží vyložit HW nebo vazebnou nerovnováhu zavedením populační struktury
- Místo přímého výpočtu – odhad pomocí Markov chain Monte Carlo

Turdus helleri

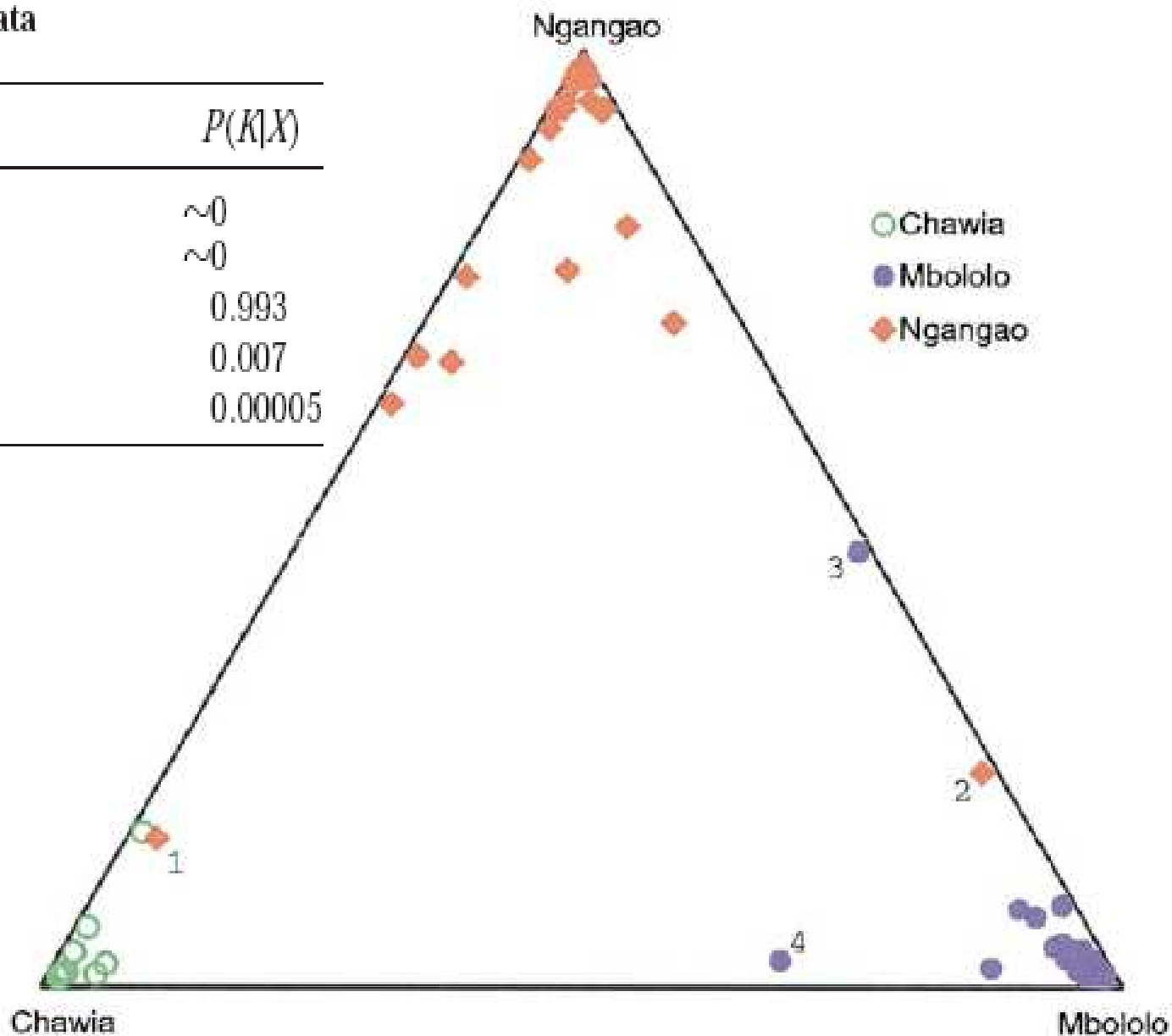


- Fragmenty deštného pralesa
- Lokality Chawia, Ngangao, Mbololo, Yale (Kenya)
- 7 mikrosatelitových lokusů
- Neighbour-joining
- * špatně zařazení jedinci

Program STRUCTURE - Bayesiánský přístup

Inferring the value of K , the number of populations,
for the *T. helleri* data

K	$\log P(X K)$	$P(K X)$
1	-3144	~ 0
2	-2769	~ 0
3	-2678	0.993
4	-2683	0.007
5	-2688	0.00005

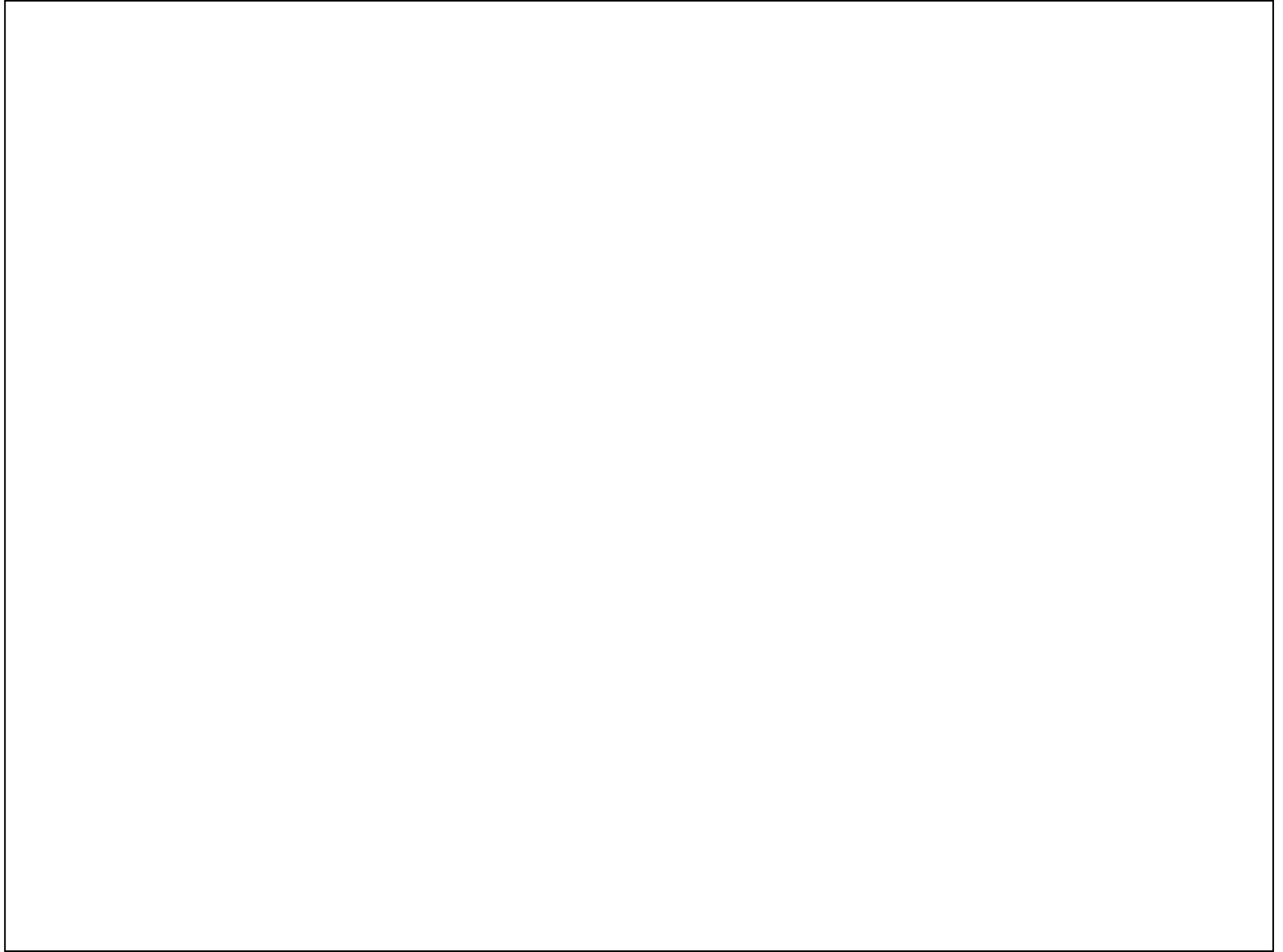


Příklad: Hořavka duhová (*Rhodeus sericeus*)

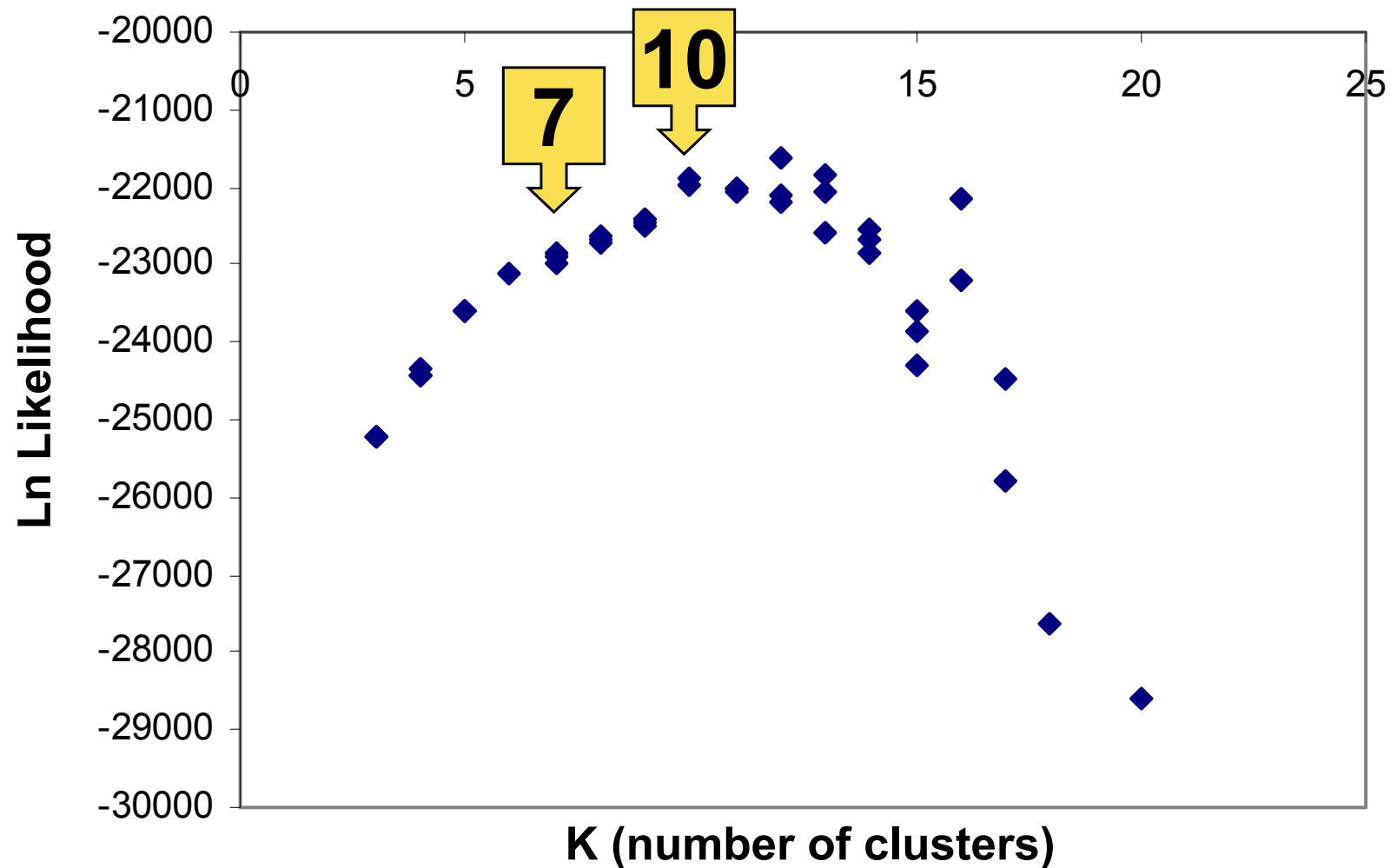


téměř 600 jedinců, 21 lokalit, 12 mikrosatelitů

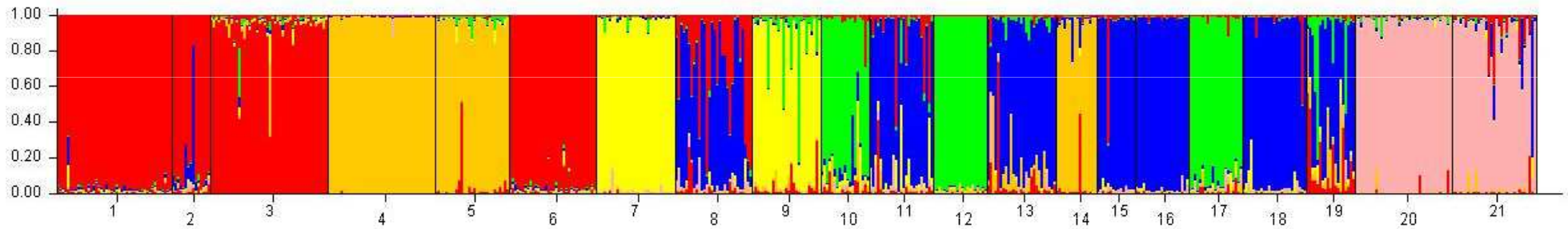
Bryja & Reichard, nepubl. data



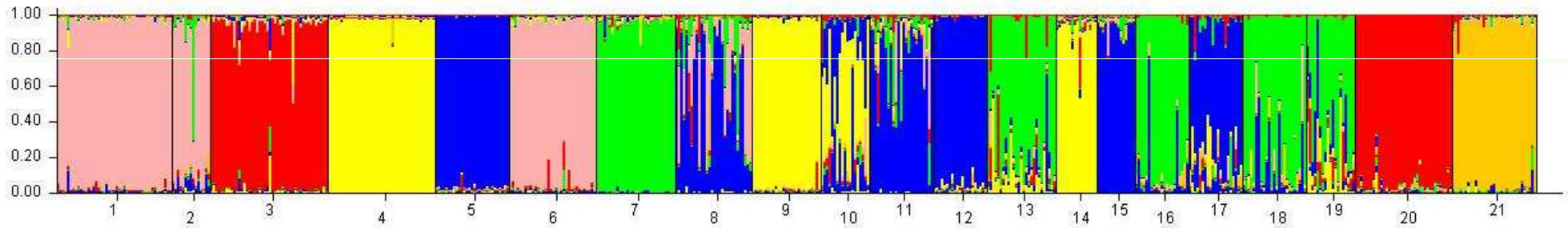
Stanovení počtu „přirozených“ subpopulací



K = 7

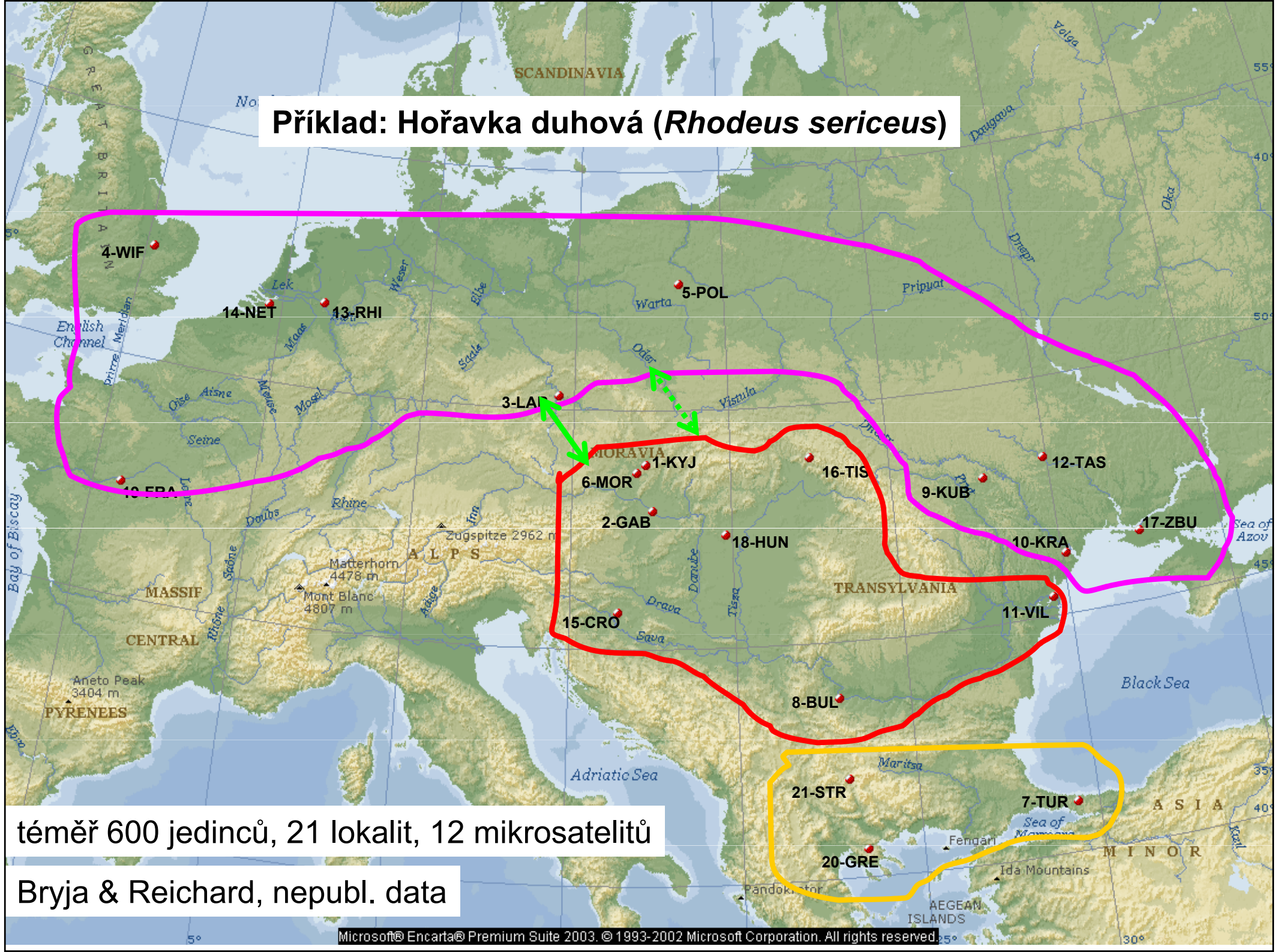


K = 10



Proporce genomu každého jedince náležející určitému „clusteru“

Příklad: Hořavka duhová (*Rhodeus sericeus*)



téměř 600 jedinců, 21 lokalit, 12 mikrosatelitů

Bryja & Reichard, nepubl. data