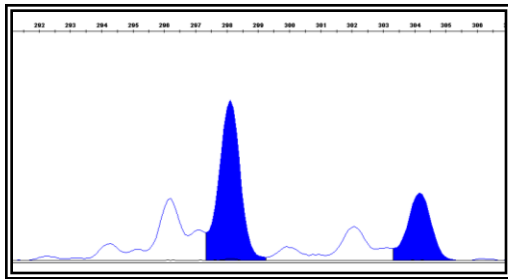
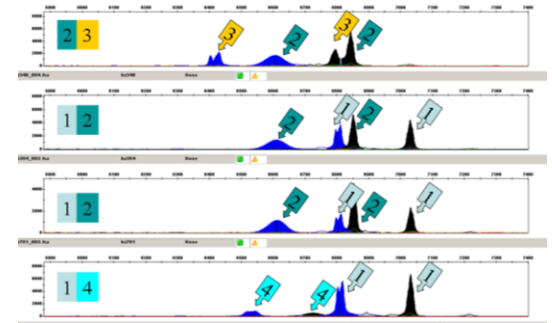


Populačně-genetická data

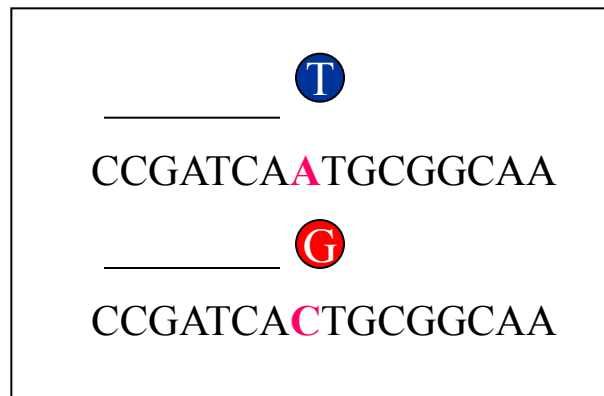
Základy analýzy diploidních kodominantních znaků (Mendelovská dědičnost)



mikrosatelity

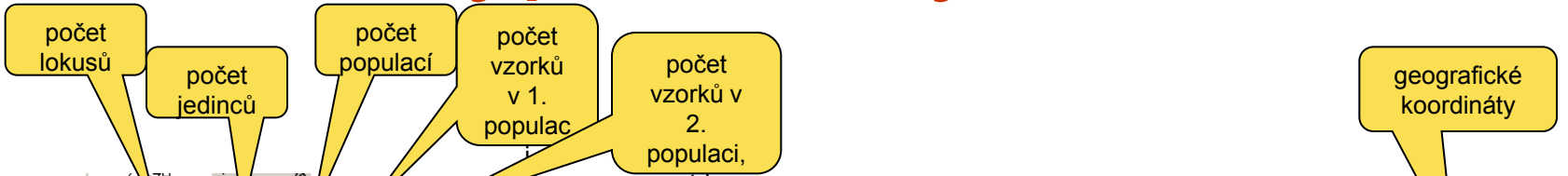


jaderné sekvence (např. SSCP)



SNPs

Typ získaných dat



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB									
1	11	233	10	30	22	25	22	18	23	12	27	25	29																								
2	Allele frequency																																				
3	ID	Pop	Ppip01	Ppip02	Ppip04	Ppip06	EF1	EF4	EF6	Paur05	NN18	NnP217	NnP219	X	Y																						
4	TYN1	TYN	174	176	128	128	213	215	120	132	166	175	236	247	174	188	249	251	287	291	205	209	140	170	48.76500	17.00750											
5	TYN2	TYN	167	167	132	132	207	207	130	136	169	175	230	245	182	188	247	249	287	287	205	218	166	166	48.76500	17.00750											
6	TYN3	TYN	176	178	128	128	207	207	130	130	169	169	0	184	190	190	249	249	283	293	209	213	166	170	48.76500	17.00750											
7	TYN4	TYN	180	180	130	132	207	207	118	124	171	173	220	243	184	190	247	249	283	285	209	214	140	166	48.76500	17.00750											
8	TYN5	TYN	174	174	128	128	207	207	128	138	179	181	241	241	182	188	247	253	0	0	205	209	158	162	48.76500	17.00750											
9	TYN6	TYN	180	182	128	130	209	209	136	136	173	175	241	243	184	186	253	253	295	325	209	213	162	170	48.76500	17.00750											
10	TYN7	TYN	174	180	128	130	215	215	130	130	173	173	241	241	190	190	249	249	285	325	209	213	158	162	48.76500	17.00750											
11	TYN8	TYN	174	175	128	134	207	221	128	138	171	177	220	240	178	186	249	249	287	293	213	213	162	166	48.76500	17.00750											
12	TYN9	TYN	156	178	126	130	209	209	128	149	177	177	220	230	184	188	249	255	287	287	217	218	158	166	48.76500	17.00750											
13	TYN10	TYN	167	178	128	130	209	209	132	134	171	175	243	245	184	186	245	247	285	287	205	218	158	174	48.76500	17.00750											
14	TYN11	TYN	170	174	130	130	217	217	130	138	173	175	0	0	182	190	251	257	289	291	213	213	166	170	48.76500	17.00750											
15	TYN12	TYN	174	176	130	132	0	0	134	134	175	177	220	220	186	186	249	251	293	297	209	209	158	162	48.76500	17.00750											
16	TYN13	TYN	166	166	130	130	213	213	130	130	173	173	241	241	190	190	249	249	285	325	209	213	162	170	48.76500	17.00750											
17	TYN14	TYN	162	162	130	130	213	213	130	130	173	173	241	241	190	190	249	249	285	325	209	213	162	170	48.76500	17.00750											
18	TYN15	TYN	156	156	130	130	213	213	130	130	173	173	241	241	190	190	249	249	285	325	209	213	162	170	48.76500	17.00750											
19	TYN16	TYN	168	178	128	128	215	217	132	134	173	175	243	243	178	178	247	249	289	289	205	209	162	170	48.76500	17.00750											
20	TYN17	TYN	178	178	126	130	207	211	132	132	171	175	220	238	176	180	249	255	289	289	178	209	162	166	48.76500	17.00750											
21	TYN18	TYN	161	174	130	130	213	213	132	132	173	177	245	245	178	178	243	249	289	289	205	209	162	182	48.76500	17.00750											
22	TYN19	TYN	174	180	128	132	213	213	134	134	169	169	220	220	174	182	249	249	285	285	209	213	131	131	48.76500	17.00750											
23	TYN20	TYN	176	176	130	130	0	0	138	151	171	171	247	247	186	188	249	255	287	287	209	209	131	166	48.76500	17.00750											
24	TYN21	TYN	178	178	128	134	213	213	132	134	171	173	245	245	180	188	247	249	285	287	209	209	144	154	48.76500	17.00750											
25	TYN22	TYN	156	175	128	132	217	217	124	132	177	177	220	220	180	182	251	253	287	297	209	217	170	170	48.76500	17.00750											
26	TYN23	TYN	168	176	128	128	213	213	120	153	173	173	220	236	178	180	245	253	285	287	209	213	154	162	48.76500	17.00750											
27	TYN24	TYN	176	176	130	130	219	219	132	149	175	175	216	245	182	184	255	255	283	297	205	209	131	174	48.76500	17.00750											
28	TYN25	TYN	177	179	128	130	209	213	126	126	156	173	241	241	188	188	249	255	289	297	198	213	158	166	48.76500	17.00750											
29	TYN26	TYN	177	179	126	130	201	201	118	124	171	179	218	245	174	184	249	255	295	295	205	205	158	166	48.76500	17.00750											
30	TYN27	TYN	176	176	126	128	207	207	120	153	169	173	220	220	182	188	249	255	291	291	205	209	136	162	48.76500	17.00750											
31	TYN28	TYN	168	172	128	132	219	219	130	130	169	175	249	249	186	188	245	253	285	289	209	213	132	170	48.76500	17.00750											
32	TYN29	TYN	170	180	128	130	219	219	130	130	171	171	243	243	184	184	249	255	285	289	209	213	166	178	48.76500	17.00750											
33	TYN30	TYN	170	177	124	130	215	215	138	138	171	177	232	236	184	186	249	253	285	291	209	209	162	174	48.76500	17.00750											
34	NOV1	NOV	172	176	128	134	209	219	120	120	169	175	253	253	182	184	251	255	289	295	178	209	162	166	48.83320	16.50610											
35	NOV2	NOV	178	178	130	130	209	209	128	132	173	173	238	243	182	184	249	251	285	287	205	209	136	170	48.83320	16.50610											
36	NOV3	NOV	163	165	126	126	207	213	122	124	173	175	220	245	178	186	247	247	285	285	205	213	133	179	48.83320	16.50610											
37	NOV4	NOV	167	178	128	134	205	205	118	124	173	177	220	220	186	188	253	255	287	289	205	213	162	166	48.83320	16.50610											
38	NOV5	NOV	176	178	128	130	201	201	130	130	173	177	243	245	174	184	247	249	287	293	201	209	162	170	48.83320	16.50610											
39	NOV6	NOV	176	178	126	130	209	215	138	134	166	173	0	0	184	184	245	245	287	287	205	213	162	166	48.83320	16.50610											

pop1
pop2

genotypy, tj. velikosti fragmentů v populaci

formát GenAlex



Microsoft Excel - PYG_GenAlex.xls

Soubor Úpravy Zobrazit Vložit Formát Nástroje Data Okno GenALEX Nápověda

A1 = 11

	A	B	C	D	E	F	G	H	I	J
1	11	233	10	30	22	25	22	18	23	12
2	Allele frequency									
3	ID	Pop	Ppip01	Ppip02	Ppip04	Ppip06				
4	TYN1	TYN	174	176	128	128	213	215	120	13
5	TYN2	TYN	167	167	132	132	207	207	130	13
6	TYN3	TYN	176	178	128	128	207	207	130	13
7	TYN4	TYN	180	180	130	132	207	207	118	12
8	TYN5	TYN	174	174	128	128	207	207	128	13
9	TYN6	TYN	180	182	128	130	209	209	136	13
10	TYN7	TYN	174	180	128	130	215	215	130	13
11	TYN8	TYN	174	175	128	134	207	221	128	13
12	TYN9	TYN	156	178	126	130	209	209	128	14
13	TYN10	TYN	167	178	128	130	209	209	132	13
14	TYN11	TYN	170	174	130	130	217	217	130	13
15	TYN12	TYN	174	176	130	132	0	0	134	13
16	TYN13	TYN	166	176	126	132	215	217	122	15
17	TYN14	TYN	162	178	128	128	219	219	128	14
18	TYN15	TYN	156	176	128	132	209	209	118	12
19	TYN16	TYN	168	178	128	128	215	217	132	13
20	TYN17	TYN	178	178	126	130	207	211	132	13
21	TYN18	TYN	161	174	130	130	213	213	132	13
22	TYN19	TYN	174	180	128	132	213	213	134	13
23	TYN20	TYN	176	176	130	130	0	0	138	15
24	TYN21	TYN	178	178	128	134	213	213	132	13
25	TYN22	TYN	156	175	128	132	217	217	124	132
26	TYN23	TYN	168	176	128	128	213	213	120	153
27	TYN24	TYN	176	176	130	130	219	219	132	149
28	TYN25	TYN	177	179	128	130	209	213	126	126
29	TYN26	TYN	177	179	126	130	201	201	118	124
30	TYN27	TYN	176	176	126	128	207	207	120	153
31	TYN28	TYN	168	172	128	132	219	219	130	130
32	TYN29	TYN	170	180	128	130	219	219	130	130
33	TYN30	TYN	170	177	124	130	215	215	138	138
34	NOV1	NOV	172	176	128	134	209	219	120	120
35	NOV2	NOV	178	178	130	130	209	209	128	132
36	NOV3	NOV	163	165	126	126	207	213	122	124
37	NOV4	NOV	167	178	128	134	205	205	118	124
38	NOV5	NOV	176	178	128	130	201	201	130	130
39	NOV6	NOV	176	178	126	130	209	215	128	134

GenALEX menu options:

- Frequency...
- HWE...
- Distance
- AMOVA...
- Mantel...
- PCA...
 - EF6
 - Paur05
 - NN18
- Assignment
 - 247 174 188 249 251 287
 - 245 182 188 247 249 287
 - 0 184 190 249 249 283
- Spatial
- Relatedness
 - 243 184 190 247 249 283
- Multilocus
 - 241 182 188 247 253 0
 - 243 184 186 253 253 295
- Template
 - 241 190 190 249 249 285
- Create
 - 240 178 186 249 249 287
- Parameters
 - 230 184 188 249 255 287
 - 245 184 186 245 247 285
- Data
 - 0 182 190 251 257 289
- Import Data
 - 220 186 186 249 251 293
 - 243 178 186 249 251 287
- Raw Data
 - 220 178 182 253 253 285
- Edit Raw Data
 - 220 178 182 249 251 285
- Export Data
 - Arlequin... 249 289
 - Cervus... 255 289
 - Famos... 249 285
 - Fdist2... 255 287
 - GeneClass... 249 285
 - GenePop... 253 287
 - Kingroup... 253 285
 - MEGA... 255 283
 - MSA... 255 289
 - MsVar... 255 295
 - Nexus... 255 291
 - Phylip... 253 285
 - Phylip... 255 285
 - Phylip... 253 285
 - Phylip... 255 289
 - PopGene... 251 285
 - SPAGeDI... 247 285
 - SPAGeDI... 255 287
 - Structure... 249 287
 - Structure... 245 287
- Graph
- Stats
- Options

Velké množství populačně-genetických programů

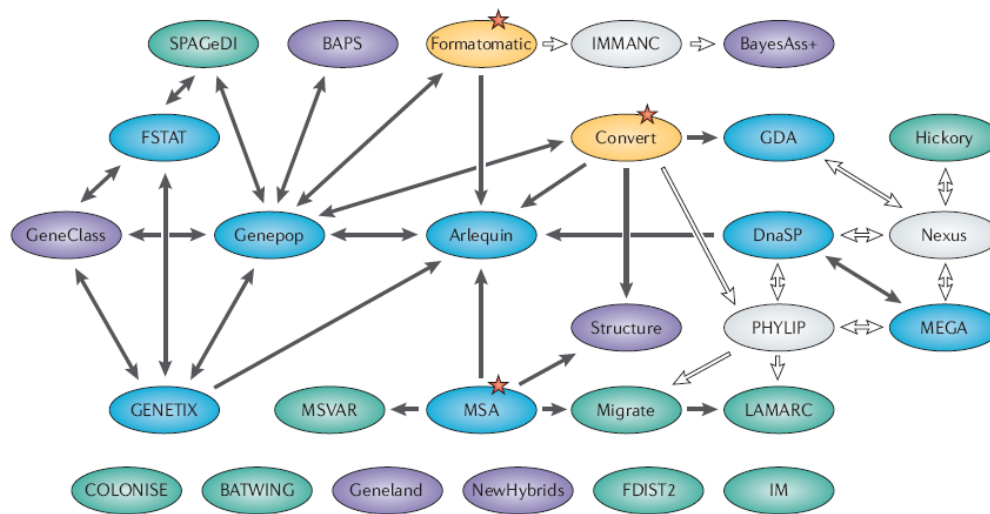
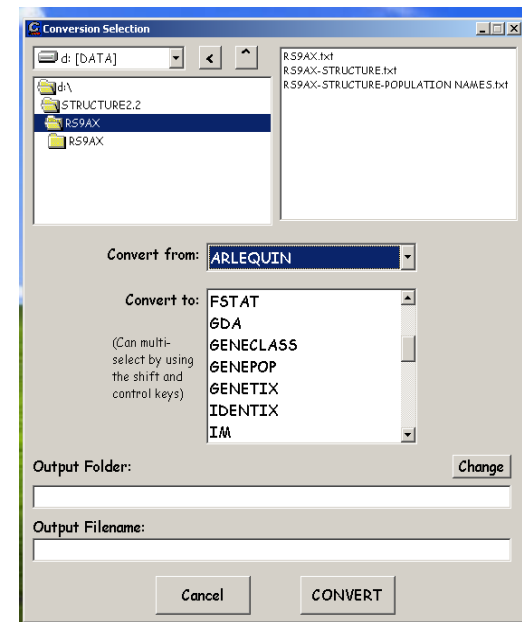
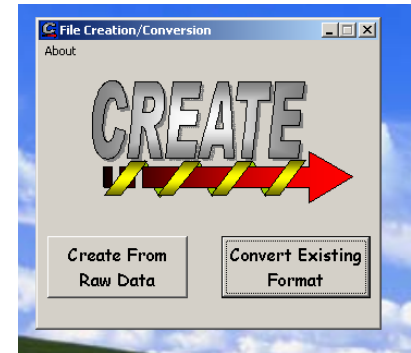


Figure 1 | **Flow chart of possible data exchange between different population genetics programs.** Although many programs have their own input-file specification, data files can still be exchanged between most programs (black arrows), avoiding tedious reformatting processes. The red stars are recommended starting points to format an initial data set. Blue ellipses represent multi-purpose packages, whereas individual-centred programs are shown in violet. The two conversion programs are shown in yellow. Specialized programs are shown in green, and light grey ellipses represent programs that are not reviewed here, but the data formats of which are used by other programs allowing indirect data exchange (white arrows). The data files associated with the programs listed on the bottom row cannot be exchanged directly with the other programs.



Computer programs for population genetics data analysis: a survival guide

Laurent Excoffier and Gerald Heckel

```

PYG - Poznámkový blok
Soubor Úpravy Formát Zobrazení Nápvěda
Title line:"Genotype file for Pipistrellus data"
Ppip01
Ppip02
Ppip04
Ppip06
EF1
EF4
EF6
Paur05
NN18
NnP217
NnP219
Pop
  Pop1, 1820 0505 0809 0511 0308 1321 0209 0506 0608 1214 0513
  Pop1, 1111 0707 0505 1013 0508 0920 0609 0405 0606 1219 1212
  Pop1, 2022 0505 0505 1010 0505 0000 0710 0505 0409 1416 1213
  Pop1, 2424 0607 0505 0407 0607 0419 0710 0405 0405 1417 0512
  Pop1, 1818 0505 0505 0914 1011 1717 0609 0407 0000 1214 1011
  Pop1, 2426 0506 0606 1313 0708 1719 0708 0707 1014 1416 1113
  Pop1, 1824 0506 0909 1010 0707 1717 1010 0505 0514 1416 1011
  Pop1, 1819 0508 0512 0914 0609 0416 0408 0505 0609 1616 1112
  Pop1, 0222 0406 0606 0919 0909 0409 0709 0508 0606 1819 1012
  Pop1, 1122 0506 0606 1112 0608 1920 0708 0304 0506 1219 1014
  Pop1, 1418 0606 1010 1014 0708 0000 0610 0609 0708 1616 1213
  Pop1, 1820 0607 0000 1212 0809 0404 0808 0506 0911 1414 1011
  Pop1, 1020 0407 0910 0621 0505 0419 0408 0506 0606 1416 1113
  Pop1, 0722 0505 1111 0919 0709 0404 0406 0707 0505 0212 1212
  Pop1, 0220 0507 0606 0409 0910 0404 0406 0506 0514 1616 1214
  Pop1, 1222 0505 0910 1112 0708 1919 0404 0405 0707 1214 1113
  Pop1, 2222 0406 0507 1111 0608 0414 0305 0508 0707 0214 1112
  Pop1, 0618 0606 0808 1111 0709 2020 0404 0205 0707 1214 1118
  Pop1, 1824 0507 0808 1212 0505 0404 0206 0505 0505 1416 0101
  Pop1, 2020 0606 0000 1420 0606 2121 0809 0508 0606 1414 0112
  Pop1, 2222 0508 0808 1112 0607 2020 0509 0405 0506 1414 0609
  Pop1, 0219 0507 1010 0711 0909 0404 0506 0607 0611 1418 1313
  Pop1, 1220 0505 0808 0521 0707 0413 0405 0307 0506 1416 0911
  Pop1, 2020 0606 1111 1119 0808 0220 0607 0808 0411 1214 0114
  Pop1, 2123 0506 0608 0808 0207 1717 0909 0508 0711 0916 1012
  Pop1, 2123 0406 0202 0407 0610 0320 0207 0508 1010 1212 1012
  Pop1, 2020 0405 0505 0521 0507 0404 0609 0508 0808 1214 0411
  Pop1, 1216 0507 1111 1010 0508 2222 0809 0307 0507 1416 0213
  Pop1, 1424 0506 1111 1010 0606 1919 0707 0508 0507 1416 1216
  Pop1, 1421 0306 0909 1414 0609 1113 0708 0507 0508 1414 1114
Pop
  Pop2, 1620 0508 0611 0505 0508 2323 0607 0608 0710 0214 1112
  Pop2, 2222 0606 0606 0911 0707 1419 0607 0506 0506 1214 0413
  Pop2, 0809 0404 0508 0607 0708 0420 0408 0404 0505 1216 0317
  Pop2, 1122 0508 0404 0407 0709 0404 0809 0708 0607 1216 1112
  Pop2, 2022 0506 0202 1010 0709 1920 0207 0405 0609 1014 1113
  Pop2, 2022 0406 0609 0912 0307 0000 0707 0303 0606 1216 1112
  Pop2, 1414 0506 0000 1111 0709 2020 0206 0307 0606 1012 1212
  Pop2, 1818 0505 0710 1111 0708 2020 0709 0305 0606 1516 1213
  Pop2, 2020 0405 0208 1021 0606 0412 0607 0305 0606 1212 1012
  Pop2, 1820 0505 0808 0809 0709 0319 0710 0407 0000 1212 0212
  Pop2, 0220 0406 0707 0909 0909 0304 0808 0208 1617 1112 0213
  Pop2, 2020 0206 0505 0509 0307 0517 0308 0105 0914 1014 0213
  Pop2, 2024 0607 1010 1212 0307 1920 0809 0407 0505 1414 1213
  Pop2, 2024 0509 0202 1515 0206 0420 0307 0508 0707 1414 1112
  Pop2, 1624 0507 0205 1115 0208 0420 0709 0205 0707 1214 1215
  Pop2, 1621 0607 1010 0921 0609 2222 0708 0707 1415 1416 1113
  Pop2, 1616 0506 0508 1010 0607 0517 0306 0509 0608 1117 1215
  Pop2, 1420 0507 0909 1010 0708 0416 0710 0506 0506 1214 0912
  Pop2, 2024 0405 1111 0910 0507 1921 0709 0105 0607 1014 0912
  Pop2, 1620 0508 0611 0406 0511 0319 0610 0506 0610 0212 0211
  Pop2, 1220 0505 0609 0909 0809 1717 0208 0708 0606 1416 0412
  Pop2, 2222 0506 0808 1010 0608 1919 0708 0307 0506 1114 0210
Pop
  Pop3, 2125 0407 0000 0606 0606 0000 0409 0708 0406 1616 0513
  Pop3, 1620 0405 0609 1414 0708 1418 0203 0405 0505 1015 1011
  Pop3, 1818 0205 0910 1010 0307 0505 0609 0411 0410 1115 1112
  Pop3, 1823 0406 1010 0306 0810 1420 0208 0303 0809 1212 0712

```

Genepop file format – jednoduchý ASCII kód (.txt)

- jednotlivé alely pro daný lokus jsou seřazeny podle velikosti a očíslovány
- tj. např. 128/130 je převedeno na 10/11

Účel populačně-genetické analýzy

Table 5 | List of computer programs suited for a given analysis and genetic marker

	Multi-allelic markers*	STR	Dominant markers (AFLP)	SNP	DNA sequences
Descriptive statistics			SPAGeDi		Arlequin, DnaSP, MEGA
Linkage disequilibrium					
Analysis of population subdivision			Hickory		Arlequin, DnaSP, MEGA
Detection of new immigrants: known populations					
Detection of new immigrants: inferred populations					
Demographic expansion or decline					Arlequin, DnaSP, IM, LAMARC, Migrate
Population size					IM, LAMARC, Migrate
Divergence time					DnaSP, IM, LAMARC, Migrate
Migration rates					DnaSP, IM, LAMARC, Migrate
Neutrality tests					Arlequin, DnaSP, MEGA
Spatially explicit analyses					

*By multi-allelic markers, we mean loci for which no specific mutation model is assumed. Computations are based on allele frequencies only, as well as nucleotide frequencies. Some software packages (e.g., Arlequin) are able to detect new immigrants. AFLP, amplified fragment length polymorphism; IM, isolation by migration; LAMARC, linkage disequilibrium and genetic structure analysis; Migrate, migration analysis; MEGA, Molecular Evolutionary Genetics Analysis; SPAGeDi, SPeAr Geographical Diversity Index.

ed. In the latter case, computations are based on a sequence, STR and SNP allele frequencies, and to detect new immigrants.

frekvence alel

frekvence alel + mutační model

1. Deskriptive statistics = intrapopulation variation

- genetická variabilita
 - počet variabilních lokusů (polymorfismus)
 - počet alel
 - heterozygotnost
-
- Hardy-Weiberg equilibrium

Genetická variabilita

Polymorfismus

- podíl polymorfních lokusů (znaků) – 95 % nebo 99 % (např. 0,8 = 4 z pěti zkoumaných mikrosatelitů mají v populaci alespoň 2 alely, z nichž ta vzácnější dosahuje frekvence alespoň 1% nebo 5%)

Počet alel (number of alleles)

- počet alel na lokus

Alelická bohatost (allelic richness)

- počet alel na lokus vztažený k velikosti vzorku (metodou „rarefaction“)

Pozorovaná heterozygotnost (observed heterozygosity)

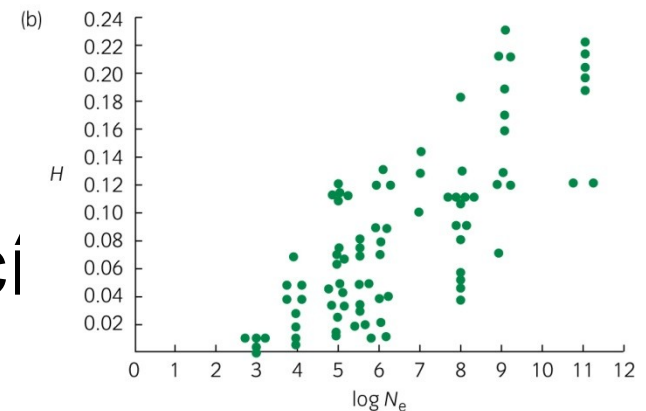
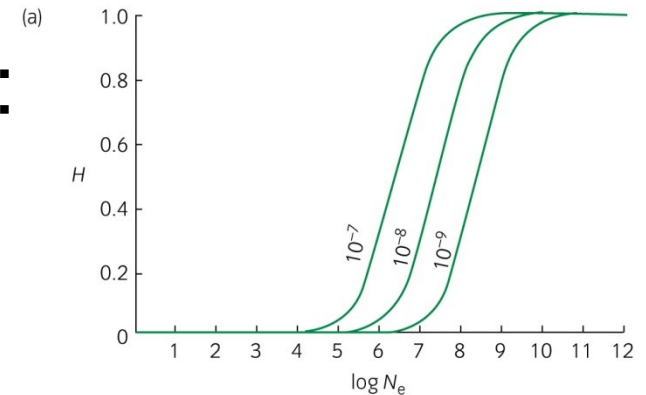
- průměrná četnost heterozygotů v jednotlivých lokusech

Očekávaná heterozygotnost (expected heterozygosity)

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

Použití údajů o genetické variabilitě

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



Hardy-Weinbergova rovnováha

Př. Jeden lokus se 2 alelami

Alela	Četnost alely
A	p
a	q

$$p + q = 1$$

p, q - zjistíme genetickou analýzou

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

= Hardy-Weinbergova rovnováha

- četnosti genotypů zjistíme genetickou analýzou
- odchylky od očekávaných četností \Rightarrow např. χ^2 test

Odchylky od HW rovnováhy

- **nadbytek heterozygotů** = negativní asortativní páření (tj. cílené rozmnožování nepodobných jedinců) – použité lokusy mohou být výhodné v heterozygotním stavu (např. geny MHC)
- **nedostatek heterozygotů**
 - inbreeding (postihuje všechny lokusy stejně)
 - nulové alely (jen na některých lokusech bude deficit heterozygotů)

Příklad

- Genepop on the web:
- <http://genepop.curtin.edu.au/>
- FSTAT

2. Analysis of population subdivision



Hierarchická populační struktura

Druh → populace → subpopulace (demy)

Předpoklady studia populačně-genetické struktury

- lokusy používané pro analýzu populační struktury jsou neutrální vůči selekci
- **klasický populačně-genetický přístup** = jednotlivé populace jsou předem známy (např. chceme zjistit úroveň genetických rozdílů mezi dvěma lokalitami)

Genetická struktura populací drift, mutace a migrace

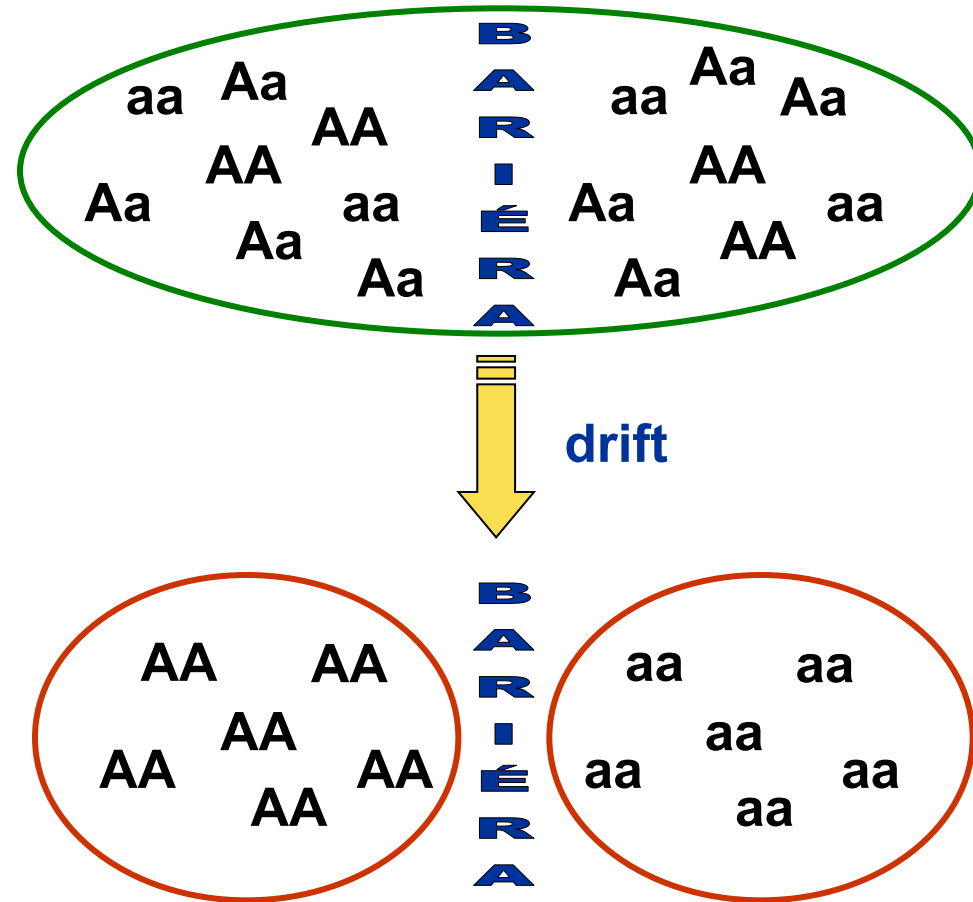
- **Drift**

→ diferenciaci subpopulací
díky fixaci alternativních alel

- **Mutace**

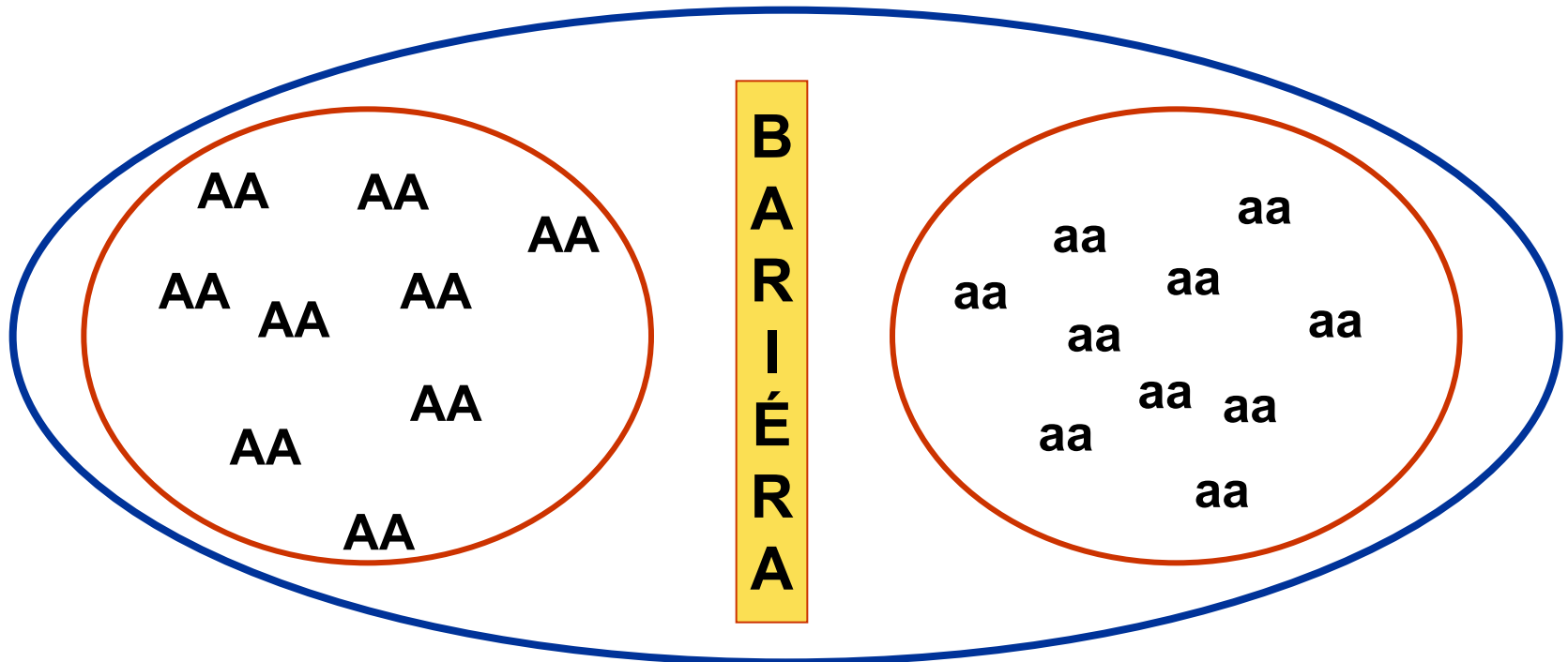
mohou zvýšit diferenciaci
(odlišit subpopulace)
ale riziko homoplázií

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



Vliv populační struktury na heterozygotnost

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



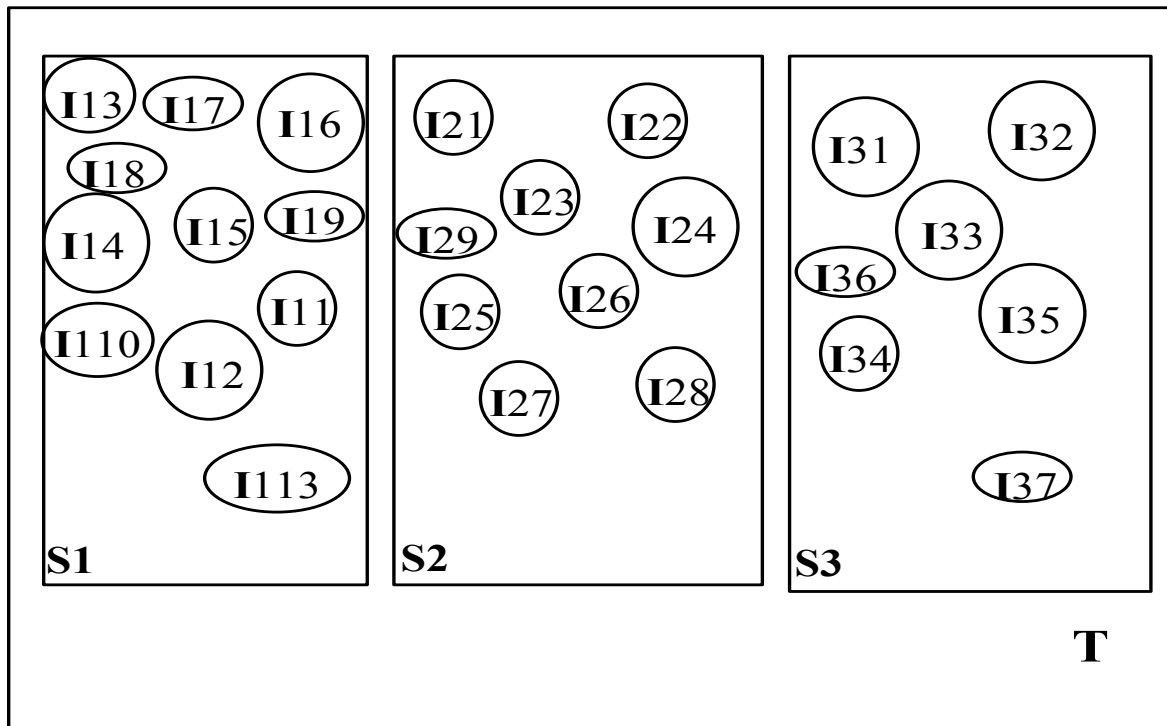
F-statistika

- tzv. fixační indexy
- 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

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

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

$$H_T = 2 p_0 q_0$$

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

F statistiky

$$F_{IS} = \frac{\overline{H}_s - I_I}{\overline{H}_s}$$

Snížení heterozygotnosti jedince kvůli nenáhodnému páření v subpopulaci (~ HWE)

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

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 odlišnosti od nuly – nejčastěji permutace

Výpočet F statistik - příklad

Průměrná frekvence alely A v celé populaci

Lokus	Subpopulace 1 (N ₁ =40)				Subpopulace 2 (N ₂ =20)				p _{0(j)}	Pozn.
	AA	AB	BB	p _{1(j)}	AA	AB	BB	p _{2(j)}		
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(j)}	H _{2(j)}	H _{I(j)}	H _{S(j)}	H _{T(j)}	F _{IS(j)}	F _{ST(j)}	F _{IT(j)}
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

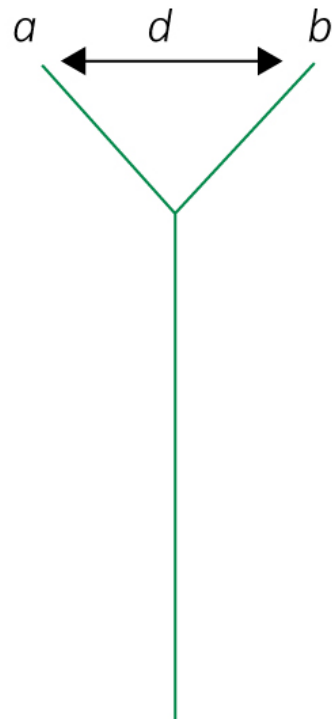
Hodnoty F_{ST}

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

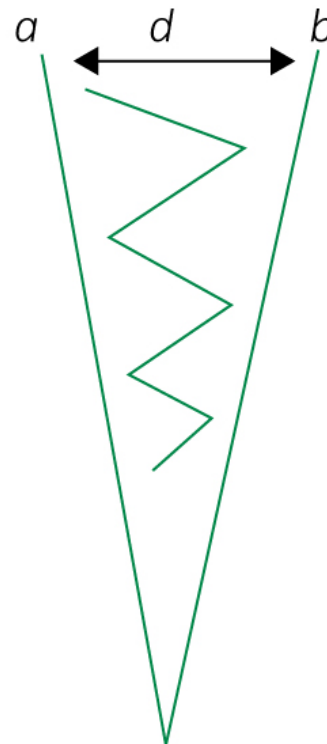
F statistiky

popis výsledku nikoliv příčin → možná alternativní vysvětlení

Recent separation,
no gene flow



Old separation, but
continuous (low)
gene flow



Time

Příklad

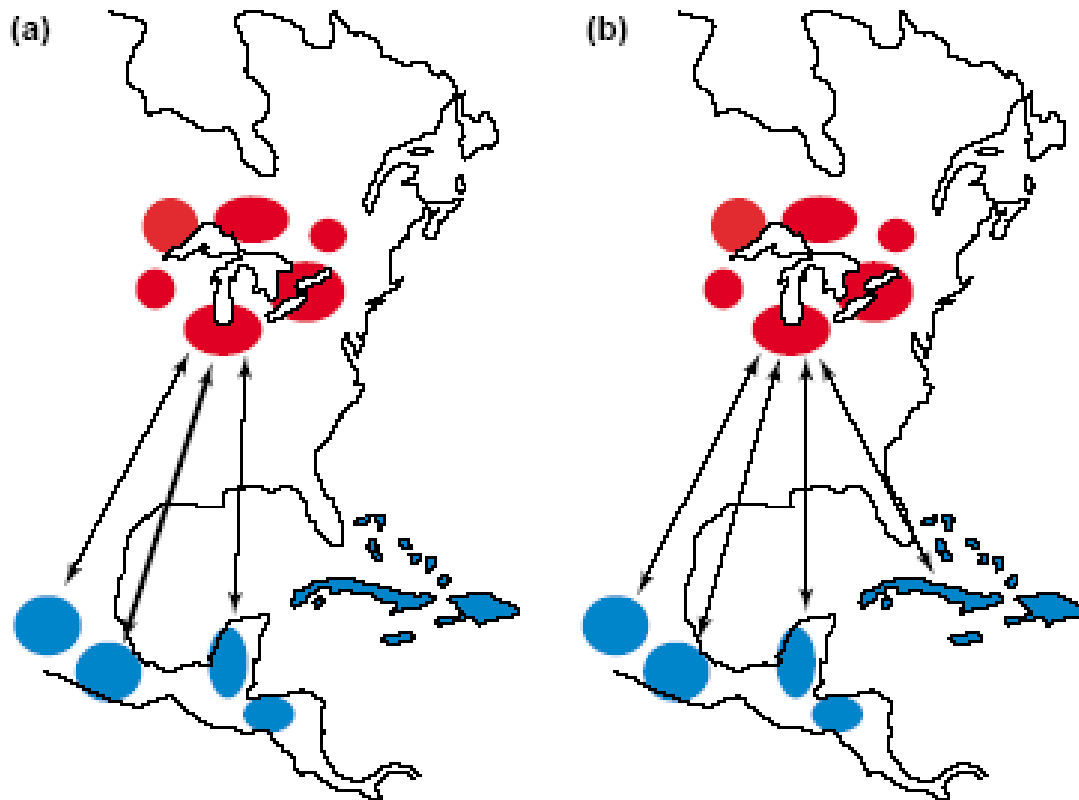
- FSTAT
- Genetix

3. Population assignments

Klasické problémy populační genetiky

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

Unraveling migratory connectivity



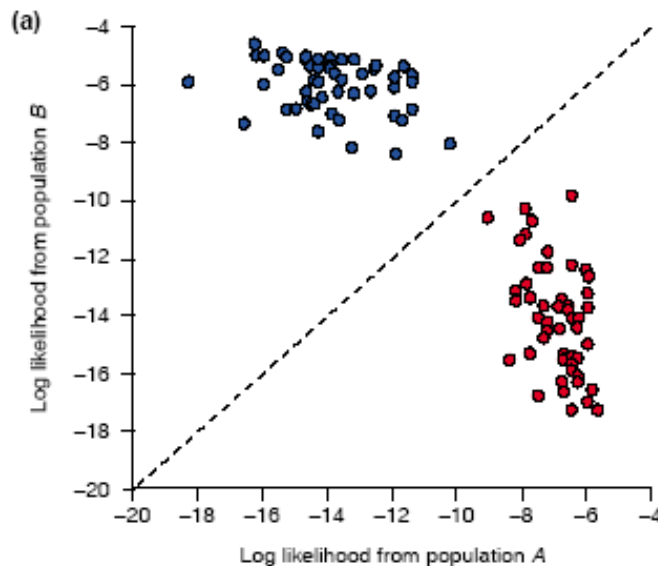
(a) strong vs. (b) weak connectivity

Genetická analýza

- « 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

Population assignment tests

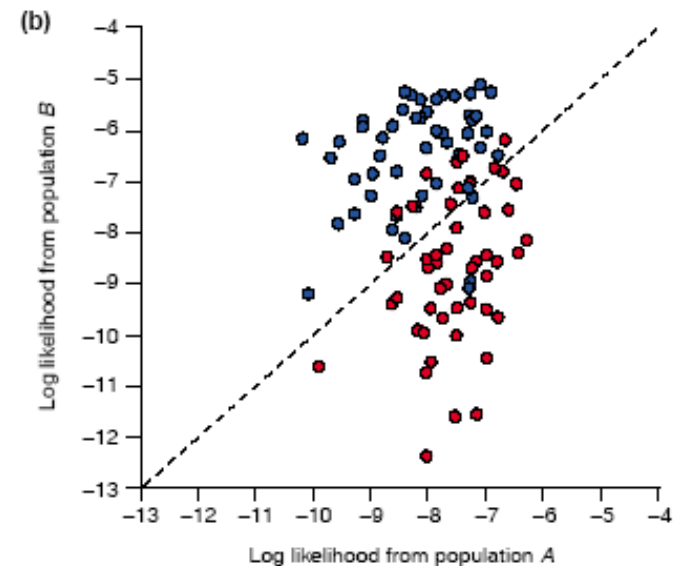
- program GeneClass (Piry et al. 2004)
- calculates the probability that an individual's genotype might exist in a particular population
- can combine data from multiple genetic marker types



5 microsatellite loci

$$F_{st} = 0.14$$

99.9% assigned correctly



5 microsatellite loci

$$F_{st} = 0.04$$

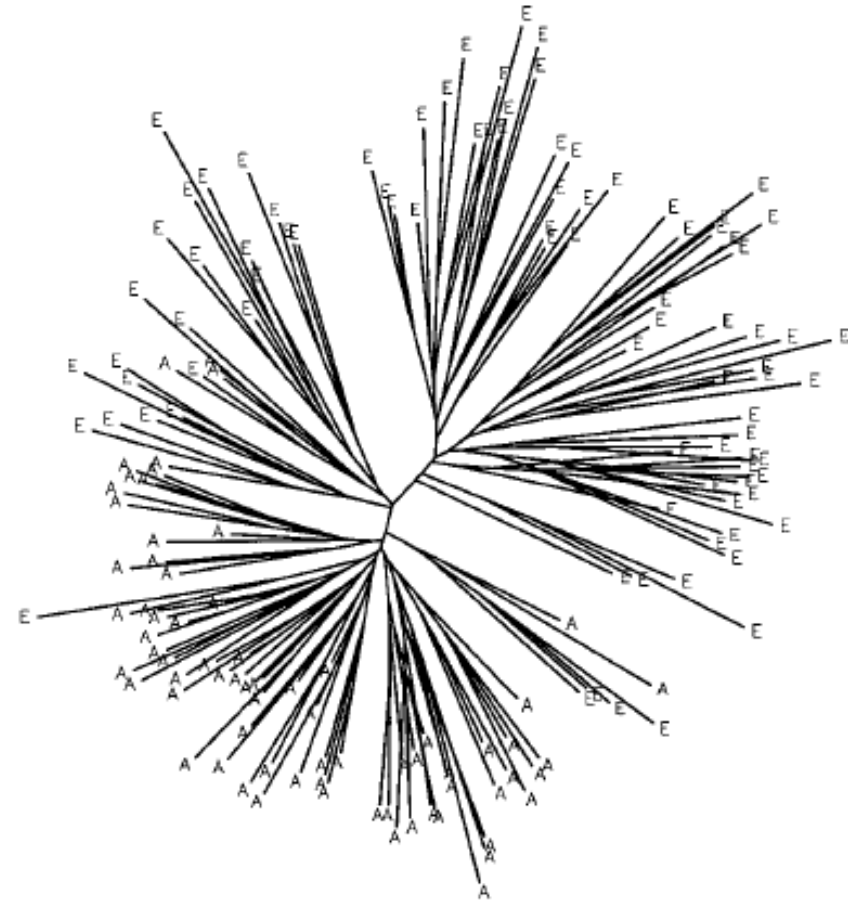
90.2% assigned correctly

Klastrování – hledání „přirozených populací“

- **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

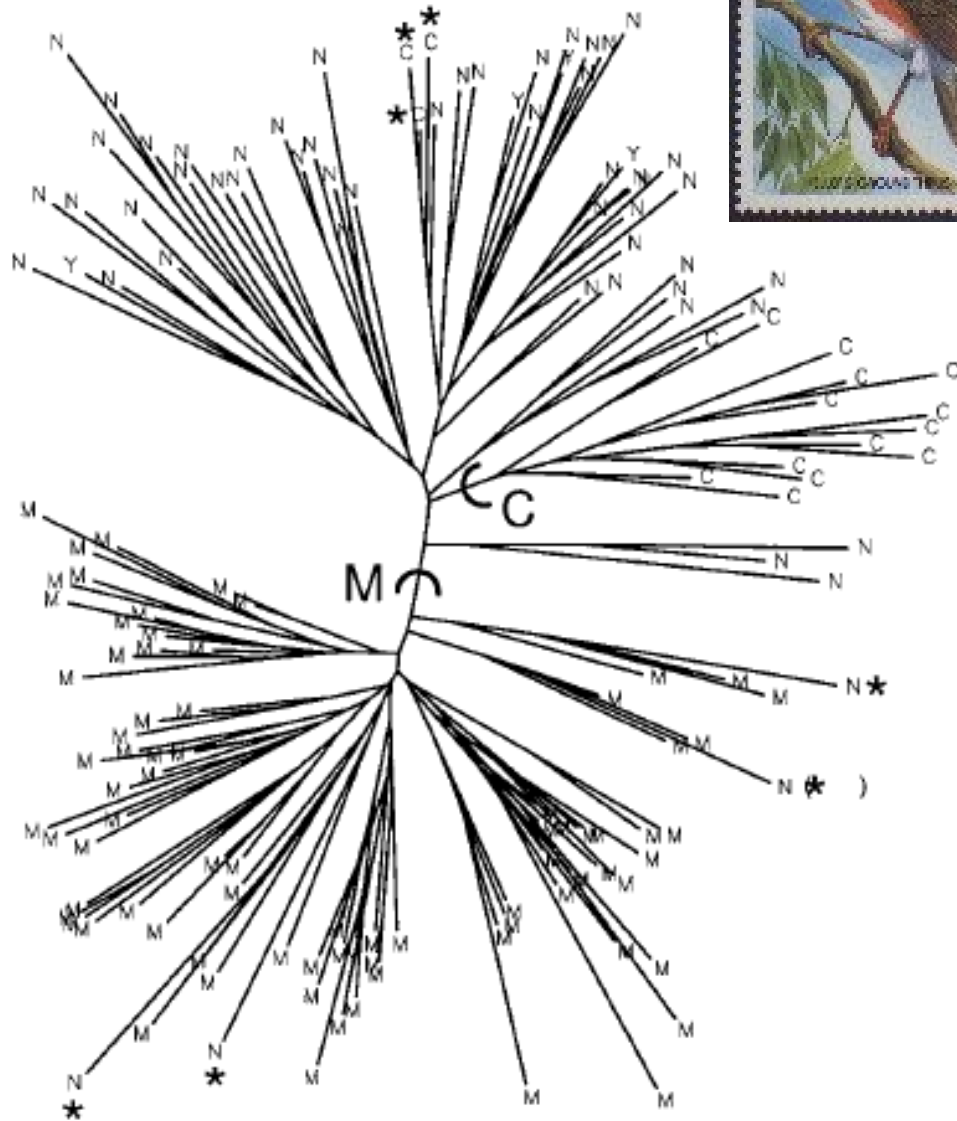
(Maximum likelihood, Bayesianá
metoda)

Bayesian clustering approach

STRUCTURE - *Pritchard et al. 2000*

- Neznámý počet populací charakterizovaných různými frekvencemi alel → počet populací a frekvence zjišťuji
- Současně přiřazuji individua do populací
- Lokusy, které nejsou ve vazbě, HW uvnitř subpopulací (např. mikrosatelity, 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

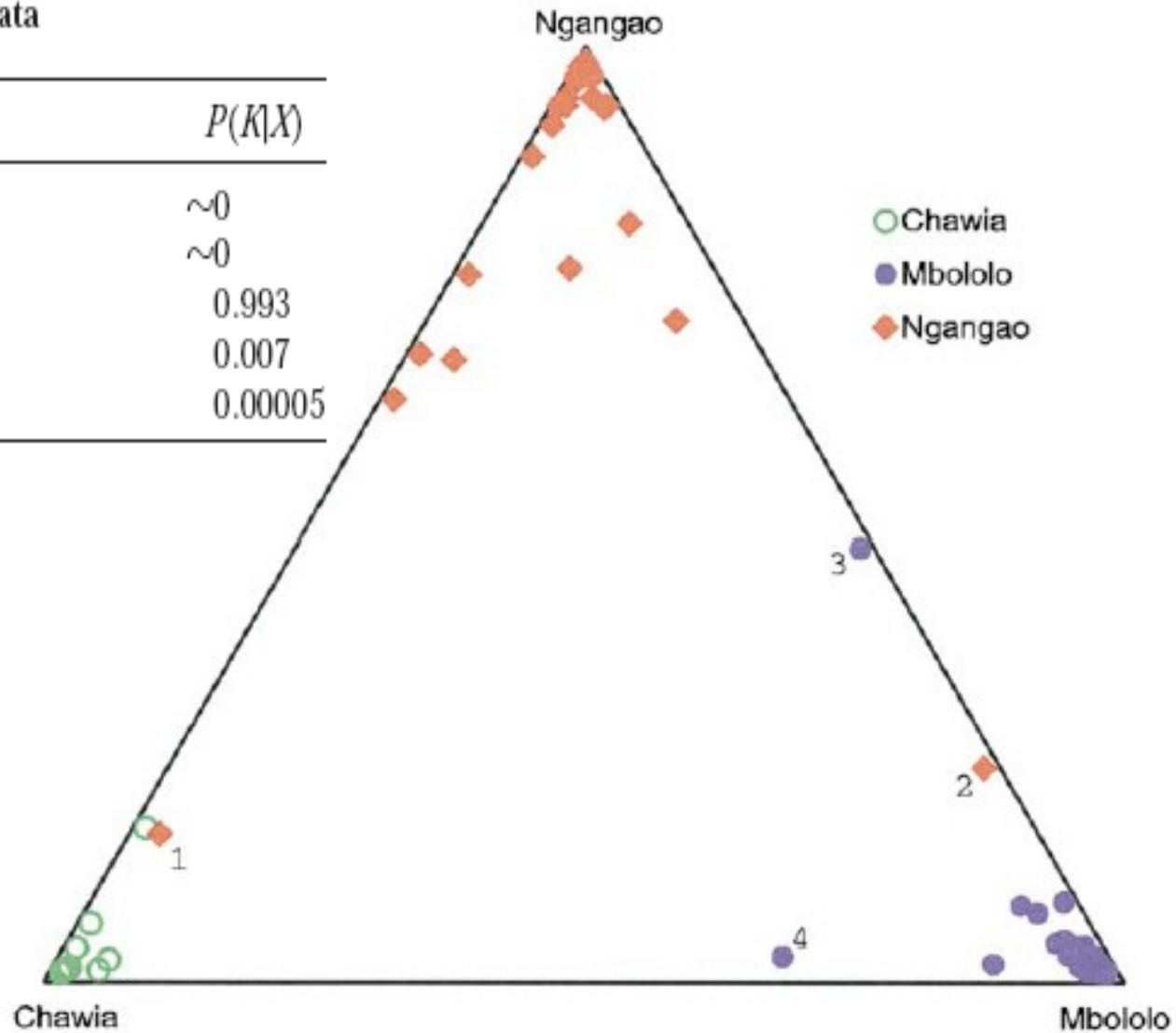


- Fragmenty deštného pralesa
- Lokality Chawia, Ngangao, Mbololo, Yale (Kenya)
- 7 mikrosatelitových lokusů
- Neighbour-joining
- * špatně zařazení jedinci

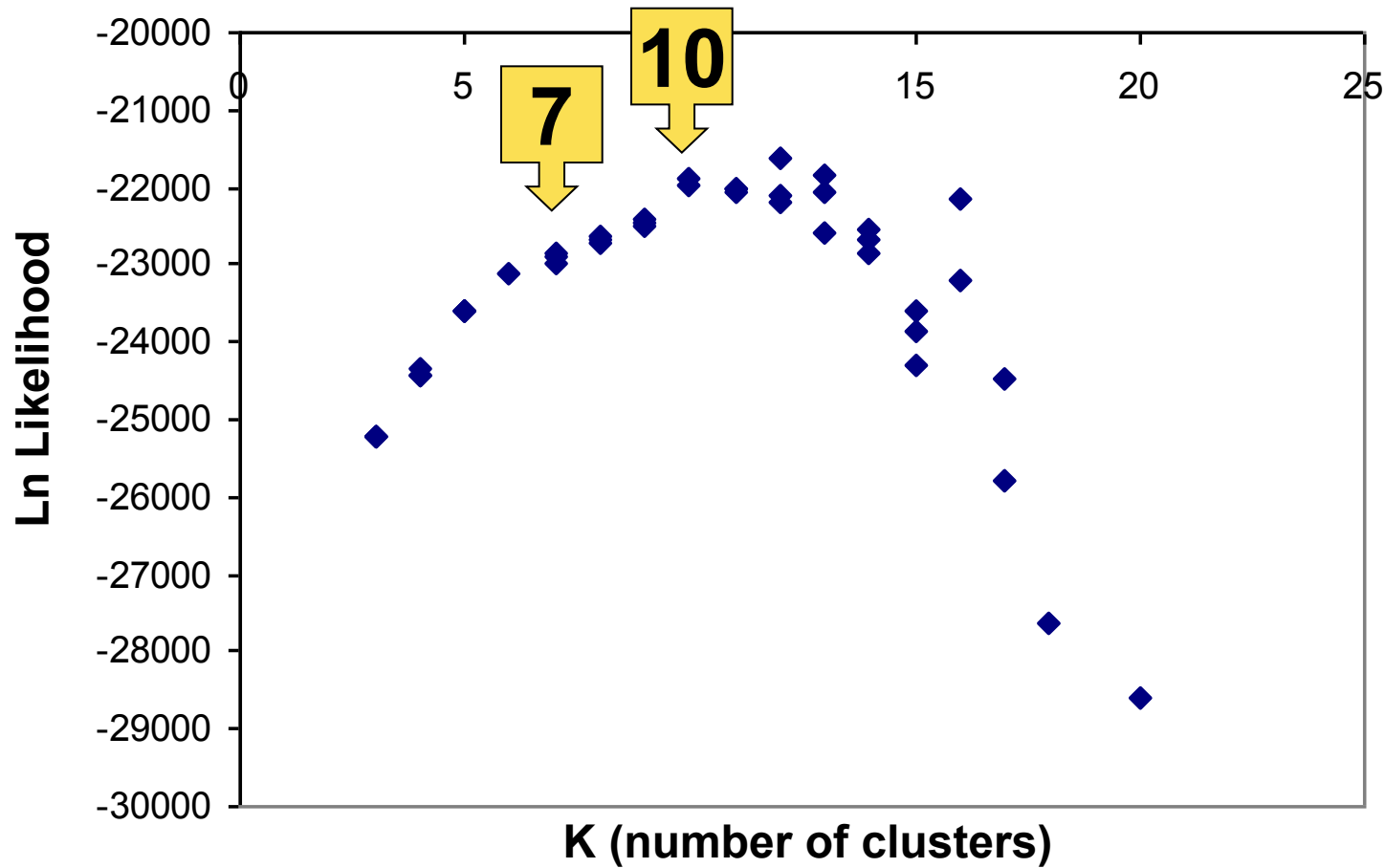
Program STRUCTURE - Bayesiánský přístup

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

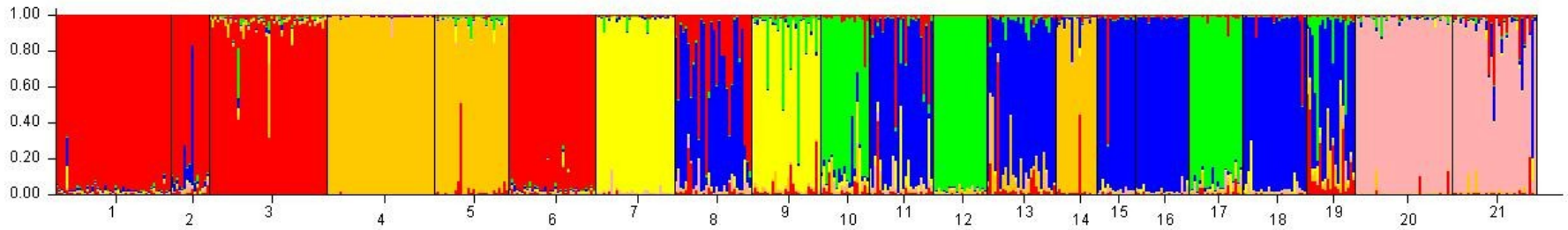
K	$\log P(X K)$	$P(K X)$
1	-3144	~ 0
2	-2769	~ 0
3	-2678	0.993
4	-2683	0.007
5	-2688	0.00005



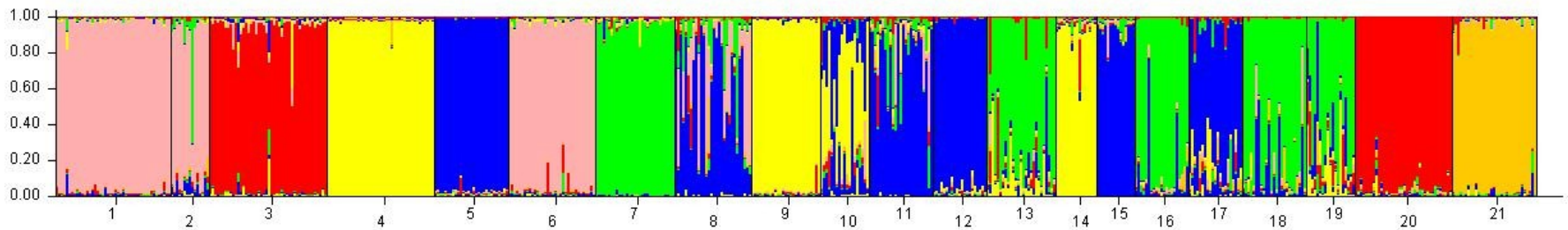
Stanovení počtu „přirozených“ subpopulací



K = 7



K = 10

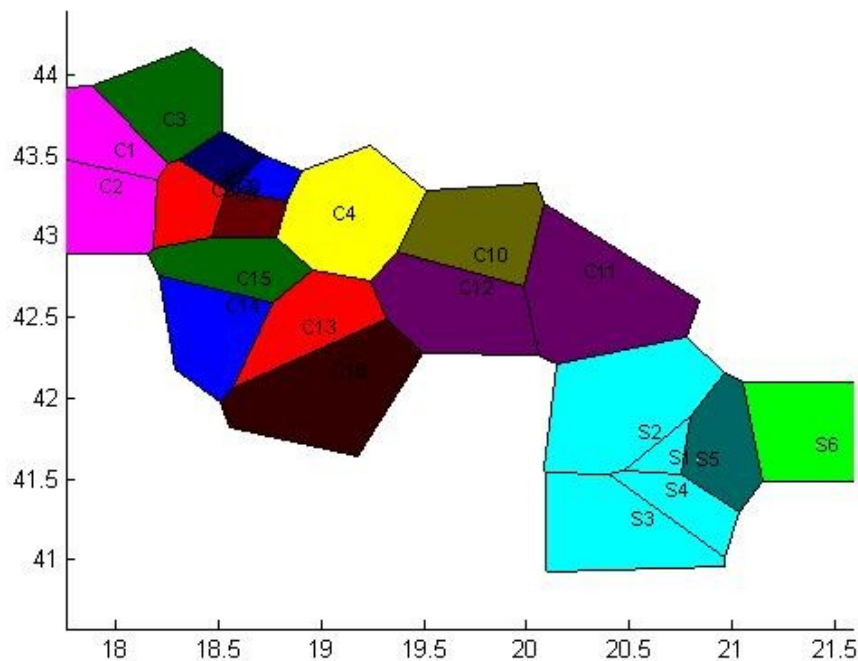


Proporce genomu každého jedince náležející určitému „clusteru“

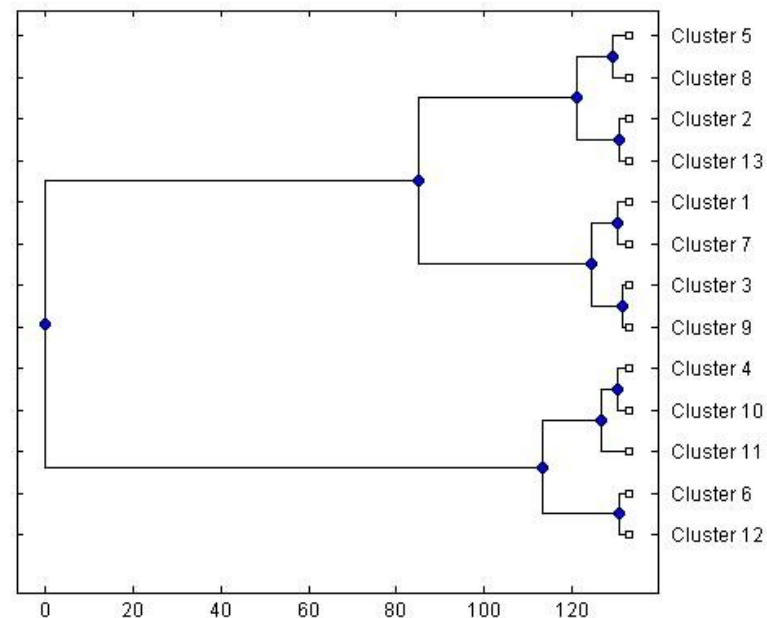
4. Spatially explicit analyses = spatial genetics = landscape genetics

- vychází z Bayesian clustering approach (typu STRUCTURE) – **individual based models**
- do modelování genetické informace přidává i geografické koordináty
- např. programy BAPS, TESS, Geneland (automaticky stanovují nejlepší počet populací K)

The example of very fragmented populations: the best model in BAPs for Central and Southern *Dinaromys* populations (spatial clustering of groups of individuals): K=13 (i.e. evidence of very high structuration)

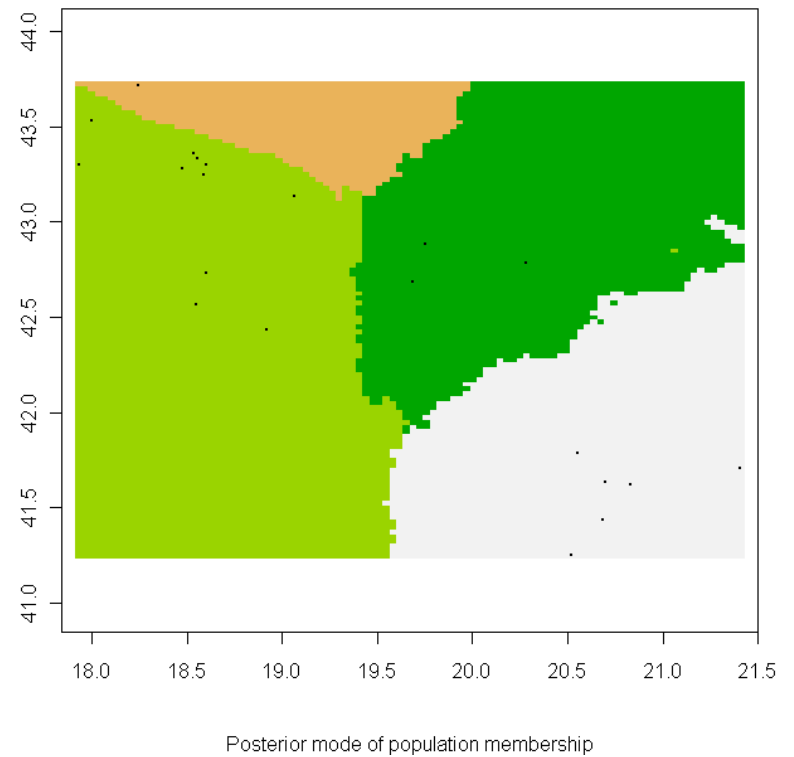
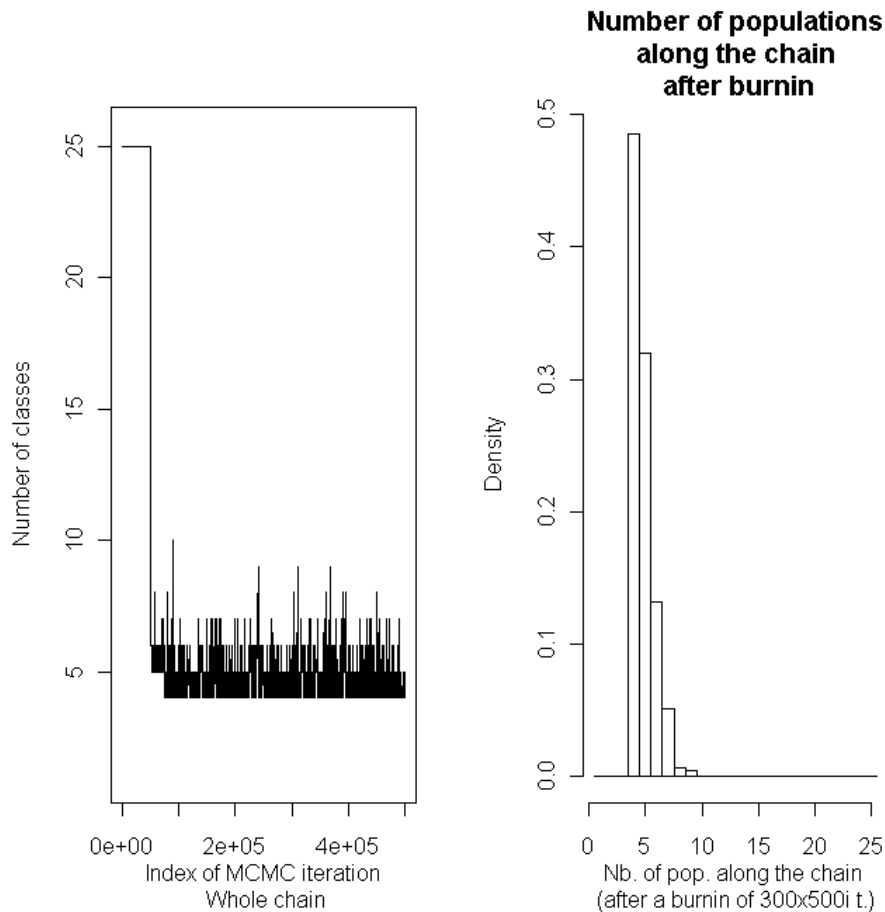


- Best Partition:
- Cluster 1: {C9, C13}
 - Cluster 2: {S6}
 - Cluster 3: {C8, C14}
 - Cluster 4: {C4}
 - Cluster 5: {C1, C2}
 - Cluster 6: {S1, S2, S3, S4}
 - Cluster 7: {C6}
 - Cluster 8: {C3, C15}
 - Cluster 9: {C5, C7}
 - Cluster 10: {C10}
 - Cluster 11: {C11, C12}
 - Cluster 12: {S5}
 - Cluster 13: {C16}



Př. program BAPS

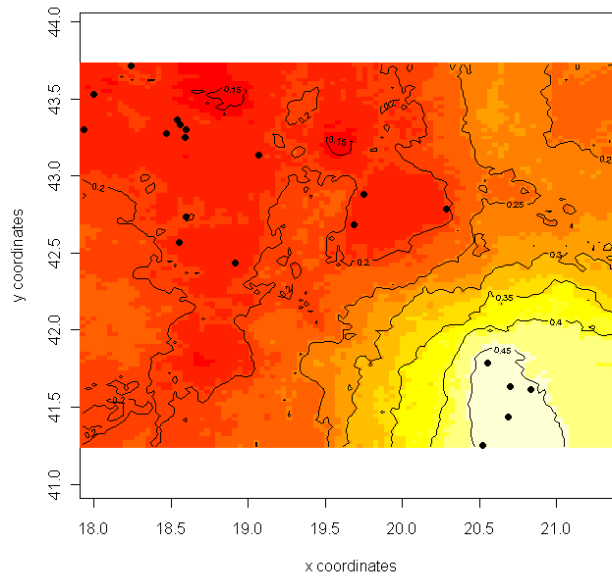
Př.: Geneland



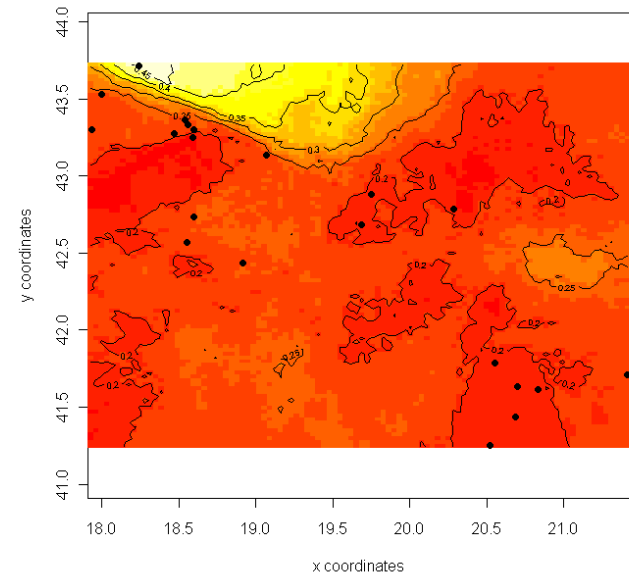
R platform

Posterior probability maps

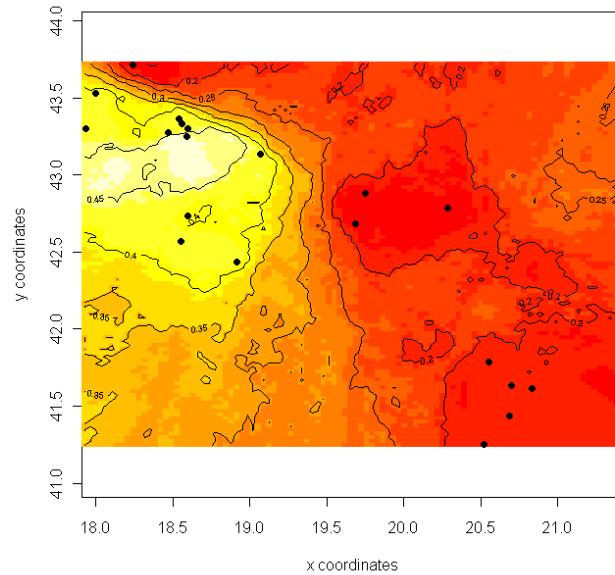
Map of posterior probability to belong to class 4



Map of posterior probability to belong to class 3



Map of posterior probability to belong to class 2



Map of posterior probability to belong to class 1

