



Mend
The
Gap

Smart Integration of Genetics with Sciences of the Past in Croatia: Minding and Mending the Gap



ANALIZA MOLEKULARNE RAZNOLIKOSTI

Prof. dr. sc. Zlatko Šatović

Sveučilište u Zagrebu, Agronomski fakultet

Znanstveni centar izvrsnosti

za bioraznolikost i molekularno oplemenjivanje bilja (CroP-BioDiv)

E-mail: zsatovic@agr.hr

PLAN

- (1) Ljekovita kadulja (*Salvia officinalis* L.)
- (2) Analiza genetske raznolikosti
- (3) Analiza genetske udaljenosti
- (4) Analiza genetske strukture
- (5) Modeliranje ekoloških niša
- (6) Analiza evolucijske povijesti
- (7) Buduća istraživanja

PLAN

- (1) Ljekovita kadulja (*Salvia officinalis* L.)
- (2) Analiza genetske raznolikosti
- (3) Analiza genetske udaljenosti
- (4) Analiza genetske strukture
- (5) Modeliranje ekoloških niša
- (6) Analiza evolucijske povijesti
- (7) Buduća istraživanja

LJEKOVITA KADULJA

- vrsta:

Ljekovita kadulja (*Salvia officinalis L.*)



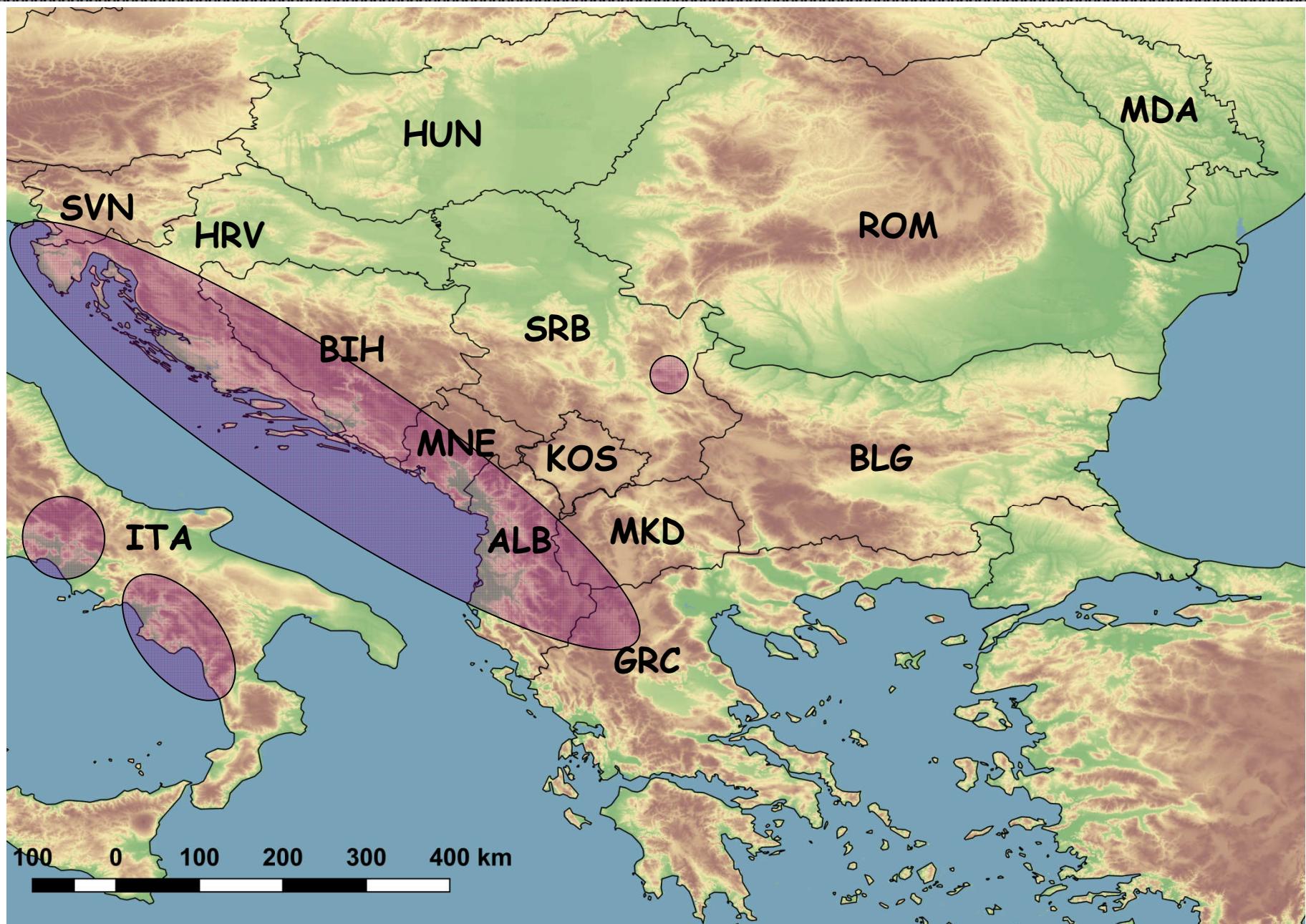
- višegodišnja, stranooplodna biljna vrsta
- eterično ulje: cis- i trans-tujon, kamfor
- raznolikost kemotipova

- rasprostranjenost
obalno područje zapadnog Balkana
središnji i južni Apenini



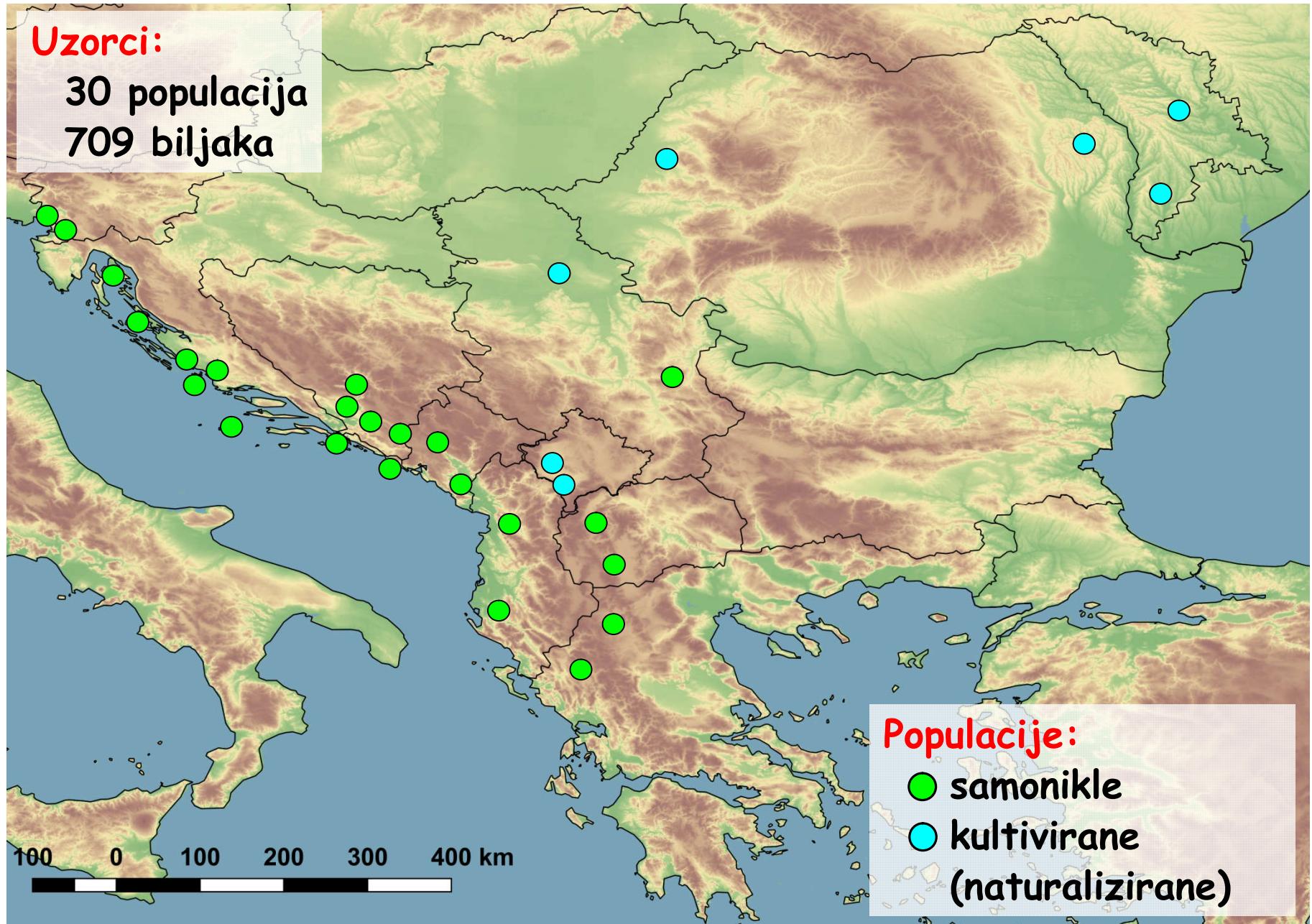
- uzgoj: Sredozemlje, Balkan, Njemačka, SAD, Australija
- prikupljanje u prirodi: Albanija, BiH, Hrvatska
 - negativan utjecaj na bioraznolikost
- naturalizirane (feralne) populacije: Sredozemlje, Balkan
 - biljke koje su pobjegle iz uzgoja

RASPROSTRANJENOST LJEKOVITE KADULJE



UZORKOVANJE

Uzorci:
30 populacija
709 biljaka





Cavtat

MIKROSATELITNI BILJEZI (i)

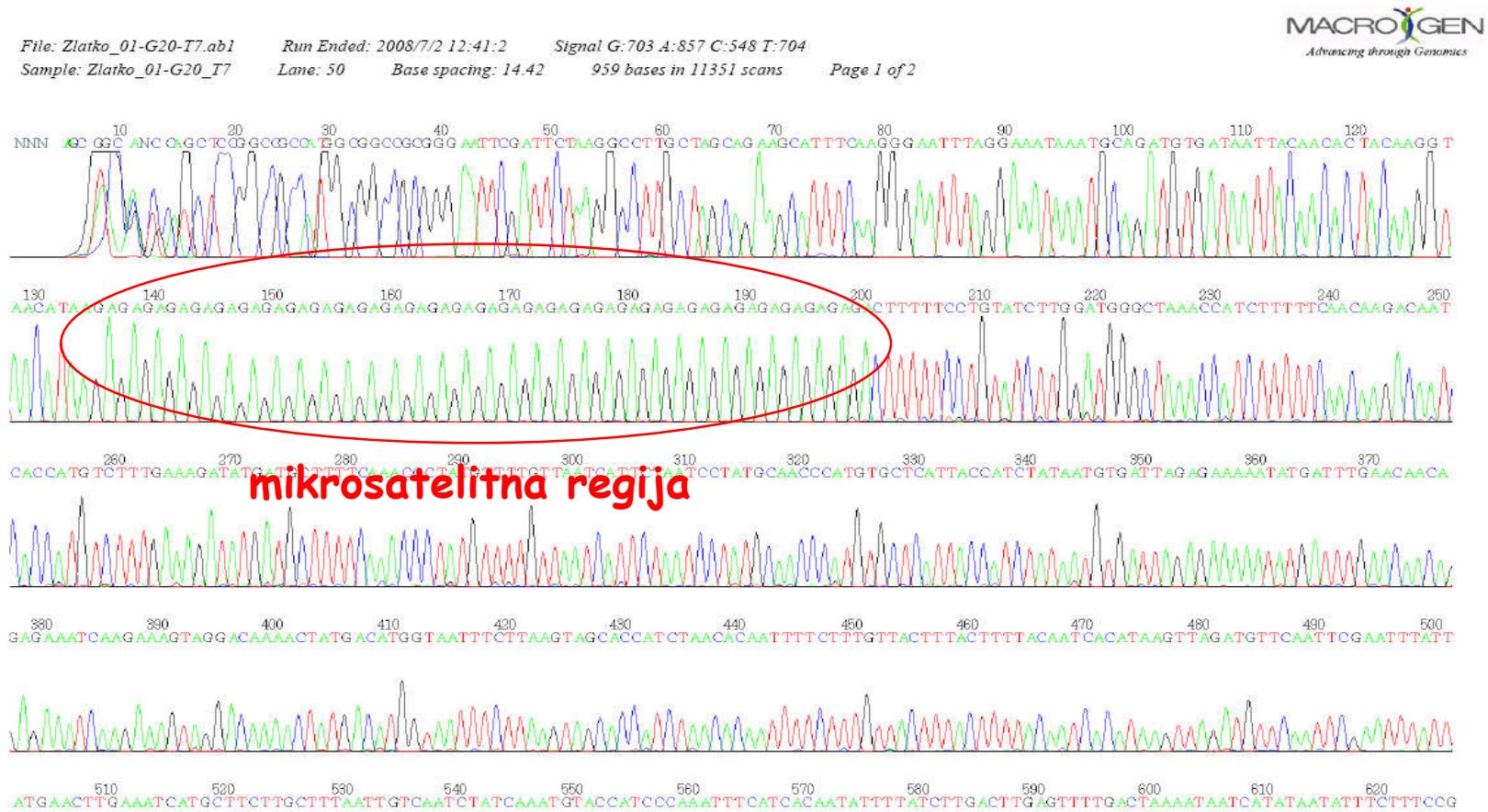
- upotreba para oligonukleotidnih početnica (**primers**) komplementarnima sekvencama DNA koje omeđuju ponavljaču regiju



- početnice su specifične za vrstu (ili skupinu srodnih vrsta)
- ukoliko ne postoji opisane mikrosatelitne početnice za određenu vrstu biljezi se moraju izolirati iz genomnih knjižnica (**genomic library**):
 1. detekcija ulomaka koji sadrže ponavljaču regiju (**microsatellite-enriched genomic library**)
 2. sekvenciranje ulomaka
 3. izrada specifičnih početnica

MIKROSATELITNI BILJEZI (ii)

- izrada genomne knjižnice obogaćene ulomcima koji sadržavaju mikrosatelitne motive GA, GT, AGA i ACA
- izrada početnica koje omeđuju mikrosatelitne regije
- 29 mikrosatelitnih biljega testirano na 25 biljaka



BAZA PODATAKA NUKLEOTIDNIH SEKVENCI



The National Center for
Biotechnology Information
- održava brojne baze podataka

GenBank

- baza podataka sekvenci DNA
- omogućava pohranjivanje i pretraživanje sekvenci
- slobodno dostupna (sekvence se mogu slobodno koristiti u istraživanjima)

Salvia officinalis microsatellite SoUZ001 sequence

GenBank: GQ376512.1

LOCUS GQ376512 782 bp DNA linear PLN 20-DEC-2009
DEFINITION Salvia officinalis microsatellite SoUZ001 sequence.

SOURCE Salvia officinalis (garden sage)

ORGANISM Salvia officinalis

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiids; Lamiales; Lamiaceae; Nepetoideae; Mentheae;
Salvia.

REFERENCE 1 (bases 1 to 782)

AUTHORS Radosavljevic,I., Jakse,J., Javornik,B., Satovic,Z. and Liber,Z.

TITLE Isolation and characterization of polymorphic microsatellites loci
from common sage (Salvia officinalis L.; Lamiaceae)

JOURNAL Mol Ecol Resour (2010)

ORIGIN

1 tgttcagtca acgttagatttataatagac tatttttttata tgattttaca ctaaaaaaaac
61 cccaaaatgt tagttacatt aatttatgtt aaaggcccggt tggatcatgtttatcatt
121 cctgtttttac caaaaactaaa tcaacacgggaaaagaaaag caaatatata tttatgtatgt
181 tatagttgaa ccaaaactgtac taaatcttaa ggggcgcata atttcaagaa caaccacacc
241 aaatcaccac ccaaaaacggaa aagaattttgcataaatttcttacaaatcc tgaccgc
301 tcatgccacg ttt
361 acggatggatccacccaaacttcgttg ggtccccacaa acccaactttttaatcatcgat
421 caccaaccgc ggccggccggccggccggccggccggccggccggccggccggccggccggcc
481 cacccggccggccggccggccggccggccggccggccggccggccggccggccggccggccggcc
541 atcagcaaac geactttccc acaacaacaa gagagagagagagagagagagagagag
601 ttggcagaa ccagtt
661 caggcatcg gagggtggc acaagagatcttttgcataatgtt gatgacatcttaccaggtat
721 cgtcataaat gcatt
781 ca

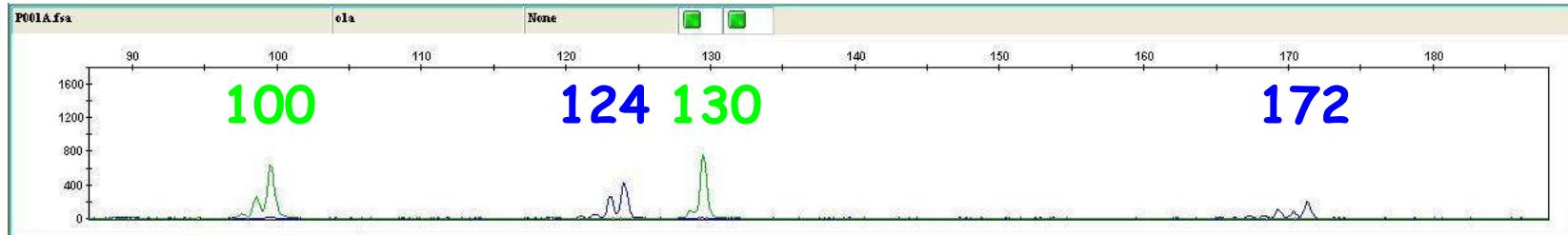
//

OČITAVANJE POLIMORFIZAMA

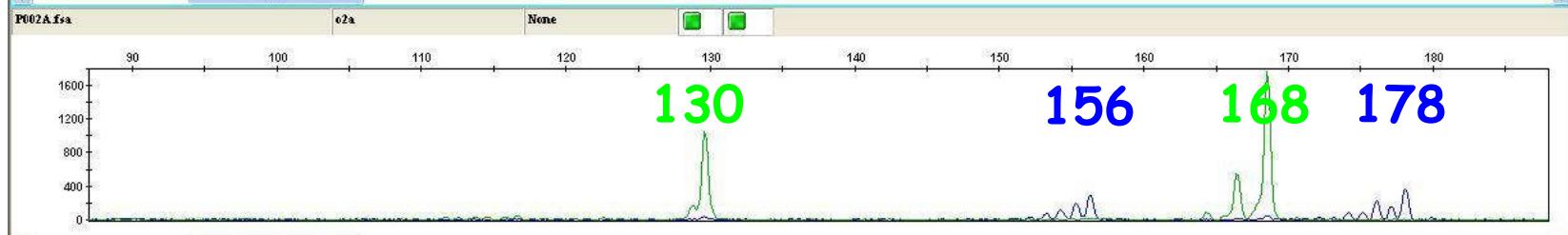
Biljezi SSR

Jedinke

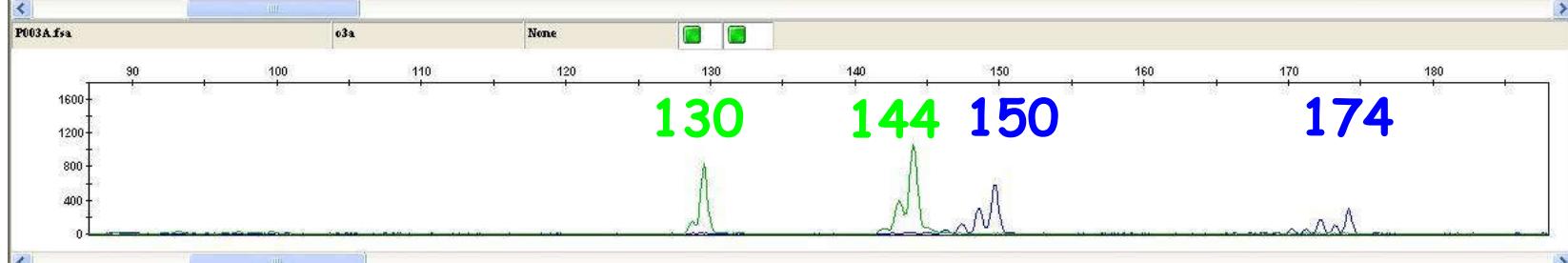
S1



S2



S3



Matrica podataka:

Jedinke

Biljezi

M1

M2

S1

S2

S3

	100	130	130	168	130	144
	124	172	156	178	150	174

UZORCI / ANALIZA

- uzorci:

30 populacija / 709 biljaka

- 23 samonikle populacije

- Slovenija (2), Hrvatska (8), BiH (4), Crna Gora (2),
Albanija (2), Makedonija (2), Grčka (2), Srbija (1)

- 7 kultiviranih/naturaliziranih populacija

- Srbija (1), Kosovo (2), Rumunjska (2), Moldavija (2)

- analiza:

8 biljega SSR (odabrani su najpolimorfni)

- dinukleotidni motiv

- SoUZ001, SoUZ002, SoUZ003, SoUZ007, SoUZ011

- trinukleotidni motiv

- SoUZ013, SoUZ014, SoUZ019

ukupno 165 alela

PODACI

Populacija	Jedinka	SoUZ001	SoUZ001	SoUZ002	SoUZ002	...	SoUZ019	SoUZ019
P01	S001	173	189	183	185		134	150
P01	S002	165	199	185	185		150	153
P01	S003	183	185	195	197		150	159
P01	S004	163	187	189	195		150	150
P01	S005	165	167	187	195		134	150
P01	S006	165	183	185	187		134	156
P01	S007	173	183	185	195		150	159
P01	S008	165	167	183	195		134	134
P01	S009	165	171	181	183		147	150
P01	S010	165	173	195	201		150	159
...								
P30	S709	161	165	185	195		134	150

LJEKOVITA KADULJA: INFORMATIVNOST BILJEGA

- informativnost biljega na temelju 30 populacija / 709 biljaka

Biljeg	Ponavljači motiv	Raspon (pb)	N_a	PIC
SoUZ001	(AG) ₁₅	159-221	30	0.937
SoUZ002	(TG) ₁₁	177-217	15	0.771
SoUZ003	(GT) ₁₃	160-216	23	0.778
SoUZ007	(GT) ₁₁	138-210	13	0.633
SoUZ011	(GA) ₂₅	156-212	29	0.923
SoUZ013	(AAC) ₈	179-215	13	0.821
SoUZ014	(AGA) ₁₀	175-244	22	0.883
SoUZ019	(AGA) ₁₆	132-201	20	0.750
Prosjek			20.63	0.812
Ukupno				165

- N_a - broj alela

- PIC - informacijski sadržaj polimorfizma

PLAN

- (1) Ljekovita kadulja (*Salvia officinalis* L.)
- (2) Analiza genetske raznolikosti**
- (3) Analiza genetske udaljenosti
- (4) Analiza genetske strukture
- (5) Modeliranje ekoloških niša
- (6) Analiza evolucijske povijesti
- (7) Buduća istraživanja

(2) ANALIZA GENETSKE RAZNOLIKOSTI

Mjerila unutarpopulacijske raznolikosti

- (1) Prosječan broj alela u populaciji (N_{av})
- (2) Alelno bogatstvo (N_{ar})
- (3) Broj jedinstvenih alela (N_{pr})
- (5) Zapažena heterozigotnost (H_O)
- (6) Očekivana heterozigotnost (H_E)
- (7) Fiksacijski indeks (f)

MJERILA UNUTARPOPULACIJSKE RAZNOLIKOSTI (i)

(1) Prosječan broj alela u populaciji (N_{av})

	Samonikle	Kultivirane
Broj populacija	23	7
N_{av}	8.696	3.821

(2) Alelno bogatstvo (N_{ar})

- mjerilo alelne raznolikosti neovisno o veličini populacije
- najmanja uzorkovana populacija: 20 jedinki

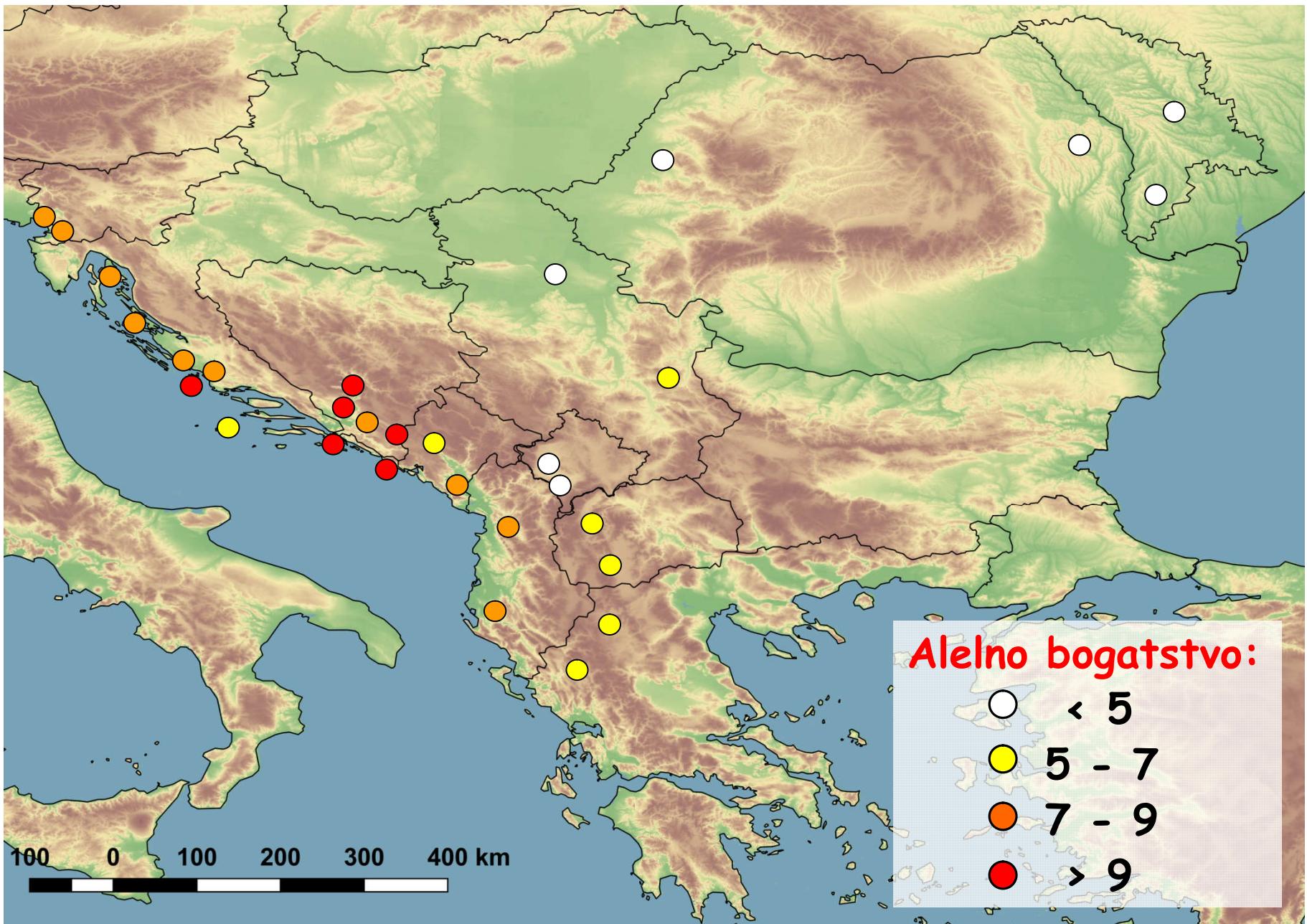
	Raspon	5.13 >> 10.30	2.71 >> 4.19
N_{ar}	7.920	3.672	

(3) Broj jedinstvenih alela (N_{pr})

- aleli zapaženi samo u jednoj populaciji ili skupini populacija (samonikle vs. kultivirane/naturalizirane)

	P(N_{ar})	P < 0.001
N_{pr} (ukupno po populacijama)	20	0
N_{pr} (samonikle vs. kultivirane)	115	0

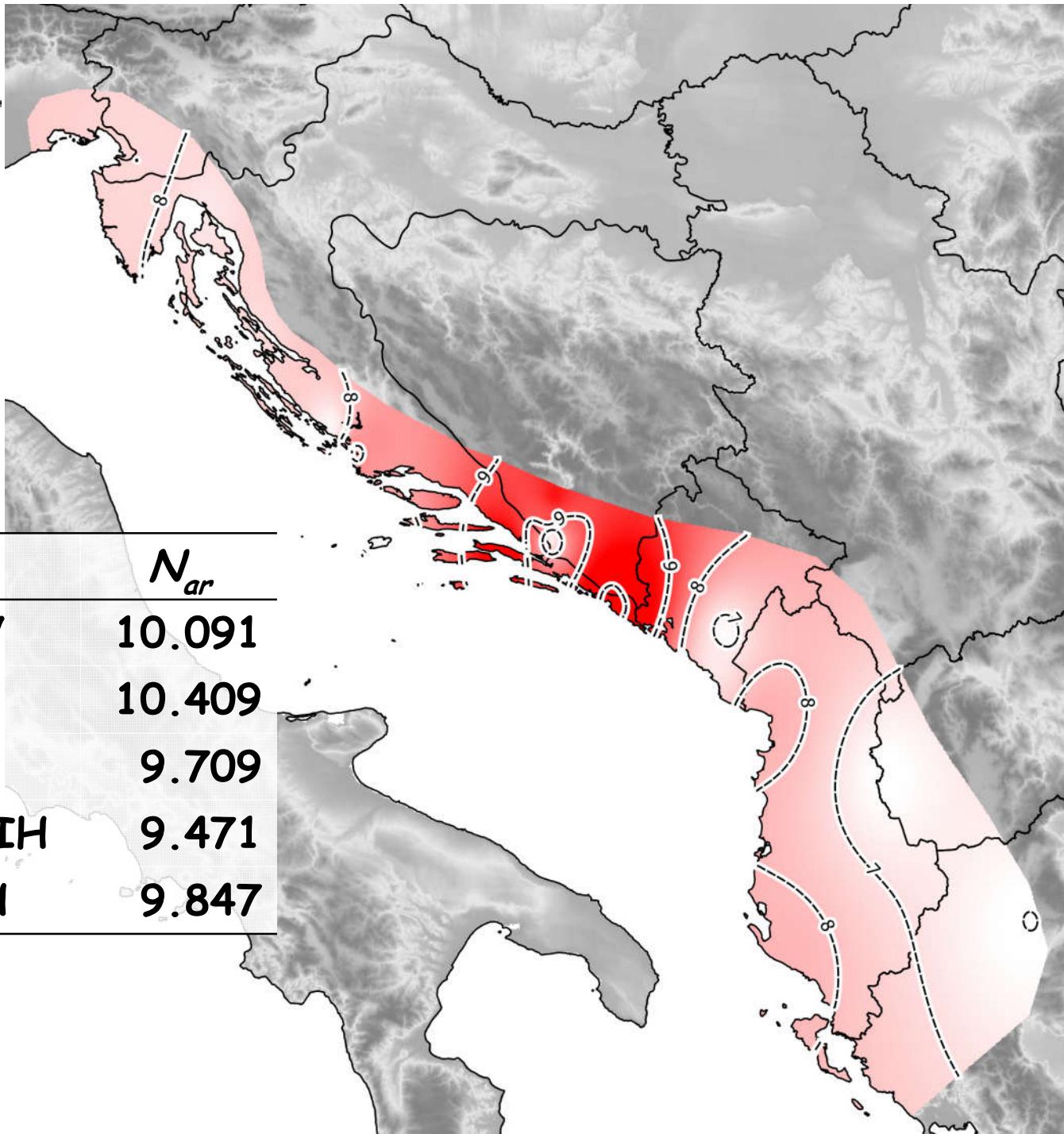
ALELNO BOGATSTVO (N_{ar})



Alelno bogatstvo:

- < 5
- 5 - 7
- 7 - 9
- > 9

N_{ar}



Broj	Populacija	N_{ar}
P09	Pelješac, HRV	10.091
P10	Konavle, HRV	10.409
P12	Mostar, BIH	9.709
P13	Međugorje, BIH	9.471
P14	Trebinje, BIH	9.847

MJERILA UNUTARPOPULACIJSKE RAZNOLIKOSTI (ii)

(3) Zapažena heterozigotnost (H_o)

- prosječan udio heterozigotnih jedinki u populaciji

	Samonikle	Kultivirane
Broj populacija	23	7
H_o	0.733	0.590
Raspon	0.54 >> 0.85	0.31 >> 0.79
H_E	0.751	0.545
Raspon	0.61 >> 0.85	0.38 >> 0.61
$P(H_E)$	$P < 0.001$	

$$H_E = 1 - \sum_{i=1}^I p_i^2$$

p_i - učestalost alela i
 I - ukupni broj alela

FIKSACIJSKI INDEKS

(6) Fiksacijski indeks (f)

- mjerilo odstupanja od HWE

$$f = 1 - \frac{H_O}{H_E}$$

- H_O - zapažena heterozigotnost
- H_E - očekivana heterozigotnost

$f = 0$: populacija je u HWE

$f > 0$: nedostatak heterozigota

$f < 0$: višak heterozigota

- testiranje signifikantnosti:
algoritma Markovljevih lanaca
(Markov chain algorithm)
>> program Genepop

Broj	Populacija	H_O	H_E	f	
P01	SVN_1	0.745	0.795	0.063	ns
P02	SVN_2	0.854	0.795	-0.075	ns
P03	HRV_1	0.717	0.734	0.024	ns
P04	HRV_2	0.754	0.759	0.008	ns
P05	HRV_3	0.707	0.708	0.002	ns
P06	HRV_4	0.698	0.743	0.061	ns
P07	HRV_5	0.736	0.771	0.045	ns
P08	HRV_6	0.717	0.712	-0.007	ns
P09	HRV_7	0.754	0.769	0.019	*
P10	HRV_8	0.825	0.847	0.026	*
P11	BIH_1	0.817	0.758	-0.077	ns
P12	BIH_2	0.762	0.796	0.043	*
P13	BIH_3	0.818	0.815	-0.004	ns
P14	BIH_4	0.777	0.833	0.067	ns
P15	MNE_1	0.693	0.745	0.070	ns
P16	MNE_2	0.699	0.718	0.026	ns
P17	ALB_1	0.744	0.766	0.029	ns
P18	ALB_2	0.743	0.802	0.073	*
P19	MKD_1	0.683	0.699	0.024	ns
P20	MKD_2	0.720	0.725	0.007	ns
P21	GRC_1	0.679	0.673	-0.009	ns
P22	GRC_2	0.672	0.702	0.042	ns
P23	KOS_1	0.598	0.561	-0.067	ns
P24	KOS_2	0.505	0.561	0.100	*
P25	SRB_1	0.313	0.377	0.171	ns
P26	SRB_2	0.537	0.612	0.123	**
P27	ROU_1	0.783	0.610	-0.283	***
P28	ROU_2	0.793	0.612	-0.296	***
P29	MDA_1	0.592	0.595	0.004	ns
P30	MDA_2	0.545	0.497	-0.096	ns
	Prosjek	0.699	0.703	0.004	

PLAN

- (1) Ljekovita kadulja (*Salvia officinalis* L.)
- (2) Analiza genetske raznolikosti
- (3) Analiza genetske udaljenosti**
- (4) Analiza genetske strukture
- (5) Modeliranje ekoloških niša
- (6) Analiza evolucijske povijesti
- (7) Buduća istraživanja

ANALIZA GENETSKE UDALJENOSTI

1. Genetska udaljenost između populacija

- na temelju alelnih učestalosti

2. Genetska udaljenost između jedinki

- na temelju multilokusnih genotipova
- ovisi o tipu biljega u analizi:
 - kodominantni biljezi (SSR; SNP)
 - dominantni biljezi (AFLP)

LJEKOVITA KADULJA: ALELNA UČESTALOST

Population	138	140	146	192	194	196	198	199	200	202	204	206	208	209	210
P01						0.348	0.022		0.239	0.391					
P02						0.438	0.021		0.021	0.229	0.292				
P03						0.680			0.120	0.080	0.120				
P04						0.620	0.060		0.060	0.180	0.080				
P05						0.375	0.021		0.292	0.292	0.021				
P06						0.125	0.708	0.021	0.042	0.083	0.021				
P07			0.022			0.174	0.304		0.174	0.109	0.152	0.022	0.044		
P08						0.326	0.022		0.087	0.413	0.065		0.087		
P09					0.020	0.640	0.060		0.020	0.240			0.020		
P10						0.520	0.040		0.060	0.060	0.260	0.020	0.040		
P11						0.500	0.021		0.125	0.042	0.229		0.042	0.042	
P12						0.625	0.125		0.042	0.125	0.083				
P13						0.580	0.080		0.220	0.100	0.020				
P14	P17 samonikla populacija iz Albanije					0.500			0.023	0.091	0.386				
P15						0.479	0.146		0.125	0.021	0.229				
P16						0.500			0.091	0.023	0.386				
P17	0.042 0.042 0.042				0.021	0.604			0.146	0.083	0.021				
P18						0.239			0.022	0.304	0.370		0.065		
P19						0.479			0.021	0.271	0.229				
P20						0.455				0.432	0.114				
P21						0.604				0.271	0.125				
P22						0.659				0.227	0.114				
P23	Fiksacija alela u kultiviranim populacijama					0.348				0.652					
P24						0.458				0.542					
P25						0.225				0.775					
P26						0.729			0.167	0.042			0.063		
P27						0.625			0.083		0.292				
P28			0.023			0.568					0.409				
P29						0.696			0.044	0.022	0.239				
P30						0.875					0.125				

Biljeg
SoUZ007

$N_a = 15$

Raspon
(pb):
138-210

Jedinstveni
aleli

$N_{pr} = 5$

Najčešći
aleli
196, 204

MJERILA UDALJENOSTI IZMEĐU POPULACIJA

(A) Evolucijska mjerila udaljenosti

- pretpostavka: genetske razlike su uzrokovane novonastalim mutacijama i genetskim otklonom

- (1) Standardna genetska udaljenost po Neiju (Nei, 1972)
- (2) Nejeva nepristrana udaljenost (Nei, 1978)
- (3) Hillisova udaljenost (Hillis, 1984)
- (4) Hillisova nepristrana udaljenost (Swofford i Olsen, 1990)

(B) Geometrijska mjerila udaljenosti

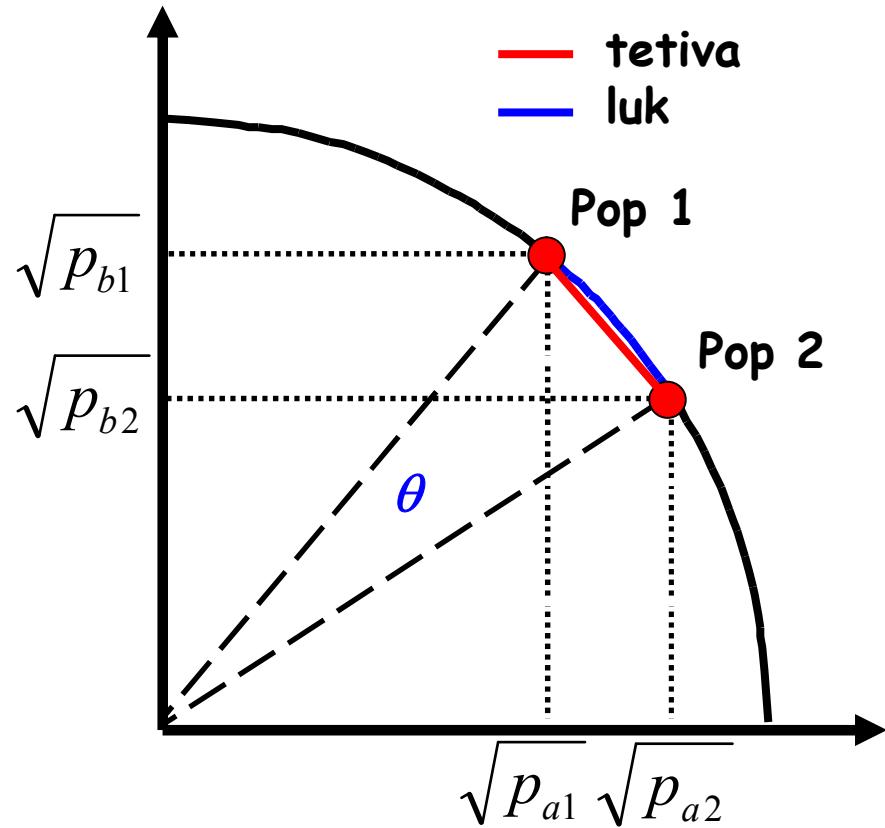
- pretpostavka: genetske razlike su uzrokovane isključivo genetskim otklonom

- (1) Tetivna udaljenost (Cavalli-Sforza i Edwards, 1967)
- (2) Lučna udaljenost (Cavalli-Sforza i Edwards, 1967)
- (3) Rogersova udaljenost (Rogers, 1972)

TETIVNA I LUČNA UDALJENOST

Cavalli-Sforza and Edwards' (1967) chord and arc distances

- populacije su točke u višedimenzionalnoj hipersferi
- m -dimenzija: broj dimenzija ovisi o ukupnom broju alela u obje populacije
- hipersfera je definirana korijenom alelnih učestalosti na svakom lokusu
- **tetiva** - dužina koja spaja dvije točke kružnice
- **luk** - dio kružnice omeđen dvjema točkama



Tetivna udaljenost:

$$D_{CHORD} = \frac{2}{\pi} \left(1 - \sum_{l=1}^L \sum_{a=1}^A \sqrt{p_{al1} p_{al2}} \right)$$

p_{al1} - učestalost alela a lokusa / u populaciji 1

LJEKOVITA KADULJA: MATRICA TETIVNE UDALJENOSTI

P01	P02	P03	P04	P05	P06	P07	P08	P09	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28	P29	P30		
P01																															
P02	0.019																														
P03	0.033	0.026																													
P04	0.045	0.048	0.049																												
P05	0.058	0.054	0.057	0.059																											
P06	0.054	0.054	0.062	0.064	0.055																										
P07	0.053	0.050	0.062	0.062	0.037	0.036																									
P08	0.083	0.084	0.083	0.069	0.076	0.074	0.068																								
P09	0.053	0.050	0.052	0.052	0.054	0.061	0.053	0.061																							
P10	0.044	0.047	0.061	0.054	0.070	0.052	0.052	0.075	0.051																						
P11	0.062	0.062	0.067	0.069	0.086	0.068	0.073	0.078	0.064	0.055																					
P12	0.059	0.052	0.063	0.055	0.072	0.057	0.055	0.075	0.048	0.038	0.051																				
P13	0.061	0.060	0.062	0.052	0.062	0.065	0.063	0.069	0.047	0.047	0.052	0.036																			
P14	0.055	0.051	0.063	0.063	0.075	0.061	0.062	0.090	0.063	0.027	0.060	0.049	0.054																		
P15	0.068	0.064	0.066	0.090	0.086	0.081	0.066	0.091	0.074	0.072	0.071	0.067	0.064	0.064																	
P16	0.060	0.061	0.068	0.067	0.082	0.075	0.065	0.066	0.068	0.054	0.052	0.058	0.065	0.055	0.067																
P17	0.086	0.087	0.089	0.082	0.080	0.090	0.088	0.090	0.068	0.067	0.102	0.080	0.068	0.067	0.104	0.089															
P18	0.065	0.061	0.069	0.074	0.078	0.074	0.074	0.083	0.065	0.045	0.071	0.062	0.075	0.048	0.082	0.049	0.089														
P19	0.084	0.077	0.097	0.097	0.101	0.083	0.081	0.085	0.095	0.070	0.090	0.080	0.089	0.074	0.087	0.070	0.093	0.072													
P20	0.094	0.088	0.113	0.096	0.114	0.098	0.092	0.088	0.091	0.079	0.095	0.078	0.086	0.085	0.103	0.075	0.095	0.078	0.030												
P21	0.069	0.071	0.072	0.076	0.084	0.087	0.079	0.091	0.088	0.068	0.096	0.077	0.088	0.081	0.097	0.067	0.067	0.089	0.059	0.085	0.090										
P22	0.073	0.077	0.078	0.076	0.093	0.084	0.078	0.083	0.084	0.062	0.082	0.071	0.080	0.073	0.084	0.047	0.088	0.057	0.077	0.081	0.021										
P23	0.074	0.082	0.080	0.100	0.110	0.095	0.100	0.118	0.094	0.090	0.096	0.110	0.105	0.083	0.123	0.104	0.124	0.099	0.128	0.132	0.125	0.130									
P24	0.070	0.079	0.076	0.094	0.108	0.097	0.098	0.109	0.091	0.087	0.092	0.107	0.104	0.082	0.114	0.098	0.125	0.099	0.130	0.137	0.120	0.123	0.009								
P25	0.083	0.091	0.087	0.111	0.102	0.116	0.110	0.127	0.110	0.113	0.117	0.124	0.128	0.105	0.131	0.114	0.137	0.113	0.139	0.138	0.138	0.142	0.042	0.045							
P26	0.081	0.081	0.094	0.091	0.092	0.092	0.084	0.089	0.072	0.075	0.079	0.075	0.092	0.087	0.085	0.071	0.106	0.073	0.099	0.105	0.097	0.086	0.126	0.122	0.131						
P27	0.068	0.080	0.073	0.084	0.093	0.086	0.092	0.107	0.087	0.082	0.086	0.101	0.090	0.077	0.109	0.091	0.110	0.093	0.123	0.126	0.113	0.116	0.013	0.016	0.038	0.117					
P28	0.082	0.093	0.093	0.103	0.119	0.107	0.108	0.124	0.106	0.091	0.094	0.110	0.103	0.086	0.121	0.090	0.134	0.095	0.126	0.126	0.121	0.127	0.028	0.031	0.060	0.120	0.027				
P29	0.070	0.085	0.079	0.090	0.094	0.091	0.087	0.108	0.086	0.088	0.090	0.103	0.098	0.083	0.108	0.101	0.120	0.099	0.123	0.129	0.109	0.113	0.024	0.022	0.052	0.114	0.017	0.036			
P30	0.075	0.083	0.078	0.092	0.097	0.101	0.097	0.114	0.086	0.096	0.091	0.107	0.103	0.084	0.110	0.098	0.119	0.108	0.133	0.136	0.117	0.121	0.025	0.020	0.048	0.115	0.019	0.035	0.024		

- prosječna udaljenost: 0.081
- minimalna udaljenost: 0.009 između dvije naturalizirane populacije s Kosova (P23/P24)
- maksimalna udaljenost: 0.142 između samonikle populacije iz Grčke (P22) i kultivirane populacije iz Srbije (P26)

MULTIVARIJATNE METODE

(1) Prikaz podataka u obliku stabla

- razvrstavanje jedinki/populacija u skupine sličnih
- grafičko predstavljanje jedinki/populacija u obliku stabla

(1.1) Neponderirana metoda za sparivanje skupina na temelju prosječnih vrijednosti

(Unweighted pair-group method with arithmetic average; UPGMA)

(1.2) Metoda sparivanja susjeda

(Neighbour-joining method; NJ)

(2) Prikaz podataka u koordinatnom sustavu

- sažimanje i analiza povezanosti većeg broja varijabli
- omogućuju grafičko predstavljanje jedinki/populacija u reduciranim broju dimenzija tako da su njihove razlike i sličnosti naglašene na najbolji način

(2.1) Analiza glavnih koordinata

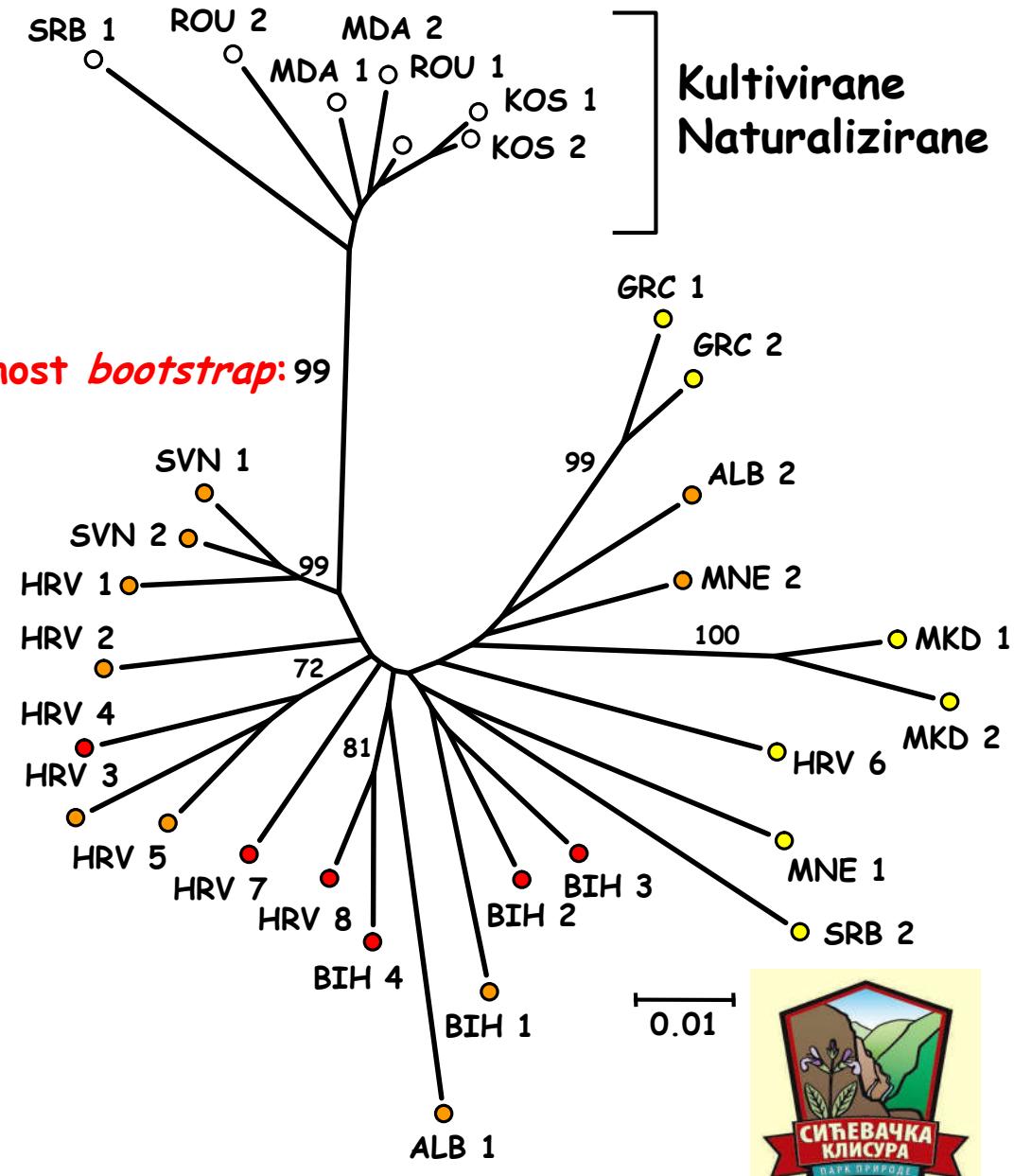
(Principal co-ordinate analysis; PCoA)

(2.2) Faktorijalna analiza korespondencije

(Factorial correspondence analysis; FCA)

LJEKOVITA KADULJA: STABLO

- matrica udaljenosti:
tetivna udaljenost
- algoritam za izradu stabla:
Metoda sparivanja susjeda
(**Neighbour-joining; NJ**)
- nezakorijenjeno stablo
- analiza pouzdanosti
grananja na stablu:
metoda *bootstrap*
- kultivirane/naturalizirane
populacije su se odvojile od
prirodnih populacija u
zasebnu skupinu uz visoku
vrijednost *bootstrap* (99%)



PLAN

- (1) Ljekovita kadulja (*Salvia officinalis* L.)
- (2) Analiza genetske raznolikosti
- (3) Analiza genetske udaljenosti
- (4) Analiza genetske strukture**
- (5) Modeliranje ekoloških niša
- (6) Analiza evolucijske povijesti
- (7) Buduća istraživanja

(4) ANALIZA GENETSKE STRUKTURE

Metoda:

Bayesovska analiza genetske strukture
(*Bayesian model-based clustering method*)

Software:

STRUCTURE (Pritchard et al., 2000)

Prepostavka:

- u analiziranom uzorku postoje jedinke koje pripadaju različitim skupinama sa svojstvenim alelnim učestalostima

Cilj:

- svrstati jedinke u skupine na takav način da je unutar svake skupine maksimalizirana
 - (A) Hardy-Weinbergova ravnoteža (**HWE**)
 - (B) ravnoteža vezanosti gena (**linkage equilibrium**)

Postupak:

- Bayesovska statistika (**Bayesian statistics**)
izračun posteriornih vjerojatnosti (= vjerodostojnosti)
- neparametrijska metoda 'Markovljev lanac Monte Carlo'
(**Markov chain Monte Carlo; MCMC**): poduzorkovanje na temelju modela

RAVNOTEŽA VEZANOSTI GENA

Učestalost
gameta
dvaju gena

gen
Aleli
 A
 a

		Aleli	učestalost		
		B	p_B	p_A	q_a
		b	q_b	$p_A p_B$	$q_a p_B$
				AB	aB
				$p_A q_b$	$q_a p_B$
				Ab	ab
				$p_A q_b$	$q_a q_b$

Ukoliko u određenoj populaciji postoji ravnoteža vezanosti gena A i B:

$$P_{AB} = P_A \times P_B$$

Učestalost gameta AB jednaka je umnošku alelnih učestalosti alela A i B

NERAVNOTEŽA VEZANOSTI GENA

- neravnoteža vezanosti gena (**linkage disequilibrium**: LD)
- pojava neravnomjernog udjela genotipskih klasa dvaju gena u populaciji
- uzroci:

1. geni su vezani

- geni se nalaze blizu jedan drugom na istom kromosomu tako da rijetko dolazi do rekombinacije
- temelj za izračunavanje učestalosti rekombinacija (**recombination frequency**; r) i izradu genetskih karata (**linkage map**; **genetic map**)

2. uzorkovana populacija je strukturirana

- populacija se sastoji od više podpopulacija tako da se sve uzorkovane jedinke ne pare nasumično
- učestalost alela u pojedinim podpopulacijama su različite od učestalosti alela cjelokupne uzorkovane populacije

VJEROJATNOST I VJERODOSTOJNOST

- vjerojatnost (**probability**) podataka
 - vjerojatnost pojave određenog rasporeda jedinki po određenom (zadanom) broju izvornih populacija na temelju multilokusnih genotipova
 - ovisi o hipotezi: broj K
- vjerodostojnost (**likelihood**) hipoteza
 - ukoliko postoji više hipoteza, za svaku pojedinačnu hipotezu možemo izračunati njezinu vjerodostojnost

$$L(H = h_K) = \Pr(D|h_K)$$

- $L(H=h_K)$ - vjerodostojnost hipoteze h_K (za određeni broj K)
- $\Pr(D|h_K)$ - uvjetna vjerojatnost (**conditional probability**) podataka (D: multilokusni genotipovi jedinki) uz danu hipotezu (h_K : hipoteza uz određeni broj K)

POSTUPAK

(1) Procjena broja izvornih populacija (K)

izvorne populacije (population of origin; ancestral population)
odnosno genski skupovi (gene pools)

(1.1) Izaberi različite hipoteze
(različite vrijednosti K)

(1.2) Izračunaj vjerodostojnost svake hipoteze
($H_{K=1}$; $H_{K=2}$; $H_{K=3}, \dots$)

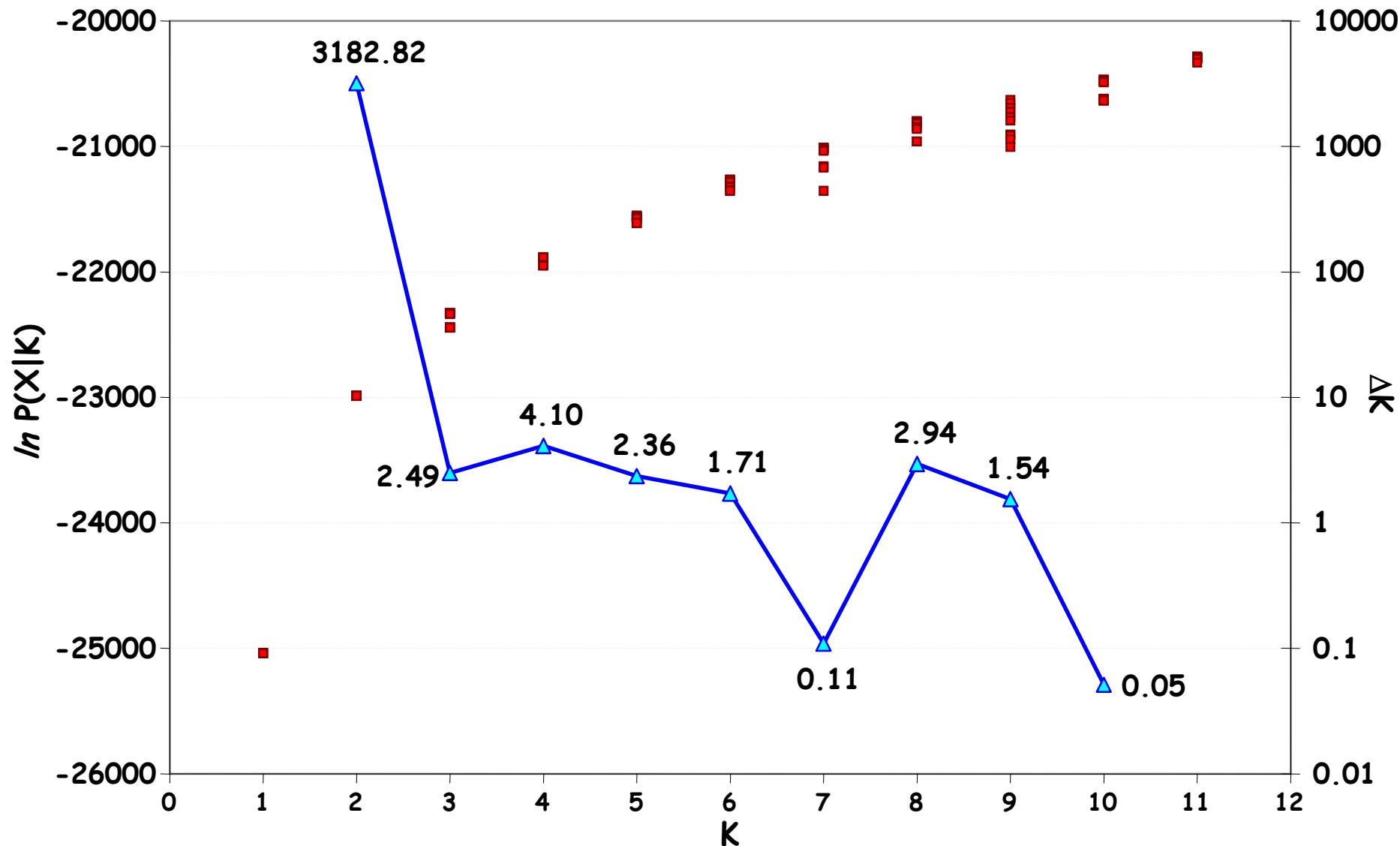
(1.3) Izaberi najvjerojatniju hipotezu

(2) Procjena pripadnosti jedinki izvronim populacijama

(2.1) Pridruži jedinke izvornim populacijama na temelju
najvjerojatnije hipoteze

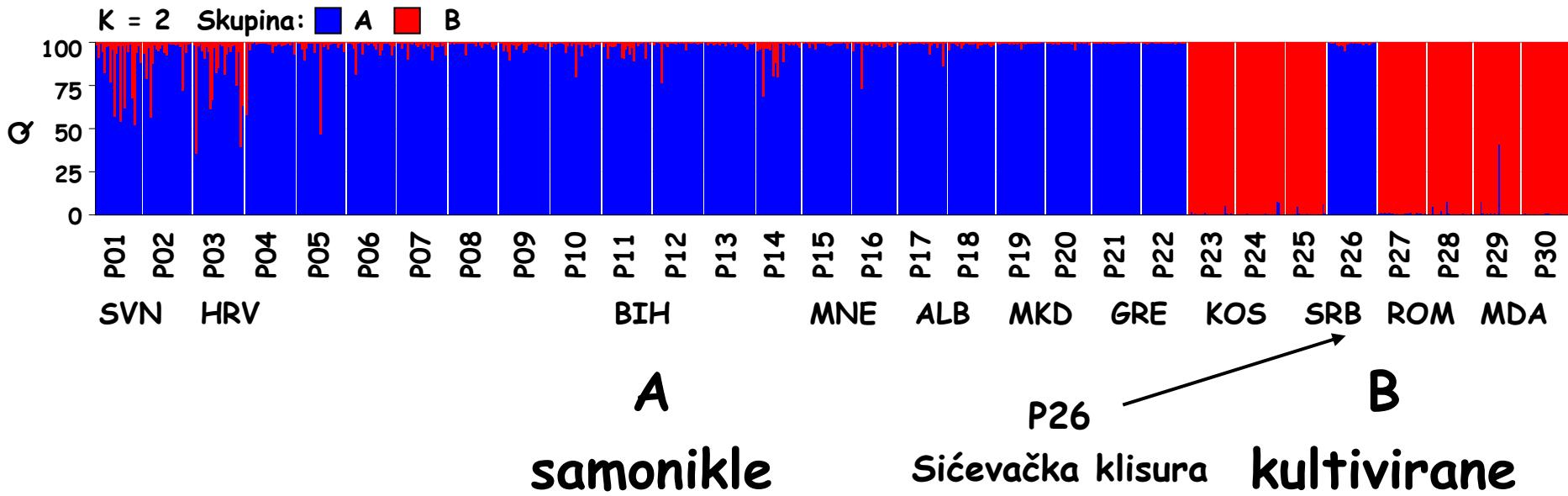
(2.2) Izračunaj % genoma svake jedinke koji potječe iz
određenog genskog skupa

REZULTAT: $\ln P(X|K)$ i ΔK



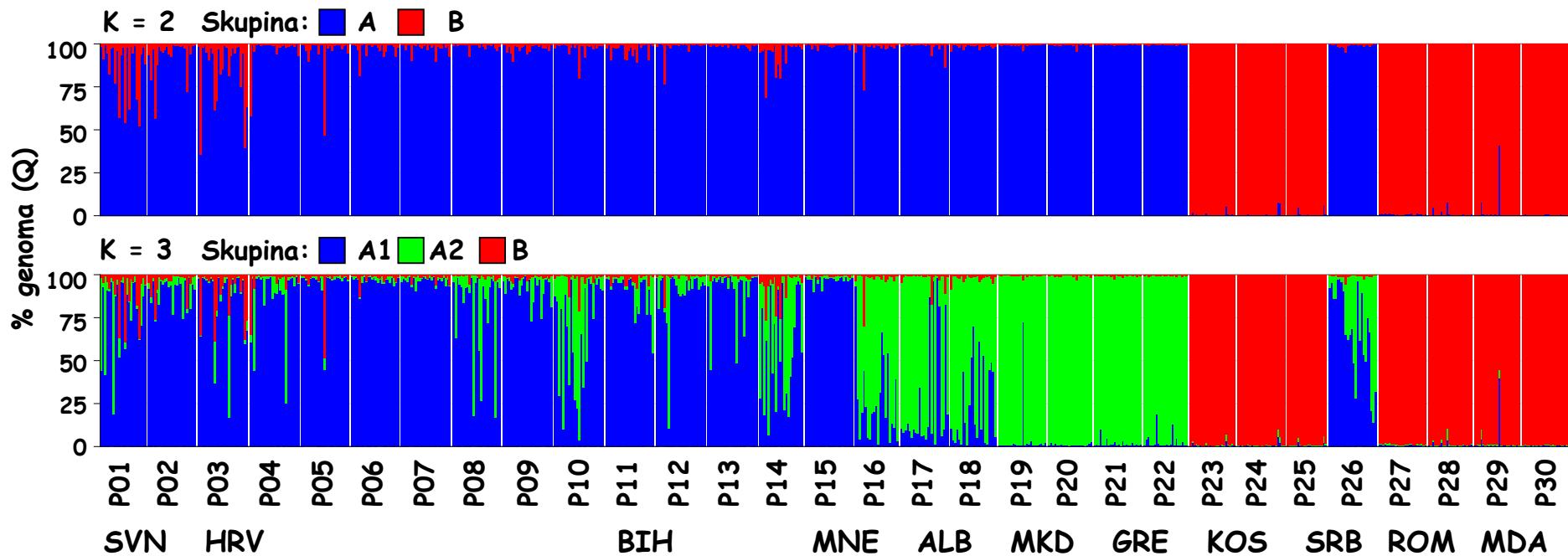
- $\ln P(X|K)$: vjerodostojnost (posteriorna vjerojatnost; 10 prohoda za svaki K)
- ▲ ΔK : stopa promjene vjerodostojnosti između uzastopnih vrijednosti K (Evanno et al., 2005)

LJEKOVITA KADULJA PRI $K = 2$



- izračun % genoma (Q) svake jedinke koji potječe iz određenog genskog skupa
- jedinke mogu potpuno pripadati jednoj skupini ili pak imati mješovito ('hibridno') podrijetlo iz dvije ili više skupina
- svaka je jedinka predstavljena stupcem, a boja odgovara % genoma (Q) jedinke koji potječe iz određenog genskog skupa

LJEKOVITA KADULJA PRI $K = 3$



A1

sjeverozapadne
populacije

A2

jugoistočne
populacije

B

kultivirane

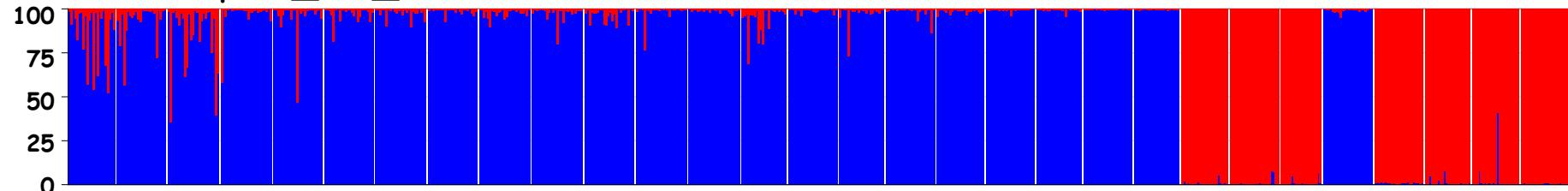
- skupina A (divlje populacije) se pri $K = 3$ odvaja na dvije na temelju zemljopisnih lokacija mjesto prikupljanja

A1 sjeverozapadne populacije: SVN-HRV-BIH-MNE-SRB

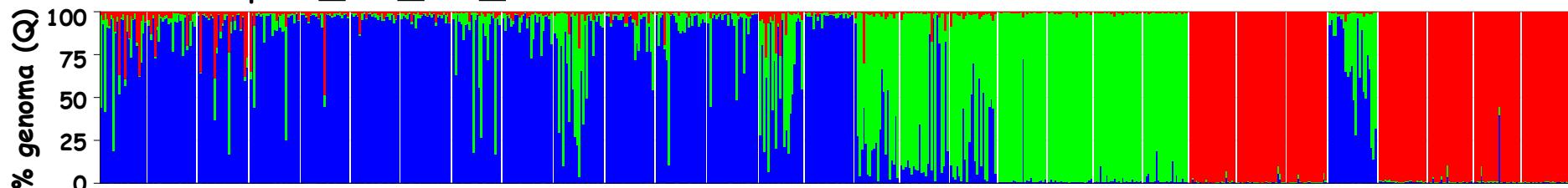
A2 jugoistočne populacije: MNE-ALB-MKD-GRE

LJEKOVITA KADULJA PRI $K = 4$

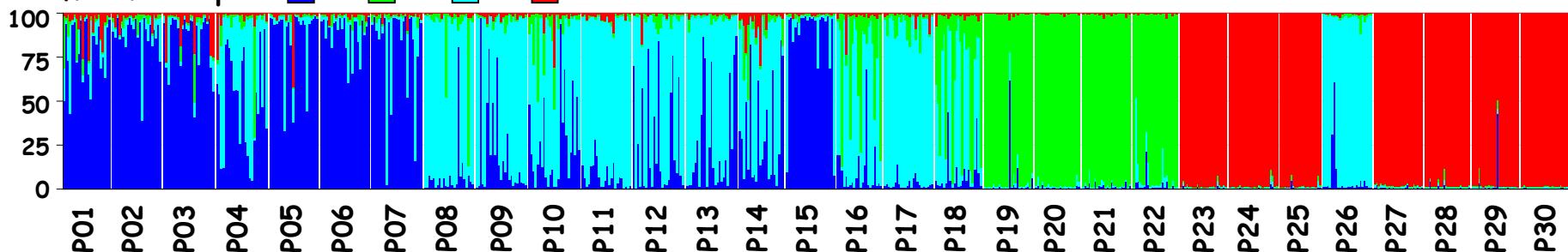
$K = 2$ Skupina: █ A █ B



$K = 3$ Skupina: █ A1 █ A2 █ B



$K = 4$ Skupina: █ A1 █ A2 █ A3 █ B



A1

Sjeverni
Jadran

A2

Južni
Jadran

A3

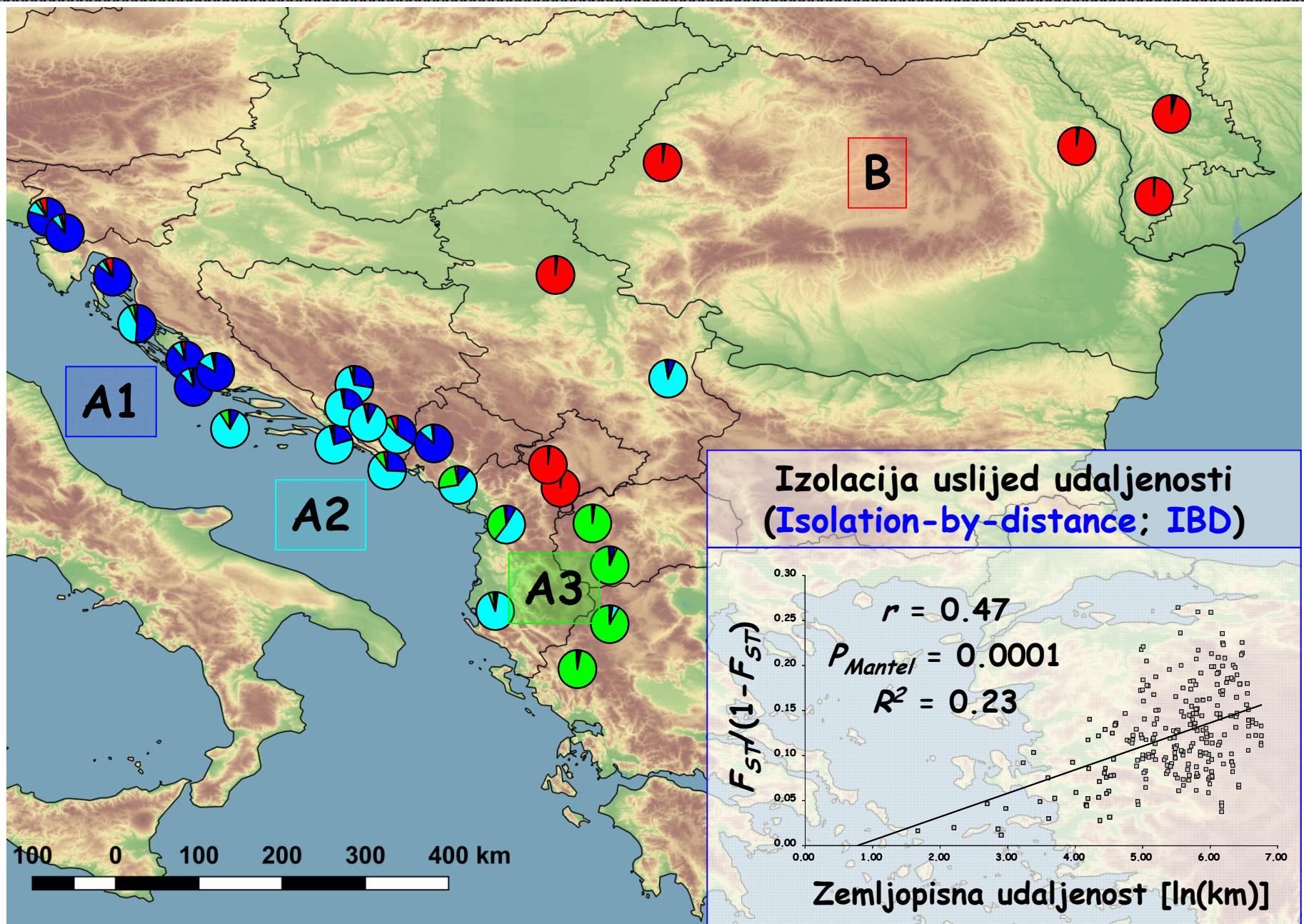
Makedonija
Grčka

B

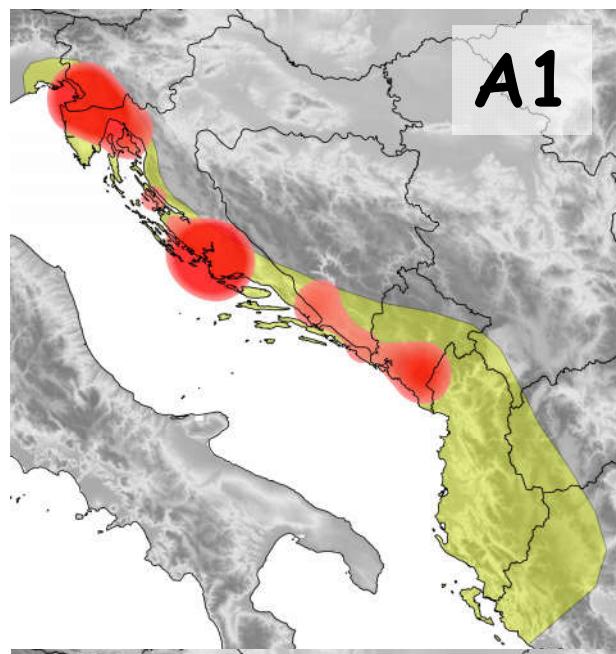
kultivirane
populacije

- divlje populacije se razvrstavaju u tri zemljopisne skupine

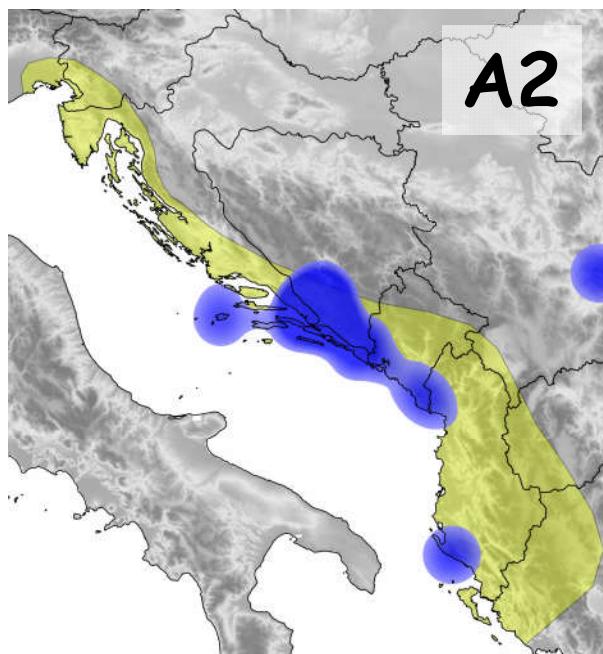
LJEKOVITA KADULJA: PROSJEČNI Q PRI K = 4



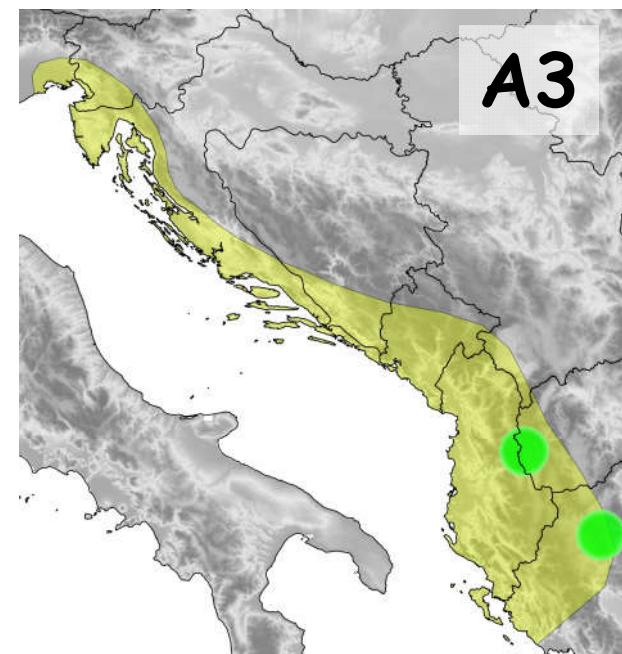
AELNO BOGATSTVO



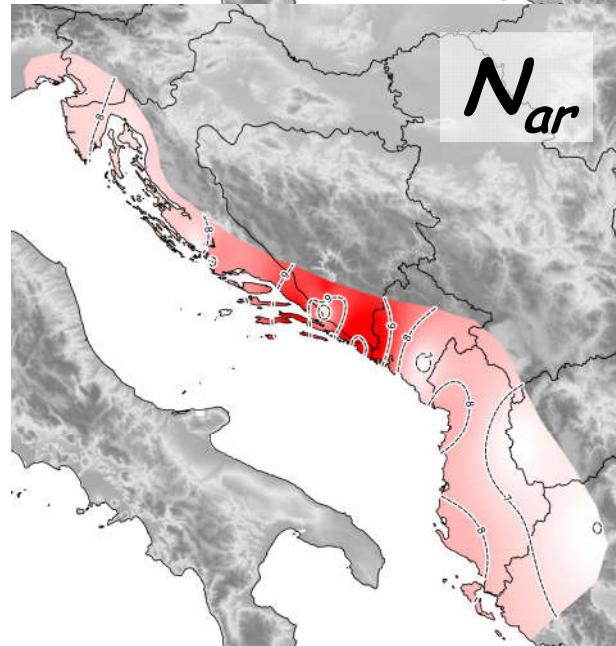
A1



A2



A3



N_{ar}

Izvorna populacija

Broj
uzorkovanih
populacija

N_{ar}

A1 Sjeverni Jadran

8

7.97

A2 Južni Jadran

11

8.49

A3 Makedonija/Grčka

4

6.25

$P(N_{ar})$

< 0.001

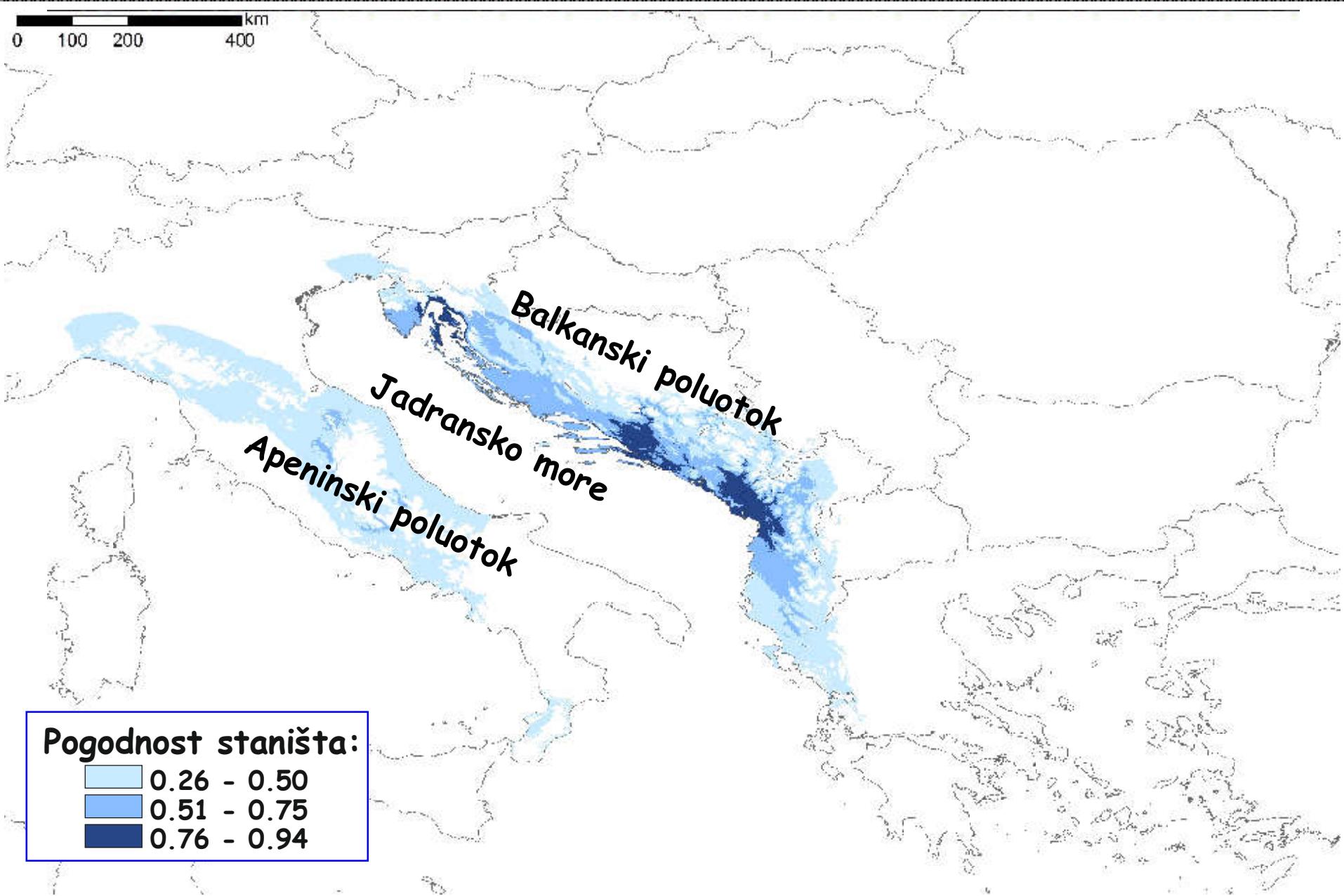
PLAN

- (1) Ljekovita kadulja (*Salvia officinalis* L.)
- (2) Analiza genetske raznolikosti
- (3) Analiza genetske udaljenosti
- (4) Analiza genetske strukture
- (5) Modeliranje ekoloških niša**
- (6) Analiza evolucijske povijesti
- (7) Buduća istraživanja

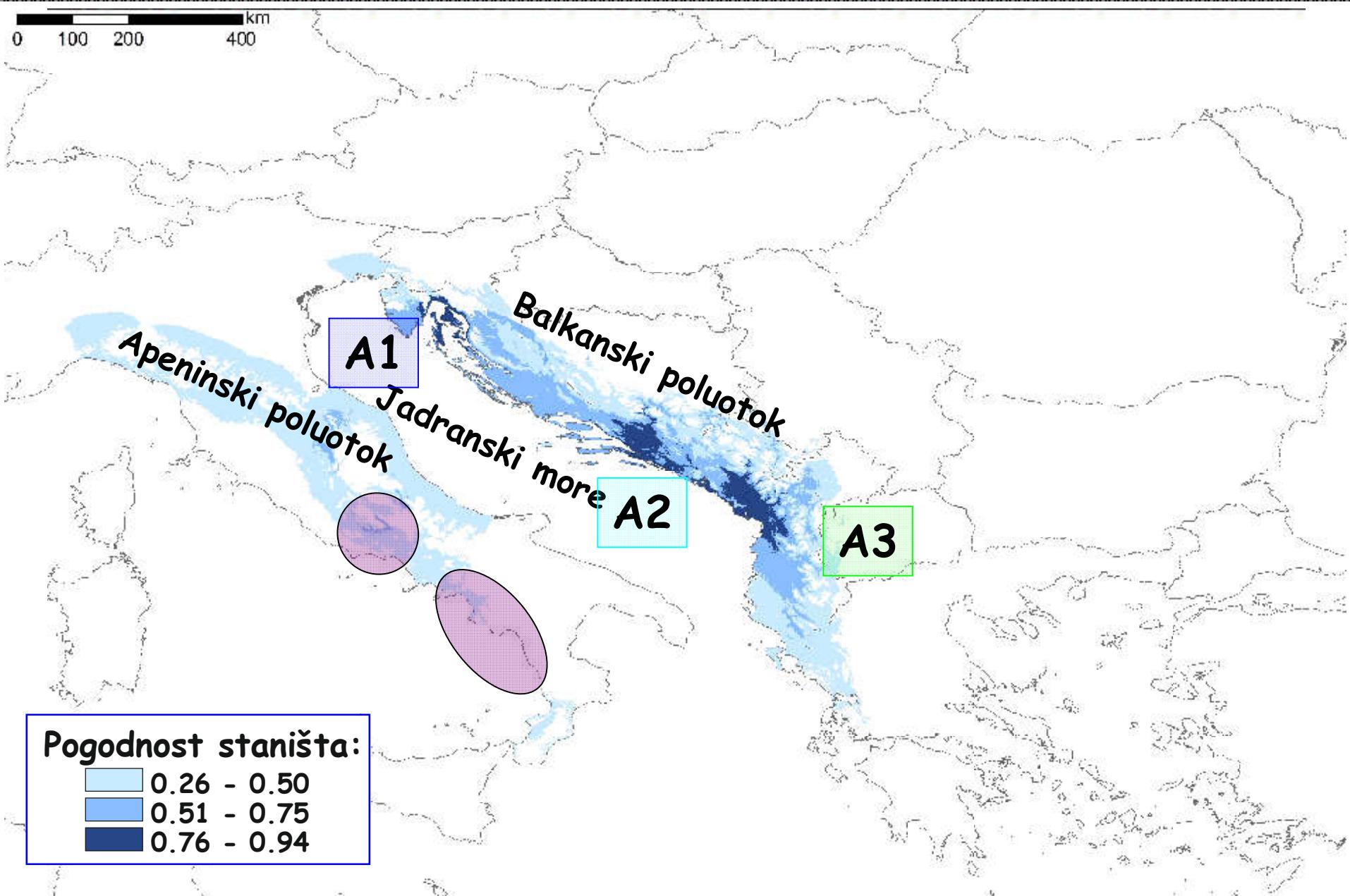
(5) MODELIRANJE EKOLOŠKIH NIŠA

- Modeliranje ekoloških niša
(Ecological Niche Modelling; ENM)
- Modeliranje rasprostranjenosti vrsta
(Species Distribution Modelling; SDM)
 - procijeniti odnos između rasprostranjenosti vrste po određenim nalazištima i okolišnih uvjeta tih nalazišta
 - procijeniti pogodnost staništa za vrstu
 - izraditi model rasprostranjenosti vrste
(= utvrditi povoljno stanište za vrstu)
- **ulazni podaci:**
 - (1) zemljopisna rasprostranjenost
68 podataka o rasprostranjenosti
 - (2) okolišni uvjeti nalazišta
19 bioklimatskih parametara (WorldClim database)
 - 11 temperaturnih- i 8 oborinskih parametara
 - predstavljaju godišnja kretanja, sezonske promjene, te temperaturne i oborinske ekstreme

ENM: SADAŠNJI KLIMATSKI UVJETI (i)



ENM: SADAŠNJI KLIMATSKI UVJETI (ii)



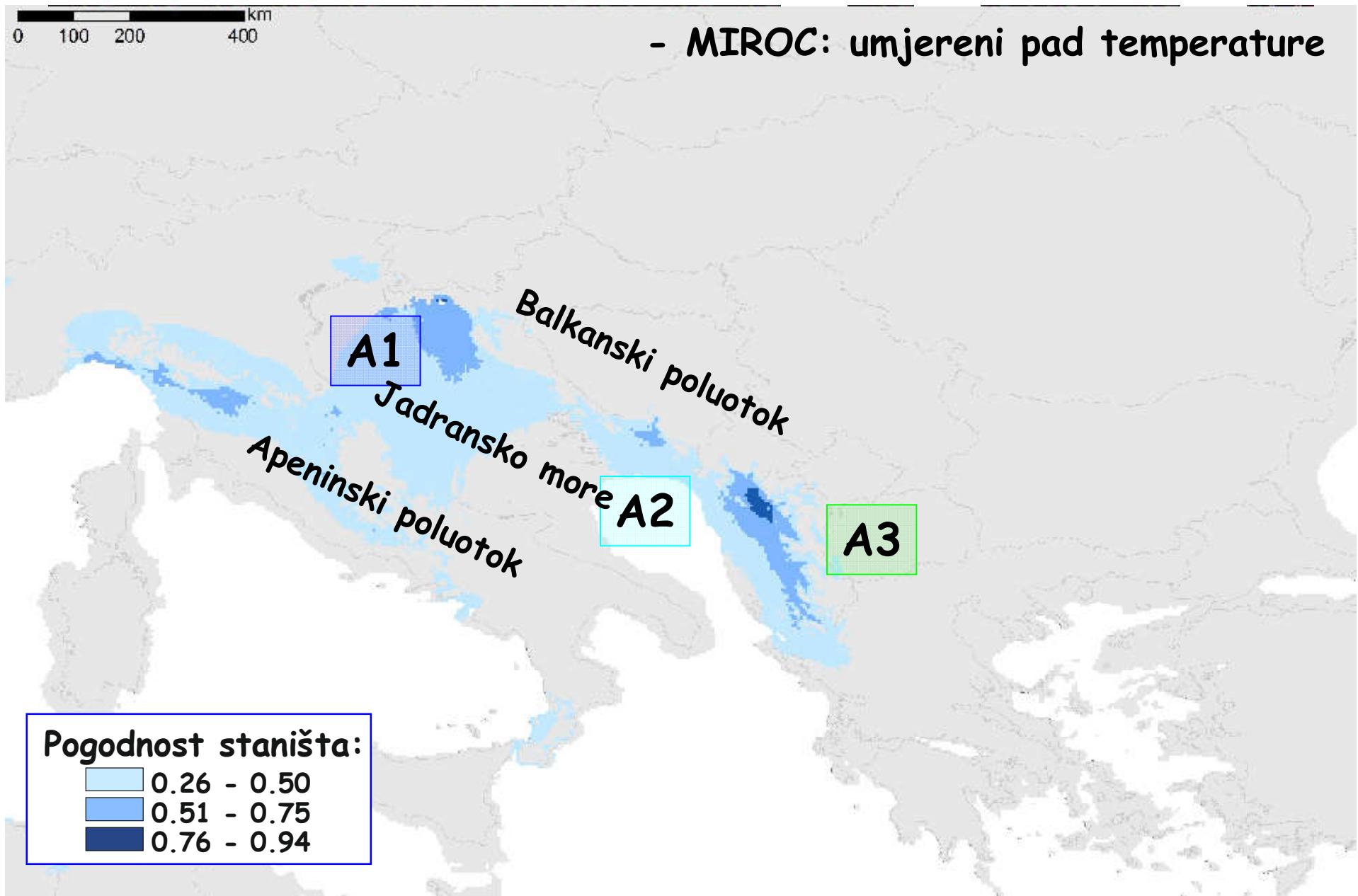
POVIJESNO RASPROSTRANJENJE

- modeliranje rasprostranjenosti vrste u vrijeme posljednjeg glacijalnog maksimuma
(Last Glacial Maximum; LGM; ~21,000 g.pr.n.e.)
- otkrivanje mogućih glacijalnih refugija vrste
- glacijalni refugij (**glacial refugium**) - regija u kojoj je određena vrsta preživjela ledeno doba
- ulazni podaci:
 - (1) model za sadašnje stanja
 - (2) klimatski uvjeti u prošlosti
 - 19 bioklimatskih varijabli
 - bioklimatski podaci za LGM razvijeni u okviru projekta Paleoclimate Modelling Intercomparison Project Phase II
- dva modela:
 - CCSM (Community Climate System Model; USA)
 - MIROC (Model for Interdisciplinary Research on Climate; Japan)

ENM: POSLJEDNJI GLACIJALNI MAKSIMUM (CCSM)



ENM: POSLJEDNJI GLACIJALNI MAKSIMUM (MIROC)



PLAN

- (1) Ljekovita kadulja (*Salvia officinalis* L.)
- (2) Analiza genetske raznolikosti
- (3) Analiza genetske udaljenosti
- (4) Analiza genetske strukture
- (5) Modeliranje ekoloških niša
- (6) Analiza evolucijske povijesti**
- (7) Buduća istraživanja

(6) ANALIZA EVOLUCIJSKE POVIJESTI

- Apsoksimativna Bayesovske analiza
(Approximate Bayesian Computation; ABC)

Ulagni podaci:

- molekularni podaci
- scenariji koji opisuju demografsku povijest

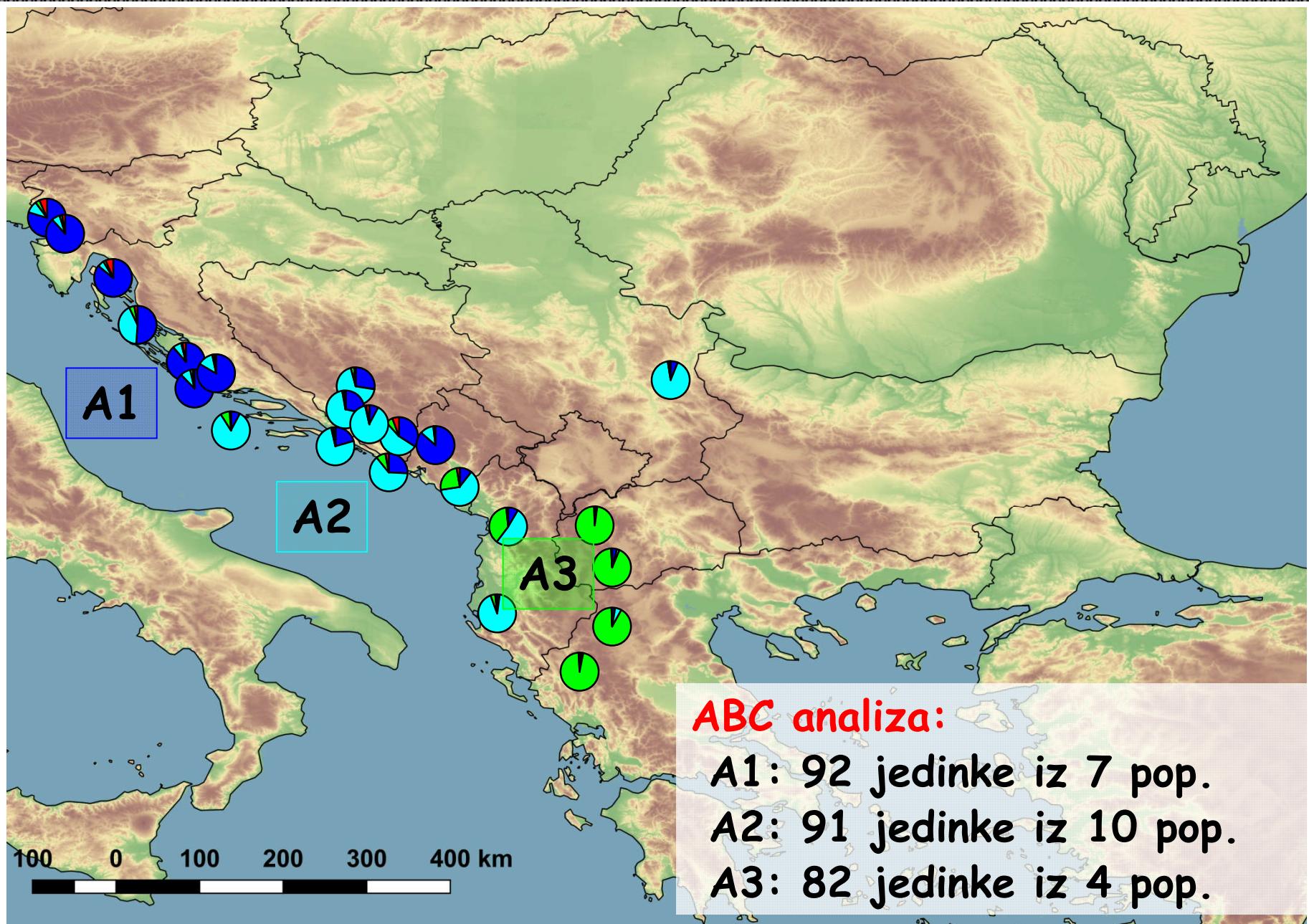
Postupak:

- simulirati veliki broj setova podataka za svaki scenarij
- usporediti simulirane setove s stvarnim podacima
- simulirani set najsličniji stvarnim podacima je najvjerojatniji

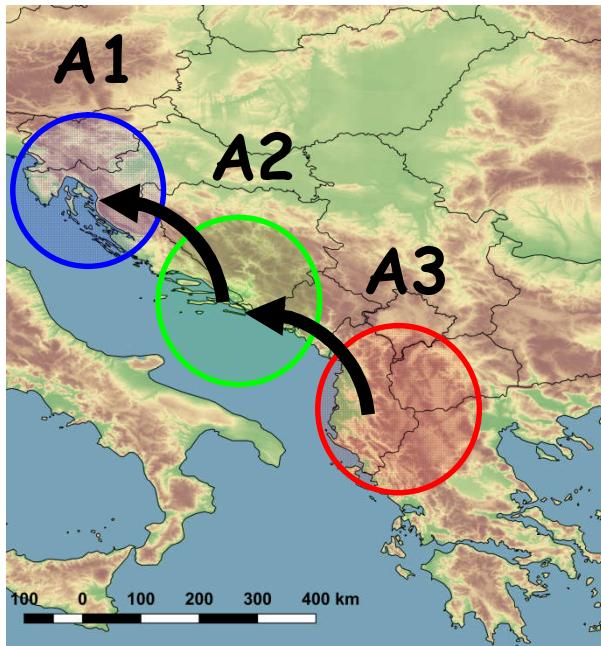
Cilj:

- usporediti alternativne scenarije na temelju posteriorne vjerojatnosti svakog scenarija
- procijeniti parametre za jedan ili više scenarija:
efektivna veličina populacije; vrijeme razdvajanja

IZVORNE POPULACIJE

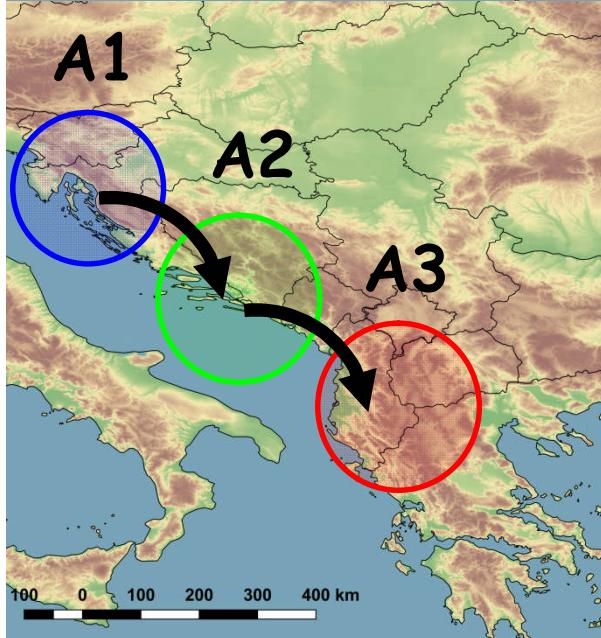
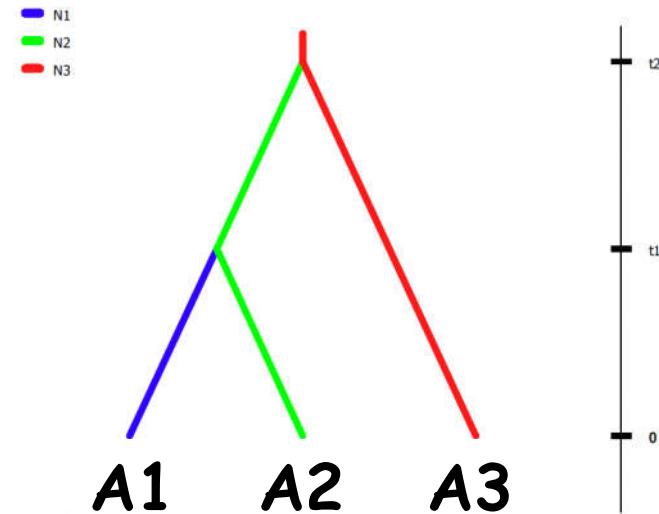


PET JEDNOSTAVNIH POVIJESNIH SCENARIJA (i)



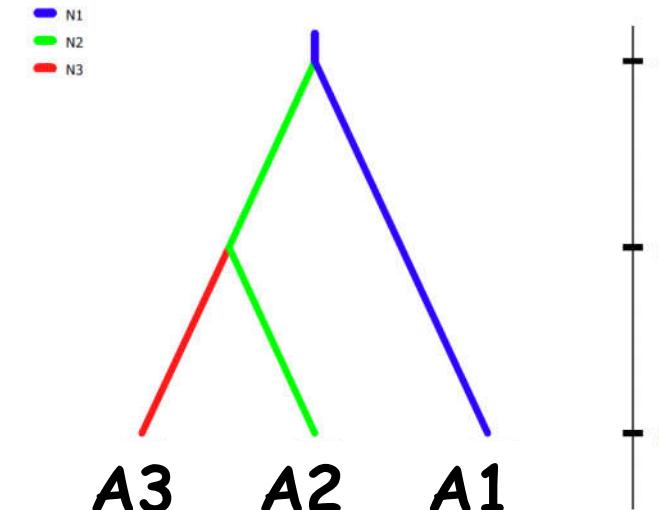
Scenarij 1

Populacija A1 je nastala iz populacije A2, koja je nastala iz populacije A3

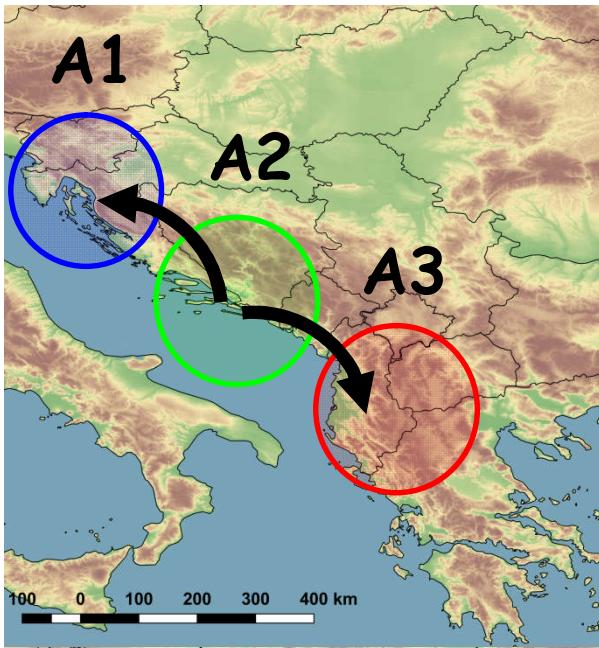


Scenarij 2

Populacija A3 je nastala iz populacije A2, koja je nastala iz populacije A1

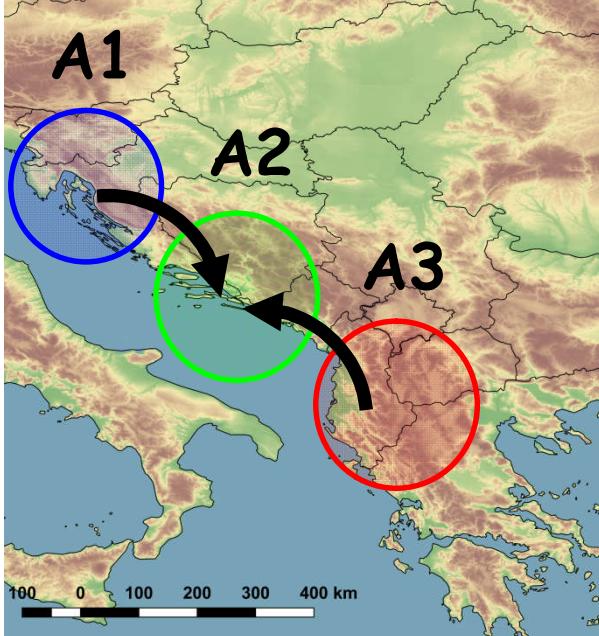
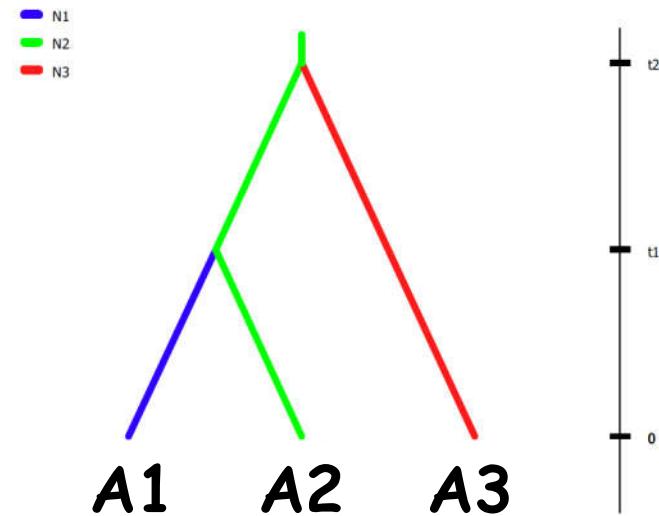


PET JEDNOSTAVNIH POVIJESNIH SCENARIJA (ii)



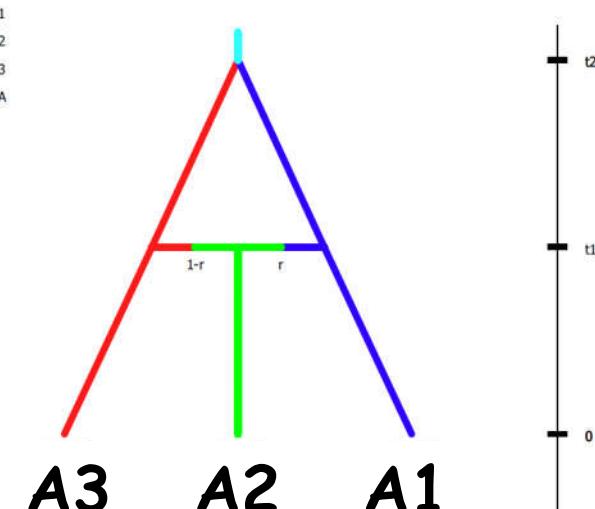
Scenarij 3

Populacije A1 i A3
su nastale iz
populacije A2

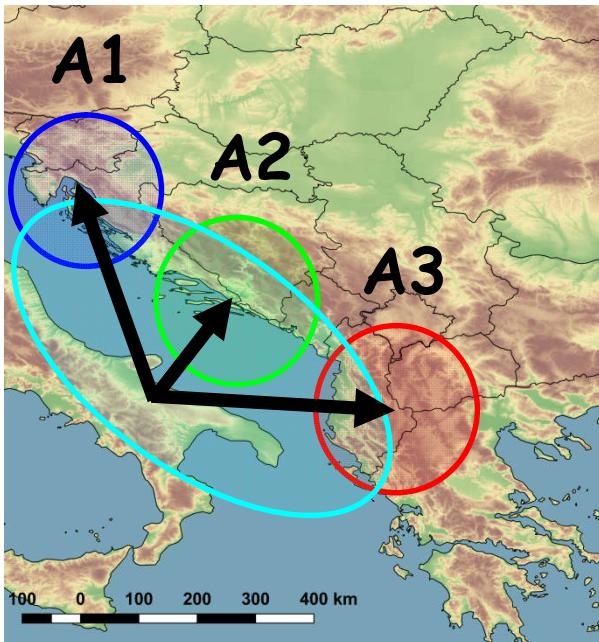


Scenarij 4

Populacija A2 je
nastala miješanjem
populacija
A1 i A3

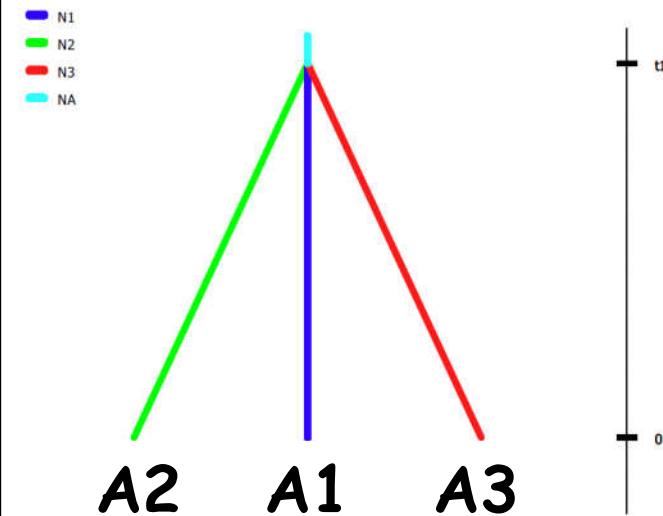


PET JEDNOSTAVNIH POVIJESNIH SCENARIJA (iii)



Scenarij 5

Sve su se
tri populacije
razdvojile u isto
vrijeme



Posteriorna vjerojatnost
(PP)

Scenarij

REZULTATI:

1	0.166
2	0.193
3	0.075
4	0.126
5	0.440

POVIJESNA REKONSTRUKCIJA: DAVNA PROŠLOST



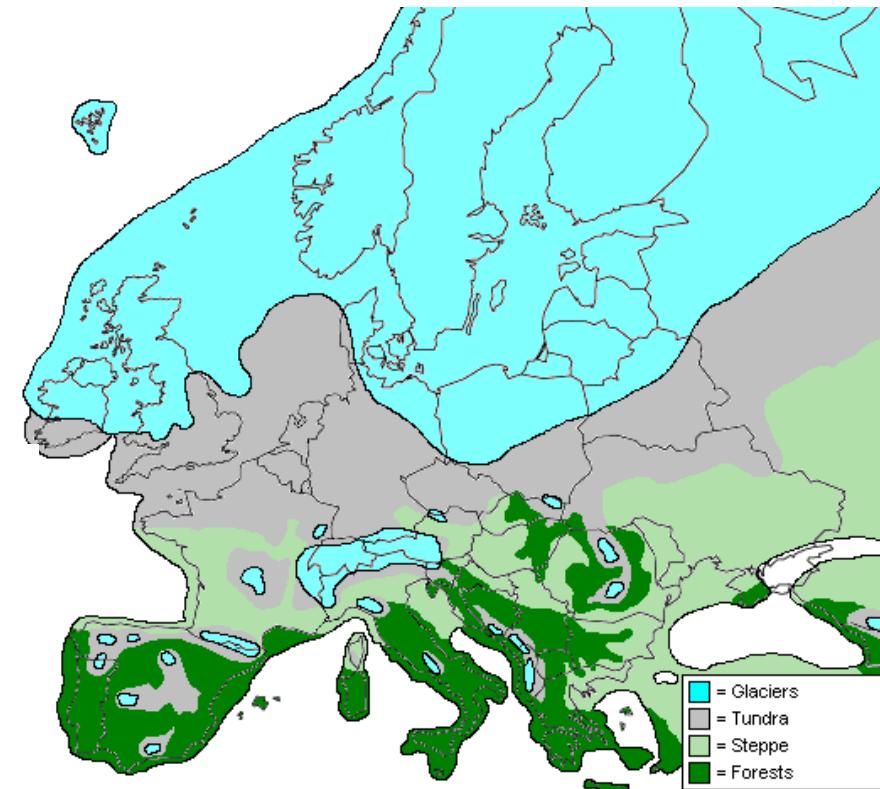
Rasprostranjenost ljekovite kadulje prije posljednjeg ledenog doba (?)

ABC Efektivna veličina populacije:

Izvorna pop. 8,550 jedinki

A1	4,330 jedinki
A2	7,190 jedinki
A3	2,440 jedinki

Vegetacijska karta Europe u posljednjem ledenom dobu

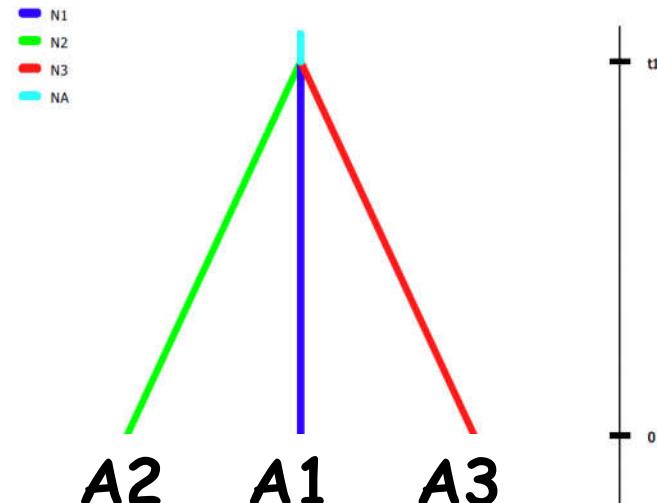
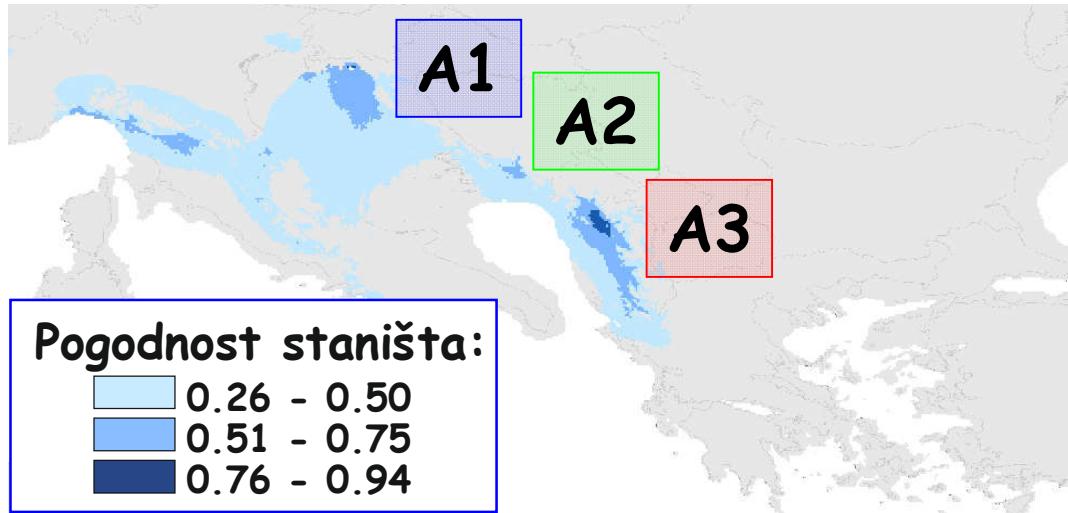


Glacijalne refugije:

1. Iberijski poluotok
2. Apeninski poluotok
3. Balkanski poluotok

POVIJESNA REKONSTRUKCIJA: RAZDVAJANJE

Rasprostranjenost ljekovite kadulje u vrijeme posljednjeg glacijalnog maksimuma ($\sim 21,000$ g.pr.n.e.)

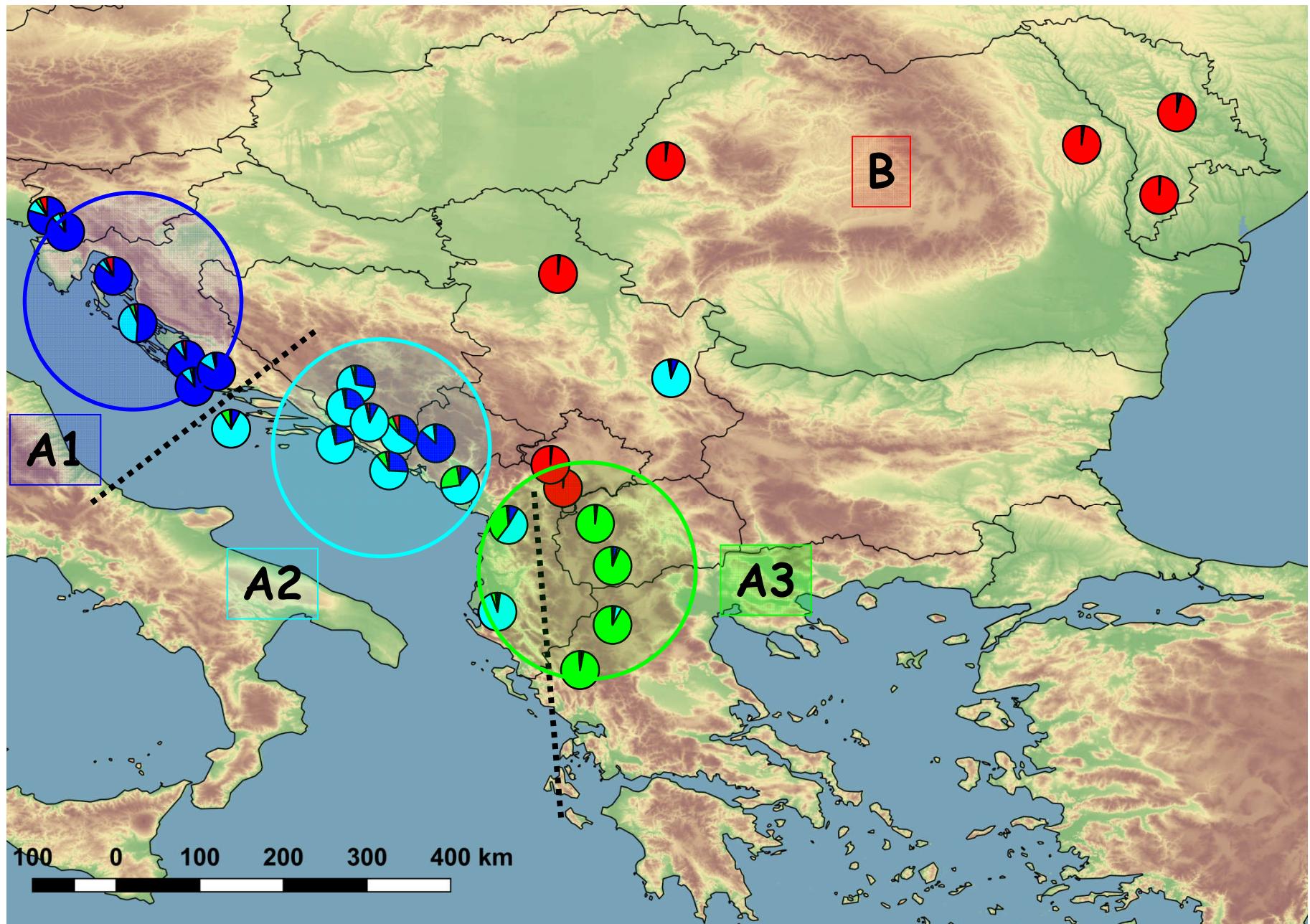


ABC Vrijeme razdvajanja:
prije 572 generacije
(95% CI: 157-1,540)

Generacijsko vrijeme
(*generation time*)
 ~ 30 godina (?)

Glacijalni refugiji u Sredozemlju (Médail i Diadema, 2009)

POVIJESNA REKONSTRUKCIJA: SADAŠNJOST



PLAN

- (1) Ljekovita kadulja (*Salvia officinalis* L.)
- (2) Analiza genetske raznolikosti
- (3) Analiza genetske udaljenosti
- (4) Analiza genetske strukture
- (5) Modeliranje ekoloških niša
- (6) Analiza evolucijske povijesti
- (7) Buduća istraživanja**

(7) BUDUĆA ISTRAŽIVANJA

(1) Ljekovita kadulja (*S. officinalis* L.)

62 populacije (13 Italija) / 1350 jedinki
8 SSR / 2 kloroplastne regije

(2) Španjolska kadulja (*S. lavandulifolia* Vahl)

25 populacija / 533 jedinke
7 SSR / 2 kloroplastne regije

(3) Uškasta kadulja (*S. x auriculata* Mill.)

prirodni križanac između ljekovite i grčke kadulje
(*S. fruticosa* Mill.) na Visu

75 jedinki

7 SSR / AFLP / MSAP / morfološka svojstva

(4) Kratkozupčasta kadulja (*S. brachyodon* Vandas)

endemična vrsta: 3 populacije / 332 jedinke
vegetativno razmnažanje

8 SSR / MSAP

ANALIZA MOLEKULARNE RAZNOLIKOSTI

- (1) Ljekovita kadulja (*Salvia officinalis* L.)
- (2) Analiza genetske raznolikosti
- (3) Analiza genetske udaljenosti
- (4) Analiza genetske strukture
- (5) Modeliranje ekoloških niša
- (6) Analiza evolucijske povijesti
- (7) Buduća istraživanja

GOTOVO!



Mend
The
Gap

Smart Integration of Genetics with Sciences of the Past in Croatia: Minding and Mending the Gap



ANALIZA MOLEKULARNE RAZNOLIKOSTI

Prof. dr. sc. Zlatko Šatović

Sveučilište u Zagrebu, Agronomski fakultet

Znanstveni centar izvrsnosti

za bioraznolikost i molekularno oplemenjivanje bilja (CroP-BioDiv)

E-mail: zsatovic@agr.hr