Genetic Diversity and Evolutionary History of *Salvia officinalis* L.

Zlatko Šatović

University of Zagreb, Faculty of Agriculture, Zagreb, Croatia

Centre of Excellence for Biodiversity and Molecular Plant Breeding (CoE CroP-BioDiv), Zagreb, Croatia

E-mail: zsatovic@agr.hr

Europska unija
Zajedno do fondova EU

EUROPSKI STRUKTURNI I INVESTICIJSKI FONDOVI

Operativni program KONKURENTNOST I KOHEZIJA

Republika Hrvatska

CoE CroP-BioDiv

+ 5 Member Institutions

Host Institution

FAZ

University of Zagreb, Faculty of Agriculture

IPTPO

Institute for Agriculture and Tourism Poreč

PMF

University of Zagreb, Faculty of Science

PIO

Agricultural Institute Osijek

IJK

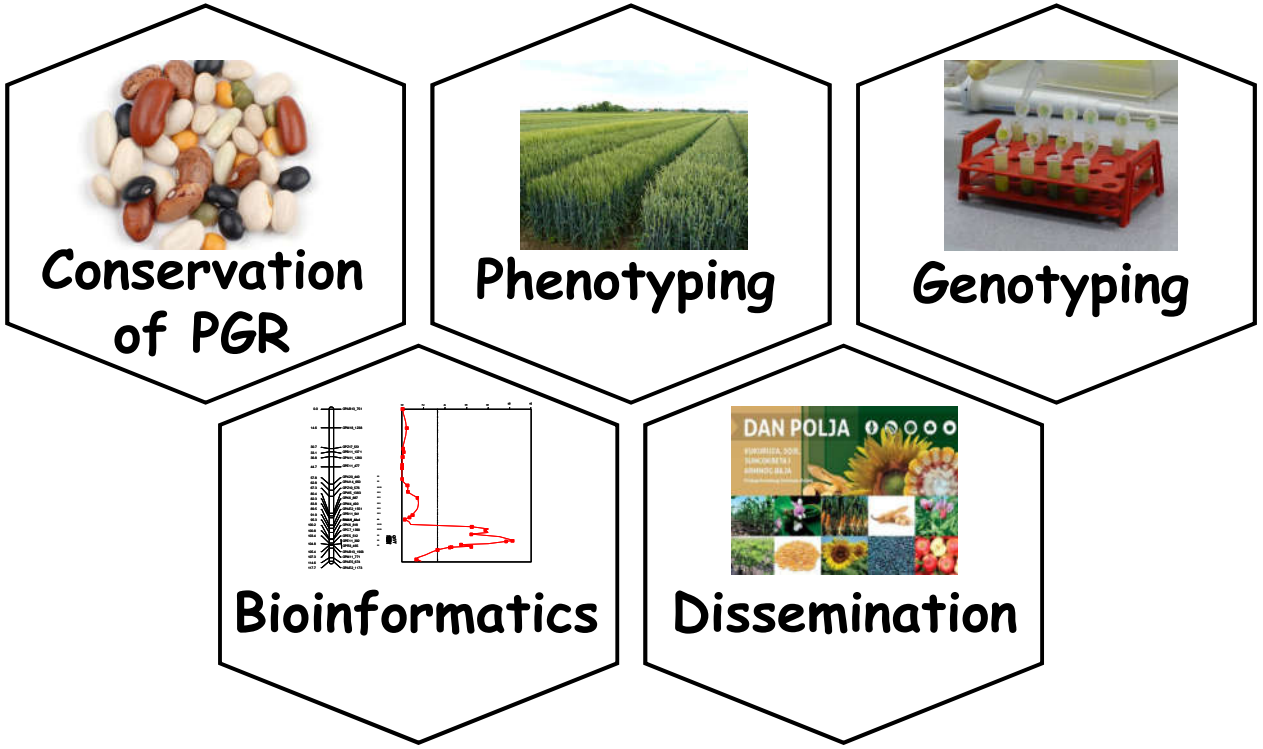
Institute for Adriatic Crops Split

PFOS

University of Osijek, Faculty of Agriculture

CoE CroP-BioDiv

OBJECTIVES



Conservation of Plant Genetic Resources >> increase the benefits

Phenotyping >> high-throughput phenotyping (HTP) platforms

Genotyping >> next-generation sequencing (NGS) techniques

Bioinformatics >> genome-wide association studies (GWAS)

Dissemination >> workshops, summer schools

MODEL PLANT SPECIES

Major Crops

Maize

Wheat

Soybean

Traditional Crops

Grape vine

Olives

Promising Crops

Brassicas
Alliums

Common bean

Dalmatian pyrethrum
Dalmatian sage

Centre of Excellence for Biodiversity and Molecular Plant Breeding (CoE CroP-BioDiv)

CroP-BioDiv

Operativni program KONKURENTNOST I KOHEZIJA

Europska unija
Zajedno do fondova EU

EUROPSKI STRUKTURNI I INVESTICIJSKI FONDOVI

PLAN

(1) Dalmatian sage (*Salvia officinalis* L.)

(2) Genetic diversity

(3) Genetic structure

(4) Ecological niche modelling

(5) Demographic history

PLAN

(1) Dalmatian sage (*Salvia officinalis* L.)

(2) Genetic diversity

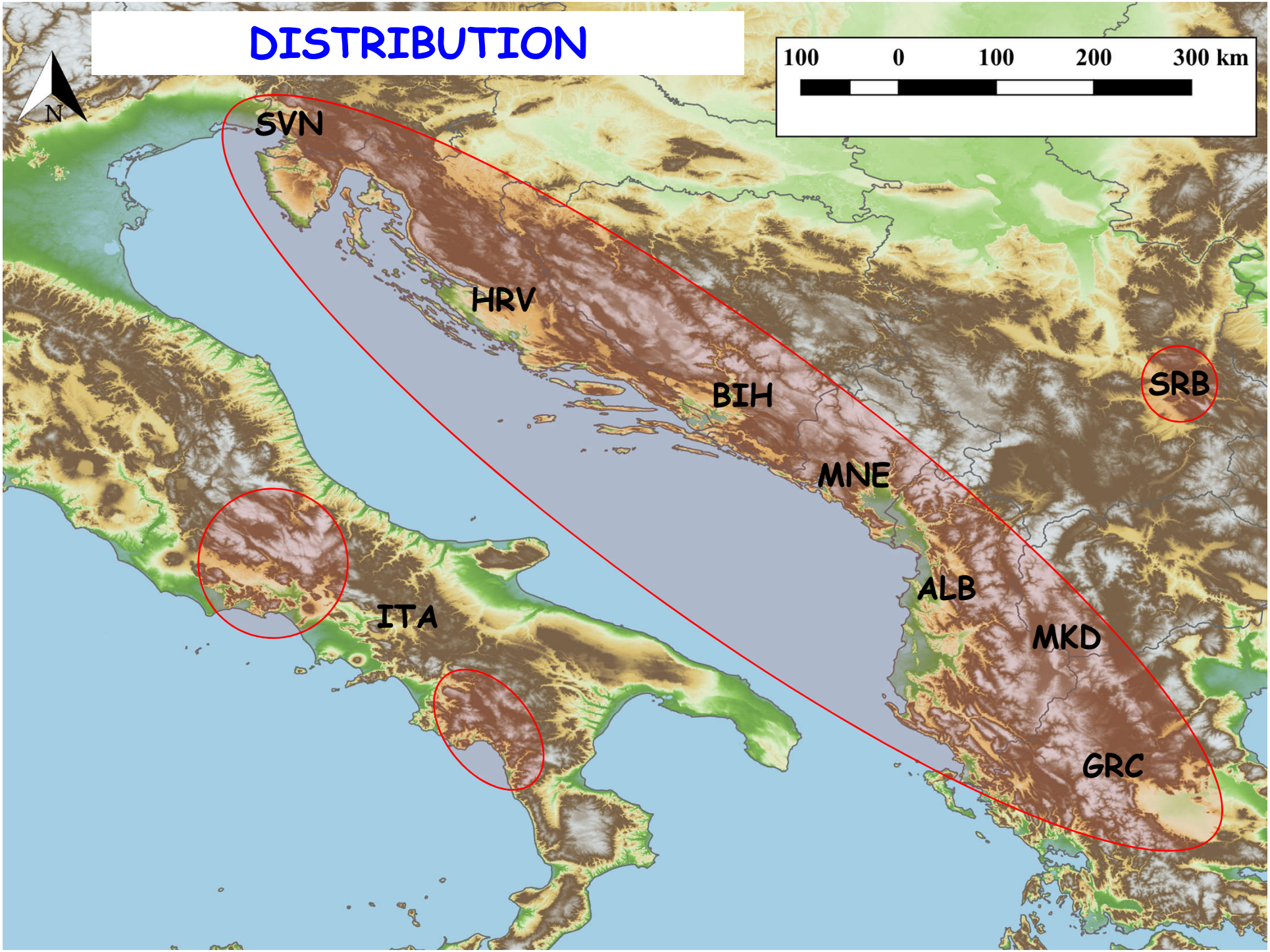
(3) Genetic structure

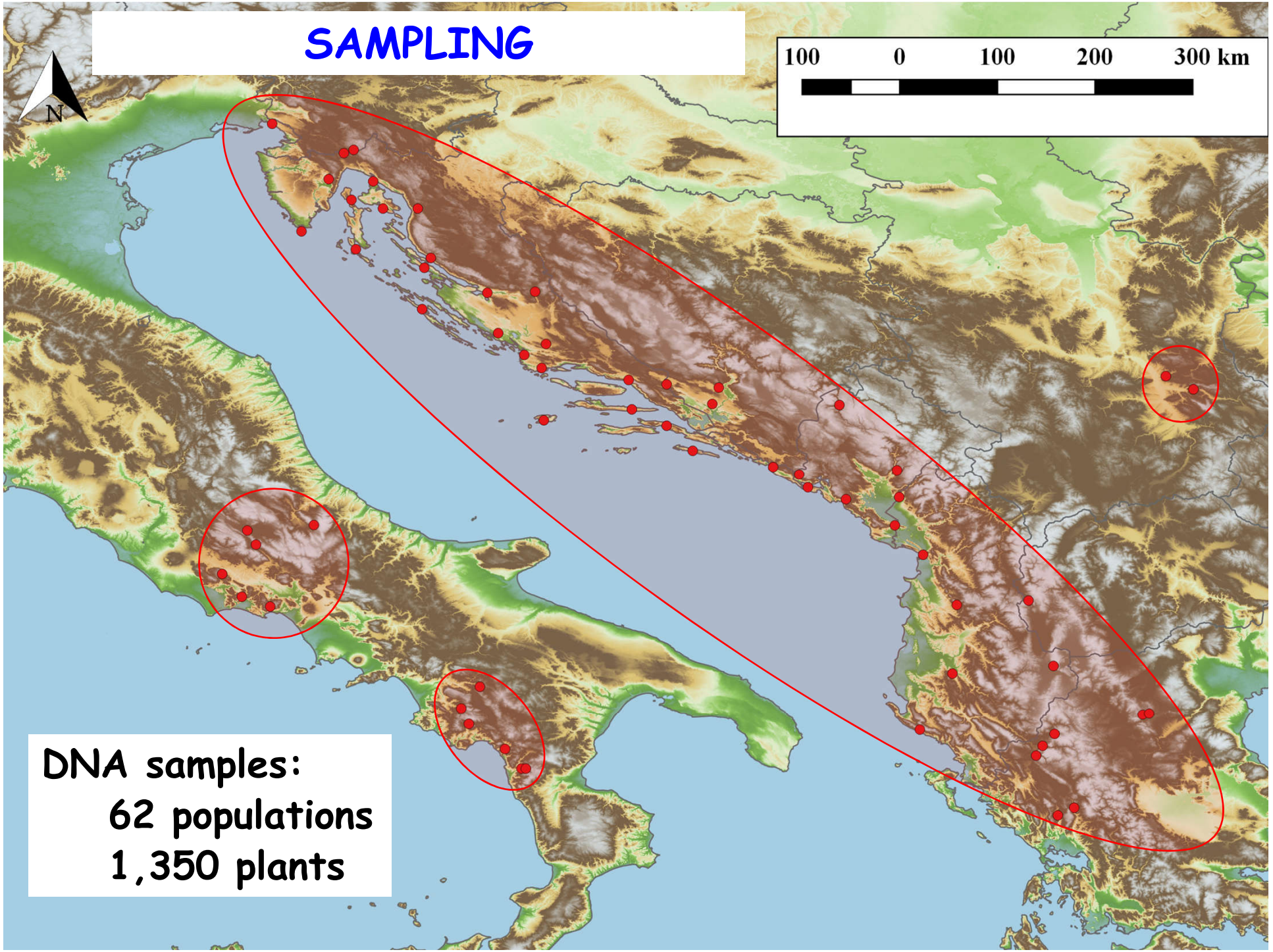
(4) Ecological niche modelling

(5) Demographic history

(1) DALMATIAN SAGE

- *Salvia officinalis* L.
- outcrossing, insect-pollinated
- perennial subshrub
- widely used since ancient times
  for medicinal, culinary and ornamental purposes
- natural distribution:
  coastal region of the western Balkan and central and southern Apennine Peninsulas
- cultivation:
  throughout the Mediterranean region, in Australia, Germany, USA etc.
- naturalized (feral) populations:
  plants that have escaped from earlier cultivation


DISTRIBUTION
N
100 0 100 200 300 km
SVN
HRV
BIH
SRB
MNE
ALB
MKD
ITA
GRC


SAMPLING
N
100 0 100 200 300 km
DNA samples:
62 populations
1,350 plants

SAMPLES / ANALYSES

- samples:
  - 62 populations / 1,350 plants
    - Italy (13), Slovenia (2), Croatia (23),
      Bosnia and Herzegovina (2), Montenegro (7), Serbia (2),
      Albania (5), Macedonia (1), Greece (7)

- molecular analysis:
  - 8 microsatellite markers
  - 191 alleles

PLAN

(1) Dalmatian sage (*Salvia officinalis* L.)

(2) Genetic diversity

(3) Genetic structure

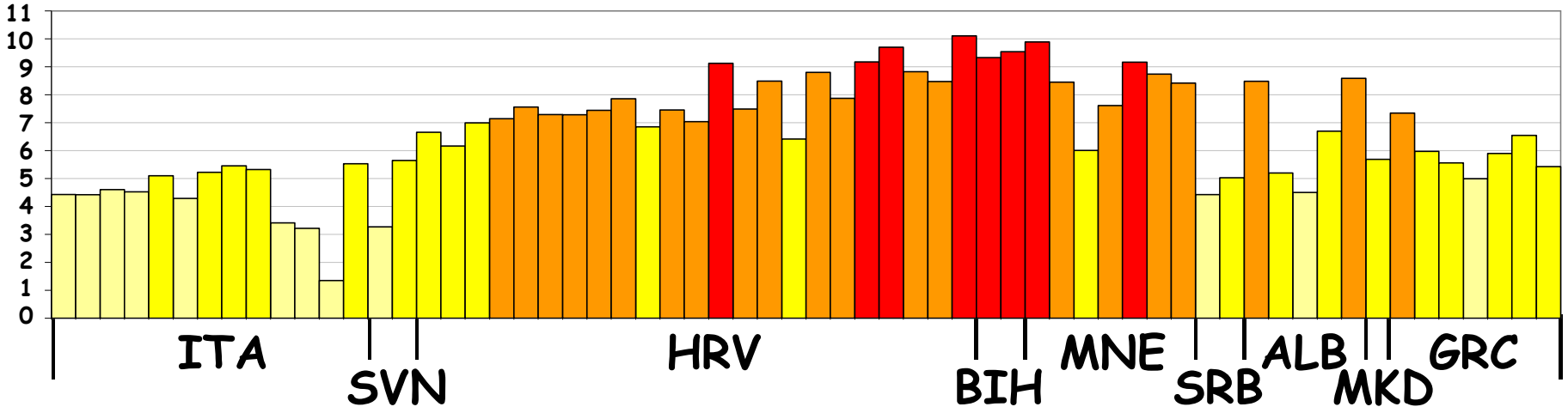
(4) Ecological niche modelling

(5) Demographic history

(2) GENETIC DIVERITY

Parameter	Mean	Range
Allelic richness (N_{ar})	6.669	1.35-10.11
Observed heterozygosity (H_O)	0.665	0.081-0.832
Expected heterozygosity (H_E)	0.683	0.087-0.846

Allelic richness (N_{ar})



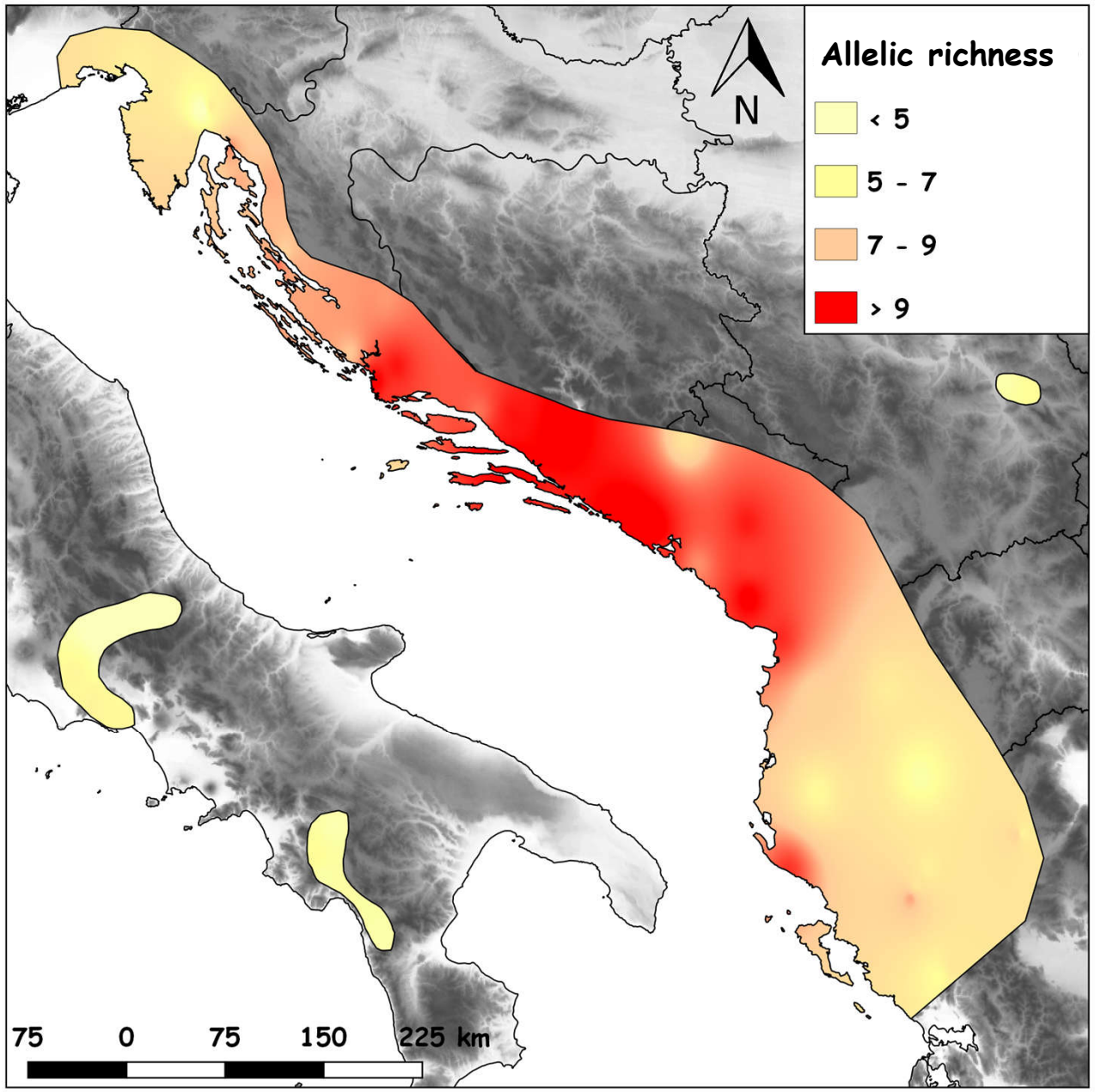
ALLELIC RICHNESS

N_{ar} > 9

Br.	Population	Country
P38	Konavle	HRV
P41	Vrbanj	MNE
P35	Pelješac	HRV
P40	Mostar	BIH
P39	Međugorje	BIH
P34	Hvar	HRV
P45	Rumija	MNE
P28	Šparadići	HRV

N_{ar} < 5

P59	Arta	GRC
P03	Maratea	ITA
P04	Cilento	ITA
P52	Prespa	ALB
P01	Pollino 2	ITA
P48	Miljkovac	SRB
P02	Pollino 1	ITA
P06	Savoia	ITA
P10	d'Antino	ITA
P14	Novokračine	SVN
P11	M. Salviano	ITA
P12	Maiella	ITA



PLAN

(1) Dalmatian sage (*Salvia officinalis* L.)

(2) Genetic diversity

(3) Genetic structure

(4) Ecological niche modelling

(5) Demographic history

(3) GENETIC STRUCTURE

- Bayesian model-based clustering method for inferring population structure

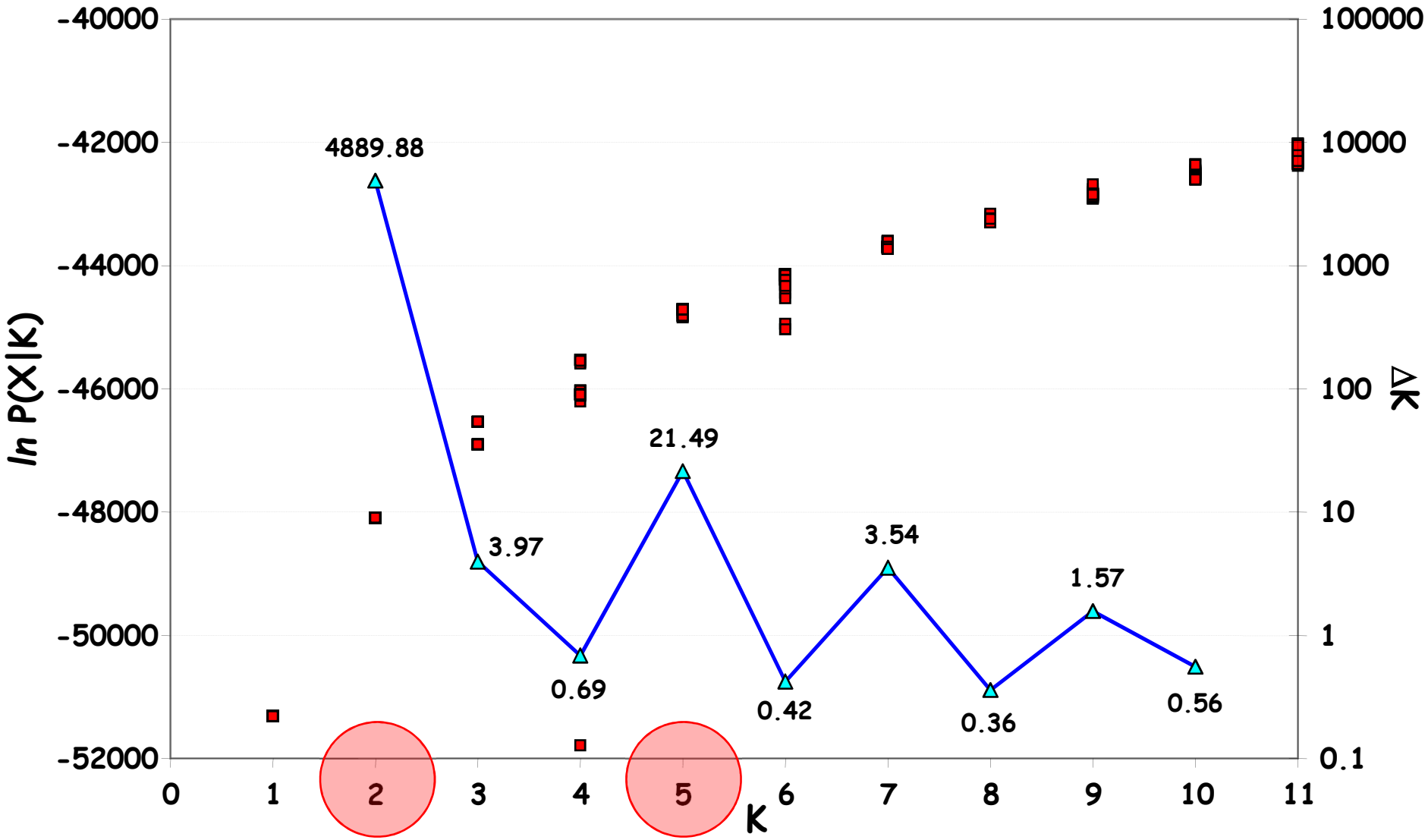
Assumption:

- there are K populations of origin each of which is characterized by a set of allele frequencies at each locus

Goal:

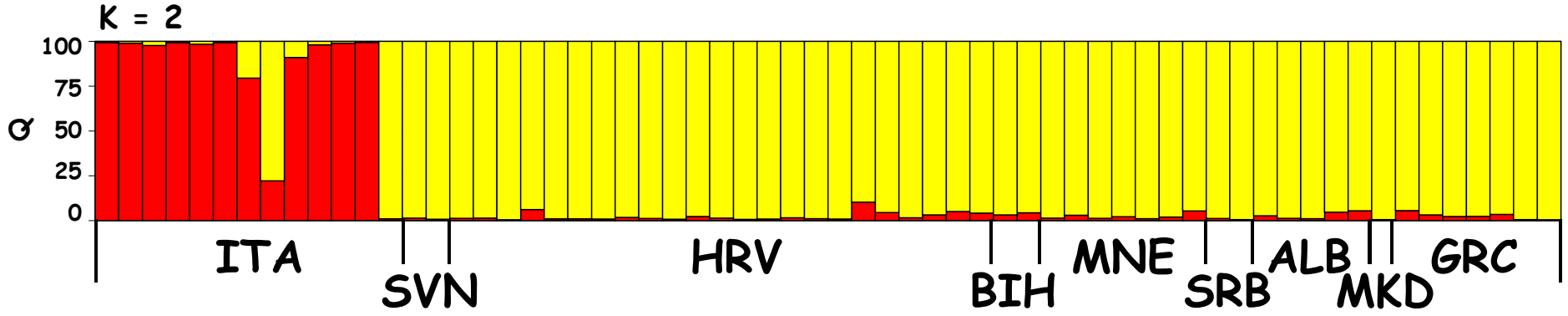
- assign individuals to populations of origin in such a way that within each population the departures from:
  (1) Hardy-Weinberg equilibrium (HWE) and
  (2) linkage equilibrium (LE)
  are minimized

NUMBER OF CLUSTERS: *ln* P(X|K) and ΔK



- *ln* P(X|K) values (20 runs per each K; primary y axis - left)
- ΔK values for each K (secondary y axis - right)

K = 2

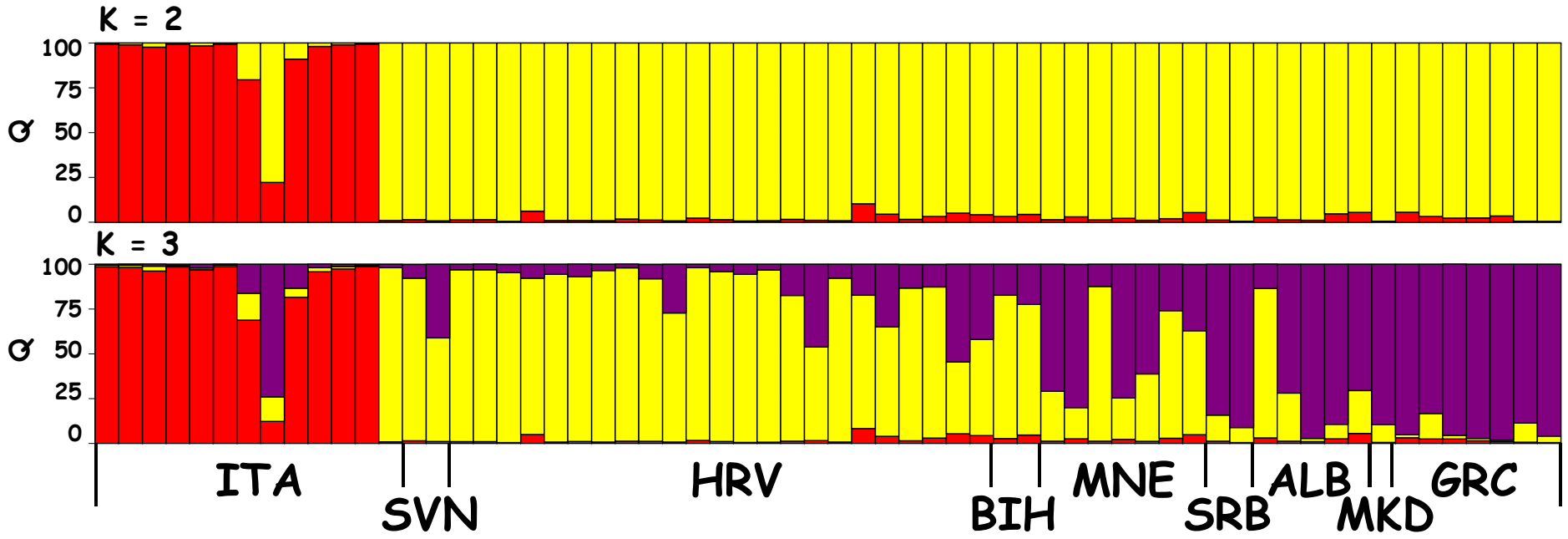


A
Apennine Peninsula

B
Balkan Peninsula

- mean proportion of membership (Q) of each population in each of the two clusters (populations of origin)

K = 3



Apennines

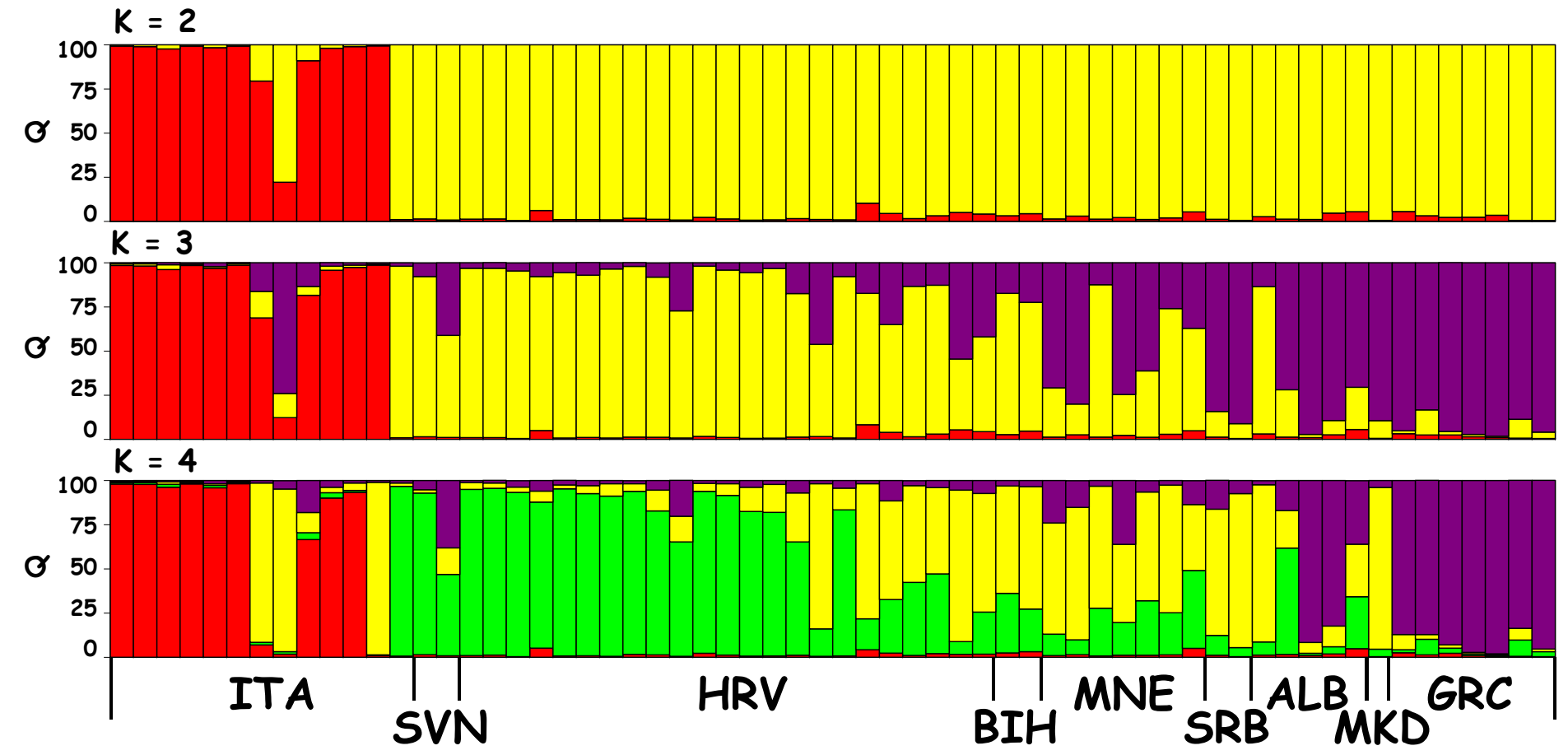
Balkans

North-West populations

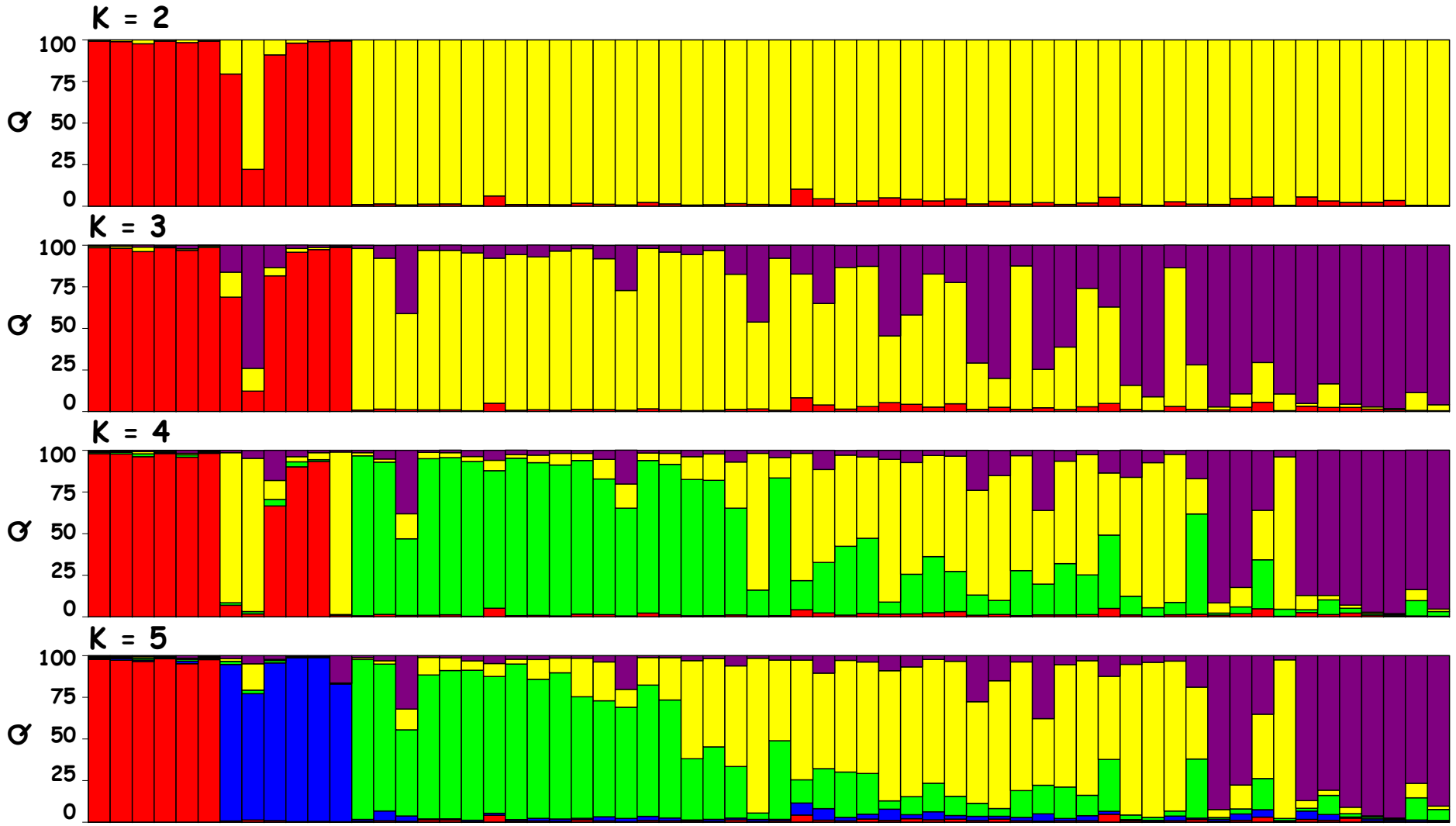
South-East populations

- the cluster B (Balkan) is at K = 3 split into two clusters according to geographical locations of the sampling sites:
  North-West: Slovenia-Croatia-BiH
  South-East: Montenegro-Albania-Macedonia-Greece

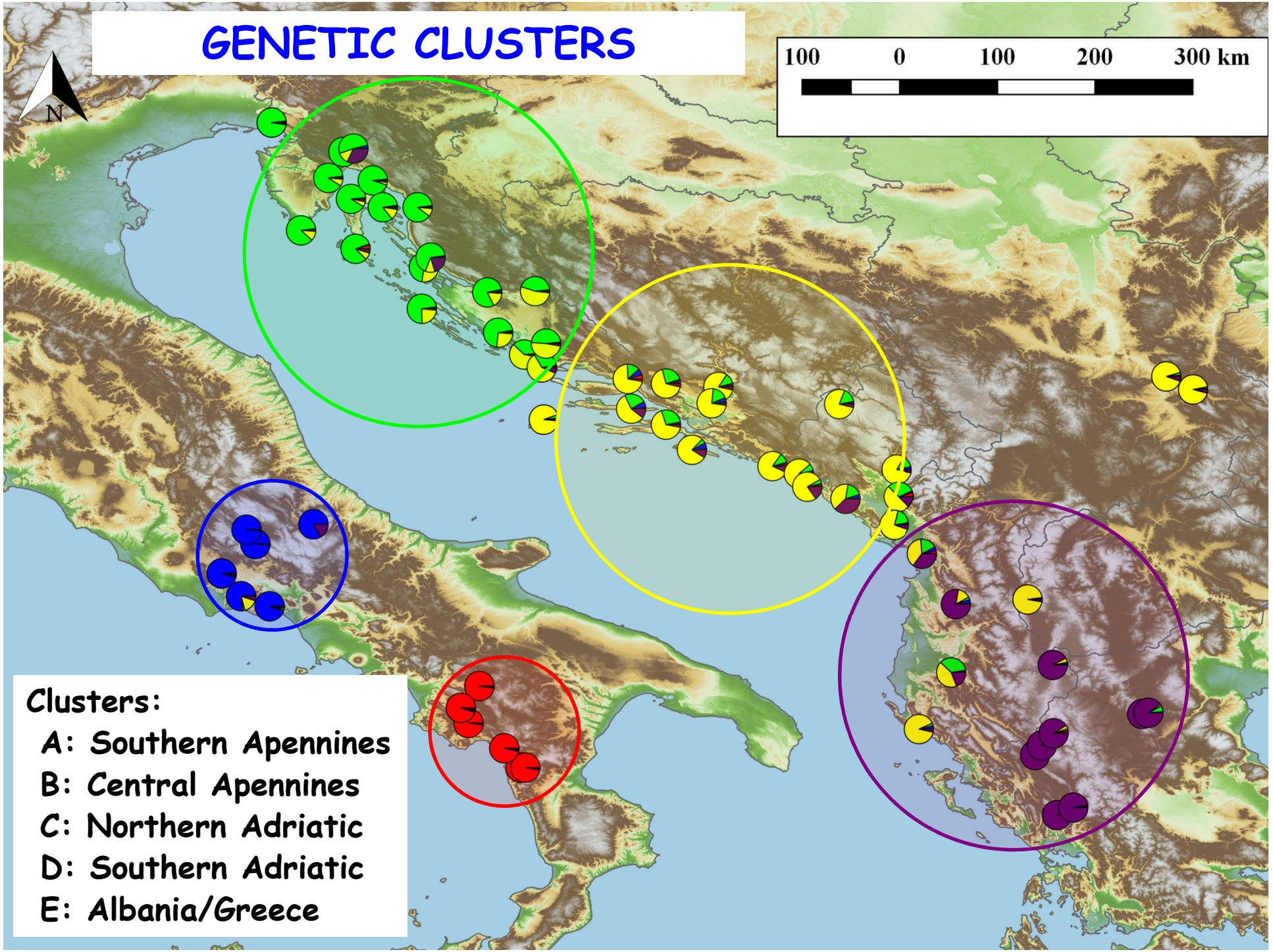
K = 4


K = 2
K = 3
K = 4
100
75
50
25
0
Q
ITA
SVN
HRV
BIH
MNE
SRB
ALB
MKD
GRC
Apennines
Balkans
Northern Adriatic
Southern Adriatic
Albania Greece

K = 5



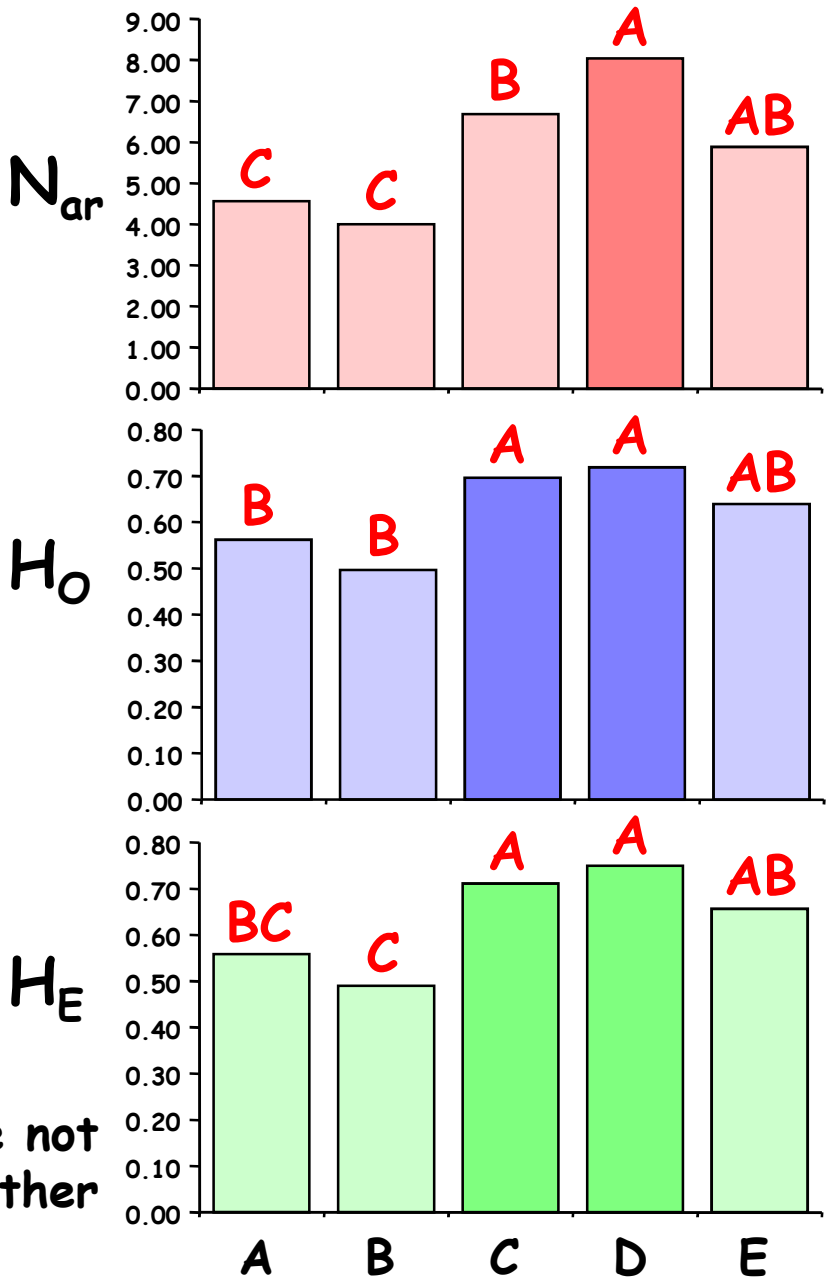
A B C D E


GENETIC CLUSTERS
N
100 0 100 200 300 km
Clusters:
A: Southern Apennines
B: Central Apennines
C: Northern Adriatic
D: Southern Adriatic
E: Albania/Greece

CLUSTER DIVERSITY

	Cluster	No. populations	N_{pr}
A	Southern Apennines	6	6
B	Central Apennines	6	0
C	Northern Adriatic	15	2
D	Southern Adriatic	26	26
E	Albania/Greece	9	9

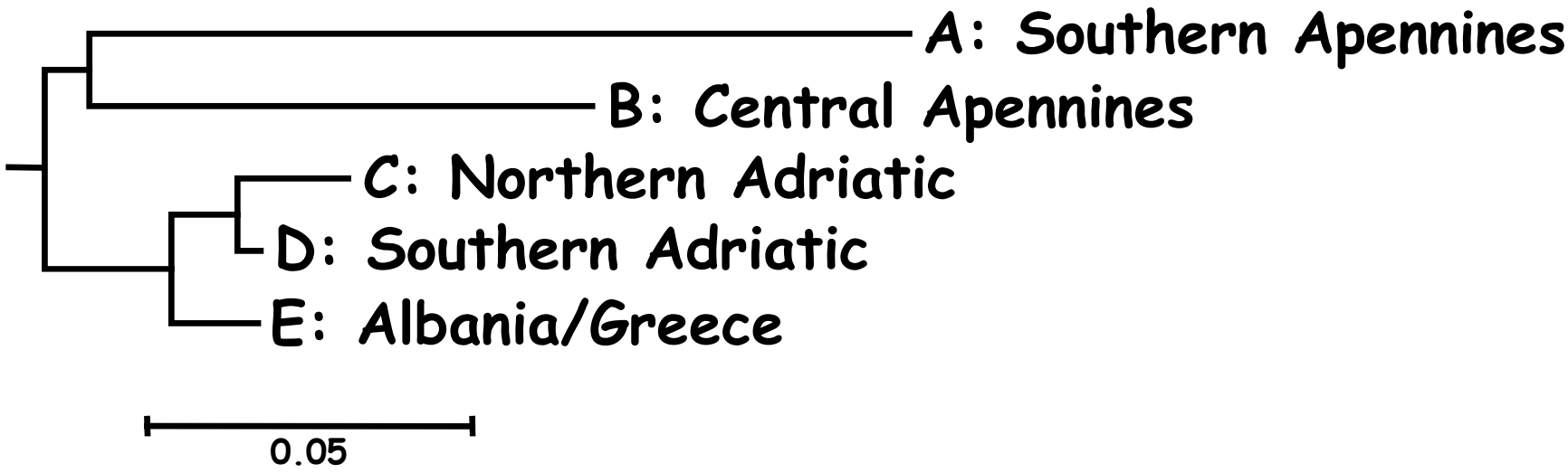
N_{pr} - No. of private alleles
N_{ar} - Allelic richeness
H_O - Observed heterozygosity
H_E - Expected heterozygosity



Means with the same letter are not significantly different from each other

RELATIONSHIPS AMONG CLUSTERS

Distance measure: net nucleotide distance
Tree algorithm: Neighbor-Joining



PLAN

(1) Dalmatian sage (*Salvia officinalis* L.)

(2) Genetic diversity

(3) Genetic structure

(4) Ecological niche modelling

(5) Demographic history

(4) ECOLOGICAL NICHE MODELLING

- species distribution modelling (SDMs)
  - estimate the relationship between species records at sites and the environmental characteristics of those sites
  - predict the suitability of sites for occupation or persistence of the species
  - produce a modelled distribution of the species (= identify species' suitable environmental space)
- input data:

  (1) geographic distribution

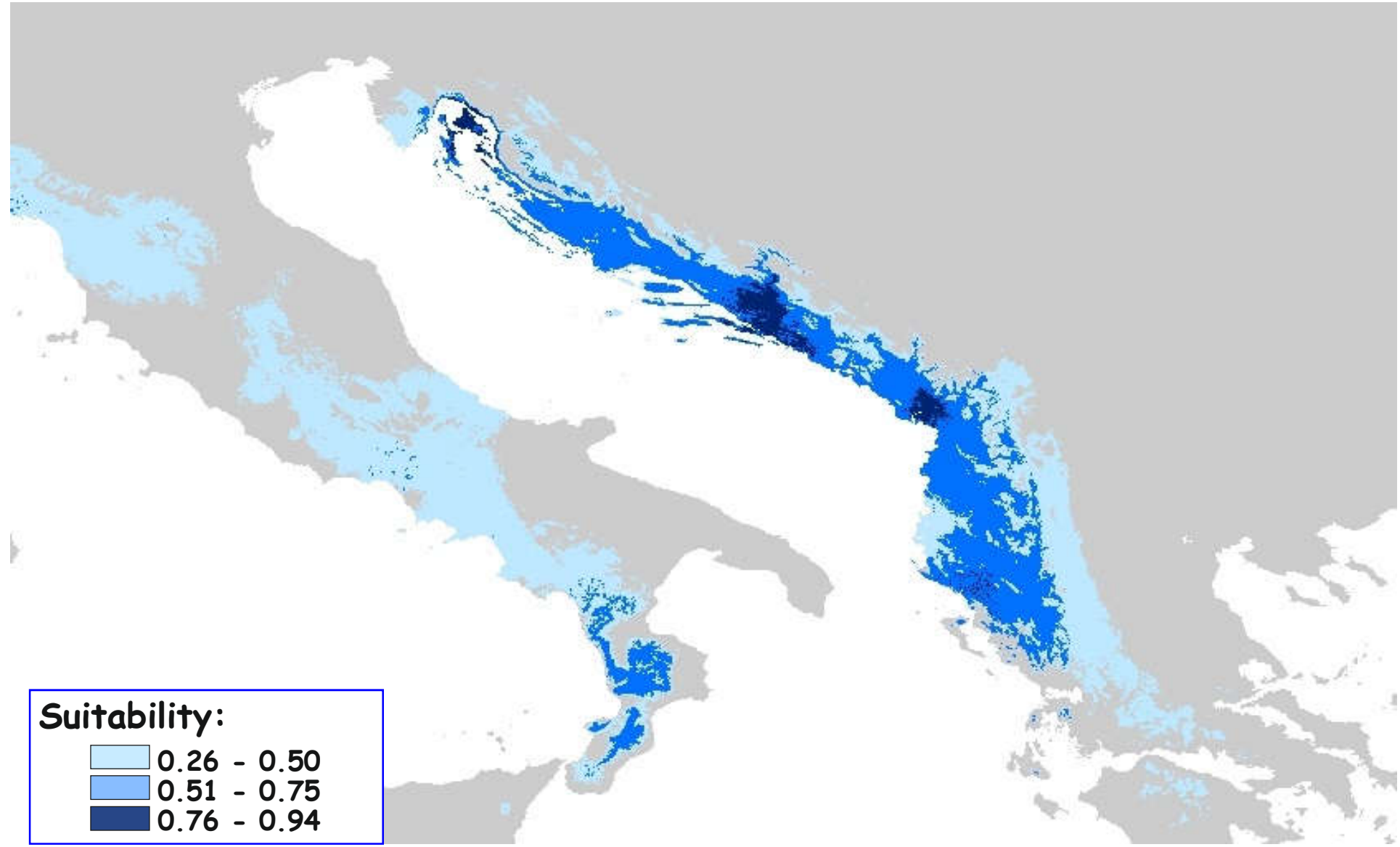
  68 data on occurrence (evenly distributed)

  (2) environmental characteristics of the sites

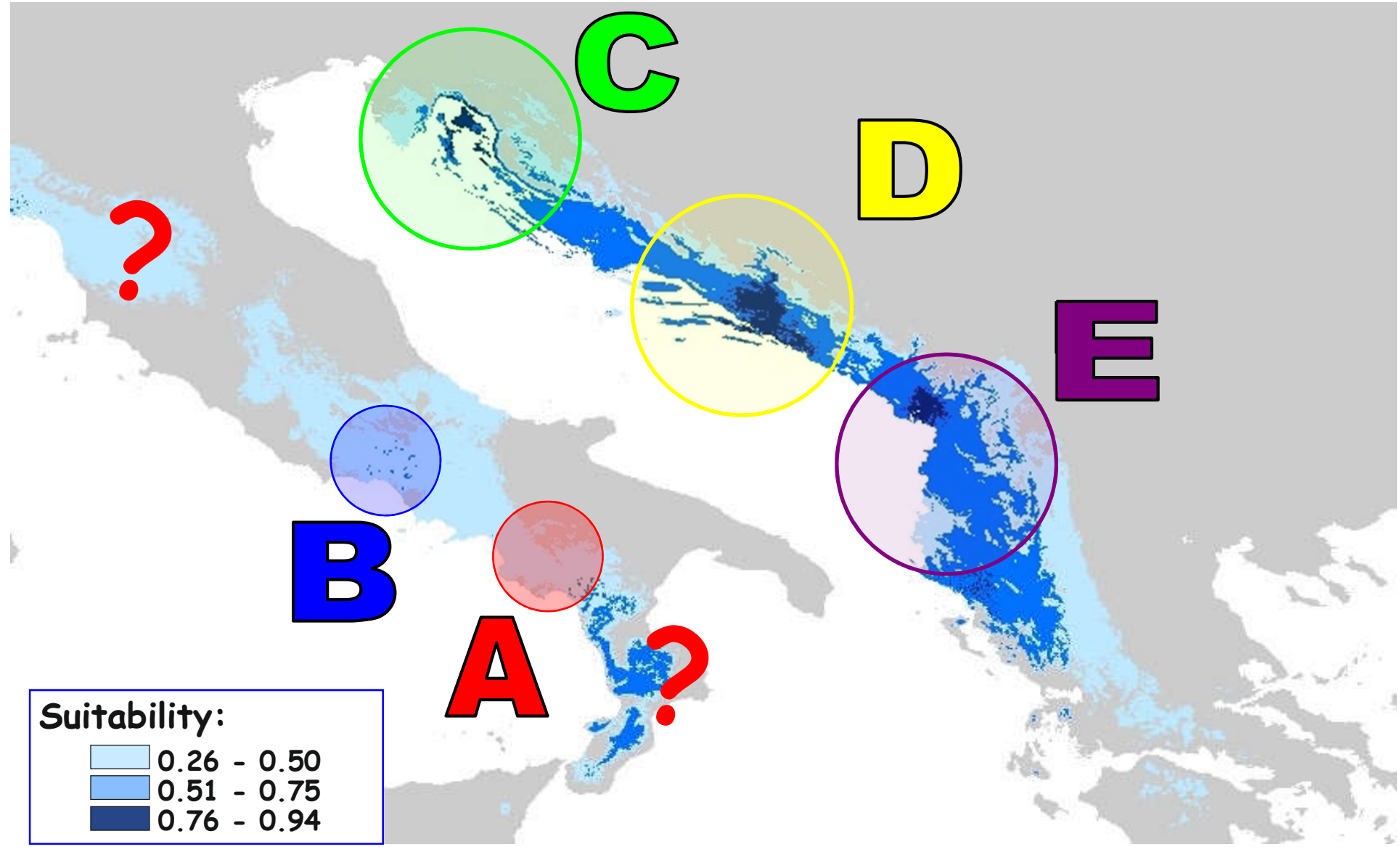
  19 bioclimatic variables (WorldClim database)

  - 11 temperature- and 8 precipitation-related
  - representing the annual trends, seasonal variations and extremes in temperature and precipitation

ENM: PRESENT DAY CONDITIONS



ENM: PRESENT DAY CONDITIONS



SUITABILITY: EDAPHIC FACTORS

Soil Atlas of Europe

Classification: The World Reference Base for Soil Resources (WBR)

(1) Balkans

- Leptosols

Shallow soils over limestone

(2) Apennines

- Cambisols

Moderatelly developed brown soils

PAST DISTRIBUTION

- to model potential species distribution during the Last Glacial Maximum (LGM; ~26,500 years BP)
- to identify putative glacial refugia of the species (= region which made possible the survival of the species and allowed a post-glacial re-colonization)
- input:

(1) the present model

(2) past environmental characteristics of the sites

19 bioclimatic variables

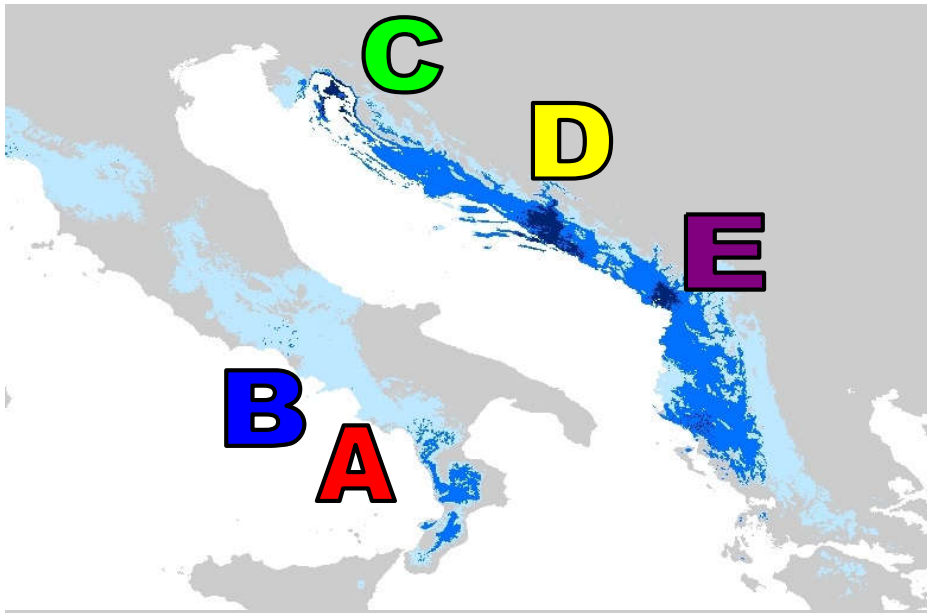
bioclimatic data for the LGM developed by Paleoclimate Modelling Intercomparison Project Phase II

- two models:

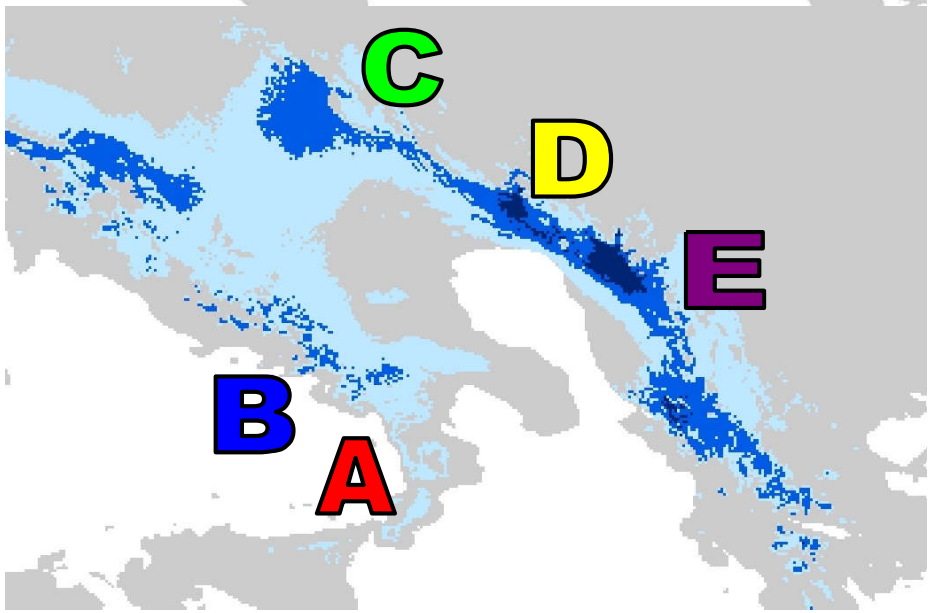
CCSM (Community Climate System Model; USA)

MIROC (Model for Interdisciplinary Research on Climate; Japan)

ECOLOGICAL NICHE MODELLING



- ENM: present day



- ENM: ~26,500 years BP
- distribution during the Last Glacial Maximum
- model MIROC
- putative glacial refugia ?

PLAN

(1) Dalmatian sage (*Salvia officinalis* L.)

(2) Genetic diversity

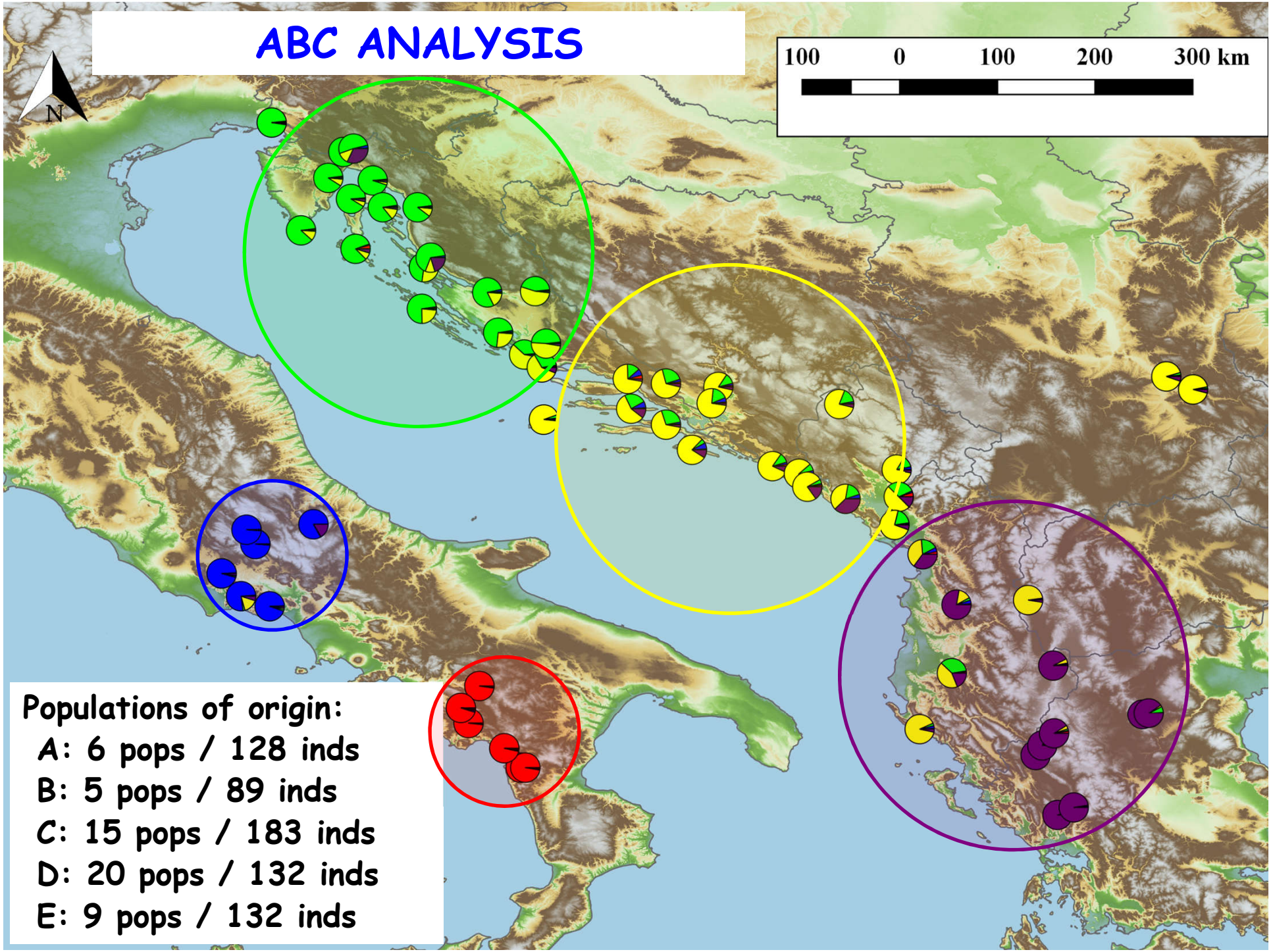
(3) Genetic structure

(4) Ecological niche modelling

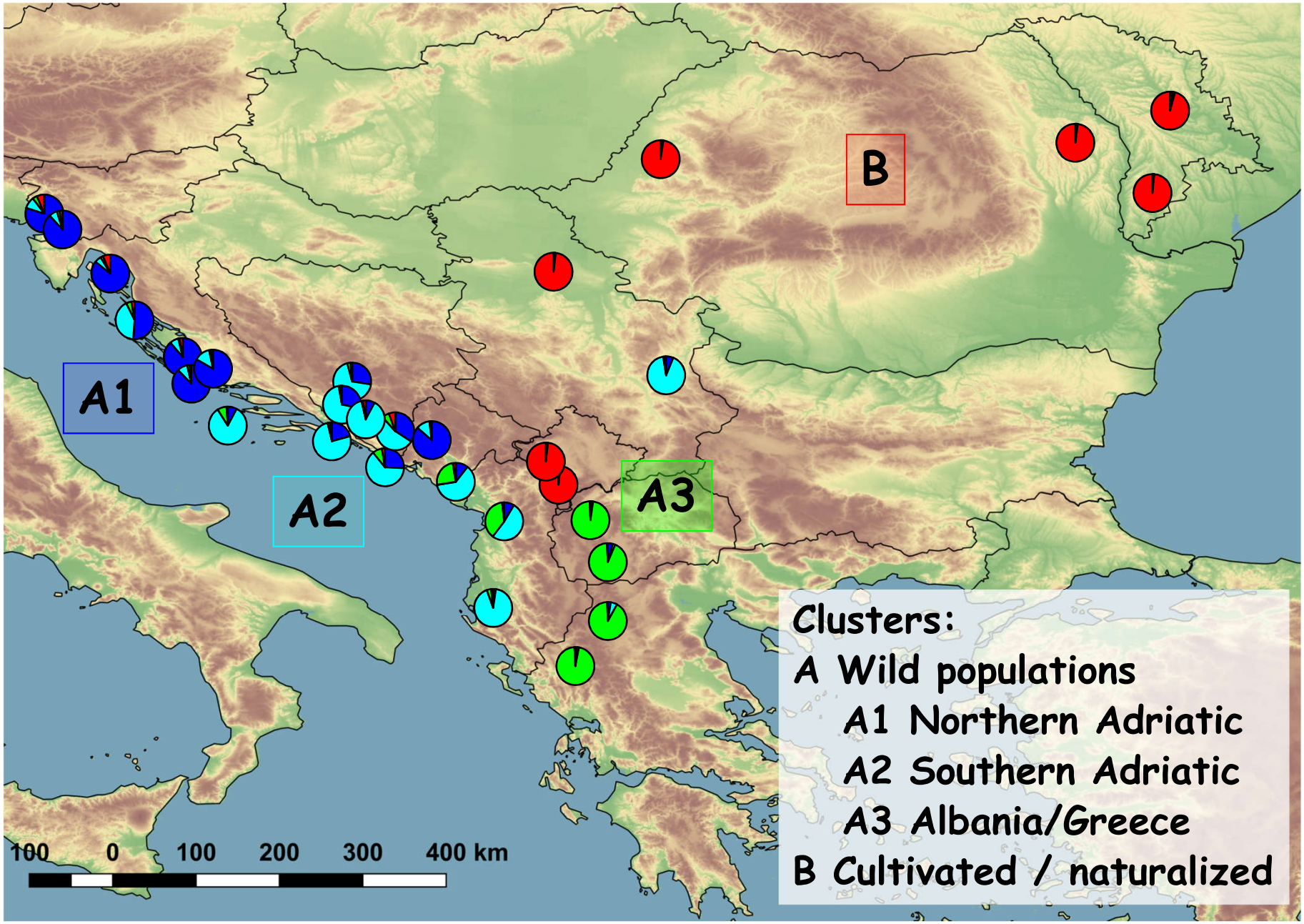
(5) Demographic history

(5) DEMOGRAPHIC HISTORY

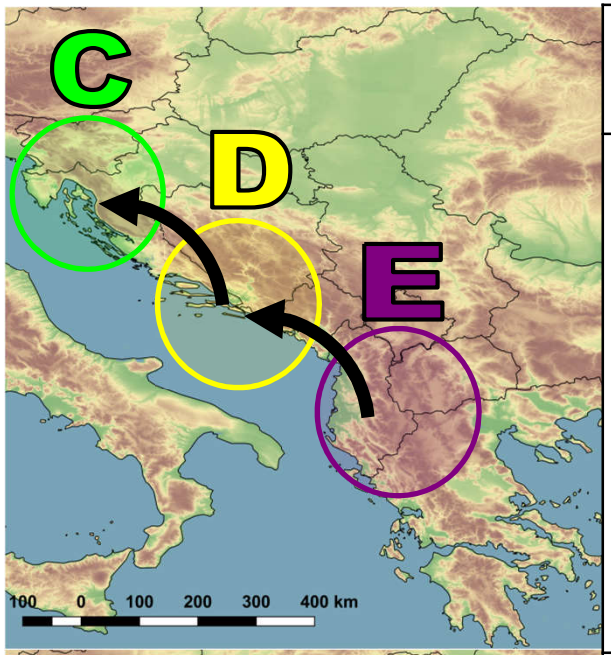
- inference on demographic history of a species based on approximate Bayesian computation (ABC)
- input:
  - molecular data
  - scenarios describing demographic history
- procedure:
  - simulate (a large number of) datasets for each scenario
  - compare simulated and observed datasets
  - the most similar simulated dataset is the most likely
- goal:
  - compare competing scenarios
    posterior probability of each scenario
  - estimate parameters for one or more scenarios
    effective population size; time of divergence


ABC ANALYSIS
N
100 0 100 200 300 km
Populations of origin:
A: 6 pops / 128 inds
B: 5 pops / 89 inds
C: 15 pops / 183 inds
D: 20 pops / 132 inds
E: 9 pops / 132 inds

PREVIOUS RESEARCH: Rešetnik et al. (2016)

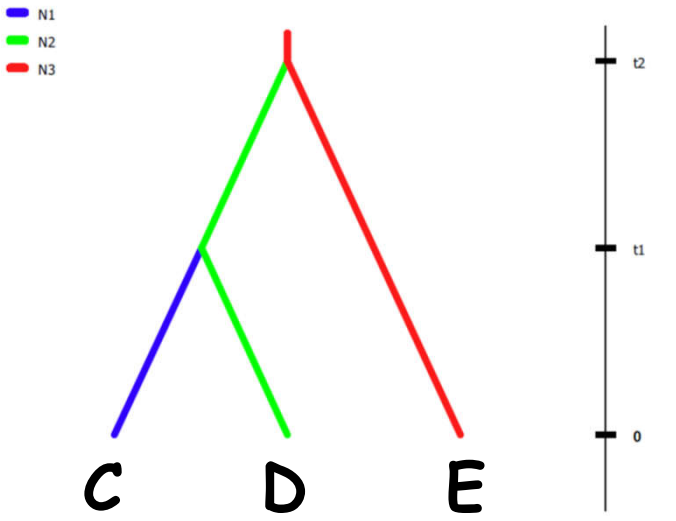


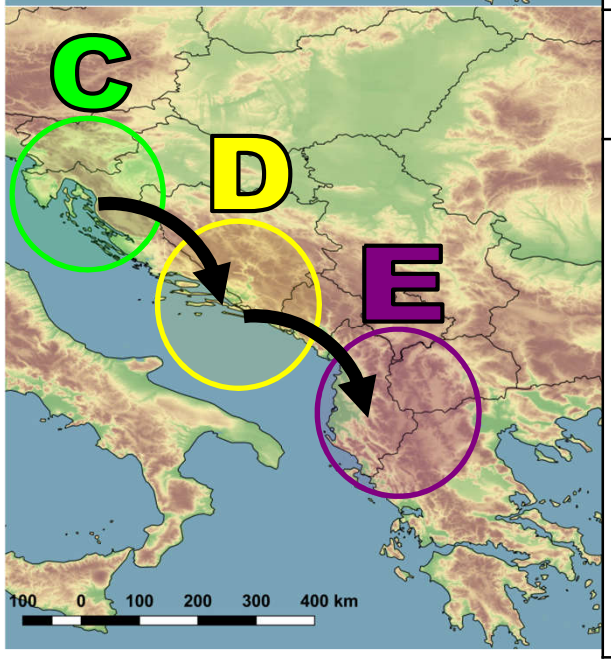
BALKANS: FIVE SIMPLE HISTORIC SCENARIOS



Scenario 1

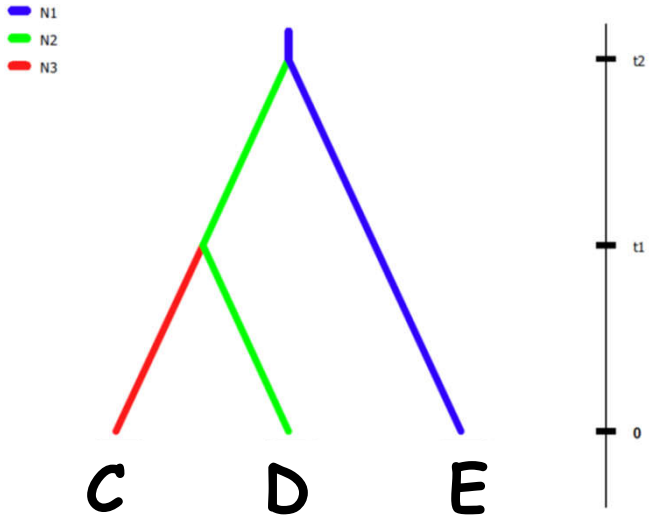
Population C is derived from population D, itself derived from population E



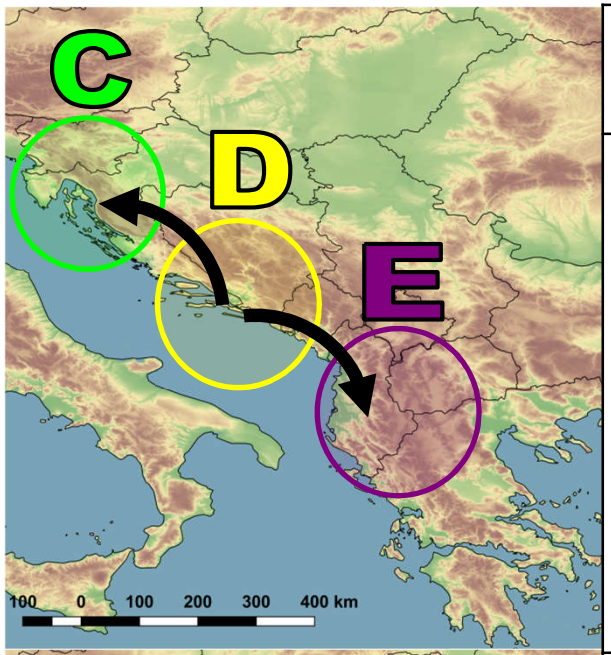


Scenario 2

Population E is derived from population D, itself derived from population C

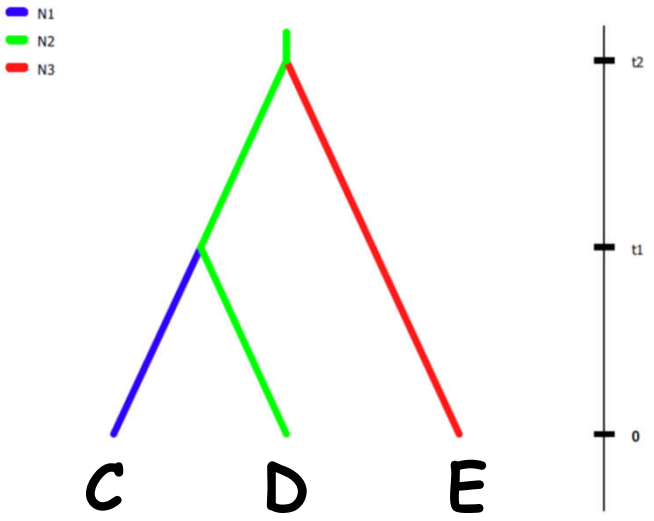


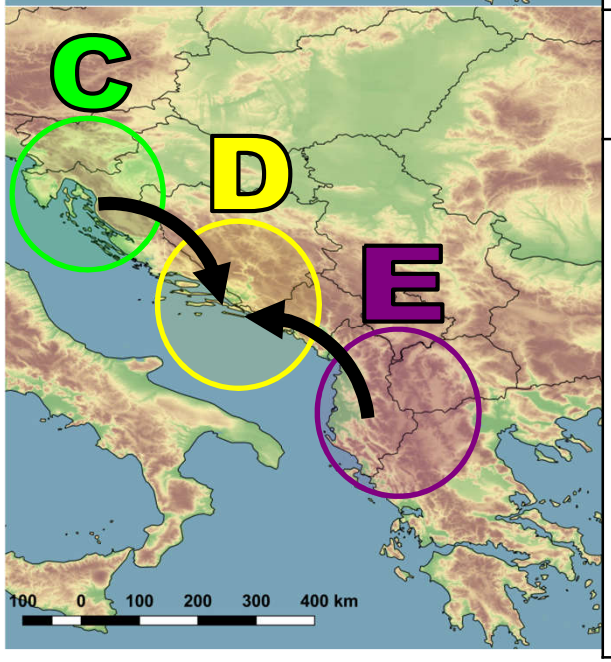
BALKANS: FIVE SIMPLE HISTORIC SCENARIOS



Scenario 3

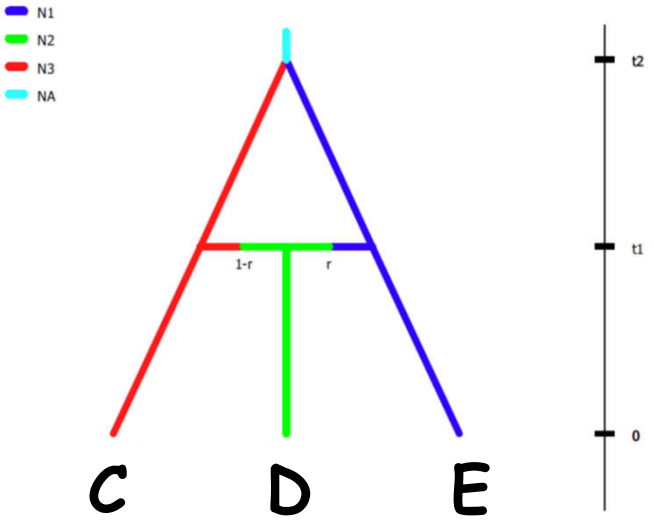
Both populations C and E derived independently from population D



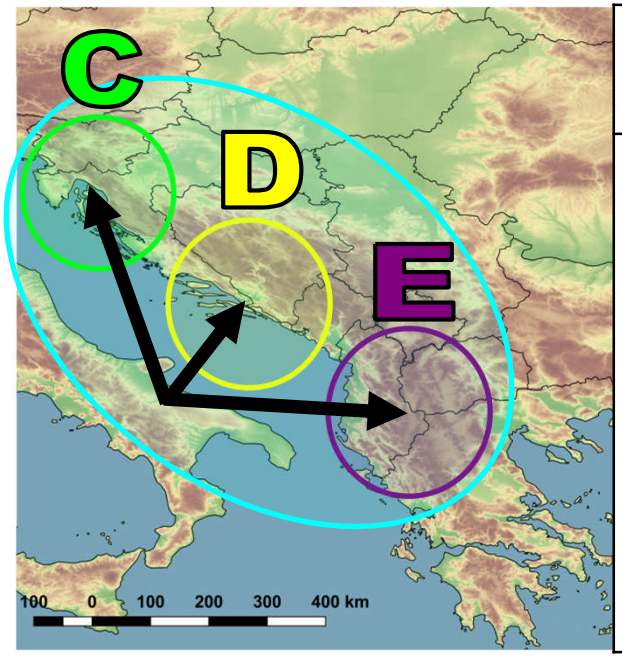


Scenario 4

Population D was generated by admixture of populations C and E

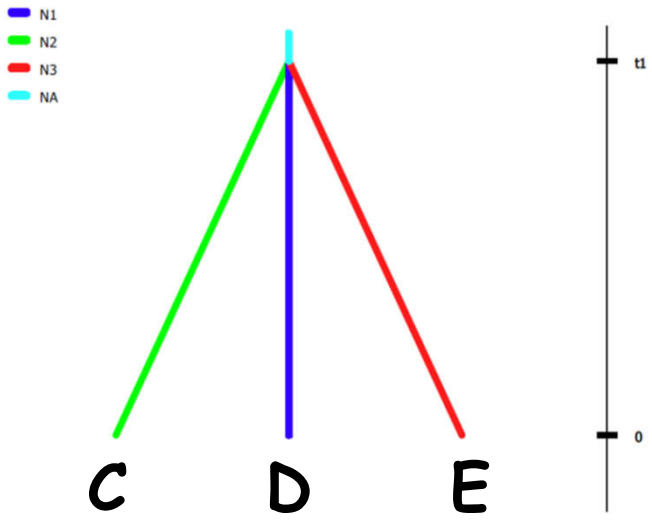


BALKANS: FIVE SIMPLE HISTORIC SCENARIOS



Scenario 5

All three populations diverged at the same time

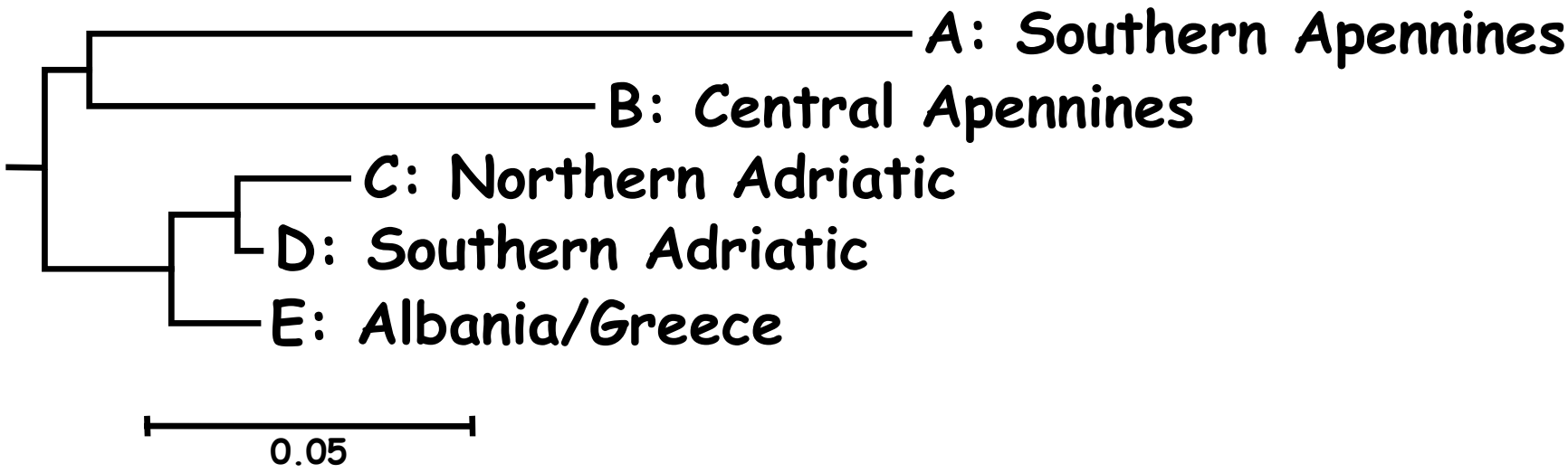


RESULTS:

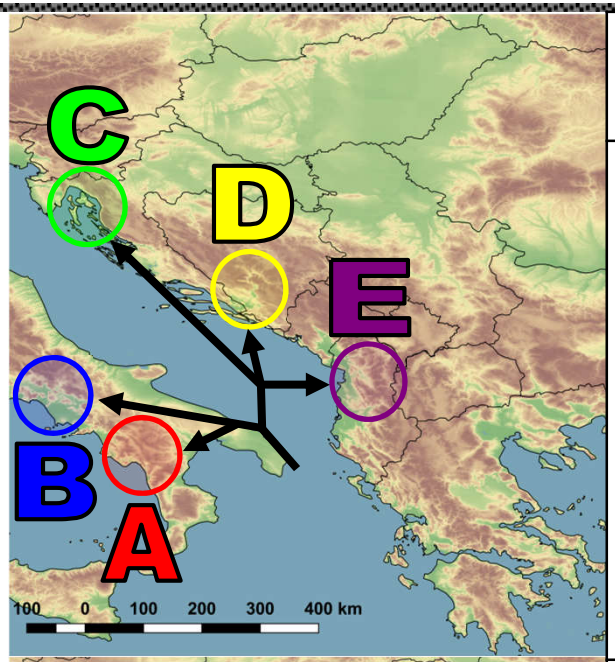
Scenario	Posterior probability (PP)
1	0.114
2	0.009
3	0.107
4	0.007
5	0.764

APENNINES + BALKANS: HISTORIC SCENARIOS ?

Distance measure: net nucleotide distance
Tree algorithm: Neighbor–Joining

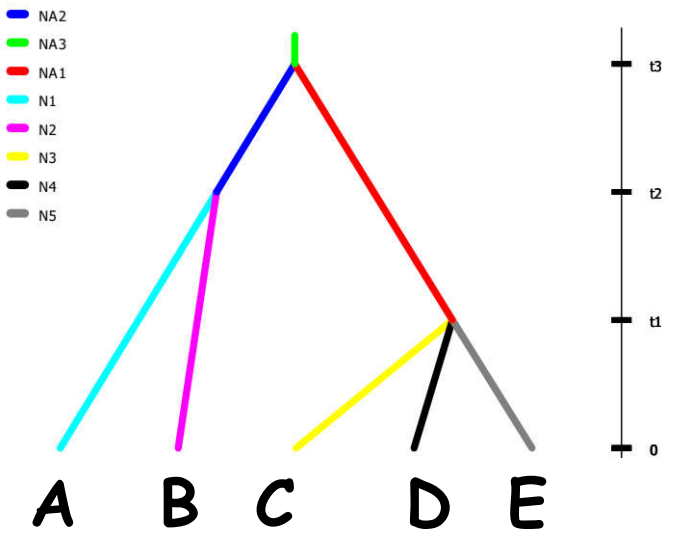


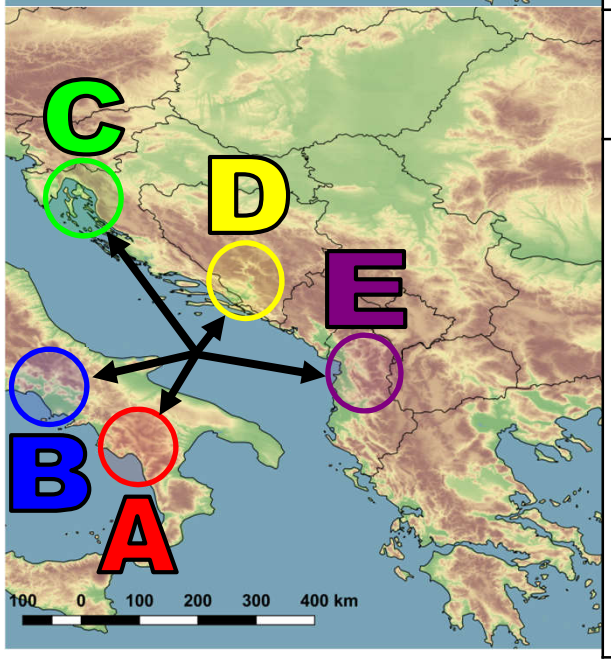
APENNINES + BALKANS: HISTORIC SCENARIOS



Scenario 1

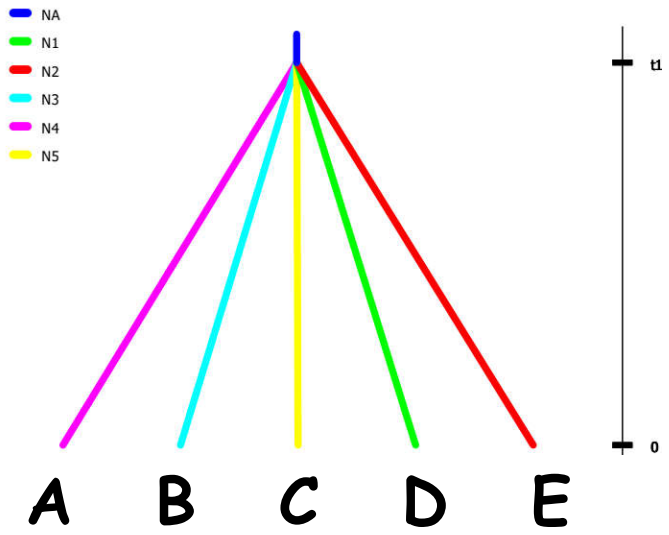
Apennine population group diverged from Balkan population group





Scenario 2

All five populations diverged at the same time



POSTERIOR PROBABILITIES OF SCENARIOS

Scenario	Posterior probability (Confidence interval)
1	0.384 (0.307-0.462)
2	0.616 (0.538-0.693)

- Scenario 2: All five populations diverged at the same time

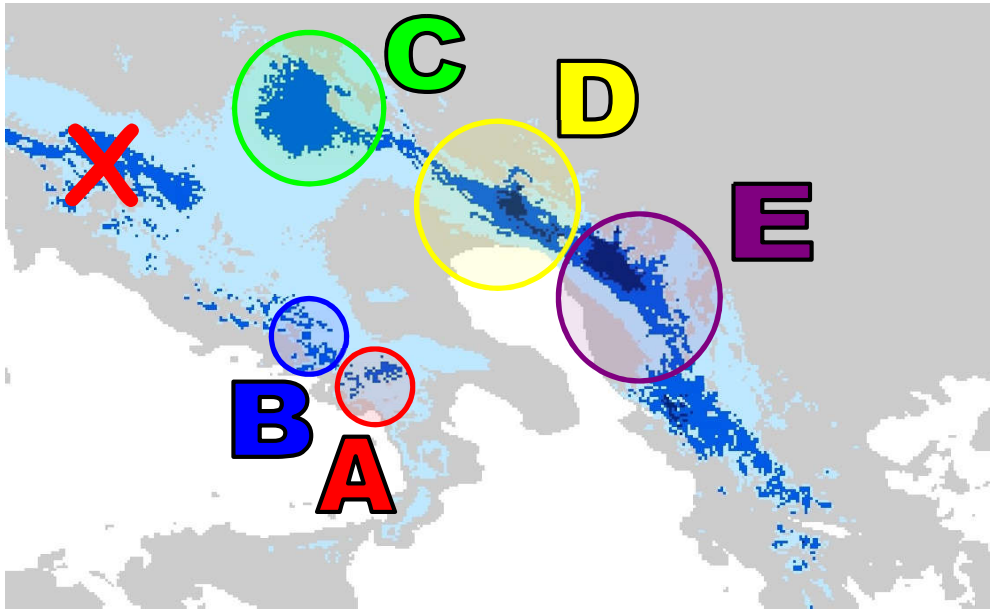
- estimation of parameters under Scenario 2
  - effective population size
    - of the ancestral population
    - of the populations of origin (A-E)
  - time of divergence
    - in number of generations

HISTORICAL RECONSTRUCTION: ANCIENT PAST

Species distribution before the last glaciation (?)

Effective populations sizes:

Ancestral	9,280 individuals
A	1,140 individuals
B	940 individuals
C	4,360 individuals
D	8,330 individuals
E	5,220 individuals

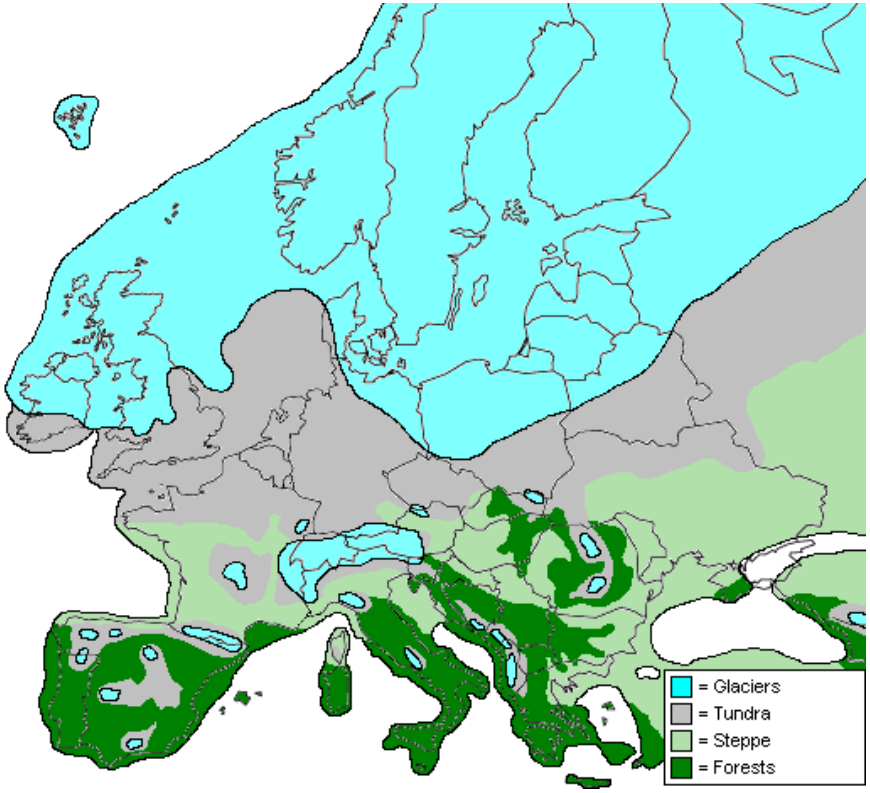


Species distribution during the last glaciation

(Last Glacial Maximum; LGM; ~26,500 years BP)

HISTORICAL RECONSTRUCTION: GLACIATION

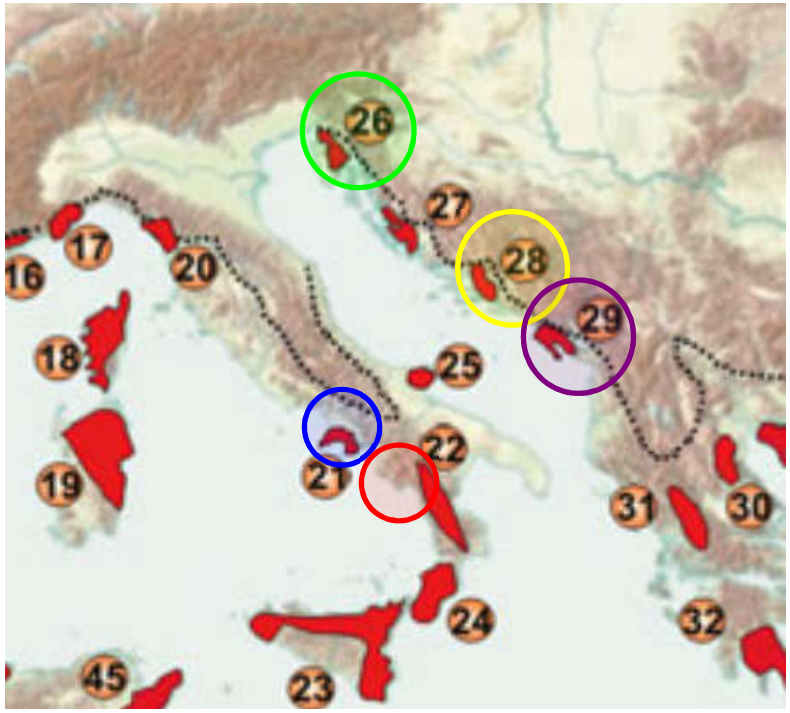
Vegetation map of Europe during the last glaciation



Glacial refugia:

1. Iberian Peninsula
2. Apennine Peninsula
3. Balkan Peninsula

Glacial refugia in the Mediterranean Basin

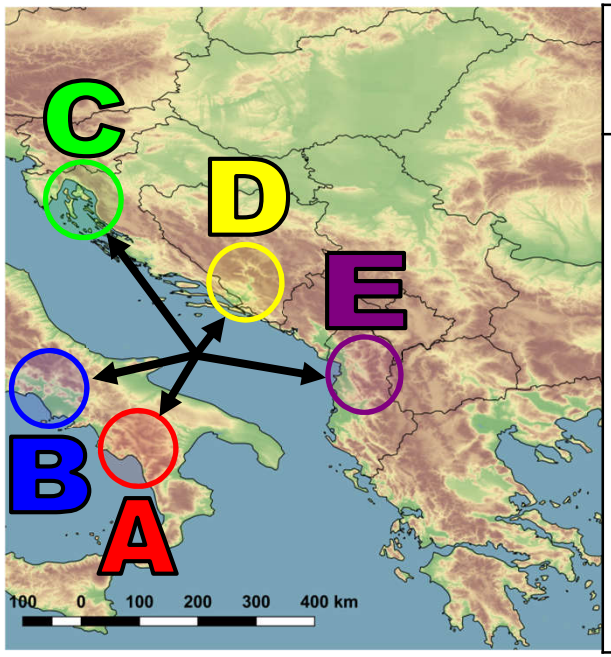


Médail and Diadema (2009)

Glacial (micro)refugia:

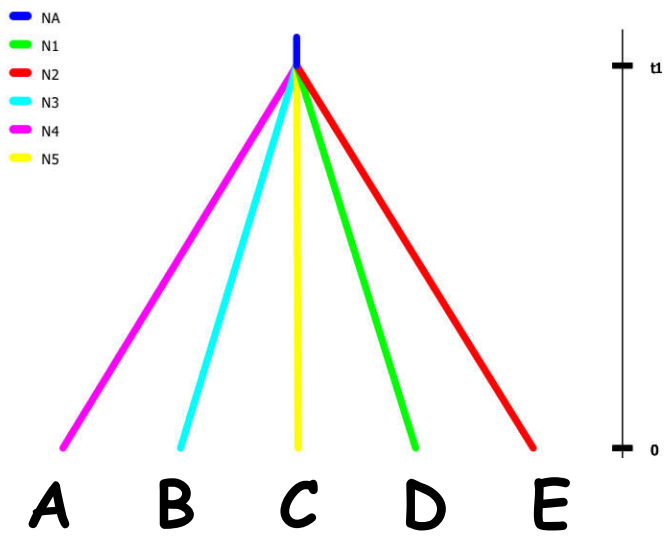
A?	22. S. Apennines
B?	21. Campania
C	26. N. Istria
D	28. S. Bosnia / Biokovo
E	29. Montenegro

HISTORICAL RECONSTRUCTION: DIVERGENCE



Scenario 2

All five populations diverged at the same time



Time of divergence (t1):

1,220 generations ago

Last glacial maximum (LGM):

26,500 years BP

Generation time of Dalmatian sage:

~20 years (?)

Genetic Diversity and Evolutionary History of *Salvia officinalis* L.

Zlatko Šatović

University of Zagreb, Faculty of Agriculture, Zagreb, Croatia

Centre of Excellence for Biodiversity and Molecular Plant Breeding (CoE CroP-BioDiv), Zagreb, Croatia

E-mail: zsatovic@agr.hr

Europska unija
Zajedno do fondova EU

EUROPSKI STRUKTURNI I INVESTICIJSKI FONDOVI

Operativni program KONKURENTNOST I KOHEZIJA

Republika Hrvatska