# GENETIC DIVERSITY OF WILD vs. CULTIVATED/NATURALIZED POPULATIONS OF DALMATIAN SAGE


**Zlatko Šatović**

**University of Zagreb, Faculty of Agriculture, Zagreb, Croatia**

**Centre of Excellence for Biodiversity and Molecular Plant Breeding, Zagreb, Croatia**

**E-mail: zsatovic@agr.hr**

# PLAN

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

(2) Genetic diversity and relationships

(3) Genetic structure

(4) Ecological niche modelling

(5) Demographic history

# PLAN

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

(2) Genetic diversity and relationships

(3) Genetic structure

(4) Ecological niche modelling

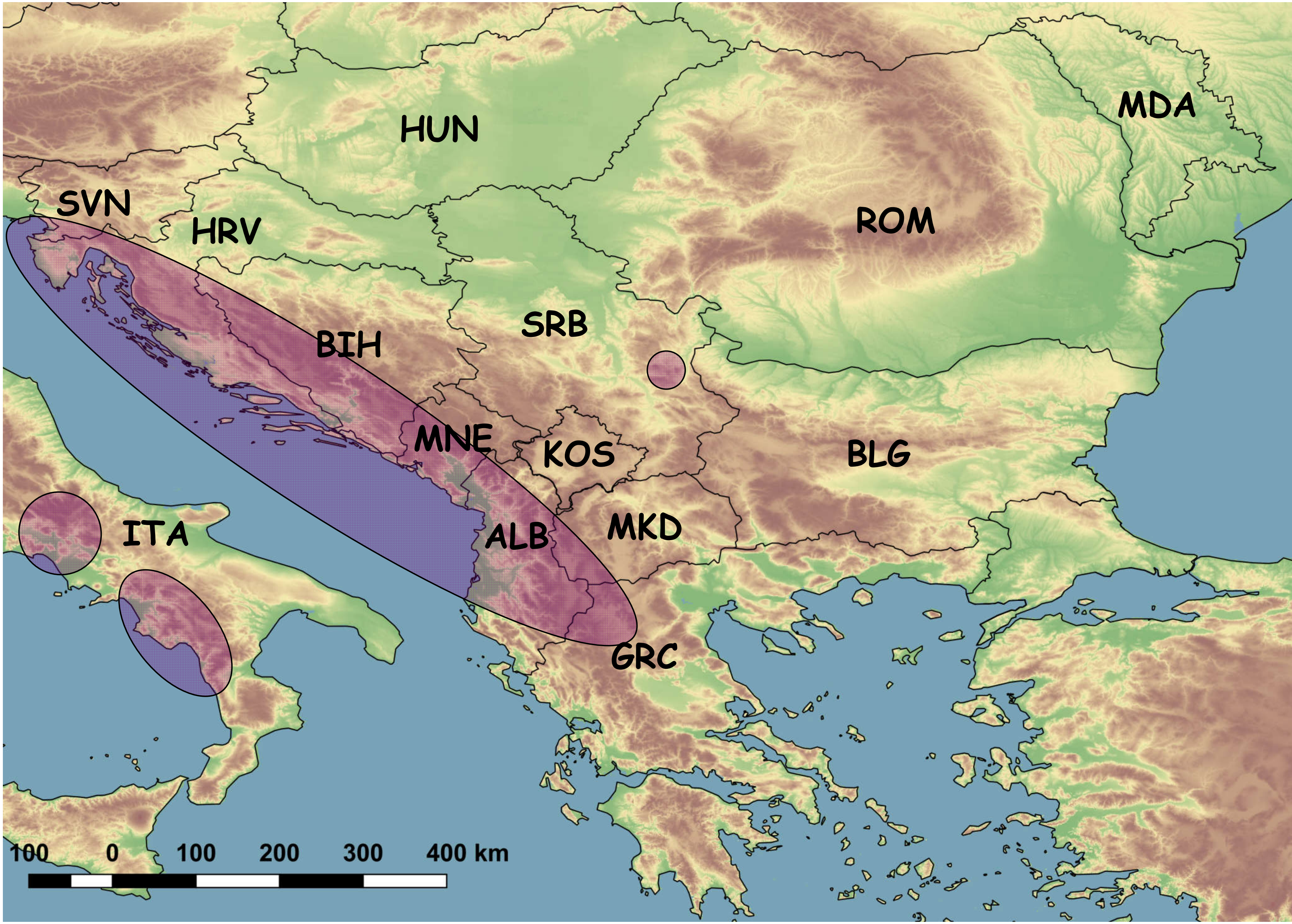
(5) Demographic history

# 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 populations:
  plants that have escaped from earlier cultivation

# NATURAL DISTRIBUTION OF DALMATIAN SAGE



# PLAN

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

(2) Genetic diversity and relationships

(3) Genetic structure

(4) Ecological niche modelling

(5) Demographic history

# GENETIC DIVERSITY

## (1) Collecting

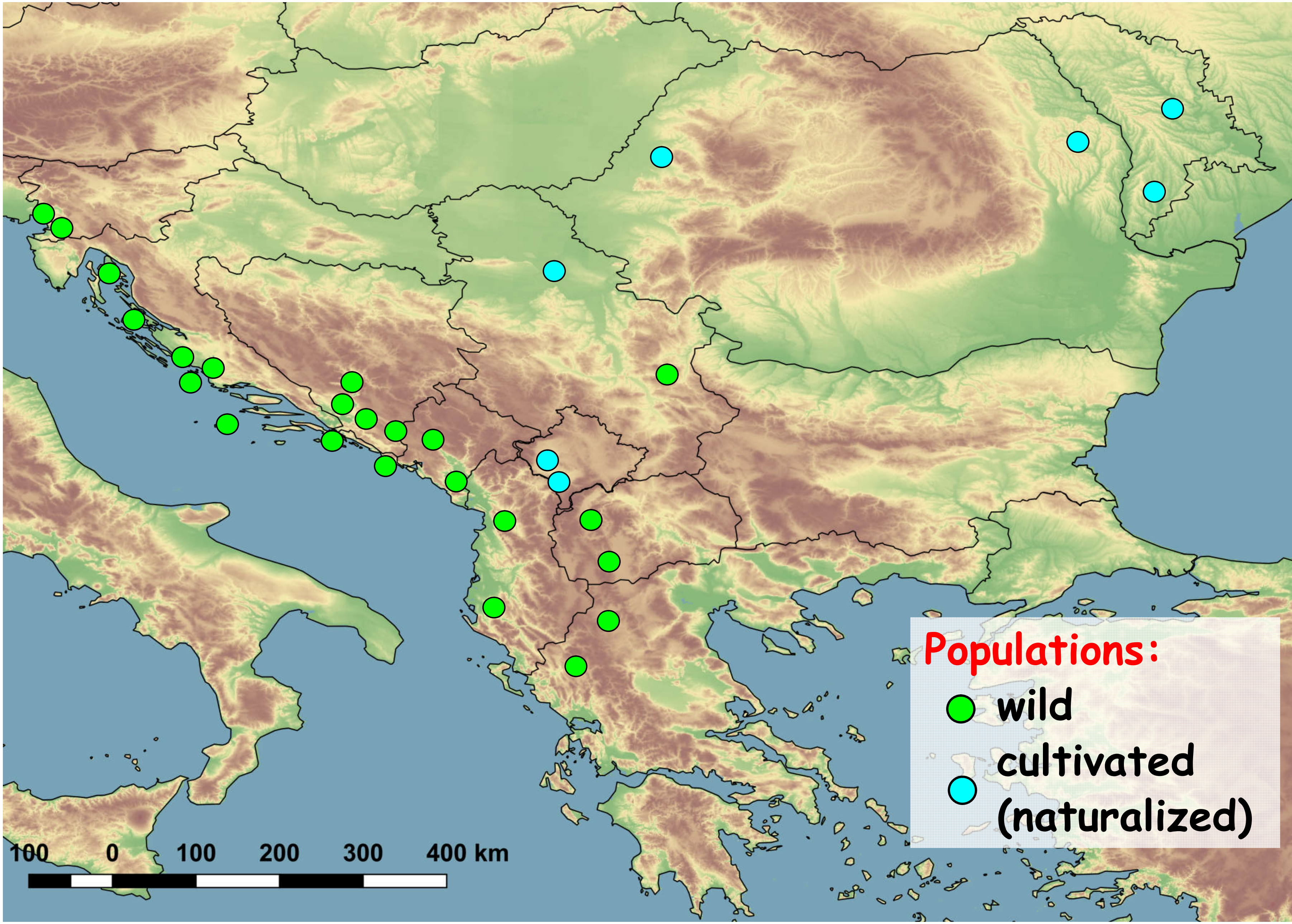
- leaf tissue for DNA extraction
- 30 populations / 709 samples (20 to 25 per populations)
- origin:
    - 11 Balkan countries
- status:
    - 23 wild populations
    - 7 cultivated/naturalized populations

## (2) DNA extraction

## (3) Molecular marker analysis

- 8 microsatellite markers developed for Dalmatian sage (simple sequence repeats; SSRs)
- 165 alleles

# COLLECTING


Populations:
wild
cultivated (naturalized)
100 0 100 200 300 400 km


Komiža, island of Vis


Cavtat

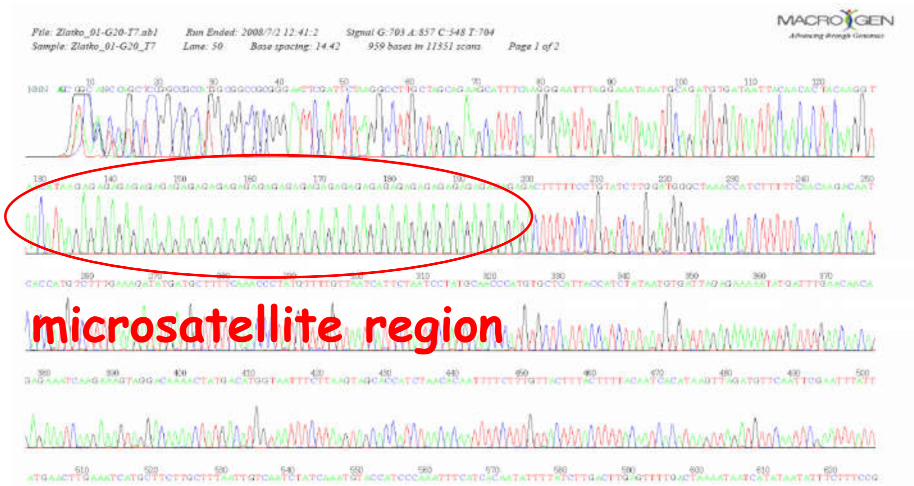
# MOLECULAR MARKER ANALYSIS

## (A) Isolation and characterization of microsatellites

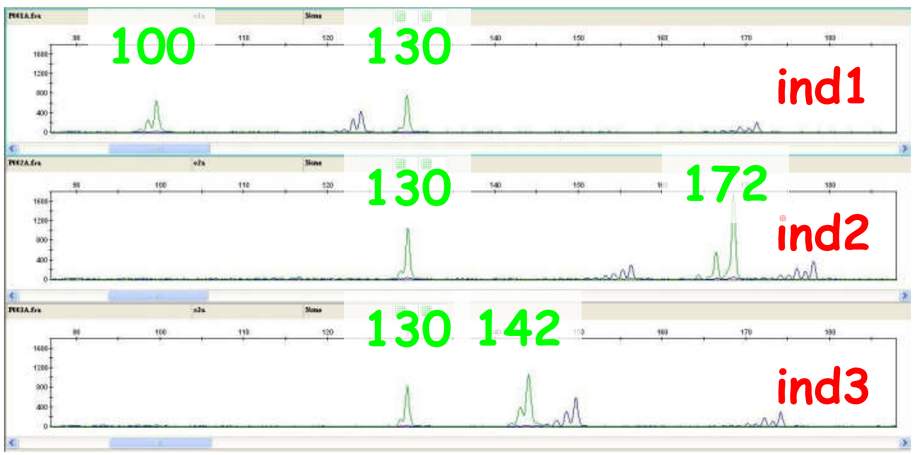
- construction of microsatellite libraries from genomic DNA for GA, GT, AGA and ACA repeats
- primer pairs flanking microsatellite repeats were designed for 29 clones
- 29 microsatellites were tested using 25 plants

## (B) Microsatellite marker analysis

- eight most polymorphic markers: 165 alleles




sequencing




scoring

# MOLECULAR DATA

| Population | Individual | SoUZ001 | SoUZ001 | SoUZ002 | SoUZ002 | ... | SoUZ019 | SoUZ019 |
|---|---|---|---|---|---|---|---|---|
| P01 | S001 | 173 | 189 | 183 | 185 | | 135 | 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 | | 135 | 150 |
| P01 | S006 | 165 | 183 | 185 | 187 | | 135 | 156 |
| P01 | S007 | 173 | 183 | 185 | 195 | | 150 | 159 |
| P01 | S008 | 165 | 167 | 183 | 195 | | 135 | 135 |
| P01 | S009 | 165 | 171 | 181 | 183 | | 147 | 150 |
| P01 | S010 | 165 | 173 | 195 | 201 | | 150 | 159 |
| ... | | | | | | | | |
| P30 | S709 | 161 | 165 | 185 | 195 | | 135 | 150 |

# GENETIC DIVERSITY

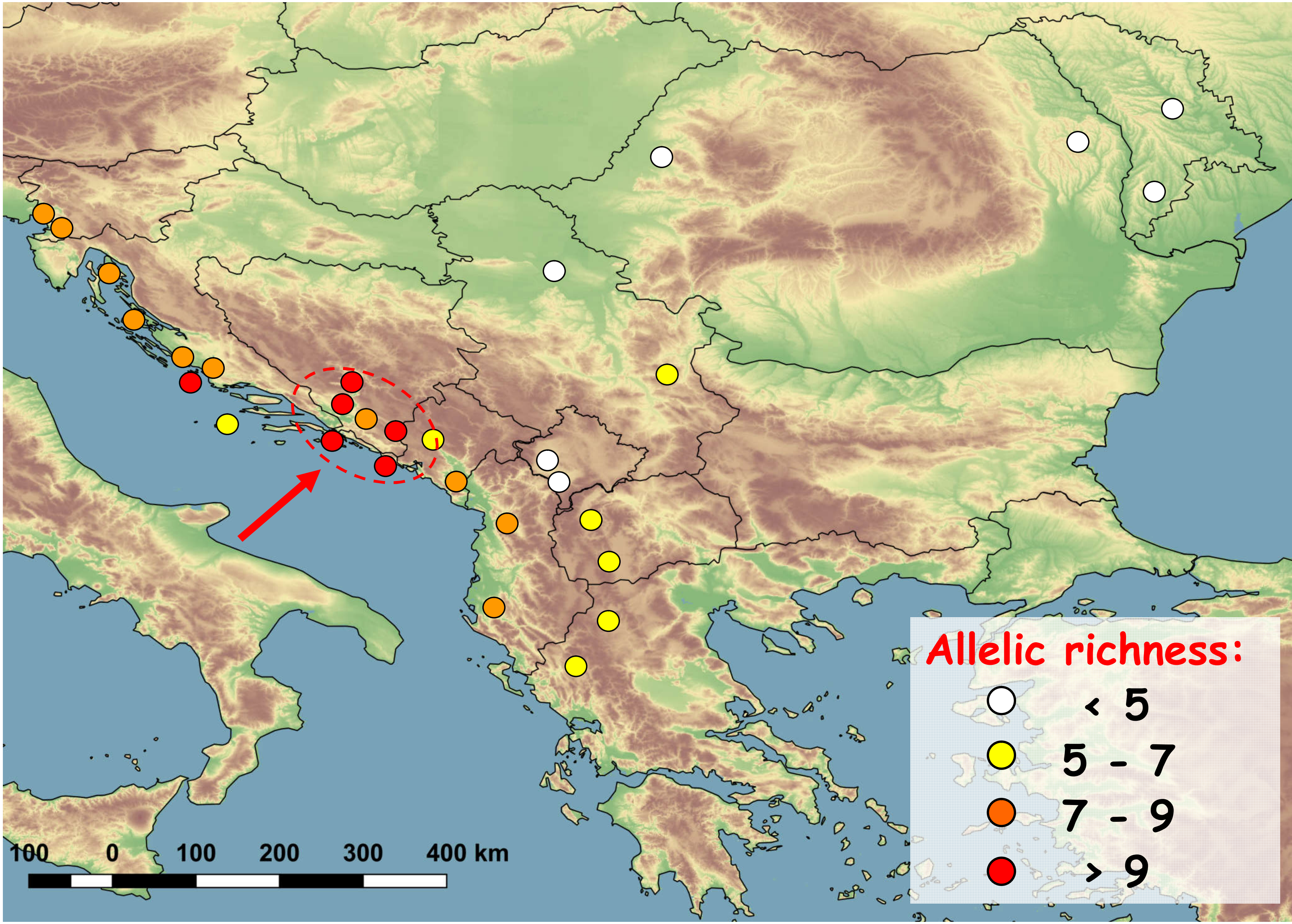
(1) Average no. of alleles per population ($N_{av}$)

(2) Allelic richness ($N_{ar}$)
- average no. of alleles per population independent of sample size

(3) No. of private alleles ($N_{pr}$)
- number of alleles detected in a single population
  (or in a group of populations: wild vs. cultivated)

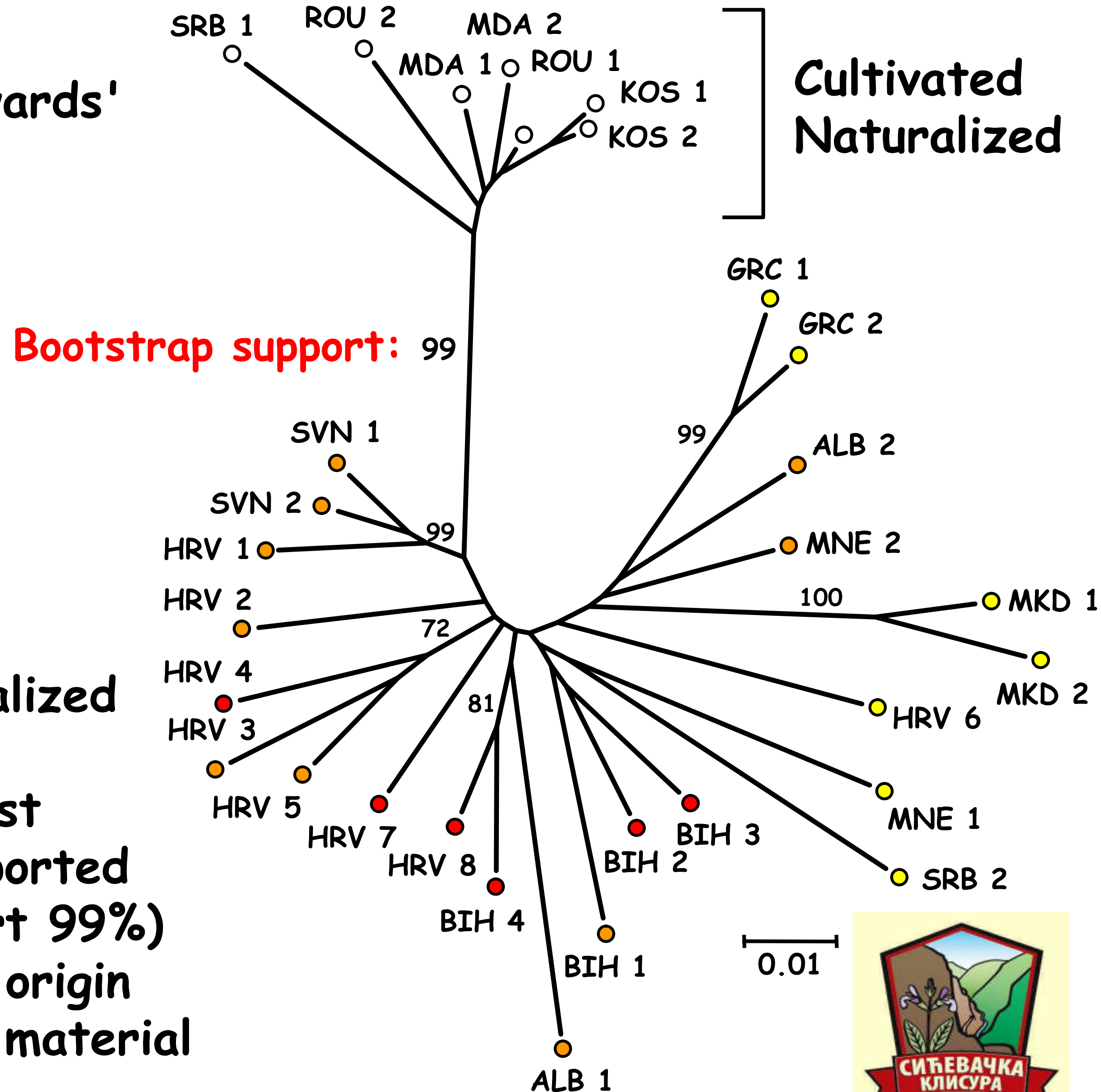
| | Wild | Cultivated |
|---|---|---|
| No. | 23 | 7 |
| $N_{av}$ | 8.696 | 3.821 |
| $N_{ar}$ | 7.920 | 3.672 |
| Range | 5.13-10.30 | 2.71-4.19 |
| P($N_{ar}$) | P < 0.001 | |
| $N_{pr}$ (total per population) | 20 | 0 |
| $N_{pr}$ (wild vs. cultivated) | 115 | 0 |

# ALLELIC RICHNESS



# GENETIC RELATIONSHIPS

- distance matrix: Cavalli-Sforza and Edwards' chord distance
- tree algorithm: Neighbour-joining
- unrooted tree




- seven cultivated/naturalized populations grouped separately from the rest and formed a well-supported clade (bootstrap support 99%) suggesting the common origin of the cultivated plant material

# PLAN

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

(2) Genetic diversity and relationships

(3) Genetic structure

(4) Ecological niche modelling

(5) Demographic history

# 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

# PROCEDURE

## (1) Estimation of the number of Ks

populations of origin

(1.1) Choose different values of K

(1.2) Compute posterior probabilities for each K

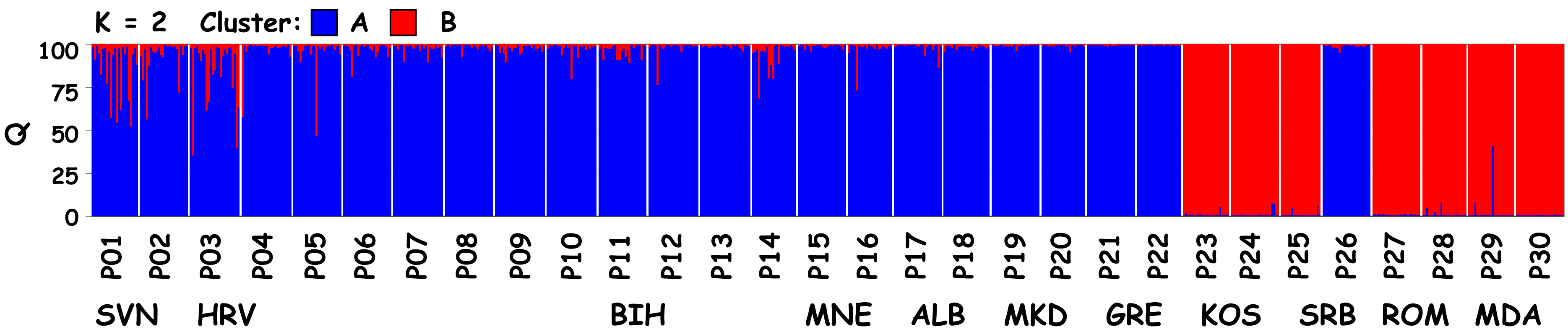
(1.3) Choose the most likely number of populations (K)

## (2) Assignment

(2.1) Assign individuals to populations (K)

(2.2) For each individual, estimate the proportion of genome derived from the different population of origin

# GENETIC STRUCTURE AT K = 2



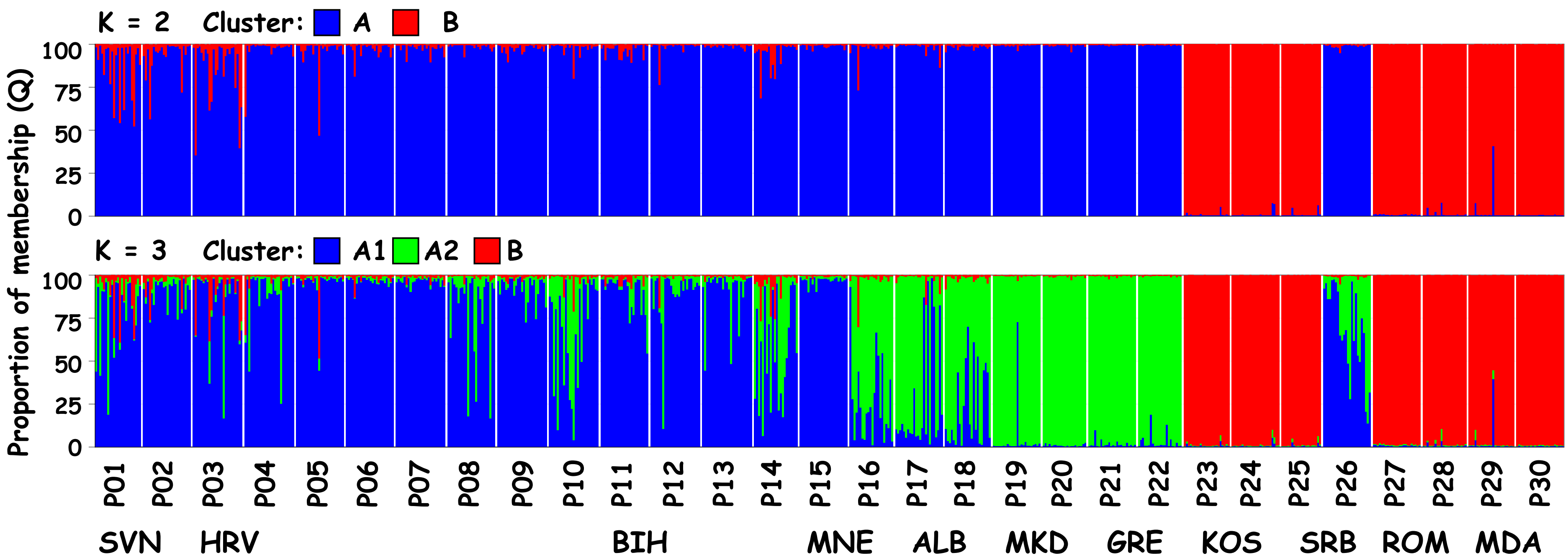

**A**
**Wild populations**

**P26**
**Sićevo Gorge**
**Serbia**

**B**
**Cultivated populations**

- proportions of membership (Q) of each individual plant in each of the two clusters (population of origin):
  - each individual plant is represented by a single vertical line divided into colors representing different clusters
  - the length of the colored segment shows the individual's estimated proportion of membership in that cluster

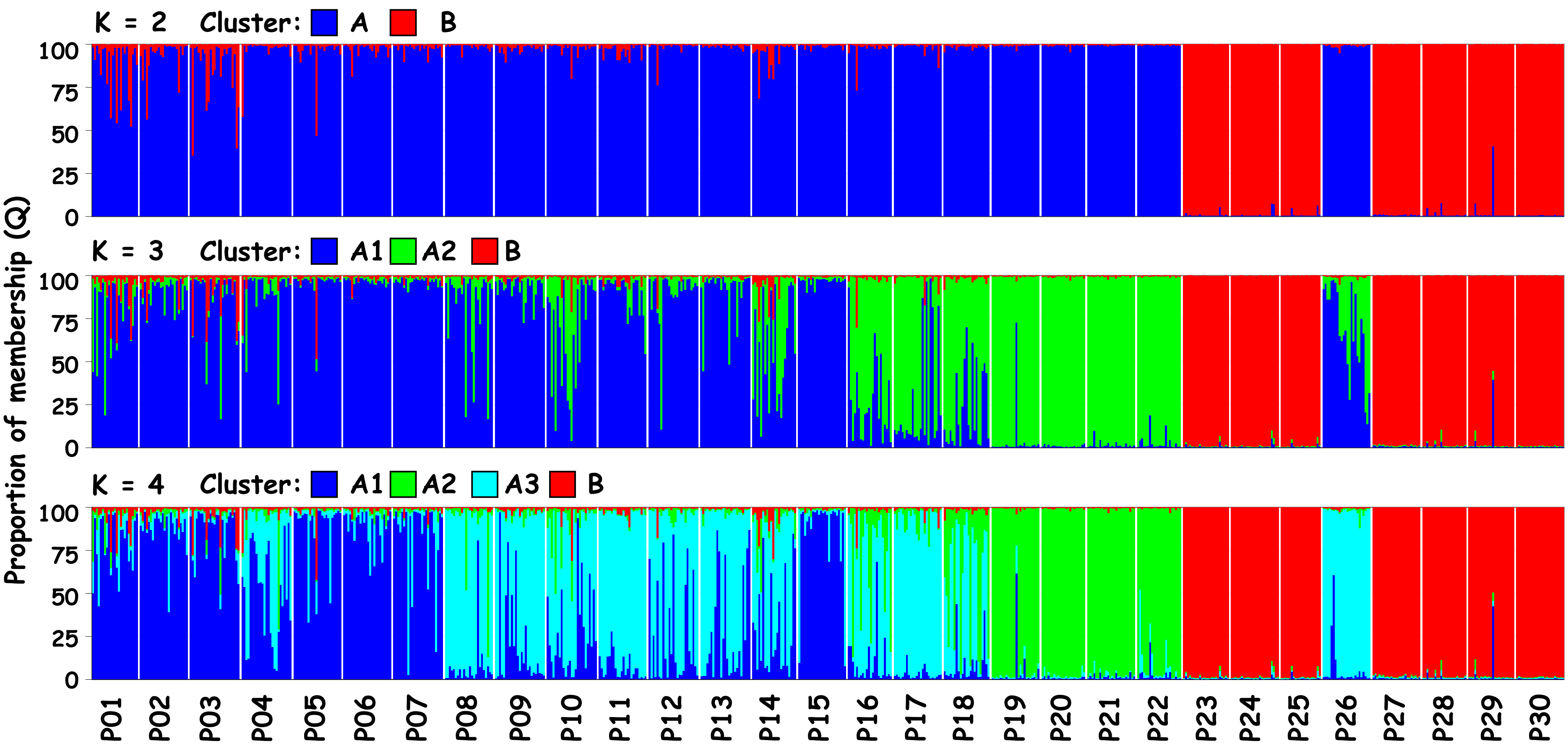
# GENETIC STRUCTURE AT K = 3




| A1 | A2 | B |
|---|---|---|
| North-West populations | South-East populations | Cultivated populations |

- the cluster A (Wild populations) is at K = 3 split into two clusters according to geographical locations
  A1 North-West: Adriatic populations
  A2 South-East: Montenegro-Albania-Macedonia-Greece

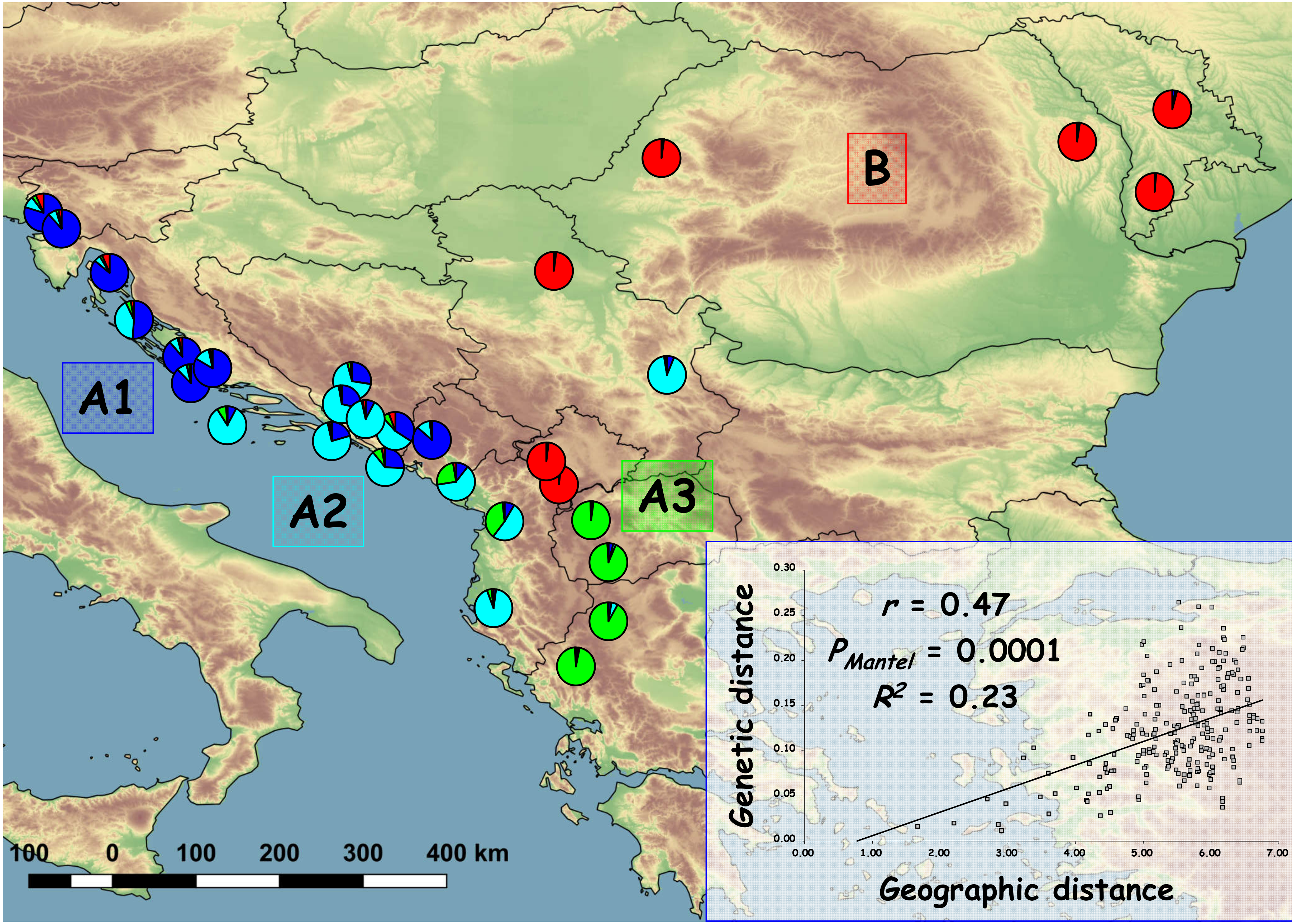
# GENETIC STRUCTURE AT K = 4




| A1 | A2 | A3 | B |
|---|---|---|---|
| Nothern Adriatic | Southern Adriatic | Macedonia Greece | Cultivated populations |

- wild populations split further into three clusters

# GENETIC STRUCTURE AT K = 4



# PLAN

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

(2) Genetic diversity and relationships

(3) Genetic structure

(4) Ecological niche modelling

(5) Demographic history

# 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

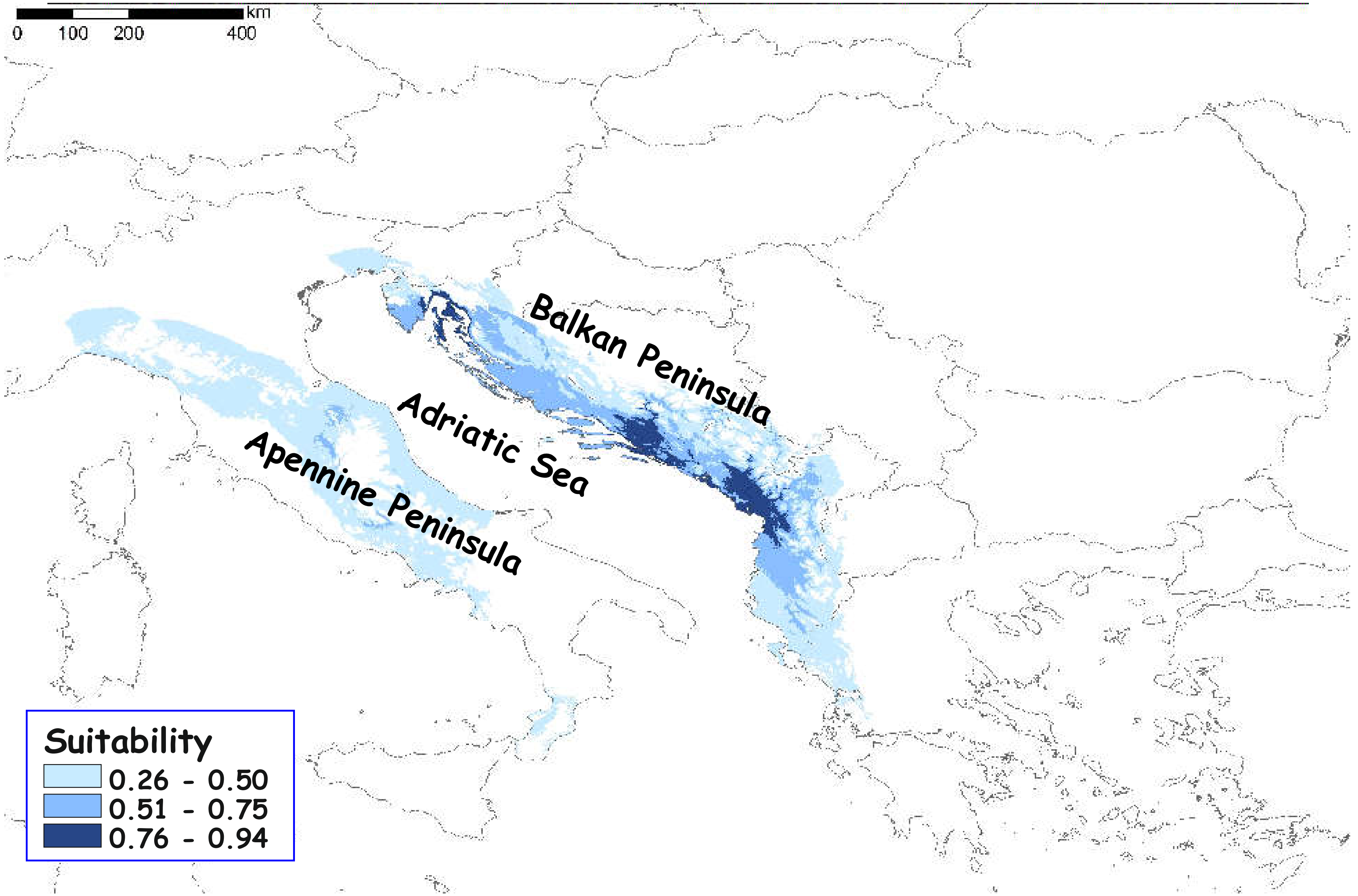
  68 data on occurrence (evenly distributed)

  (2) environmental characteristics of the sites

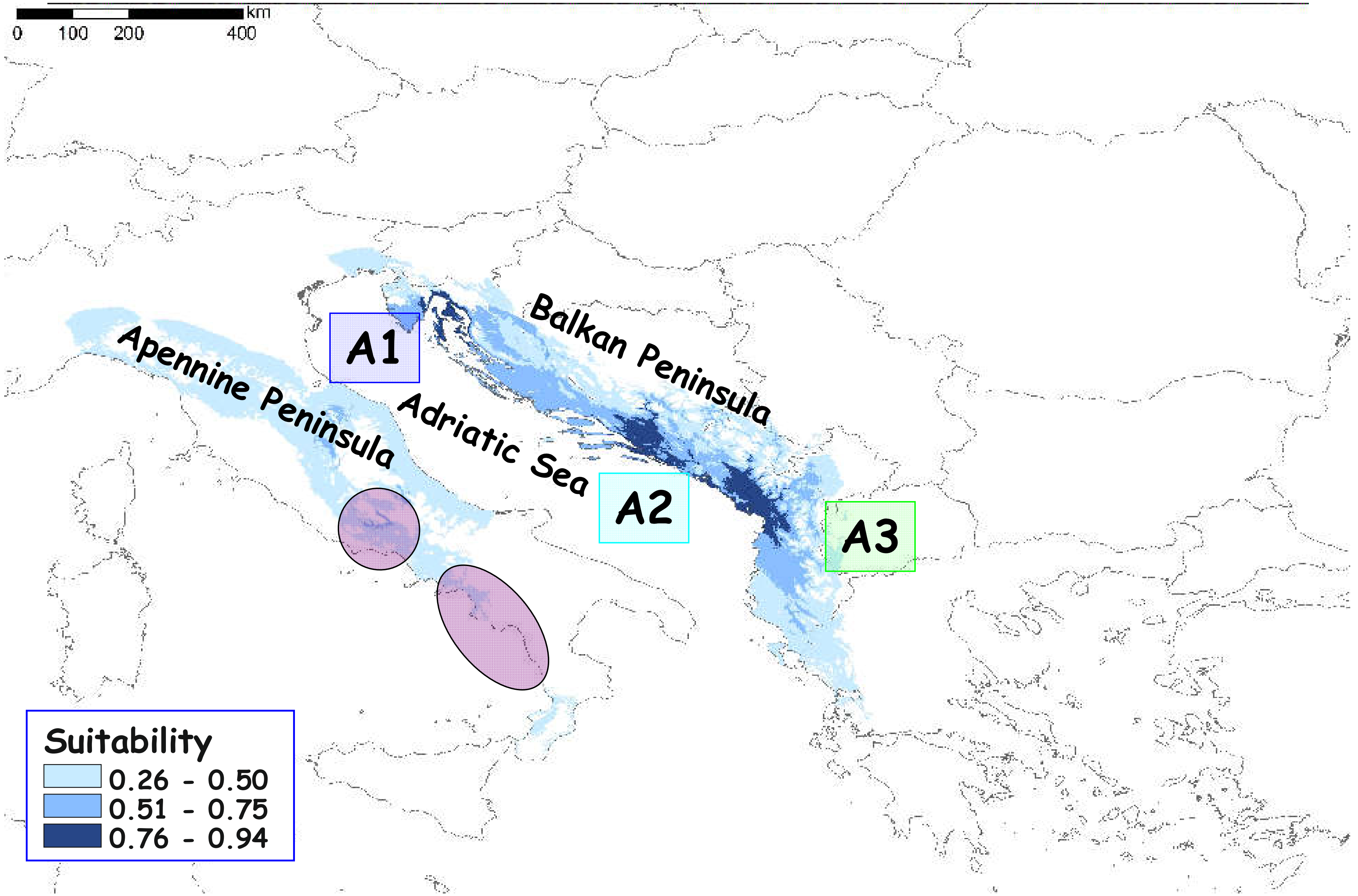
  19 bioclimatic variables (WorldClim database)

  - 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



# PAST DISTRIBUTION

- to model potential species distribution during the Last Glacial Maximum (LGM; ~21,000 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

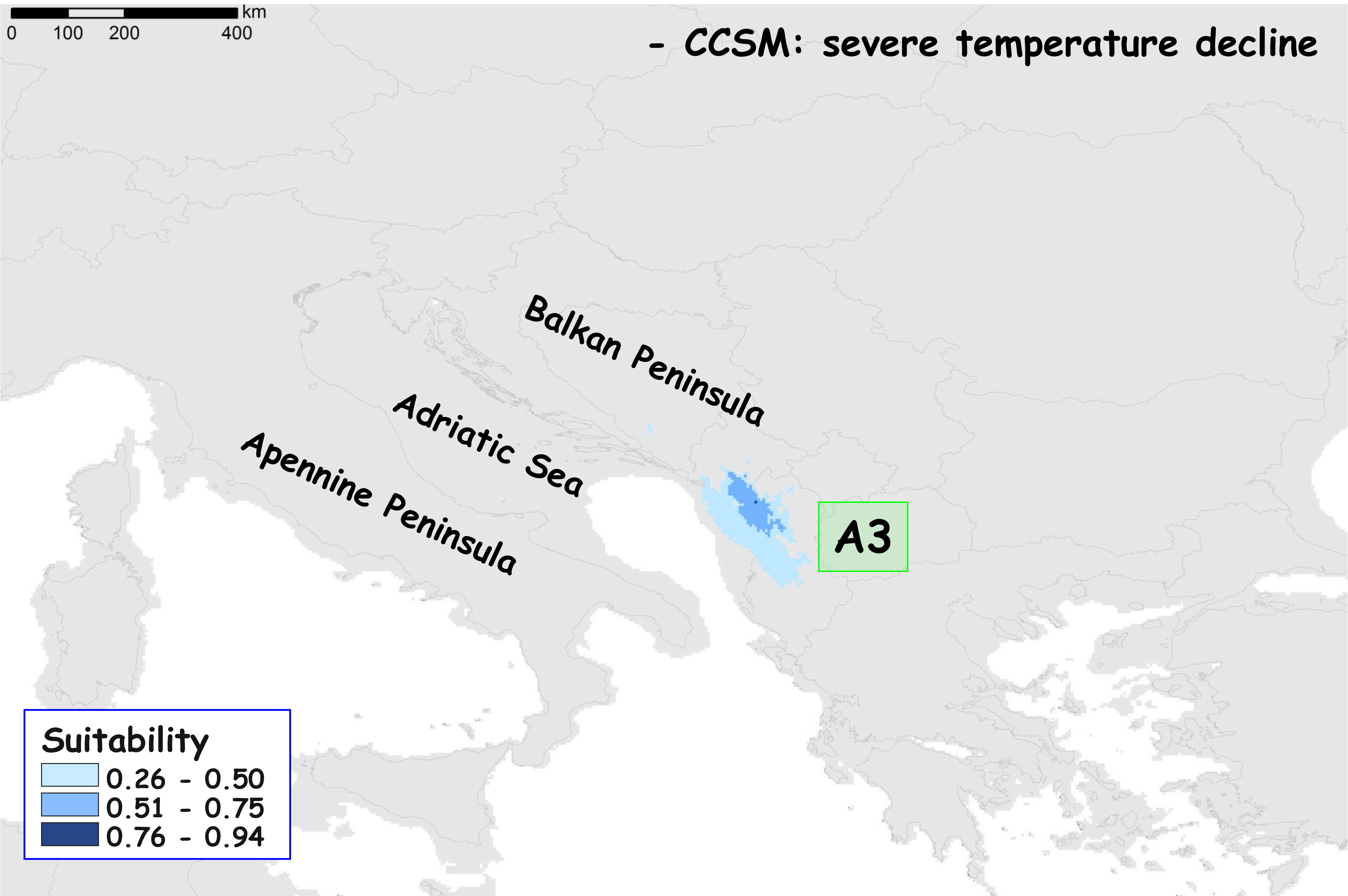
  - two models:

  CCSM (Community Climate System Model; USA)

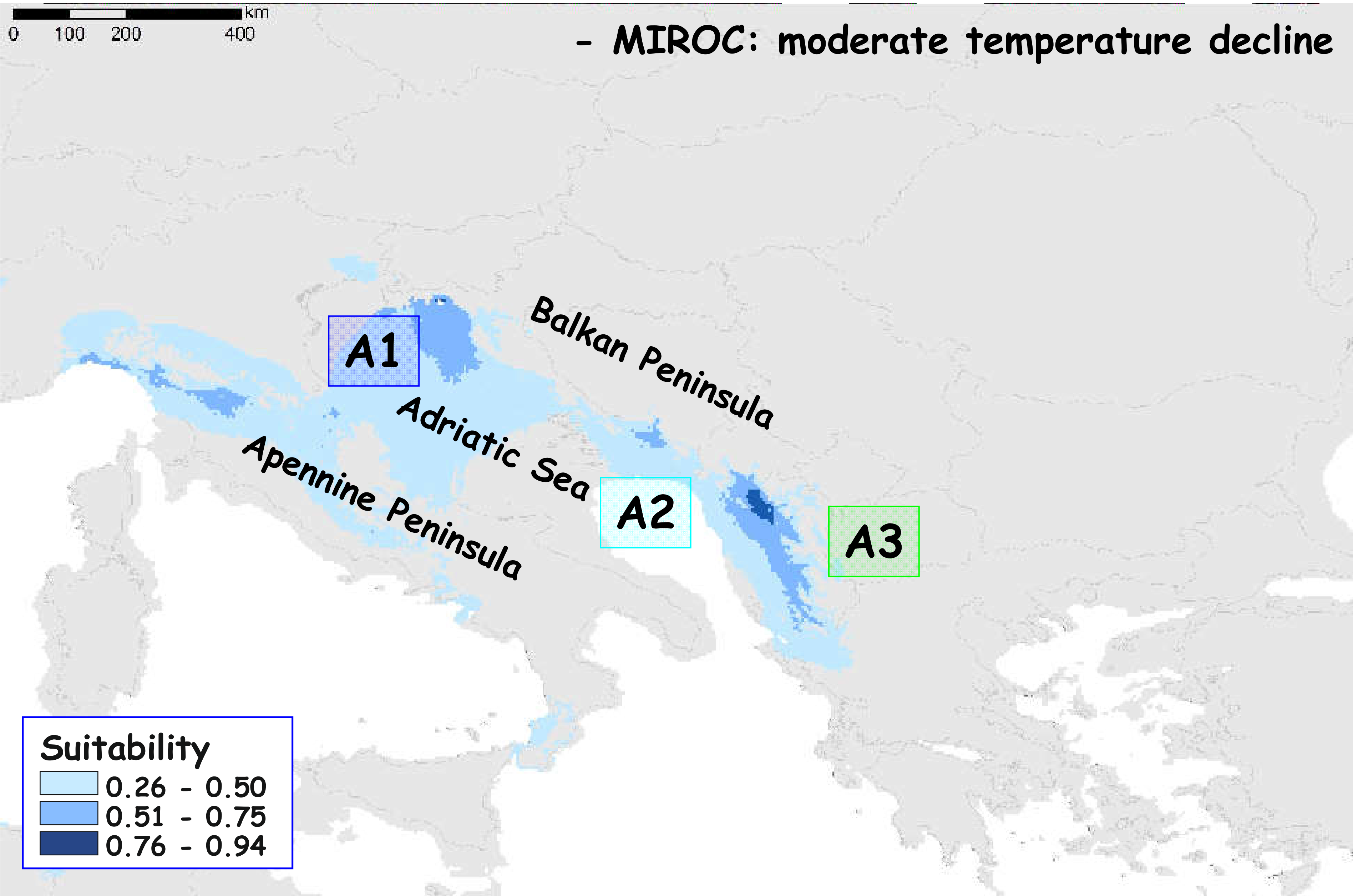
  MIROC (Model for Interdisciplinary Research on Climate; Japan)

# ENM: LAST GLACIAL MAXIMUM CCSM CONDITIONS

- CCSM: severe temperature decline



# ENM: LAST GLACIAL MAXIMUM MIROC CONDITIONS



# PLAN

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

(2) Genetic diversity and relationships

(3) Genetic structure

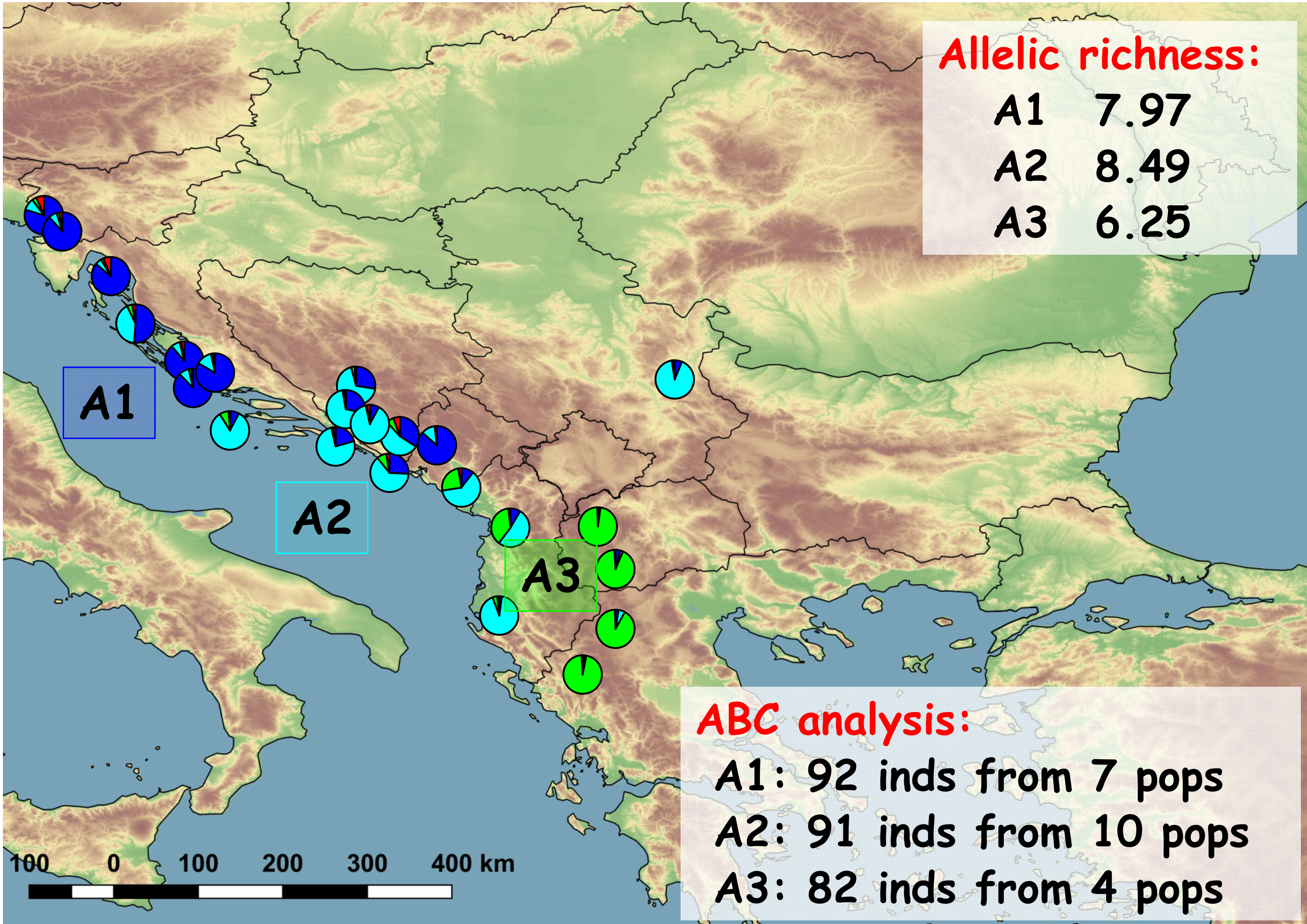
(4) Ecological niche modelling

(5) Demographic history

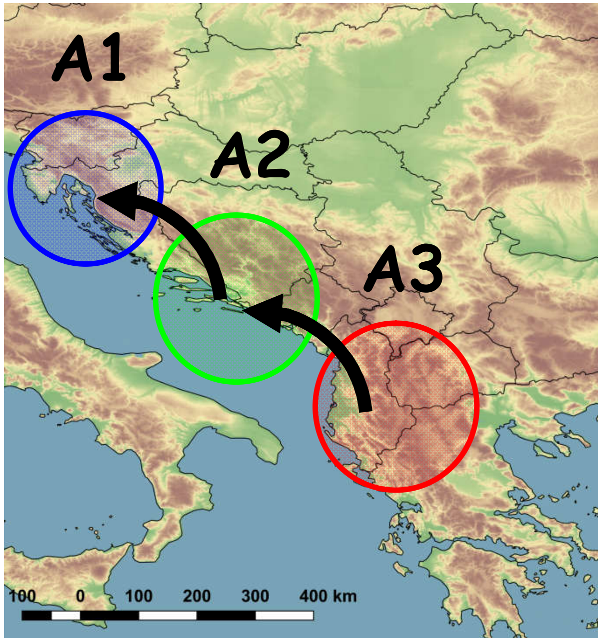
# DEMOGRAPHIC HISTORY

- 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

# POPULATIONS

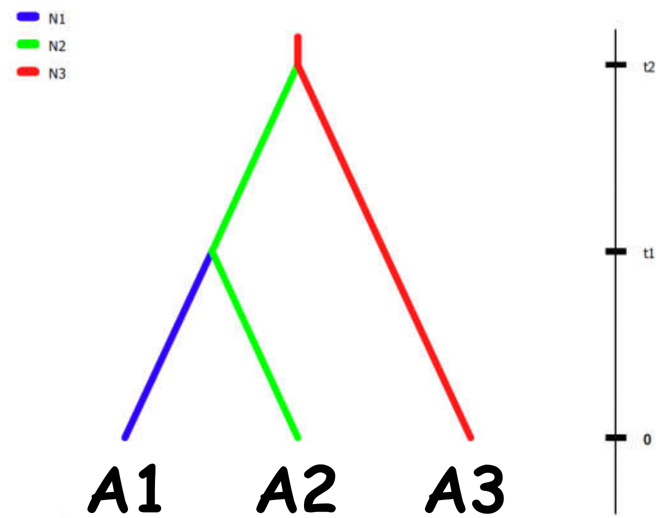


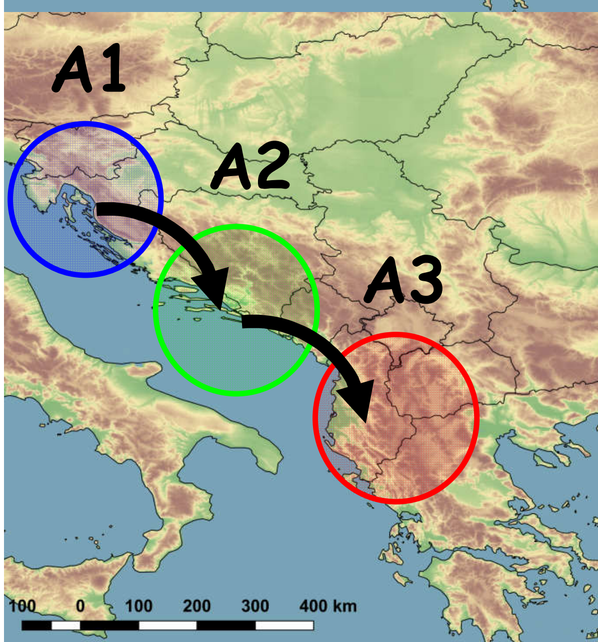
# FIVE SIMPLE HISTORIC SCENARIOS




**Scenario 1**

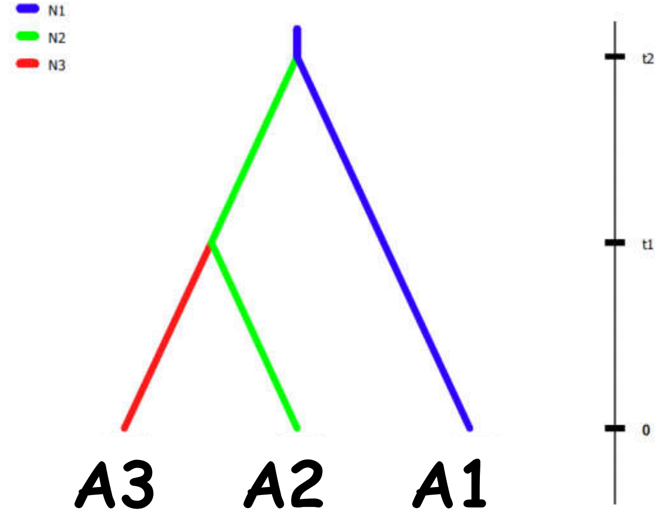
Population A1 is derived from population A2, itself derived from population A3



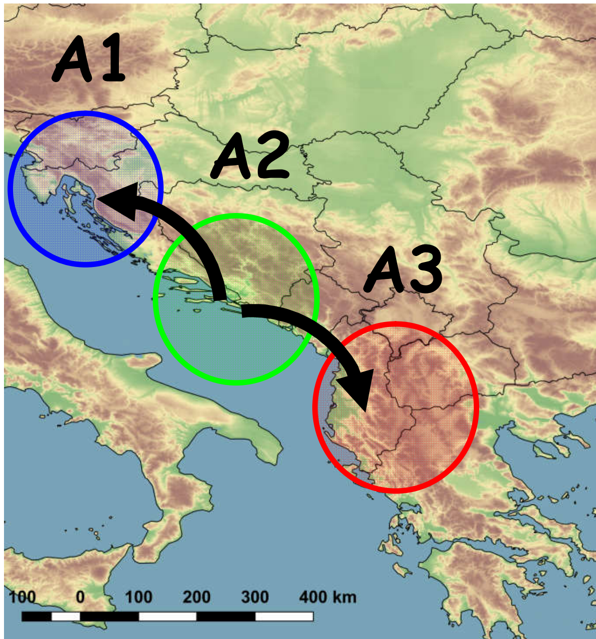




**Scenario 2**

Population A3 is derived from population A2, itself derived from population A1

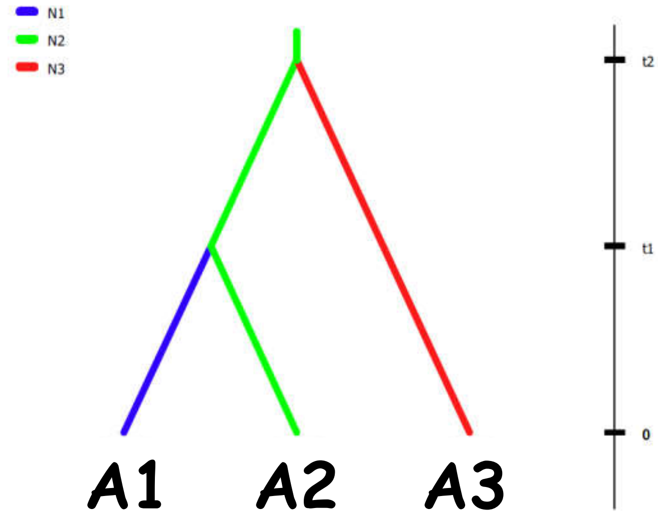


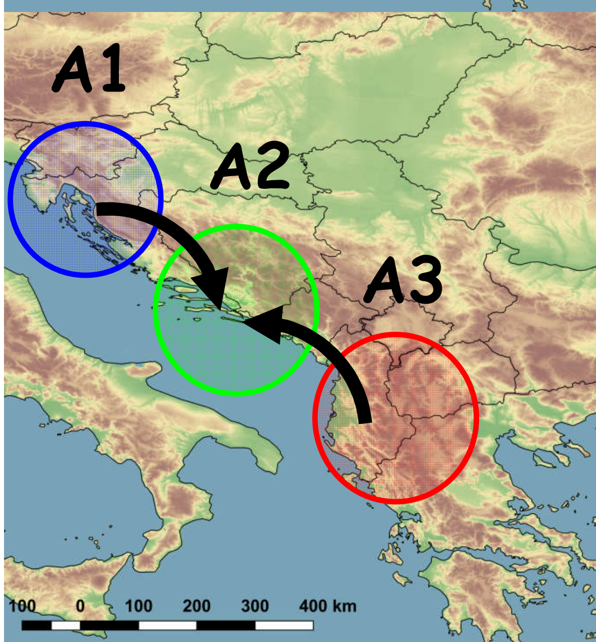
# FIVE SIMPLE HISTORIC SCENARIOS




Scenario 3

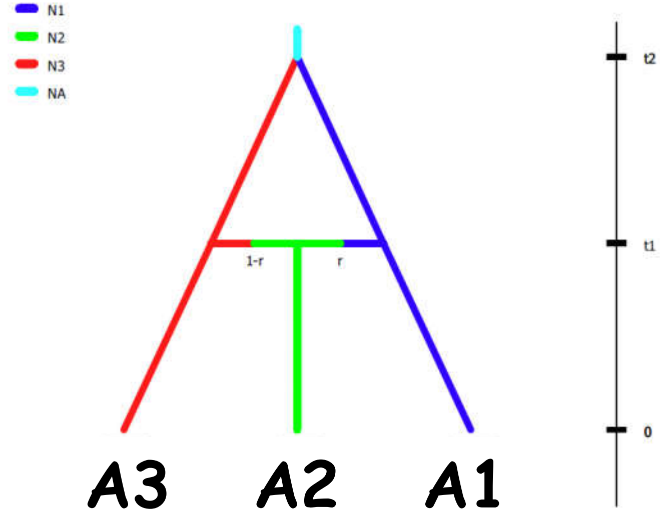
Both populations A1 and A3 derived independently from population A2



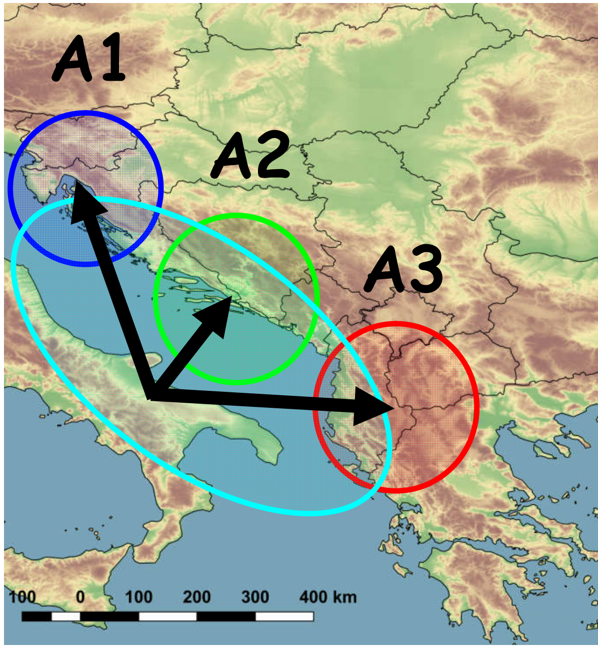




Scenario 4

Population A2 was generated by admixture of populations A1 and A3

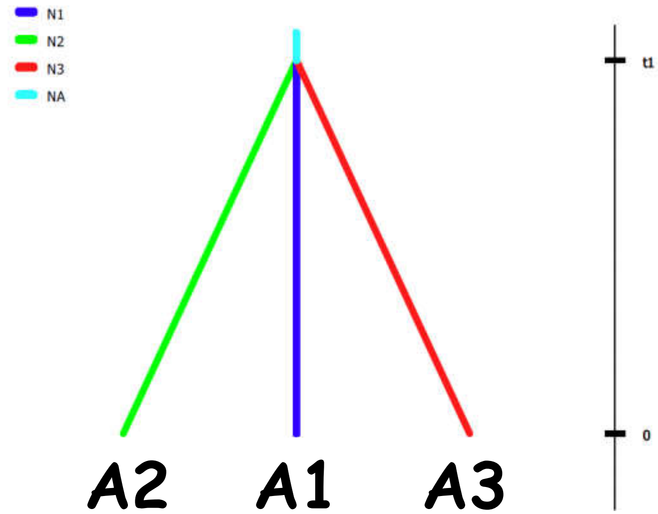


# FIVE SIMPLE HISTORIC SCENARIOS




**Scenario 5**

All three populations diverged at the same time




**RESULTS:**

| Scenario | Posterior probability (PP) |
|---|---|
| 1 | 0.166 |
| 2 | 0.193 |
| 3 | 0.075 |
| 4 | 0.126 |
| 5 | 0.440 |

# HISTORICAL RECONSTRUCTION: ANCIENT PAST

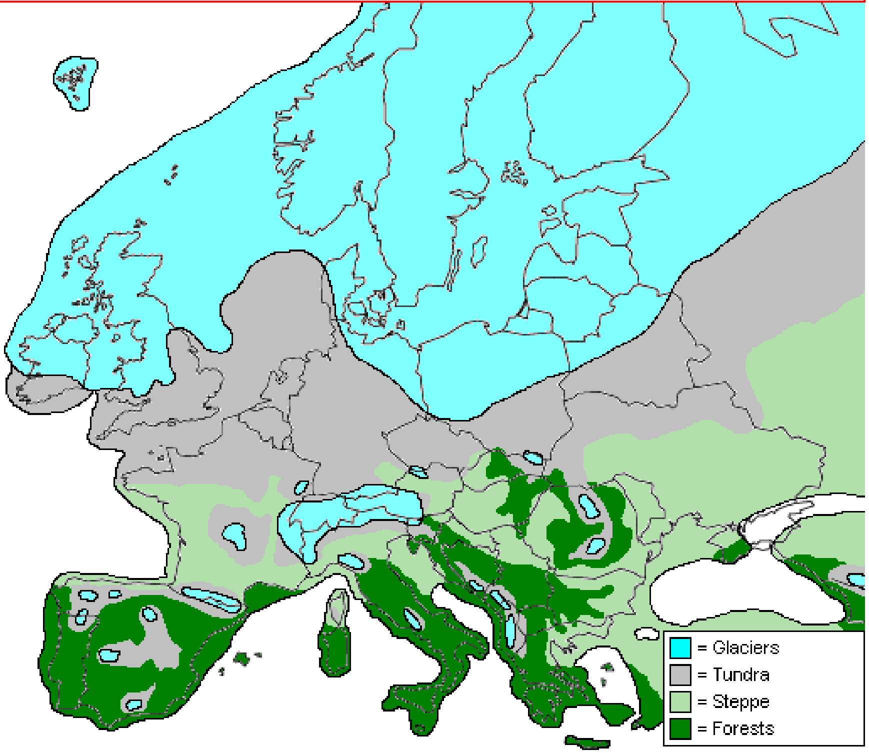

Distibution of Dalmatian sage before the last glaciation ?

ABC Effective population sizes:

| | |
|---|---|
| Ancestral | 8,550 inds |
| A1 | 4,330 inds |
| A2 | 7,190 inds |
| A3 | 2,440 inds |

Vegetation map of Europe during the last glacial period

Source: http://www.metatech.org/07/ice_age_global_warming.html

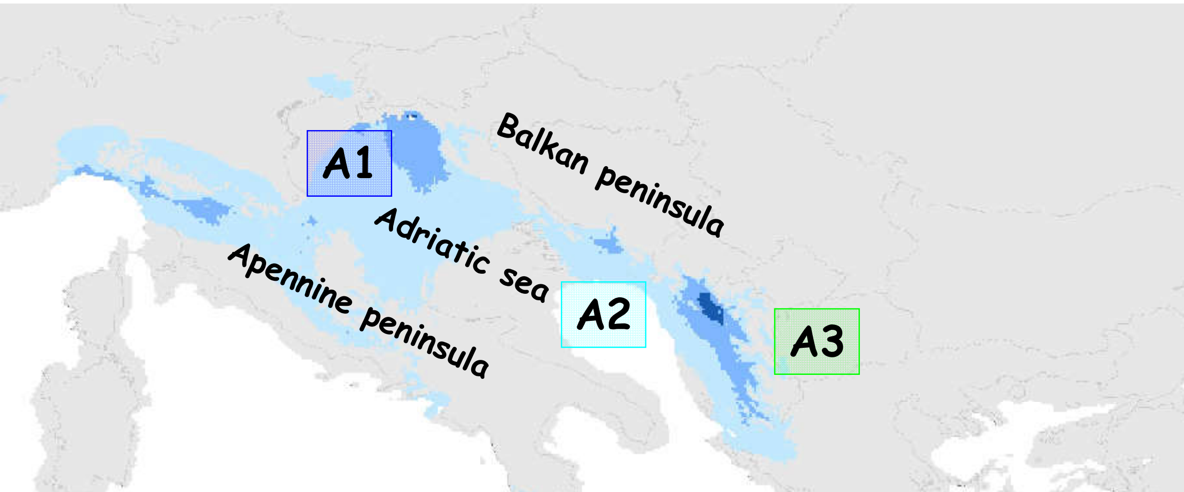


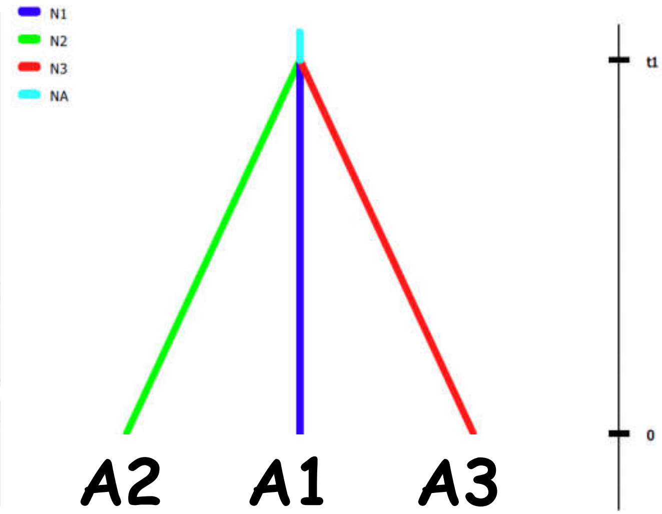

Glacial refugia:

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

# HISTORICAL RECONSTRUCTION: DIVERGENCE

## Distribution of Dalmatian sage during the Last Glacial Maximum





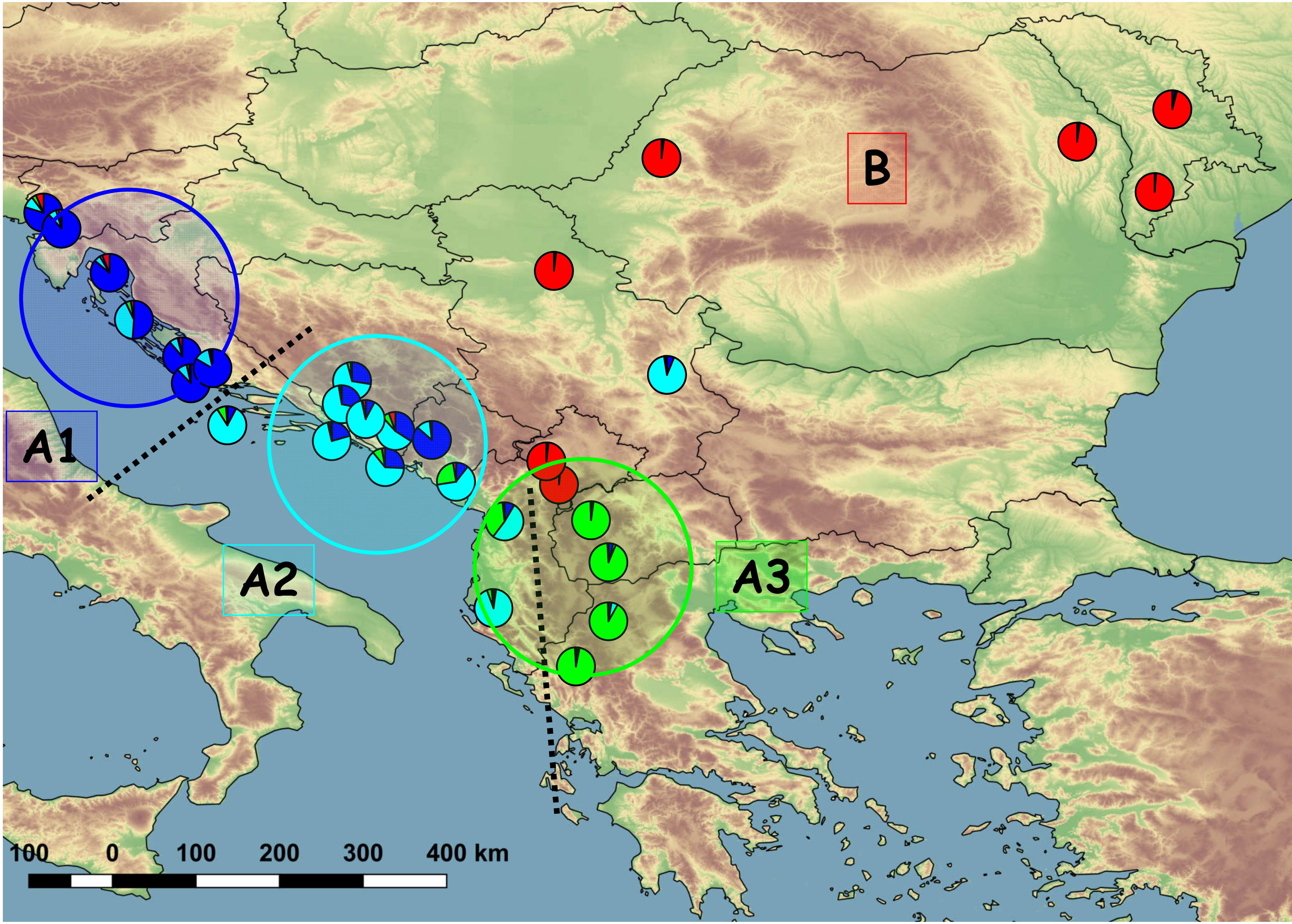
ABC Time of divergence:
572 generations ago
(95% CI: 157-1,540)

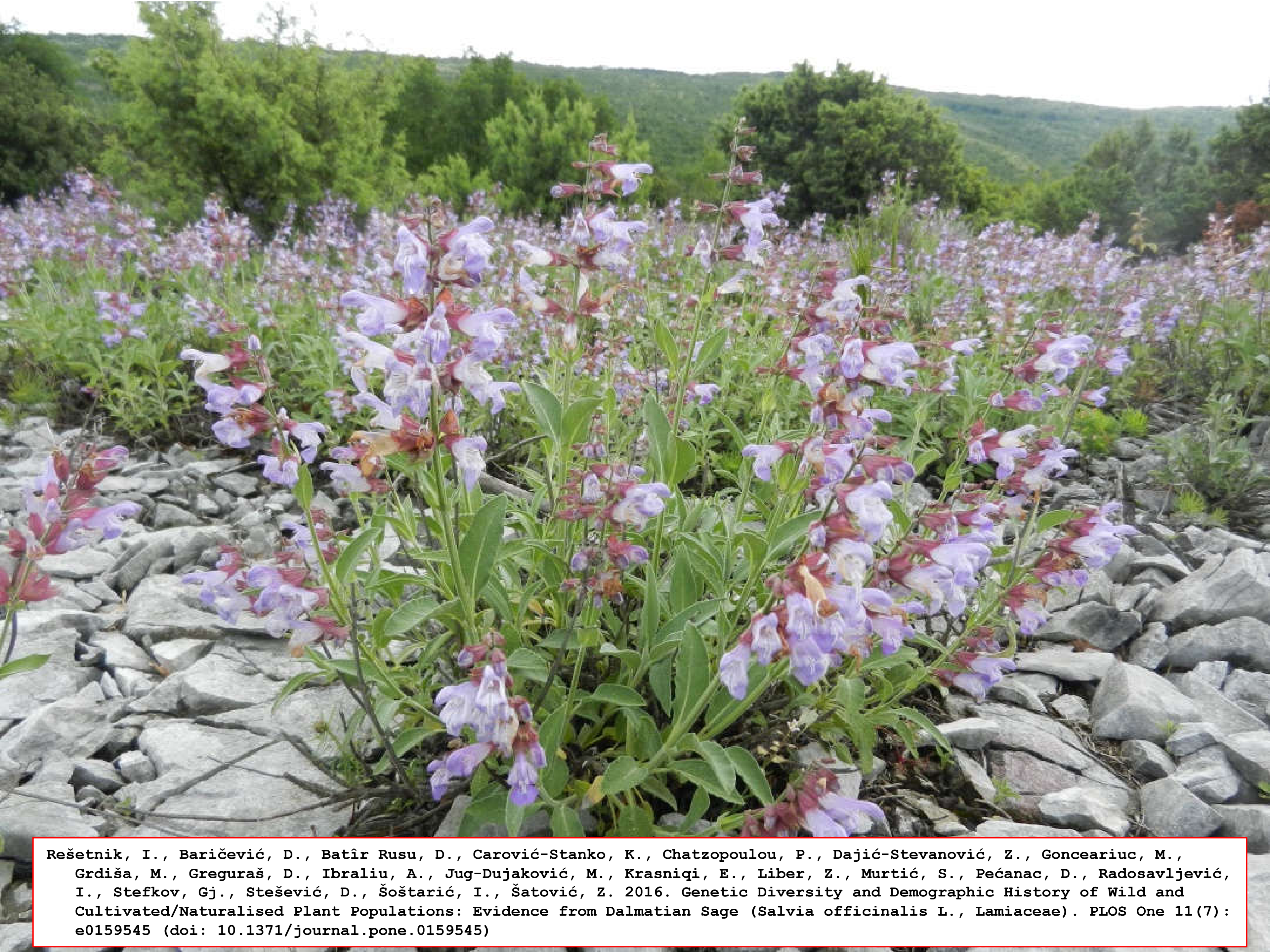
Generation time (life span):
up to 300 years

Map of the 52 putative refugia within the Mediterranean region

Source: http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2699.2008.02051.x/abstract

# HISTORICAL RECONSTRUCTION: PRESENT TIME





Rešetnik, I., Baričević, D., Batîr Rusu, D., Carović-Stanko, K., Chatzopoulou, P., Dajić-Stevanović, Z., Gonceariuc, M., Grdiša, M., Greguraš, D., Ibraliu, A., Jug-Dujaković, M., Krasniqi, E., Liber, Z., Murtić, S., Pećanac, D., Radosavljević, I., Stefkov, Gj., Stešević, D., Šoštarić, I., Šatović, Z. 2016. Genetic Diversity and Demographic History of Wild and Cultivated/Naturalised Plant Populations: Evidence from Dalmatian Sage (Salvia officinalis L., Lamiaceae). PLOS One 11(7): e0159545 (doi: 10.1371/journal.pone.0159545)