

ISTITUTO DI BIOLOGIA
AGROAMBIENTALE E
FORESTALE

Claudia Mattioni

Ilaria Lusini

Marcello Cherubini

Paola Pollegioni

Fiorella Villani

Project

*State and prospects of the *Castanea sativa* population in Belasitsa mountain: climate change adaptation; maintenance of biodiversity and sustainable ecosystem management.*

Funded by: EEA and Norway Grants.

Beneficiary: Forest Research Institute at the Bulgarian Academy of Sciences

Report

*Genetic analysis performed on samples of *Castanea sativa* from Belasitsa and Slavyanka mountains.*

Genetic analysis performed on samples of *Castanea sativa* from Belasitsa and Slavyanka mountains

Claudia Mattioni, Ilaria Lusini, Marcello Cherubini, Paola Pollegioni, Fiorella Villani

Istituto di Biologia Agroambientale e Forestale

Via Marconi 2, 05010 Porano (TR), Italy

Tel. (+39) (0)763 374 911 - 927

Fax: (+39) (0)763 374 980

E-Mail:segreteria@ibaf.cnr.it

claudia.mattioni@ibaf.cnr.it

Recommended citation:

Mattioni C, I Lusini, M Cherubini, P Pollegioni, F Villani. 2011. Genetic analysis performed on samples of *Castanea sativa* from Belasitsa and Slavyanka mountains. In: Zlatanov T, I Velichkov, B Nikolov (eds.) State and prospects of the *Castanea sativa* population in Belasitsa mountain: climate change adaptation; maintenance of biodiversity and sustainable ecosystem management. Project BG 0031 EEA report.

Material and Methods

DNA extraction

Total genomic DNA was isolated by grinding 50-60 mg of leaves tissue in 2 ml microcentrifuge tube contained a 5 mm steel bead. The tissue cooled with liquid nitrogen was homogenized in a Mixer Mill 300 (Quiagen). Genomic DNA was extracted and purified using the DNeasy96 Plant Kit (QIAGEN) according to the manufacturer's instructions.

Microsatellite analysis

A set of eight polymorphic microsatellites developed in *C. sativa* (Buck *et al.*, 2003; Marinoni *et al.*, 2003) was tested (Table 1).

The reaction was performed in 20 μ L total volume containing 20 ng of genomic DNA following the Qiagen multiplex kit protocol (Qiagen, Valencia).

Cycling parameters were as follows: 15 min at 95°C, 30 cycles for 30 s at 94°C, 90 s at 57°C and 1 min at 72°C, and a final step of 30 min at 72°C. Amplification products (0.1–1 µL) were added to 20 µL formamide and 0.3 µL Genescan-500 ROX and denatured at 95°C for 5 min. The samples were run on ABI PRISM 3100 DNA sequencer. Alleles scoring was performed using the GeneScan 3.5 and Genotyper 3.7 softwares (Applied Biosystems).

Data analysis

Genetic diversity of populations

A set of measures of intra and inter population genetic statistics were calculated. The observed (N_a) and effective (N_e) number of alleles, the observed (H_o) and expected (H_e) heterozygosity were calculated using the software Popgene 3.2 (Yeh et al., 1997) and GeneAIEx 6 (Peakall and Smouse, 2005). The within-population inbreeding coefficient F_{is} (Weir and Cockerham 1984) and the statistical significance of F_{is} was tested using a non-parametric approach described in Excoffier et al. (1992) with 1000 permutations using the software Arlequin (Schender et al 2000).

Population structure analysis

The population structure and proportion of membership (Q-value) for each predefined population and each individual sample in each of the predicted clusters were inferred using the Markov Chain Monte Carlo (MCMC) and Bayesian clustering algorithms implemented in STRUCTURE software 2.3.3 (Pritchard et al. 2000).

STRUCTURE analysis was performed using the admixture model on the whole dataset with no previous population information and the correlated allele frequencies between populations option. Based on the initial results, a series of six independent runs were performed for K between 1 and 6 with a burn in period of 10000 steps followed by 10^5 MCMC replicates. Furthermore, the ad-hoc statistic ΔK defined by Evanno et al. (2005)

was used to detect the most likely number of populations. The six runs from the most probable number of clusters were averaged applying *FullSearch* algorithm provided by CLUMPP software 1.1.2 (Jakobsson and Rosenberg 2007). The corresponding Q-matrices were graphically displayed by DISTRUCT software (Rosenberg 2004). Unweighted pair-group clustering was performed with NTSYS 3.2 software (Exeter Software, Setauker, NY USA) using dissimilarity matrix based on Nei's genetic distance.

Results and Discussion

The values of genetic diversity observed in the 4 different groups of tree (Western-Middle-Eastern Belasitsa and Slavyanka mountains) are comparable. Only the Middle and Eastern populations of Belasitsa mountain showed a F_{is} value significantly different from the equilibrium (Tab.2). An higher number of private alleles was observed in the Middle and Eastern populations (Tab. 2).

The PCA analysis (Fig.1) reveals the separation of the samples in the two principal groups: the populations coming from the Belasitsa mountain and the population from Slavyanka. Looking at the results obtained using the software STRUCTURE all the samples appeared clustering in two most probable groups (Blue and Yellow) Fig.2. There was no evidence of a clear geographical separation of the two groups. The membership of each genotypes to the two clusters are represented in a map considering the geographic location of each individuals (Fig.3a, 3b, 3c). No relations were observed between altitude and the different genotypes. To have an indication of the possible origin of the Bulgarian populations and their relationships with the other European populations previously analysed, an unweighted pair-group clustering analysis and population Structure analysis were performed.

As showed in the Fig. 4 the Bulgarian samples are in the same cluster of the Greece samples indicating a congruence between geographical and genetic distance. The analysis

performed using the software STRUCTURE shows that the most probable clusters were $K=2$, $K=3$. For $K=2$ the populations are grouping in two main clusters: one cluster including populations from Turkey, Greece and Bulgaria, the other main cluster includes populations from Italy and Slovakia. For $K=3$ the populations from Bulgarian are more close to the Greece populations confirming the results obtained with the unweighted pair-group analysis.

The observed geographic distribution of the three main gene pools are congruent with the location of the likely refugium areas inferred from palynological data (Krebs et al. 2004). Further genetic analysis on a larger number of Bulgarian populations may give indications on the genetic diversity of the chestnut tree in Bulgaria and possible migration routes of the species from refuge areas.

References

- Buck EJ, Russell K, Hadonou M, James CJ, Blakesley D (2003) Isolation and characterization of polymorphic microsatellites in European chestnut (*Castanea sativa* Mill.). *Mol Ecol Notes* 3: 239-241
- Yeh FC, Yang RC, Boyle TBJ, Ye ZH, Mao JX (1997). Popgene ver 1.32. The user-friendly software for population genetic analysis. *Molecular Biology and Biotechnology*
Center, University of Alberta, Canada
- Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 14:1801-1806
- Krebs P, Conedera M, Pradella M, Torriani D, Felber M, Tinner W (2004) Quaternary refugia of the sweet chestnut (*Castanea sativa* Miller): an extended palynological approach. *Veg Hist Archaeobot* 13:145-160
- Marinoni D, Akkak A, Bounous G, Edwards KJ, Botta R (2003) Development and characterization of microsatellite markers in *Castanea sativa* (Mill.). *Mol Breeding* 11:127-136.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multi-locus genotype data *Genetics* 155:945-959
- Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Mol Ecol Notes* 4: 137-138
- Schneider S, Roessli D, Excoffier L (2000) Arlequin: a software for population genetics data analysis, version 3.1. Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva, Switzerland

Tab 1

Locus	Sequenza del primer (5' → 3')	Motivo ripetuto	Range (pb)	T (°C)	Ne
CsCAT1	F GAGAATGCCCATTTTGCA R GCTCCCTTATGGTCTCG	(TG) ₅ TA(TG) ₂₄	194-223	50°C	12
CsCAT2	F GTAACCTTGAAGCAGTGTGAAC R CGCATCATAGTGAGTGACAG	(AG) ₁₆	207-231	55°C	20
CsCAT3	F CACTATTTTATCATGGACGG R CGAATTGAGAGTTCATACTC	(AG) ₂₀	190-269	50°C	14
CsCAT6	F AGTGCTCGTGGTCAGTGAG R CAACTCTGCATGATAAC	(AC) ₂₄ AT(AC) ₄	161-197	50°C	19
CsCAT14	F CGAGGTTGTTGTTTCATCATTAC R GATCTCAAGTCAAAAGGTGTC	(CA) ₂₂	130-164	58°C	7
CsCAT16	F CTCCTTGACTTTGAAGTTGC R CTGATCGAGAGTAATAAAG	(TC) ₂₀	127-157	50°C	12
EMCs25	FATGGGAAAATGGGTAAAGCAGTAA RAACCGGAGATAGGATTGAACAGAA	(GA) ₁₂	127-157	58°C	8
EMCs38	F TTTCCCTATTTCTAGTTTGTGATG R ATGGCGCTTTGGATGAAC	(AG) ₃₁	228-270	56°C	17

Tab 2

Popolazioni	Na	Ne	I	Ho	He	UHe	Fis ^a	Probability Identity PIunb PIsib	Private alleles
West	5,125 (0,549)	3,844 (0,435)	1,428 (0,102)	0,750 (0,120)	0,719 (0,028)	0,775 (0,030)	0,034	3,2x10 ⁻⁰⁸ 9,6x10 ⁻⁰⁴	0,375 (0,183)
Middle	11,125 (0,990)	4,669 (0,405)	1,756 (0,076)	0,690 (0,077)	0,773 (0,022)	0,776 (0,022)	0,110 [*]	1,2x10 ⁻⁰⁹ 4,6x10 ⁻⁰⁴	1,500 (0,423)
Est	11,125 (1,432)	4,744 (0,454)	1,767 (0,127)	0,692 (0,081)	0,761 (0,043)	0,766 (0,043)	0,096 [*]	1,2x10 ⁻⁰⁹ 5,0x10 ⁻⁰⁴	1,500 (0,655)
Slavyanka	4,625 (0,498)	3,376 (0,336)	1,274 (0,142)	0,625 (0,113)	0,658 (0,063)	0,692 (0,066)	0,101	2,5x10 ⁻⁰⁷ 1,9x10 ⁻⁰³	0,625 (0,183)

Fig. 1 PCA analysis shows a separation between samples of the two mountainous areas

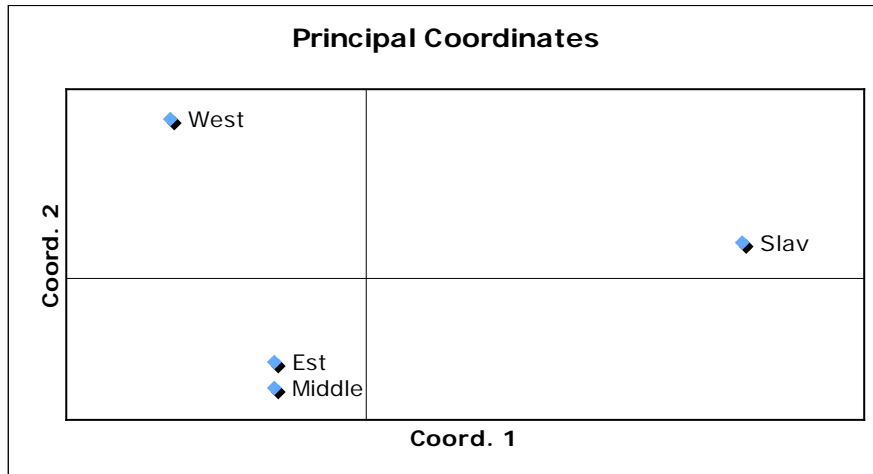


Fig.2 The analysis carried out with the software STRUCTURE indicates the presence of two main groups (Yellow and Blue)

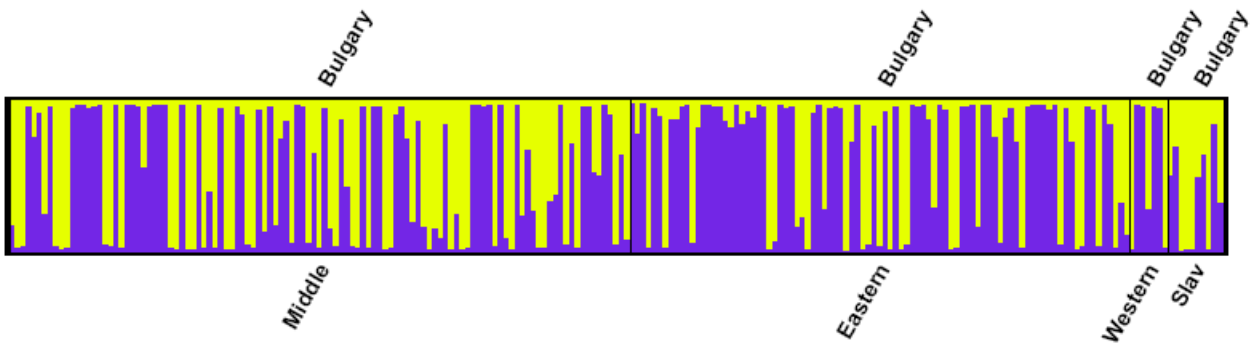
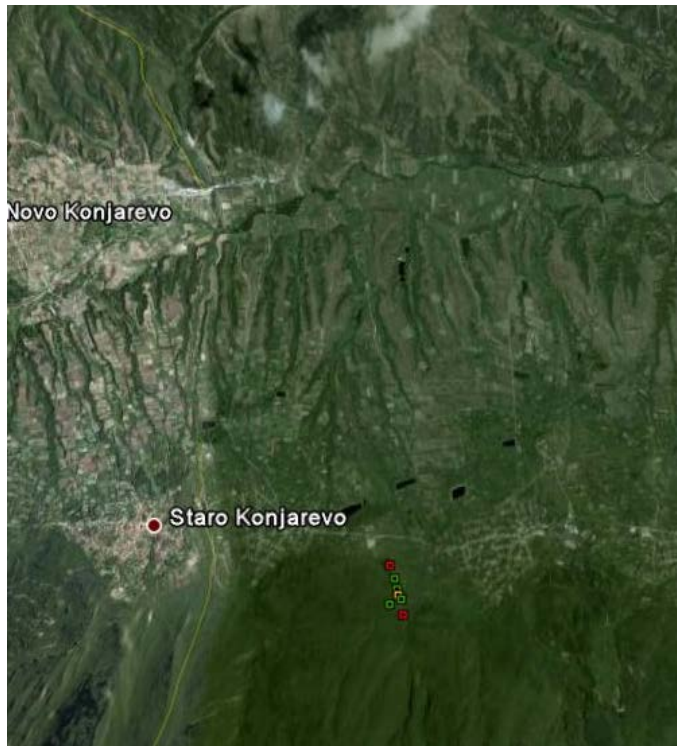


Fig.3a



Geographic representation of the genotypes

- Red=genotype 1
- Green = genotype 2
- Orange =genotype hybrid

Fig.3b

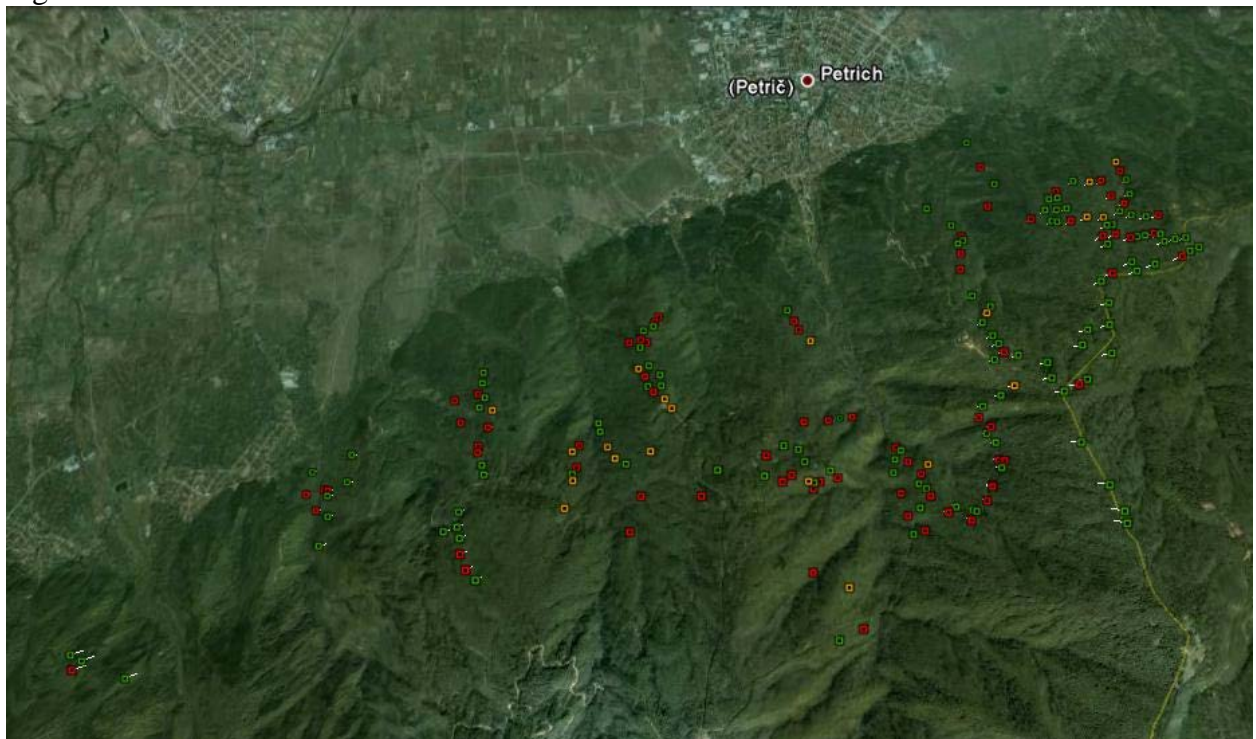


Fig.3c

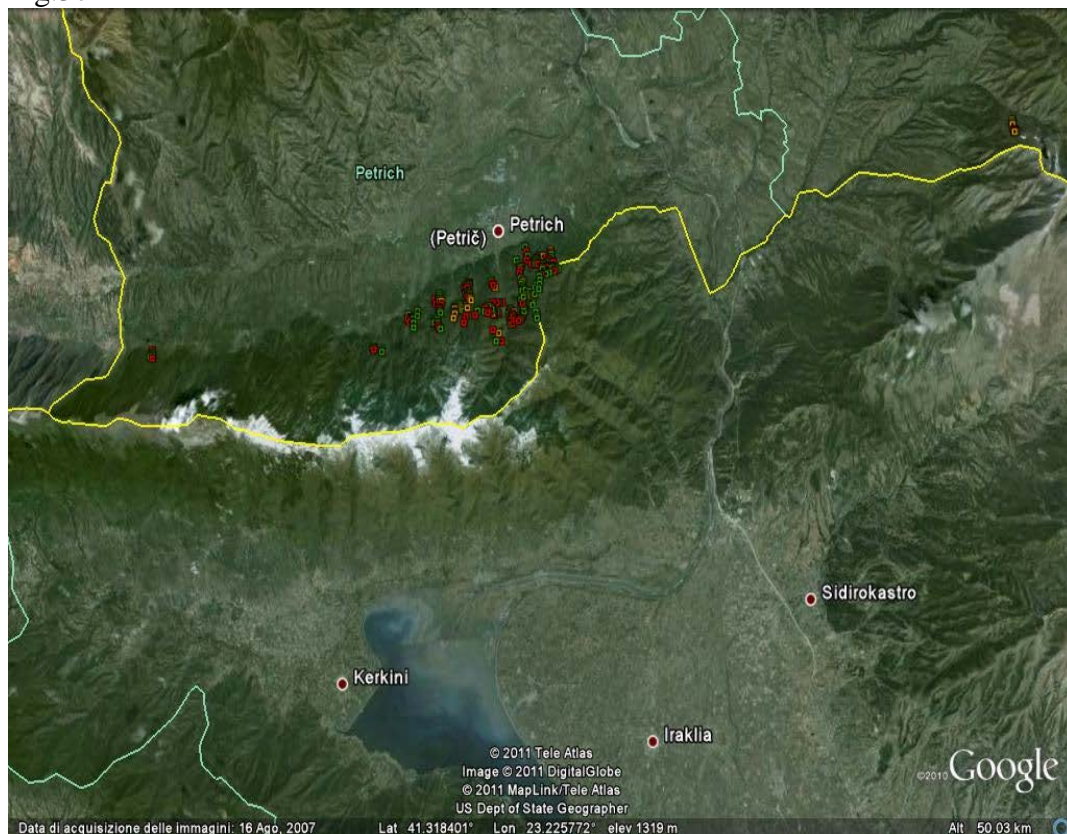


Fig.4 The samples coming from Bulgarian are analyzed with samples from European population. The Bulgarian samples are in the same cluster with the Greece populations

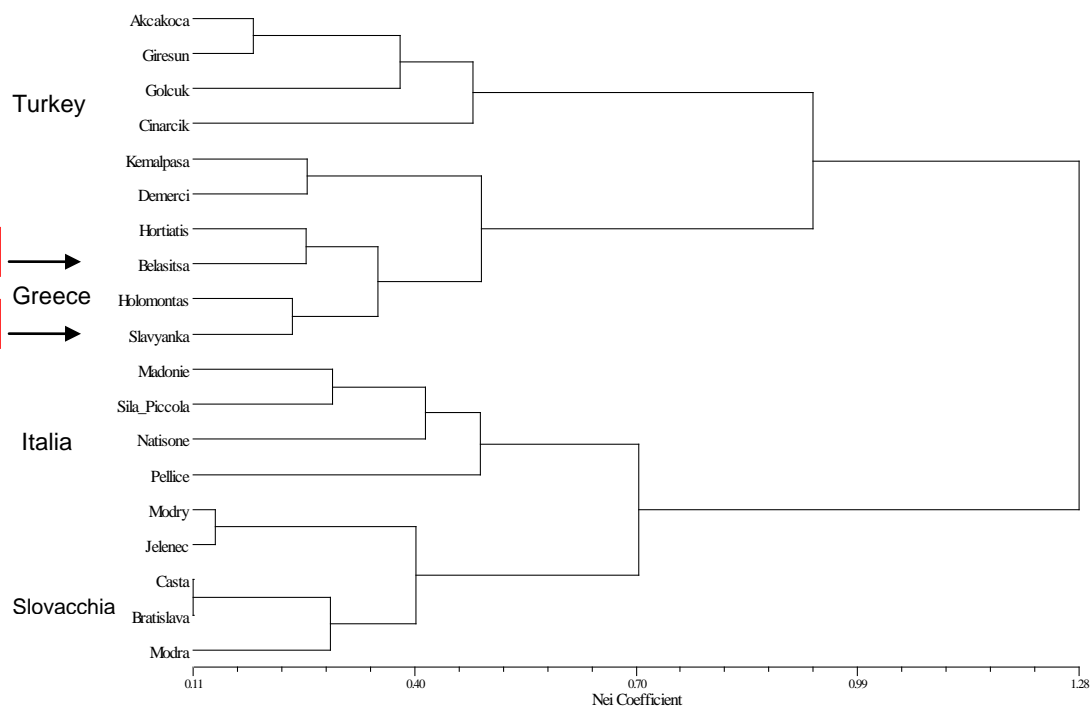


Fig.5 Structure of European populations Using the software STRUCTURE the main probable cluster found is K=2. The Bulgarian population appeared to have the same genotypes of the Turkish and Greek populations

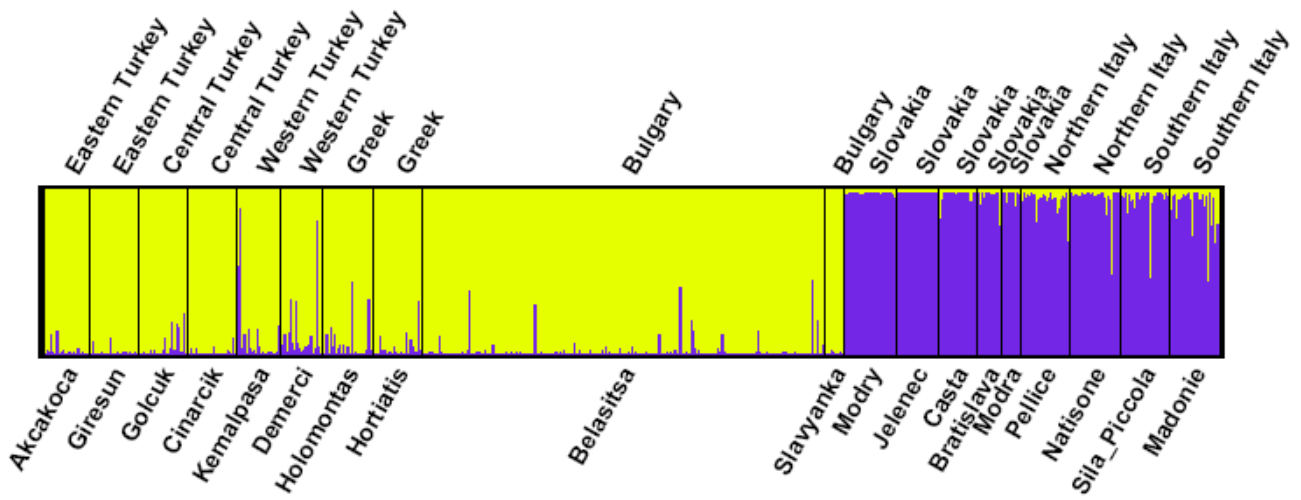


Fig.6 The other most probable grouping K=3. Bulgarian populations are more close to the Greek populations (red clusters)

