Short Communication

Genetic variability and population structure of chamois in Greece (*Rupicapra rupicapra balcanica*)

Haritakis Papaioannou[2], Margarita Fernández[1], Trinidad Pérez[1], Ana Domínguez[1]

[1] Departamento de Biología Funcional (Genética), Facultad de Medicina 6ª Planta, Universidad de Oviedo, C/Julián Clavería 6, 33006 Oviedo, Spain

[2] Balkan Chamois Society, 44004 Papingo Ioannina, Greece

Corresponding author: Ana Domínguez

e-mail: sanjurjo@uniovi.es

telephone (34) 985 102 719

fax (34) 985 102 535

Abstract

Balkan chamois (*Rupicapra rupicapra balcanica*) is the southernmost subspecies within the distribution of the genus in Europe. In Greece, which is its marginal area of distribution, the population presents a fragmented pattern. This is the first study that investigates genetic variability and structure of Greek chamois. We collected samples from the wider Pindus mountain range, Mount Olympus, the Rhodope mountains and from the North-Northwestern mountains. Individuals were screened for mitochondrial (mt) sequences, cytochrome b (*cytb*) and control region (*CR*), and 18 microsatellite loci. Only one haplotype of *cytb* was observed. Sequences of the *CR* showed extensive variability grouping into three differentiated clades, one of them including specimens of the subspecies *asiatica* and *caucasica*. The GenBank haplotypes of *balcanica* from the Dinarides form a different clade. There is differentiation among geographical areas both for the *CR* as well as for microsatellites. In particular, the Olympus population is clearly distinct from the rest and shows low diversity. This differentiation can be related to recent isolation and small population size more than to a singular long evolutionary history, given that the haplotypes present there are shared by the Pindus populations. The chamois in Greece harbor an outstanding amount of variability within the species *R. rupicapra* and hence merit the implementation of special conservation measures. We propose actions to prevent further fragmentation in the wider area of Pindus and the North-Northwestern mountains. For the isolated populations of Olympus and the Rhodopes, conservation must focus on actions to maintain a viable population size.

Keywords *Rupicapra rupicapra balcanica*, microsatellites, mtDNA, population structure, Balkans, glaciations, conservation

Introduction

Balkan chamois (*Rupicapra rupicapra balcanica*) is the southernmost subspecies within the distribution of the genus in Europe. Its geographical distribution extends across nine countries in the Balkan Peninsula, forming usually small and often isolated populations with different conservation and management statuses (Corlatti et al. 2011). The population of chamois in Greece is on the edge within the subspecies range. Populations occupying marginal areas tend to be fragmented and are susceptible to the effects of genetic drift and inbreeding (Frankham et al. 2010). In Greece, after a decreasing trend that lasted until the year 2000 with 477-750 individuals across the whole Greek mainland, the total population of chamois is now increasing and counts around 1500 individuals (Papaionnou 2015, 2016). Hunting has been officially forbidden since 1969. Reintroductions have never been attempted in Greece, the increase in population size can be attributed to the implementation of conservation measures. The chamois is now a strictly protected species and most of its range is situated within the borders of protected areas (Natura 2000 sites and National Parks). The chamois distribution is fragmented, with 24 subpopulations all over Greece (Figure 1). These populations were grouped in six main blocks, with no confirmed animal movement between them (Papaioannou 2015, 2016). Three blocks in the Pindus mountain range (A+B+C), Olympus (D) Rhodope mountain range (E) and North-Northwestern mountains (F). One of the most robust populations, but also one of the most isolated, is the population in Mount Olympus, surrounded by extensive agricultural lowlands. This population was even suggested to be a separate subspecies under the name of *Rupicapra rupicapra olympica*, on the basis of its distinct skull parameters (Koller 1929).

The subspecies *balcanica* was included in several phylogenetic studies (Pérez et al. 2002; Rodríguez et al. 2009; Rodríguez et al. 2010; Pérez et al. 2017). In addition, two studies deal with the populations of *balcanica* inhabiting the Dinarides (Šprem and Buzam 2016) and the western Rhodope in Bulgaria (Markov et al. 2016). However, the status of the subspecies in the periphery of its range, where it is prone to the effects of drift and inbreeding, has not been addressed. Hence, we investigated the genetic variability and population structure of chamois in Greece in order to obtain data relevant to its conservation.

Methods

For this study, we collected 76 new samples. Tissue samples were collected over a long period (1994-2016). In addition, we carried out foot surveys (11/2016) to collect fresh droppings from Mount Olympus and from the Rhodope mountain range. Droppings were preserved in sterilized 15ml Falcon tubes with silica gel. Only 53 out of the 76 new samples yielded DNA.

After the comparison of the microsatellite profiles of these 53 DNA samples, we identified 48 specimens (several profiles corresponded to repeated sampling of four individuals). We added to the dataset fourteen specimens included in previous studies to attain a total of 62 individuals (see Table S1).

Due to the limited number of samples, individuals were classified into four population groups for the analysis, one including the samples of blocks A, B and C (wider area of Pindus mountain range, with similar habitat) and the other three as previously defined.

To isolate DNA we used different methods, depending on the nature of the samples (Pérez et al. 2002). For hairs and stool, we used chelex extraction (using 5-10 rooted hairs) and QIAamp DNA Stool Mini Kit (Qiagen, Hilden, Germany), respectively.

We sequenced two mitochondrial regions (*CR* and *cytb*) and typed eighteen microsatellite loci (SR-CRSP-1, SR-CRSP-3, SR-CRSP-4, SR-CRSP-5, SR-CRSP-6, SR-CRSP-8, SR-CRSP-9, SR-CRSP-11, SR-CRSP-12, SR-CRSP-13, SR-CRSP-15, ETH10, ETH225, INRA003, INRA005, INRA011, INRA036, INRA063). Amplification, sequencing and genotyping were as described (Rodríguez et al. 2010). For stool samples, amplifications were performed independently for each marker in a 10 µl volume reaction containing 1 µl template DNA, 0.5 µM of each primer and 5 µl of ‘‘Qiagen Multiplex PCR Kit’’ (Qiagen, Hilden, Germany). Following suggestions for limiting genotyping errors (Bonin et al. 2004), we performed three PCR repetitions per sample and obtained a consensus genotype using GIMLET 1.3.3 (Valière 2002). Errors were almost exclusively dropout and the error rate was 2.45%. Genotype reliability was checked with RELIOTYPE (Miller et al. 2002) and was 94.31% on average.

Microsatellite data were arranged in a matrix of 18 loci per 52 individuals classified into the four population groups previously indicated. We tested linkage disequilibrium (LD) with Genepop 4.2 (Raymond and Rousset 1995). We used MICRO-CHECKER (Van Oosterhout et al. 2004) to check for the presence of null alleles. Basic descriptive statistics were obtained with GenAlex 6.5 (Peakall and Smouse 2012). Allele richness and pairwise Fst values were estimated with FSTAT (Goudet 1995). P-

values of the Fst estimates are based on 6000 permutations, although it must be taken into account that SE for divergence can be biased when sample size is very small.

The clustering of individuals into groups was investigated both with a Principal Coordinate Analysis (PCoA) and with the software STRUCTURE. PCoA, based in the codominant genotypic distances between individuals, was performed with GenALEx 6.5. The three-dimensional output was represented with the program Plotly of the R-Software (R-Core-Team 2010). We ran STRUCTURE (Pritchard et al. 2000) using the admixture model, and frequencies correlated among populations. We tested different values of K (from 1 to 6) 20 times, using a burn-in period of 500,000 steps followed by 1,000,000 Markov Chain Monte Carlo repeats. We obtained the most likely value of K using STRUCTURE HARVESTER (Earl and VonHoldt 2012). Average Q-matrices were obtained with CLUMPAK (Kopelman et al. 2015) and imported into EXCEL for graphical representation.

We obtained a sequence of 349 nucleotides of *cytb* for 24 individuals and 457 nucleotides of the left hypervariable region (HVR-1) of the *CR* for 55 individuals (GenBank Accession Numbers in Table S2). Basic sequence analysis was done with DnaSP v5 (Librado and Rozas 2009). Estimates of diversity and the net distance between population groups were quantified with MEGA 7 (Kumar et al. 2016) under Jukes-Cantor (JC) model. Standard errors were based on 1000 bootstrap replicates and significance was tested with the Z-test. *CR* haplotypes of the Clade West of *Rupicapra,* were downloaded from the GenBank and included in phylogenetic analysis. We constructed a Neighbor Joining (NJ) tree using JC distance. Besides, we obtained Maximun Likelihood (ML) and Minimum Evolution (ME) trees based on the optimal substitution model (Tamura 3-parameter) found with MEGA. Reliability of the nodes was obtained from 1000 bootstrap replicates. The haplotype *CR10* of *Rupicapra pyrenaica pyrenaica* (GenBank Acc. Nr. GU951852) was used as outgroup to root the trees.

Results and Discussion

The number of alleles for microsatellites in the total population was 115, with a mean number of 6.39 alleles per locus. LD was non-significant after Bonferroni correction ($P>0.05$). MICRO-CHECKER suggested null alleles at frequencies slightly exceeding 20% at four loci (SR-CRSP-6, SR-CRSP-8, INRA003, INRA036). The effect of excluding these loci in diversity estimates and F-statistics was limited, therefore we retained them for the analysis. Measures of diversity and allelic richness were lower

for the Olympus population than for the other three (Table 1). Total genetic diversity (53.6%) parallels the values observed for the large populations from the Pyrenees or the Alps (Pérez et al. 2002; Rodríguez et al. 2010; Soglia et al. 2010). The observed heterozygosities were lower than the expected, denoting a general deficit of heterozygotes, but the differences are significant only for the groups D and A+B+C, presumably due to the small sample size of other groups. High microsatellite diversity and deficit of heterozygotes was similarly observed in the *balcanica* populations from the Dinarides (Sprem and Buzam 2016) and from Bulgaria (Markov et al. 2016) consistent with recent and fragmentation.

Pairwise Fst values (Table 1) between the defined groups indicated differentiation between all of them. In addition, we checked for differentiation without an *a priori* definition of populations. The three major axes obtained after the PCoA explained 68.9% of the variation. Tridimensional representation of individuals shows two clearly differentiated clusters (Fig. 1) that correspond with the Olympus population on one side and the rest of specimens in the other. Following the program STRUCTURE, the most likely number of clusters was four (K=4) that relate to the defined population groups (Fig. 1). However, the regions of Pindos and the Northwestern border were not clearly delimited. The differentiation between groups of populations for microsatellites can be related to its recent population history and geographical distribution.

All the 24 sequences of *cytb* were the haplotype *cytb-07* (GenBank Acc. Nr. EU836156), the same identified previously in the samples from *asiatica* and *caucasica*. Regarding the *CR,* we identified 17 haplotypes (Table S2). Estimates of diversity within population groups (Table 1) reveal the low variability within Olympus (D) that contrasts with a high variability in the Pindos Mountains (A+B+C). Estimates of net evolutionary divergence between groups of sequences were all significant even though it must be said that standard errors for divergence computed by bootstrapping can be biased when sample size is very small. The overall diversity for the *CR* among Greek chamois was outstanding (haplotype diversity=74.7%, SE=5.7; nucleotide diversity=3.46%, SE=0.50) close to the diversity of the larger populations in the Alps (Crestanello et al. 2009; Rodríguez et al. 2010). This picture coincides with the findings in other mammals, such as roe deer (Randi et al. 2004) or wild boar (Alexandri et al. 2012), among others. The sequences of the subspecies *balcanica* inhabiting Greece group into three main clades (Fig. 2); one of them (in red) includes all the samples from Olympus as well as some samples from Pindus and the two samples from North Macedonia. This clade also includes the haplotypes from the Caucasus and Asia Minor. A second clade (in green) comprises the sequences from the Rhodopes, and a

third clade (in blue) contains most of the sequences from Pindos. The haplotypes of the subspecies *balcanica* inhabiting the south of the Dinarides, obtained by Sprem and Buzam (2016), together with the only sequence from that region in our previosu study, form a different clade closer to the subspecies *rupicapra*. The outstanding diversity present in the subspecies *balcanica* can be noted and, in particular, the concurrence of differentiated lineages in Greece.

There is differentiation between the Olympus population and the rest, both for mtDNA and for microsatellites. However, the mtDNA lineage of Olympus is shared with other regions. From these findings, it can be argued that the observed differentiation of chamois in Olympus is due to recent segregation and not to a singular long evolutionary history. Thus, the idea of a different subspecies (Koller 1929) is not supported by genetic data. The low diversity in Olympus is related to isolation together with small population size, consistent with the geographical distance and the orographic and ecological characteristics of the surrounding areas.

The data show that the chamois in Greece maintain an important deal of variability within the species *R. rupicapra* and consequently merits special conservation. The variation is structured among the different mountain ranges, coherent with the subdivision of the populations that has been reported (Papaioannou and Kati 2007; Papaioannou 2015, 2016). It is well known that the fragmentation of the populations leads to local inbreeding, loss of diversity and demographic stochasticity, factors of prime importance for population viability (Frankham et al. 2010). To avoid these effects, a suite of conservation measures such as the establishment of new wildlife reserves, the reinforcement of road control and guarding against poaching, should be implemented in the areas of the current populations. The populations of blocks A, B, C and F are little differentiated and the mountainous zones connecting these areas could function as genetic corridors. Such measures are expected to increase subpopulation sizes, and gene flow among subpopulations, as well as natural recolonization of previously occupied areas (a-i in the map). The latter is expected where neighboring subpopulations are quite large, as it was the case for the extinct subpopulations c and g, which are now under natural recolonization (Papaioannou 2016). This is deemed necessary especially for the small isolated population of NNW mountains (F), where additional conservation endeavors might be needed, including translocations with individuals from the genetically close population block (A). For populations of the Mount Olympus and the Rhodopes, which are more isolated and genetically differentiated, conservation actions should focus on maintaining a viable

population size. Finally, a systematic chamois monitoring scheme is advised to report on the trends of all subpopulations and on the success of implementing the suggested conservation measures.

Acknowledgments

We thank Sara de Albornoz and Victoria Coupe for the correction and improvement of our English. This work was funded by grant CGL2007-64315 from the Spanish *Ministerio de Educación y Ciencia* (FEDER support included). Special thanks to Ass. Prof V. Kati for her contribution to the MS writing, to Rhodope National Park Management Body and to the members of Balkan Chamois Society for their support in samples' collection as well as to the Greek Ministry of Environment and Energy (Department of Wild Life and Game) for providing the relative license for this study.

References

Alexandri P, Triantafyllidis A, Papakostas S, Chatzinikos E, Platis P, Papageorgiou N, Larson G, Abatzopoulos TJ, Triantaphyllidis C (2012) Balkans and the colonization of Europe: the post- glacial range expansion of the wild boar, *Sus scrofa*. J Biogeogr 39:713-723

Bonin A, Belllemain E, Eidesen PB, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. Mol Ecol 13:3261-3273

Corlatti L, Lorenzini R, Lovari S (2011) The conservation of the chamois *Rupicapra* spp. Mamm Rev 41:163-174.

Crestanello B, Pecchioli E, Vernesi C, Mona S, Martínková N, Janiga M, Hauffe HC, Bertorelle G (2009) The Genetic Impact of Translocations and Habitat Fragmentation in Chamois (*Rupicapra*) spp. J Hered 100:691-708

Earl DA, VonHoldt B (2012) Structure Harvester : a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:1-3

Frankham R, Briscoe DA, Ballou JD (2010) *Introduction to conservation genetics*, 2 edn. Cambridge University Press, Cambridge.

Goudet J (1995) FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. J Hered 86:485-486

Koller O (1929) Das Vorkommen von *Rupicapra rupicapra* auf dem Berge Olymp (Griechenland). Mammologische Mitteilun III: 46

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Mol Ecol Resour 15:1179-1191

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33:1870-1874

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451-1452

Markov G, Zhelev P, Ben Slimen H, Suchentrunk F (2016) Population genetic data pertinent to the conservation of Bulgarian chamois (*Rupicapra rupicapra balcanica*). Conserv Genet 17:155-164

Miller CR, Joyce P, Waits LP (2002) Assessing Allelic Dropout and Genotype Reliability Using Maximum Likelihood. Genetics 160:357-366

Papaioannou H (2015) Current status and conservation management of Balkan chamois (*R.r. balcanica*) in Greece. In: *Chamois International Congress* (eds. Antonucci A, Di Domenico G), pp. 111-122, Lama dei Peligni, Majella National Park, Italy.

Papaioannou H, Kati V (2007) Current status of the Balkan Chamois (*Rupicapra rupicapra balcanica*) in Greece: Implications for conservation. Belg J Zool 137:33-39

Papaioannou H (2016) The Balkan Chamois (*Rupicapra rupicapra balcanica* Bolkay, 1925) in Greece. Phd Thesis. In: *Department of Environmental and Natural Resources*, p. 264. University of Patras, Patras, Greece.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. Bioinformatics 28:2537-2539
Pérez T, Albornoz J, Domínguez A (2002) Phylogeography of chamois (*Rupicapra* spp.) inferred from microsatellites. Mol Phyl Evol 25:524-534
Pérez T, Fernandez M, Hammer SE, Dominguez A (2017) Multilocus Intron Trees Reveal Extensive Male-Biased Homogenization of Ancient Populations of Chamois (Rupicapra spp.) across Europe during Late *Pleistocene*. PLOS One 12: e0170392
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959
R-Core-Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
Randi E, Alves PC, Carranza J, Milosevic-Zlatanovic S, Sfougaris A, Mucci N (2004) Phylogeography of roe deer (*Capreolus capreolus*) populations: the effects of historical genetic subdivisions and recent nonequilibrium dynamics. Mol Ecol 13:3071-3083
Raymond M, Rousset F (1995) GENEPOP (version 1.2): Population genetics software for exact tests and ecumenism. J Hered 86:248-249
Rodríguez F, Hammer S, Perez T, Suchentrunk F, Lorenzini R, Michallet J, Martinkova N, Albornoz J, Domínguez A (2009) Cytochrome b Phylogeography of Chamois (*Rupicapra* spp.). Population Contractions, Expansions and Hybridizations Governed the Diversification of the Genus. J Hered 100:47-55
Rodríguez F, Perez T, Hammer SE, Albornoz J, Domínguez A (2010) Integrating phylogeographic patterns of microsatellite and mtDNA divergence to infer the evolutionary history of chamois (genus *Rupicapra*). BMC Evol Biol 10: 222
Soglia D, Rossi L, Cauvin E, Citterio C, Ferroglio E, Maione S, Meneguz PG, Spalenza V, Rasero R, Sacchi P (2010) Population genetic structure of Alpine chamois (*Rupicapra r. rupicapra*) in the Italian Alps. Eur J Wildl Res 56:845-854
Šprem N, Buzan E (2016) The genetic impact of chamois management in the dinarides. J Wildlife Manage 80:783-793
Valière N (2002) Gimlet: a computer program for analyzing genetic individual identification data. Mol Ecol Notes 2:377-379
Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICROCHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535-538

Figure legends

Fig. 1. a) Distribution of chamois in Greece (from Papaioannou 2015, 2016). Capital letters indicate the isolated population blocks, the numbers besides the letters correspond to estimated population sizes. Filled areas with numbers (1-24) indicate actual subpopulations and areas without filling (identified with letters) indicate locally extinct subpopulations or being currently under natural reconization. Filling colours indicate the groups considered for the analysis: wider area of Pindus mountain range (A+B+C), Olympus (D) Rhodope mountain range (E) and North-Northwestern mountains (F). b) Clustering of individuals without prior information on the basis of microsatellite genotypes obtained from Principal coordinates analysis (PCoA) of the multidimensional data and the program STRUCTURE (each vertical bar represents estimated membership coefficients, Q, for each individual in each cluster)

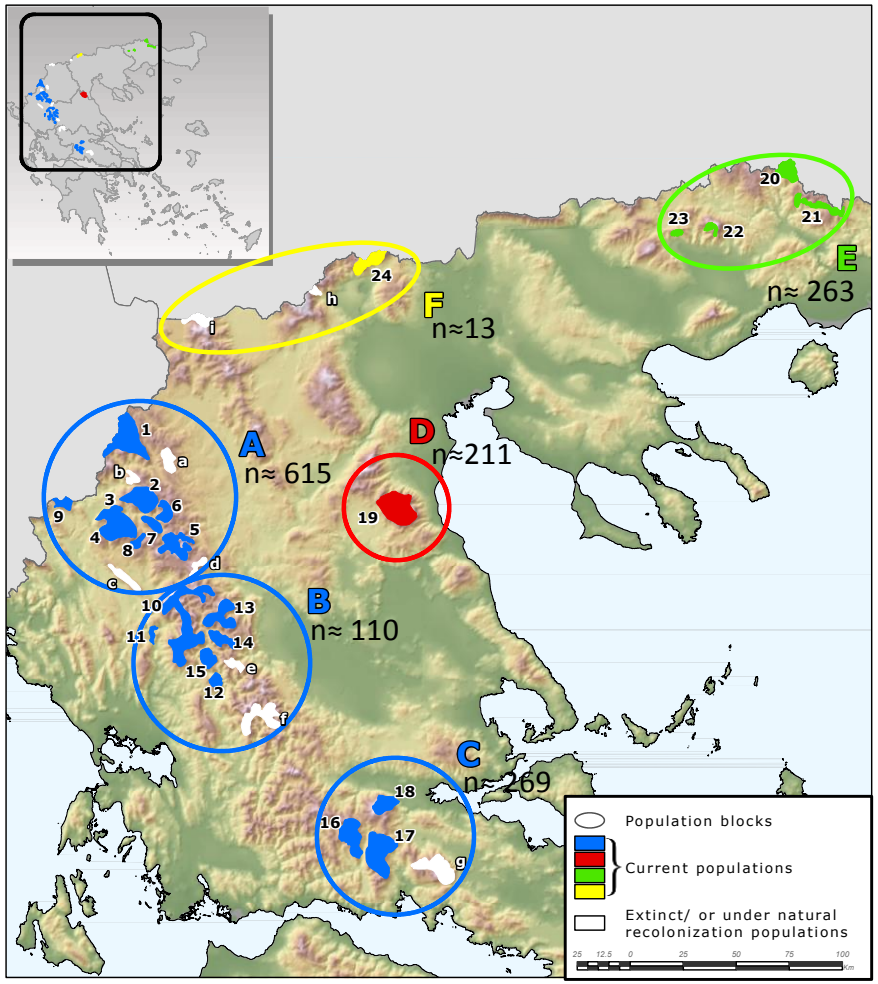
Fig. 2. Neighbor-Joining tree of the *CR* hablotypes based on the number of substitutions per nucleotide under the model of Jukes-Cantor. Next to each haplotype are the numbers of individuals in which it presents in each population block. Bootstrap supports larger than 50% are shown at the nodes. In

addition, ML and ME bootstrap support indices for the clades including *balcanica* are presented. GenBank Acc. Nr. and the initials denoting the subspecies of *R. rupicapra* are indicated (rup, *rupicapra*; tat, *tatrica*; cap, *carpatica*; bal, *balcanica*; asi, *asiatica*; and cau, *caucasica*). The population of origin of the *balcanica* haplotypes is specified.

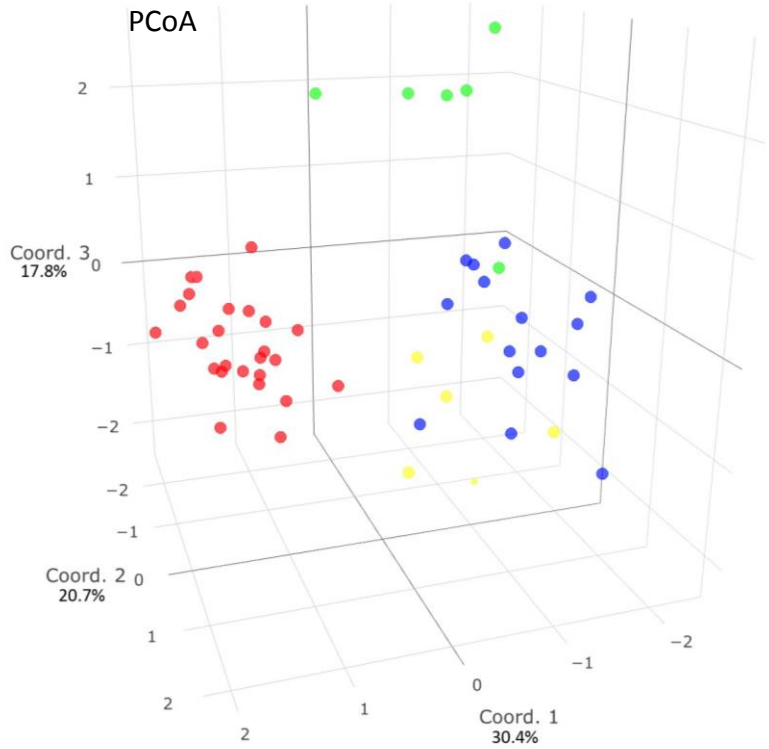
Figure 1

Click here to access/download;Figure;1902_Fig. 1.pptx

a)



b)



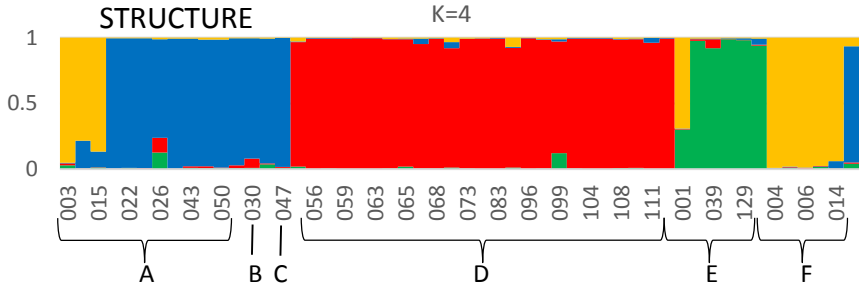


Figure 2

Click here to access/download;Figure;Fig2.pdf

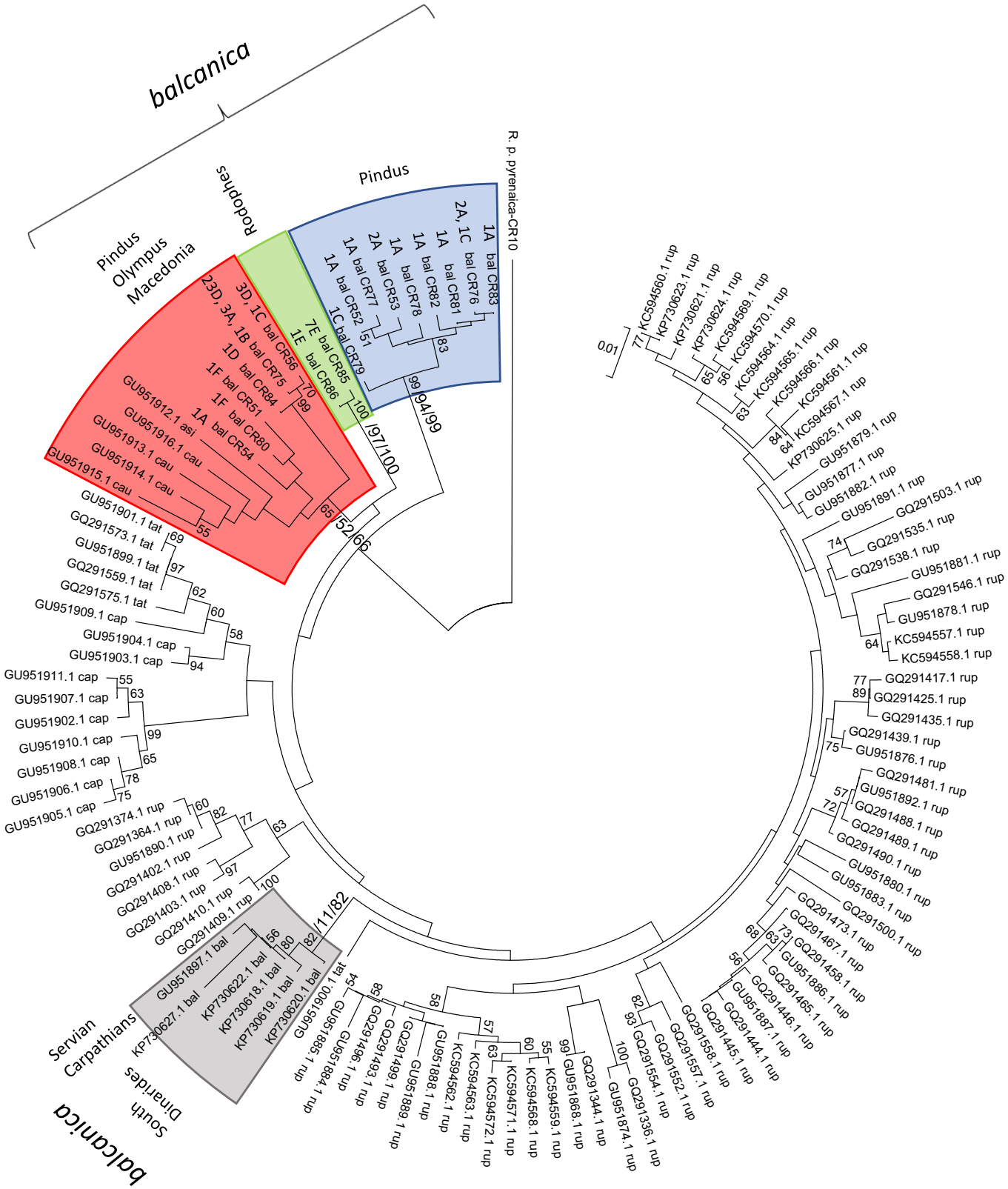


Table 1. Estimates of diversity (a) and differentiation among population groups (b) for microsatellites and the mitochondrial CR

a) Estimates of diversity in the four population groups

	Microsatellites				Control Region	
Population Group	n-µsats	Ar	%Ho	%He	n-CR	%π
A+B+C	15	3.50 (0.41)	40.0 (6.6)	49.9 (6.6)	18	3.49 (0.59)
D	25	2.49 (0.26)	35.1 (6.3)	38.9 (5.7)	27	0.06 (0.05)
E	6	3.61 (0.34)	47.2 (6.4)	51.5 (5.6)	8	0.05 (0.05)
F	6	3.44 (0.44)	42.6 (7.3)	48.8 (6.5)	2	1.10 (0.49)
Total	52	3.65 (0.42)	38.8 (1.3)	53.6(6.4)	55	3.46 (0.50)

n- µsats -sample size microsatellites, Ar = Allele richness. Ho = Observed Heterozygosity, He = Expected Heterozygosity. n-CR number of sequences obtained. π=nucleotide diversity using Jukes-Cantor model of substitution. In brackets, standard error of each statistic (SE).

b) Pairwise values of differentiation between population groups. %Fst values for microsatellites below the diagonal and net evolutionary distance in number of substitutions per 100 nucleotides in the CR above the diagonal. ** P<0.01; *** P<0.001

Population Group	A+B+C	D	E	F
A+B+C	-	2.94 ***	3.32 ***	1.61 ***
D	18.9 ***	-	4.67 ***	2.43 ***
E	17.1***	26.4***	-	3.79 ***
F	9.9 **	22.3 ***	16.6**	-

Supplementary Material 1

Click here to view linked References

Click here to access/download
Supplementary Material
Table S1-2.xls

Supplementary Material 2

Click here to view linked References