Glacial allopatry vs. postglacial parapatry and peripatry: the case of hedgehogs

Barbora Černá Bolfíková (1, §), Kristýna Eliášová (2,3 *), Miroslava Loudová (2), Boris Kryštufek (4), Petros Lymberakis (5), Attila D. Sándor (6), Pavel Hulva (2,7)

(1) Faculty of Tropical AgriSciences, Czech University of Life Sciences Prague, Prague, Czech Republic

(2) Faculty of Science, Charles University in Prague, Prague, Czech Republic

(3) Department of Zoology, National Museum, Prague, Czech Republic

(4) Slovenian Museum of Natural History, Ljubljana, Slovenia

(5) Natural History Museum of Crete, University of Crete, Heraklion Crete, Greece

(6) Faculty of Veterinary Medicine, University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, Cluj-Napoca, Romania

(7) Faculty of Science, University of Ostrava, Ostrava, Czech Republic

§ corresponding author

Barbora Černá Bolfíková
Faculty of Tropical AgriSciences, Czech University of Life Sciences Prague,
Kamýcká 129
165 21 Prague
Czech Republic
Email: bolfikova@ftz.czu.cz
Tel. 00420 22438 2497
Fax. 00420 22438 2497

2 **Abstract**

3 Although hedgehogs are well-known examples of postglacial recolonisation, the specific

4 processes that shape their population structures have not been examined by detailed sampling

5 and fast-evolving genetic markers in combination with model based clustering methods. This

6 study aims to analyse the impacts of isolation within glacial refugia and of postglacial expansion

7 on the population structure of the Northern White-breasted hedgehog (*Erinaceus roumanicus*). It

8 also discusses the role of the processes at edges of species distribution in its evolutionary history.

9 The maternally inherited mitochondrial control region and the bi-parentally inherited nuclear

10 microsatellites were used to examine samples within the Central Europe, Balkan Peninsula and

11 adjacent islands. Bayesian coalescent inference and neutrality tests proposed a recent increase in

12 the population size. The most pronounced pattern of population structure involved differentiation

13 of the insular populations in the Mediterranean Sea and population within contact zone with *E.*

14 *europaeus* in Central Europe. An inter-species hybrid was detected for the first time in Central

15 Europe. A low genetic diversity was observed in Crete, while the highest genetic distances

16 among individuals were found in Romania. Recent population in post-refugial area related to

17 Balkan Peninsula shows a complex pattern with pronounced subpopulations located mainly in

18 the Pannonian basin and at the Adriatic and Pontic coasts. Detailed analyses indicate that

19 parapatry and peripatry may not be the only factors that limit range expansion, but also strong

20 microevolutionary forces that may change the genetic structure of the species. Here we present

21 evidence showing that population differentiation may occur not only during the glacial restriction

22 of the range into the refugia, but also during the interglacial range expansion. Population

23 differentiation at Balkan Peninsula and adjacent regions could be ascribed to diversification in

24 steppe / forest biomes and complicated geomorphology, including pronounced geographic

25 barriers as Carpathians.

Background

27 Consequences of geographic isolation and gene flow limitation are highly debated topics in

28 biogeography (Coyne & Orr, 2004). Strong empirical evidence for the role of isolation in

29 different modes of speciation has been obtained by analysis of temperate species that underwent

30 episodes of refugial isolations in the Mediterranean region during the Pleistocene climate

31 oscillations (Hewitt, 2000). Balkan Peninsula acted as one of the main glacial refugium for

32 European biota (Hewitt, 2000) and is considered as an important source of endemism and genetic

33 diversity due to its position, historical development, topography and climatic variability

34 (Kryštufek & Reed, 2004). After the Last Glacial Maximum (LGM), many species expanded

35 from this area northward (Hewitt, 2000). Range expansion may be limited not just by various

36 abiotical factors, but also by formation of a contact zone and parapatry with a related species.

37 Interspecies interactions may lead to quick reinforcement, i.e. evolution of adaptations

38 preventing formation of hybrids. Terrestrial taxa whose ranges are restricted by the sea barrier

39 may colonise particular islands, providing an opportunity for peripatric evolution which is

40 characterised by a strong founder effect, genetic drift and other island phenomena.

41 Hedgehogs from the *Erinaceus* genus are a classical example of postglacial recolonisation of the

42 Western Palearctic (Hewitt, 2000). Recently, we recognize three species occupying this area.

43 Northern White-breasted hedgehog (*Erinaceus roumanicus* Barrett-Hamilton, 1900) currently

44 inhabits Central and Eastern Europe, the Baltic, and the Balkan Peninsula, and its eastern range

45 stretches to western Siberia. The current data indicate that the Balkan Peninsula used to be the

46 only refugium for *E. roumanicus* (Seddon et al., 2001; Hewitt, 2001). The Caucasus Mountains

47 and the Bosporus strait currently form a barrier between *E. roumanicus* and its sister species

48 Southern White-breasted hedgehog (*E. concolor* Martin, 1837), which inhabits Asia Minor. It

49 has been estimated that these two species diverged approximately 0.4 − 1.4 million years ago

50 (Bannikova et al., 2014). Western Europe and Scandinavia are populated by Western European

51 hedgehog (*E. europaeus* Linne 1758) whose refugial area used to exist in the Iberian and

52 Apennine Peninsulas. After the LGM, hedgehog species expanded from the southern refugial

53 areas and created contact zones in the Baltics and in Central Europe (Seddon et al., 2001;

54 Santucci, Emerson & Hewitt, 1998). Central European contact zone was probably established

55 during the Neolithic deforestation and occurrence of interspecies hybridization is probably rare

56 or absent (Bolfíková & Hulva, 2012). On the other hand one hybrid individual was reported from

57 Baltic contact zone and the authors speculate that hybridization must be common in this area

58 (Bogdanov et al., 2009). Seddon et al. (2001) suggest the presence of at least three mitochondrial

59 lineages of *E. roumanicus*. The first lineage, which previously expanded from the Balkans to

60 Russia and Ukraine, colonises the northern foothills of the Caucasus Mountains. The second

61 lineage colonises western parts of Europe, and the third lineage is present in Greece.

62 A large morphologic variability has been observed for *E. roumanicus* throughout its wide

63 population range, and multiple subspecies have been described. A study has recently proved that

64 body size variability is strongly associated with the environment, and it clinally varies in the

65 mainland (Kryštufek et al.*,* 2009). A morphometric study by Kryštufek et al. (2009) revealed a

66 positive correlation between body size and temperature and a negative correlation between body

67 size and summer precipitation.

68 Association between particular species and their habitats give rise to the need of comparative

69 phylogeographic studies. Dependency of *Erinaceus* hedgehogs on deciduous woodland (Rautio

70 et al., 2014; Jensen, 2004) is in concordance with similar phylogeography observed in Oak and

71 Silver Fir (Hewitt, 2000; Seddon et al., 2001) and implies that the extension of forests was

72 followed by the hedgehogs. Some literature highlights an *E. roumanicus* preference for lowlands

73 and open landscapes, particularly during the spatial expansion (Bolfíková & Hulva, 2012;

74 Seddon et al., 2002). However, modern studies based on past environmental niche modelling are

75 missing. Moreover, ecological considerations are based rather on knowledge of *Erinaceus*

76 *europaeus*, while data on *E. roumanicus* are scarce. Different position of these two species

77 ranges on often overlooked oceanic–continental axis, represented by more humid and less

78 seasonal climate on the west and the opposite in the east, should be also considered (Stewart et

79 al., 2010). The southern parts of *E. roumanicus* range are located mainly in drier biomes in the

80 Mediterranean basin and Pontic-Caspian steppe. Northern parts of the range are situated mainly

81 in the temperate deciduous forests. Adaptations to the more dry open and seasonal habitats in *E.*

82 *roumanicus* could not be excluded (Bolfíková & Hulva, 2012).

83 Do the classical models have a potential to further develop phylogeogeographic paradigm using

84 up to date tools of the landscape genetics? Here, we provide new information on the hedgehog's

85 evolutionary history using a combination of fast evolving markers, including nuclear

86 microsatellites and mitochondrial control region. By analysing multiple *E. roumanicus* samples

87 ranging from the post-refugial area in the Balkan Peninsula to the parapatric contact zone with *E.*

88 *europaeus* in central Europe, we aim to:

89 1) assess the population structure of *E. roumanicus* within the studied area using the spatial and

90 non-spatial Bayesian clustering methods and discuss observed patterns within landscape genetic

91 framework in the light of geomorphology and habitat quality of the region;

92 2) reveal the history of island populations and suggest possible way of colonisation of these

93 islands. He hypothesize, that insular effects like founder effect, genetic drift and other factors

94 which decrease genetic variation highly affected these populations;

95 3) analyse the populations in the secondary contact zone from Central Europe. We hypothesize

96 that respective populations were affected by species interactions, including introgression and

97 reinforcement;

98 4) analyse the populations from the post-refugial areas at Balkan peninsula and assess the

99 hypothesis of occurrence of the glacial subrefugia.

100

101 **Material and Methods**

102 *Sampling and genotyping*

103 Samples obtained from muscle tissue or digits or ears were collected from road-killed animals or

104 from museum collections. No special permit is needed from the Institutional Review Board. The

105 tissues were fixed in 96% ethanol and stored at -20 °C. The samples were obtained in

106 cooperation with the Natural History Museum of Crete (Petros Lymberakis), the Slovenian

107 Museum of Natural History (Boris Kryštufek), and the University of Agricultural Sciences and

108 Veterinary Medicine Cluj Napoca (Attila Sandor). For the study, 314 samples were analysed

109 from the following European countries: Czech Republic (49), Slovakia (26), Hungary (9),

110 Romania (43), Slovenia (44), Italy (2), Croatia (20), Bosnia and Herzegovina (25), Montenegro

111 (6), Serbia (10), Bulgaria (4), Macedonia (17) and Greece (59), which included several islands

112 (the Ionian islands, Euboea, Cyclades, and Crete). Each sample's origin, GPS coordinates, and

113 date of collection along with the name of the person who collected it are listed in Supplemental

114   information: Table S1. A map (Fig. 1), indicating the geographic origins of the samples, was

115   created using Geographic Information System (ESRI, 2014).

116   DNA was extracted using a commercial DNA Blood and Tissue Kit (Qiagen, Hilden, Germany).

117   A combination of nine microsatellite markers and mtDNA control region sequences was used.

118   Laboratory procedures were performed according to the protocols by Bolfíková & Hulva (2012),

119   details may be found in Supplemental information: Table S3. Each dataset was analysed

120   separately as specified below.

121

122   *Non-spatial analyses of genetic variability*

123   The   sizes   of   the   microsatellite   alleles   were   determined   using   GeneMarker   V1.85

124   (www.softgenetics.com).   Data   binning   was   performed   with   the   AutoBin   software

125   (http://www4.bordeauxaquitaine.inra.fr/biogeco/layout/set/print/Ressources/Logiciels/Autobin).

126   The presence of the null alleles (false homozygotes), stutter bands and large allele dropouts was

127   tested using Micro-Checker (Van Oosterhout et al., 2004). Bayesian statistics for recognising

128   hybrid categories was run in NewHybrids 1.1 (Anderson & Thompson, 2002).

129   Evaluation of the genetic structuring of the microsatellite dataset was performed with a Bayesian

130   Clustering   Analysis   in   Structure   2.2   (Pritchard,   Stephens   &   Donnely,   2000).   Structure

131   determines   the   most   likely   assignment   of   individuals   to   clusters   (K)   based   on   analysis   of

132   likelihood. Structure is powerful to detect clinal variability and high geographic admixture in the

133   dataset (Chen et al., 2007). The parameters of the final run were as follows: 1 000 000 MCMC

134   steps,   with   burn-in-period   100 000   and   5   iterations   for   each K.   An   Admixture   Model   of

135   Correlated Frequencies was used. The number of tested population was within the $K = 1 - 14$

136   interval. The results of the analyses were combined and evaluated by Structure Harvester (Earl &

137 VonHoldt, 2012) using the Evanno method (Evanno, Regnaut & Goudet, 2005). Graphic

138 visualisation was provided using Distruct 1.1 (Rosenberg, 2004). The inbreeding coefficient (F_{IS})

139 for each population was tested by GenePop (Rousset, 2008). Departures from the Hardy-

140 Weinberg equilibrium (*HWE*), number of alleles (N_a) and allelic richness (A_R) were estimated in

141 FSTAT (Goudet, 1995). Allelic richness was used to estimate an allele numbers, across all loci,

142 corrected for equal sample size in all analysed populations. The population with the smallest

143 number of individuals was $N = 24$ from Crete. The expected (H_E) and observed (H_O)

144 heterozygosities were computed using Genetix (Belkhir et al., 1999).

145 Sequences of the mtDNA partial control region were edited using Geneious (Kearse et al., 2012)

146 and aligned using MAFFT (Katoh & Standley, 2013). Particular haplotypes were identified in

147 DnaSP v.5 (Rozas et al., 2003). The haplotype data were submitted to GenBank® (Accession

148 numbers are from KY489901 to KY489953). The relationships among the haplotypes were

149 visualised by the median-joining network (Bandelt, Forster & Röhl, 1999) using Network 4.5

150 (www.fluxus-engineering.com). The demographic parameters were estimated by summary

151 statistics of the genetic variability for each population (recognised by the Bayesian clustering

152 analysis of the haplotype data using the Geneland software; details in spatial analyses below).

153 Tests for haplotype diversity (*h*; represents the probability that two randomly chosen haplotypes

154 are different), nucleotide diversity (π; the average number of nucleotide substitutions per site

155 between two sequences) and neutrality tests (Tajima's *D*, Fu's F_S and R^2) were computed for each

156 population in DnaSP v.5 (Rozas et al., 2003). Fu's F_S (for big sample sizes) and R^2 (for small

157 sample sizes) are considered as the most powerful tests (Ramos-Onsins & Rozas, 2002, Ramírez-

158 Soriano et al., 2008). Significantly negative results indicate presence of selective sweep or strong

159 population growth signal in the data. The course of the past population size was estimated with

160   coalescent-based Bayesian skyline plots (Drummond et al., 2005). We inferred the model of

161   sequence evolution in jModelTest (Posada & Crandall, 1998) using Bayesian information

162   criterion (BIC). The Bayesian skyline plot analysis was conducted using the BEAST 1.4.8

163   program (Drummond & Rambaut, 2007) with a strict molecular clock and a GTR+I+G

164   substitution model. The MCMC procedure was run three times for each species with 30 000 000

165   iterations, and the genealogy and parameters of the model were stored every 1 000 iterations.

166   The results were combined in LogCombiner, and burn-in was set to 10 000 000 iterations in each

167   run. The Bayesian skyline plot and the convergence of chains were assessed by Tracer v.1.4

168   (Rambaut et al., 2014).

169

170   *Spatial analyses of genetic variability*

171   To analyse the spatial genetic architecture of *E. roumanicus* in the area of interest, we used two

172   spatial model based clustering approaches, including software Geneland 4.0.4 (Guillot, Mortier

173   & Estoup, 2005) and TESS (Chen et al., 2007; Durand et al., 2009). Geneland encompasses

174   Bayesian clustering of individual genotypes based on separation of random mating

175   subpopulations and produces Dirichlet cells centred on subpopulation territories. Geneland can

176   analyse both haplotype data and codominant data in independent runs. The parameters for the

177   Geneland analysis were set for each marker type as follows: 500 000 MCMC iterations, thinning

178   = 100, five independent runs and K = 1 – 15. The posterior probability of the subpopulation

179   membership for the most supported run was computed for each pixel of the spatial domain, using

180   a burn-in of 50 iterations. An uncorrelated Frequency Model was chosen. TESS is based on

181   minimizing of Wahlund effect and produces Dirichlet cells based on individuals and is using

182   microsatellite data. We chose BYM model due to lower Deviance Information Criterion (DIC)

183 and higher stability of algorithm. The first run of program was set to *Kmax* = 2 - 10 populations,

184 then we plotted *Kmax* versus DIC values and selected the optimal *Kmax* (see details in

185 Supplemental information: Figure S1). The final run of TESS was set to *Kmax* = 5 - 6, with 50

186 000 sweeps, a burn-in 30 000, the interaction parameter Ψ = 0.6 and 100 iteration for each K. The

187 lowest DIC values were performed for *Kmax* = 6 and we chose as representative the run within

188 30% of runs with lowest DIC values, displaying five colours.

189 Combination of these approaches may be useful in exploring spatial population structure

190 (Balkenhol et al., 2015), as Geneland performs well in detecting recent linear barriers to gene

191 flow (Blair et al., 2012), while TESS addresses situations with barriers of complex shapes and

192 permeability for migrants, as well as secondary fusions of subpopulations (Chen et al., 2007).

193 The Mantel test (Mantel, 1967) was used to characterise the presence of isolation-by-distance to

194 evaluate the correlation between congruent similarity/dissimilarity matrices, which addressed the

195 relationship between the geographic and genetic distances of observations across a landscape

196 provided by the Alleles in Space (AIS) software (Miller, 2005). The number of iterations used

197 was 1 000. We used the AIS software for the Genetic Landscape Shape Analysis to compute the

198 genetic distances, which were assigned to the connector centres of each locality using

199 Delauney´s triangulation. The interpolation procedure, which determined the presumed genetic

200 distances in a lattice frame containing 15-km units, covered the area of interest, and a graph of

201 the interpolated distances was created. Peaks in this graph represent an area of high genetic

202 distances and should be correlated with areas of restricted gene flow. Both marker types were

203 independently analysed using AIS software.

204

Results

206   We did not detect any genotyping errors that resulted from a large allele dropout or stutter bands

207   in 260 samples that were successfully genotyped at nuclear markers. All genotyped individuals

208   are indicated in Supplemental information: Table S1. Micro-Checker detected null alleles in

209   several microsatellite loci, and the estimated frequencies differed between loci and populations.

210   The estimated frequencies were lower than 0.2432 for all populations examined. In the Cretan

211   population, two loci, EEU12H (0.2432) and EEU3 (0.2099), showed high null allele frequencies

212   using the methods described by Van Oosterhout et al. (2004). Primers were designed for *E.*

213   *europaeus* (Becher & Griffiths, 1997; Henderson et al., 2000) and tested by Bolfíkova & Hulva

214   (2012). All genotypes were originally obtained from the tissues, which are less susceptible to the

215   presence of null alleles than non-invasive samples. Further analyses proved that the Cretan

216   population was influenced by the founder effect, genetic drift and inbreeding. Thus, the null

217   allele estimation was probably influenced by the high homozygote frequency in the population.

218   Sequences produced by Sanger sequencing were of good quality and some individuals with

219   ambiguous sites were re-sequenced or skipped from the analyses. A total of 296 partial control

220   region sequences (404 bp) were obtained and represented 53 haplotypes among the analysed

221   samples (see details in Supplemental information: Table S1).

222

223   *Non-spatial analyses of genetic variability*

224   The nine selected microsatellite loci were polymorphic. The total number of alleles in a single

225   locus varied from 5 to 23 for the entire dataset. One sample from the Slovakia (Trencianske

226   Teplice) was omitted from further analysis because it was suspected of having a hybrid origin. A

227   follow-up analysis with the NewHybrids program (Anderson & Thompson, 2002) using the

228   original dataset from Bolfíková & Hulva (2012) confirmed that the sample was a backcrossed

229   hybrid of *E. europaeus* and *E. roumanicus*; the mitochondrial DNA originated from *E.*

230   *roumanicus*.

231   From the total number of tested clusters ($K = 1 - 14$), division into the three populations ($K = 3$)

232   obtained the highest support according to Evanno et al. (2005). The populations were

233   differentiated as follows: (i) samples from Crete and from the Hellenic peninsula in southern

234   Greece; (ii) samples from the Balkan area (Serbia, Bulgaria, Macedonia, Greece, Montenegro,

235   Croatia and Bosnia and Herzegovina) plus Slovenia and Romania; and (iii) samples from Central

236   Europe (Czech Republic, Slovak Republic and Hungary) (Fig. 2a). The second highest ΔK value

237   was observed for a division into the seven populations ($K = 7$). An increasing number of clusters

238   resulted in further differentiation of the Balkan population by separation of the population from

239   Pannonia (at $K = 4$), differentiation of the Slovenian population ($K = 5$) and separation of

240   populations from the west and east of the Balkans (from $K = 6$). However, there is certain degree

241   of admixture among particular clusters. Two different landscape genetic algorithms supported

242   similar spatial pattern with three (Geneland, Fig. 5d) and five (TESS, Fig. 2b) clusters.

243   The aforementioned differentiation was also tested without the Czech and Slovakian samples,

244   whose genetic architecture could have been influenced by phenomena associated with the

245   recently developed contact zone. This analysis of Structure supported the previous subdivision of

246   our dataset.

247   The genetic variability parameters are given for the seven populations (Table 1) and three

248   populations (Supplemental information: Table S2). H_O was significantly lower than H_E and

249   GenePop analyses showed a significant lack of heterozygotes ($p < 0.01$) for each population.

250   Crete had lowest A_R and H_E. Highest F_{IS} was detected in Romanian population.

251 The median-joining network didn't demonstrate reciprocally monophyletic lineages in the

252 studied area (Fig. 3). Haplotypes are shared across populations, but some trends are visible. One

253 unique haplotype shared with majority of individuals from Crete and few other haplotypes

254 carried by one or two individuals are representing population of South Balkan (in blue on Fig. 3).

255 Similar pattern is representing Central Europe (in orange on Fig. 3). North Balkan population (in

256 pink on Fig. 3) has two locally common haplotypes, but individuals from Central Europe and

257 Dinaric area are also sharing them. Carpathian population (in yellow on Fig. 3) had several

258 haplotypes in common with other populations but also several unique haplotypes. Samples from

259 Dinaric area (in green on Fig. 3) were wide-spread in the network and produced star-like pattern

260 around haplotype that is common in Slovenia, Hungary and Croatia. Black sea coast population

261 (in violet on Fig. 3) has several unique haplotypes which are widespread within the network.

262 The genetic variability parameters for mitochondrial data were estimated for six populations

263 according to the results of analysis in Geneland (Table 2). All tests of neutrality yielded negative

264 values indicating recent population growth, only Central European (containing Czech Republic,

265 East Slovakia and North Hungary) population neutrality tests were not significant (Table 2). The

266 Bayesian skyline plot indicated a long period with a stable population size and its recent increase

267 (Fig. 4.).

268

269 *Spatial analyses of genetic variability*

270 The Mantel test indicated a significant isolation by distance (r = 0.303; p < 0.001) within our

271 dataset. Geneland analysis of microsatellite markers determined that the division of our dataset

272 into three populations (Central Europe, mainland area, and south Greece plus islands) was the

273 most likely to occur (Fig. 5d). Analysis implemented by TESS revealed the most fitting *Kmax* =

274 6, which spatially represents five clusters corresponding to population from Crete, Czech

275 Republic, Pannonian area, Eastern and Western part of Balkan Peninsula (Fig. 2b). This pattern

276 is in concordance with results of Structure analysis for $K = 5$ (Fig. 2a). Mitochondrial analysis

277 divided our dataset into the following six clusters: (i) South Balkan (southern and central Greece,

278 Peloponnese, Crete and the adjacent islands); (ii) Dinaric area (north Greece, Macedonia, south

279 Bosnia and Hercegovina, south Croatia, Montenegro and south Serbia); (iii) Black sea coast

280 (Thrace and most of Bulgaria and southeast Romania); (iv) Carpathian area (northwest Romania,

281 east Slovak Republic and east Hungary); (v) North Balkan (north Serbia, north Bosnia and

282 Herzegovina, north Croatia, Slovenia, southwest Hungary and north Italy); and (vi) Central

283 Europe (northwest Hungary, west Slovak Republic and Czech Republic) (Fig. 5c; Supplemental

284 information: Table S1). Genetic Landscape Shape Analysis using AIS software detected very

285 low genetic distances in Crete and the Czech Republic based on the mitochondrial data and

286 nuclear data in particular. However, high genetic distances were observed with the

287 microsatellites among samples in regions to the south and east of the Balkan Peninsula, with the

288 highest peaks observed in Romania (Fig. 5b). The mitochondrial data showed flat genetic

289 landscapes (Fig. 5a).

290

Discussion

291

292 Based on the results obtained using both non-spatial and spatial individual based clustering

293 approaches, the population from the southernmost region of the range (including southern

294 Greece and the adjacent islands) and the population from the contact zone with *E. europaeus*

295 were separated from the rest of the range. This pattern indicates that processes acting at

296 peripatric and parapatric range edges shaped the genetic architecture of the species substantially.

297   Further differentiation in clustering hierarchy provides evidence for occurrence of

298   subpopulations within southeast part of the range. This result could be ascribed to complicated

299   geomorphology and land cover of the region and may imply also subdivision of refugial

300   populations in the past.

301

302   *Insular populations in the Mediterranean Sea*

303   All analyses indicate strong differentiation of the individuals from Crete, Cyclades and from

304   Peloponnese, Euboea and southern and central Greece. As the largest island of Greece and the

305   fifth largest Mediterranean island, Crete is typified by impoverished native mammalian fauna

306   and the high impact of recently introduced species. However, Crete is also known as a generator

307   of novelty in the numerous lineages that were able to colonise the island. From a

308   macroevolutionary viewpoint, Crete is the place where pygmy forms of deer, elephants and

309   hippopotamuses occurred (Van der Geer et al., 2010). From the microevolutionary viewpoint, the

310   cryptic bat species occur there (Hulva et al.*,* 2007; Hulva et al., 2010). Peloponnesus is also

311   typified by its cryptic variability in vertebrates (Gvoždík et al., 2010). The isolation of this region

312   was recently enhanced by the 1893 construction of the Corinth Canal, which effectively

313   separated Peloponnesus from the mainland. However, our analyses do not confirm a reciprocal

314   monophyly for the *E. r. nesiotes* subspecies, which supports the findings by Schaschl,

315   Lymberakis & Suchentruck (2002). Moreover, we ascertain that in addition to Crete, the range of

316   this population includes other islands and reaches continental Greece. The assessment of the

317   geographic origins and dispersal scenarios of this population will require more detailed analysis.

318   However, the presumably absent land bridge between Crete and the mainland during the

319   Pleistocene period (Schule, 1993), the absence of hedgehogs in Crete's fossil record (Lax &

320   Strasser, 1992) and the frequent introductions of hedgehogs onto the islands by humans

321   (Bolfíková et al., 2013) indicate that the role of the anthropogenic factor in the phylogeography

322   of the southern cluster is highly probable.

323   The first archaeological evidence of hedgehogs in Crete was collected from an Early Minoan

324   tomb (Jarman, 1996), but the author concluded that the hedgehog might have been an intrusion

325   (e.g., entered the tomb much later). Using allozymes and partial cytochrome *b* sequences,

326   Schaschl, Lymberakis & Suchentruck (2002) proposed that hedgehogs were introduced to Crete

327   by archaic settlers from the mainland. This population was previously described as a subspecies,

328   *E.r. nesiotes*, by Dorotha Bate in 1906 based on phenotypic data. These facts imply the

329   occurrence of insular syndrome in the Cretan population. The markedly decreased genetic

330   variation values in the insular population imply that factors such as the founder effect and

331   genetic drift played substantial roles. The agreement between the genetic variation patterns and

332   the phenotypic data indicates that insular syndrome, including dwarfism in hedgehogs from the

333   Greek islands, has occurred (Kryštufek et al., 2009; Giagia-Athanasopoulou & Markakis, 1996).

334   Despite the uncertainty surrounding the origins of differentiation for the southern cluster, the

335   aforementioned results indicate a pronounced role for peripatric processes in the microevolution

336   of *E. roumanicus*. Elevated evolutionary rates in islands are obvious also in phenotypic

337   evolution. The trends in body size changes of insular hedgehogs are negatively correlated with

338   the island's distance from the mainland, however, the insular response was not uniform

339   (Kryštufek et al., 2009).

340

341   *Secondary contact zone with E. europeaus in Central Europe*

342   The population from Central Europe is genetically different from the rest of the mainland, as

343   suggested by all Bayesian clustering analyses. This cluster is also characterised by its lower

344   genetic variation, even though, in some parameters still comparable to post-refugial populations.

345   This may be the outcome of spatial expansion, which is usually followed by a decrease in genetic

346   variability due to the bottleneck effect. More complex phenomena, such as allele surfing

347   (increase of an allele frequency in the wave front of an expanding population), might alter the

348   genetic architecture of the population's expanding edge and change the genetic variation as well

349   (Klopfstein, Currat & Excoffier, 2006; Excoffier & Ray, 2008; Excoffier, Foll & Petit, 2009).

350   Demographic analyses did not confirmed statistically significant population size increase. Thus,

351   we suspect that the aforementioned phenomena played important roles in shaping the population

352   genetics of Central European hedgehogs. Other factors that caused population differentiation in

353   this case may be associated with parapatry and its consequent species interactions. Ecological

354   interactions and their consequent selections may have influenced the diversity of the studied

355   microsatellite loci by genetic hitchhiking, while hybridisation and introgression in the secondary

356   contact zones may have further shaped the allopatrically evolved lineages and potentially cover

357   the signal of recent population expansion. These potential factors will require further study and

358   genomic tools.

359   We are the first to report a single hybrid sample from the range margins of *E. europaeus* in the

360   Central Europe (Slovakia - Trenčianske Teplice) using traditional genetic markers. Interspecies

361   interactions in the sympatry zone may have previously played roles in forming reproductive

362   barriers. We hypothesised that hybridisation occurred when the two species came into contact for

363   the first time, when the isolating mechanisms were not yet formed. Afterwards, during the

364   reinforcement phase, species developed mechanisms for reproductive isolation and the hybrid

365 event frequency decreased for follow-up generations. The history of possible hybridisation

366 events in the Central European population should be tested using approaches that include whole

367 genome data because the introgression modes may be complex. Future investigations should

368 address the contact zone at the Slovenian/Italian border (located to the south of the Alps) and the

369 effect of altitudinal heterogeneity on both species. This region involves the contact zone for both

370 species and the transition zone between the Slovenian mountains and the Padan Plain.

371 The lower genetic diversity associated with the range margins and the novel selective pressures

372 applied by interactions with related species may be why the Central European population was

373 highly divergent from the other populations in the dataset.

374

375 *Post-refugial areas at the Balkan Peninsula and adjacent regions*

376 Separation of the Pannonian population represents the most pronounced pattern after the split of

377 parapatric and peripatric populations. Pannonian basin is confined area from the view of

378 geographical isolation by Carpathian arch and vegetation cover typical by grass steppe, i.e.

379 possess habitat similarity with the part of species range in Pontic-Caspian steppe. Consequently,

380 it represents pronounced biogeographical region of Europe and presumed continental interglacial

381 refugium of continental and steppe faunal elements (Stewart et al., 2010, Říčanová et al., 2011).

382 In *E. roumanicus*, this pattern might mirror geographical isolation, ecological differentiation, and

383 also the quality of ancestral habitat and larger extent of steppe biome during glacial periods.

384 Further differentiation occurs between populations from the west of the peninsula at the Adriatic

385 coast and the east at Black sea coast. This differentiation might be ascribed to isolation by

386 distance and barrier role of Carpathian Mountains. For example, based on the results provided by

387 AIS, the largest genetic distances between the samples in our dataset are located in Romania.

388 This pattern implies a restricted level of gene flow within the population, which might be

389 ascribed to the high altitudinal heterogeneity of the Carpathian Mountains. However, this pattern

390 may also mirror occurrence of forest refugia at Adriatic and Pontic regions. The Black Sea shore

391 likely served as a glacial refugium for multiple animal species, such as *Bombina bombina*

392 (Fijarczyk et al., 2011) and *Emys orbicularis* (Joger et al., 2007). The presence of trees during

393 the LGM has also been proven (Tzedakis, 2004). Palynological analyses have confirmed

394 occurrence of isolated forests within the refugia in response to climatic and spatial variabilities

395 (Huntley, 1999), for example in the Pindos Mountains in northwestern Greece, where levels of

396 precipitation were higher (Tzedakis, 2004). In the case of the European oak *Quercus*, two

397 separate Balkan subrefugia (Greece and the west coast of the Black Sea) have been located

398 (Brewer et al., 2002). King & Ferris (1998) discovered unique haplotypes of the European alder

399 (*Alnus glutinosa*) in the regions around Bulgaria and Greece. Based on these examples, Tzedakis

400 (2004) presumed the existence of temperate forests during the last glacial maximum in the

401 mountainous and coastal areas, whereas two previously identified subrefugia occurred in western

402 Greece and in eastern Bulgaria near the Black Sea.

403 The population from Slovenia is separated from the Balkan genotypes in Structure analyses. This

404 pattern suggests a gene flow limitation due to the geographical isolation, caused by the Alps, the

405 Dinarids and/or specific climatic conditions. For example, a fully grown forest was present

406 during the LGM in Slovenia (Tzedakis, 2004). Another hypothesis suggests a possible species

407 interaction with *E. europaeus* and, consequently, a local parapatic evolutionary process (see

408 above).

409

410 **Conclusions**

411   The most pronounced pattern within our dataset was represented by the differentiation of

412   populations at the edges of the recent range. This study indicates that peripatry and parapatry

413   might be not only limiting factors to range expansion, but may also provide strong

414   microevolutionary forces that shape the genetic structure of the species. Traditionally, population

415   differentiation in temperate species has been ascribed to allopatric processes during glacial

416   periods. Here, we provide an alternative example, showing that population differentiation may

417   occur not only during the glacial restriction of the range into refugia, but also during the phase of

418   interglacial range expansion. During the interglacial periods, peripatric modes of evolution for

419   many taxa might be more frequent due to the rise in sea level, which cause increased shelf island

420   formation. Parapatric processes are accelerated by the range expansions of lineages isolated in

421   the refugia and by the formation of potential suture zones.

422   Recent population of *E. roumanicus* in post-refugial area at Balkan Peninsula and adjacent

423   regions shows differentiation with particular subpopulation located in areas of interglacial

424   steppic refugia and glacial forest refugia. Detailed habitat model and past niche modelling will

425   be necessary to infer, which type of habitat was crucial for survival of *E. roumanicus* during

426   glacial periods. However, certain degree of admixture among particular clusters is obvious,

427   mirroring complex geomorphology and shape of geographical barriers, as well as altitudinal

428   heterogenetity within Balkan Peninsula. The situation might be further complicated by the fact,

429   that a mosaic of particular biomes often existed in refugia (Huntley & Birks, 1983), increasing

430   the proportion of ecotones, and the biotas were often "disharmonious", i.e. particular ecosystems

431   often contain species that inhabit different habitats today (Tyrberg, 1991).

432

433   ***Acknowledgement***

434   We would like to thank all researchers and institutions which provided samples for this study (all

435   are listed in Supplemental information: Table S1). We thank to P. Matějů for help in laboratory.

436

437   **References**

438   **Amori G, Hutterer R, Kryštufek B, Yigit N, Mitsain G, Palomo LJ. 2016.** *Erinaceus*
439   *roumanicus*. The IUCN Red List of Threatened Species 2016: e.T136344A22325720.
440   DOI: 10.2305/IUCN.UK.2016-3.RLTS.T136344A22325720.en.

441   **Anderson EC, Thompson E. 2002.** A model-based method for identifying species
442   hybrids using multilocus genetic data**.** *Genetics* **160:** 1217-1229.

443

444   **Balkenhol N, Cushman S, Storfer A, Waits L. 2015.** *Landscape genetics: concepts,*
445   *methods, applications*. Hoboken, New Jersey: Wiley-Blackwell.

446   **Bandelt HJ, Forster P, Röhl A. 1999.** Median-Joining Networks for Inferring
447   Intraspecific Phylogenies. *Molecular Biology and Evolution* **16:** 37-48.

448   **Bannikova AA, Lebedev VS, Abramov AV, Rozhnov VV. 2014.** Contrasting
449   evolutionary history of hedgehogs and gymnures (Mammalia: Erinaceomorpha) as
450   inferred from a multigene study. *Biological Journal of Linnean Society* **112:** 499-519.
451   DOI: 10.1111/bij.12299.

452   **Becher SA, Griffiths R. 1997.** Isolation and characterization of six polymorphic
453   microsatellite loci in the European hedgehog *Erinaceus europaeus*. *Molecular Ecology* **6:**
454   89–90. DOI: 10.1046/j.1365-294X.1997.00159.x.

455   **Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F. 1999.** GENETIX v.4.05
456   logiciel sous Windows pour la génétique des populations. *Available at http://kimura.univ-*
457   *montp2.fr/genetix/.*

458   **Blair C, Weigel DE, Balazik M, Keeley AT, Walker FM, Landguth E, Cushman S,**
459   **Murphy M, Waits L, Balkenhol N. 2012.** A simulation-based evaluation of methods for
460   inferring linear barriers to gene flow. *Molecular Ecology Resources* **12:** 822-33. DOI:
461   10.1111/j.1755-0998.2012.03151.x.

462   **Bogdanov AS, Bannikova AA, Pirusskii YM, Formozov NA. 2009.** The first genetic
463   evidence of hybridization between West European and Northern White-breasted
464   hedgehogs (*Erinaceus europaeus* and *E. roumanicus*) in Moscow region. *Biological*
465   *Bulletin* **36:** 647-651. DOI:10.1134/S106235900906017X.

466   **Bolfíková B, Hulva P. 2012.** Microevolution of sympatry: landscape genetics of
467   hedgehogs *Erinaceus europaeus* and *E. roumanicus* in Central Europe. *Heredity* 108: 248-
468   255. DOI: 10.1038/hdy.2011.67.

469   **Bolfíková B, Konečný A, Pfäffle M, Skuballa J. 2013.** Population biology of
470   establishment in New Zealand hedgehogs inferred from genetic and historical data:
471   conflict or compromise? *Molecular Ecology* **22**: 3709-3720. DOI: 10.1111/mec.12331.

472   **Brewer S, Cheddadi R, de Beaulieu JL, Reille M. 2002.** The spread of deciduous
473   Quercus throughout Europe since the last glacial period. *Forest Ecology and Management*
474   **156:** 27-48.

475   **Chen C, Durand E, Forbes F, Francois O. 2007.** Bayesian clustering algorithms
476   ascertaining spatial population structure: a new computer program and a comparison
477   study. *Molecular Ecology Notes* **7:** 747–756. DOI: 10.1111/j.1471-8286.2007.01769.x.

478   **Coyne JA, Orr HA. 2004.** *Speciation*. Sunderland, Massachusetts: Sinauer Associates.

479   **Drummond AJ, Rambaut A. 2007.** BEAST: Bayesian evolutionary analysis by sampling
480   trees. *BMC Evolutionary Biology* **7:** 214. DOI: 10.1186/1471-2148-7-214.

481   **Drummond AJ, Rambaut A, Shapiro B, Pybus OG. 2005.** Bayesian Coalescent
482   Inference of Past Population Dynamics from Molecular Sequences. *Molecular Biology*
483   *and Evolution* **22:** 1185-1192. DOI: 10.1093/molbev/msi103.

484   **Durand E, Jay F, Gaggiotti OE, François O. 2009.** Spatial inference of admixture
485   proportions and secondary contact zones. *Molecular Biology and Evolution* **26:** 1963–
486   1973. DOI: 10.1093/molbev/msp106.

487   **Earl DA, VonHoldt BM. 2012.** STRUCTURE HARVESTER: A website and program
488   for visualizing STRUCTURE output and implementing the Evanno method. *Conservation*
489   *Genetic Resources* **4:** 359-361. DOI: 10.1007/s12686-011-9548-7.

490   **ESRI ArcGIS Desktop: Release 10. 2014.** Redlands, CA: Environmental Systems
491   Research Institute.

492   **Evanno G, Regnaut S, Goudet J. 2005.** Detecting the number of clusters of individuals
493   using the software STRUCTURE: A simulation study. *Molecular Ecology* **14:** 2611-2620.
494   DOI: 10.1111/j.1365-294X.2005.02553.x.

495   **Excoffier L, Ray N. 2008.** Surfing during population expansions promotes genetic
496   revolutions and structuration. *Trends in Ecology & Evolution* **23:** 347-351. DOI:
497   10.1016/j.tree.2008.04.004.

498 **Excoffier L, Foll M, Petit RJ. 2009.** Genetic Consequences of Range Expansions.
499 *Annual Review of Ecology, Evolution, and Systematics* **40:** 481-501. DOI:
500 10.1146/annurev.ecolsys.39.110707.173414.

501 **Fijarczyk A, Nadachowska K, Hofman S, Litvinchuk SN, Babik W, Stuglik M,**
502 **Gollmann G, Choleva L, Cogălniceanu D, Vukov T, Džukić G, Szymura JM. 2011.**
503 Nuclear and mitochondrial phylogeography of the European fire-bellied toads *Bombina*
504 *bombina* and *Bombina variegata* supports their independent histories. *Molecular Ecology*
505 **20:** 3381-3398. DOI: 10.1111/j.1365-294X.2011.05175.x.

506 **Giagia-Athanasopoulou EB, Markakis G**. **1996.** Multivariate analysis of morphometric
507 characters in the Eastern hedgehog *Erinaceus concolor* from Greece and adjacent areas.
508 *Zeitschrift Fuer Saeugetierkunde* **61:1** 29-139.

509 **Goudet J. 1995.** FSTAT (Version 1.2): A Computer Note Computer Program to Calculate
510 F-Statistics. *Journal of Heredity* **86:** 485-486.

511 **Guillot G, Mortier F, Estoup A. 2005.** GENELAND: A computer package for landscape
512 genetics. *Molecular Ecology Notes* **5:** 712-715. DOI: 10.1111/j.1471-8286.2005.01031.x.

513 **Gvoždík V, Jandzik D, Lymberakis P, Jablonski D, Moravec J. 2010.** Slow worm,
514 *Anguis fragilis* (Reptilia: Anguidae) as a species complex: Genetic structure reveals deep
515 divergences. *Molecular Phylogenetics and Evolution* **55:** 460-472. DOI:
516 10.1016/j.ympev.2010.01.007.

517 **Henderson M, Becher SA, Doncaster CP, Maclean N. 2000.** Five new polymorphic
518 microsatellite loci in the European hedgehog, *Erinaceus europaeus*. *Molecular Ecology* **9:**
519 1949–1950. DOI: 10.1046/j.1365-294x.2000.01098-18.x.

520 **Hewitt GM. 2000.** The genetic legacy of the Quaternary ice ages. *Nature* **405:** 907-913.
521 DOI: 10.1038/35016000.

522 **Hewitt GM. 2001.** Speciation, hybrid zones and phylogeography - Or seeing genes in
523 space and time. *Molecular Ecology* **10**: 537-549. DOI:
524 10.1046/j.1365294x.2001.01202.x.

525 **Hulva P, Benda P, Hanák V, Evin A, Horáček I. 2007.** New mitochondrial lineages
526 within the *Pipistrellus pipistrellus* complex from Mediterranean Europe. *Folia Zoologica*
527 **56:** 378-388.

528 **Hulva P, Fornůsková A, Chudárková A, Evin A, Allegrini B, Benda P, Bryja J. 2010**.
529 Mechanisms of radiation in a bat group from the genus *Pipistrellus* inferred by
530 phylogeography, demography and population genetics. *Molecular Ecology* **19:** 5417-
531 5431. DOI: 10.1111/j.1365-294X.2010.04899.x.

532   **Huntley B. 1999.** Climatic change and reconstruction. Journal of Quarternary Science **14:**
533   513-520. DOI: 10.1002/(SICI)1099-1417(199910)14:6<513::AID-JQS486>3.0.CO;2-E.

534   **Huntley B, Birks HJB. 1983.** *An atlas of past and present pollen maps for Europe: 0–*
535   *13000 years ago*. Cambridge: Cambridge University Press.

536   **Jarman MR. 1996.** Human Influence in the Development Of the Cretan Mammalian
537   Fauna. In: Reese DS, ed. *Pleistocene and Holocene Fauna of Crete and Its First Settlers.*
538   Madison: Prehistoric Press.

539   **Jensen AB. 2004.** Overwintering of European hedgehogs *Erinaceus europaeus* in a
540   Danish rural area. *Acta Theriologica* **49:** 145–155. DOI: 10.1007/BF03192516.

541   **Joger U, Fritz U, Guicking D, Kalyabina-Hauf S, Nagy ZT, Wink M. 2007.**
542   Phylogeography of western Palaearctic reptiles–Spatial and temporal speciation patterns.
543   *Zoologischer Anzeiger - A J. of Comp. Zool.* **246:** 293-313. DOI:
544   10.1016/j.jcz.2007.09.002.

545   **Katoh K, Standley DM. 2013.** MAFFT Multiple Sequence Alignment Software, Version
546   7: Improvements in Performance and Usability. *Molecular Biology and   Evolution* **30:**
547   772-780. DOI: 10.1093/molbev/mst010.

548   **Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S,**
549   **Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A.**
550   **(2012).** Geneious Basic: An integrated and extendable desktop software platform for the
551   organization and analysis of sequence data. *Bioinformatics* **28:** 1647-1649. DOI:
552   10.1093/bioinformatics/bts199.

553   **King RA, Ferris C. 1998.** Chloroplast DNA phylogeography of *Alnus glutinosa* (L.)
554   Gaertn. *Molecular Ecology* **7:** 1151-1161. DOI: 10.1046/j.1365-294x.1998.00432.x.

555   **Klopfstein S, Currat M, Excoffier L. 2006.** The fate of mutations surfing on the wave of
556   a range expansion. *Molecular Biology and Evolution* **23:** 482-490. DOI:
557   10.1093/molbev/msj057.

558   **Kryštufek B, Reed JM. 2004.** Pattern and process in Balkan biodiversity - an overview.
559   In: Griffiths HI, ed. *Balkan Biodiversity: Pattern and Process in the European Hotspot*.
560   Springer Netherlands, 1-8. DOI: 10.1007/978-1-4020-2854-0_1.

561   **Kryštufek B, Tvrtković N, Paunović M, Ozkan B. 2009.** Size variation in the Northern
562   white-breasted hedgehog *Erinaceus roumanicus*: Latitudinal cline and the island rule.
563   *Mammalia* **73:** 299-306. DOI: 10.1515/MAMM.2009.055.

564   **Lax E, Strasser TF. 1992.** Early Holocene Extinctions on Crete; the Search for the
565   Cause. *Journal of Mediterranean Archaeology* **5:** 203-224. DOI: 10.1558/jmea.v5i1.203.

566   **Mantel N. 1967.** The detection of disease clustering and a generalized regression
567   approach. *Cancer Research* **27:** 209-220.

568   **Miller MP. 2005.** Alleles In Space (AIS): Computer software for the joint analysis of
569   interindividual spatial and genetic information. *Journal of Heredity* **96:** 722-724. DOI:
570   10.1093/jhered/esi119.

571   **Posada D, Crandall K. 1998.** MODELTEST: testing the model of DNA substitution.
572   *Bioinformatics* **14:** 817-818. DOI: 10.1093/bioinformatics/14.9.817.

573   **Pritchard JK, Stephens M, Donnely PJ. 2000.** Inference of population structure using
574   multilocus genotype data. *Genetics* **155:** 945-959.

575   **Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014.** Tracer v1.6. *Available from:*
576   *http://beast.bio.ed.ac.uk/Tracer.*

577   **Ramírez-Soriano A, Ramos-Onsins SE, Rozas J, Calafell F, Navarro A. 2008.**
578   Statistical power analysis of neutrality tests under demographic expansions, contractions
579   and bottlenecks with recombination. *Genetics* **179:** 555-567. DOI:
580   10.1534/genetics.107.083006.

581   **Ramos-Onsins SE, Rozas J. 2002.** Statistical properties of new neutrality tests against
582   population growth. *Molecular Biology and Evolution* **19:** 2092-2100. DOI:
583   10.1093/molbev/msl052.

584   **Rautio A, Valtonen A, Auttila M, Kunnasranta M. 2014.** Nesting patterns of European
585   hedgehogs (*Erinaceus europaeus*) under northern conditions. *Acta Theriologica* **59:**173–
586   181. DOI: 10.1007/s13364-013-0150-0.

587   **Rosenberg N. 2004.** Distruct: a Program for the graphical display of population structure.
588   *Molecular Ecology Notes* **4:** 137-138. DOI: 10.1046/j.1471-8286.2003.00566.x.

589   **Rousset F. 2008.** GENEPOP'007: A complete re-implementation of the GENEPOP
590   software for Windows and Linux. *Molecular Ecology Resources* **8:** 103-106. DOI:
591   10.1111/j.1471-8286.2007.01931.x.

592   **Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R. 2003.** DnaSP, DNA
593   polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19:** 2496-
594   2497. DOI: 10.1093/bioinformatics/btg359.

595   **Říčanová Š, Bryja J, Cosson JF, Gedeon C, Choleva L, Ambros M, Sedláček F. 2011.**
596   Depleted genetic variation of the European ground squirrel in Central Europe in both
597   microsatellites and the major histocompatibility complex gene: implication for
598   conservation. *Conservation Genetics* **12:** 1115-1129. DOI: 10.1007/s10592-011-0213-1.

599 **Santucci F, Emerson BC, Hewitt GM. 1998.** Mitochondrial DNA phylogeography of
600 European hedgehogs. *Molecular Ecology* **7:** 1163-1172. DOI: 10.1046/j.1365-
601 294x.1998.00436.x.

602 **Schaschl H, Lymberakis P, Suchentruck F. 2002.** On allozyme and cyt-b gene
603 characteristics of Cretan hedgehogs, *Erinaceus consolor nesiotes* Bate, 1906. *Mammalian*
604 *Biology* **67:** 257-267. DOI:10.1078/1616-5047-00040.

605 **Schule W. 1993.** Mammals, vegetation and the initial human settlement of the
606 Mediterranean islands: a palaeoecological approach. *Journal of Biogeography* **20:** 399-
607 412. DOI: 10.2307/2845588.

608 **Seddon JM, Santucci F, Reeve NJ, Hewitt GM. 2001.** DNA footprints of European
609 hedgehogs, *Erinaceus europaeus* and *E. concolor*: Pleistocene refugia, postglacial
610 expansion and colonization routes. *Molecular Ecology* **10:** 2187-2198. DOI:
611 10.1046/j.0962-1083.2001.01357.x.

612 **Seddon JM, Santucci F, Reeve N, Hewitt GM. 2002.** Caucasus Mountains divide
613 postulated postglacial colonization routes in the white-breasted hedgehog, *Erinaceus*
614 *concolor*. *Journal of Evolutionary Biology* **15:** 463-467. DOI: 10.1046/j.1420-
615 9101.2002.00408.x.

616 **Stewart JR, Lister AM, Barnes I, Dalén L. 2010.** Refugia revisited: individualistic
617 responses of species in space and time. *Proceedings of the Royal Society* **277:** 661-6. DOI:
618 10.1098/rspb.2009.1272.

619 **Tyrberg T. 1991.** Arctic, montane and steppe birds as glacial relicts in the west palearctic.
620 *Ornithologische Verhandlungen* **25:** 29-49.

621 **Tzedakis PC. 2004.** The Balkans as prime glacial refugial territory of European temperate
622 trees. In: Griffiths HI, ed. *Balkan Biodiversity: Pattern and Process in the European*
623 *Hotspot*. Springer Netherlands, 1-8. DOI: 10.1007/978-1-4020-2854-0_1.

624 **Van der Geer A, Lyras G, de Vos J, Dermitzakis M. 2010.** Evolution of Island
625 Mammals. London: Wiley-Blackwell.

626 **Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004.** MICRO-
627 CHECKER: S ftware for identifying and correcting genotyping errors in microsatellite
628 data. *Molecular Ecology Notes* **4:** 535-538. DOI: 10.1111/j.1471-8286.2004.00684.x.

629

none

The descriptive statistics of microsatellite genetic diversity in seven populations of *Erinaceus roumanicus*, which were recognised by the Bayesian clustering analysis using Structure.

Number of alleles (N_a), allelic richness (A_R), expected heterozygosity (H_E), observed heterozygosity (H_O), inbreeding coefficient (F_{IS}) are given for each population.

	N_a	A_R	H_E	H_O	F_{IS}
Czech Republic	7	6.187	0.7120	0.6151	0.134
Slovakia, Hungary	9	7.318	0.7032	0.6291	0.097
Romania	9	6.575	0.7521	0.6171	0.242
Slovenia	8	6.317	0.6844	0.6071	0.102
Croatia, Bosnia and Herzegovina	9	7.045	0.7125	0.6442	0.114
Bulgaria, Greece, Macedonia, Serbia, Montenegro	10	7.912	0.7628	0.6923	0.102
Crete	6	4.994	0.5243	0.4292	0.167

The summary statistics of the partial control region of mitochondrial DNA in six populations of *Erinaceus roumanicus*, which were recognised by Bayesian analysis using Geneland.

Colours in bracelets correspond to colours used in Fig. 5c. Number of individuals (N), number of haplotypes (N_h), haplotype diversity (H_d), number of polymorphic sites (N_p), nucleotide polymorphism (π), Tajima's D (D), Fu's Fs (F_S) and R^2 are given for each population and for whole dataset. The significant values tests are shown with asterisks (*$p<0.05$, **$p<0.01$, ***$p<0.001$)

	n	N_h	H_d	N_p	π	D	R^2	F_S
Central Europe (orange)	58	6	0.3134	10	0.00204	-0.92848	0.0688	-0.561
Carpathians (yellow)	41	10	0.8269	9	0.00292	-0.81891	0.0874	-2.779*
North Balcan (pink)	63	15	0.6912	17	0.00272	-1.35223	0.0534	-7.052***
Black sea cost (violet)	23	9	0.7787	14	0.00548	-1.25805	0.0817*	-2.731
Dinarids (green)	59	17	0.8954	19	0.00568	-0.90052	0.0730	-5.470**
South Balkan (blue)	52	11	0.3808	18	0.00229	-1.82115*	0.0492*	-5.937*
all data	296	53	0.9185	38	0.00644	-1.38337	0.0415	-3.638***

Sampling localities for the 314 samples of the Northern White-breasted Hedgehog (*Erinaceus roumanicus*) in central Europe and the Balkan region.

Each sampled individual is represented by a pink dot. The distribution map of the species is shown in the smaller square. Map was done using ArcGIS (ESRI 2017) using publicly available layers (Bathymetry - EMODnet, Administrative areas - GADM database, Elevation - WorldClim database, Rivers - EEA hydrographic data set). Inset of the figure is based on an IUCN distribution map of species (Amori et al., 2016) edited for the purposes of this article in Photoshop CS3 (Adobe).

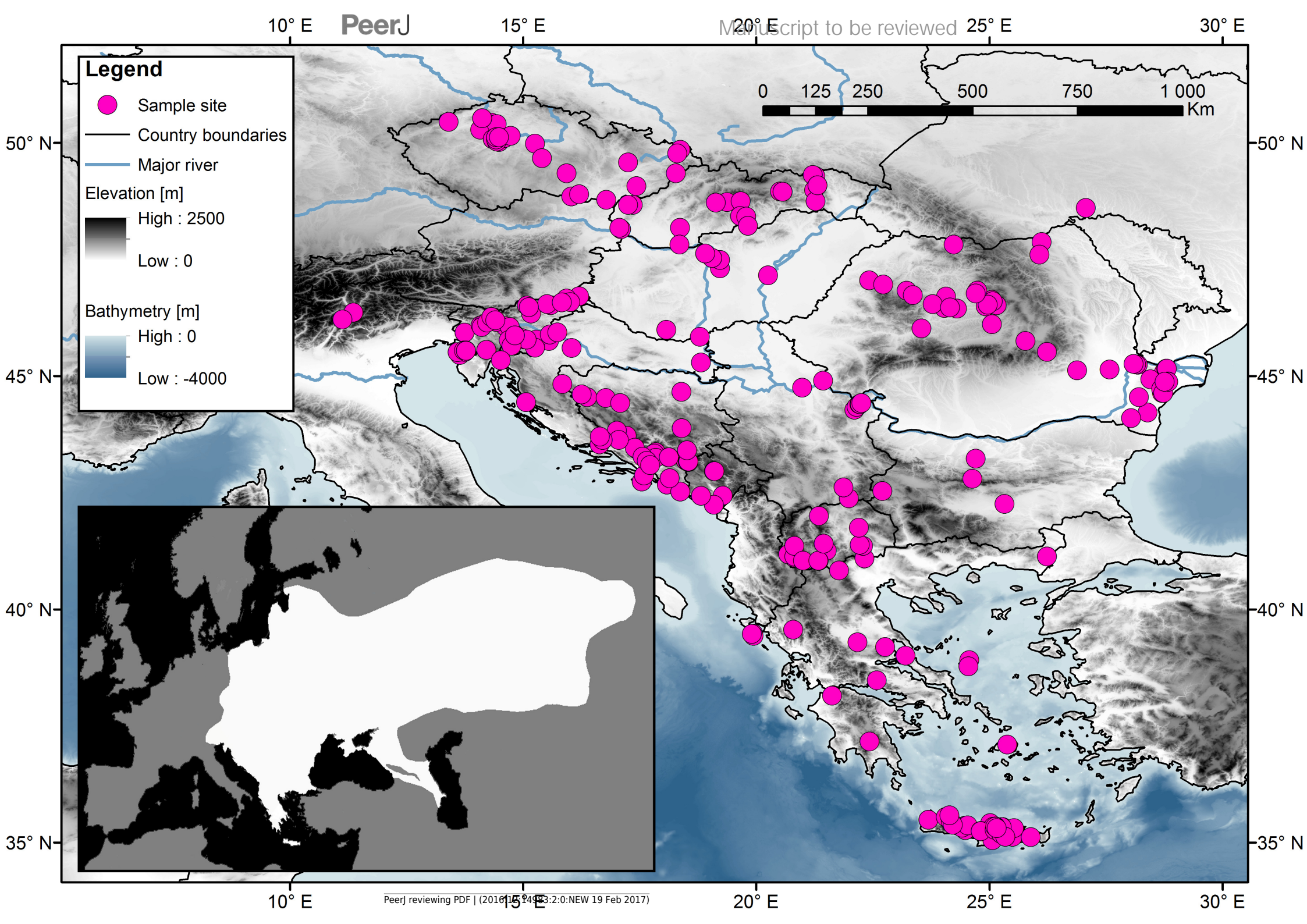
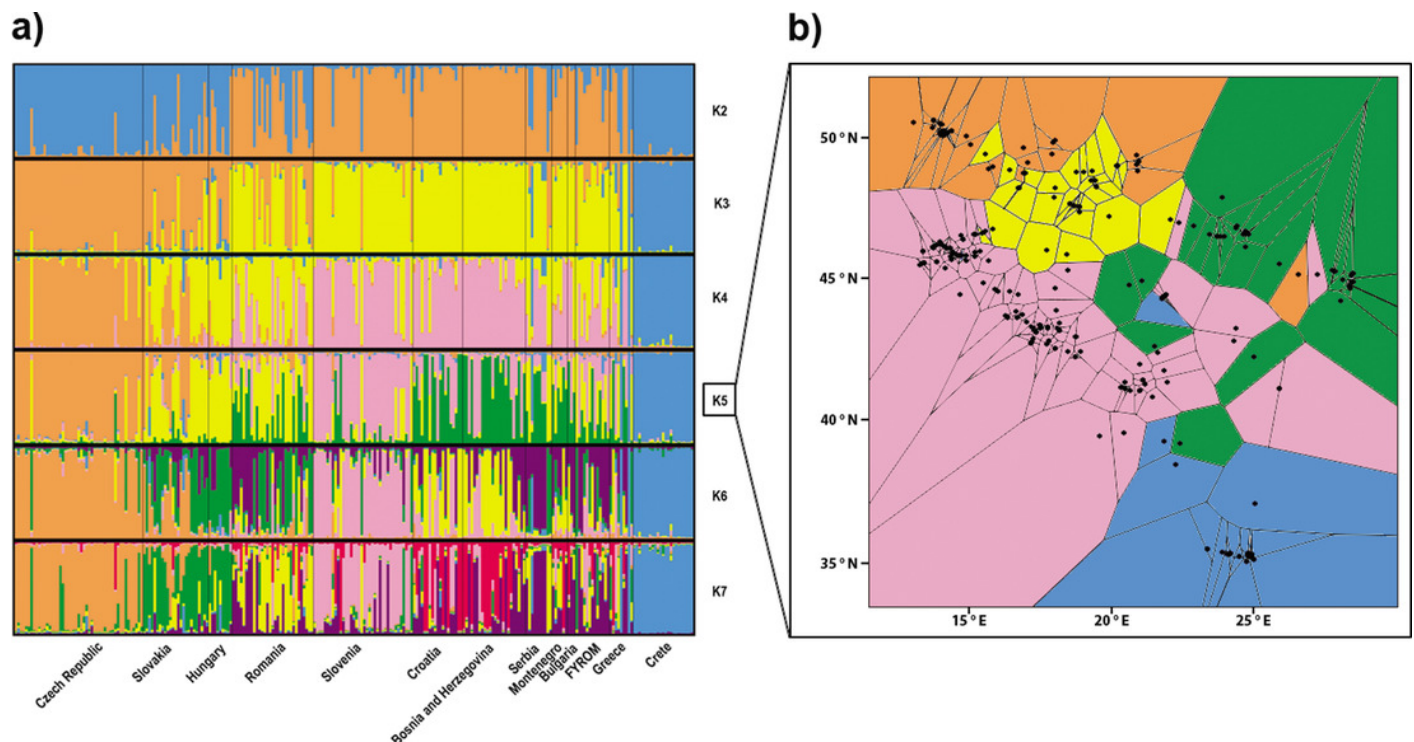



Figure 2

Structuring of the nuclear data

(a) Population assignment using Bayesian clustering analysis in Structure for the 260 *Erinaceus roumanicus* samples in Central Europe and the Balkan region. The results are shown for $K = 2$ to $K = 7$. The highest ΔK was for $K = 3$, and the second highest ΔK was for $K = 7$. (b) Geographical cluster assignments implemented by TESS. The results display five colours for Kmax = 6.

Median-joining network of the mitochondrial control region haplotypes for 296 *Erinaceus roumanicus* samples.

Haplotypes are represented by circles, which are proportional to the haplotype frequency. The hypothesized haplotypes are represented by red dots. Colour codes are used according to the population assignments of the samples using mtDNA data from the Geneland analysis (see Fig. 5c or Supporting information: Table S1). Lines between nodes represent a nucleotide changes.

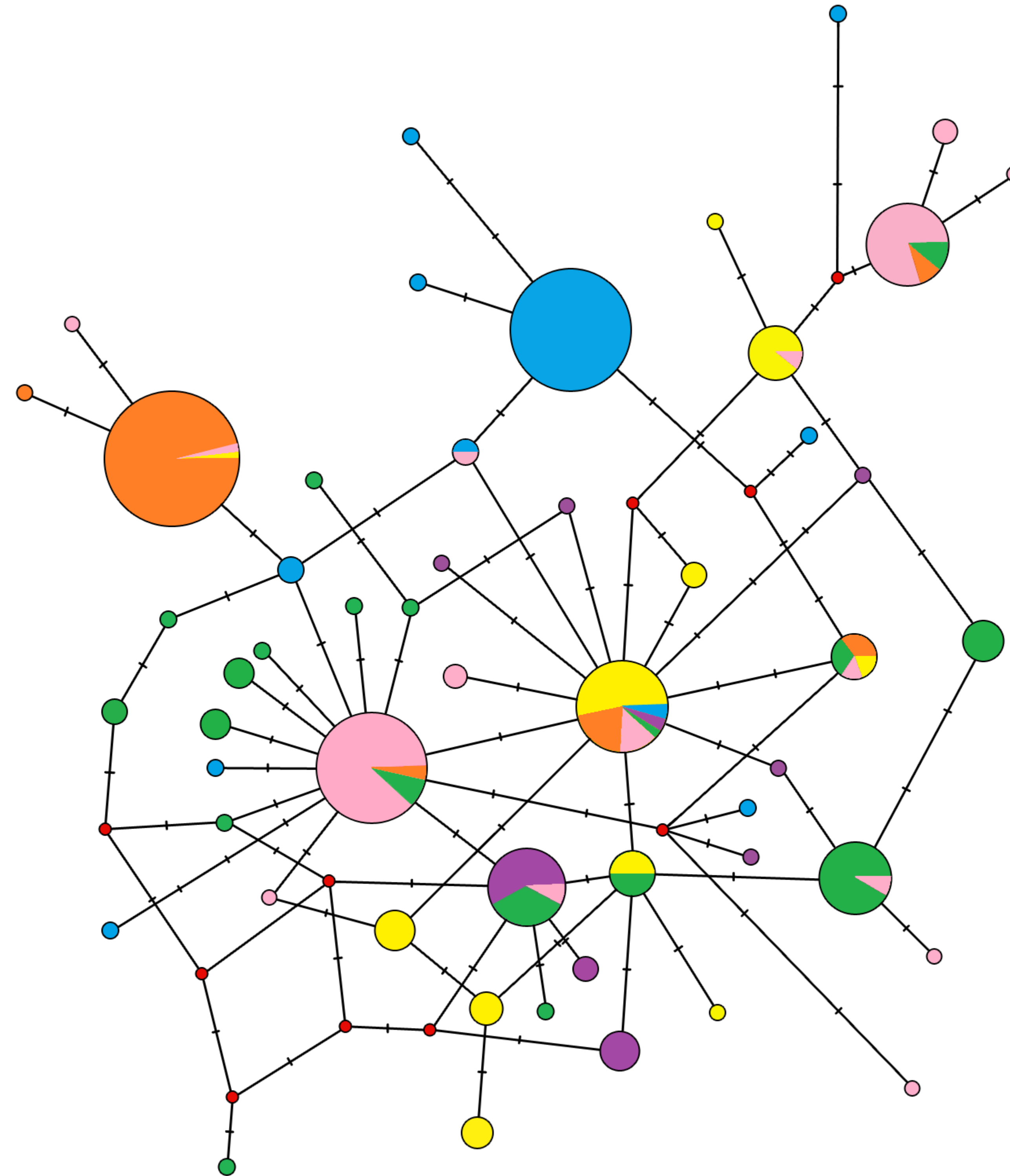



Figure 4

Bayesian skyline plot based on mtDNA of 296 Erinaceus roumanicus samples.

The X axis represents the time in mutation units per nucleotide position. The Y axis represents a correlate of the population size (Ne). The black line shows the median of the Ne estimation, and the 95% confidence intervals are indicated by the blue areas.

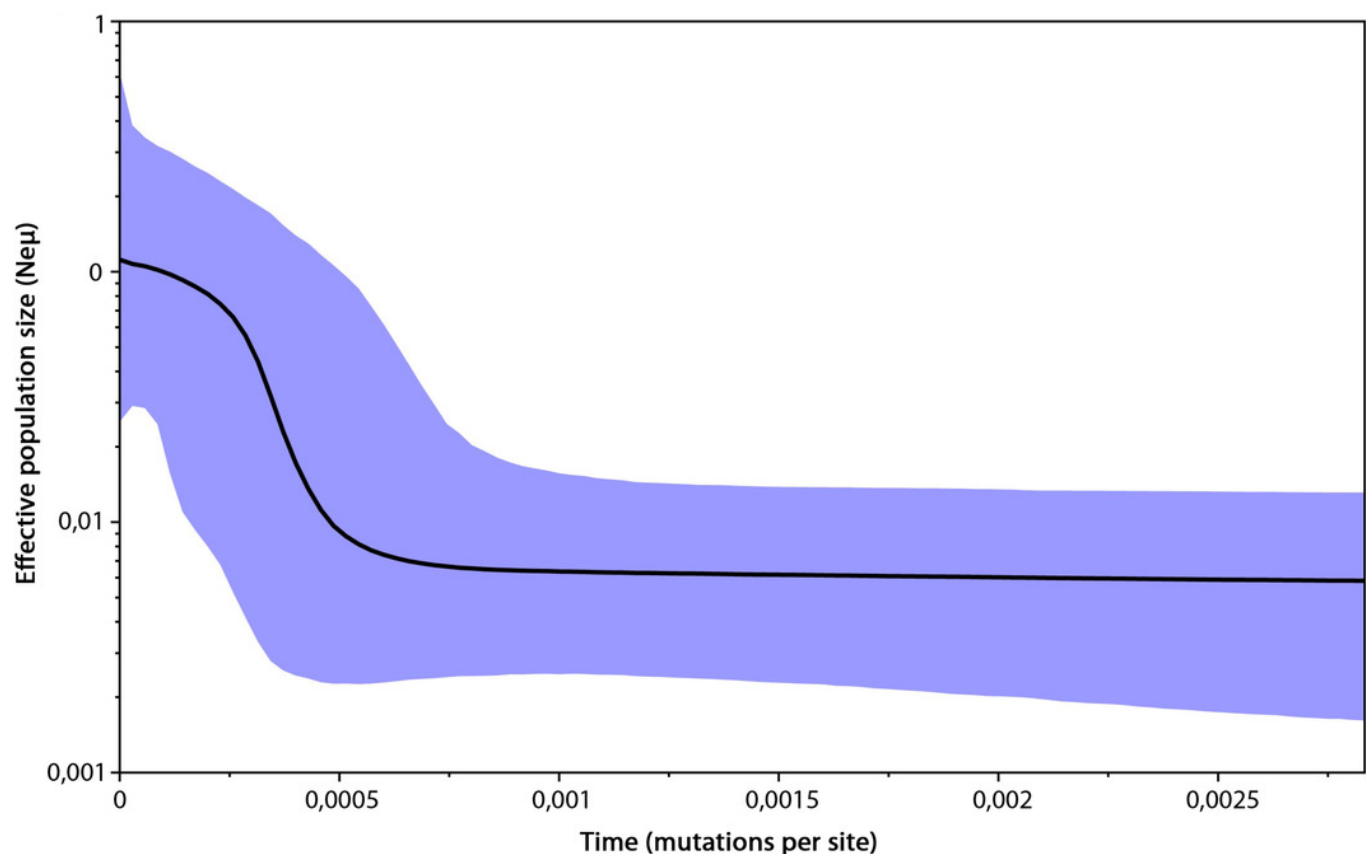



Figure 5

Genetic landscape shape interpolation analysis.

Genetic landscape shape interpolation analysis based on variation of (a) the control region of mtDNA using 296 Erinaceus roumanicus samples and (b) nine nuclear microsatellite loci using 260 E. roumanicus samples. The X and Y axes reflect the geographic coordinates within the study area, while the Z axis represents the average genetic distances between the analysed samples. The Bayesian landscape genetic analysis using Geneland illustrates the spatial distribution of the populations using (c) the control region of mtDNA and (d) nine nuclear microsatellite loci. Figure was done using publicly available outlines (https://CRAN.R-project.org/ package=rworldmap) and edited in Photoshop CS3 (Adobe).

a)

South/North axis Y
(eastern edge)

West/East axis X (southern edge)

Genetic distances (Z axis)

b)

South/North axis Y
(eastern edge)

West/East axis X (southern edge)

Genetic distances (Z axis)

c)

d)