1   **Distinctive insular forms of threespine stickleback (*Gasterosteus***

2   ***aculeatus*) from western Mediterranean islands**

3

4   Kay Lucek[1,2,3*] & Ole Seehausen[1,2]

5

6   [1] Aquatic Ecology and Evolution, Institute of Ecology & Evolution, University of

7   Bern, Baltzerstrasse 6, CH-3012 Bern, Switzerland

8   [2] Department of Fish Ecology and Evolution, EAWAG Swiss Federal Institute of

9   Aquatic Science and Technology, Center for Ecology, Evolution and

10   Biogeochemistry, CH-6047 Kastanienbaum, Switzerland

11   [3] Current address: Department of Animal and Plant Sciences, University of

12   Sheffield, Sheffield S10 2TN, United Kingdom

13

14   * kay.lucek@eawag.ch, telephone: 0041316313016

15

16 **Abstract**

17 Neutral and adaptive variation among populations within a species is a major

18 component of biological diversity and may be pronounced among insular

19 populations due to geographical isolation and island specific evolutionary forces

20 at work. Detecting and preserving potential evolutionary significant units below

21 the species rank has become a crucial task for conservation biology. Combining

22 genetic, phenotypic and ecological data, we investigated evolutionary patterns

23 among the enigmatic threespine stickleback populations from western

24 Mediterranean islands, all of which are threatened by habitat deterioration and

25 climate change. We find indications that these populations derive from different

26 genetic lineages, being genetically highly distinct from the stickleback of

27 mainland Europe and the northern Atlantic as well as from each other.

28 Mediterranean island stickleback populations are also phenotypically distinct

29 from mainland populations but interestingly stickleback from Iceland have

30 converged on a similar phenotype. This distinctive island stickleback phenotype

31 seems to be driven by distinct selective regimes on islands versus continents.

32 Overall, our results reveal the status of western Mediterranean island stickleback

33 as evolutionarily distinct units, important for conservation of biodiversity.

34

35 Keywords: Island rule, *Gasterosteus aculeatus*, glacial refugium, Mediterranean

36 biota

37

38

39

Introduction

Species and even entire ecosystems are currently threatened by an array of anthropogenic changes to the environment, including global warming (Pereira et al. 2010), eutrophication (Seehausen et al. 1997; Taylor et al. 2006; Vonlanthen et al. 2012) or the introduction of nonnative species (Elton 1958; Lockwood et al. 2007). As a result, many endemic species and distinct populations with small geographic ranges have either become extinct, experienced severe genetic bottlenecks that threaten their ability to adapt to changing environments (Lande 1998), or become genetically homogenized with introduced populations of related species (Olden et al. 2004). In particular, insular ecosystems, which in many cases harbor the phenotypically and genetically most distinctive populations within otherwise widely distributed species (Foster 1964; Case 1978; Whittaker and Fernández-Palacios 2007), are vulnerable to climatic and human-induced perturbations (Ricketts et al. 2005; Courchamp et al. 2014). Freshwater organisms on Mediterranean islands belong to the most threatened species, but assessment of genetic and phenotypic distinctiveness is in most cases still lacking (Kottelat and Freyhof 2007; Cuttelod et al. 2008; Geiger et al. 2014).

Islands may generally harbor genetically and phenotypically distinct populations more often than mainland sites, following the "island rule" (Foster 1964; Case 1978; Lomolino 1985). The latter suggests that insular populations of small vertebrates evolve an increased body size, whereas large vertebrates evolve dwarf populations on islands, potentially due to partial release from competitors and predators, often leading to further phenotypic adaptation (Case 1978;

3

65   Lomolino et al. 2012). Albeit the island rule has been validated in many

66   vertebrate taxa, examples from freshwater fish are relatively rare (Lomolino

67   2005, but see MacColl et al. 2013).

68

69   The threespine stickleback (*Gasterosteus aculeatus* species complex) is an

70   ancestrally marine species with a circumpolar distribution in the Northern

71   Hemisphere. Throughout its range, it has repeatedly colonized freshwater

72   systems, and most extant freshwater populations are young, having been formed

73   after the last glacial maximum about 12,000 years ago (McKinnon and Rundle

74   2002). In Europe, freshwater stickleback exist from Iceland down to southern

75   Italy and Greece, having been relatively widespread in freshwaters along the

76   Mediterranean coast (e.g. Crivelli and Britton 1987; Lobón-Cerviá et al. 1988;

77   Mäkinen et al. 2006; Mäkinen and Merilä 2008; Araguas et al. 2012) and

78   reaching as far south as Syria (Krupp 1985). The repeated evolution of

79   freshwater populations from marine ancestors was often facilitated by the

80   maintenance of freshwater-adapted alleles at low frequencies in the ancestral

81   marine gene pool due to some ongoing gene flow from freshwater adapted

82   populations to the ancestral marine populations (Schluter and Conte 2009; Jones

83   et al. 2012). Consequently marine populations harbor a large standing genetic

84   variation, upon which selection can rapidly act during the colonization of new

85   freshwater bodies. However, marine stickleback do not occur farther south in the

86   Atlantic ocean than the Bay of Biscay and are absent from the Mediterranean Sea

87   (Munzing 1963; Mäkinen et al. 2006). Freshwater stickleback populations from

88   southern Europe are thus geographically strongly isolated from marine

89   stickleback and cannot have exchanged genes for a significant length of time.

90   Despite their patchy but geographically wide distribution in the Mediterranean

91   bioregion, only three Mediterranean islands are known to harbor freshwater

92   stickleback, all in the western Mediterranean: Mallorca (Riera 1980; Lobón-

93   Cerviá et al. 1988), Corsica (Gauthier and Rose 1974) and Sardinia (Bertin 1925;

94   Munzing 1963; Bianco 1980; Kottelat and Freyhof 2007; Orru et al. 2010).

95

96   The absence of a contemporary marine stickleback population has rendered the

97   current biogeographic distribution of Mediterranean freshwater stickleback

98   somewhat of an enigma (Bertin 1925; Munzing 1963). Some have suggested that

99   the existing populations may represent glacial relics, where marine fish from

100 the Atlantic Ocean had occurred in the Mediterranean Sea and colonized

101 Mediterranean freshwater sites from there during the last glacial maximum

102 when water temperatures and salinities were lower (Munzing 1963). Genetic

103 analyses of mainland populations imply that Mediterranean freshwater

104 stickleback may indeed be remnants of several glacial refugia (Mäkinen et al.

105 2006). The absence of potential gene flow through the Mediterranean Sea among

106 these freshwater mainland populations has moreover led to high genetic

107 differentiation amongst them, even between geographically close populations

108 (Cano et al. 2008; Mäkinen and Merilä 2008; Araguas et al. 2012; DeFaveri et al.

109 2012).

110

111 Albeit stickleback are listed as a species of least concern for conservation by the

112 IUCN, they are locally threatened, especially in the Mediterranean region

113 (Araguas et al. 2012, IUCN 2015). Recent conservation-related assessments of

114 mainland Mediterranean stickleback suggested that many of the contemporary

populations represent distinct evolutionary significant units, i.e. forming unique

adaptive groups due to their genetic and ecological distinctiveness (Ryder 1986;

Moritz 1994). They differ from other European populations genetically and

phenotypically (Mäkinen et al. 2006; Cano et al. 2008; Mäkinen and Merilä 2008;

DeFaveri et al. 2012) as well as on a few occasions in their life history (Crivelli

and Britton 1987; Clavero et al. 2009). Many of the populations have declined or

even become extinct since the beginning of the 20th century. The remaining

populations are currently threatened by an array of factors, including habitat

fragmentation, the release of non-native fish species or anthropogenic changes of

the flow regime due to the withdrawal of water for agricultural irrigation

(Clavero et al. 2009; Orru et al. 2010; Araguas et al. 2012). Moreover, the

Mediterranean climate itself with its alternate dry and wet periods and irregular

flash floods may negatively affect the remaining populations (Gasith and Resh

1999), with climate change likely to result in extended dry periods with often

very low residual water flow (Palmer and Räisänen 2002).

Much of the current knowledge on the genetic and phenotypic distinctiveness of

Mediterranean stickleback is based on mainland populations, whereas island

populations are understudied. Using microsatellite markers, we first infer the

genetic relationships between populations from the three Mediterranean islands

where stickleback were known to occur in relation to other mainly mainland

populations from across Europe. Based on mitochondrial DNA, we further assess

if the observed population genetic structure may be a result of different

colonization histories and/or glacial refugias. Subsequently, we test to which

degree these Mediterranean island populations may differ in their ecology and

140   phenotype from their mainland counterparts and moreover compare them to an

141   island population in northern Europe (Iceland).

142

143   **Materials & Methods**

144   Stickleback were collected from all Mediterranean islands where their presence

145   had previously been indicated, i.e. Corsica (Gauthier and Rose 1974), Mallorca

146   (Lobón-Cerviá et al. 1988) and Sardinia (Bianco 1980). To assess the degree of

147   genetic and phenotypic distinctiveness of these Mediterranean island stickleback

148   in relation to other European populations, ten populations from across Europe

149   were included (Figure 1, Table 1). Individuals were collected using minnow traps

150   and hand nets. In all cases, fish were sacrificed with an overdose of clove oil and

151   preserved in 70% ethanol. A fin clip was additionally taken for genetic analysis

152   and preserved in absolute ethanol. Our studied populations originate from

153   different freshwater habitats (i.e. lakes and streams) in order to cover much of

154   the ecological variation among European freshwater stickleback to which the

155   Mediterranean populations can be compared. In addition, one marine population

156   from Iceland, representing a putative ancestral state of all European freshwater

157   stickleback was included (Orti et al. 1994; Mäkinen et al. 2006). Populations

158   were further assigned to belong either to northern or southern Europe,

159   depending on whether the respective catchment was in reach of the

160   contemporary population of marine stickleback (populations A-F in Figure 1) or

161   not (populations G-M). Because only subadult individuals were caught in

162   Sardinia, which differ phenotypically from adults (Bell 1981), the respective

163   phenotypic data was collected from adult museum specimens that were collected

164   from a different site in the same catchment ~20 km away (natural history

165   museum in Vienna ID: NHW-83118). Due to formaldehyde preservation, none of

166   the museum specimens were available for DNA extraction.

167

168   *Morphology*

169   Sixteen linear morphological traits that are known to be linked or involved in

170   ecological diversification of stickleback (e.g. Kristjánsson et al. 2002; Mori and

171   Takamura 2004; Berner et al. 2008) were measured to the nearest 0.01 mm

172   using a digital caliper. These traits were either related to anti-predator defense

173   (length of the first and second dorsal spine, the pelvic spine and the pelvic girdle),

174   feeding ecology (head length, upper jaw length, snout length, eye diameter) or to

175   general body shape and swimming performance (standard length, width of the

176   pelvic girdle, body depth measured after the first and second dorsal spine, caudal

177   peduncle length, basal length of the anal and dorsal fin, total length of the pelvic

178   fin). Because all linear traits were significantly correlated with standard length

179   (results not shown), a size correction was applied by using the residuals of a

180   regression of each trait against standard length, pooling all individuals and

181   populations. In addition to the linear morphological measurement, the number

182   of lateral plates was counted for each individual on its left flank.

183

184   To estimate the overall phenotypic differentiation among individuals, a principal

185   component analysis (PCA) was conducted combining all size-corrected linear

186   measurements. A second PCA was performed using only size-corrected defense-

187   related linear measurements, as these traits commonly differ between

188   stickleback from northern and southern Europe (e.g. Gross 1978; Lobón-Cerviá

189   et al. 1988; Cano et al. 2008). To subsequently test if *geography* (northern or

8

190   southern Europe) or *insularity* (island or mainland) may account for the

191   observed phenotypic variation among freshwater populations along the two

192   leading PC axes of either analysis, a linear mixed model was employed using the

193   population identity as a random factor. The best fitting model was determined

194   using a backward elimination procedure starting with a model including both

195   factors. Given the low levels of statistical replication for each factor (Table S1),

196   no interaction was tested. Similarly, both the number of lateral plates and

197   standard length were tested for potential effects of *geography* or *insularity*

198   among freshwater populations. All statistical analyses were performed in R

199   2.15.1 (R Core Team, 2012).

200

201   *Feeding ecology*

202   To study potential trophic specialization among Mediterranean stickleback, the

203   stomachs of 214 adult individuals from eight populations (Corsica, Mallorca,

204   Galicia, France Stream 1, Geneva Lake and Stream, Constance Lake and Stream)

205   were extracted and all food items were counted using a dissection microscope.

206   Due to access restrictions, stomachs could not be extracted for the Sardinian

207   museum specimens. Food items were assigned to the following taxonomic

208   classes: *Amphipoda, Isopoda, Cladocera, Copepoda, Ostracoda, Decapoda, Diptera*

209   imagos, *Chironomidae, Ephemeroptera, Trichoptera, Odonata, Acaria, Lumbricidae,*

210   *Gastropoda,* parasites and stickleback fry and eggs. To test for feeding-related

211   divergence along the limnetic-benthic feeding axis, the percentage of planktonic

212   prey (%PPP) was then calculated as the fraction of *Cladocera* and *Copepoda* to

213   the total number of all food items present for each individual following Lucek et

214   al. (2012). Similarly the percentage of each prey category per population was

215   calculated to allow a qualitative comparison of feeding habits across populations.

216

217   *Genetic analysis*

218   DNA for all individuals was extracted using a 10% Chelex solution, following the

219   manufacturer's protocol (Biorad, California, USA). Seven species-specific

220   microsatellite markers (Gaest66, Stn26, Stn30, Stn96, Stn130, Stn173 and

221   Stn196) were amplified in one multiplex kit. Detailed information on the

222   multiplexing setup and the PCR protocol can be found in Lucek et al. (2014b).

223   Alleles were visualized on an ABI 3130XL and scored with GENEMAPPER 4.0

224   (Applied Biosystems, Zug, Switzerland).

225

226   Deviation from Hardy-Weinberg equilibrium was calculated using GENODIVE 2.0

227   using 10,000 bootstrap replicates (Meirmans and Van Tienderen 2004). To

228   estimate the genetic variation within a population, a number of genetic indices

229   were calculated in GENODIVE 2.0: the average number of observed alleles (A_N), the

230   effective number of alleles, i.e. the number of alleles weighted for their frequency

231   (A_R), the observed heterozygosity (H_O) and the expected heterozygosity within a

232   population (H_S). The inbreeding coefficient (F_{IS}) was estimated for each

233   population. All genetic indices were subsequently compared between

234   populations from northern and from southern Europe (Table S1) as well as

235   between insular populations from the south, south European mainland

236   populations and populations from northern Europe using Wilcoxon tests. In

237   addition, pairwise genetic differentiation (F_{ST}) among all populations was

238   estimated in GENODIVE using 10,000 bootstrap replicates to assess significance.

239   Pairwise F_{ST} were also compared between populations from northern and

240   southern Europe as well as between the marine population and the freshwater

241   populations using an ANOVA with a Tukey HSD *post hoc* test. Because F_{ST}

242   estimates can be biased towards lower values in cases where the level of

243   heterozygosity within populations is high (Meirmans and Hedrick 2011), the

244   analysis was repeated using F'_{ST} (Hedrick 2005). F'_{ST} was calculated in R with the

245   package DIVERSITY (Keenan et al. 2013), calculating its 95% confidence interval

246   using 10,000 bootstrap replicates. The genetic structuring was estimated using

247   an admixture model implemented in STRUCTURE 2.3.3 (Falush et al. 2007) with

248   30,000 burnin steps followed by 300,000 MCMC steps. The simulation was

249   performed assuming 1-13 genetic clusters (K) with 10 replicates for each

250   assumed K. The optimal number of clusters was determined based on the

251   estimated log likelihood of each run and its variation among runs for the same K,

252   following Evanno et al. (2005). To further infer the genetic relationship among

253   populations, a neighbour joining tree was established based on Cavalli-Sforza

254   distances among populations based on microsatellite allele frequencies.

255   Statistical support for each node of the inferred tree was obtained using a

256   bootstrap procedure with 1000 replicates in PHYLIP 3.69 (Felsenstein 2011).

257   Lastly, a PCA was conducted on population-based allele frequencies in GENODIVE.

258

259   To assess the phylogenetic relationships between the west Mediterranean island

260   populations and other European stickleback, and to estimate the divergence time

261   among different lineages, a fragment of the mitochondrial control region

262   spanning over 426 consecutive base pairs was amplified with primers and

263   protocols described in Mäkinen & Merilä (2008) using one individual from

264   Mallorca and Sardinia and including the previously identified haplotype from

265   Corsica (Lucek et al. 2010). The phylogenetic analyses were performed in Beast

266   2.2.1 (Mäkinen and Merilä 2008; Bouckaert et al. 2014), adding all available

267   stickleback sequences from GenBank, including samples from Europe, Atlantic

268   North America (Mäkinen and Merilä 2008; Lucek et al. 2010) and Pacific North

269   America (California; Richmond et al. 2014). The best model of molecular

270   evolution was determined in Jmodeltest 2.1.7 (Darriba et al. 2012) using a

271   Bayesian Information Criterion (BIC). An uncorrelated lognormal relaxed clock

272   model was selected with a constant size coalescent tree prior. Substitution rate

273   was set to 0.028 substitutions/site/million years following (Aldenhoven et al.

274   2010). The analysis was run over 50 million generations, sampling every 5000

275   generations. Tracer 1.6.0 (Rambaut *et al.* 2014) was then used to examine

276   convergence and mixing based on effective sample sizes (ESS). Lastly,

277   Treeannotator 2.2.1 was used to generate a maximum clade credibility tree

278   using 50% of trees as burnin. In addition, a maximum likelihood tree was

279   estimated in Mega 5 (Tamura et al. 2011) with 1000 bootstrap replicates.

280
281   **Results**

282   *Feeding ecology*

283   Out of the 214 extracted stomachs, 13 were empty (6.1%) and were

284   subsequently excluded from all analyses. The percentage of planktonic prey

285   (%PPP) was highest in the two lake populations from Lake Geneva and Lake

286   Constance, feeding predominantly on copepods in comparison to all stream

287   populations ($F_{1,199}$ = 52.8, $p < 0.001$; Figure 2). Stream-dwelling individuals in

288   contrast fed predominantly on chironomid larvae. In both populations from

12

289   Galicia and the France Stream 1, several individuals had fish eggs or even

290   stickleback fry in their stomachs.

291

292   *Genetics*

293   Deviation from Hardy-Weinberg equilibrium occurred in three out of 91

294   estimates following a sequential Bonferroni correction (STN26: Iceland Lake;

295   STN96: Sardinia and Corsica). Genetic diversity was overall significantly reduced

296   in land-locked south European populations, showing a lower number of alleles

297   (A_N: $W = 6$, $p = 0.045$; A_R: $W = 0$, $p = 0.002$; Table 2) and decreased heterozygosity

298   (H_O: $W = 1$, $p = 0.007$; H_S: $W = 0$, $p = 0.004$). But there was no overall difference in

299   the degree of inbreeding between populations from northern and southern

300   Europe (F_{IS}: $W = 21$, $p = 0.943$). Comparing populations from northern Europe

301   with insular Mediterranean populations, the latter showed a significantly lower

302   effective number of alleles (A_R: $W = 0$, $p = 0.036$) but not observed alleles (A_N: W

303    $= 4$, $p = 0.365$) and decreased heterozygosity (H_O: $W = 0$, $p = 0.036$; H_S: $W = 0$, $p =$

304   0.036), where inbreeding coefficients did not differ (F_{IS}: $W = 10$, $p = 0.571$). None

305   of the genetic indices differed between insular Mediterranean populations and

306   the populations from the south European mainland (all $p > 0.100$).

307

308   Pairwise F_{ST}s (Table S1) and F'_{ST}s (Table S2) were highly correlated ($R^2 = 0.884$;

309    $F_{1,76} = 588.6$, $p < 0.001$), where F'_{ST} values were on average significantly higher

310   than F_{ST} (paired *t*-test: $t_{1,77} = 24.8$, $p < 0.001$). The pairwise comparisons differed

311   overall significantly among the geographic groups for both F_{ST} ($F_{4,73} = 18.1$, $p <$

312   0.001; Figure 3a) and F'_{ST} ($F_{4,73} = 5.6$, $p < 0.001$; Figure 3b). Comparisons

313   involving the marine population yielded significantly lower F_{ST} (average F_{ST}

314 marine vs. north European populations = 0.178 ± 0.016 SE; average F_{ST} marine vs.

315 south European populations = 0.309 ± 0.021 SE) than allopatric comparisons

316 between freshwater populations from north and south Europe (average F_{ST} =

317 0.483 ± 0.013 SE, *post hoc p* < 0.001). They were also lower than comparisons

318 among south European populations (average F_{ST} = 0.579 ± 0.032 SE, *post hoc p* <

319 0.001) but was not statistically different from comparisons among freshwater

320 populations from northern Europe only (average F_{ST} = 0.296 ± 0.060 SE, *post hoc*

321 *p* = 0.960). The pairwise genetic differentiation among south European

322 populations was furthermore significantly higher than all other comparisons (all

323 *p* < 0.018), but did not differ between insular populations and populations from

324 the south European mainland ($t_{1,11}$ = 1.87, *p* = 0.092). The increased level of

325 heterozygosity in the marine population (Table 2) may bias the F_{ST} estimation

326 towards lower values (Hedrick 2005; Meirmans and Hedrick 2011). Using F'_{ST} to

327 correct this bias (Hedrick 2005), the level of genetic differentiation between the

328 marine population and populations from northern Europe (average F'_{ST} = 0.406

329 ± 0.038 SE) was still significantly lower than that among populations from

330 southern Europe (average F'_{ST} = 0.783 ± 0.047 SE, *post hoc p* = 0.005) and

331 between populations from northern and southern Europe (average F'_{ST} = 0.726 ±

332 0.026 SE, *post hoc p* = 0.023). Also the differentiation among south European

333 populations was higher than that among north European populations (average

334 F'_{ST} = 0.505 ± 0.100 SE, *post hoc p* = 0.018). Lastly, F'_{ST} did also not differ between

335 insular populations and populations from the south European mainland ($t_{1,11}$ =

336 1.99, *p* = 0.074).

337

338   Ten genetic clusters (K=10) was the best supported K in STRUCTURE based on the

339   observed variation of the estimated log likelihood values within a run and the

340   increase in the log likelihood for each K (Figure 4a; Figure S1). Consistent with

341   previous findings (Lucek et al. 2014a), both the Geneva lake and stream

342   populations as well as the France Stream 2 population were all assigned to a

343   single genetic cluster with evidence for some introgression from the Constance

344   cluster, comprising both the Constance lake and stream site. All other

345   populations formed their own distinct genetic cluster except for the Iceland lake

346   population showing some evidence for introgression from the marine population.

347

348   The two leading PC axes that were constructed using population-based

349   microsatellite allele frequencies accounted for 30.4% and 15.5% of the total

350   variance among populations respectively (Figure 4b). The first PC axis separated

351   the Mediterranean island populations from all other populations, where the

352   populations from Geneva and France Stream 2 are genetically most distinct from

353   the latter. The second PC axis separated the France Stream 1 population from all

354   other populations, whereas populations from Iceland, Constance, Poland and

355   Galicia were intermediate to the other populations along the two PC axes.

356   Interestingly the France Stream 1 population clusters closely with the population

357   from Mallorca in the population based neighbor-joining tree (Figure 4c), but is

358   otherwise consistent with the PC analysis, where Corsica, Sardinia and Galicia

359   form a distinct clade.

360

361   JMODELTEST selected the HKY+G+I model of molecular evolution as the best fitting

362   model for the partial mitochondrial control region. All ESS values for the BEAST

363   analysis were >200, suggesting an adequate sampling of the posterior. The BEAST

364   analysis shows a deep divergence between a clade that comprises all haplotypes

365   from the North American Pacific, North American Atlantic, the Black Sea, and the

366   Mediterranean Lake Skadar and a clade that includes all other haplotypes from

367   across Europe plus one North American Atlantic haplotype (Figure 5). However,

368   whereas this deeper split is not statistically supported, several clades within

369   these groups are well supported (see below). Mediterranean stickleback

370   populations occur widely scattered across both clades with several

371   Mediterranean specific clusters, including ancient splits as well as recent ones.

372   Two ancient populations from the Balkans (Ska, Ner) are strongly supported

373   monophyletic groups with deep divergence from other populations, confirming

374   previous studies (DeFaveri et al. 2012). In particular, the haplotypes of Lake

375   Skadar fall into a relatively old clade (divergence from other stickleback 0.0419-

376   0.3886 mya; diversification within the Skadar clade started 0.0069-0.1212 mya

377   as suggested by the 95% highest posterior density for this node height). Its

378   closest relatives are haplotypes from North America, which is consistent with an

379   ancient colonization of Europe that may be independent of that of the main clade

380   (Bell and Foster 1994).

381

382   The haplotypes found on the three Mediterranean islands are spread across the

383   European main clade: The haplotype from Mallorca (estimated age: 0.0070-

384   0.1160 mya) forms a clade together with haplotypes from the Font Dame springs

385   in southern France and the River Sorgue in the southern Rhone valley albeit with

386   only weak support (BPP = 0.65). This clade has an estimated age of 0.0426-

387   0.2360 mya. The Corsican haplotype with a divergence time of 0.0001-0.0724

388   mya, clustered closely with the EU43 haplotype, which has formerly been

389   reported from the Loire drainage in Central France that drains into the Atlantic

390   (Mäkinen and Merilä 2008). The Sardinian haplotype clustered closest to the

391   EU19 and EU20 haplotypes, which derive from the Baltic Sea and the Northern

392   Atlantic. However, the age of the Sardinian haplotype could not be reliable dated.

393

394   *Phenotypic differentiation*

395   The PC analysis comprising all size-corrected linear measurements accounted for

396   39.2% and 17.2% of the overall phenotypic variation on the first and second PC

397   axis respectively (Figure 6). Defense-related traits accounted for most of the

398   variation observed along the first PC axis, whereas the second PC axis was

399   mainly driven by differences in fin sizes and body depth (Table S3). The best

400   statistical model explaining the variation in PC scores along PC1 for freshwater

401   populations using all linear traits combined included only the factor *island*,

402   where island populations differed significantly from mainland populations ($F_{1,10}$

403   = 19.65, p = 0.001). The difference between island and mainland population was

404   also the best fitting model to explain variation on the second axis albeit being not

405   significant ($F_{1,10}$ = 4.57, p = 0.058). Using only defense-related linear morphology,

406   the two leading PC axes accounted for 72.2% and 16.6% of the total phenotypic

407   variation respectively, where both axes were mainly driven by variation in pelvic

408   girdle and spine length (Table S3). The best fitting statistical model to explain the

409   variation among freshwater populations on the first PC axis also included only a

410   significant effect of *island* ($F_{1,10}$ = 22.17, p < 0.001). Similarly to the analysis

411   comprising all traits, all the island populations are separated from all the

412   mainland populations and from the marine population, where mainland south

17

413　European populations seem to resemble the insular populations most (Figure 6).

414　None of the tested factors was significantly associated with the variation

415　observed along the second PC axis for defense-related traits (all p > 0.1, results

416　not shown).

417

418　The best statistical model for body size (standard length) retained only

419　*geography* as a factor, where south European individuals were significantly

420　smaller than individuals from the north ($F_{1,10}$ = 11.23, p = 0.007; Figure 7). For

421　the number of lateral plates, the best fitting model showed significant effects of

422　both *geography* and *island*, where individuals from the south had significantly

423　fewer plates ($F_{1,9}$ = 44.31, p < 0.001), which was equally true for individuals

424　originating from islands ($F_{1,9}$ = 15.08, p = 0.004).

425

426　**Discussion**

427

428　Many insular ecosystems are currently threatened by climatic changes and other

429　human-induced perturbations (Ricketts et al. 2005; Courchamp et al. 2014).

430　Detecting and preserving insular endemics thus is pressing as, in many cases,

431　islands harbor populations of otherwise widely distributed species that are

432　phenotypically and genetically most distinct (Foster 1964; Case 1978; Lomolino

433　2005). Combining genetic, phenotypic and ecological data we aimed to

434　determine the status of the little known and biogeographically exceptional

435　populations of the threespine stickleback on Mediterranean islands. We find that

436　these populations are both phenotypically and genetically distinct from

437　mainland populations underlining their status as evolutionary significant units

438 for conservation (Ryder 1986; Moritz 1994). Convergence on a similar

439 phenotype was furthermore observed between these populations and an insular

440 population from Iceland, northern Europe, suggesting that similar evolutionary

441 forces may act during phenotypic adaptation to insular ecosystems.

442

443 *Genetic relationship and divergence of Mediterranean populations*

444 For many organisms, the Mediterranean region has acted as a glacial refugium

445 from which subsequent postglacial expansions have taken place (Hewitt 2000).

446 The contemporary distribution of freshwater-dwelling threespine stickleback in

447 the Mediterranean region is likewise thought to originate from a distinct glacial

448 refugium, but expansion to northern Europe is thought to not have occurred

449 from this refugium. In contrast, freshwater stickleback populations from

450 northern Europe are believed to derive from recolonization events from Atlantic

451 marine populations (Mäkinen et al. 2006). Analyses of mitochondrial haplotypes

452 further suggested that some Mediterranean populations may have survived the

453 last Pleistocene glaciation (Mäkinen and Merilä 2008; DeFaveri et al. 2012).

454 Whereas earlier phylogeographic work focused on populations from the Eastern

455 Mediterranean, we focused on island populations in the western Mediterranean.

456 Our phylogeographic estimate based on mitochondrial control region sequences

457 suggests that each western Mediterranean island population represents a

458 distinct genetic lineage (Figure 5). Interestingly, the haplotype found on Mallorca

459 clusters closely to haplotypes originating from a population in southern France,

460 and together may be remnants of one glacial freshwater refugium as indicated by

461 our Bayesian estimate of divergence. The haplotypes of Corsica and Sardinia, not

462 known from anywhere else, are most closely related to haplotypes found in

19

463 central or northern Europe. They may either originate from a more recent

464 colonization event of the Mediterranean region, or may represent lineages that

465 expanded northward again after the glacial maximum.

466

467 Threespined stickleback likely originated in the Pacific Ocean, with the most

468 divergent lineage being the Japanese Pacific clade (Watanabe et al. 2003; Jones et

469 al. 2012; Rezansoff et al.). We could not include this lineage in our tree because

470 sequences of the mitochondrial control region are not present in GenBank.

471 However, studies based on other mitochondrial genes have shown that this

472 lineage diverged from the clade that we studied about 2.5 mya (Rezansoff et al.).

473 Our tree is consistent with the hypothesis that stickleback of Europe derive from

474 several distinct colonization events (Bell and Foster 1994). Remnants of one

475 relatively ancient colonization event seem to persist in the Black Sea and in Lake

476 Skadar, being known otherwise from both sides of North America (Figure 5). The

477 other colonization event, similarly ancient, would have formed the Neretva River

478 clade from the Balkans and to the main European clade of freshwater stickleback.

479 The deep divergence between these two, and the lack of statistical support for

480 nodes deeper in the tree make it plausible that Europe was colonized more than

481 twice. Indeed, stickleback from the Black Sea, Lake Skadar, Neretva and the clade

482 comprising all other European haplotypes may derive from four colonization

483 events. Clearly much more sequencing effort is needed to resolve the

484 phylogeography of the genus *Gasterosteus* in Europe and beyond.

485

486 The potential for a Mediterranean-based refugium is further indicated by the

487 levels of pairwise genetic differentiation found with nuclear markers, where the

488   levels of genetic differentiation involving populations from southern Europe

489   tend to be higher than comparisons involving north European or the marine

490   population (Figure 3, Table S1 & S2). Also, the levels of pairwise genetic

491   differentiation between populations from southern Europe and the marine

492   population tend to be higher than the levels of genetic differentiation between

493   the marine population and populations from northern Europe. This may reflect

494   the absence of contemporary gene flow from the Atlantic marine population into

495   the Mediterranean region or could indicate that the overall split between the

496   Atlantic marine and the freshwater populations from southern Europe predates

497   the one between the marine and northern European freshwater populations. In

498   contrast, the pattern found in northern Europe is consistent with a recent

499   postglacial recolonization of freshwater bodies from the Atlantic, with a

500   significantly lower genetic differentiation between marine and freshwater (Jones

501   et al. 2012). Lastly, almost all populations from southern Europe cluster closely

502   together in the distance-based *tree* based on microsatellites (Figure 4). A second

503   genetic southern European cluster is formed by the three populations in the

504   upper Rhone drainage (i.e. France Stream 2, Geneva Lake and Stream). This is in

505   agreement with another study showing that the latter populations originate from

506   a very recent (less than 140 years) expansion event following the translocation

507   of stickleback within this region (Lucek et al. 2014a).

508

509   Despite clustering together in the distance-based tree, almost all populations

510   from southern Europe represent distinct genetic clusters in the individual based

511   assignment (Figure 4). Moreover, the same populations show the highest degree

512   of pairwise genetic differentiation, which in most cases even surpasses the level

513   of genetic differentiation between populations from southern and northern

514   Europe (Figure 3, Table S1 & S2). This is consistent with other studies on

515   Mediterranean stickleback, where genetic differentiation was found to be high

516   even between geographically proximate populations (Cano et al. 2008; Mäkinen

517   and Merilä 2008; Araguas et al. 2012; DeFaveri et al. 2012). This pattern,

518   combined with the significantly decreased genetic variation within south

519   European populations is consistent with a reduction of gene flow due to the

520   absence of a connecting marine population. The reduced genetic variation may

521   furthermore be explained by a decrease in their population size due to habitat

522   loss, fragmentation or the introduction of non-native species that compete with

523   stickleback (Clavero et al. 2009; Orru et al. 2010; Araguas et al. 2012).

524

525   *Evolution of phenotypically distinct island populations*

526   The island rule predicts the evolution of phenotypically distinct insular

527   populations (Foster 1964; Case 1978; Lomolino 1985), where on the one hand

528   insular populations of small vertebrate species evolve an increased body size as

529   a consequence of partial release from competitors and predators. On the other

530   hand, large vertebrate species may evolve dwarfism due to resource limitations.

531   These changes may in both cases trigger subsequent phenotypic adaptations

532   (Foster 1964; Case 1978; Lomolino 1985; Lomolino et al. 2012). In stickleback,

533   differences in anti-predator related phenotypes have been commonly attributed

534   to distinct selection pressures imposed by differences in the strength and

535   composition of the predation regimes (Reimchen 1994). Specifically, stickleback

536   may evolve long spines and an increased number of lateral plates in

537   environments with strong predation pressure dominated by piscivorous fishes

and birds, such as in the marine environment as opposed to freshwater

populations (Reimchen 1994). The evolution of low armour plate phenotypes

among freshwater stickleback can occur rapidly through selection on standing

genetic variation from the marine population (Colosimo et al. 2005). Indeed, the

best example so far for insular phenotypes in stickleback derives from Scotland,

where stickleback reduced their body size and defense-related traits, following

the island rule (MacColl et al. 2013). The evolution these insular phenotypes was

potentially mediated by a distinct water chemistry (i.e. a low pH and few

dissolved alkaline metals (Spence et al. 2013)) and the presence of small

predators (MacColl et al. 2013).

The phenotypic differentiation observed in our study is consistent with the

evolution of distinct insular phenotypes, where populations from islands,

including Iceland, differ phenotypically from populations on the European

continent. This pattern is largely driven by a reduction in defense-related traits

in insular populations (Figures 6 & 7, Table S3). Albeit the number of lateral

plates differs overall between northern and southern Europe, where populations

from the latter region have a much lower plate number (Figure 7b),

Mediterranean island populations show nevertheless more extreme phenotypes,

in comparison to populations from the southern European mainland (Figure 7).

Similarly reduced lateral plate numbers, below the common freshwater

phenotype with 5-7 lateral plates, are rare across Europe, where they occur

mainly in the south (Gross 1977; Gross 1978) or among insular populations in

Scotland (Spence et al. 2013). Thus, the reduction in anti-predator related traits

among our studied insular populations and insular populations from Scotland

563 (Spence et al. 2013), may reflect common features in terms of the reduced

564 predator and competitor communities on islands (Case 1978; Lomolino et al.

565 2012).

566

567 Body size itself differed overall mainly between freshwater populations from

568 northern and southern Europe, where populations from the latter were

569 significantly smaller (Figure 7a). The evolution of distinct body sizes among

570 freshwater stickleback has been associated previously with differences in water

571 chemistry (McGuigan et al. 2011) and foraging strategies (Nagel and Schluter

572 1998). Whereas we cannot exclude differences in water chemistry, we found no

573 support for feeding-related ecological differentiation among populations from

574 southern Europe (Figure 2). Remarkably, adult individuals from Mallorca

575 showed a more reduced body size, resembling subadult phenotypes from

576 northern Europe (i.e. Frommen et al. 2011; Lucek et al. 2012). Similar dwarfism

577 has otherwise only rarely been recorded in stickleback. The best examples are

578 the insular populations from Scotland that have reduced their body size

579 potentially as a response to water chemistry and the presence of small predators

580 (MacColl et al. 2013). The Mallorcan population is moreover exceptional as we

581 observed breeding individuals in November (Lucek *personal observation*),

582 suggesting either an extended (from spring to November) or shifted breeding

583 period. Analogous life-history adaptations in stickleback where stickleback shift

584 their breeding effort to periods with decreased water temperatures have only

585 been recorded in populations from Japan (Mori 1987) and the southern Rhone

586 valley in France (Crivelli and Britton 1987). The latter is in the range of the other

587   French populations that form a distinct cluster with the Mallorcan haplotype in

588   the phylogenetic analysis (Figure 5).

589

590   *Conclusions*

591   Taken together, our results suggest that endemic insular forms exist in

592   threespine stickleback, where the populations from the Mediterranean islands

593   are both phenotypically and genetically highly distinct from all mainland

594   populations including those from southern Europe, and from each other. The

595   genetic evidence suggests that they likely derive from glacial refugia in the

596   Mediterranean region. Thus given their genetic distinctiveness (Figure 3 & 4)

597   and phenotypic differentiation (Figure 6 & 7), Mediterranean island populations

598   should be considered as evolutionary significant units (Ryder 1986; Moritz

599   1994), each endemic to one island. The status as evolutionary significant unit is

600   particularly highlighted in the Mallorcan population, which shows specific life-

601   history adaptations that has otherwise only reported in a single population in

602   Southern France in the Mediterranean region (Crivelli and Britton 1987).

603   However, no such data was available for Sardinian and Corsican populations.

604

605   The low genetic variation observed within these populations, combined with

606   their restricted range, and ecological requirements (such as year round cool

607   water which is rare in the Mediterranean region), renders them especially

608   vulnerable to habitat loss, fragmentation and alteration (including deforestation)

609   (Araguas et al. 2012) or non-native fish species that outnumber native

610   stickleback ((Orru et al. 2010); *Gambusia holbrooki* in Mallorca (Lucek *personal*

611   *observation*)). For instance the Corsican population is historically only known

612   from two neighboring rivers in a small part of the island's coastal lowlands

613   (Gauthier and Rose 1974). Despite intensive search, we found it in only one of

614   them. This river is exceptional among Corsican coastal lowland rivers in having

615   relatively cold water coming from the nearby mountains, and is well shaded by a

616   thin strip of riparian forest. Further protective measures are thus needed that

617   allow the monitoring of these populations, the safe-guarding of their habitats,

618   and aim to decrease the threat made by non-native species.

619

629

References

630

631 Aldenhoven JT, Miller MA, Corneli PS, Shapiro MD (2010) Phylogeography of
632   ninespine sticklebacks (*Pungitius pungitius*) in North America: glacial refugia
633   and the origins of adaptive traits. Mol Ecol 19:4061–4076. doi:
634   10.1111/j.1365-294X.2010.04801.x

635 Araguas RM, Vidal O, Pla C, Sanz N (2012) High genetic diversity of the
636   endangered Iberian three-spined stickleback (*Gasterosteus aculeatus*) at the
637   Mediterranean edge of its range. Freshwater Biol 57:143–154. doi:
638   10.1111/j.1365-2427.2011.02705.x

639 Bell MA (1981) Lateral plate polymorphism and ontogeny of the complete plate
640   morph of threespine sticklebacks (*Gasterosteus aculeatus*). Evolution 35:67–
641   74.

642 Bell MA, Foster SA (1994) The Evolutionary Biology of the Threespine
643   Stickleback. Oxford University Press, Oxford, UK.

644 Berner D, Adams DC, Grandchamp A-C, Hendry AP (2008) Natural selection
645   drives patterns of lake-stream divergence in stickleback foraging
646   morphology. J Evol Biol 21:1653–1665. doi: 10.1111/j.1420-
647   9101.2008.01583.x

648 Bertin L (1925) Recherches bionomiques, biométriques et sysématiques sur les
649   épinoches (*Gastérostéidés*). Ann Inst Océanogr Monaco 2:1–204.

650 Bianco PG (1980) Areale Italico, rinvenimento in Calabria e origini delle
651   popolationi metiterranee di *Gasterosteus aculeatus* L. Boll Mus Vic St Nat
652   Verona 7:197–216.

653 Bouckaert R, Heled J, Kühnert D, et al. (2014) BEAST 2: a software platform for
654   Bayesian evolutionary analysis. PLoS Comput Biol 10:e1003537. doi:
655   10.1371/journal.pcbi.1003537

656 Cano JM, Mäkinen HS, Leinonen T, et al. (2008) Extreme neutral genetic and
657   morphological divergence supports classification of Adriatic three-spined
658   stickleback (*Gasterosteus aculeatus*) populations as distinct conservation
659   units. Biol Cons 141:1055–1066. doi: 10.1016/j.biocon.2008.01.015

660 Case TJ (1978) A general explanation for insular body size trends in terrestrial
661   vertebrates. Ecology 59:1–18. doi: 10.2307/1936628

662 Clavero M, Pou-Rovira Q, Zamora L (2009) Biology and habitat use of three-
663   spined stickleback (*Gasterosteus aculeatus*) in intermittent Mediterranean
664   streams. Ecol Freshw Fish 18:550-559:

665 Colosimo PF, Hosemann KE, Balabhadra S, et al. (2005) Widespread parallel
666   evolution in sticklebacks by repeated fixation of Ectodysplasin alleles.
667   Science 307:1928–1933. doi: 10.1126/science.1107239

668 Courchamp F, Hoffmann BD, Russell JC, et al. (2014) Climate change, sea-level
669   rise, and conservation: keeping island biodiversity afloat. Trends Ecol Evol
670   29:127–130. doi: 10.1016/j.tree.2014.01.001

671 Crivelli AJ, Britton RH (1987) Life history adaptations of *Gasterosteus aculeatus*
672   in a Mediterranean wetland. Environ Biol Fish 18:109–125.

673 Cuttelod A, Garcia N, Malak DA, et al. (2008) The Mediterranean: a biodiversity
674   hotspot under threat. The 2008 Review of The IUCN Red List of Threatened
675   Species 1–16.

676 Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models,
677   new heuristics and parallel computing. Nat Methods 9:772. doi:

678   10.1038/nmeth.2109

679   DeFaveri J, Zanella LN, Zanella D, et al. (2012) Phylogeography of isolated
680   freshwater three-spined stickleback *Gasterosteus aculeatus* populations in
681   the Adriatic Sea basin. J Fish Biol 80:61–85. doi: 10.1111/j.1095-
682   8649.2011.03147.x

683   Elton CS (1958) The ecology of invasions by animals and plants. University Of
684   Chicago Press, Chicago, IL, USA.

685   Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of
686   individuals using the software STRUCTURE: a simulation study. Mol Ecol
687   14:2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x

688   Falush D, Stephens M, Pritchard J (2007) Inference of population structure using
689   multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes
690   7:574–578. doi: 10.1111/j.1471-8286.2007.01758.x

691   Foster JB (1964) Evolution of mammals on islands. Nature 202:234-235.

692   Frommen JG, Herder F, Engqvist L, et al. (2011) Costly plastic morphological
693   responses to predator specific odour cues in three-spined sticklebacks
694   (*Gasterosteus aculeatus*). Evol Ecol 25:641–656. doi: 10.1007/s10682-010-
695   9454-6

696   Gasith A, Resh VH (1999) Streams in mediterranean climate regions: abiotic
697   influences and biotic responses to predictable seasonal events. Annu Rev
698   Ecol Evol S 30:51–81. doi: 10.2307/221679

699   Gauthier A, Rose B (1974) Presence de l'épinoche en Corse. Bull Sco Sci Hist Nat
700   Corse 44:41-48.

701   Geiger MF, Herder F, Monaghan MT, et al. (2014) Spatial heterogeneity in the
702   Mediterranean Biodiversity Hotspot affects barcoding accuracy of its
703   freshwater fishes. Mol Ecol Resour 14:1210–1221. doi: 10.1111/1755-
704   0998.12257

705   Gross HP (1978) Natural selection by predators on defensive apparatus of the
706   three-spined stickleback, *Gasterosteus aculeatus* L. Can J Zool 56:398–413.

707   Gross HP (1977) Adaptive trends of environmentally sensitive traits in the three-
708   spined stickleback, *Gasterosteus aculeatus* L. Z Zool Syst Evol 15:252-278.

709   Hedrick PW (2005) A standardized genetic differentiation measure. Evolution
710   59:1633–1638.

711   Hewitt G (2000) The genetic legacy of the Quaternary ice ages. Nature 405:907–
712   913. doi: 10.1038/35016000

713   Jones FC, Chan YF, Schmutz J, et al. (2012) A genome-wide SNP genotyping array
714   reveals patterns of global and repeated species-pair divergence in
715   sticklebacks. Curr Biol 22:83–90. doi: 10.1016/j.cub.2011.11.045

716   Keenan K, McGinnity P, Cross TF, et al. (2013) diveRsity: An R package for the
717   estimation and exploration of population genetics parameters and their
718   associated errors. Methods Ecol Evol 4:782–788. doi: 10.1111/2041-
719   210X.12067

720   Kottelat M, Freyhof J (2007) Handbook of European freshwater fishes. Kottelat,
721   Cornol and Freyhof, Berlin, Germany.

722   Kristjánsson BK, Skulason S, Noakes D (2002) Morphological segregation of
723   Icelandic threespine stickleback (*Gasterosteus aculeatus* L). Biol J Linn Soc
724   76:247–257.

725 Krupp F, Coad BW (1985) Notes on a population of the threespine stickleback,
726 Gasterosteus aculeatus, from Syria (Pisces: Osteichthyes: Gasterosteidae). Senck
727 Biol 66:35–39.

728 Lande R (1998) Anthropogenic, ecological and genetic factors in extinction and
729 conservation. Res Popul Ecol 40:259–269.
730 Lobón-Cerviá J, Penczak T, de Sostoa A (1988) Morphological variability and
731 distribution of stickleback (*Gasterosteus aculeatus* L.) in Spain. Cybium
732 12:219–227.
733 Lockwood JL, F Hoopes M, P Marchetti M (2007) Invasion Ecology. Blackwell,
734 Malden, MA, USA.

735 Lomolino MV (1985) Body size of mammals on islands: the island rule
736 reexamined. Am Nat 125:310–316. doi: 10.2307/2461638
737 Lomolino MV (2005) Body size evolution in insular vertebrates: generality of the
738 island rule. J Biogeogr 32:1683–1699. doi: 10.1111/j.1365-
739 2699.2005.01314.x
740 Lomolino MV, Sax DF, Palombo MR (2012) Of mice and mammoths evaluations of
741 causal explanations for body size evolution in insular mammals. J Biogeogr
742 40:1427-1439.

743 Lucek K, Lemoine M, Seehausen O (2014a) Contemporary ecotypic divergence
744 during a recent range expansion was facilitated by adaptive introgression. J
745 Evol Biol 27:2233–2248. doi: 10.1111/jeb.12475
746 Lucek K, Roy D, Bezault E, et al. (2010) Hybridization between distant lineages
747 increases adaptive variation during a biological invasion: stickleback in
748 Switzerland. Mol Ecol 19:3995–4011. doi: 10.1111/j.1365-
749 294X.2010.04781.x
750 Lucek K, Sivasundar A, Kristjánsson BK, et al. (2014b) Quick divergence but slow
751 convergence during ecotype formation in lake and stream stickleback pairs
752 of variable age. J Evol Biol 27:1878–1892. doi: 10.1111/jeb.12439
753 Lucek K, Sivasundar A, Seehausen O (2012) Evidence of Adaptive Evolutionary
754 Divergence during Biological Invasion. PLoS ONE 7:e49377. doi:
755 10.1371/journal.pone.0049377
756 MacColl ADC, Nagar El A, de Roij J (2013) The evolutionary ecology of dwarfism
757 in three-spined sticklebacks. J Anim Ecol 82:642–652. doi: 10.1111/1365-
758 2656.12028
759 Mäkinen HS, Cano JM, Merilä J (2006) Genetic relationships among marine and
760 freshwater populations of the European three-spined stickleback
761 (*Gasterosteus aculeatus*) revealed by microsatellites. Mol Ecol 15:1519–1534.
762 doi: 10.1111/j.1365-294X.2006.02871.x
763 Mäkinen HS, Merilä J (2008) Mitochondrial DNA phylogeography of the three-
764 spined stickleback (*Gasterosteus aculeatus*) in Europe-Evidence for multiple
765 glacial refugia. Mol Phylogenet Evol 46:167–182. doi:
766 10.1016/j.ympev.2007.06.011
767 McGuigan K, Nishimura N, Currey M, et al. (2011) Cryptic genetic variation and
768 body size evolution in threespine stickleback. Evolution 65:1203–1211. doi:
769 10.1111/j.1558-5646.2010.01195.x
770 McKinnon JS, Rundle H (2002) Speciation in nature: the threespine stickleback
771 model systems. Trends Ecol Evol 17:480–488.

772   Meirmans PG, Hedrick PW (2011) Assessing population structure: FST and
773   related measures. Mol Ecol Resour 11:5-18.
774   Meirmans PG, Van Tienderen P (2004) GENOTYPE and GENODIVE: two
775   programs for the analysis of genetic diversity of asexual organisms. Mol Ecol
776   Notes 4:792–794.
777   Mori S (1987) Divergence in reproductive ecology of the three-spined
778   stickleback, *Gasterosteus aculeatus*. Jap J Ichth 34:165–175.
779   Mori S, Takamura N (2004) Changes in morphological characteristics of an
780   introduced population of the threespine stickleback *Gasterosteus aculeatus* in
781   Lake Towada, northern Japan. Ichthyol Res 51:295–300. doi:
782   10.1007/s10228-004-0232-8
783   Moritz C (1994) Defining "Evolutionarily significant units" for conservation.
784   Trends Ecol Evol 9:373–375. doi: 10.1016/0169-5347(94)90057-4
785   Munzing J (1963) The evolution of variation and distributional patterns in
786   European populations of the three-spined stickleback, *Gasterosteus aculeatus*.
787   Evolution 320–332.
788   Nagel L, Schluter D (1998) Body size, natural selection, and speciation in
789   sticklebacks. Evolution 52:209–218.
790   Olden JD, Leroy Poff N, Douglas MR, et al. (2004) Ecological and evolutionary
791   consequences of biotic homogenization. Trends Ecol Evol 19:18–24. doi:
792   10.1016/j.tree.2003.09.010
793   Orru F, Deiana AM, Cau A (2010) Introduction and distribution of alien
794   freshwater fishes on the island of Sardinia (Italy): an assessment on the basis
795   of existing data sources. J Appl Ichthyol 26:46–52. doi: 10.1111/j.1439-
796   0426.2010.01501.x
797   Orti G, Bell MA, Reimchen TE, Meyer A (1994) Global survey of mitochondrial
798   DNA sequences in the threespine stickleback: evidence for recent migrations.
799   Evolution 48:608–622.
800   Palmer TN, Räisänen J (2002) Quantifying the risk of extreme seasonal
801   precipitation events in a changing climate. Nature 415:512–514. doi:
802   10.1038/415512a
803   Pereira HM, Leadley PW, Proença V, et al. (2010) Scenarios for global
804   biodiversity in the 21st century. Science 330:1496–1501. doi:
805   10.1126/science.1196624
806   Reimchen TE (1994) Predators and morphological evolution in threespine
807   stickleback. In Bell M and Foster S (eds) The evolutionary biology of the
808   threespine stickleback, 1st edn, Oxford University Press, Oxford, UK, pp 240– 276.

809   Rezansoff AM, Crispo E, Blair C, et al. Toward the genetic origins of a potentially
810   non-native population of threespine stickleback (*Gasterosteus aculeatus*) in
811   Alberta. Conserv Genet. doi: 10.5061/dryad.pc1th
812   Richmond JQ, Jacobs DK, Backlin AR, et al. (2014) Ephemeral stream reaches
813   preserve the evolutionary and distributional history of threespine
814   stickleback in the Santa Clara and Ventura River watersheds of southern
815   California. Conserv Genet. doi: 10.1007/s10592-014-0643-7
816   Ricketts TH, Dinerstein E, Boucher T, et al. (2005) Pinpointing and preventing
817   imminent extinctions. P Natl Acad Sci Usa 102:18497–18501. doi:
818   10.1073/pnas.0509060102
819   Riera F (1980) Breves notas y primera cita del Espinoso (*Gasterosteus aculeatus*

820   L.) en S'Albufera, Mallorca. Boll Soc Hist Nat Balears 24:109-111.

821   Ryder OA (1986) Species conservation and systematics - the dilemma of

822   subspecies. Trends Ecol Evol 1:9–10.

823   Schluter D, Conte GL (2009) Genetics and ecological speciation. P Natl Acad Sci

824   Usa 106 Suppl 1:9955–9962. doi: 10.1073/pnas.0901264106

825   Seehausen O, vanAlphen J, Witte F (1997) Cichlid fish diversity threatened by

826   eutrophication that curbs sexual selection. Science 277:1808–1811.

827   Spence R, Wootton RJ, Barber I, et al. (2013) Ecological causes of morphological

828   evolution in the three-spined stickleback. Ecol Evol 3:1717–1726. doi:

829   10.1002/ece3.581

830   Tamura K, Peterson D, Peterson N, et al. (2011) MEGA5: molecular evolutionary

831   genetics analysis using maximum likelihood, evolutionary distance, and

832   maximum parsimony methods. Mol Biol Evol 28:2731–2739. doi:

833   10.1093/molbev/msr121

834   Taylor EB, Boughman JW, Groenenboom M, et al. (2006) Speciation in reverse:

835   morphological and genetic evidence of the collapse of a three-spined

836   stickleback (*Gasterosteus aculeatus*) species pair. Mol Ecol 15:343–355. doi:

837   10.1111/j.1365-294X.2005.02794.x

838   Vonlanthen P, Bittner D, Hudson AG, et al. (2012) Eutrophication causes

839   speciation reversal in whitefish adaptive radiations. Nature 482:357–362.

840   doi: 10.1038/nature10824

841   Watanabe K, Mori S, Nishida M (2003) Genetic relationships and origin of two

842   geographic groups of the freshwater threespine stickleback, "hariyo." Zool

843   Sci 20:265–274. doi: 10.2108/zsj.20.265

844   Whittaker RJ, Fernández-Palacios JM (2007) Island Biogeography: Ecology,

845   Evolution and Conservation. Oxford University Press, Oxford, UK.

846

847

Figure Captions

849 Figure 1

850 Overview of the studied populations: a) Map of Europe (© Wikimedia) with the

851 populations included in this study highlighted. Populations from islands are

852 highlighted in color (orange – Iceland, red – Mediterranean). b) Representative

853 examples of each sex for the different stickleback populations: A – Iceland

854 Marine, B – Iceland Lake, C – Poland, D – Constance Lake, E – Constance stream, F

855 – Geneva Lake, G – Geneva stream, H – France Stream 2, I – France Stream 1, J –

856 Galicia, K – Mallorca, L – Corsica, M – Sardinia. Pictures for populations C, J and M

857 were taken from ethanol preserved specimens. Note that individuals are not

858 scaled by size. See Table 1 for details.

859

860 Figure 2

861 Stomach content analysis for eight populations. Top: percentage of planktonic

862 prey found within a population sample (see main text for details). Bottom:

863 percentage of each prey category found in each population. N indicates the

864 number of individuals that had at least one prey item in their stomachs.

865

866 Figure 3

867 Boxplots summarizing of the pairwise genetic differentiation among different

868 geographic groups based on a) F_{ST} and b) F'_{ST}. Genetic differentiation was

869 calculated between the marine and freshwater populations from either northern

870 or southern Europe, between freshwater populations from northern Europe,

871 between freshwater populations from northern and southern Europe and

872 between south European populations respectively (see Table S1 & S2 for the

873   actual F_{ST} and F'_{ST} values respectively). Asterisks indicate the level of statistical

874   significance of comparisons between geographic groups, indicated by horizontal

875   bars, based on *post hoc* ANOVA comparisons (* $p < 0.05$, ** $p < 0.01$, *** $p <$

876   0.001). The number of pairwise comparisons are indicated for each geographic

877   group (N).

878

879   Figure 4

880   Genetic relationship between populations based on seven microsatellites: a)

881   Individual genetic assignment using STRUCTURE. Shown is the best run, where

882   ten genetic clusters were inferred (K=10), which showed the highest posterior

883   likelihood. b) Principal component analysis, where the two leading axes capture

884   45.9% of the genetic variation among populations. c) Phylogram showing the

885   genetic relationship among populations based on Cavalli-Sforza distances.

886   Numbers indicate statistical support based on 1000 bootstrap replicates (only

887   values >50% are shown). Populations from islands are highlighted in color

888   (orange – Iceland, red – Mediterranean).

889

890   Figure 5

891   Bayesian maximum clade credibility tree based on a 426 base pair fragment of

892   the mitochondrial control region showing the genealogical relationships among

893   threespine stickleback. Node values represent Bayesian posterior probabilities ≥

894   0.50. Additional percentages reflect bootstrap support from an additional

895   maximum likelihood tree for the same node, based on 1000 bootstrap replicates

896   (only values ≥ 50% are shown). Error bars depict 95 % highest posterior density

897   for node heights (see main text for details).

898

899   Figure 6

900   Phenotypic variation across populations: Principal component analyses using

901   either all size-corrected linear morphological traits (left) or only size-corrected

902   defense-related traits (right, see main text for details). Shown are the mean PC

903   scores for each population ± 1 SE. Populations from islands are highlighted in

904   color (orange – Iceland, red – Mediterranean).

905

906   Figure 7

907   Phenotypic differentiation among the sampled populations for a) size (standard

908   length) and b) the number of lateral plates. Populations from islands are

909   highlighted in color (orange – Iceland, red – Mediterranean).

910

911   Figure S1

912   Summary statistics for the individual based assignments using STRUCTURE: a)

913   estimated likelihood for each run ± 1 SD; b) estimation of Delta K following

914   Evanno et al. (2005).
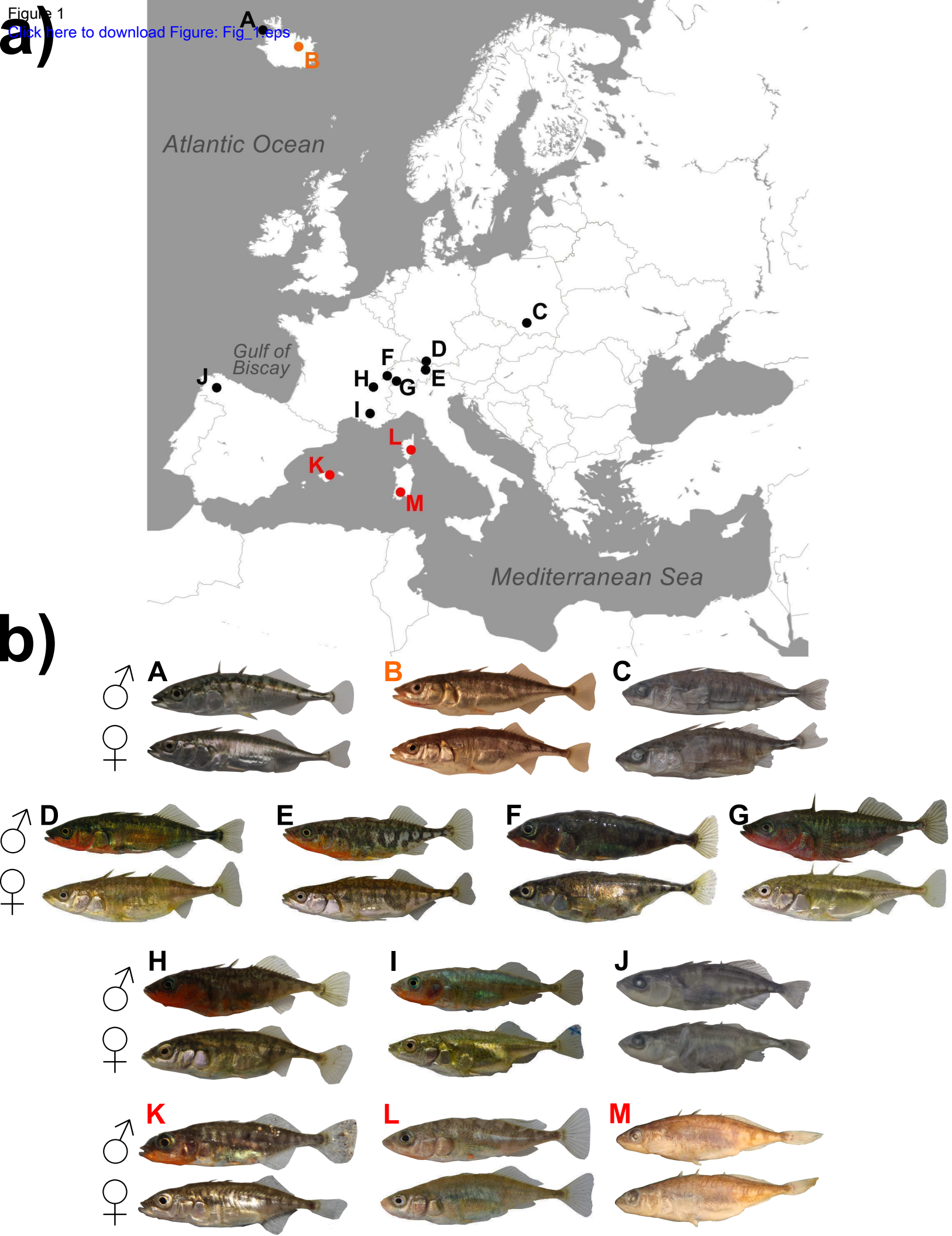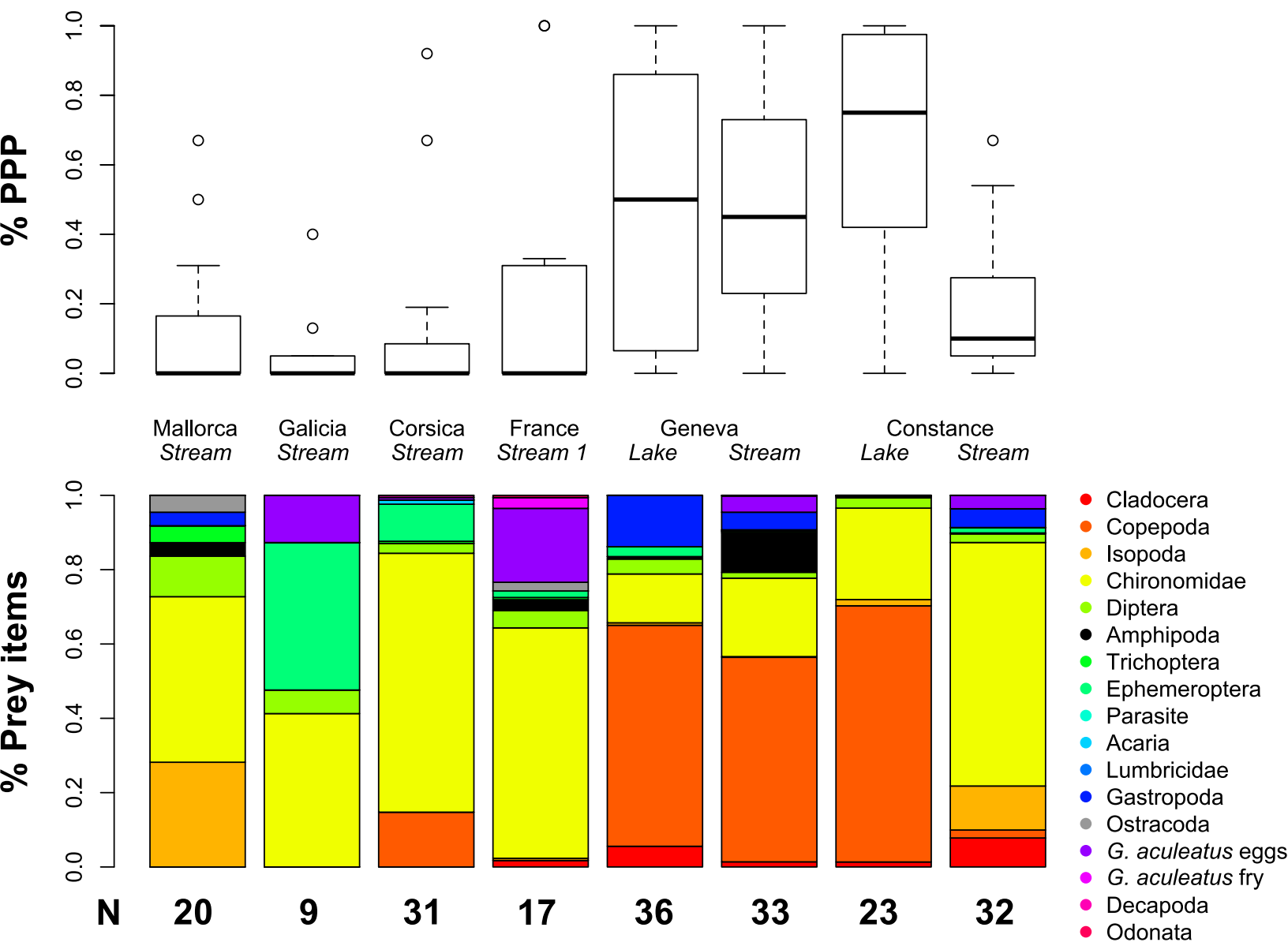
Figure 1
Click here to download Figure: Fig_1.eps

a)

Atlantic Ocean

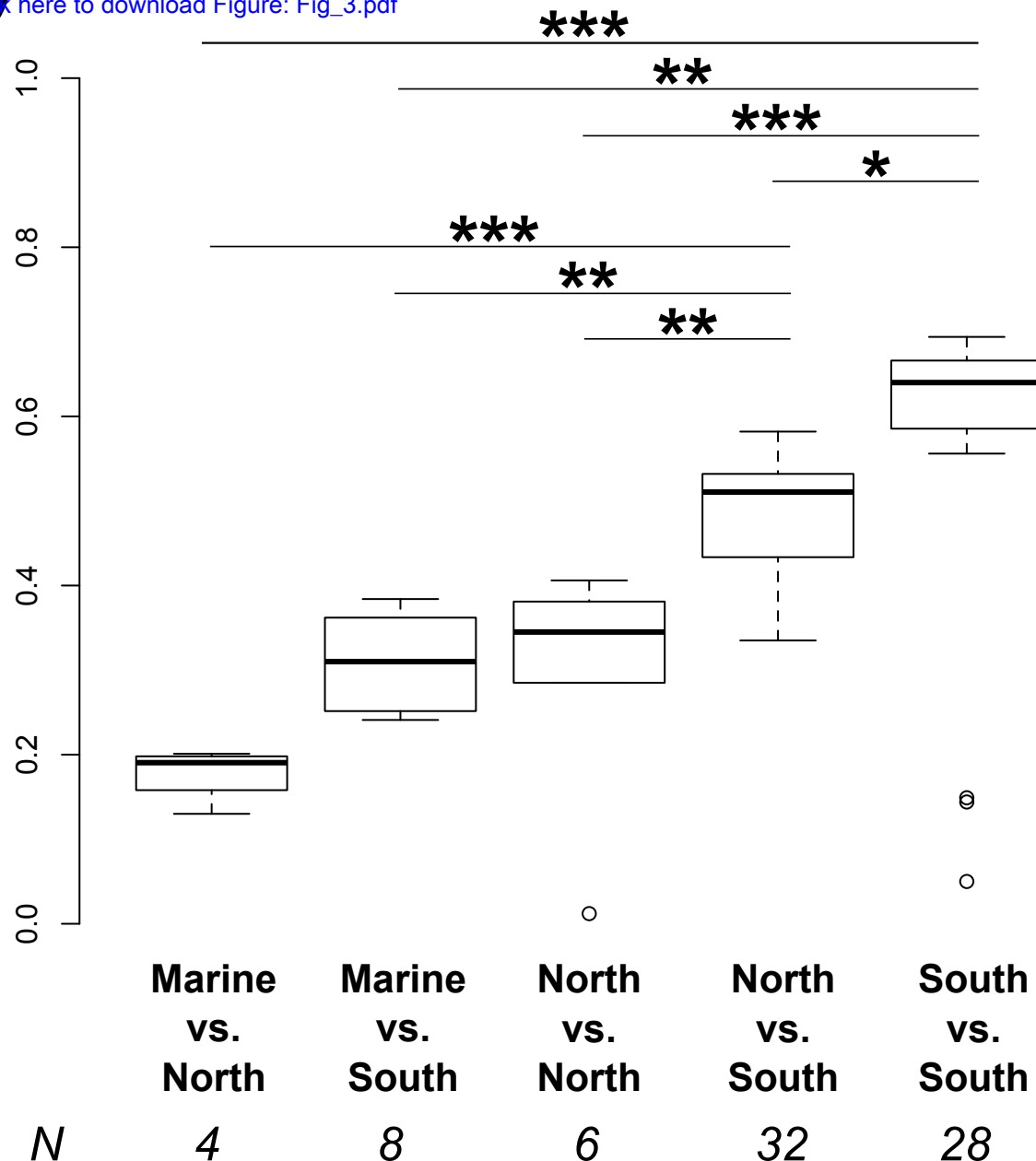Gulf of Biscay

Mediterranean Sea

b)

Figure 2
Click here to download Figure: Fig_2.eps
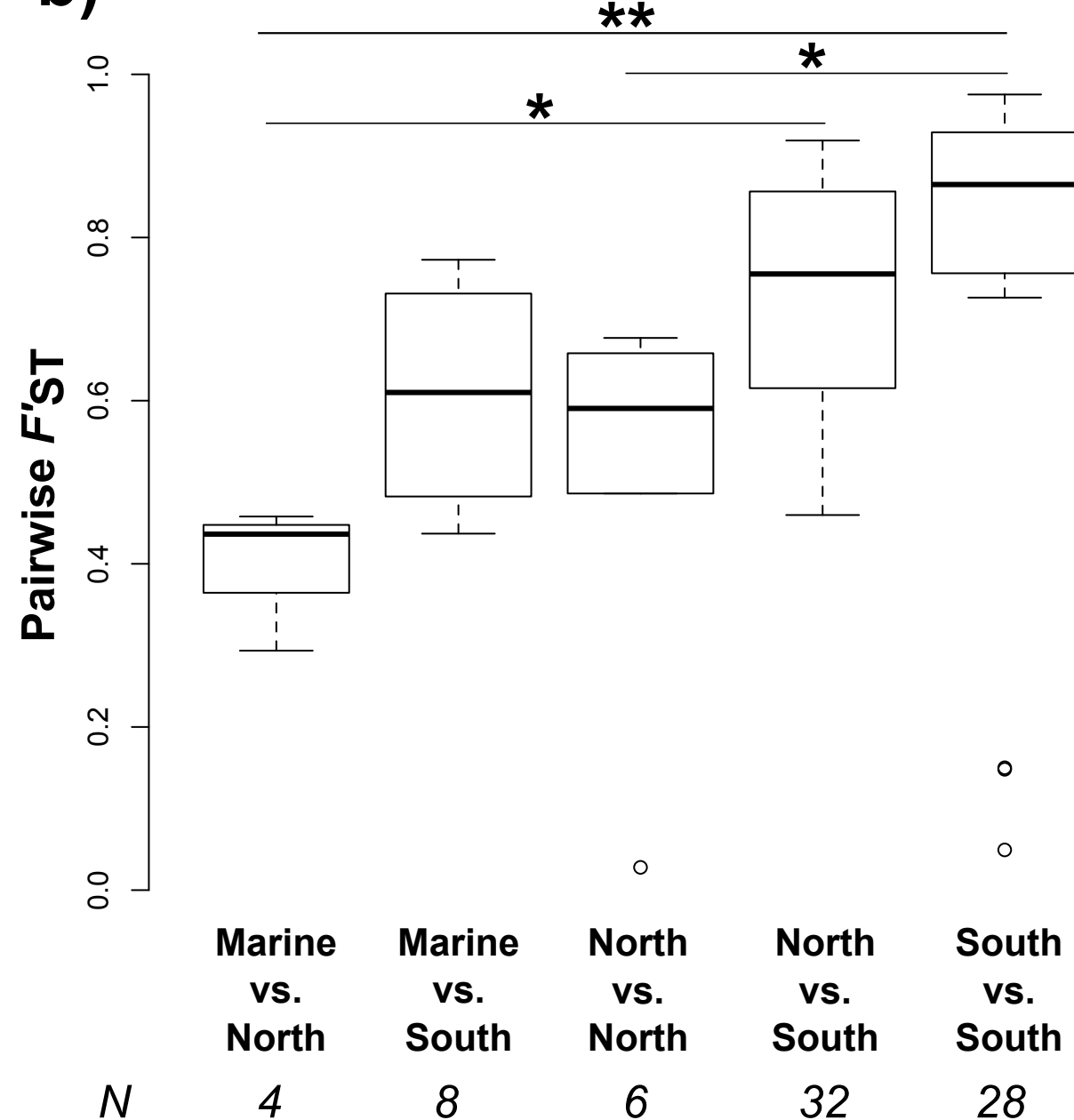
Figure 2

Figure 3

Figure 4
Click here to download Figure: Fig_4.eps

Figure 5

Million years ago

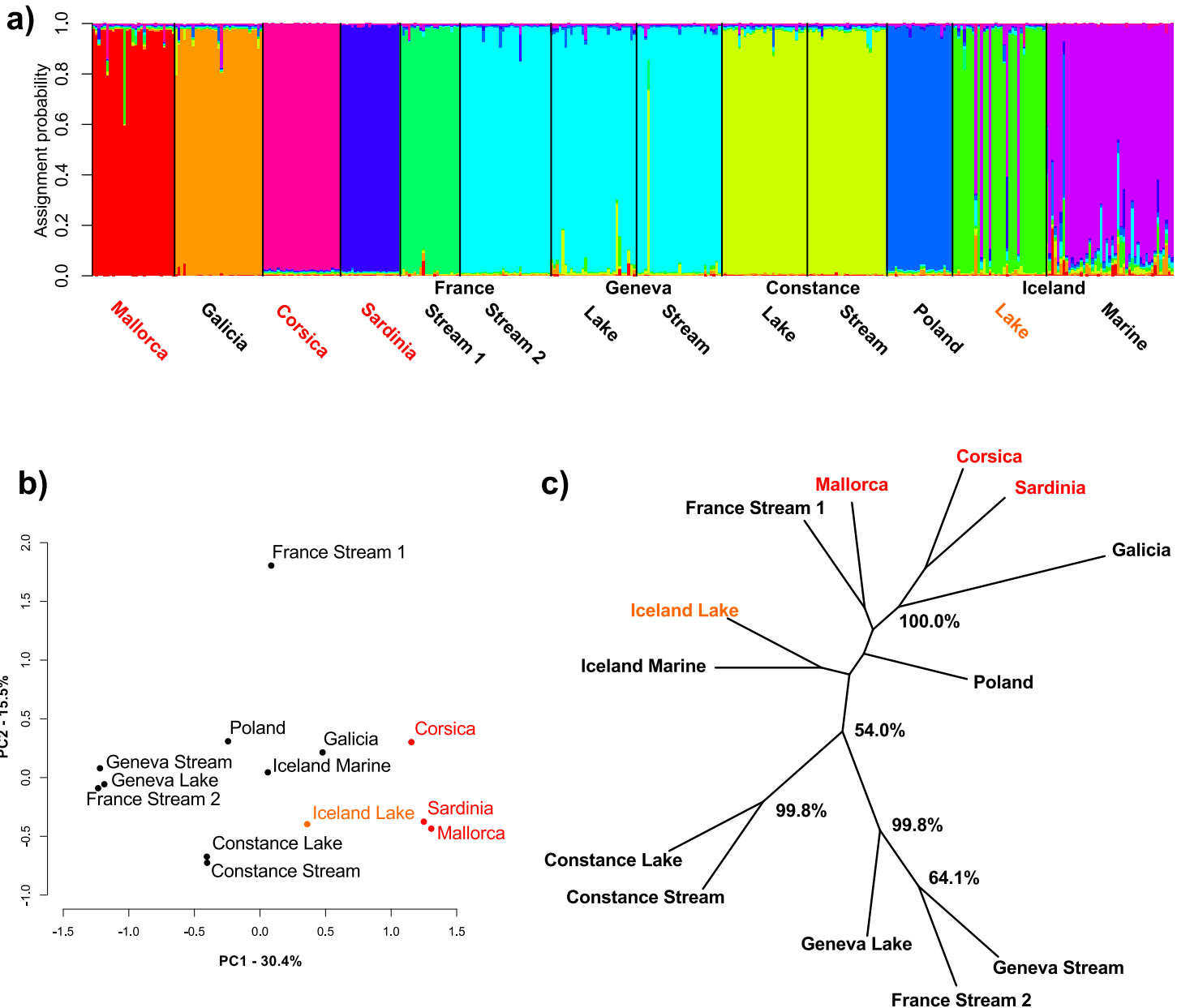Mediterranean Sea   Atlantic (Europe)   Black Sea   Atlantic (America)   Pacific
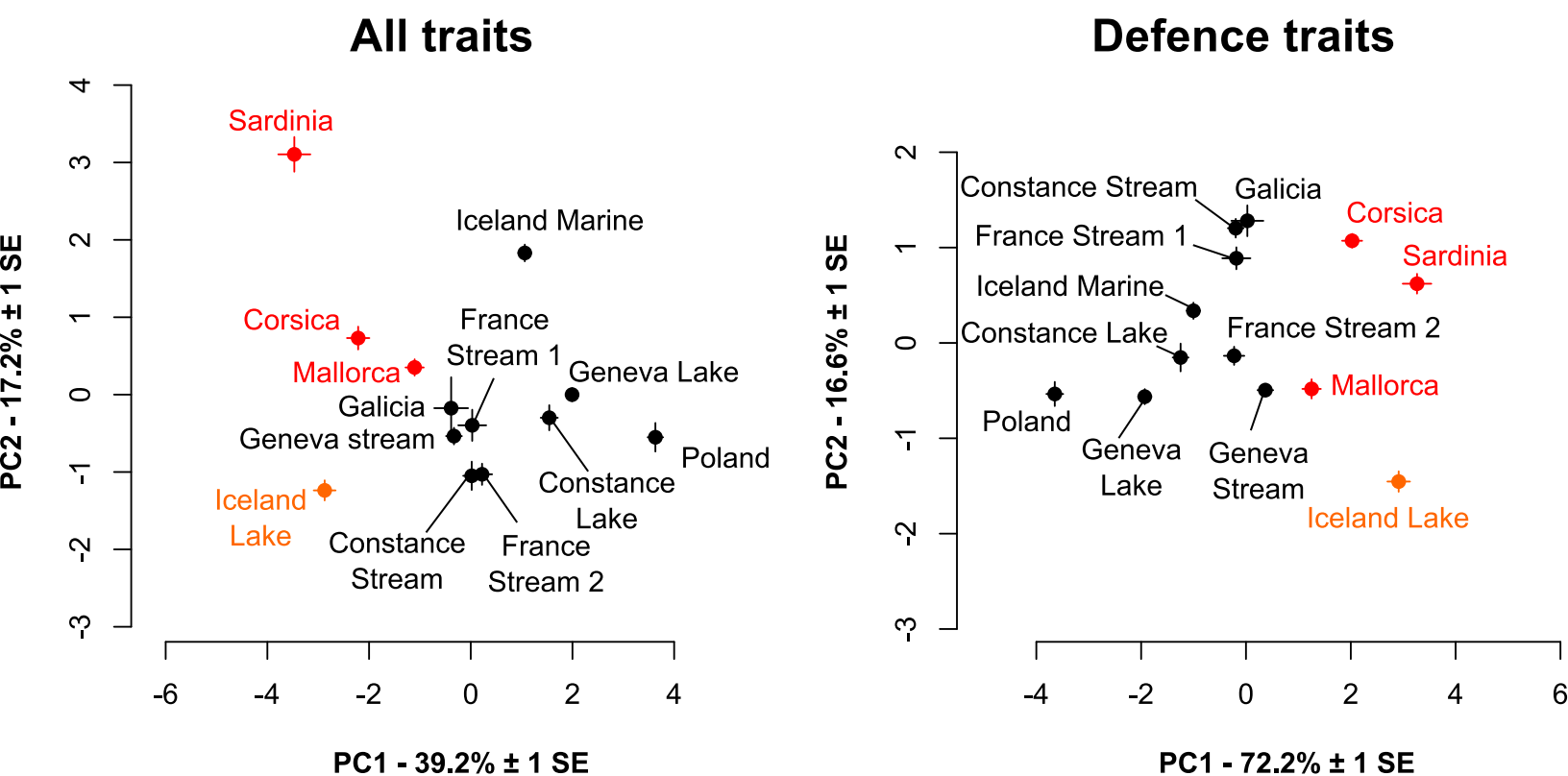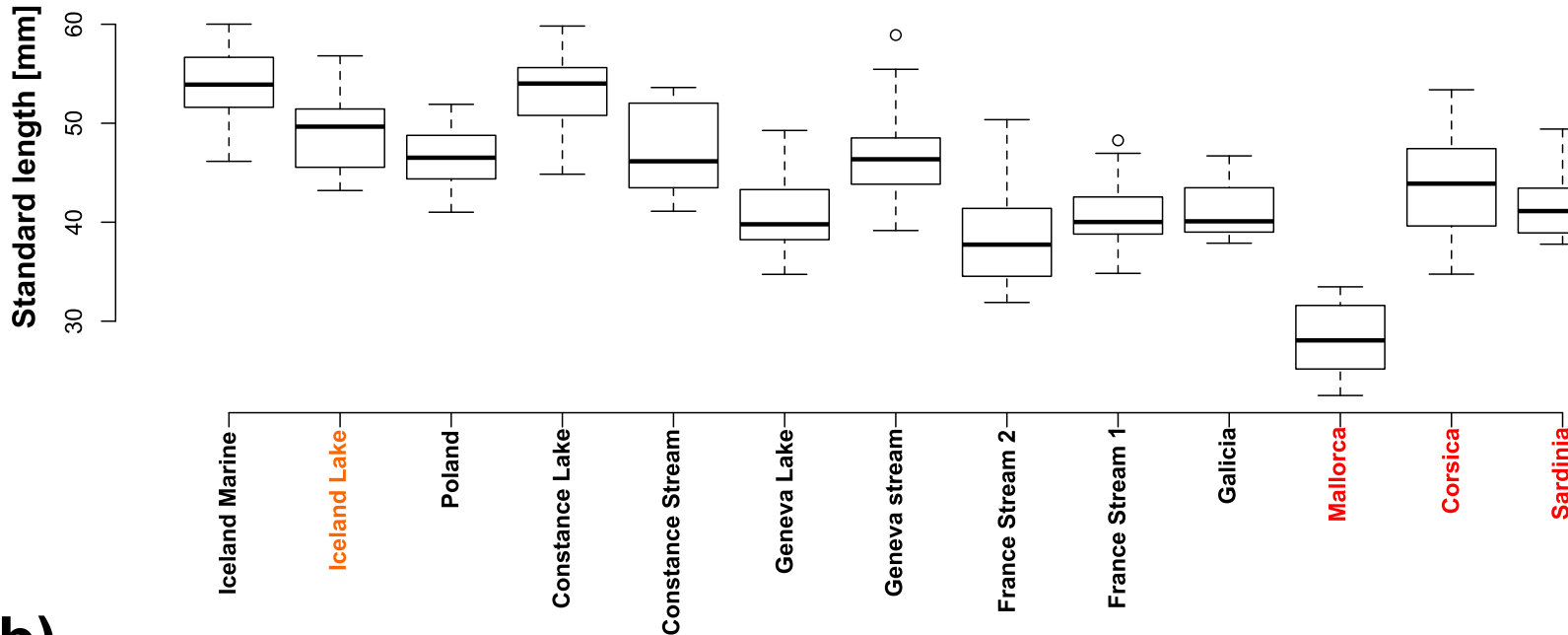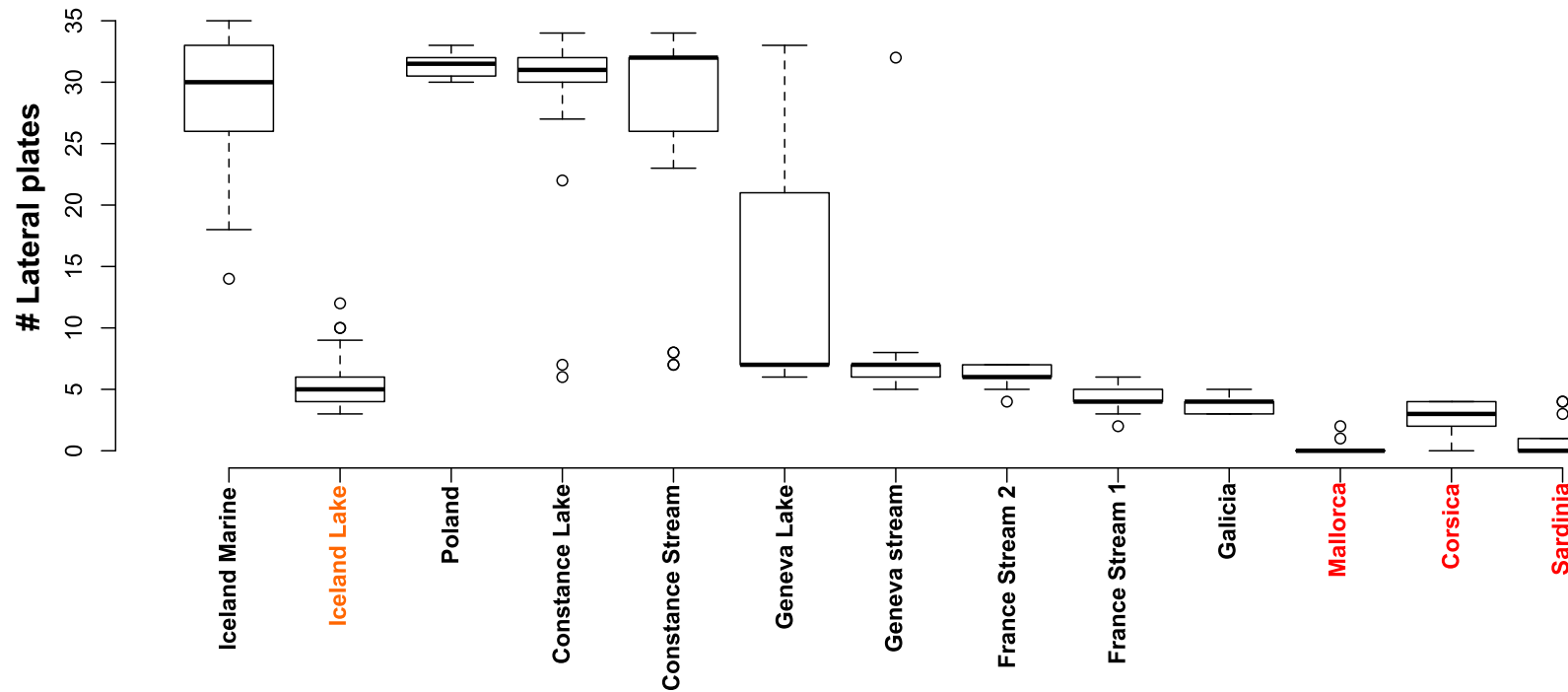
All traits

Defence traits
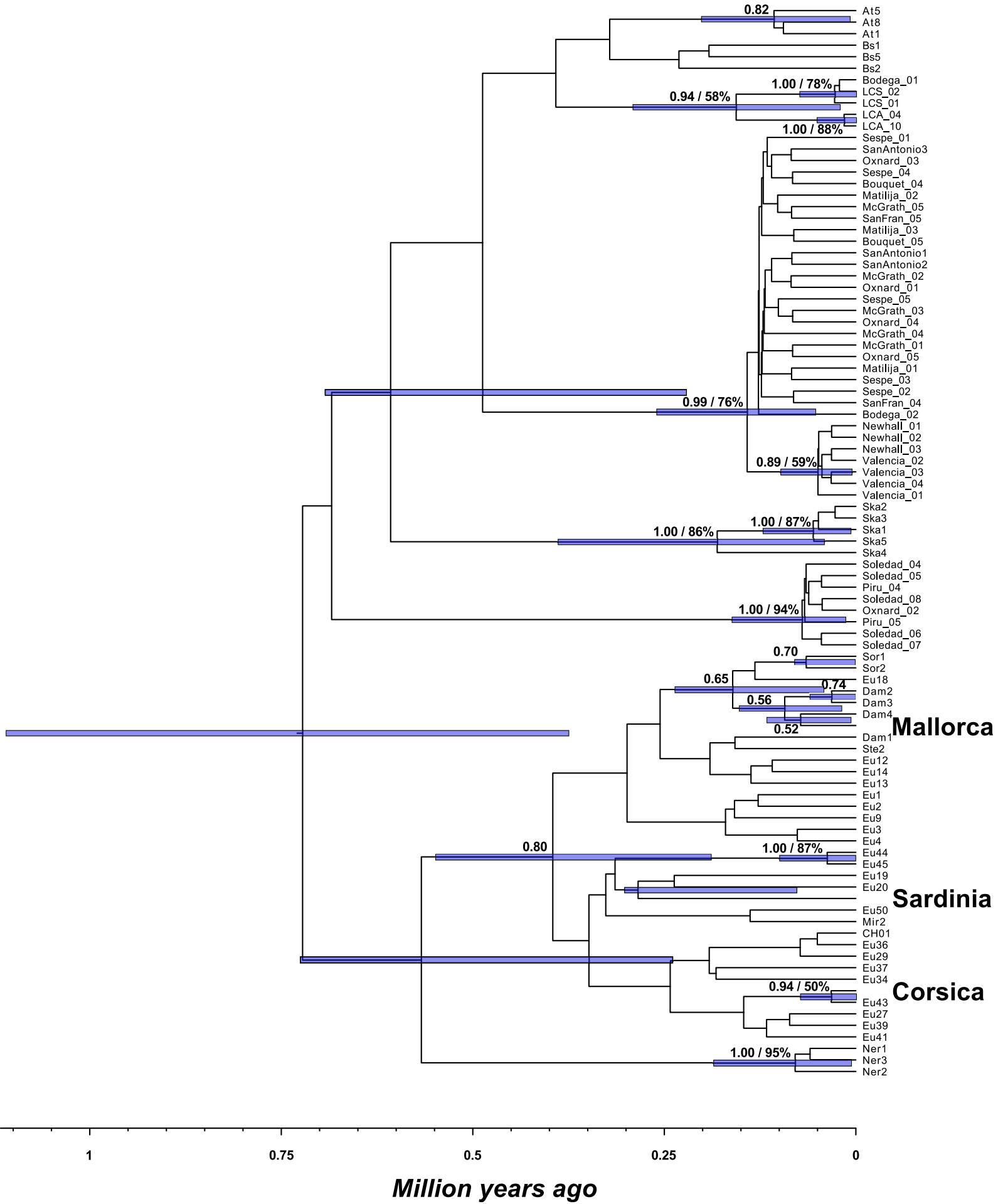
Figure 7
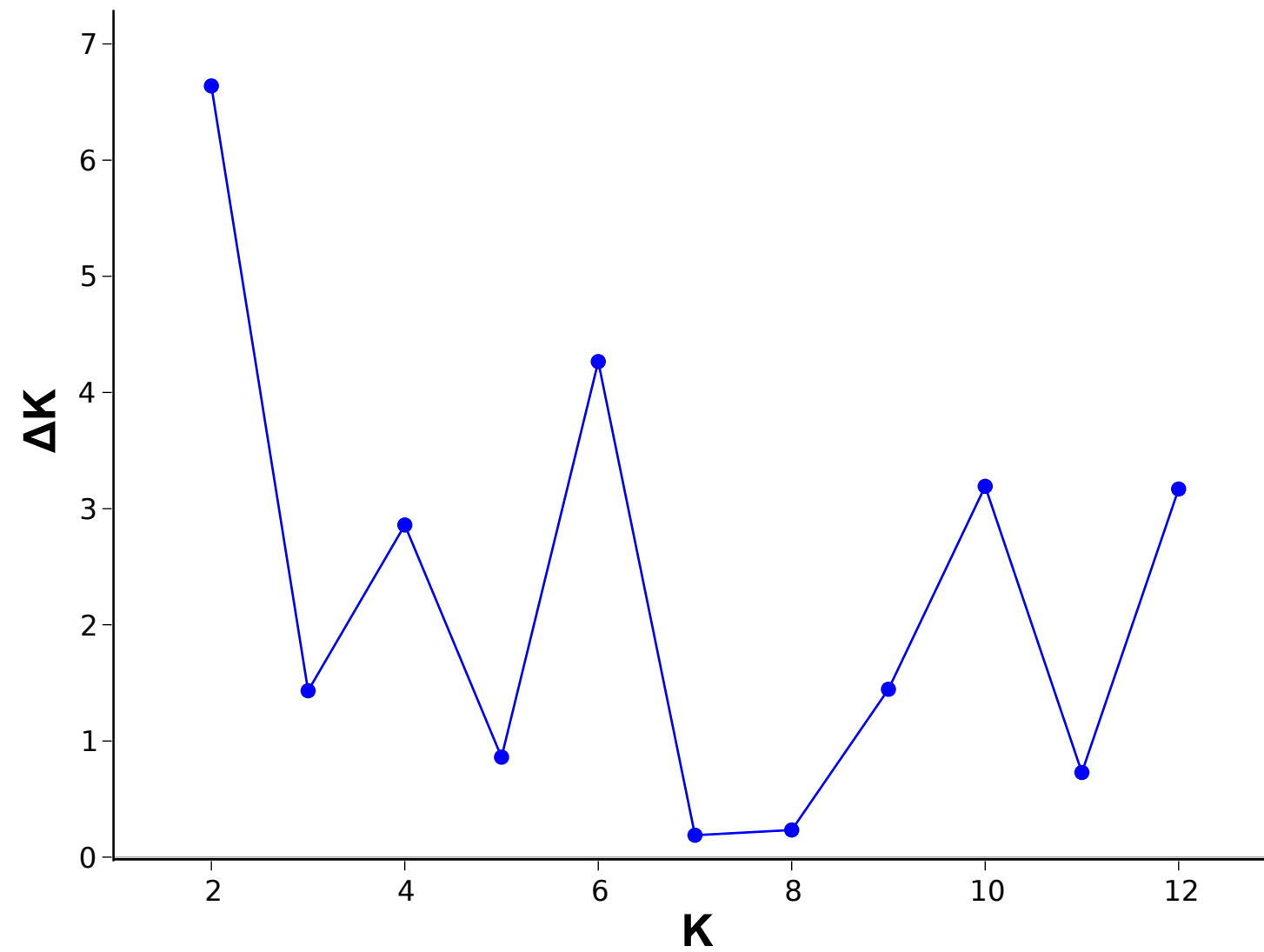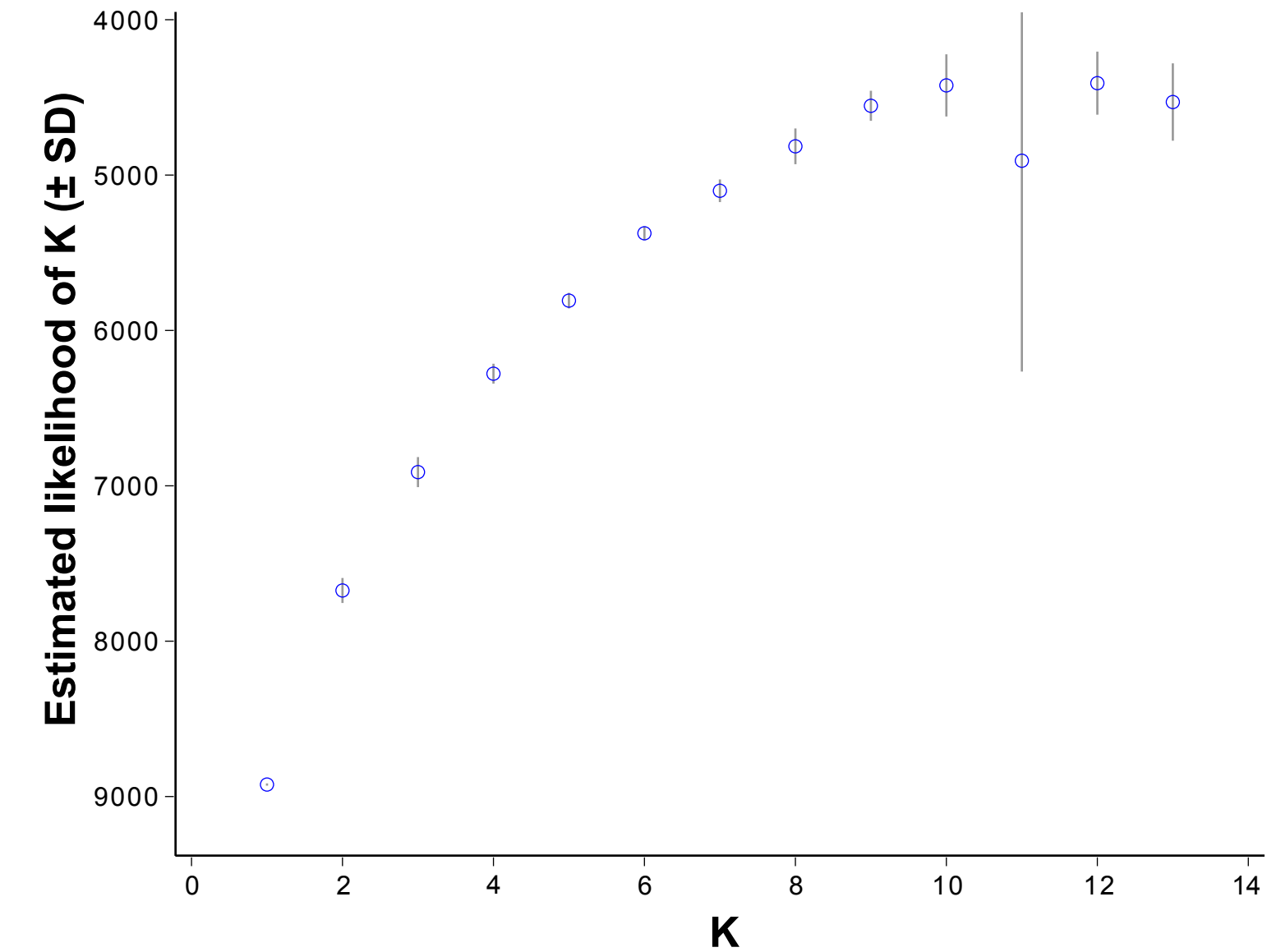Click here to download Figure: Fig_7.eps

Table

[Click here to download Table: Tables.docx](Tables.docx)

Table 1

Summary of all populations used in this study with their coordinates and country of origin, their relative geographical origin within Europe and their connectivity as well as the habitat they were sampled in and the sample size for morphology, stomach content analysis and microsatellites respectively (see main text for details).

Population	Country	Geography	Connectivity	Habitat	Latitude (N)	Longitude		$N_{Morphology}$	$N_{Stomachs}$	$N_{\mu sats}$
Mallorca	Spain	South	Island	Stream	39°47'18.5"	3°04'39.1"	E	26	26	29
Galicia	Spain	South	Mainland	Stream	42°36'13.2"	7°43'16.1"	W	9	9	30
Corsica	France	South	Island	Stream	41°59'46.1"	9°22'41.5"	E	32	32	29
Sardinia	Italy	South	Island	Stream	39°26'06.2"	8°45'13.0"	E	-	-	21
Sardinia*	Italy	South	Island	Stream	39°16'9.32"	8°48'40.6"	E	14	-	-
France Stream 1	France	South	Mainland	Stream	43°54'33.6"	5°05'08.7"	E	20	20	21
France Stream 2	France	South	Mainland	Stream	45°58'03.6"	5°17'39.6"	E	32	-	26
Geneva Lake	Switzerland	South	Mainland	Lake	46°31'02.0"	6°34'40.7"	E	40	37	29
Geneva Stream	Switzerland	South	Mainland	Stream	46°12'47.7"	7°18'25.5"	E	51	33	30
Constance Lake	Switzerland	North	Mainland	Lake	47°29'08.4"	9°32'38.4"	E	30	25	27
Constance Stream	Switzerland	North	Mainland	Stream	47°19'36.8"	9°34'04.5"	E	30	32	23
Poland	Poland	North	Mainland	Stream	50°03'13.1"	19°14'46.4"	E	20	-	22
Iceland Lake	Iceland	North	Island	Lake	65°36'52.1"	17°03'30.1"	W	30	-	31
Iceland Marine	Iceland	North	Island	Marine	65°02'39.5"	22°27'28.6"	W	45	-	41

* Samples from the Natural History Museum Vienna (NHW-83118)

Table 2

Measures of genetic diversity: A_N – average number of observed alleles within a population, A_R – effective number of alleles (number of alleles weighted for their frequencies), H_O – observed heterozygosity, H_S – within population heterozygosity, G_{IS} inbreeding coefficient.

Population	A_N	A_R	H_O	H_S	G_{IS}
Mallorca	2.857	1.609	0.291	0.315	0.078
Galicia	2.714	1.669	0.261	0.282	0.076
Corsica	3.286	1.457	0.207	0.286	0.277
Sardinia	3.000	1.863	0.279	0.308	0.094
France Stream 1	2.571	1.557	0.291	0.286	-0.017
France Stream 2	2.571	1.504	0.211	0.270	0.219
Geneva Lake	2.857	1.683	0.324	0.327	0.011
Geneva Stream	3.000	1.851	0.290	0.314	0.075
Constance Lake	3.571	2.356	0.506	0.502	-0.009
Constance Stream	3.000	2.131	0.461	0.497	0.073
Poland	2.857	1.999	0.414	0.465	0.110
Iceland Lake	4.571	2.170	0.307	0.445	0.310
Iceland Marine	8.571	4.350	0.690	0.713	0.033

Table S1

Pairwise genetic relationship among populations based on F_{ST} (lower triangle) with their respective p value based on 10,000 bootstrap replicates (upper triangle).

	France Stream 1	Sardinia	Poland	Galicia	Mallorca	Geneva Stream	Geneva Lake	Constance Lake	Corsica	Constance Stream	Iceland Marine	Iceland Lake	France Stream 2
France Stream 1	-	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Sardinia	**0.694**	-	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Poland	**0.487**	**0.520**	-	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Galicia	**0.654**	**0.587**	**0.401**	-	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Mallorca	**0.671**	**0.594**	**0.580**	**0.613**	-	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Geneva Stream	**0.606**	**0.660**	**0.391**	**0.572**	**0.670**	-	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Geneva Lake	**0.609**	**0.642**	**0.336**	**0.556**	**0.661**	**0.050**	-	0.001	0.001	0.001	0.001	0.001	0.001
Constance Lake	**0.552**	**0.529**	**0.285**	**0.503**	**0.527**	**0.384**	**0.335**	-	0.001	0.098	0.001	0.001	0.001
Corsica	**0.681**	**0.584**	**0.545**	**0.576**	**0.638**	**0.657**	**0.647**	**0.560**	-	0.001	0.001	0.001	0.001
Constance Stream	**0.560**	**0.526**	**0.313**	**0.521**	**0.531**	**0.383**	**0.344**	0.012	**0.568**	-	0.001	0.001	0.001
Iceland Marine	**0.349**	**0.334**	**0.130**	**0.241**	**0.375**	**0.262**	**0.241**	**0.186**	**0.384**	**0.195**	-	0.001	0.001
Iceland Lake	**0.582**	**0.520**	**0.406**	**0.462**	**0.492**	**0.483**	**0.459**	**0.381**	**0.533**	**0.377**	**0.201**	-	0.001
France Stream 2	**0.662**	**0.675**	**0.449**	**0.590**	**0.690**	**0.149**	**0.144**	**0.432**	**0.681**	**0.435**	**0.286**	**0.518**	-

Table S2

Pairwise genetic relationship between populations based on Hedrick's F'_{ST} with their respective 95% confidence intervals based on 10,000 bootstrap replicates.

Pairwise relationship between		F'_{ST}	95%CI
France Stream 1	Constance Lake	0.887	0.865-0.908
France Stream 1	Constance Stream	0.893	0.869-0.917
France Stream 1	Corsica	0.917	0.876-0.949
France Stream 1	France Stream 2	0.881	0.853-0.903
France Stream 1	Galicia	0.859	0.818-0.893
France Stream 1	Geneva Lake	0.821	0.792-0.848
France Stream 1	Geneva Stream	0.801	0.766-0.832
France Stream 1	Iceland Lake	0.896	0.864-0.925
France Stream 1	Iceland Marine	0.694	0.654-0.734
France Stream 1	Mallorca	0.929	0.902-0.952
France Stream 1	Poland	0.698	0.651-0.747
France Stream 1	Sardinia	0.964	0.939-0.984
Sardinia	Constance Lake	0.873	0.849-0.898
Sardinia	Constance Stream	0.855	0.823-0.886
Sardinia	Corsica	0.760	0.721-0.799
Sardinia	France Stream 2	0.934	0.915-0.952
Sardinia	Galicia	0.760	0.737-0.784
Sardinia	Geneva Lake	0.916	0.894-0.936
Sardinia	Geneva Stream	0.938	0.918-0.957
Sardinia	Iceland Lake	0.792	0.745-0.832
Sardinia	Iceland Marine	0.700	0.648-0.748
Sardinia	Mallorca	0.806	0.781-0.832
Sardinia	Poland	0.802	0.762-0.842
Poland	Constance Lake	0.486	0.433-0.545
Poland	Constance Stream	0.546	0.493-0.605
Poland	Corsica	0.810	0.730-0.868
Poland	France Stream 2	0.621	0.583-0.660
Poland	Galicia	0.532	0.480-0.582
Poland	Geneva Lake	0.460	0.422-0.501
Poland	Geneva Stream	0.543	0.510-0.580
Poland	Iceland Lake	0.677	0.605-0.743
Poland	Iceland Marine	0.294	0.240-0.353
Poland	Mallorca	0.919	0.895-0.941
Galicia	Constance Lake	0.763	0.730-0.796
Galicia	Constance Stream	0.789	0.753-0.826
Galicia	Corsica	0.726	0.655-0.785
Galicia	France Stream 2	0.753	0.722-0.781
Galicia	Geneva Lake	0.727	0.695-0.757
Galicia	Geneva Stream	0.742	0.718-0.765
Galicia	Iceland Lake	0.644	0.586-0.699
Galicia	Iceland Marine	0.437	0.387-0.486

Galicia	Mallorca	0.816	0.787-0.844
Mallorca	Constance Lake	0.857	0.824-0.888
Mallorca	Constance Stream	0.856	0.823-0.886
Mallorca	Corsica	0.871	0.843-0.897
Mallorca	France Stream 2	0.975	0.963-0.986
Mallorca	Geneva Lake	0.964	0.947-0.978
Mallorca	Geneva Stream	0.968	0.952-0.981
Mallorca	Iceland Lake	0.730	0.677-0.777
Mallorca	Iceland Marine	0.773	0.729-0.813
Geneva Stream	Constance Lake	0.561	0.498-0.623
Geneva Stream	Constance Stream	0.544	0.469-0.618
Geneva Stream	Corsica	0.910	0.826-0.959
Geneva Stream	France Stream 2	0.148	0.089-0.213
Geneva Stream	Geneva Lake	0.049	0.016-0.090
Geneva Stream	Iceland Lake	0.706	0.639-0.768
Geneva Stream	Iceland Marine	0.499	0.447-0.551
Geneva Lake	Constance Lake	0.479	0.411-0.545
Geneva Lake	Constance Stream	0.484	0.413-0.556
Geneva Lake	Corsica	0.903	0.819-0.953
Geneva Lake	France Stream 2	0.150	0.090-0.219
Geneva Lake	Iceland Lake	0.674	0.599-0.743
Geneva Lake	Iceland Marine	0.466	0.415-0.518
Constance Lake	Constance Stream	0.028	-0.011-0.081
Constance Lake	Corsica	0.895	0.829-0.941
Constance Lake	France Stream 2	0.621	0.560-0.682
Constance Lake	Iceland Lake	0.658	0.598-0.716
Constance Lake	Iceland Marine	0.435	0.385-0.489
Corsica	Constance Stream	0.904	0.837-0.947
Corsica	France Stream 2	0.929	0.854-0.972
Corsica	Iceland Lake	0.786	0.717-0.839
Corsica	Iceland Marine	0.763	0.673-0.829
Constance Stream	France Stream 2	0.610	0.540-0.678
Constance Stream	Iceland Lake	0.635	0.576-0.692
Constance Stream	Iceland Marine	0.458	0.409-0.513
Iceland Marine	France Stream 2	0.526	0.471-0.578
Iceland Marine	Iceland Lake	0.437	0.362-0.515
Iceland Lake	France Stream 2	0.747	0.683-0.806

Table S3

Trait loadings based on principal component (CV) analyses that were conducted using either all available linear measurements or on defense related traits only.

Trait	Overall PCA		Defence traits	
	PC1	PC2	PC1	PC2
% Variance explained	39.2	17.2	72.2	16.6
Head length	0.069	-0.097		
Eye diameter	0.010	-0.077		
Body depth at first dorsal spine	-0.090	0.315		
Body depth at second dorsal spine	-0.141	0.401		
Caudal peduncle length	0.091	0.302		
First dorsal spine length	0.337	0.080	-0.341	-0.417
Second dorsal spine length	0.328	0.109	-0.333	-0.385
Pelvic spine length	0.520	0.063	-0.531	-0.418
Pelvic girdle width	-0.109	0.279	0.103	-0.003
Pelvic girdle length	0.646	-0.074	-0.693	0.710
Total length of pelvic fin	0.098	-0.368		
Basal length of anal fin	-0.055	-0.474		
Basal length of dorsal fin	-0.150	-0.405		
Snout length	0.035	-0.053		
Upper jaw length	0.025	-0.026		