Invasion genetics of the mummichog (*Fundulus heteroclitus*): recent anthropogenic introduction in Iberia

Teófilo Morim [1] , **Grant R Bigg** [2] , **Pedro M Madeira** [1] , **Jorge Palma** [1] , **David D Duvernell** [3] , **Enric Gisbert** [4] , **Regina L Cunha** [1] , **Rita Castilho** Corresp. [1]

[1] Centre for Marine Sciences (CCMAR), University of Algarve, Faro, Portugal

[2] Department of Geography, University of Sheffield, Sheffield, United Kingdom

[3] Department of Biological Sciences, Missouri University of Science and Technology, Rolla, Missouri, United States of America

[4] RTA, Sant Carles de la Ràpita, Spain

Corresponding Author: Rita Castilho
Email address: rita.castil@gmail.com

Human activities such as trade and transport have increased considerably in the last decades, greatly facilitating the introduction and spread of non-native species at a global level. In the Iberian Peninsula, *Fundulus heteroclitus*, a small euryhaline coastal fish with short dispersal, was found for the first time in the mid-1970s. Since then, *F. heteroclitus* underwent range expansions, colonizing the southern region of Portugal, southwestern coast of Spain and the Ebro Delta in the Mediterranean Sea. Cytochrome *b* sequences were used to elucidate the species invasion pathway in Iberia. Three Iberian locations (Faro, Cádiz and Ebro Delta) and 13 other locations along the native range of *F. heteroclitus* in North America were sampled. Results revealed a single haplotype, common to all invasive populations, which can be traced to the northern region of the species' native range. We posit that the origin of the founder individuals is between New York and Nova Scotia. Additionally, the lack of genetic structure within Iberia is consistent with a recent invasion scenario and a strong founder effect. We suggest the most probable introduction vector is associated with the aquarium trade. We further discuss the hypothesis of a second human-mediated introduction responsible for the establishment of individuals in the Ebro Delta supported by the absence of adequate muddy habitats linking Cádiz and the Ebro Delta. Although the species has a high tolerance to salinity and temperature, ecological niche modelling indicate that benthic habitat constraints prevent along-shore colonisation suggesting that such expansions would need to be aided by human release.

1   **Invasion genetics of the mummichog (*Fundulus heteroclitus*): recent anthropogenic**

2   **introduction in Iberia**

5   Teófilo Morim[1], Grant R. Bigg[2], Pedro M. Madeira[1], Jorge Palma[3], David D. Duvernell[4], Enric

6   Gisbert[5], Regina L. Cunha[1], and Rita Castilho*[1]

8   [1] Centre for Marine Sciences (CCMAR), University of Algarve, Campus de Gambelas, Faro

9   8005-139, Portugal.

10   [2] Department of Geography, University of Sheffield, Winter Street, Sheffield S10 2TN, UK.

11   [3] Department of Biological Sciences, Missouri University of Science and Technology, Rolla,

12   Missouri 65409, USA

13   [4] RTA - Sant Carles de la Ràpita, Crta Poble Nou km 5.5, P.O. Box 200, 43540 Sant Carles de la

14   Ràpita, SPAIN

17   *Corresponding author:

18   Email: rcastil@ualg.pt

19   Telephone: +351 918397282

21   Runing head: invasive Iberian mummichog

25   **Abstract**

26   Human activities such as trade and transport have increased considerably in the last decades,

27   greatly facilitating the introduction and spread of non-native species at a global level. In the

28   Iberian Peninsula, *Fundulus heteroclitus*, a small euryhaline coastal fish with short dispersal, was

29   found for the first time in the mid-1970s. Since then, *F. heteroclitus* underwent range

30   expansions, colonizing the southern region of Portugal, southwestern coast of Spain and the Ebro

31   Delta in the Mediterranean Sea. Cytochrome *b* sequences were used to elucidate the species

32   invasion pathway in Iberia. Three Iberian locations (Faro, Cádiz and Ebro Delta) and 13 other

33   locations along the native range of *F. heteroclitus* in North America were sampled. Results

34   revealed a single haplotype, common to all invasive populations, which can be traced to the

35   northern region of the species' native range. We posit that the origin of the founder individuals is

36   between New York and Nova Scotia. Additionally, the lack of genetic structure within Iberia is

37   consistent with a recent invasion scenario and a strong founder effect. We suggest the most

38   probable introduction vector is associated with the aquarium trade. We further discuss the

39   hypothesis of a second human-mediated introduction responsible for the establishment of

40   individuals in the Ebro Delta supported by the absence of adequate muddy habitats linking Cádiz

41   and the Ebro Delta. Although the species has a high tolerance to salinity and temperature,

42   ecological niche modelling indicate that benthic habitat constraints prevent along-shore

43   colonisation suggesting that such expansions would need to be aided by human release.

Introduction

44

45 As a consequence of human activities involving large distance marine transportation and trade,

46 worldwide marine biological invasion rates have increased dramatically in the last 30 years (e.g.,

47 Hulme 2009). Along the European coasts, there are reports of over 850 invasive species of

48 which, 237 occur along the eastern Atlantic coastal areas, 680 in the Mediterranean Sea and the

49 remaining in the Baltic (Galil et al. 2014). Marine invasive species pose a significant

50 environmental threat as they are recognized as one of the major drivers of biodiversity loss

51 (Millennium Ecosystem Assessment 2005), altering ecosystems and their dynamics, shifting the

52 community structure and displacing endemic species (Molnar et al. 2008). Negative impacts may

53 also be registered at the economic and social levels, affecting fisheries, aquaculture, tourism or

54 human health (Bax et al. 2003). Invasive species spread and occupy new marine and coastal

55 ecosystems through several maritime introduction vectors such as ballast water, biofouling of

56 vessels, aquaculture escape or ornamental species trade (see Williams et al. 2013 and references

57 therein).

58 Reconstructing the invasion pathways and identifying the putative source populations

59 with historical and contemporary vector records is a difficult task (Estoup & Guillemaud 2010;

60 Lawson Handley et al. 2011). In this context, molecular genetic data is a powerful tool to

61 reconstruct invasive history by identifying putative source populations and genetic bottlenecks

62 (Bock et al. 2015; Cristescu 2015). Although genetic data does not always allow for successful

63 identification of these processes, there are three conditions which increase the probability of

64 accurate reconstruction of the invasive pathway: (1) extensive sampling along the native range to

65 ensure all the putative source populations are known, (2) the presence of genetic structure in the

66 native range to narrow down the putative source regions and (3) a short amount of time passed

67 since the invasion so that processes such as genetic drift do not increase the genetic

68 differentiation between native and invasive populations (Geller et al. 2010).

69 The advent of ecological niche modelling (ENM) has contributed to the building of

70 environmental risk maps for biological invasion by allowing formal, quantitative, inclusion of

71 varying environmental characteristics, both in space and time, to be added to such

72 considerations. Such new approaches to invasion risk monitoring provide useful insights by

73 predicting both potential colonization routes and the probability of new invasions occurring

74 (Hulme 2009; Molnar et al. 2008).

75   The mummichog, *Fundulus heteroclitus* (Linnaeus, 1766), is a small teleost naturally

76   occurring almost continuously in saltmarshes of the North American east coast, from

77   Newfoundland to Florida (Hardy Jr 1978). This species is extremely resistant to a wide range of

78   salinities and temperatures, and can be found in freshwater, brackish or saltwater, inhabiting

79   sheltered coastal areas such as saltmarshes, tidal creeks, estuaries or bays all year-round

80   (Bigelow & Schroeder 1953; Hardy Jr 1978). *F. heteroclitus* is one of the most stationary marine

81   species (Bigelow & Schroeder 1953), with short dispersal distances (1-2 km, Fritz et al. 1975),

82   high site fidelity closely related to the presence of saltmarshes (Kneib 1984) and short home

83   ranges (36-38 m, Lotrich 1975). Enzyme-coding loci (Powers & Place 1978; Powers et al. 1986;

84   Ropson et al. 1990), mitochondrial DNA (Bernardi et al. 1993; González-Vilaseñor & Powers

85   1990; Smith et al. 1998), putative neutral nuclear microsatellite loci (Adams et al. 2006;

86   Duvernell et al. 2008), and a variety of nuclear single-nucleotide polymorphisms (see McKenzie

87   et al. 2016 and references therein) demonstrated concordance of clinal patterns, where a break is

88   placed between the meridians 40º and 41º N, along the coast of New Jersey. These studies concur

89   with Morin and Able (1983), supporting the division into two subspecies: *F. heteroclitus*

90   *macrolepidotus* (Walbaum, 1792)to the north of the cline and *F. heteroclitus heteroclitus*

91   (Linnaeus, 1766) to the south of the cline.

92   In the Iberian Peninsula, *F. heteroclitus* was first detected in the 1970s, in the

93   Guadalquivir and Guadiana saltmarshes (southwestern coast of Spain) (Hernando 1975) and in

94   the Guadiana Delta (Coelho et al. 1976). More recently, its presence was also reported in the

95   Ebro Delta, northeastern coast of Spain (Gisbert & López 2007) and in the Ria Formosa, south of

96   Portugal (e.g., Catry et al. 2006) where it can reach high densities (Gonçalves et al. 2017).The

97   species was probably introduced between 1970 and 1973 in the Spanish saltmarshes (Fernández-

98   Delgado 1989) either involuntarily via aquarium trade (e.g. Elvira & Almodóvar 2001), ballast

99   water (Fernández-Delgado 2010; García-Revillo & Fernández-Delgado 2009), or intentionally

100 for purposes of biological control (Gozlan 2010). The Ebro Delta individuals were probably

101 caught in the southern Spanish saltmarshes to be used in aquaculture and the aquarium trade

102 (Gisbert & López 2007) and were later released from captivity. Another possibility for the Ebro

103 Delta introduction is an accidental escape from a research centre nearby, where this species had

104 already been used as a scientific model (Gisbert & López 2007). Two previous studies have

105 analysed the origin of the southwestern Spanish populations using mitochondrial DNA

106 (mtDNA), concluding the founder individuals were original from the northern native region

107 between Maine and Nova Scotia where the northern subspecies *F. h. macrolepidotus* is

108 distributed (Bernardi et al. 1995; Fernández-Pedrosa et al. 1996).

109   In the present study, we aim to build on previously published studies on the invasive

110 range of *F. heteroclitus* by using (1) more sampling locations (one in the Mediterranean Sea and

111 two locations in the eastern Atlantic, and thirteen native locations), (2) a significantly larger

112 number of individuals (248 in total), and (3) a three times larger fragment of the mitochondrial

113 DNA cytochrome *b* gene; and to evaluate the genetic diversity and invasion pathways. Given

114 that the species was recently reported in the Iberian Peninsula and has a limited adult dispersal

115 capability, we tested the hypothesis of a human-mediated single Iberian introduction followed by

116 dispersal promoted along the main oceanographic currents. and we mapped the environments

117 compatible with the species' ecological requirements to evaluate the dispersion potential through

118 suitable continuous habitat. This hypothesis leads to the expectation of an Iberian invasion based

119 on a few founder individuals, with consequent lower genetic diversity than the putative identified

120 source population. Also, the Iberian populations are expected to show no evident genetic

121 structure given the short time since invasion.

122

123 **Materials and methods**

124 **Sampling.** A total of 248 *Fundulus heteroclitus* individuals from 16 locations: 13 sites in the

125 western Atlantic, one in the Mediterranean Sea and two locations in the eastern Atlantic (Table 1

126 and Fig. 1a,b) were obtained and stored in 96% ethanol and kept at -20 ºC. The populations from

127 the western Atlantic sampled above 40º N are hereafter referred to as northern locations (ID 1-8),

128 while samples collected below that latitude are referred to as southern locations (ID 9-13).

129 Samples collected from the Mediterranean and eastern Atlantic are referred to as Iberian (ID 14-

130 16).

131

132 **DNA extraction, PCR amplification and sequencing.** Total genomic DNA was extracted from

133 caudal fin tissue following a standard Chelex 100 protocol (Walsh et al. 1991). Extraction results

134 were checked by electrophoresis in 0.8% agarose gel stained with GelRed. Polymerase Chain

135 Reactions (PCR) were conducted in a total volume of 25 μL, with 1X buffer, 10 mM dNTPs,

136 10mM of each primer, 1U Taq Advantage 2 Polymerase mix DNA polymerase (CLONTECH-

137   TaKaRa), 2 μL of DNA and Milli-Q water to the final volume. A fragment of the cytochrome *b*

138   (cyt *b*) gene (1000 base pairs) was amplified with the forward primer GludG-L14724 (Palumbi et

139   al. 1991) and the reverse primer cb6b.h (Martin & Bermingham 1998). PCR amplification

140   consists of an initial 4 min denaturation step at 95 ºC, followed by 40 cycles of 1 min at 94 ºC

141   (denaturation), 1 min at 50 ºC (annealing) and 1.5 min at 72 ºC (extension), and a 5 min final

142   extension step. When amplification was not successful the following profile was used: initial 3

143   min denaturation step at 95 ºC, followed by 32 cycles of denaturation for 30 s at 95 ºC, annealing

144   for 30 s at 54 ºC and extension for 1 min at 68 ºC, and a final extension step for 4 min at 68 ºC.

145   PCR products were checked afterwards by electrophoresis in a 1% agarose gel stained with

146   GelRed. Mitochondrial DNA was purified by ethanol/sodium acetate precipitation (Sambrook &

147   Russel 2001). Its purity and quantity were analysed using a NanoDrop1000 spectrophotometer

148   (Thermo Fisher Scientific, USA). Sequencing was performed on an ABI 3130xl capillary

149   sequencer (Applied Biosystems – CCMAR, Portugal) using the forward primer from the PCR

150   amplification (GludG-L14724).

151

152   **Genetic analysis.** Cyt *b* sequences were aligned and manually checked using the software

153   Geneious v4.8.2 (Biomatters, Ltd., Auckland, New Zealand). The number of haplotypes (n),

154   number of private haplotypes (n_p), and the haplotype (h) (Nei & Tajima 1981) and nucleotide

155   diversities (π) (Nei 1987) were calculated for each location using the DNAsP v5.10.1 (Librado &

156   Rozas 2009). All sequences were uploaded to GenBank (MH809691 - MH809938). To represent

157   the phylogeographic relationships among haplotypes, a haplotype network was constructed using

158   the Median Joining algorithm implemented in NETWORK v5.0 (Bandelt et al. 1999, fluxus-

159   engineering.com).

160   We used two approaches to infer the most probable source area of the introduced Iberian

161   populations within the native range of *F. heteroclitus*: (1) the geographical distribution of

162   haplotypes in native populations, and (2) the phylogeographical relationships among haplotypes.

163

164   **Ecological niche modelling.** To explore the potential spread of *F. heteroclitus* within European

165   waters beyond its currently known locations we used an ecological niche model. We selected the

166   interactive AquaMaps system in FISHBASE (www.fishbase.org). Using observed locations of

167   adults to construct a range of acceptable environmental conditions within which a species can

168 exist, this model uses a c-squares distribution modelling approach (Rees 2003) to predict the

169 probability of occurrence of the adult fish for a particular location, possessing specific

170 environmental parameters. While this is normally computed just within the native range of a

171 species it can be extended within the interactive system to consider all possible locations, and the

172 associated probability of occurrence linked to the location's mean environmental state if the

173 species was able to reach that region. The basic statistical approach is given in Kaschner et al.

174 (2006) and the AquaMaps version is described in Kesner-Reyes et al. (2012). Another ecological

175 niche model, based on maximum entropy principles (MAXENT, Phillips & Dudík 2008), was

176 also tested using environmental parameters of sea surface temperature (SST) and sea surface

177 salinity (SSS), but the specialist ecological substrate niche of *F. heteroclitus* led to poor solutions

178 for range prediction using ENM. This latter approach is therefore not considered further here.

179   A set of favourable and extreme environmental conditions compatible with the native

180 occurrence of *F. heteroclitus* is automatically specified by AquaMaps, relying heavily on Page

181 and Burr (2011). It includes parameter ranges for water depth, SST, SSS, primary production

182 rates and sea-ice cover, leading to probabilities of occurrence exceeding 0.6 along the whole

183 eastern seaboard of North America from South Carolina to the Canadian Maritime Provinces

184 south of the Gulf of St. Lawrence. This matches well to the known range (Page & Burr 2011).

185 However, the long-term survival of *F. heteroclitus* in two Iberian estuaries has implications for

186 extending the extreme salinity level that this species can tolerate beyond that automatically

187 specified. The extreme salinity tolerance has therefore been raised to 39.6 psu, compared to the

188 automatic 36.47 psu initially prescribed by AquaMaps, consistent with the values found for the

189 Ria Formosa (Cristina et al. 2016). In addition, the sea-ice variable has been excluded from

190 constraining the AquaMaps solution, as this is not relevant to the area being considered in this

191 study, and also the primary production variable, as coastal environments satisfy the automatic

192 requirement almost uniformly. The environmental variables used for the AquaMaps simulation

193 shown here are given in Table 2.

194   AquaMaps was re-run with these environmental constraints and using the combined

195 occurrence data from the native range, along the North American eastern seaboard, and the two

196 Iberian sites of Ria Formosa and the Ebro River Delta. There is no change to the North American

197 predictions, so here we concentrate only on those for European waters. The occurrence of *F.*

198 *heteroclitus* was further constrained to coastal zones with the muddy benthic habitats in which *F.*

199   *heteroclitus* is found. These are defined as the coastal fine muds to muddy sands benthic

200   environments of the EMODnet Seabed Habitats project (www.emodnet-seabedhabitats.eu).

201   These environments are formally defined by Long (2006), but basically include sediments with

202   more than a 10% mud component.

203

204

205   **Results**

206   Data analysis on the sampled specimens resulted in cyt *b* sequences with a final length of 700

207   base pairs (bp), which comprised 77 (11%) polymorphic sites and 32 (41.6%) parsimony-

208   informative sites. These polymorphisms defined 70 haplotypes, of which 62 (88.6%) are private

209   haplotypes (present in one location only) and 55 (78.6%) are singletons (present in one

210   individual only). Overall haplotype diversity was high (0.74 ± SD 0.03), ranging from null in

211   Iberia (ID 14-16) to 0.93 in Tuckerton (ID 9) and Roanoke Island (ID 12), whereas mean

212   nucleotide diversity was low (0.26% ± SD 0.06%) ranging from null diversity in Iberia to 0.91%

213   in Woods Hole (ID 4) (Table 1 and Fig. 2).

214   The most abundant haplotype in North American locations is shared by 50.4% ($N = 125$)

215   of the individuals and is present in all northern group locations, in one southern location and in

216   all Iberian locations (Fig. 1c). This is the only haplotype detected in the invasive range of the

217   species (Faro, Cádiz and Ebro Delta). The second most frequent haplotype in North American

218   locations is shared by 9.7% ($N = 24$) individuals in six locations, although in higher frequency in

219   the southern group locations. Two other haplotypes were found in 3.2% ($N = 8$) and in 2.4% ($N = 6$)

220   6) of the individuals from two locations. All other haplotypes were present in five or less

221   individuals and in less than three locations (Fig.1c).

222   The cyt *b*-based haplotype network (Fig. 3) displays two haplogroups separated by 9

223   mutational steps. Haplogroup A is constituted by all eight northern group locations (Bridgewater

224   to Red Bank, ID 1-8), two individuals from the southern group (Suffolk, ID 11) and by the

225   Iberian locations (Faro to Ebro Delta, ID 14-16). Haplogroup B is formed by all the other

226   individuals from all southern locations (Tuckerton to Skidaway Island, ID 9-13) and includes

227   eight individuals from northern locations: one from Bridgewater (ID 1) and seven from Woods

228   Hole (ID 4). Overall, both haplogroups display star-like configurations with different levels of

229   complexity. The northern haplogroup (A) is simpler with 90% of the haplotypes separated by a

230   single mutation, while the southern haplogroup (B) is more complex, with 3 mini-stars

231   interconnected by one mutation each, and haplotypes separated by up to five mutations. The

232   haplotype found in the Iberian Peninsula belongs to the northern haplogroup.

233   The AquaMaps ENM shows that the basic environment for the spread of *F. heteroclitus*

234   is fundamentally favourable (probability>0.75), along much of the Atlantic coastline of western

235   Europe (Fig. 4). Conditions become less favourable in the Mediterranean, although the Alboran

236   Sea, east of the Strait of Gibraltar, has an environment that is at least acceptable

237   (probability>0.25). There are also a small number of estuaries along the Balearic Sea coastline of

238   NE Spain and southern France where acceptable conditions are also found. The main constraint

239   on spread of *F. heteroclitus*, however, is the absence of muddy benthic habitats. The latter are

240   shown by the solid lines in Fig. 4. All the European coastlines within the zone where *F.*

241   *heteroclitus* is now found have a small number of muddy estuaries separated by long stretches of

242   unsuitable, rockier, benthic habitats. It is only along the environmentally unfavourable North

243   African coast that extensive areas of favourable habitats are found. The ENM shows

244   conclusively that lateral spread of *F. heteroclitus* along European shorelines by natural means is

245   very unlikely.

246

247

248

249   **Discussion**

250   Results revealed the presence of a single haplotype common to all individuals in the Iberian

251   Peninsula. This haplotype is the most abundant in the northern group of the native distribution,

252   indicating the most probable origin of the invasion. We posit as the most likely introduction

253   vector the aquarium trade and propose, through a combination of DNA and ENM evidence, that

254   the Ebro Delta colonization results from an independent human-mediated secondary

255   introduction. Before addressing the main interpretations and conclusions of these results, one

256   main caveat must be addressed. Successful understanding of the invasion pathway relies on (1)

257   comparable genetic data retrieved from an adequate number of sampled individuals throughout

258   the entire native range, (2) presence of genetic clines within the native populations and (3) the

259   use of adequate molecular markers (Geller et al. 2010). While the first two points are fulfilled for

260   *F. heteroclitus*, the use of a single mitochondrial DNA marker constitutes the main caveat of this

261   study. Mitochondrial DNA has been a widely used molecular marker in population genetics

262   studies (Ballard & Whitlock 2004). Nevertheless, the use of high variable nuclear markers, such

263   as microsatellites, provides an opportunity to perform assignment tests based on their multiple-

264   locus genotypes, to test for recent reductions in population sizes and to estimate effective

265   population sizes. However, previously published studies show relatively low microsatellite

266   genetic variation in the northern group (Duvernell et al. 2008). Single nucleotide polymorphisms

267   (SNPs, Morin et al. 2004), extend the previously referred analytical possibilities improving their

268   statistical power because of the sheer number of existing loci genome wide.

269

270

271   **Genetic diversity**

272   The presence of a single haplotype common to all *F. heteroclitus* sampled in the Iberian

273   Peninsula lends support to the hypothesis of an extremely recent introduction of the species

274   which has not allowed the accumulation of mutations at the mtDNA level, and with a single

275   introduction event composed by a very small number of individuals (Roman & Darling 2007).

276   Theoretically, invasive species are expected to suffer loss of genetic variation since the new

277   established populations are often based on a few individuals, which by definition, have lower

278   genetic diversity than the native source populations (Dlugosch & Parker 2008). The single-

279   haplotype characteristic can be found in other invasive species, such as *Equulites elongatus,* the

280   slender pony fish (Sakinan et al. 2017); *Cercopagis pengoi*, the fishhook waterflea, a planktonic

281   cladoceran crustacean (Cristescu et al. 2001); *Corbicula fluminea*, the Asian clam (Gomes et al.

282   2016) and *Didemnum perlucidum,* a sea squirt (Dias et al. 2016). However, many successful

283   invasive species do not display significant erosion of genetic diversity (Dlugosch & Parker

284   2008). For example, in a recent review of the literature on European seas invasion genetics, in

285   54% of studies that compared the genetic diversity between introduced species and their native

286   range, 74% reported comparable levels of diversity between them, while only 23% displayed a

287   reduction in the genetic diversity of introduced species, and the remaining 2% showed an

288   increase in diversity (Rius et al. 2014).

289

Population sources

290   **Population sources**

291   We identified a single Iberian haplotype present in all northern populations in high frequency

292   (between 47 and 93%) and in two individuals in Suffolk, one of the southern locations.

293   According to our results, the northern group is the most probable source of the founder

294   individuals, which corresponds to the natural range of the subspecies *F. h. macrolepidotus*.

295   However, we cannot definitely exclude Suffolk as a presumptive population source. The absence

296   of genetic diversity in the Iberian Peninsula populations prevents the precise determination of the

297   putative source population. The low spatial resolution of our data arises from the lack of genetic

298   variability in the invasive range, rather than insufficient sampling of *F. heteroclitus* individuals

299   in the Iberian Peninsula or in its native range (Muirhead et al. 2008).

300   Nevertheless, our findings are consistent with two previous studies on the origin of invasive

301   individuals found in the Guadalquivir. First, based on mtDNA restriction fragment length

302   polymorphisms (RFLP) (Fernández-Pedrosa et al. 1996) reported the presence of two

303   haplotypes: the most abundant corresponding to the northern haplotype 1, dominant between

304   Maine and Nova Scotia in North America; the other haplotype did not match any of the sampled

305   native haplotypes and we found no evidence of its presence in the present study, using a larger

306   number of individuals. It was previously suggested to be either a native unsampled haplotype or

307   an endemic haplotype from Iberia, which is rather unlikely due to its recent invasion

308   (Fernández-Pedrosa et al. 1996) and total absence of records in the area. Although there are no

309   reports of hybridization between *F. heteroclitus* and any of the Spanish endemic species, the

310   presence of a new haplotype could nonetheless be due to hybridization (Rius et al. 2014).

311   Secondly, a study based on cyt *b* sequences (Bernardi et al. 1995) concluded the individuals from

312   the Guadalquivir originated in the region between Maine and Nova Scotia. While we cannot

313   discount the possibility of a few rogues from Suffolk being responsible for the invasion, the

314   weight of probability falls heavily on the side of Maine to Nova Scotia origin.

315

Introduction vector

316   **Introduction vector**

317   Since it was first recorded in Iberian saltmarshes, several studies linked *F. heteroclitus*

318   introduction to different vectors. The aquarium trade has been suggested as the most important

319   vector responsible for the introduction of this species (e.g. Gozlan 2010 and references therein),

320   followed by ballast water (Fernández-Delgado 2010; García-Revillo & Fernández-Delgado

321   2009), biological control (Gozlan 2010) and unknown origins (Fernández-Delgado 1989;

322   Gutiérrez-Estrada et al. 1998). We posit that the introduction via the aquarium trade or animal

323   acquisition for scientific purposes is the most probable scenario responsible for the establishment

324   of the first individuals in Iberia, given the genetic results obtained.

325   It is well known that vessels can transport large numbers of organisms from several

326   species at the same time in their ballast water (Gollasch 2007). Since there are countless vessels

327   active around the world (e.g., Kaluza et al. 2010), ballast water-mediated transport allows the

328   possibility of multiple introduction events, each with large groups of individuals (Hulme 2009).

329   Considering such large groups usually comprise higher genetic diversity than fewer individuals

330   alone, the assemblages transported are likely to display levels of genetic diversity similar to the

331   levels found within their native range (Wilson et al. 2009). Our findings, however, suggest that

332   this was not the introduction vector responsible for the spread of the species; in contrast to the

333   expectation of similar levels of genetic diversity between the native and invasive range, the

334   Iberian locations display a strong founder effect, with all samples sharing a single haplotype.

335   Similarly to ballast water, the aquarium and ornamentals trade transport many species at a

336   global scale (Padilla & Williams 2004). For instance, at least 19% of the invasive fishes found in

337   the Iberian Peninsula were introduced via the aquarium trade (Maceda-Veiga et al. 2013).

338   However, there are two main differences between introductions that follow ballast water or the

339   aquarium trade. First, each introduction event after aquarium release is likely to comprise a small

340   number of individuals (Duggan et al. 2006). Thus, even though this vector may be responsible

341   for the establishment of several non-native species (Padilla & Williams 2004), multiple

342   introductions would be necessary for an invasive species to display high genetic diversity

343   (Roman & Darling 2007). Secondly, the individuals released by aquarists are usually adults of

344   higher fitness, which makes them better adapted to survive in a natural environment (Padilla &

345   Williams 2004). Thus, not only are the *F. heteroclitus* invasive genetic diversity and structure

346   consistent with an introduction of a low number of individuals via the aquarium trade, but it is

347   also plausible that a few resistant individuals would manage to survive, reproduce and colonize

348   the environment in which they were released.

349

350   **Human mediated introduction in the Ebro Delta**

351   Although the absence of genetic structure within Iberia limits possible insights into the invasion

352   pathway, our data support the hypothesis of a human-mediated introduction episode responsible

353   for the establishment of *F. heteroclitus* in the Ebro Delta, as previously suggested by Gisbert and

354   López (2007) based on taxonomic identification. The hypothesis of a long-distance colonization

355   via natural dispersal is quite unlikely, as is strongly shown by the ENM analysis (Fig. 4).

356   According to a review of the geographical distribution of Cyprinodontiformes along the

357   northeastern coast of Spain by (García-Berthou & Moreno-Amich 1991), no *F. heteroclitus*

358   individuals were found at the Ebro Delta in 1989; the first record of the species was only

359   registered 16 years later by Gisbert and López (2007). Thus, we estimate the date of

360   establishment in the Ebro Delta ranges between 12 and 28 years ago. Assuming this estimate is

361   correct, a natural colonization hypothesis implies individuals would have taken roughly two

362   decades to travel more than 1000 km from their southernmost limit located in the Guadalquivir

363   saltmarshes (Gutiérrez-Estrada et al. 1998) to the Ebro Delta. However, when compared with the

364   natural colonization of the Ria Formosa, this hypothesis seems quite improbable. While *F.*

365   *heteroclitus* was never collected during sampling events that happened in the Ria Formosa

366   between 1980 and 2006 (França et al. 2009; Ribeiro et al. 2006; Ribeiro et al. 2008), analysis of

367   prey remains left by Little Terns (*Sterna albifrons*) in the salt-pans and barrier islands revealed

368   this prey species was present in the salt-pans and adjacent channels at least since 2002 (Catry et

369   al. 2006). Although this may sound contradictory, *F. heteroclitus* could in fact have been present

370   in the Ria Formosa in specific unsampled locations or at extremely low densities, avoiding

371   capture. Nonetheless, assuming that colonization happened around 2002, it seems that *F.*

372   *heteroclitus* took no more than ca. 20 years to travel around 50 km from the Guadiana Delta,

373   where it was first detected in 1976 (Coelho et al. 1976)}, despite both areas being on a stretch of

374   coast with a favourable benthic habitat (Fig. 4). This estimate indicates that if the Ebro

375   colonization happened via natural dispersal, it must have happened 20 times faster than the

376   natural colonization of the Ria Formosa. Given that *F. heteroclitus* has very low dispersal

377   abilities (e.g., Fritz et al. 1975; Lotrich 1975), the natural colonization scenario for the Ebro

378   seems unlikely.

379   Furthermore, if establishment in the Ebro Delta followed a natural range expansion one

380   would expect to find several established populations between the Barbate marshes in the

381   Guadalquivir, its southernmost location (Gutiérrez-Estrada et al. 1998), and the Ebro Delta.

382 However, no invasive individuals were found at Mar Menor, a coastal lagoon with suitable

383 conditions for successful establishment (Pérez-Ruzafa et al. 2006), located midway between the

384 Guadalquivir and the Ebro Delta.

385   Finally, the Strait of Gibraltar would represent a barrier to dispersal and gene flow,

386 limiting *F. heteroclitus* natural range expansion towards the Mediterranean Sea (Doadrio et al.

387 2002) because of the strong currents prevailing in the area and the absence of suitable habitats.

388 Although uncertain for *F. heteroclitus,* this movement has already been documented for two

389 other Iberian toothcarps, where restricted gene flow in this region led to speciation of the

390 *Aphanius iberus* in the Mediterranean Iberian coast and *Aphanius baeticus* in the southwestern

391 Atlantic Spanish coast (Doadrio et al. 2002).

392

393 **Ecological niche modelling**

394 Conditions for the spread of *F. heteroclitus* are limited by the existence of benthic muddy

395 saltmarsh environments compatible with the species' ecological requirements. These habitats are

396 only found near major estuaries or lagoons areas along the Atlantic and western Mediterranean

397 coastlines. The exception to this is along the North African coast (Fig. 4), where such

398 environments are more continuous, although other aspects may be less favourable there, as

399 suggested by the rapid decline in occurrence probability eastwards. The consequence of this

400 benthic habitat constraint is to make along-shore colonisation unlikely in most areas, suggesting

401 that any such sudden expansion would need to be aided by human release.

402

403

404 **Conclusions**

405 *Fundulus heteroclitus* invasive Iberian populations revealed the presence of a single cyt *b*

406 haplotype common to all individuals. This haplotype is predominant in the northern group of the

407 native distribution, and although we cannot determine which exact location was at the origin of

408 the introduced individuals, one can identify the northern end of range as the source of the

409 introduction. The lack of genetic diversity is consistent with a strong founder effect at the origin

410 of *F. heteroclitus* in Iberia. Although there is no direct evidence, we infer that the most likely

411 vector was the aquarium trade, and that the Ebro Delta colonization results from a human-

412 mediated secondary introduction isolated from the rest of Iberia. Considering the tolerance of the

413 species to high salinity, its temperature range, and the significant amount of colonized area in the

414 Iberian southern region, we predict that *F. heteroclitus* will most likely keep on expanding its

415 invasive range until it faces unfavourable environmental conditions. However, natural

416 colonisation in Europe will be strongly restricted by its restrictive requirement for suitable

417 muddy, benthic habitats and human-mediated transfer is its most likely means of range

418 expansion.

419

420

421 **Acknowledgements**

422 Occurrence data for the AquaMaps modelling was provided through the AquaMaps website

423 (www.aquamaps.org), and the model simulation was run through the same site. Information on

424 seabed habitats contained in this study has been derived from data that is made available under

425 the European Marine Observation Data Network (EMODnet) Seabed Habitats project

426 (http://www.emodnet-seabedhabitats.eu/), funded by the European Commission's Directorate-

427 General for Maritime Affairs and Fisheries (DG MARE). Authors are thankful to N. Franch and

428 J.M. Queral (Parc Natural del Delta de l'Ebre, PNDE) for providing animals from the Ebro River

429 Delta.

430

431

432 **References**

433 Adams SM, Lindmeier JB, and Duvernell DD. 2006. Microsatellite analysis of the
434 phylogeography, Pleistocene history and secondary contact hypotheses for the killifish,
435 *Fundulus heteroclitus*. *Molecular Ecology* 15:1109-1123. https://doi.org/10.1111/j.1365-
436 294x.2006.02859.x
437 Ballard JWO, and Whitlock MC. 2004. The incomplete natural history of mitochondria.
438 *Molecular Ecology* 13:729-744. https://doi.org/10.1046/j.1365-294X.2003.02063.x
439 Bandelt HJ, Forster P, and Röhl A. 1999. Median-joining networks for inferring intraspecific
440 phylogenies. *Molecular Biology and Evolution* 16:37–48.
441 https://doi.org/10.1093/oxfordjournals.molbev.a026036
442 Bax N, Williamson A, Aguero M, Gonzalez E, and Geeves W. 2003. Marine invasive alien species:
443 a threat to global biodiversity. *Marine Policy* 27:313-323.
444 https://doi.org/10.1016/s0308-597x(03)00041-1
445 Bernardi G, Fernandez-Delgado C, Gomez-Chiarri M, and Powers DA. 1995. Origin of a Spanish
446 population of *Fundulus heteroclitus* inferred by cytochrome b sequence analysis. *Journal
447 of Fish Biology* 47:737-740. https://doi.org/10.1111/j.1095-8649.1995.tb01939.x

448   Bernardi G, Sordino P, and Powers DA. 1993. Concordant mitochondrial and nuclear DNA
449   phylogenies for populations of the teleost fish *Fundulus heteroclitus*. *Proceedings of the*
450   *National Academy of Sciences* 90:9271-9274. https://doi.org/10.1073/pnas.90.20.9271
451   Bigelow HB, and Schroeder WC. 1953. Fishes of the Gulf of Maine. *Available at*
452   *https://archive.org/stream/fishesofgulfofma1953bige/fishesofgulfofma1953bige_djvu.t*
453   *xt* (accessed 17 June 2017).
454   Bock DG, Caseys C, Cousens RD, Hahn MA, Heredia SM, Hübner S, Turner KG, Whitney KD, and
455   Rieseberg LH. 2015. What we still don't know about invasion genetics. *Molecular*
456   *Ecology* 24:2277-2297. https://doi.org/10.1002/9781119072799.ch20
457   Catry T, Ramos JA, Paiva VH, Martins J, Almeida A, Palma J, Andrade PJ, Peste F, Trigo S, and
458   LuÍs A. 2006. Intercolony and annual differences in the diet and feeding ecology of little
459   tern adults and chicks in Portugal. *The Condor* 108:366-376.
460   https://doi.org/10.1650/0010-5422(2006)108[366:IAADIT]2.0.CO;2
461   Coelho M, Gomes J, and Re PB. 1976. *Valencia hispanica*, a new fish to Portugal. *Archivos do*
462   *Museu Bocage (2ª Série) Notas e suplementos*:5.
463   Cristescu ME. 2015. Genetic reconstructions of invasion history. *Molecular Ecology* 24:2212-
464   2225. https://doi.org/10.1111/mec.13117
465   Cristescu MEA, Hebert PDN, Witt JDS, MacIsaac HJ, and Grigorovich IA. 2001. An invasion
466   history for *Cercopagis pengoi* based on mitochondrial gene sequences. *Limnology and*
467   *Oceanography* 46:224-229. https://doi.org/10.4319/lo.2001.46.2.0224
468   Cristina S, Sequeira A, Newton A, JG F, and Icely J. 2016. ECASA Study Site Report Ria Formosa
469   coastal lagoon, Portugal. p 48.
470   Dias JP, Rocha R, Godwin S, Tovar-Hernández MA, Delahoz MV, McKirdy S, de Lestang P,
471   McDonald JI, and Snow M. 2016. Investigating the cryptogenic status of the sea squirt
472   *Didemnum perlucidum* (Tunicata, Ascidiacea) in Australia based on a molecular study of
473   its global distribution. *Aquatic Invasions* 11:239-245.
474   https://doi.org/10.3391/ai.2016.11.3.02
475   Dlugosch KM, and Parker IM. 2008. Founding events in species invasions: genetic variation,
476   adaptive evolution, and the role of multiple introductions. *Molecular Ecology* 17:431-
477   449. https://doi.org/10.1111/j.1365-294x.2007.03538.x
478   Doadrio I, Carmona JA, and Fernandez-Delgado C. 2002. Morphometric study of the Iberian
479   *Aphanius* (Actinopterygii, Cyprinodontiformes), with description of a new species. *Folia*
480   *Zoologica* 51:67-79.
481   Duggan IC, Rixon CAM, and MacIsaac HJ. 2006. Popularity and propagule pressure:
482   determinants of introduction and establishment of aquarium fish. *Biological Invasions*
483   8:377-382. https://doi.org/10.1007/s10530-004-2310-2
484   Duvernell DD, Lindmeier JB, Faust KE, and Whitehead A. 2008. Relative influences of historical
485   and contemporary forces shaping the distribution of genetic variation in the Atlantic
486   killifish, *Fundulus heteroclitus*. *Molecular Ecology* 17:1344-1360.
487   https://doi.org/10.1111/j.1365-294x.2007.03648.x
488   Elvira B, and Almodóvar A. 2001. Freshwater fish introductions in Spain: facts and figures at the
489   beginning of the 21st century. *Journal of Fish Biology* 59:323-331.
490   https://doi.org/10.1111/j.1095-8649.2001.tb01393.x

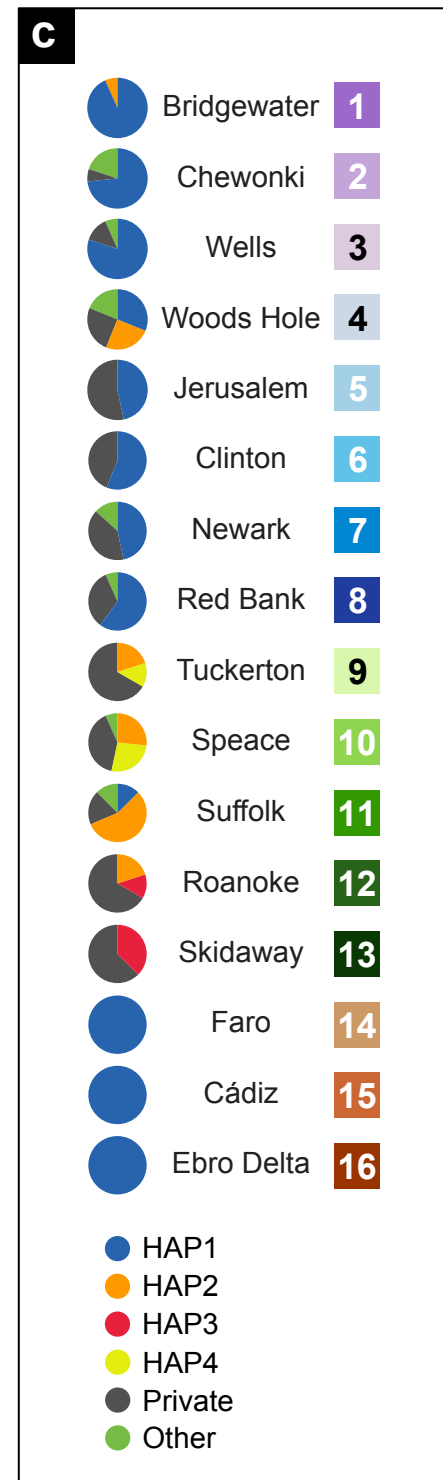
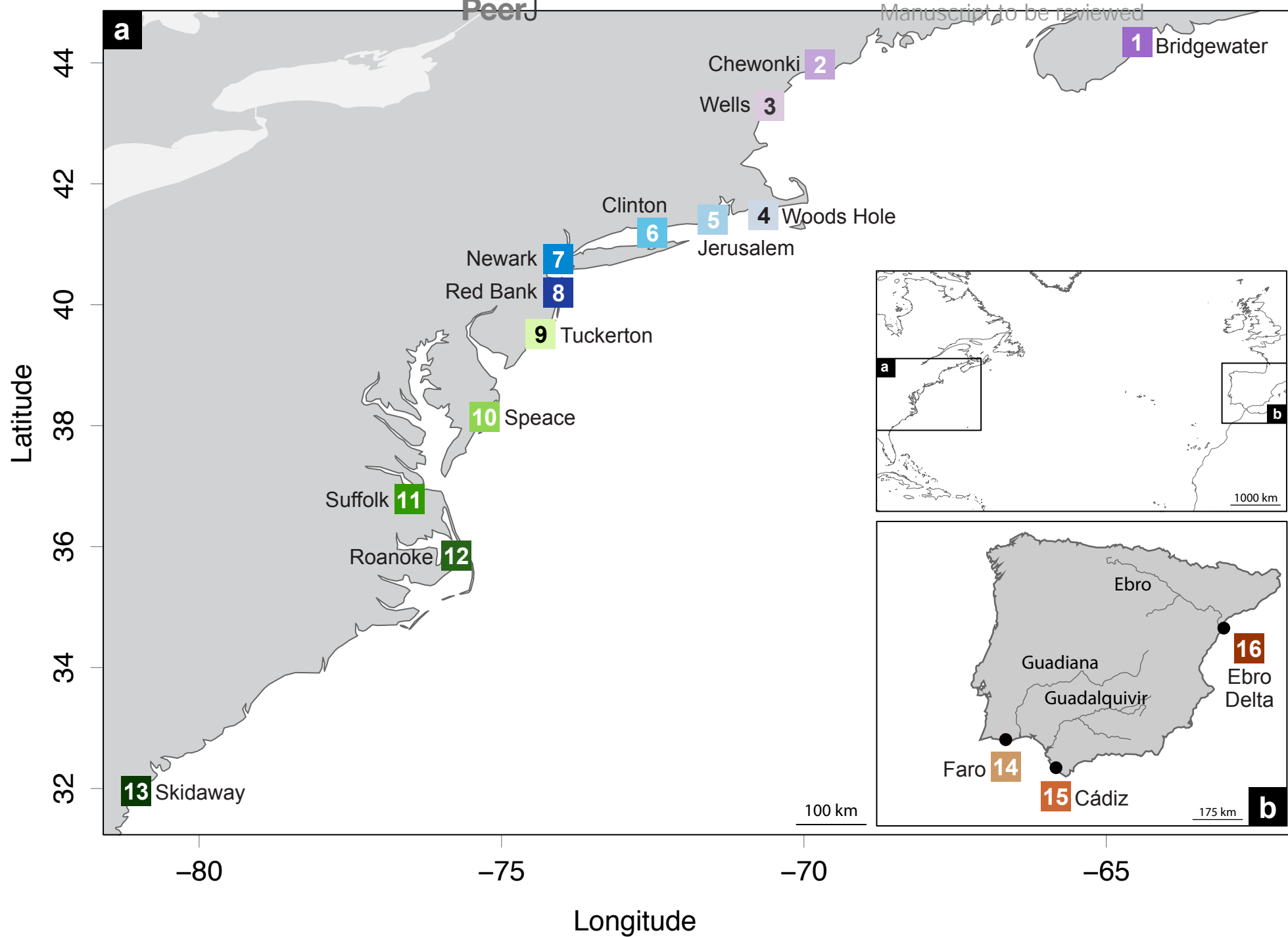
491   Estoup A, and Guillemaud T. 2010. Reconstructing routes of invasion using genetic data: why,
492   how and so what? *Molecular Ecology* 19:4113-4130. https://doi.org/10.1111/j.1365-
493   294x.2010.04773.x
494   Fernández-Delgado C. 1989. Life-history patterns of the salt-marsh killifish *Fundulus heteroclitus*
495   (L.) introduced in the estuary of the Guadalquivir River (South West Spain). *Estuarine,*
496   *Coastal and Shelf Science* 29:573-582. https://doi.org/10.1016/0272-7714(89)90011-5
497   Fernández-Delgado C. 2010. Especies introducidas en Andalucía: El caso de los peces. *Especies*
498   *exóticas invasoras en Andalucía Talleres provinciales 2004-2006*: Consejería de Medio
499   Ambiente. Junta de Andalucía, 132-137.
500   Fernández-Pedrosa V, Latorre A, and González A. 1996. Evidence from mtDNA RFLP analysis for
501   the introduction of *Fundulus heteroclitus* to southwestern Spain. *Journal of Fish Biology*
502   48:1278-1282. https://doi.org/10.1111/j.1095-8649.1996.tb01822.x
503   França S, Costa MJ, and Cabral HN. 2009. Assessing habitat specific fish assemblages in
504   estuaries along the Portuguese coast. *Estuarine, Coastal and Shelf Science* 83:1-12.
505   https://doi.org/10.1016/j.ecss.2009.03.013
506   Fritz ES, Meredith WH, and Lotrich VA. 1975. Fall and winter movements and activity level of
507   the mummichog, *Fundulus heteroclitus*, in a tidal creek. *Chesapeake Science* 16:211-215.
508   https://doi.org/10.2307/1350898
509   Galil BS, Marchini A, Occhipinti-Ambrogi A, Minchin D, Narščius A, Ojaveer H, and Olenin S.
510   2014. International arrivals: widespread bioinvasions in European Seas. *Ethology*
511   *Ecology & Evolution* 26:152-171. https://doi.org/10.1080/03949370.2014.897651
512   García-Berthou E, and Moreno-Amich R. 1991. New records of *Aphanius iberus* (Pisces:
513   Cyprinodontidae) and review of the geographical distribution of cyprinodontiform fishes
514   in Catalonia (NE-Spain). *Scientia gerundensis* 17:69-76.
515   García-Revillo MG, and Fernández-Delgado C. 2009. *La introducción por mar de especies*
516   *exóticas invasoras a través del agua de lastre de los barcos: el caso de Doñana*. Córdoba:
517   UCOPress, Editorial Universidad de Córdoba.
518   Geller JB, Darling JA, and Carlton JT. 2010. Genetic perspectives on marine biological invasions.
519   *Annual Review of Marine Science* 2:367-393.
520   https://doi.org/10.1146/annurev.marine.010908.163745
521   Gisbert E, and López MA. 2007. First record of a population of the exotic mummichog *Fundulus*
522   *heteroclitus* (L., 1766) in the Mediterranean Sea basin (Ebro River delta). *Journal of Fish*
523   *Biology* 71:1220-1224. https://doi.org/10.1111/j.1095-8649.2007.01579.x
524   Gollasch S. 2007. Is ballast water a major dispersal mechanism for marine organisms? In:
525   Nentwig W, ed. *Biological Invasions*: Springer-Verlag Berlin Heidelberg, 49-57.
526   Gomes C, Sousa R, Mendes T, Borges R, Vilares P, Vasconcelos V, Guilhermino L, and Antunes A.
527   2016. Low genetic diversity and high invasion success of *Corbicula fluminea* (Bivalvia,
528   Corbiculidae)(Müller, 1774) in Portugal. *PLOS ONE* 11:e0158108.
529   https://doi.org/10.1371/journal.pone.0158108
530   Gonçalves R, Teodósio M, Cruz J, Ben-Hamadou R, Correia A, and Chícharo L. 2017. Preliminary
531   insight into winter native fish assemblages in Guadiana estuary salt marshes coping with
532   environmental variability and non-indigenous fish introduction. *Fishes* 2.
533   https://doi.org/10.3390/fishes2040019

534   González-Vilaseñor LI, and Powers DA. 1990. Mitochondrial-DNA restriction-site polymorphisms
535   in the teleost *Fundulus heteroclitus* support secondary intergradation. *Evolution* 44:27-
536   37. https://doi.org/10.1111/j.1558-5646.1990.tb04277.x
537   Gozlan RE. 2010. The cost of non-native aquatic species introductions in Spain: fact or fiction?
538   *Aquatic Invasions* 5:231-238. https://doi.org/10.3391/ai.2010.5.3.02
539   Gutiérrez-Estrada JC, Prenda J, Oliva F, and Fernández-Delgado C. 1998. Distribution and habitat
540   preferences of the introduced mummichog *Fundulus heteroclitus* (Linnaeus) in South-
541   western Spain. *Estuarine, Coastal and Shelf Science* 46:827-835.
542   https://doi.org/10.1006/ecss.1997.0318
543   Hardy Jr JD. 1978. *Development of fishes of the mid-Atlantic Bight: an atlas of egg, larval, and*
544   *juvenile stages. Volume II. Anguillidae through Syngnathidae.*: United States Fish and
545   Wildlife Service. FWS/OBS-78/12.
546   Hernando J. 1975. Nuevas localidades de *Valencia hispanica* (Pisces: Cyprinodontidae) en el
547   suroeste de España. *Doñana Acta Vertebrata* 2:265-267.
548   Hulme PE. 2009. Trade, transport and trouble: managing invasive species pathways in an era of
549   globalization. *Journal of Applied Ecology* 46:10-18. https://doi.org/10.1111/j.1365-
550   2664.2008.01600.x
551   Kaluza P, Kölzsch A, Gastner MT, and Blasius B. 2010. The complex network of global cargo ship
552   movements. *Journal of the Royal Society Interface* 7:1093-1103.
553   https://doi.org/10.1098/rsif.2009.0495
554   Kaschner K, Watson R, Trites A, and Pauly D. 2006. Mapping world-wide distributions of marine
555   mammal species using a relative environmental suitability (RES) model. *Marine Ecology*
556   *Progress Series* 316:285-310. 10.3354/meps316285
557   Kesner-Reyes K, Kaschner K, Kullander S, Garilao C, Baril J, and Froese R. 2012. AquaMaps:
558   algorithm and data sources for aquatic organisms. *Available at www.fishbase.org*
559   (accessed March 2018).
560   Kneib RT. 1984. Patterns in the utilization of the intertidal salt marsh by larvae and juveniles of
561   *Fundulus heteroclitus* (Linnaeus) and *Fundulus luciae* (Baird). *Journal of Experimental*
562   *Marine Biology and Ecology* 83:41-51. https://doi.org/10.1016/0022-0981(84)90116-3
563   Lawson Handley L-J, Estoup A, Evans DM, Thomas CE, Lombaert E, Facon B, Aebi A, and Roy HE.
564   2011. Ecological genetics of invasive alien species. *BioControl* 56:409-428.
565   https://doi.org/10.1007/s10526-011-9386-2
566   Librado P, and Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA
567   polymorphism data. *Bioinformatics* 25:1451–1452.
568   https://doi.org/10.1093/Bioinformatics/Btp187
569   Long D. 2006. *BGS detailed explanation of seabed sediment modified folk classification*.
570   Keyworth, UK.
571   Lotrich VA. 1975. Summer home range and movements of *Fundulus heteroclitus* (Pisces:
572   Cyprinodontidae) in a tidal creek. *Ecology* 56:191-198. https://doi.org/10.2307/1935311
573   Maceda-Veiga A, Escribano-Alacid J, de Sostoa A, and García-Berthou E. 2013. The aquarium
574   trade as a potential source of fish introductions in southwestern Europe. *Biological*
575   *Invasions* 15:2707. https://doi.org/10.1007/s10530-013-0485-0

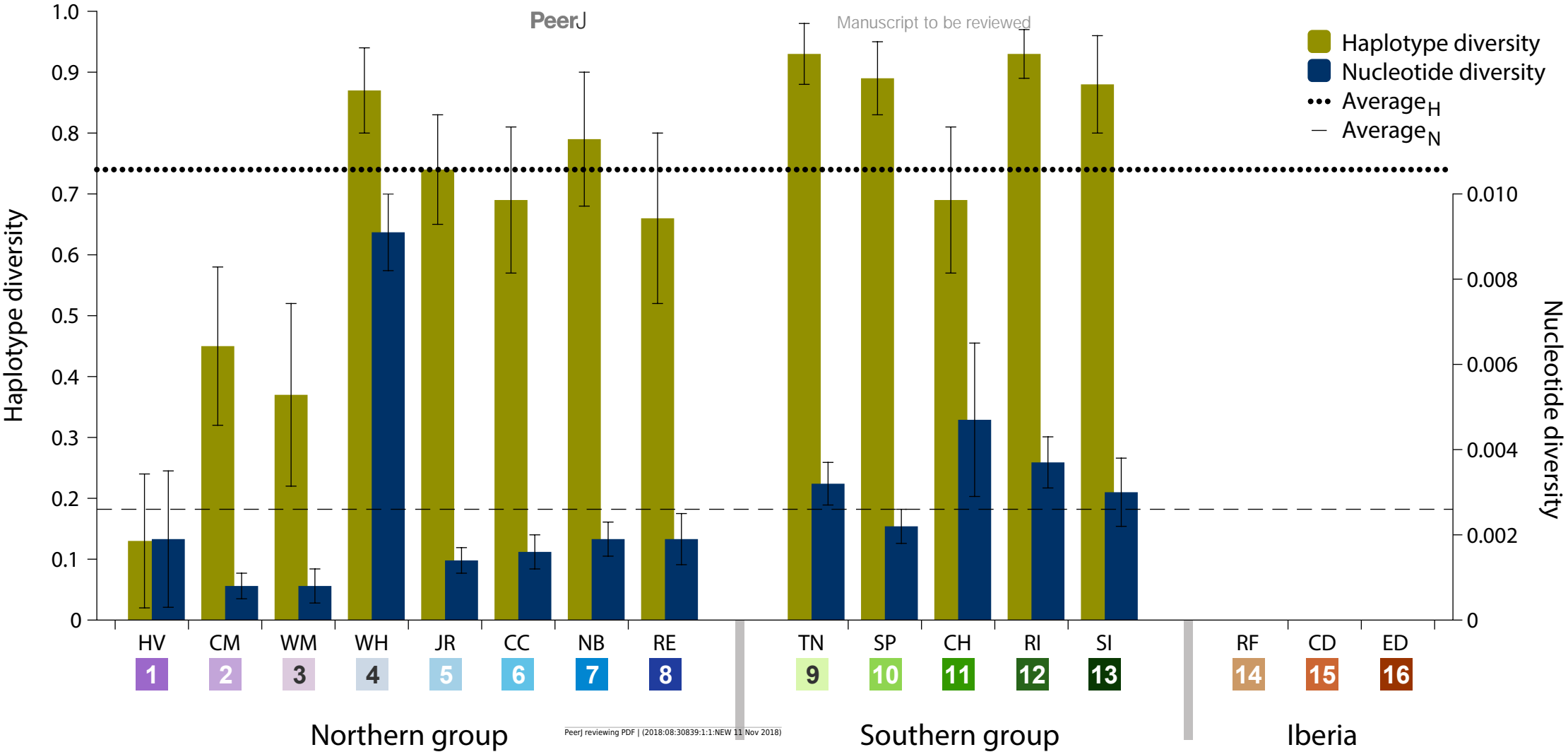
576   Martin AP, and Bermingham E. 1998. Systematics and evolution of lower Central American
577   cichlids inferred from analysis of cytochrome b gene sequences. *Molecular*
578   *Phylogenetics and Evolution* 9:192-203. https://doi.org/10.1006/mpev.1997.0461
579   McKenzie JL, Dhillon RS, and Schulte PM. 2016. Steep, coincident, and concordant clines in
580   mitochondrial and nuclear-encoded genes in a hybrid zone between subspecies of
581   Atlantic killifish, *Fundulus heroclitus*. *Ecology and Evolution* 6:5771-5787.
582   https://doi.org/10.1002/ece3.2324
583   Millennium Ecosystem Assessment. 2005. Ecosystems and Human Well-being: Biodiversity
584   Synthesis. *Available at http://millenniumassessment.org/en/Synthesis.html* (accessed 2
585   August 2017).
586   Molnar JL, Gamboa RL, Revenga C, and Spalding MD. 2008. Assessing the global threat of
587   invasive species to marine biodiversity. *Frontiers in Ecology and the Environment* 6:485-
588   492. https://doi.org/10.1890/070064
589   Morin PA, Luikart G, Wayne RK, and the SNP workshop group. 2004. SNPs in ecology, evolution
590   and conservation. *Trends in Ecology & Evolution* 19:208-216.
591   https://doi.org/10.1016/j.tree.2004.01.009
592   Morin RP, and Able KW. 1983. Patterns of geographic variation in the egg morphology of the
593   fundulid fish, *Fundulus heteroclitus*. *Copeia* 1983:726-740.
594   https://doi.org/10.2307/1444339
595   Muirhead JR, Gray DK, Kelly DW, Ellis SM, Heath DD, and Macisaac HJ. 2008. Identifying the
596   source of species invasions: sampling intensity vs. genetic diversity. *Molecular Ecology*
597   17:1020-1035. https://doi.org/10.1111/j.1365-294x.2008.03669.x
598   Nei M. 1987. Genetic distance and molecular phylogeny. In: Ryman N, and Utter FW, eds.
599   *Population Genetics & Fishery Management*. Seattle: Washington Sea Grant Program,
600   University of Washington, 193–223.
601   Nei M, and Tajima F. 1981. DNA polymorphism detectable by restriction endonucleases.
602   *Genetics* 97:145–163.
603   Padilla DK, and Williams SL. 2004. Beyond ballast water: aquarium and ornamental trades as
604   sources of invasive species in aquatic ecosystems. *Frontiers in Ecology and the*
605   *Environment* 2:131-138. https://doi.org/10.1890/1540-
606   9295(2004)002[0131:bbwaao]2.0.co;2
607   Page L, and Burr B. 2011. *A field guide to freshwater fishes of North America north of Mexico*.
608   Boston: Houghton Mifflin Harcourt.
609   Palumbi S, Martin A, Romano S, McMillan W, Stice L, and Grabowski G. 1991. *The Simple Fool's*
610   *Guide to PCR, Version 2.0*. Honolulu, HI, 96822: University of Hawaii Press.
611   Pérez-Ruzafa A, Garcıa-Charton J, Barcala E, and Marcos C. 2006. Changes in benthic fish
612   assemblages as a consequence of coastal works in a coastal lagoon: The Mar Menor
613   (Spain, Western Mediterranean). *Marine Pollution Bulletin* 53:107-120.
614   https://doi.org/10.1016/j.marpolbul.2005.09.014
615   Phillips SJ, and Dudík M. 2008. Modeling of species distributions with Maxent: new extensions
616   and a comprehensive evaluation. *Ecography* 31:161–175.
617   Powers DA, and Place AR. 1978. Biochemical genetics of *Fundulus heteroclitus* (L.). I. Temporal
618   and spatial variation in gene frequencies of *Ldh-B*, *Mdh-A*, *Gpi-B*, and *Pgm-A*.
619   *Biochemical Genetics* 16:593-607. https://doi.org/10.1007/bf00484222

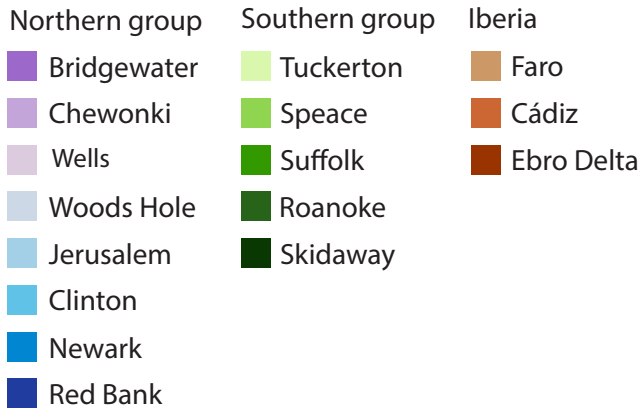
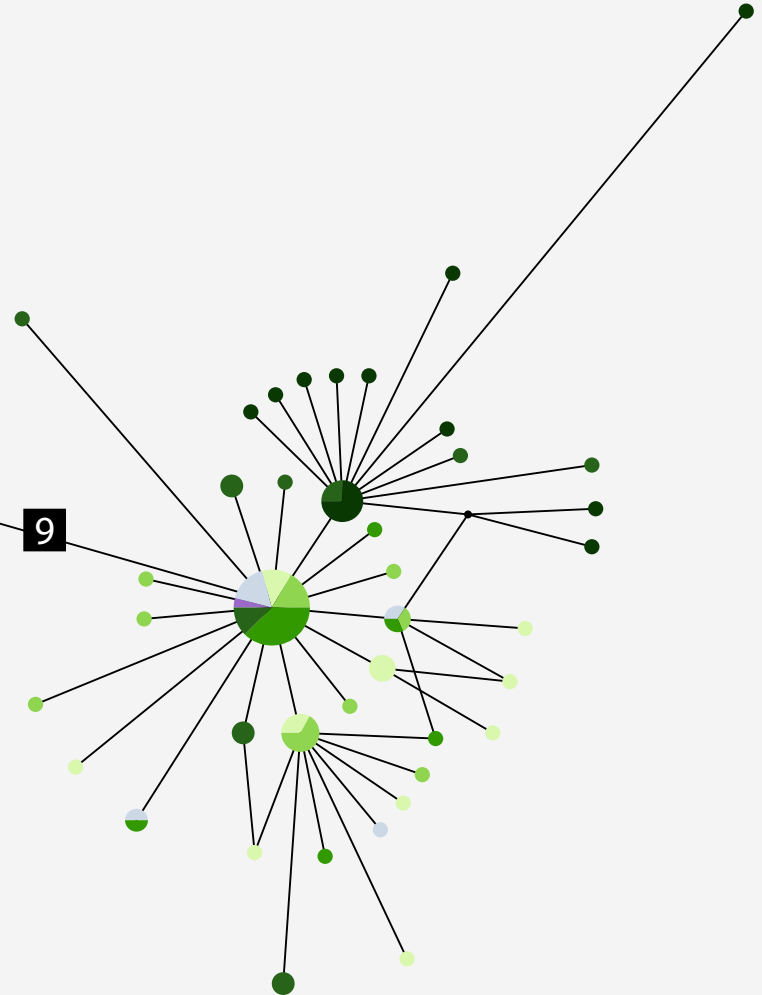
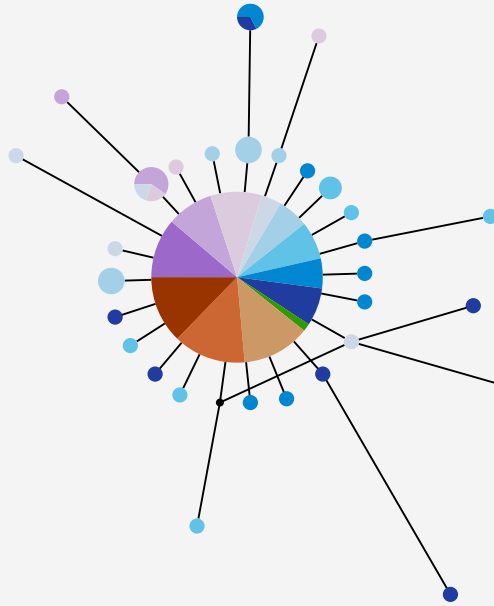
620   Powers DA, Ropson I, Brown DC, Van Beneden R, Cashon R, González-Villaseñor LI, and
621   DiMichele JA. 1986. Genetic variation in *Fundulus heroclitus*: geographic distribution.
622   *American Zoologist* 26:131-144. https://doi.org/10.1093/icb/26.1.131
623   Rees T. 2003. "C-Squares", a new spatial indexing system and its applicability to the description
624   of oceanographic datasets. *Oceanography* 16:11-19.
625   Ribeiro J, Bentes L, Coelho R, Gonçalves JM, Lino PG, Monteiro P, and Erzini K. 2006. Seasonal,
626   tidal and diurnal changes in fish assemblages in the Ria Formosa lagoon (Portugal).
627   *Estuarine, Coastal and Shelf Science* 67:461-474.
628   https://doi.org/10.1016/j.ecss.2005.11.036
629   Ribeiro J, Monteiro CC, Monteiro P, Bentes L, Coelho R, Gonçalves JMS, Lino PG, and Erzini K.
630   2008. Long-term changes in fish communities of the Ria Formosa coastal lagoon
631   (southern Portugal) based on two studies made 20 years apart. *Estuarine, Coastal and*
632   *Shelf Science* 76:57-68. https://doi.org/10.1016/j.ecss.2007.06.001
633   Rius M, Turon X, Bernardi G, Volckaert FA, and Viard F. 2014. Marine invasion genetics: from
634   spatio-temporal patterns to evolutionary outcomes. *Biological Invasions* 17:869-885.
635   https://doi.org/10.1007/s10530-014-0792-0
636   Roman J, and Darling JA. 2007. Paradox lost: genetic diversity and the success of aquatic
637   invasions. *Trends in Ecology & Evolution* 22:454-464.
638   https://doi.org/10.1016/j.tree.2007.07.002
639   Ropson IJ, Brown DC, and Powers DA. 1990. Biochemical genetics of *Fundulus heteroclitus* (L.).
640   VI. Geographical variation in the gene frequencies of 15 loci. *Evolution* 44:16-26.
641   https://doi.org/10.1111/j.1558-5646.1990.tb04276.x
642   Sakinan S, Karahan A, and Ok M. 2017. Integration of DNA barcoding for the initial recordings of
643   Lessepsian fishes: a case study of the Indo-Pacific slender ponyfish *Equulites elongatus*.
644   *Journal of Fish Biology* 90:1054-1061. https://doi.org/10.1111/jfb.13207
645   Sambrook J, and Russel DW. 2001. *Molecular cloning: a laboratory manual*. Cold Spring Harbor,
646   New York: Cold Spring Harbor Laboratory Press.
647   Smith MW, Chapman RW, and Powers DA. 1998. Mitochondrial DNA analysis of Atlantic coast,
648   Chesapeake Bay, and Delaware Bay populations of the teleost *Fundulus heteroclitus*
649   indicates temporally unstable distributions over geologic time. *Molecular Marine*
650   *Biology and Biotechnology* 7:79-87.
651   Walsh PS, Metzger DA, and Higuchi R. 1991. Chelex 100 as a medium for simple extraction of
652   DNA for PCR-based typing from forensic material. *BioTechniques* 10:506-513.
653   https://doi.org/10.2144/000114018
654   Williams SL, Davidson IC, Pasari JR, Ashton GV, Carlton JT, Crafton RE, Fontana RE,
655   Grosholz ED, Miller AW, and Ruiz GM. 2013. Managing multiple vectors for marine
656   invasions in an increasingly connected world. *BioScience* 63:952-966.
657   https://doi.org/10.1525/bio.2013.63.12.8
658   Wilson JRU, Dormontt EE, Prentis PJ, Lowe AJ, and Richardson DM. 2009. Something in the
659   way you move: dispersal pathways affect invasion success. *Trends in Ecology &*
660   *Evolution* 24:136-144. https://doi.org/10.1016/j.tree.2008.10.007
661
662

Distribution map of localities and haplotypes

Distribution map of localities along the coast of (a) North America, and (b) Iberian Peninsula of *Fundulus heroclitus*. Locations and sample details can be found in Table 1. (c) Coloured circles display the distribution of the relative proportions of the cytochrome*b*haplotypes from each location: the four most common haplotypes, the private haplotypes, and all the other shared, but less frequent haplotypes.

Haplotype and nucleotide diversity

Haplotype and nucleotide diversity of *Fundulus heteroclitus*from all sampled locations.

Abbreviations in the legend are defined in Table 1, and colours are the same as in Fig. 1.

Median-joining cytochrome *b* haplotype network for *Fundulus heteroclitus*

Median-joining cytochrome*b*haplotype network for *Fundulus heteroclitus*. Each circle represents a different haplotype with size proportional to the frequency of the haplotype within the sample. Line length is proportional to the number of mutations between haplotypes. Each colour corresponds to a different location. Colours are the same as in Figure 1. Locations details can be found in Table 1. The two black dots represent putative unsampled haplotypes, and the number in black square shows the number of mutations separating haplogroup A from the haplogroup B.

Haplogroup A

Haplogroup B

Northern group

Bridgewater
Chewonki
Wells
Woods Hole
Jerusalem
Clinton
Newark
Red Bank

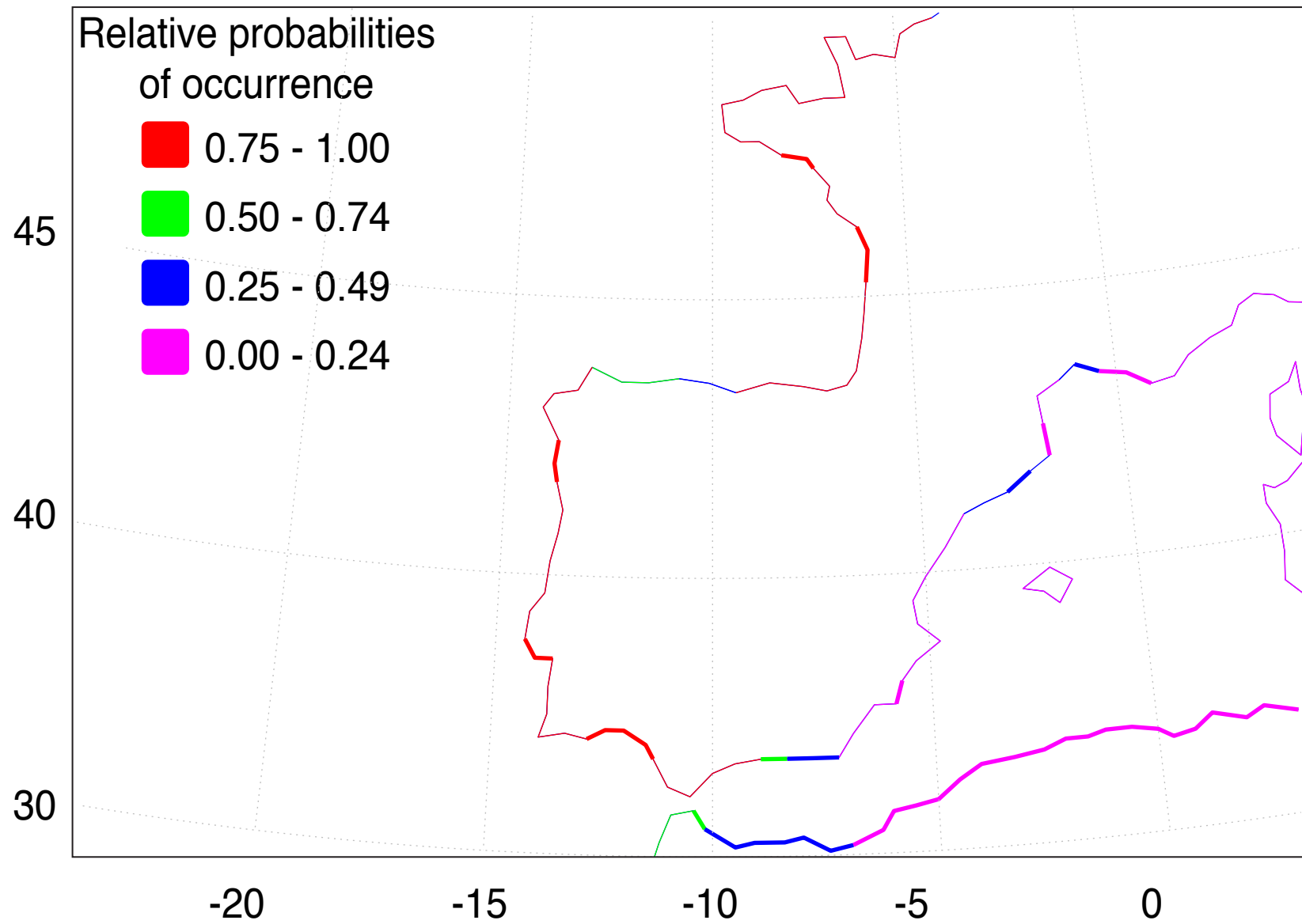
Southern group

Tuckerton
Speace
Suffolk
Roanoke
Skidaway

Iberia

Faro
Cádiz
Ebro Delta

Western European and Mediterranean coastal environments showing AquaMap probabilities of occurrence of *F. heteroclitus*.

Western European and Mediterranean coastal environments showing AquaMap probabilities of occurrence of *F. heteroclitus*. Those areas in bold show coastal seabed habitats with a mud content > 10%.

Sample details

Sample location, sample abbreviations and summary statistics for a cytochrome *b*sequence fragment from *Fundulus heroclitus*. ID refers to numbers in Fig 1

1 **Table 1** Sample location, sample abbreviations and summary statistics for a cytochrome *b*
2 sequence fragment from *Fundulus heteroclitus*. ID refers to numbers in Fig 1

Location	ID	Code	Latitude / longitude	n	n_h	n_p
Bridgewater	1	HV	44°22.0'N / 64°31.0'W	15	2	0
Chewonki	2	CM	43°57.3'N / 69°43.2'W	15	3	1
Wells	3	WM	43°19.2'N / 70°34.2'W	15	4	2
Woods Hole	4	WH	41°31.5'N / 70°40.4'W	16	9	4
Jerusalem	5	JR	41°23.1'N / 71°31.5'W	15	5	4
Clinton	6	CC	41°15.3'N / 72°32.8'W	16	7	6
Newark Bay	7	NB	40°41.2'N / 74°06.7'W	15	8	6
Red Bank	8	RE	40°20.9'N / 74°05.0'W	15	7	5
Tuckerton	9	TN	39°32.2'N / 74°19.4'W	15	10	8
Speace	10	SP	38°09.1'N / 75°17.2'W	15	9	6
Suffolk	11	CH	36°51.8'N / 76°28.7'W	16	7	3
Roanoke Island	12	RI	35°53.8'N / 75°36.9'W	15	9	7
Skidaway Island	13	SI	31°56.8'N / 81°04.2'W	16	11	1
Faro	14	RF	37°00.3'N / 07°58.0'W	16	1	0
Cádiz	15	CD	36°31.4'N / 06°11.4'W	17	1	0
Ebro Delta	16	ED	40°37.38'N / 0°39.44'E	16	1	0

4 n, number of individuals; n_h, number of haplotypes; n_p, number of private haplotypes
5

Aquamaps environmental envelope

Aquamaps default environmental envelope for *F. heteroclitus*

1 **Table 2** Aquamaps default environmental envelope for *F. heteroclitus*

Variable	Absolute	Preferred	Preferred	Absolute
Depth [m]	0	0	3	5
Sea Surface Temperature	4.76	6.41	23.53	26.1
Sea Surface Salinity [SSS;	28.96	29.94	35.74	39.6