1   **Effects of water pollution and river fragmentation on population genetic structure**

2   **of invasive mosquitofish**

3

4

5   David Díez-del-Molino[a,b], Emili García-Berthou[c], Rosa-Maria Araguas[a], Carles

6   Alcaraz[d], Oriol Vidal[a], Nuria Sanz[a], Jose-Luis García-Marín[a,*]

7

8

9   [a] *Laboratori d'Ictiologia Genètica (LIG), University of Girona. Edifici AC-LEAR,*

10   *Carrer M. Aurèlia Capmany, 40, 17003 Girona, Spain*

11   [b] *Dept. of Bioinformatics and Genetics, Swedish Museum of Natural History, Box*

12   *50007, 104 05 Stockholm, Sweden*

13   [c] *GRECO, Institute of Aquatic Ecology, University of Girona. 17003 Girona, Spain*

14   [d] *IRTA Marine and Continental Waters, Carretera Poble Nou, km 5.5, 43540 Sant*

15   *Carles de la Ràpita, Spain*

16

17   *Corresponding author

18   telephone: +34 616603335

19   Fax: +34 972418150

20   *E-mail addresses:*diez.molino@gmail.com (D. Díez-del-Molino),

21   emili.garcia@udg.edu (E. García-Berthou), rosa.araguas@udg.edu (R.-M. Araguas),

22   carles.alcaraz@gmail.com (C. Alcaraz), oriol.vidal@udg.edu (O. Vidal),

23   nuria.sanz@udg.edu (N. Sanz), joseluis.garcia@udg.edu (J.-L. García-Marín)

24 **Highlights**

25 We examined genetic variation of *Gambusia holbrook*i in a polluted river reservoir

26 Mosquitofish at the most polluted sites showed genetic distinction from other studied
27 sites along the river basin

28 Changes at the GPI-2 locus agree with predicted response to acute exposure to mercury

29 Microsatellite loci also distinguished fish at the vicinity of most polluted sites

30 Immigration, even from sites beneath the dam, mitigates genetic losses at polluted sites

31   **Graphical Abstract**
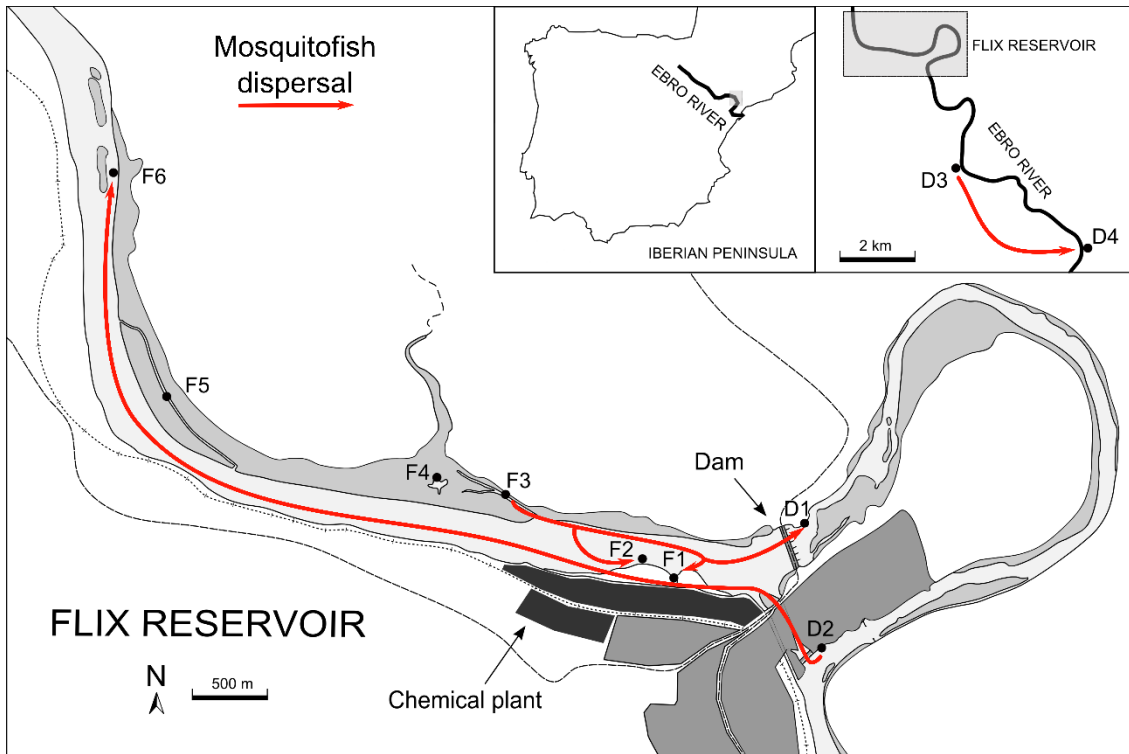
32

33 **Abstract**

34 We analyzed variation at the *GPI-2* locus and eleven microsatellite loci of eastern

35 mosquitofish *Gambusia holbrooki* populations introduced to the Ebro River (Spain),

36 sampling above and below a dam (Flix Reservoir) where severe chronic pollution has

37 been well documented. Allele frequency changes at the *GPI-2* locus in the sites nearest

38 to the polluted sediments agree with previous results from studies in mercury-exposed

39 populations of this highly invasive fish. Genetic distinction of the mosquitofish

40 collected close to the polluted sediments was detected at the GPI locus but also at the

41 presumptive neutral microsatellite loci. Recent migration rates estimated from

42 microsatellites indicated that around 30 percent of fish collected in a specific location

43 were immigrants from upstream and downstream sources. Such high migration rates

44 probably contribute to the mosquitofish's invasive success and suggest that the

45 consequences on the mosquitofish regional genetic structured of high levels of water

46 toxicants could be mediated by immigration from other sites, but the effect of

47 pollutants on local diversity might be higher than observed here.

48

49

50 **Keywords:** Water pollution; genetic population structure; river fragmentation;

51 *Gambusia holbrooki*; invasive species.

52

53

1. Introduction

Genetic diversity patterns have been advocated as crucial in determining the invasion success of non-native species. Identifying the spatial scales of genetic divergence among invaded locations is fundamental to understand biological processes involved in successful invasions and hence to define management actions to control invasive populations (Blanchet, 2010; Palsboll et al., 2007; Purcell and Stockwell, 2015). Freshwater ecosystems are among the most affected by invasions, in part because of anthropogenic actions leading to habitat alterations (McLellan et al., 2014). In freshwater systems, genetic population structure within species often results from the combination of spatial dispersal (a linear isolation by distance, IBD), fragmentation, and population size fluctuations. Natural barriers such as waterfalls and hydrological regimes showing periodical severe droughts or floods can modulate connectivity and gene flow between populations (Crispo et al., 2006; Díez-del-Molino et al., 2016). Moreover, anthropogenic perturbations such as dams, weirs, and pollution might also modify the amount and distribution of genetic variants between populations along streams (Horreo et al., 2013; Roberts et al., 2013), either by increasing isolation or by changing directional selective pressures (Bélanger-Deschênes et al., 2013; Vega-Retter et al., 2015). However, widespread freshwater invaders, such as common carp (*Cyprinus carpio)*, tilapias (*Oreochromis* spp.), or mosquitofishes (*Gambusia affinis* and *G. holbrooki)*, are species generally more tolerant to water pollution and habitat degradation than the native ones (García-Berthou et al., 2005), and hence might suffer less from these anthropogenic perturbations.

The eastern mosquitofish, *Gambusia holbrooki,* is a poeciliid fish native to the eastern U.S.A., which has been introduced worldwide in marshlands, lakes, and rivers,

and is considered as one of the worst invasive freshwater fishes, producing diverse

impacts on ecosystem structure and functioning (Alcaraz et al., 2008; Pyke, 2005).

Several factors contribute to the invasive success of mosquitofish, including broad

thermal and salinity tolerance (Stockwell and Weeks, 1999), population size recovery

in a few months (Chapman and Warburton, 2006; Deacon et al., 2011), large brood

sizes (sometimes over 100 newborn fish), multiple paternity resulting in an average of

four sires per female (Zeng et al., 2017), and high dispersal capabilities (Rehage and

Sing, 2004). Both native and invasive mosquitofish populations display high levels of

genetic diversity (Sanz et al., 2013), probably as a response of the huge reproductive

potential (Zane et al., 1999). In addition, gene flow induced from sporadic individual

exchanges among hydrographically close populations has been suggested to maintain

and even increase local diversity in both native (Smith et al., 1983) and invasive

populations (Ayres et al., 2010; Díez-del-Molino et al., 2013, 2016).

Along river longitudinal gradients, *G. holbrooki* often displays complex patterns of

population structure due to demographic fluctuations, breeding strategies, and sex and

cohort-specific dispersal abilities (Díez-del-Molino et al., 2016; Kennedy et al., 1986).

Active dispersal and downstream individual transport during floods often results in an

isolation by distance (IBD) pattern among mosquitofish populations at distances of 6-

150 km along river basins (Congdon, 1995; Díez-del-Molino et al., 2013; Hernandez-

Martich and Smith, 1997). Selection can also play an important role in the divergence

among neighboring mosquitofish populations. For example, selection resulting from

periods of saltwater inundation in coastal marshlands can promote allelic variation at

the glucosephosphate isomerase *GPI-2* locus (Congdon, 1994). Similarly, a significant

genetic divergence between mosquitofish (*G. affinis*) populations in a transect of 10 km

in the upper San Antonio River (Texas, USA) likely arose from the combined effects of

103 limited gene flow due to numerous dams and the selective response to low dissolved

104 oxygen during summers in highly altered river locations (Roark et al., 2001). In fact,

105 mosquitofish quickly adapt in response to changing biotic and abiotic factors such as

106 presence of predators or salinity and thermal gradients (Congdon, 1994; Langerhans et

107 al., 2007; Meffe et al., 1995; Purcell et al., 2012a).

108 Toxicants may affect the genetic diversity within populations by directly increasing

109 the mutation rate, forcing selection in favor of tolerant genotypes, causing local

110 bottlenecks, or altering migration patterns (Van Straalen and Timmermans, 2002). In

111 mosquitofish, chemical contaminants induced mutations in the mitochondrial DNA of

112 introduced populations in Azerbaijan (Rinner et al., 2010). The significant changes in

113 allele frequencies at the *GPI-2* locus when mosquitofish are subject to pollutants have

114 been extensively studied (Tatara et al., 1999, 2002). For example, rare homozygous

115 genotypes at the *GPI-2* locus displayed significantly shorter median time to death

116 (TTD) when mosquitofish were exposed to sublethal concentrations of mercury

117 (Diamond et al., 1989; Heagler et al.,1993). Mulvey et al. (1995) showed that

118 homozygous *GPI-2*$^{100/100}$ females were less likely to be gravid and had fewer

119 developed embryos under chronic mercury exposition.

120 Flix Reservoir, located 90 km upstream from the Ebro River delta (Spain), is an

121 invaluable opportunity to test the relationships between pollutants and genetic diversity

122 within and among wild mosquitofish populations. The first and main introduction of

123 mosquitofish in Europe was of 12 specimens in 1921 in a pond in SW Spain and soon

124 after its acclimatization, mosquitofish was spread into many rivers and lagoons, but in

125 1945 its presence in the Ebro River basin was limited to the river delta and lowest

126 reaches (reviewed in Navarro-Garcia, 2013).   Flix dam was built in 1941-1948

127   although a medieval weir was already present before. During the second half of the 20th

128   century, a chemical plant has deposited more than 20–36×10^4 tons of industrial waste

129   contaminated with heavy metals, organochlorine pesticides and radionuclides into the

130   Flix Reservoir (Bosch et al., 2009; Navarro et al., 2009). Floods have transported part

131   of these pollutants along the river course down to the delta (Alcaraz et al., 2011;

132   Navarro et al., 2009; Quirós et al., 2008; Suárez-Serrano et al., 2010, Quesada et al.

133   2014). Mercury is one of the most dangerous heavy metals deposited into the aquatic

134   environment because microbes can transform it into most toxic organic forms, such as

135   methylmercury that bioaccumulates throughout the aquatic food chain (Carrasco et al.,

136   2008, 2011). The suspended particulate matter in Flix Reservoir has very high

137   concentrations of mercury (7.24–$31.14 \, \mu g \, g^{-1}$ dw), largely exceeding the

138   concentrations at close upstream areas (0.07–$0.17 \, \mu g \, g^{-1}$ dw). Mercury is still abundant

139   (1.53–$1.88 \, \mu g \, g^{-1}$ dw) in downstream areas located as far as 21 km downstream and

140   has bioaccumulated in aquatic biota of Flix Reservoir and downstream (Alcaraz et al.,

141   2011; Cid et al., 2010; Soto et al., 2016; Suárez-Serrano et al., 2010, Palanques et al.

142   2015). As a consequence, several fish species showed physiological responses (Navarro

143   et al., 2009) and reduced reproductive fitness and body condition (Benejam et al., 2010)

144   in Flix Reservoir and also downstream.

145   Almost all previous studies have used neutral markers (e.g. microsatellites) to

146   measure genetic diversity within and among invasive mosquitofish populations (Ayres

147   et al., 2010; Díez-del-Molino et al. 2013, 2016; Purcell et al., 2012b; Sanz et al., 2013).

148   The comparison of genetic diversity patterns measured with neutral and non-neutral

149   markers, such as protein-coding loci, may allow to uncover the role of adaptive

150   responses to anthropogenic perturbations in invasive populations. Therefore, we

151   analyzed genetic variation at eleven microsatellite loci and the protein coding *GPI-2*

152 locus of invasive mosquitofish populations along the Ebro River, upstream and

153 downstream of the polluted sediments of Flix Reservoir. Our main aim was to test

154 whether the presence of a dam and of a vast amount of contaminants resulted in

155 reduced local diversity levels and altered the expected linear population structure along

156 the river.

157 **2. Material and methods**

158 2.1 Sampling area

159 Fish were sampled during October, 2007, in Flix Reservoir and downstream (Ebro

160 River), more or less at the beginning of an increased concern on biota and humans of

161 the effects of the pollutants deposited in this reservoir (e.g. Carrasco et al. 2008). The

162 Ebro River is the second largest river (in terms of drainage area and discharge) in the

163 Iberian Peninsula, with 928 km of length and 85,550 km² of drainage area. Available

164 published information on mercury was used as a proxy for pollutants abundance in the

165 study area (see Table 1). In short, total mercury showed a peak at the polluted

166 sediments of Flix Reservoir, but also abounds throughout the reservoir and in

167 downstream biota (Bosch et al., 2009). Sampling locations included two sites at the

168 riverbank beside the chemical plant and near to the toxic sediments (F1 and F2), a close

169 site in the opposite riverbank (F3), and three sites (F4, F5 and F6) upstream of the

170 polluted sediments area (Table 1, Fig. 1). We also collected a sample (D1) just beneath

171 the Flix dam, which is 26 m high. Flix is a small reservoir for electricity production

172 through a water diversion in the right shore. In addition to the channel toward the

173 power turbines, the diversions include a navigable channel with locks that have not

174 been opened in the last 30 years (see *Esquemas* tab at

175 http://195.55.247.237/saihebro/index.php?url=/datos/ficha/ estacion:E002), and a fish

176   ladder that has been proven to be ineffective for native fish species (CHE, 2009). We

177   collected a sample at the riverside in the downstream mouth of these channels, at the

178   end of Flix meander (D2). Two additional downstream samples were collected at Ascó

179   (D3), 13 km apart from Flix Reservoir, and at the riverbank of Garcia village (D4), 23

180   km apart from the reservoir. After capture with dip nets and with scientific permits of

181   the relevant authorities (Generalitat de Catalunya), fish were euthanized *in situ* by

182   lethal sedation with tricaine methanosulphonate (MS-222) and then preserved on ice

183   and kept frozen until tissue preparation for electrophoretic analysis and DNA extraction

184   in the laboratory. From each site, 40 *G. holbrooki* specimens were genotyped for the

185   molecular markers described below.

186   2.2 GPI polymorphisms

187   In 2008, the head of each fish was homogenized in an Eppendorf tube with 0.5 ml of

188   distilled water, while a small portion of muscular tissue in the caudal part of each

189   individual was used for DNA isolation by using the "Real Pure DNA extraction toolkit"

190   (Durvitz®) following the manufacturer's instructions. DNA extractions were stored at -

191   20 ℃ until use in Polymerase Chain Reactions (PCR). Starch gel electrophoresis

192   buffered with an amino-citrate (pH 7.0) was used following Araguas et al. (2007) in

193   order to resolve gene products of the glucosephosphate isomerase enzyme (GPI,

194   E.C.5.3.1.9) from head homogenates. The *GPI-2* locus was codominantly expressed

195   and permitted direct counts of alleles from gel phenotypes. In addition, to check for no

196   selective population divergences at the glucosephosphate isomerase gene of *G.*

197   *holbrooki*, we designed a routine primer extension analysis protocol to genotype a CxT

198   transition in position 85 of the no coding intron 17 of this gene, using the SNaPshot®

199   Multiplex kit (Applied Biosystems), with amplification primers GPIe17F (5'-

200　CAGTGGGGGTGAGTTTATTCTT-3') and GPIi17R (5'-

201　TTAAGGATGTCAGACAAAAACTTTAAT-3'), and extension primer GPIi17snp1

202　(5'-taatatataaaatattACTGTCAGTTGGTCTTAAATCTTAT-3'). PCR was performed at

203　50 °C in a final volume of 30 μl with MgCl₂, and 200 μM dNTPs, 0.2 μM of each

204　amplification primer and 5-15 ng of DNA. Genetic linkage between the intron variants

205　and the alleles observed at the *GPI-2* allozyme locus are unclear.

206　2.3 Microsatellites

207　Each specimen was genotyped for eleven microsatellite loci: Pooc-G49, M*f*13,

208　Gafμ3, Gafμ5, Gafμ6, Gafμ7, Gaaf7, Gaaf9, Gaaf10, Gaaf13 and Gaaf15, following the

209　procedures described in Díez-del-Molino et al. (2013). Forward primers were

210　fluorescently labeled and genotype peaks were resolved on a 3130 Genetic Analyzer

211　and using the GeneMapper 4.0 software (Applied Biosystems, Foster City, CA, USA).

212　2.4 Data analyses

213　Genetic diversity within locations was expressed as average observed (H_O), expected

214　heterozygosity (H_E), and allele richness (Ar), either for the microsatellite loci and the

215　*GPI-2* polymorphisms (allozyme and SNP). The exact probability test of Guo and

216　Thompson (1992) implemented in GENEPOP 4 software (Rousset, 2008) was applied

217　to check for conformance of genotype frequencies with Hardy-Weinberg expectations

218　(HWE) in all analyzed loci. MICRO-CHECKER (van Oosterhout et al., 2004) was used

219　to potentially identify null alleles as responsible for HWE deviations. Null allele

220　frequencies were estimated according to the estimator based on the Expectation-

221　Maximization algorithm of Dempster et al. (1997) using FreeNA (Chapuis and Estoup,

222　2007). Composite linkage disequilibrium between pairs of loci was assessed from

223　genotypic contingency tables using GENEPOP. The levels of significance for multiple

224   comparisons were adjusted according to Bonferroni's correction. Recent population

225   bottlenecks were tested using BOTTLENECK 1.2.02 (Piry et al., 1999). Effective

226   population sizes (Ne) at each location were estimated using the linkage disequilibrium

227   approach of the LDNe program included in the NeEstimator V2 software (Do et al.,

228   2014).

229   Differentiation among collections and their significance were estimated from F_{ST}

230   (Weir and Cockerham, 1984), with and without correcting for null alleles as

231   implemented in FreeNA. In the latter approach, correction for null alleles involved the

232   ENA method described by Chapuis and Estoup (2007). To check for a pattern of

233   divergence among locations resulting from isolation by distance, we used a non-

234   parametric Mantel test between genetic distances ($F_{ST}/(1-F_{ST})$) and hydrographical

235   distances (km) between locations using GENEPOP 4 and 10^6 permutations.

236   Hierarchical analyses of molecular variance (AMOVA) were carried out as

237   implemented in ARLEQUIN 3.5 (Excoffier and Lischer, 2010) to test for the effects of

238   Flix dam and polluted sediments area on the population structure of mosquitofish along

239   the study river. In both cases genetic diversity was partitioned into three hierarchical

240   levels: within locations, among locations within sections, and among sections (above

241   and below the dam, or heavily polluted – F1 and F2 sites- and less polluted areas).

242   Group-level Bayesian analysis in BAPS 5.4 (Corander et al., 2008) was used to

243   cluster locations that frequently exchange individuals. In addition, to identify genetic

244   discontinuities in specific geographical areas, we applied Monmonier's algorithm using

245   BARRIER 2. 2 (Manni et al., 2004), which detects hidden barriers to genetic exchange

246   among sites according to their geographical coordinates and relative to the genetic

247   differentiation (F_{ST}). This approach allows the identification of the main genetic

248　discontinuities existing between sampling sites (Gagnon and Angers, 2006). BARRIER

249　analyses were conducted using a F_{ST} matrix from each GPI polymorphism and from

250　each microsatellite locus obtained from the FreeNA software and incorporating null

251　allele frequencies. We identified the two main barriers for each locus and retained those

252　confirmed by at least six loci. Recent migration rates among locations were assessed by

253　the Bayesian multilocus method implemented in BayesAss 3.0 (Wilson and Rannala,

254　2003). This method does not assume migration-drift or HWE. Because of the low

255　differentiation detected among locations, long runs of a total of 10×10^6 iterations were

256　performed to ensure that the MCMC chains reached stationarity (Faubet et al., 2007).

257　Migration parameters were estimated by sampling every 1000 iterations, after a burn-in

258　of 5×10^6 iterations. Delta values were adjusted following the user's manual

259　recommendations. We combined results from the 10 runs to estimate migration rates

260　between locations, using Tracer v1.5 (Rambaut and Drummond, 2007) .

261　　To formally evaluate whether variation at the *GPI-2* locus was significantly different

262　than the expected from the levels of neutral genetic variation, we used the F_{ST} -outlier

263　approach implemented in LOSITAN (Antao et al., 2008). LOSITAN uses coalescent

264　simulations to generate a null distribution of F_{ST} values, allowing to compare with the

265　observed values and hence to detect selection processes in codominant markers. With

266　this approach, loci presenting unusually high F_{ST} values are candidates subject to

267　directional selection, while loci with very low F_{ST} are candidates of balancing selection.

268　We generated a null distribution using 10^5 simulations and performed the test using

269　both the stepwise mutation model (SMM) and the infinite alleles model (IAM), with a

270　significance threshold of 0.05.

271　**3. Results**

272   3.1 Genetic diversity patterns

273   Allozyme analysis detected three alleles at the *GPI-2* locus, scored as fast, slow, and

274   ultraslow (average frequencies of 0.49, 0.32, and 0.19). Based on relative mobility and

275   abundances, we considered that they corresponded respectively to the $GPI\text{-}2^{100}$, *GPI-*

276   2^{66} and $GPI\text{-}2^{38}$ alleles observed in American collections (e.g. Diamond et al., 1989).

277   Exact probabilities tests indicated no deviations from HWE in any collection for both

278   the *GPI-2* allozyme locus and the GPI-SNP. After adjusting for multiple comparisons,

279   significant genotypic linkage disequilibrium between allozyme alleles and GPI-SNP

280   variants was observed only in site D2.

281   The eleven microsatellite loci were polymorphic across all populations, except locus

282   M*f*13 in F3, where the 164 allele was fixed. The Gaaf13 locus was the most diverse,

283   showing eleven alleles. The total number of alleles across microsatellite loci varied

284   from 39 (site F1) to 54 (site D4). Site F4 had the highest mean allele richness per locus

285   (3.817) and the D1 the lowest (3.372) (Table 2). After correcting for multiple

286   comparisons, significant deviations from HWE were observed in all locations but D4.

287   MICROCHECKER analyses suggested that null alleles were responsible for these

288   significant deviations; in particular, the Gafμ6 and Gaaf10 loci showed null alleles in

289   every population (estimated null allele frequencies ranged from 0.09 to 0.21 at Gafμ6

290   and from 0.13 to 0.29 at Gaaf10). After removing these two loci, departures from the

291   HWE expectations were reduced to a single site (F3). Despite that null alleles could

292   bias population structure inference, we maintained the whole set of 11 microsatellite

293   loci for further analyses because the influence of null alleles in a reduced subset of loci

294   (2 out of 11 in our case) is slight on Bayesian analyses of individual assignment

14

295 (Carlsson, 2008) and has minimal and statistically corrigible effects on estimates of

296 population differentiation (Chapuis and Estoup, 2007).

297   Genetic diversity measured as H_E within collections ranged 0.494-0.670 for the *GPI-*

298 *2* locus, 0.304-0.496 for the GPI-SNP and averaged from 0.479 to 0.539 for the 13 loci.

299 At the *GPI-2* locus, significant lower diversity was observed in collections downstream

300 of Flix Reservoir (Mann-Whitney *U*-test, $P = 0.014$). At the microsatellite loci,

301 significant lower allele richness was indicated for the polluted sites F1 and F2 (Mann-

302 Whitney *U*-test, $P = 0.044$), but differences were not observed between locations above

303 and below the dam (Mann-Whitney *U*-test, $P = 0.853$). BOTTLENECK evidenced

304 heterozygosity excess relative to a population in mutation-drift equilibrium only at the

305 F4 location. LDNe results suggested very large effective population sizes for locations

306 upstream of the polluted area (Table 2). Finite effective sizes from 166.9 to 581.8 were

307 detected at the polluted area and downstream locations.

308 3.2 Population differentiation

309   In the SNP polymorphism, the *C* variant was more abundant than the *T* across all

310 collections (average frequency 0.73) and no significant differences were observed

311 among study sites. Significant variation was observed at the *GPI-2* locus, with the **100*

312 allele being more abundant in locations downstream of the Flix dam, where the

313 ultraslow allele *GPI-2[38]* reduced its presence (Fig. 2). Overall, a small but significant

314 global population differentiation was detected ($F_{ST} = 0.016$). Correcting for null alleles

315 did not substantially change the estimate ($F_{ST} = 0.017$). In addition to the GPI allozyme

316 locus, seven microsatellite loci (Gafµ3, Gafµ5, Gafµ6, Gafµ7, Gaaf7, Gaaf10, and

317 Gaaf13) contributed to differences among study locations. LOSITAN scans did not

318 fully confirm the *GPI-2* locus as a potential candidate for directional selection (Fig. 3).

319   Significant differentiation was detected in 31 out of 45 pairwise comparisons

320   between locations, 18 of them still significant after Bonferroni's correction. The largest

321   pairwise F_{ST} (0.058) was detected between the D1 and D4 sites. D4, the lowermost

322   location in the study area, was significantly different from every other location except

323   D3 (Table 3). No significant allele frequency divergences were detected in pairwise

324   comparisons involving collections upstream of the polluted area in Flix Reservoir (F6,

325   F5, F4 and F3 samples) but significant divergence was detected between the collection

326   from the most polluted area in the reservoir (F1) and the uppermost location F6 (Table

327   3). A Mantel test revealed a significant correlation between genetic ($F_{ST}/(1-F_{ST})$) and

328   hydrographical distances ($P = 0.013$) when all locations were included into the

329   analysis, but no significant correlation was detected ($P = 0.455$) when the D3 and D4

330   locations were not included (Fig. 4).

331   The first AMOVA, considering the dam as a barrier between upstream and

332   downstream collections, assigned a portion of the differentiation to this separation (F_{CT}

333    $= 0.00652$, $P < 0.01$). This difference was not significant when the furthermost

334   downstream D3 and D4 sites were removed from the analyses with all loci ($F_{CT} =$

335   0.00009, $P > 0.05$), but the *GPI-2* locus retained substantial distinction between the two

336   population groups (Table 4). The second AMOVA, which evaluated the relevance of

337   the polluted sediments, suggested no divergence between fish captured in the most

338   polluted sites (F1 and F2) and all the other captures ($F_{CT} = 0.00491$, $P > 0.05$).

339   However, significant divergence was indicated among locations within groups ($F_{SC} =$

340   0.01596, $P < 0.001$) when this analysis focused only on populations behind the Flix

341   dam (locations F1 to F6), a larger and significant portion of the divergence was

342   allocated between ($F_{CT} = 0.00903$) than within groups ($F_{SC} = 0.00154$) (Table 4), but

343   this structure inside the reservoir was not reflected at GPI polymorphisms.

16

344   Monmonier's algorithm suggested that the main genetic barrier for the *GPI-2*

345   allozyme locus was the isolation of the polluted F1 site. Three consensus barriers were

346   detected from the analyses involving all 13 loci: (i) one isolating the locations closest to

347   the polluted sediments (F1 and F2) from the rest; (ii) another leaving the site D1 just

348   beneath the Flix dam as a *cul-de-sac*; and (iii) a third one between the lowermost sites

349   (D3 and D4) and the rest of studied locations. In agreement with this third barrier,

350   BAPS results suggested that the sites D4 and D3 form a distinctive group. In previous

351   studies, BAPS has proved to be very conservative in identifying population groups;

352   consequently, different gene pools could only be reliably detected under situations of

353   very restricted migration between them (Waples and Gaggiotti, 2006). Hence, we

354   believe that the two population groups detected by BAPS probably represent only a

355   conservative estimate.

356   3.3 Connectivity among locations

357   Bayesian estimates of contemporary migration rates showed that in all sites a portion

358   of ~30 % of individuals represented immigrant fish from other locations (Table 5). For

359   each site, a random origin of these local immigrants from the other nine locations

360   should result on average estimates of immigration rates around 0.033 (0.3/9), which is

361   fairly coincident with the estimated average migration rate (m) among all locations

362   (0.032, 95% C.I. from -0.017 to 0.081). Relevant downstream dispersal was observed

363   from location D3 to D4 ($m = 0.110$), and from location F3 to F2 ($m = 0.089$), F1 ($m =$

364   0.096), and D1 ($m = 0.096$). Significant upstream migration was observed from D2

365   (below the dam) to the uppermost location F6 ($m = 0.092$). On average, locations F3

366   and D2 were sources of fish to several other locations around Flix Reservoir.

367   **4. Discussion**

368   In the Ebro River the amount of diversity detected at microsatellite loci was lower than

369   in North American populations. The averaged expected heterozygosity in our study

370   sites (0.479-0.539) was rather low compared to native American locations: 0.4-0.9 in

371   Zane et al. (1999), 0.5-0.78 in Spencer et al. (1999), and 0.75-0.76 in Purcell et al.

372   (2012b). Nevertheless, Mann-Whitney U tests confirmed ($P < 0.001$) higher allele

373   richness and expected heterozygosity at microsatellite loci of the study Ebro River

374   locations than other introduced populations in Spain northward of the Ebro River

375   (average $H_E = 0.432$, average $Ar = 2.83$, Díez-del-Molino et al., 2013), and in Australia

376   (average $H_E = 0.410$, average $Ar = 2.40$, Ayres et al., 2010). Reduced diversity of

377   mosquitofish in northern locations probably arised as consequence of founding events

378   from Iberian other locations, such as the one in the Ebro River delta (Díez-del-Molino

379   et al. 2013). In Australia, mosquitofish was introduced in 1926 from Italian populations

380   (Lloyd and Tomasov, 1985), and these Italian populations probably derived from the

381   ones introduced to Span in 1921 (Vidal et al. 2010). Therefore, bottelnecks and founder

382   effects related with sequential introductions are likely to play a role in reducing the

383   amount of genetic diversity in more recently invaded locations. However, Vera et al.

384   (2016) showed that balancing selection can retain specific polymorphisms to insure the

385   survival of introduced mosquitofish populations.

386   The average level of population differentiation among invasive mosquitofish

387   populations in the Ebro basin ($F_{ST} = 0.016$) was consistent with observed patterns in the

388   native range, where large transects along river basins were occupied by a single

389   mosquitofish metapopulation (McClenaghan et al., 1985; Smith et al., 1989).

390   Nevertheless, mosquitofish can disperse at rates greater than 800 m day^{-1} in unimpeded

391   corridors (Alemadi and Jenkins, 2007), and this high dispersal capability has been

392   considered responsible for the positive spatial autocorrelation of allele frequencies

393   between populations at distances of 6-150 km within basins (Smith et al., 1989). Our

394   analyses suggest a pattern of isolation by distance (IBD) between mosquitofish

395   populations along the Ebro River, in agreement with results in other invaded Iberian

396   basins at geographical scales up to 30 km (Díez-del-Molino et al., 2013). Nevertheless,

397   the average F_{ST} value in the studied area was significantly lower than in populations

398   from other river watersheds in northeastern Spain (range 0.061-0.104, Díez-del-Molino

399   et al., 2013). However, meanwhile in these other basins, connectivity between locations

400   is often lost during summer drought periods, a permanent waterflow in the Ebro River

401   probably contributes to maintain gene flow among the studied mosquitofish

402   populations.

403   Dams represent strong barriers for river fish dispersal, and may prevent the

404   expansion of invasive fish species (Haynes et al., 2009). Because of this, genetic

405   variation in populations above dams is expected to decline because of limited upstream

406   gene flow (Horreo et al. 2013; Yamamoto et al., 2004). Nevertheless, we found similar

407   average levels of genetic diversity between mosquitofish collections above and below

408   the Flix dam. In fact, a large proportion of immigrant fish to each location was

409   suggested by the Bayesian estimates of contemporary gene flow (~30 %). We detected

410   substantial rates of contemporary mosquitofish dispersal from F3 site to several other

411   locations below and above the dam, including the polluted F1 and F2 locations, as well

412   as substantial upstream migration to F6 above the dam from the downstream D2

413   location, probably because this later location is close to the mouth of water diversions

414   associated with the hydroelectric power plant. In addition to the channel to turbines, the

415   diversions includes a fish ladder proven to be ineffective for native species (CHE,

416   2009), but may be used by mosquitofish due to their recognized success in colonizing

417   ponds, irrigation ditches, and modified stream channels (Courtenay and Meffe, 1989).

418   Additionally, the fact that Flix town is located both above and below the dam, may

419   favor human-assisted translocations to upstream locations, as suggested for other

420   Iberian basins (Díez-del-Molino et al., 2016).

421   Pollutants can contribute to reduce levels of genetic diversity by declining effective

422   population sizes (Van Straalen and Timmermans, 2002). In Flix Reservoir,

423   contaminated sediments have affected several fish species for traits involved in the

424   recruitment of populations such as increased frequency of ectoparasites, reduced body

425   condition, gonadal weight, and fecundity (Benejam et al., 2010). We estimated finite

426   effective population sizes for mosquitofish in the most polluted locations (F1 and F2)

427   as well as in locations downstream (D1 to D4), but evidence for recent bottlenecks was

428   restricted to the F4 location, probably because these fish inhabit a lagoon that gets

429   isolated from the river mainstem during periods of low water flow. Meanwhile we

430   found similar levels on average heterozygosity when comparing polluted and non-

431   polluted locations either for microsatellite loci (Mann-Whitney U-test $P = 1.000$) and

432   the GPI polymorphisms ($P = 0.711$ for the allozyme and $P = 0.178$ for the SNP),

433   reduced microsatellite allele richness in polluted locations was detected (Mann-

434   Whitney U-test $P = 0.044$), probably because allele richness is more sensitive to

435   bottlenecks than average heterozygosity (Spencer et al., 2000). Similar results were

436   observed in populations of the three-spined stickleback, *Gasterosteus aculeatus*,

437   inhabiting polluted areas. Despite these stickleback populations had sometimes suffered

438   bottlenecks, these did not consistently result in significant and detectable reductions of

439   heterozygosity (Santos et al., 2013).

440   The presumptive neutral SNP-GPI polymorphism showed stable allele frequencies

441   among studied locations, which contrasted with significant allele frequency changes at

the allozyme *GPI-2* locus. BARRIER analyses indicated that the most abrupt shift in allele frequencies at this locus occurred at the F1 site, the closest to the toxic sediments in the Flix reservoir. The observed allele frequency changes at this locus generally agree with the literature on genetic effects of chronic exposition of mosquitofish to mercury. For example, small but significant reductions in the $GPI\text{-}2^{100}$ allele frequency have been related with the decrease of reproductive fitness at mercury-exposed sublethal concentrations because homozygous $GPI\text{-}2^{100/100}$ females were less likely to be gravid and had fewer developed embryos (Mulvey et al., 1995; Tatara et al., 1999). Accordingly, we observed a lower frequency of the $GPI\text{-}2^{100}$ allele (average $q = 0.400$) in the most toxic sites (F2 and F1) than in downstream locations (D1, D2, D3, D4; average $q = 0.588$; exact probability test, $P < 0.001$), where mercury bioaccumulation is reduced as indicated by results on the zebra mussel (Carrasco, et al., 2008) and crayfish *(*Suárez-Serrano et al., 2010*)*. We also detected higher frequencies of the *GPI-2^{66}* allele in polluted F1 and F2 locations (average $q = 0.400$) than in downstream D1, D2, D3, and D4 locations (average $q = 0.284$, exact probability test $P = 0.013$), as expected from selective pressures on experimental expositions of mosquitofish to acute mercury concentrations (Tatara et al., 1999, 2002). Nevertheless, our selection scans did not confirm specific selective pressures at the *GPI-2* locus. Instead, AMOVA analyses indicated that the distinction between polluted and non-polluted locations at Flix Reservoir mostly resulted from allele frequency changes at the microsatellite loci rather than at the *GPI-2* locus. Despite being considered as neutral markers, positive selective responses at microsatellite loci have been observed in animals (e.g., captive *Salmo salar* populations, Portnoy et al., 2014) and plants (e.g., Australian populations of *Eucalyptus grandis*, Song et al., 2016). LOSITAN results suggested that the Gafμ6 locus could be under directional selection. However, even though this result should be

467 considered with caution as this locus presented evidence of null alleles, Dharmarajan et

468 al. (2013) noted that throwing out loci with positive F_{IS} values (often considered to

469 have null alleles) weakens the ability to detect biological phenomena such as Wahlund

470 effects resulting from high rates of migration among locations. Newman and Jagoe

471 (1998) showed that migration partially mitigates the effect of selection on the *GPI-2*

472 locus in mosquitofish under experimental exposures to mercury of 1 mg L^{-1}. Similarly,

473 fish dispersal hampered the local selective pressure of thermal effluents on whitefish

474 species in Huron Lake, Canada (Graham et al., 2016). Another pollution-tolerant fish

475 species, the central stoneroller (*Campostoma anomalum*) showed temporal stable

476 population structure along the 45 km of a polluted and degraded catchment because of

477 persistence of source-sink dynamics, where immigration from populations in relatively

478 better quality habitats prevented local differentiation (Waits et al., 2008), and similar

479 source-sink dynamics have been suggested to be responsible of stable population

480 structure in invaded mosquitofish populations (Díez-del-Molino et al., 2016).

481 4.1 Conclusions

482   Freshwater ecosystems are amongst the most affected by invasions worldwide.

483 Dams, weirs, and polluted areas are perceived as barriers to fish populations and hence

484 to limit the spread of invasive fish. In this study, we detected allelic changes at the

485 *GPI-2* locus suggesting a genetic response of the invasive mosquitofish (*G. holbrooki*)

486 to pollutants that is partially supported by reduced diversity levels at microsatellite loci

487 in the most polluted locations. Elevated rates of gene flow between locations within the

488 study area, including from upstream sites to polluted ones and from downstream to

489 upstream sites of the Flix dam, mitigate the effects of water pollution and the dam on

490 the genetic structure of invasive mosquitofish populations. These observations raise

491 concerns about dams and pollutants being effective barriers to the expansion of

492 mosquitofish as well as other invasive fish. Currently, mosquitofish is recorded all

493 along the Ebro River mainsteam, but species distribution models suggest many

494 unoccupied suitable areas in its tributaries. These models showed that natural

495 environmental conditions rather than anthropogenic perturbations regulated the

496 presence and abundance of mosquitofish in invaded Iberian rivers (Murphy et al. 2015).

497 To dredger the waste deposited at Flix reservoir, from 2010 to 2012 a retaining wall

498 was build (Palenques et al, 2015). Dredging activities started in 2013, and soon

499 afterwards pollutants in fish sampled at downstreams locations have drastically reduced

500 (Blanco et al. 2018).

508   **References**

509   Alcaraz, C., Bisazza, A., García-Berthou, E., 2008. Salinity mediates the competitive

510   interactions between invasive mosquitofish and an endangered fish. Oecologia, 155(1),

511   205-213.

512   Alcaraz, C., Caiola, N., Ibáñez, C., 2011. Bioaccumulation of pollutants in the zebra

513   mussel from hazardous industrial waste and evaluation of spatial distribution using

514   GAMs. Science of The Total Environment, 409(5), 898-904.

515   Alemadi, S. D., Jenkins, D. G., 2007. Behavioral constraints for the spread of the

516   eastern mosquitofish, *Gambusia holbrooki* (Poeciliidae). Biological Invasions, 10(1),

517   59-66.

518   Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., Luikart, G, 2008. LOSITAN: A

519   workbench to detect molecular adaptation based on a F_{st}-outlier method. BMC

520   Bioinformatics, 9, 323.

521   Araguas, R. M., Roldán, M. I., García-Marín, J. L., Pla, C., 2007. Management of gene

522   diversity in the endemic killifish *Aphanius iberus*: revising Operational Conservation

523   Units. Ecology of Freshwater Fish, 16(2), 257-266.

524   Ayres, R., Pettigrove, V., Hoffmann, A., 2010. Low diversity and high levels of

525   population genetic structuring in introduced eastern mosquitofish (*Gambusia*

526   *holbrooki*) in the greater Melbourne area, Australia. Biological Invasions, 12(11),

527   3727-3744.

528 Bélanger-Deschênes, S., Couture, P., Campbell, P. G. C., Bernatchez, L., 2013.

529 Evolutionary change driven by metal exposure as revealed by coding SNP genome scan

530 in wild yellow perch (*Perca flavescens*). Ecotoxicology, 22(5), 938-957.

531 Benejam, L., Benito, J., García-Berthou, E., 2010. Decreases in condition and fecundity

532 of freshwater fishes in a highly polluted reservoir. Water, Air, & Soil Pollution, 190(1),

533 3-11.

534 Blanchet, S., 2012. The use of molecular tools in invasion biology: an emphasis on

535 freshwater ecosystems. Fisheries Management and Ecology, 19(2), 120-132.

536 Bosch, C., Olivares, A., Faria, M., Navas, J. M., del Olmo, I., Grimalt, J. O., Piña, B.,

537 Barata, C., 2009. Identification of water soluble and particle bound compounds causing

538 sublethal toxic effects. A field study on sediments affected by a chlor-alkali industry.

539 Aquatic Toxicology, 94(1), 16-27.

540 Carlsson, J., 2008. Effects of microsatellite null alleles on assignment Testing. Journal

541 of Heredity, 99 (6), 616-623.

542 Carrasco, L., Díez, S., Soto, D. X., Catalan, J., Bayona, J. M., 2008. Assessment of

543 mercury and methylmercury pollution with zebra mussel (*Dreissena polymorpha*) in

544 the Ebro river (NE Spain) impacted by industrial hazardous dumps. Science of The

545 Total Environment, 407(1), 178-184.

546 Carrasco, L., Barata, C., García-Berthou, E., Tobias, A., Bayona, J. M., Díez, S., 2011.

547 Patterns of mercury and methylmercury bioaccumulation in fish species downstream of

548 a long-term mercury-contaminated site in the lower Ebro river (NE Spain).

549 Chemosphere, 84(11), 1642-1649.

550 Chapman, P., Warburton, K., 2006. Postflood movements and population connectivity

551 in gambusia (*Gambusia holbrooki*). Ecology of Freshwater Fish, 15(4), 357-365.

552 Chapuis, M. P., Estoup, A., 2007. Microsatellite null alleles and estimation of

553 population differentiation. Molecular Biology and Evolution, 24(3), 621-631.

554 Cid, N., Ibáñez, C., Palanques, A., Prat, N., 2010. Patterns of metal bioaccumulation in

555 two filter-feeding macroinvertebrates: Exposure distribution, inter-species differences

556 and variability across developmental stages. Science of The Total Environment,

557 408(14), 2795-806.

558 Confederacion Hidrográfica del Ebro (CHE), 2009. Asistencia técnica para el estudio

559 de propuestas de mejora de la conectividad para los peces en la parte baja del río Ebro.

560 http://195.55.247.234/webcalidad/estudios/200907_conectividad_peces_Bajo_Ebro.pdf

561 (accessed 31 Juanary 2018)

562 Congdon, B. C., 1994. Salinity-related fitness differences amongst GPI genotypes in

563 the mosquitofish *Gambusia holbrooki* (Poeciliidae: Teleostei). Biological Journal of the

564 Linnean Society, 53(4), 343-352.

565 Congdon, B. C., 1995. Unidirectional gene flow and maintenance of genetic diversity

566 in the mosquitofish *Gambusia holbrooki* (Teleostei: Poeciliidae). Copeia, 1995(1), 162-

567 172.

568 Corander, J., Marttinen, P., Sirén, J., Tang, J. 2008 . Enhanced Bayesian modelling in

569 BAPS software for learning genetic structures of populations. BMC Bioinformatics, 9,

570 539

571   Courtenay, W. R. Jr., Meffe, G. K., 1989. Small fishes in strange places: a review of

572   introduced poeciliids, in: Meffe, G.K., Snelson, F.F. Jr. (Eds), Ecology and evolution of

573   livebearing fishes (Poeciliidae). Prentice Hall, Englewood Cliffs, NJ, pp. 319-331.

574   Crispo, E., Bentzen, P., Reznick, D. N., Kinnison, M. T., Hendry, A. P., 2006. The

575   relative influence of natural selection and geography on gene flow in guppies.

576   Molecular Ecology, 15(1), 49-62.

577   Deacon, A. E., Ramnarine, I. W., Magurran, A. E., 2011. How reproductive ecology

578   contributes to the spread of a globally invasive fish. PLoS ONE, 6(9), e24416.

579   Dempster, A. P., Laird, N. M., Rubin, D. B., 1977. Maximum likelihood from

580   incomplete data via the EM algorithm. Journal of the Royal Statistical Society Series B

581   (Methodological), 39(1), 1-38.

582   Dharmarajan, G., Beatty, W.S., Rhodes Jr, O. E., 2013. Heterozygote deficiencies

583   caused by a Wahlund effect: Dispelling unfounded expectations. The Journal of

584   Wildlife Management, 77(2), 226-234.

585   Diamond, S. A., Newman, M. C., Mulvey, M., Dixon, P. M., Martinson, D., 1989.

586   Allozyme genotype and time to death of mosquitofish, *Gambusia affinis* (Baird and

587   Girard), during acute exposure to inorganic mercury. Environmental Toxicology and

588   Chemistry, 8(7), 613-622.

589   Díez-del-Molino, D., Carmona-Catot, G., Araguas, R. -M., Vidal, O., Sanz, N., García-

590   Berthou, E., García-Marín, J. -L., 2013. Gene flow and maintenance of genetic

591   diversity in invasive mosquitofish (*Gambusia holbrooki*). PLoS ONE, 8(12), e82501.

592 Díez-del-Molino, D., Araguas, R. -M., Vera, M., Vidal, O., Sanz, N., García-Marín, J. -

593 L., 2016. Temporal genetic dynamics among mosquitofish (*Gambusia holbrooki*)

594 populations in invaded watersheds. Biological Invasions, 18(3), 841-855.

595 Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., Ovenden, J. R., 2014 .

596 NEESTIMATOR V2: re-implementation of software for the estimation of contemporary

597 effective population size (*Ne*) from genetic data. Molecular Ecology Resources, 14(1)**,**

598 209-214.

599 Excoffier, L., Lischer, H. E. L., 2010. Arlequin suite ver 3. 5: A new series of programs

600 to perform population genetics analyses under Linux and Windows. Molecular Ecology

601 Resources, 10(3), 564-567.

602 Faubet, P., Waples, R. S., Gaggiotti, O. E., 2007. Evaluating the performance of a

603 multilocus Bayesian method for the estimation of migration rates. Molecular Ecology,

604 16(6), 1149-1166.

605 Gagnon, M. C., Angers, B., 2006. The determinant role of temporary proglacial

606 drainages on the genetic structure of fishes. Molecular Ecology, 15(4), 1051- 1065.

607 García-Berthou, E., Alcaraz, C., Pou-Rovira, Q., Zamora, L., Coenders, G., Feo, C.,

608 2005. Introduction pathways and establishment rates of invasive aquatic species in

609 Europe. Canadian Journal of Fisheries and Aquatic Sciences, 62(2), 453-463

610 Graham, C. F., Eberts, R. L., Morgan, T. D., Boreham, D. R., Lance, S. L., Manzon, R.

611 G., Martino, J. A., Rogers, S. M, Wilson, J. Y., Somers C. M., 2016. Fine-Scale

612 Ecological and Genetic Population Structure of Two Whitefish (Coregoninae) Species

613 in the Vicinity of Industrial Thermal Emissions. PLoS ONE, 11(1), e0146656.

614 Guo, S. W., Thompson, E. A., 1992. Performing the exact test of Hardy-Weinberg

615 proportion for multiple alleles. Biometrics, 48(2), 361-372

616 Heagler, M. G., Newman, M. C., Mulvey, M., Dixon, P. M. 1993. Allozyme genotype

617 in mosquitofish, *Gambusia holbrooki*, during mercury exposure: temporal stability,

618 concentration effects and field verification. Environmental Toxicology and Chemistry,

619 12(2), 385-395.

620 Hernandez-Martich, J. D., Smith, M. H., 1997. Downstream gene flow and genetic

621 structure of *Gambusia holbrooki* (eastern mosquitofish) populations. Heredity, 79, 295-

622 301.

623 Haynes, G. D., Gilligan, D. M., Grewe, P., Nicholas, F. W., 2009. Population genetics

624 and management units of invasive common carp *Cyprinus carpio* in the Murray-

625 Darling Basin, Australia. Journal of Fish Biology, 75(2), 295-320.

626 Horreo, J. L., Martinez, J. L., Ayllón, F., Pola, I. G., Monteoliva, J. A., Héland, M.,

627 Garcia-Vazquez, E., 2011. Impact of habitat fragmentation on the genetics of

628 populations in dendritic landscapes. Freshwater Biology, 56(12), 2567-2579.

629 Kennedy, P. K., Kennedy, M. L., Zimmerman, E. G., Chesser, R. K., Smith, M. H.,

630 1986. Biochemical genetics of mosquitofish. 5. Perturbation effects on genetic

631 organization of populations. Copeia, 1986(4), 937-945.

632 Langerhans, R. B., Gifford, M. E., Joseph, E. O., 2007. Ecological speciation in

633 *Gambusia* fishes. Evolution, 61(9), 2056-2074.

634 Lloyd, L., Tomasov, J., 1985. Taxonomic status of the mosquitofish, *Gambusia affinis*

635 (Poeciliidae), in Australia. Marine and Freshwater Research, 36 (3), 447-451.

636 Manni, F., Guérard, E., Heyer, E., 2004. Geographic patterns of (genetic, morphologic,

637 linguistic) variation: how barriers can be detected by using Monmonier's algorithm.

638 Human Biology, 76(2), 173-190.

639 McClenaghan, L. R., Smith, M. H., Smith, M. W., 1985. Biochemical genetics of

640 mosquitofish. IV. Changes of allele frequencies through time and space. Evolution,

641 39(2), 451-460.

642 McLellan, R., Iyengar, L., Jeffries, B., Oerlemans, N., 2014. Living Planet Report

643 2014: Species and spaces, people and places. World Wildlife Fund (WWF), Gland,

644 Switzerland.

645 Meffe, G. K., Weeks, S. C., Mulvey, M., Kandl, K. L., 1995. Genetic differences in

646 thermal tolerance of eastern mosquitofish (*Gambusia holbrooki*; Poeciliidae) from

647 ambient and thermal ponds. Canadian Journal of Fisheries and Aquatic Science, 52(12),

648 2704-2711.

649 Mulvey, M., Newman, M. C., Chazal, A., Keklak, M. M., Heagler, M. G., Hales, L. J.

650 1995. Genetic and demographic responses of mosquitofish (*Gambusia holbrooki* Girard

651 1859) populations stressed by mercury. Environmental Toxicology and Chemistry,

652 14(8), 1411-1418.

653 Navarro-Garcia, J., 2013. De Buen family and the introduction of the" Gambusia":

654 environmental consequences of the fight against malaria in Spain. Boletín de

655 Malariología y Salud Ambiental, LIII (1), 99-112.

656 Navarro, A., Quirós, L., Casado, M., Faria, M., Carrasco, L., Benejam, L., Benito, J.,

657 Díez, S., Raldúa, D., Barata, C., Bayona, J. M., Piña, B., 2009. Physiological responses

658 to mercury in feral carp populations inhabiting the low Ebro River (NE Spain), a

659 historically contaminated site. Aquatic Toxicology, 93(2-3), 150-157.

660 Newman, M. C., Jagoe, R. H., 1998. Allozymes reflect the population-level effect of

661 mercury: simulations of the mosquitofish (*Gambusia holbrooki* Girard) GPI-2

662 response. Ecotoxicology, 7(3), 141-150.

663 Palsbøll, P. J., Bérubé, M., Allendorf, F. W., 2007. Identification of management units

664 using population genetic data. Trends in Ecology & Evolution, 22(1), 11-16.

665 Piry, S., Luikart, G., Cournet, J. M., 1999. BOTTLENECK: a computer program for

666 detecting recent reductions in the effective population size using allele frequency data.

667 Journal of Heredity, 90(4), 502-503.

668 Portnoy, D., Hollenbeck, C., Vidal, R., Gold, J. 2014. A comparison of neutral and

669 immune genetic variation in Atlantic salmon, *Salmo salar* L. in Chilean aquaculture

670 facilities. PLoS ONE, 9 (6), 1-13.

671 Purcell, K., Stockwell, C. A., 2015. An evaluation of the genetic structure and post-

672 introduction dispersal of a non-native invasive fish to the North Island of New Zealand.

673 Biological Invasions, 17(2), 625-636.

674 Purcell, K. M., Hitch, A., Martin, S., Klerks, P. L., Leberg, P. L., 2012a. The role of

675 genetic structure in the adaptive divergence of populations experiencing saltwater

676 intrusion due to relative sea-level rise. Journal of Evolutionary Biology, 25(12), 2623-

677 2632.

678 Purcell, K. M., Ling, N., Stockwell, C. A., 2012b. Evaluation of the introduction

679 history and genetic diversity of a serially introduced fish population in New Zealand.

680 Biological Invasions, 14(10), 2057-2065.

681 Pyke, G. H., 2005. A review of the biology of *Gambusia affinis* and *G. holbrooki*.

682 Reviews in Fish Biology and Fisheries, 15(4), 339-365.

683 Quirós, L., Ruiz, X., Sanpera, C., Jover, L., Piña, B., 2008. Analysis of micronucleated

684 erythrocytes in heron nestlings from reference and impacted sites in the Ebro basin (NE

685 Spain). Environmental Pollution, 155(1), 81-87.

686 Rambaut, A., Drummon, A. J., 2007. Tracer v1. 4, Available from

687 http://tree.bio.ed.ac.uk/software/tracer/ (accessed 02 February 2018)

688 Rehage, J. S., Sih, A., 2004. Dispersal behavior, boldness, and the link to invasiveness:

689 a comparison of four *Gambusia* species. Biological Invasions, 6(3), 379-391.

690 Rinner, B. P., Matson, C. W., Islamzadeh, A., McDonald, T. J., Donnelly, K. C.,

691 Bickham, J. W., 2011. Evolutionary toxicology: contaminant-induced genetic

692 mutations in mosquitofish from Sumgayit, Azerbaijan. Ecotoxicology, 20(2), 365-376.

693 Roark, S. A., Andrews, J. F., Guttman, S. I., 2001. Population genetic structure of the

694 western mosquitofish, *Gambusia affinis*, in a highly channelized portion of the San

695 Antonio River in San Antonio, TX. Ecotoxicology, 10(4), 223-227.

696 Roberts, J. H., Angermeier, P. L., Hallerman, E. M., 2013. Distance, dams and drift:

697 what structures populations of an endangered, benthic stream fish? Freshwater Biology,

698 58(10), 2050-2064.

699 Rousset, F., 2008. GENEPOP'007: a complete reimplementation of the GENEPOP

700 software for Windows and Linux. Molecular Ecology Resources, 8(1), 103-106.

701 Santos, J., Pascual, M., Simões, P., Fragata, I., Rose, M. R., Matos, M., 2013. Fast

702 evolutionary genetic differentiation during experimental colonizations. Journal of

703 Genetics, 92(2), 183-94.

704 Sanz. N., Araguas, R. M., Vidal, O., Diez-del-Molino, D., Fernández-Cebrián, R.,

705 García-Marín, J. L., 2013. Genetic characterization of the invasive mosquitofish

706 (*Gambusia* spp.) introduced to Europe: population structure and colonization routes.

707 Biological Invasions, 15(10), 2333-2346.

708 Smith, M. H., Scribner, K. T., Hernandez, J. D., Wooten, M. C., 1989. Demographic,

709 spatial, and temporal genetic variation in Gambusia. In G. K. Meffe & F. F Snelson Jr.

710 (Eds), *Ecology and evolution of livebearing fishes (Poeciliidae)* (pp 235-257).

711 Englewood Cliffs, NJ: Prentice Hall.

712 Smith, M. W., Smith, M. H., Chesser, R. K., 1983. Biochemical genetics of

713 mosquitofish. I. Environmental correlates, and temporal and spatial heterogeneity of

714 allele frequencies within a river drainage. Copeia, 1983(1), 182-193.

715 Song, Z., Zhang, M., Li, F., Weng, Q., Zhou, C., Li, M., Li, J., Huang, H., Mo, X., Gan,

716 S., 2016. Genome scans for divergent selection in natural populations of the widespread

717 hardwood species *Eucalyptus grandis* (Myrtaceae) using microsatellites. Scientific

718 Reports 6 (October), 1-13.

719 Soto, D. X., Benito, J., Gacia, E., García-Berthou, E., Catalan, J., 2016. Trace metal

720 accumulation as complementary dietary information for the isotopic analysis of

721 complex food webs. Methods in Ecology and Evolution, 7(8), 910-918.

722   Spencer, C. C., Chlan, C. A., Neigel, J. E., Scribner, K. T., Wooten, M. C., Leberg, P.

723   L., 1999. Polymorphic microsatellite markers in the western mosquitofish, *Gambusia*

724   *affinis*. Molecular Ecology, 8(1), 157-158.

725   Spencer, C. C., Neigel, J. E., Leberg, P. L., 2000. Experimental evaluation of the

726   usefulness of microsatellite DNA for detecting demographic bottlenecks. Molecular

727   Ecology, 9(10), 1517-1528.

728   Stockwell, C. A., Weeks, S. C., 1999. Translocations and rapid evolutionary responses

729   in recently established populations of western mosquitofish (*Gambusia affinis*). Animal

730   Conservation, 2(2), 103-110.

731   Suárez-Serrano, A., Alcaraz, C., Ibáñez, C., Trobajo, R., Barata, C., 2010.

732   *Procambarus clarkii* as a bioindicator of heavy metal pollution sources in the lower

733   Ebro River and Delta. Ecotoxicology and Environmental Safety, 73(3), 280-286.

734   Tatara, C. P., Mulvey, M. E., Newman, M. C., 1999. Genetic and demographic

735   responses of mosquitofish (*Gambusia holbrooki*) populations exposed to mercury for

736   multiple generations. Environmental Toxicology and Chemistry, 18(12), 2840-2845.

737   Tatara, C. P., Mulvey, M. E., Newman, M. C., 2002. Genetic and demographic

738   responses of mercury-exposed mosquitofish (*Gambusia holbrooki*) populations:

739   temporal stability and reproductive components of fitness. Environmental Toxicology

740   and Chemistry, 21(10), 2191-2197.

741   van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., Shipley, P., 2004. MICRO-

742   CHECKER: software for identifying and correcting genotyping errors in microsatellite

743   data. Molecular Ecology Resources, 4(3), 535-538

744 van Straalen, N., Timmermans, M., 2002. Genetic Variation in Toxicant-Stressed

745 Populations: An Evaluation of the "Genetic Erosion" Hypothesis. Human and

746 Ecological Risk Assessment: An International Journal, 8(5), 983-1002.

747 Vega-Retter, C., Vila, I., Véliz, D., 2015. Signatures of directional and balancing

748 selection in the silverside *Basilichthys microlepidotus* (Teleostei: Atherinopsidae)

749 inhabiting a polluted river. Evolutionary Biology, 42(2), 156-168.

750 Vera, M., Díez-del-Molino, D., & García-Marín, J. L. (2016). Genomic survey provides

751 insights into the evolutionary changes that occurred during European expansion of the

752 invasive mosquitofish (*Gambusia holbrooki*). Molecular Ecology, 25(5), 1089-1105.

753 Vidal, O., García-Berthou, E., Tedesco, P.A., García-Marín, J.-L., 2010. Origin and

754 genetic diversity of mosquitofish (*Gambusia holbrooki*) introduced to Europe.

755 Biological Invasions, 12 (4), 841–851.

756 Waits, E. R., Bagley, M. J., Blum, M. J., McCormick, F. H., Lazorchack, J. M., 2008.

757 Source-sink dynamics sustain central stonerollers (*Campostoma anomalum*) in a

758 heavily urbanized catchment. Freshwater Biology, 53(10), 2061-2075.

759 Waples, R. S., Gaggiotti, O., 2006. What is a population? An empirical evaluation of

760 some genetic methods for identifying the number of gene pools and their degree of

761 connectivity. Molecular Ecology, 15(6), 1419-1439.

762 Weir, B. S., Cockerham, C. C., 1984. Estimating *F*-statistics for the analysis of

763 population structure. Evolution, 38 (6), 1358-1370

764 Wilson, G. A., Rannala, B., 2003. Bayesian inference of recent migration rates using

765 multilocus genotypes. Genetics, 163(3), 1177-1191.

766   Yamamoto, S., Morita, K., Koizumi, I., Maekawa, K., 2004. Genetic differentiation of

767   white-spotted charr (*Salvelinus leucomaenis*) populations after habitat fragmentation:

768   Spatial-temporal changes in gene frequencies. Conservation Genetics, 5(4), 529-538.

769   Zane, L., Nelson, W. S., Jones, A. G., Avise, J. C., 1999. Microsatellite assessment of

770   multiple paternity in natural populations of a live bearing fish, *Gambusia holbrooki*.

771   Journal of Evolutionary Biology, 12(1), 61-69.

772   Zeng, Y., Díez-del-Molino, D., Vidal, O., Vera, M., García-Marín, J.-L, 2017. Multiple

773   paternity and reproduction opportunities for invasive mosquitofish. Hydrobiologia 795:

774   139–151.

775

776   **Figure captions**

777   Figure 1. Geographical location of the ten study sites. Flix town (intermediate grey),

778   the chemical plant (black), riverbank (light grey), and the Ebro River course are shown.

779   See Table 1 or the text for further information on the location codes (F1–F6 and D1–

780   D4).

781   Figure 2. Estimated frequency for the the *GPI-2*^{100} and *GPI-2*^{38} alleles in each location.

782   The position of the Flix dam is also shown. Location codes as in Table 1.

783   Figure 3. Selection scans for each loci performed with LOSITAN. Significance was set

784   at the 5 % level. *a*) Stepwise mutation model (SMM); *b*) Infinite alleles model (IAM).

785   Figure 4. Isolation-by-distance patterns in the study area. The regression of genetic

786   differentiation ($F_{ST}/(1\text{-}F_{ST})$) against geographical distance (km) for all pairs of sampled

787   sites is shown (regression line in grey, $r^2 = 0.679$). The black circles indicate

788   comparisons around Flix dam (F1 to D2 sites, regression line in black, $r^2 = 0.005$), and

789   white circles, comparisons involving D3 and D4 sites.

Table1. Geographical location and mercury concentration in the study sites. All latitudes are north and all longitudes are east. Total mercury (THg)

concentrations (mean or observed range) in sediment are from Bosch et al. (2009), in muscle of zebra mussel (*Dreissena polymorpha)* from Carrasco et al.

(2008), in common carp (*Cyprinus carpio*) from Carrasco et al. (2011), and in mosquitofish (*Gambusia holbrooki*) from Soto et al. (2011). ww = wet weight;

dw = dry weight; — no available data

Location	Code	Latitude	Longitude	Distance to toxic	Distance to river	THg in sediment	THg in zebra mussel	THg in common carp	THg in common carp	THg in mosquitofish
Reservoir upper	F6	41º 15' 13.43"	0º 30' 15.01"	4.5	118.5	—	—	—	—	—
Riparian forest	F5	41º 14' 30.65"	0º 30' 28.57"	3.0	117.1	—	0.01–0.05	—	—	—
Lagoon	F4	41º 14' 14.94"	0º 31' 31.10"	1.4	115.4	3.0	—	—	—	—
Channel	F3	41º 14' 08.92"	0º 31' 53.22"	0.9	114.9	3.0	0.03–0.14	—	—	—
close to toxic sediments	F2	41º 13' 59.51"	0º 32' 24.30"	0.1	114.1	—	—	0.33	1.42	1.23
over toxic sediments	F1	41º 13' 56.66"	0º 32' 27.42"	0.0	113.9	15.1	0.35–0.81	—	—	—
Below dam	D1	41º 14' 06.91"	0º 32' 55.38"	0.1	113.3	2.8	0.08–0.14	—	—	—
Meander	D2	41º 13' 46.19"	0º 33' 03.02"	5.8	108.2	2.7	—	0.68	—	—
Ascó village	D3	41º 11' 04.00"	0º 34' 20.91"	13.1	100.8	1.9	—	0.65	—	—
Garcia village	D4	41º 08' 08.81"	0º 38' 52.96"	22.8	91.2	1.4–1.9	—	—	—	—

795 Table 2. Genetic diversity within collections at the GPI allozyme locus, GPI-SNP and

796 11 microsatellite loci. *Ar*: allele richness; H_E: expected heterozygosis; F_{IS}: local

797 inbreeding coefficient measuring departures from HWE (in bold significant values after

798 Bonferroni's correction); *Ne*: Effective population size. Location codes as in Table 1.

799 v.l. very large population size.

Locus		Sites F6	F5	F4	F3	F2	F1	D1	D2	D3	D4
GPI polymorphims											
GPI-SNP											
	Ar	2.000	2.000	2.000	2.000	2.000	2.000	2.000	2.000	2.000	2.000
	H_E	0.309	0.392	0.406	0.309	0.443	0.445	0.367	0.381	0.380	0.496
	F_{IS}	0.110	-0.085	0.384	0.110	-0.242	0.101	-0.023	0.212	0.079	0.092
GPI-2											
	Ar	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000
	H_E	0.659	0.670	0.611	0.663	0.605	0.667	0.517	0.599	0.607	0.494
	F_{IS}	-0.206	-0.119	-0.065	-0.208	-0.199	-0.050	-0.112	-0.044	-0.071	0.190
Microsatellites											
G49											
	Ar	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000	2.995
	H_E	0.306	0.249	0.421	0.287	0.444	0.459	0.305	0.339	0.333	0.318
	F_{IS}	-0.061	-0.103	0.110	0.389	0.324	-0.007	0.017	-0.032	0.154	0.372
Mf13											
	Ar	2.000	1.925	2.000	1.000	2.000	2.000	1.995	2.000	2.000	2.000
	H_E	0.073	0.025	0.073	0.000	0.096	0.119	0.049	0.073	0.145	0.203
	F_{IS}	-0.026	0.000	-0.026	na	-0.040	-0.054	-0.013	-0.026	0.647	0.136
Gafm3											
	Ar	4.920	5.850	4.925	4.850	5.000	3.925	3.925	3.925	3.949	3.000
	H_E	0.642	0.673	0.675	0.661	0.601	0.617	0.661	0.639	0.674	0.671
	F_{IS}	-0.052	-0.002	0.000	0.092	0.126	0.068	-0.098	0.061	-0.103	-0.191
Gafm5											
	Ar	4.000	4.925	5.000	4.925	4.925	5.916	4.000	5.000	4.897	5.850
	H_E	0.684	0.67	0.678	0.687	0.622	0.597	0.692	0.699	0.675	0.717
	F_{IS}	0.175	0.216	0.152	0.236	0.115	0.079	-0.112	0.289	-0.064	0.163
Gafm6											
	Ar	3.925	3.925	3.000	3.000	3.995	3.995	3.995	3.995	3.000	3.925
	H_E	0.567	0.593	0.641	0.542	0.61	0.568	0.603	0.608	0.667	0.684
	F_{IS}	**0.559**	**0.494**	**0.532**	0.309	**0.467**	**0.428**	0.378	**0.424**	**0.514**	**0.489**
Gafm7											
	Ar	6.000	6.991	6.000	6.000	6.000	6.000	5.925	5.995	6.000	5.925
	H_E	0.816	0.822	0.811	0.83	0.822	0.788	0.774	0.781	0.820	0.815
	F_{IS}	0.203	0.179	-0.017	0.127	0.149	0.017	0.289	0.168	0.093	0.110
Gaaf7											
	Ar	3.925	3.000	3.000	3.925	3.000	3.000	3.000	3.925	3.925	3.000
	H_E	0.649	0.664	0.652	0.681	0.639	0.648	0.625	0.680	0.661	0.631
	F_{IS}	-0.041	-0.016	-0.112	-0.027	0.256	0.130	0.040	-0.029	0.206	-0.260
Gaaf9											

	Ar	2.000	2.000	2.000	2.000	2.000	2.000	2.000	2.000	2.000	2.000
	H_E	0.309	0.240	0.308	0.202	0.183	0.281	0.073	0.240	0.182	0.373
	F_{IS}	0.110	-0.147	-0.054	-0.114	0.179	-0.002	-0.026	-0.064	-0.099	-0.031
Gaaf10											
	Ar	5.000	5.000	5.000	4.995	4.000	4.000	4.000	4.920	5.000	4.973
	H_E	0.714	0.724	0.725	0.650	0.673	0.701	0.644	0.668	0.694	0.426
	F_{IS}	**0.440**	**0.689**	**0.414**	0.269	0.406	0.342	0.340	0.401	0.315	-0.050
Gaaf13											
	Ar	5.000	5.925	8.695	8.766	5.920	5.998	4.995	5.925	7.920	6.996
	H_E	0.743	0.746	0.754	0.768	0.755	0.759	0.677	0.774	0.804	0.814
	F_{IS}	-0.077	0.062	0.039	0.024	0.006	0.121	0.114	0.225	-0.089	0.023
Gaaf15											
	Ar	2.000	2.000	2.000	2.000	2.000	2.000	2.000	2.000	2.000	2.000
	H_E	0.222	0.292	0.292	0.326	0.324	0.316	0.240	0.381	0.353	0.359
	F_{IS}	0.326	-0.026	-0.026	0.540	0.075	0.269	0.064	0.212	0.009	-0.143
Average											
	Ar	3.598	3.811	3.817	3.805	3.603	3.603	3.372	3.668	3.745	3.666
	H_E	0.515	0.520	0.542	0.508	0.524	0.536	0.479	0.528	0.538	0.539
	F_{IS}	**0.117**	**0.138**	**0.117**	**0.126**	**0.142**	**0.122**	**0.093**	**0.173**	**0.101**	0.067
	Ne	v.l.	v.l.	v.l.	v.l.	581.8	248.0	179.2	217.6	166.9	338.7

800

801 Table 3. Pairwise genetic differentiation (F_{ST}) between collections. Estimated F_{ST} values incorporated null alleles following Chapuis and Estoup

802 (2007). In bold, values that were significant ($P < 0.05$) after Bonferroni's correction. See Table 1 for location codes.

803

	Locations								
	F6	F5	F4	F3	F2	F1	D1	D2	D3
F5	0.00060								
F4	0.00102	0.00282							
F3	0.00116	0.00047	0.00986						
F2	0.01837	0.00369	0.01297	0.00744					
F1	**0.01716**	0.00228	0.00763	0.00958	0.00005				
D1	0.01752	0.01378	**0.01781**	0.01088	0.01676	0.01735			
D2	0.00441	0.00273	0.00491	0.00671	0.00850	0.01191	**0.02110**		
D3	**0.02259**	0.01155	**0.01293**	**0.01916**	**0.01760**	**0.01659**	0.02598	0.01422	
D4	**0.05012**	**0.04222**	**0.03296**	**0.05737**	**0.05079**	**0.04496**	**0.05782**	**0.02921**	0.01257

804

805 Table 4. Analyses of molecular variance (AMOVA) among study locations. F_{SC}:

806 diversity among locations within analyzed groups; F_{CT}: diversity between analyzed

807 groups; F_{ST}: diversity among locations in the analyzed model. One, two and three

808 asterisks mean significance levels at $P < 0.05$, $P < 0.01$, and $P < 0.001$, respectively.
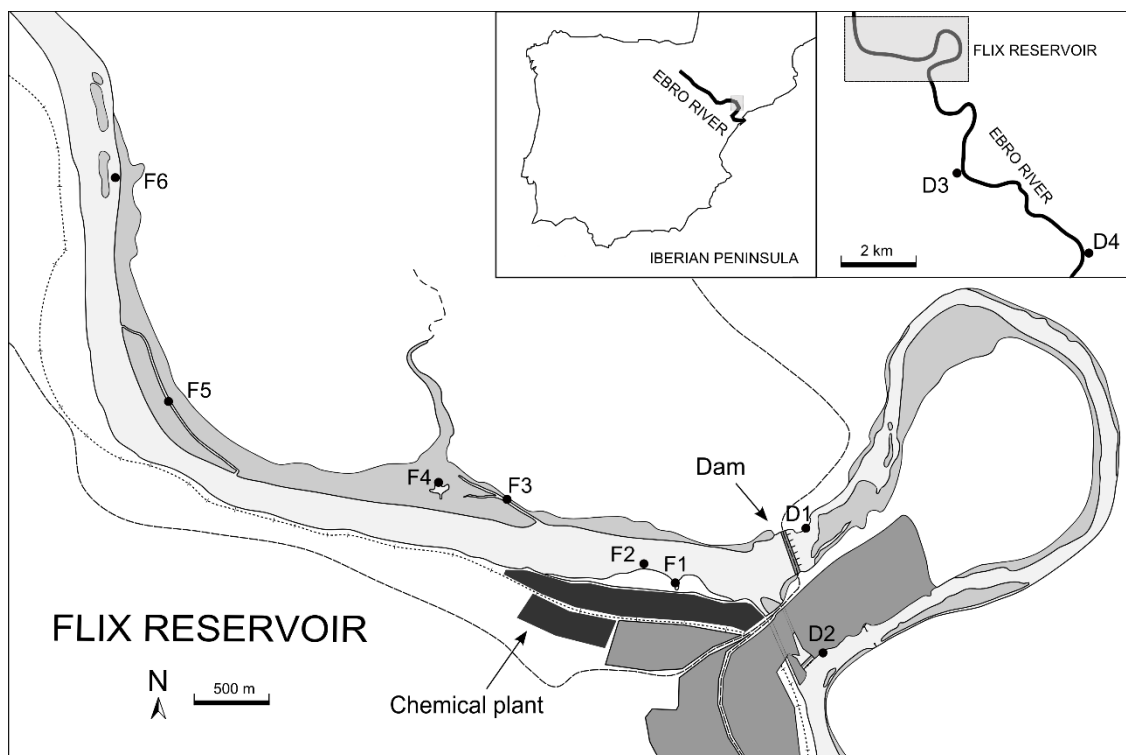
809

Model	Components of diversity		
	F_{SC}	F_{CT}	F_{ST}
Upstream *vs* downstream the dam			
All loci	0.01425***	0.00652*	0.02068***
*GPI-2**	0.00944*	0.03086*	0.04001***
GPI-SNP	0.01492*	-0.00339	0.01157*
Microsatellites	0.01470***	0.00453	0.01916***
Upstream *vs* downstream the dam (D3 and D4 excluded)			
All loci	0.00892***	0.00009	0.00901***
*GPI-2**	0.00890*	0.02748*	0.03613*
GPI-SNP	0.00434	-0.00489	-0.00050
Microsatellites	0.00923***	-0.00269	0.00656***
Polluted *vs* non-polluted			
All loci	0.01596***	0.00491	0.02079***
*GPI-2**	0.02423***	0.00508	0.02918***
GPI-SNP	0.01200	0.00323	0.01519*
Microsatellites	0.01533***	0.00501	0.02026***
Polluted *vs* non-polluted upstream dam (F1 to F6)			
All loci	0.00154	0.00903***	0.01056**
*GPI-2**	0.00838	-0.00086	0.00753*
GPI-SNP	-0.00381	0.02084	0.01711
Microsatellites	0.00112	0.00934***	0.01046*

810   Table 5. Recent migration rates among locations in the study area. In Italics (diagonal):

811   proportion on non-migrant mosquitofish. In bold: migration rates considered as
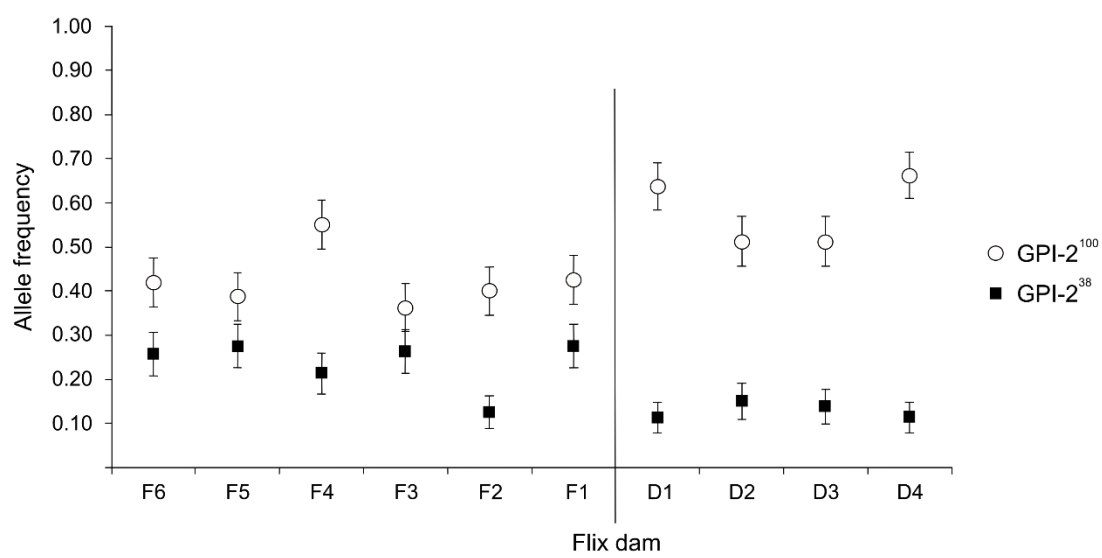
812   biologically relevant (see text).

813

	To:									
From:	F6	F5	F4	F3	F2	F1	D1	D2	D3	D4
F6	*0.678*	0.012	0.012	0.010	0.011	0.014	0.010	0.011	0.012	0.009
F5	0.037	*0.704*	0.039	0.045	0.047	0.055	0.052	0.036	0.033	0.016
F4	0.014	0.009	*0.675*	0.009	0.008	0.008	0.007	0.012	0.009	0.011
F3	0.070	0.081	0.066	*0.732*	**0.089**	**0.096**	**0.096**	0.064	0.055	0.016
F2	0.054	0.055	0.048	0.056	*0.720*	0.056	0.058	0.049	0.032	0.017
F1	0.009	0.011	0.010	0.009	0.010	*0.680*	0.009	0.007	0.011	0.011
D1	0.029	0.036	0.031	0.038	0.041	0.047	*0.710*	0.029	0.028	0.011
D2	**0.092**	0.057	0.065	0.070	0.051	0.021	0.038	*0.743*	0.032	0.033
D3	0.009	0.017	0.025	0.016	0.013	0.012	0.010	0.024	*0.709*	**0.110**
D4	0.009	0.018	0.029	0.015	0.011	0.010	0.010	0.025	0.079	*0.766*

814

Figure 1.

Figure 2.

816

Figure 3.

Figure 5.