Bird migratory flyways influence the phylogeography of the invasive brine shrimp *Artemia franciscana* in its native American range

Since Darwin's time, waterbirds have been considered an important vector for the dispersal of continental aquatic invertebrates. Bird movements have facilitated the worldwide invasion of the American brine shrimp *Artemia franciscana*, transporting cysts (diapausing eggs), and favouring rapid range expansions from introduction sites. Here we address the impact of bird migratory flyways on the population genetic structure and phylogeography of *A. franciscana* in its native range in the Americas. We examined the sequence variation for two mitochondrial gene fragments (COI and 16S for a subset of the data) in a large set of population samples representing the entire native range of *A. franciscana.* Furthermore, we performed Mantel tests and redundancy analyses (RDA) to test the role of flyways, geography and human introductions on the phylogeography and population genetic structure at a continental scale. *A. franciscana* mitochondrial DNA was very diverse, with two main clades, largely corresponding to Pacific and Atlantic populations, mirroring American bird flyways. There was a high degree of regional endemism, with populations subdivided into at least 12 divergent, geographically restricted and largely allopatric mitochondrial lineages, and high levels of population structure (Φ_{ST} of 0.92), indicating low ongoing gene flow. We found evidence of human-mediated introductions in nine out of 39 populations analysed. Once these populations were removed, Mantel tests revealed a strong association between genetic variation and geographic distance (i.e., isolation-by-distance pattern). RDA showed that shared bird flyways explained around 20% of the variance in genetic distance between populations and this was highly significant, once geographic distance was controlled for. The variance explained increased to 30% when the factor human introduction was included in the model. Our findings suggest that bird-mediated transport of brine shrimp propagules does not result in substantial ongoing gene flow; instead, it had a significant historical role on the current species phylogeography, facilitating the colonisation of new aquatic environments as

they become available along their main migratory flyways.

1   Joaquín Muñoz, Francisco Amat, Andy J. Green, Jordi Figuerola, Africa Gómez

3   Joaquín Muñoz (quini@ebd.csic.es), Andy J. Green (ajgreen@ebd.csic.es), Jordi

4   Figuerola (jordi@ebd.csic.es). Department of Wetland Ecology, Estación Biológica de

5   Doñana (CSIC), Isla de la Cartuja, Av. Américo Vespucio, s/n. 41092, Seville, Spain

7   Francisco Amat (amat@iats.csic.es). Instituto de Acuicultura de Torre de la Sal (CSIC),

8   12595, Ribera de Cabanes, Castellón, Spain

10   Africa Gómez (a.gomez@hull.ac.uk). Department of Biological Sciences, University of

11   Hull, Cottingham Rd, Hull, HU6 7RX, U.K. (corresponding author)

Introduction

13 Since Darwin (1859), the role of birds as dispersal vectors for the diapausing propagules

14 of continental aquatic organisms has been recognized (Carlquist 1983; Bilton *et al.*

15 2001; Green & Figuerola 2005). A range of studies have emphasized the importance of

16 dispersal at short and long range by waterbirds for both passively dispersed aquatic

17 invertebrates – through their diapausing eggs or other dispersing stages - and plants –

18 through their seeds – (Figuerola & Green 2002; Green & Figuerola 2005; Green *et al.*

19 2008; Brochet *et al.* 2010a, 2010b; Van Leeuwen *et al.* 2012a, 2012b). Such long-

20 distance, bird-mediated dispersal between aquatic habitats should result in high

21 population gene flow and reduced population or phylogeographic structure. In stark

22 contrast to this prediction, ongoing gene flow between populations has consistently been

23 found to be low (Boileau *et al.* 1992; De Meester *et al.* 2002) and phylogeographic

24 structures quite marked, with high levels of endemism (Gómez *et al.* 2000; Gómez,

25 Montero-Pau, *et al.* 2007; De Gelas & De Meester 2005; Muñoz *et al.* 2008). This

26 paradox has been explained through a combination of high population growth rates,

27 rapid local adaptation and a buffering effect of large egg banks accumulated in

28 sediments, resulting in a monopolisation of resources by the few initial founders,

29 reducing the impact of further immigrants on population structure – what was termed the

30 "monopolisation hypothesis" (De Meester *et al.* 2002). Consistent with this hypothesis,

31 several studies failed to uncover any relationship between the geographic distribution of

32 genetic lineages and bird migration patterns (Gómez, Montero-Pau, *et al.* 2007; Mills *et*

33 *al.* 2007; Muñoz *et al.* 2008). In contrast, the perceived similarity between bird migratory

34 pathways and the distribution of passively dispersed invertebrate genetic lineages

35 suggests that waterfowl are important dispersal vectors (Taylor *et al.* 1998; Freeland *et*

36 *al.* 2000; Hebert *et al.* 2003). In fact, Figuerola et al. (2005) tested explicitly the

37   relationship between bird movements and aquatic invertebrate population genetic

38   structure, revealing a significant association between historical ringing data – used as a

39   proxy of bird-mediated dispersal between populations – and population genetic

40   distances for two cladocerans and a bryozoan in North America, concluding that birds

41   significantly contributed to effective dispersal.

42   Given that continental aquatic invertebrates are unlikely to be in migration-drift

43   equilibrium (Boileau *et al.* 1992; Gómez, Montero-Pau, *et al.* 2007), recent studies have

44   interpreted population isolation-by-distance (IBD) patterns as a signature of historical

45   patterns due to sequential colonisation events, as newly available habitats are more

46   likely to be colonised by nearby populations, with little further impact of gene flow

47   (Gómez, Montero-Pau, *et al.* 2007; Mills *et al.* 2007; Muñoz *et al.* 2008). Therefore,

48   associations between bird movements and genetic distance in aquatic invertebrates

49   based on mitochondrial markers could result from bird-mediated historical colonisation

50   of newly available habitats, instead of ongoing gene flow (Figuerola *et al.* 2005).

51   Shedding light on the role of bird movements on the geographic distribution of genetic

52   lineages would help to understand the structuring of genetic diversity and

53   phylogeography in passively dispersed aquatic invertebrates.

54   *Artemia franciscana* (Kellogg, 1906)(Crustacea: Anostraca), the most widely

55   distributed brine shrimp in America, occurs in hypersaline habitats from Canada to Chile

56   and many Atlantic islands (Hontoria & Amat 1992). It is found in a wide diversity of

57   isolated water bodies, including coastal rock pools and lagoons, inland playas and high

58   mountain salt lakes, permanent prairie salt lakes and commercial salt works (Van

59   Stappen 2002), spanning an extreme range of water chemistry compositions and salinity

60   from high carbonate, or high sulphate athalasic waters to seawater salterns (Bowen *et*

61   *al.* 1988). It is a sexual species, and females produce two types of eggs: subitaneous

62 eggs in benign environmental conditions suitable for population growth, and diapausing

63 eggs (i.e., cysts) during adverse conditions. *Artemia* cysts are amongst the most

64 resistant animal life forms, surviving extreme environmental stresses including UV

65 radiation, desiccation, thermal extremes and anoxia (Clegg 2005). Cysts accumulate at

66 the shoreline and in egg banks in lake sediments (Moscatello & Belmonte 2009), and

67 are readily dispersed by birds, which are the main vectors between catchments. Wind

68 dispersal occurs but over much shorter distances (<1 km, Vanschoenwinkel *et al*. 2009).

69 Many migratory bird species, especially shorebirds, use *Artemia* habitats and adult brine

70 shrimp - often carrying viable cysts - can make up a substantial component of their diet

71 (Anderson 1970; Sánchez *et al.* 2005; Varo *et al.* 2011; Vest & Conover 2011). Birds can

72 disperse cysts between habitats either externally - attached to their feathers or feet - or

73 internally in their digestive tract (Brochet et al. 2010b; Green *et al.* 2005, 2013; Sánchez

74 *et al.* 2007, 2012). Research showing the internal transport of viable *A. franciscana*

75 cysts in the field by the American Avocet, *Recurvirostra americana* (AJG, unpublished

76 data), confirms shorebirds as an effective agent of dispersal in North America (see also

77 Green *et al*. 2005). Recently, Viana *et al*. (2013) estimated the maximum dispersal for

78 *Artemia* cysts via wildfowl as between 230 and 1209 km based on gut passage times of

79 cysts ingested by captive ducks and the distances moved by wild ducks.

80   Populations of *A. franciscana* have substantial levels of genetic (Abreu-Grobois &

81 Beardmore 1982; Gajardo *et al.* 1995; Maniatsi *et al.* 2009) and morphological variation

82 (Hontoria & Amat 1992), and are locally adapted to the ionic composition of their

83 habitats (Bowen *et al.* 1988). Indeed, effective reproductive isolation between some

84 populations is due to different ranges of tolerance to ionic compositions (Bowen *et al.*

85 1988), and so this taxon is regarded by some authors as a "superspecies" (Bowen *et al.*

86 1988). Nevertheless, despite half a century of research for aquaculture and

87   ecotoxicology, comprehensive large-scale phylogeographic surveys of *A. franciscana*

88   are lacking.

89   Cysts from *A. franciscana* – harvested mainly from populations in San Francisco

90   Bay saltworks and the Great Salt Lake in the U.S.A. – have been used globally as a

91   food source in aquaculture and in the pet fish trade for decades (Abatzopoulos 2002;

92   Amat *et al.* 2005, 2007). Effluents from fish farms are likely to contain cysts that can

93   potentially colonise nearby natural wetlands. In addition, the introduction of *A.*

94   *franciscana* has been and still is promoted worldwide to increase salt production or to

95   generate local sources of cysts until as recently as 1993 (Tackaert & Sorgeloos 1993;

96   Sui et al. 2012). As a result of such accidental and intentional inoculations, *A.*

97   *franciscana* has become an invasive species in saline and hypersaline wetlands

98   worldwide (Muñoz & Pacios 2010; J Muñoz, A Gómez, J Figuerola, F Amat, C Rico, AJ

99   Green, 2013 unpublished). For instance, this invasion has led to rapid local extinction of

100   native *Artemia* species in the Mediterranean region (Amat *et al.* 2005). Commercial

101   strains of *A. franciscana* were also introduced in various American sites in the 1970s

102   (Camara 2001; Amat *et al.* 2004). In Brazil, further spreading of the species, probably

103   via bird movements, was noticed within a few years of its introduction in areas where it

104   was previously absent (Camara 2001). However, the impact of these introductions on

105   the genetic diversity and structure of native American populations has yet to be

106   investigated.

107   *Artemia franciscana* represents a very interesting model to test the effect of bird

108   movements on the geographic distribution of genetic lineages and patterns of genetic

109   variation in aquatic invertebrates since (1) its distribution encompasses three

110   continental-scale bird migratory flyways spanning both North and South America (i.e.,

111   the Pacific, Central and Atlantic flyways), but is highly fragmented due to its habitat

112   requirements (hypersaline lakes), (2) its habitats are frequented by migratory shorebirds;

113   *Artemia* is an important prey of these and other waterbirds and its cysts can be readily

114   quantified in waterbird excreta (Green *et al.* 2005; Sánchez *et al.* 2007), and (3) the

115   intentional or accidental inoculations outside the native range may be affecting its

116   natural population genetic structure.

117   Here, we carry out the first comprehensive phylogeographic study of *A.*

118   *franciscana* throughout its known native range (i.e., from Central Canada to southern

119   Chile and Argentina, including the Caribbean islands) using sequence variation for two

120   mitochondrial genes (COI and 16S). Our results indicate a high level of genetic structure

121   and endemism at a continental scale, identify the impact of human introductions and

122   suggest a direct link between bird migratory routes (i.e., flyways) and the historical

123   colonization of *A. franciscana* throughout the Americas, revealing a key role for birds in

124   initial founder events.

125

126   **Methods**

127   *Samples, laboratory procedures, and sequences*

128   We obtained samples from 39 *A. franciscana* populations across its American

129   geographical distribution, from Canada to Chile and Argentina, including Caribbean

130   islands and a population from Cape Verde (Table 1 and Figure 1). Most samples were

131   cysts obtained from the 'cyst-bank' of the Instituto de Acuicultura de Torre de la Sal

132   (CSIC, Castellón, Spain), collected between 1984 and 2000. Four Canadian samples

133   were collected in the field in 2009, two of them (MANW and CHAP) as adults, which

134   were preserved in absolute ethanol until needed. An additional cyst sample from Mono

135   Lake (U.S.A.), collected in the 1970s was kindly provided by the Artemia Reference

136   Centre (ARC 270). A few cyst samples that yielded poor quality DNA extractions were

137 subject to hatching and the resulting nauplii used for DNA extractions (i.e., MexCB and

138 GUA samples).

139   DNA extractions were carried out on individual cysts (previously rinsed in distilled

140 water), whole nauplii or partial adults using a HotSHOT protocol (Montero-Pau *et al.*

141 2008). We used *Artemia*-specific primers 1/2COI_Fol-F and 1/2COI_Fol-R (Muñoz *et al.*

142 2008) to amplify and sequence a 709 bp fragment of the mitochondrial Cytochrome *c*

143 Oxidase Subunit I gene (COI). We also amplified and sequenced a 535 bp fragment of

144 the 16S ribosomal RNA gene for a subset of individuals carrying different COI

145 haplotypes to facilitate comparison with other published sequences using primers

146 16Sar-5'/16Sbr-3' (Palumbi 1996). PCRs were performed in 20 µL total volume

147 containing 1x reaction buffer, 2.0 mM $MgCl_2$, 0.2 mM dNTPs, 0.6 units *Taq* DNA

148 polymerase (Bioline, London, UK) and 0.5 µM of each primer. PCR conditions were as

149 follows: 94ºC for 3 min, followed by 35 cycles of 45 s at 94ºC, 60 s at 45ºC (60-64ºC for

150 16S locus), and 60 s at 72ºC, followed by 5 min at 72 ºC. PCR products were purified for

151 sequencing using ExoSAP-IT® (Exonuclease I and Shrimp Alkaline Phosphatase in

152 buffer; USB Corp., Ohio, USA), cleaned with Sephadex®-G50 (GE Healthcare Corp.),

153 and labelled using the BigDye Terminator Sequencing Ready Reaction v3.1 kit (Applied

154 Biosystems). The resulting fragments were separated on an ABI 3130xl genetic

155 analyzer. Sequences were checked, edited, and aligned using Sequencher™ v4.5 (Gene

156 Codes Corporation, Ann Arbor, MI, USA). All sequences were deposited in GenBank

157 [accession numbers KF662951 - KF663043]. Available published sequences of the

158 same gene fragments, to which we could assign a geographic origin, were also included

159 in our phylogenetic analyses [GenBank: DQ401259–DQ401278, GU248382-GU248387,

160 FJ007820-FJ007834, AF202735-AF202753].

161

162 *Genetic analyses*

163 Neighbour Joining (NJ) and Maximum Likelihood (ML) phylogenetic trees were inferred

164 for both COI and 16S gene fragments. NJ trees were constructed using evolutionary

165 distances computed using the Maximum Composite Likelihood method and 1000

166 bootstrap replicate tests in MEGA5 (Tamura *et al.* 2011). The best-scoring ML trees for

167 COI were estimated using RAxML-VI-HPC v. 7.2.8 (Stamatakis 2006) on the CIPRES

168 portal at the San Diego Supercomputer Center (http://www.phylo.org), optimising free

169 model parameters and executing 1000 rapid bootstraps. Average genetic distances

170 between the main COI lineages – corrected by the K2P + G substitution model – were

171 carried out using MEGA5. Additionally, to identify lineages in our COI dataset, we used

172 the Automatic Barcode Gap Discovery (ABGD) approach (Puillandre *et al*. 2012) using

173 the webtool (http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html).

174 Intra-population gene diversity *Hs* (standardized haplotype diversity) for COI was

175 computed using the program RAREFAC (Petit 1998) to account for population

176 differences in sampling size. Nucleotide diversity and pairwise Φ_{ST} values for COI from

177 all 39 sampled populations were estimated using ARLEQUIN v. 3.1. (Excoffier *et al.*

178 2005).

179

180 *Testing isolation-by-distance patterns and effect of bird flyways*

181 The significance of correlations between pairwise genetic and geographic distances

182 (isolation-by-distance or IBD patterns) was tested using Mantel tests on IBDWS v.3.21

183 (Jensen *et al.* 2005). Prior to analyses in this section, populations inferred to be

184 introduced intentionally by humans (see Table 1) were removed. For all sampled

185 locations, precise decimal longitude and latitude coordinates were obtained using

186 Google Earth (http://earth.google.com). A geographic distance matrix was then

187 computed using Geographic Distance Matrix Generator v.1.2.3 (Ersts, American

188 Museum of Natural History, Center for Biodiversity and Conservation,

189 http://biodiversityinformatics.amnh.org/open_source/gdmg). We used a population

190 geographic distance matrix (Table S1) and a population genetic distance matrix for COI

191 data (Φ_{ST} values using Kimura 2-Parameter as the evolutionary model, Table S2). The

192 99% confidence intervals for the slope and intercept were estimated using Reduced

193 Major Axis (RMA) regression with 30,000 bootstrap randomizations using log km

194 geographical distance.

195   We used canonical Redundancy Analysis (RDA) in CANOCO (ter Braak &

196 Smilauer 2002) to estimate the relative contribution of geographic distances, migratory

197 flyways and human introductions on genetic distance between populations. RDA, a

198 multiple linear regression method widely used in community ecology, has recently been

199 applied to infer the role of spatial versus environmental variables in structuring

200 population genetics data (e.g. Orsini et al 2013). As the dependent matrix we used the

201 sample loadings of a Principal Components Analysis calculated on Φ_{ST} values using

202 Kimura 2-Parameter as the evolutionary model for COI data. Environmental variables

203 were whether the flyway overlapped with the *Artemia* population (0 or 1 depending on

204 the presence of birds from the Atlantic, Central or Pacific flyways in the area) and

205 introduction history (0 or 1). We modelled spatial variables using latitude and longitude

206 (x and y).

207   We used the overlap of sampled populations with the three main migratory

208 flyways in America based on two sources (1) Boere and Stroud (2006) for shorebirds

209 and (2) Birdlife International (extracted from

210   http://www.birdlife.org/datazone/userfiles/file/sowb/flyways/). These data were used as a

211   proxy for bird movements between locations. This approach is likely to be a rough

212   approximation to the probability of bird migration between locations, given that more

213   precise estimates cannot be obtained due to the absence of sufficient shorebird ringing

214   data for the whole of the Americas. Even if extensive ringing data was available, this

215   would only estimate current bird movements, whereas climate changes since the

216   Pleistocene are likely to have strongly affected bird movements (Alerstam 1993).

217   RDA analyses were carried out using each set of migratory flyway data (i.e.,

218   Boere and Stroud (2006), and Birdlife International information) and the variance

219   partitioning calculated according to Borcard et al. (1992) when the model was

220   significant. All environmental variables contributed to the full model, so we constructed

221   two additional RDAs considering only flyway or introduction history as environmental

222   variables (data files used in RDA are deposited in Dryad doi:xxxxx).

223

224   **Results**

225   *Phylogenetic relationships and geographic distribution of lineages*

226   Once PCR primers were removed and sequences trimmed to the same length, the 604

227   bp COI alignment contained 603 individuals newly sequenced in this study (see Table 1),

228   which collapsed into 93 haplotypes. No ambiguities, insertions, deletions or stop codons

229   were present in the alignment. There were 121 variable sites, 86 of them parsimony

230   informative, and 104 synonymous and 15 non-synonymous substitutions.

231   Using the default parameters, ABGD did not find any partitions in our dataset, so we

232   reduced X, (i.e., the minimum barcode gap width) as suggested by the program, to 1.0

233   and used Kimura 80, identifying 12 groups with a prior intraspecific distance of = 0.0028

234   (we use the term 'lineages' thereafter, see below for their geographic distribution). When

235 the prior intraspecific distance increases to 0.0046 (not high by any standards) then the

236 number of partitions reduces again to 1.

237 Both phylogenetic approaches ML and NJ were highly consistent, with two highly

238 supported main branches displaying a geographic distribution along the continent, one

239 mainly Atlantic (lineages 9-12) and the other split between two sub-branches along the

240 Pacific Rim (lineages 1-7) and in Central Canada (lineage 8). Overall, there were at

241 least 12 mtDNA lineages, most of them well supported (Figure 2). With the exception of

242 lineage 1, each of the lineages showed a restricted geographic distribution indicating a

243 high level of regional endemism (Figure 3 for the geographic distribution of the

244 lineages). Lineage 2 was found in a single coastal site in NE Mexico, Lineage 3 in the

245 five locations in Peru, Lineage 4 in two high altitude populations from Central Mexico,

246 lineages 5 and 7 in Central Chile, lineage 6 in a single sulphate-rich location in NE

247 Mexico, lineage 8 in four locations in Saskatchewan, Canada, lineage 9 in Puerto Rico,

248 lineage 10 in Cape Verde, lineage 11 in several locations around the Caribbean, Baja

249 California and SW Mexico, and lineage 12 in Argentina and Chile. In stark contrast to the

250 rest, lineage 1 was genetically diverse and geographically widespread, found across a

251 large geographical area across both sides of the continent, including Brazil, Chile, Cuba,

252 Jamaica, Mexico, USA and Colombia. Out of the 27 haplotypes in lineage 1, seven

253 haplotypes were detected in SFB and GSL, the two commercialised populations in the

254 USA that have been sources for the invasion in the Mediterranean (J Muñoz, A Gómez,

255 J Figuerola, F Amat, C Rico, AJ Green, 2013 unpublished)(see section *Introduced*

256 *populations* below). In addition, lineage 1 contained six closely related haplotypes only

257 found in Mono lake – which harbours a population often considered as a separate

258 species, *A. monica* – and haplotypes found in three Colombian populations (CSC, CT

259 and CGZ), five Mexican populations, and the Jamaican population. Two other lineages

260   were distributed across the continental E-W divide, creating contact zones between

261   lineages. In lineage 12, most of the haplotypes were found in Argentinean populations,

262   but four of these haplotypes (three of them shared across populations) were found in a

263   Chilean Altiplano population (Salar de Llamara). Lineage 11, mainly Caribbean, has a

264   few haplotypes in two Mexican populations from the Pacific side (Las Coloradas and

265   Faro San José, where they coexist with lineage 1). Finally, although both lineages 1 and

266   11 are found in Colombia, they were not found together in any of the populations

267   sampled. The genetic divergence between the 12 main lineages ranged from 1.8%

268   (between lineages 1 and 4) to 6.0% (between lineages 2 and 9) (Table 2).

269   The 16S alignment contained 408 bp from 122 individuals, which collapsed into

270   59 haplotypes. There were two singleton indels, 63 variable sites and 43 parsimony

271   informative sites. In contrast to the COI analysis, the NJ and ML reconstructions were

272   poorly resolved, especially the basal branches, but most of the lineages recovered by

273   the COI analyses were also recovered for the 16S data, with variable levels of support

274   (see Figure 4). COI Lineages 3, 5, 7 and 9 were highly supported for both ML and NJ

275   analyses in the 16S analysis, whereas lineages 1 and 2 on the one hand and lineages 4

276   and 6 on the other, collapsed into poorly supported branches. The 16S analyses allowed

277   us to assign several previously sequenced populations, which we were unable to

278   sample, to COI lineages, particularly in NW America and the Caribbean. In addition, the

279   16S analysis revealed the presence of two new lineages in the Caribbean, one in the

280   Virgin Islands, related to lineage 9 (Puerto Rico) and another in Inague Island

281   (Bahamas) related to Lineage 12 (Argentina, Chile). Note that these maintain the

282   relationship with Atlantic lineages. Regarding North-Western American populations -

283   extensively sampled by Prof Sarane Bowen's group - in New Mexico, Nevada and

284   British Columbia, they hold private haplotypes which appear in a poorly supported

285   branch with Mono Lake and Mexican haplotypes. Other populations (Clinton, Basque

286   Lake, Baja California and Carrizo Soda lake) also appear in the composite lineage 1 and

287   2, underscoring the diversity of U.S.A. *A. franciscana* populations. Interestingly, 16S

288   haplotypes from Jesse Lake (Nebraska) belong to the Central Canadian lineage (lineage

289   8) together with Little Manitou, Muskiki, Meacham and Chaplin haplotypes.

290

291   *Intra- and inter-population genetic diversity in* COI

292   The number of individuals sequenced per population ranged between 4 and 37,

293   depending on available material (average 15.97; see Table 1 for estimates of π and H,

294   and details of the haplotypes present in each population). The number of haplotypes per

295   population ranged from one to seven. The highest standardized gene diversity (Hs) was

296   found in the Mexican population MexH, whereas five populations (MexCe and Mex99

297   from Mexico, BRM from Brazil, PV from Peru, AMC, from Argentina, and CHLC from

298   Chile) were fixed for a single haplotype. Most haplotypes were found within single

299   countries, except for several haplotypes shared between some countries and the two

300   commercial U.S.A. populations SFB and GSL (see Table 1).

301   Populations were strongly structured genetically (global Φ_{ST} = 0.92; 0.94 when

302   putative introduced populations were removed), with Φ_{ST} being highly significant

303   between all populations except for three lakes from Central Canada, plus one pair from

304   Chile/Argentina (see Table S2).

305

306   *Identification of introduced populations*

307   We found genetic evidence for putative non-native populations originating from the

308   commercialised SFB or GSL in nine sites from Mexico, Cuba, Jamaica, Brazil and Chile.

309   In these populations, at least one sampled individual shared a haplotype with SFB

310   and/or GSL populations (see Table 1). Those populations showed three patterns: (1) all

311   individuals sampled shared haplotypes with SFB and/or GSL (BRM, GUA, and CHLC);

312   (2) populations had haplotypes shared with the commercialised populations and a

313   further haplotype (Af19 for MEXLC and JAYA) which differs from Af18 (a common

314   haplotype in SFB and GSL) by a single substitution; and (3) populations sharing some

315   haplotypes with SFB and/or GSL, but which also had unrelated additional/private

316   haplotypes (MexT, MexSQ, MexFSJ, and CHPI). For four of the nine putative introduced

317   populations, the occurrence of intentional introductions had been previously reported

318   either in the same or nearby sites (see references in Table 1). Note that introduced *A.*

319   *franciscana* populations are likely to further expand into nearby suitable habitats due to

320   passive dispersal by birds.

321

322   *Isolation-by-distance pattern and the role of American migratory flyways*

323   Mantel tests on pairwise genetic and geographic distances for populations ranging from

324   Chile-Argentina to Canada, excluding those inferred to be introduced, revealed a strong

325   IBD pattern (Figure 5) with a highly significant correlation between pairwise geographic

326   and genetic distances, indicating that geographic distance between populations explains

327   a large proportion of the genetic variability in the sampled area (R^2-value= 0.323, *p*-

328   value <0.001).

329   RDA showed that both flyway and introduction history were significantly

330   associated to population genetic distance (p<0.02 for all the correlations with genetic

331   distance, whether or not geographical distance was controlled for). The effect of using

332   different flyway data was quite small. Flyway explained 18.7-21.2% of variance in

333   population genetic distance, depending on the flyway dataset used (Table 3). Genetic

334 distance was also affected by historical anthropogenic introductions (15.6% of variance

335 explained) and both factors together (i.e., flyway and anthropogenic introduction)

336 explained 30.6-31.2% of variance in genetic distance. In comparison, geographic

337 distance only explained 8.2-8.7% of variance.

338

339 **Discussion**

340 Our analyses revealed that *A. franciscana* has a strong regional genetic structure in its

341 native distribution range throughout the Americas, with twelve largely allopatric endemic

342 lineages. Such high level of population structure, supported by a very high overall Φ_{ST}

343 value, high number of private haplotypes and significant IBD patterns, indicate that the

344 populations studied are not connected by high ongoing gene flow, pointing instead to the

345 effects of genetic drift and persistent founder effects during historical colonisation

346 processes and development of local adaptation (i.e., the Monopolisation hypothesis; De

347 Meester et al. 2002). The few population pairs with non-significant population

348 differentiation (three lakes from Central Canada, plus one pair from Chile/Argentina) are

349 likely to reflect recent colonisation, rather than ongoing gene flow. Our data also reveals

350 the impact of *A. franciscana* introductions on the phylogeography of the species, as the

351 lineage including the commercialised SFB and GSL populations has now achieved the

352 widest distribution across the continent, in some cases coexisting – and presumably

353 hybridizing – with pre-existing native populations. In addition, our results suggest that

354 migratory birds have an important role in the colonisation of new habitats and are

355 associated with range expansions both in the history of *Artemia* colonisation across the

356 Americas, and also at a local scale, where birds facilitate the expansion of introduced

357 lineages.

358

359 *The role of bird migratory flyways*

360 Our study provides new evidence supporting the key historical role of waterbirds as the

361 main factor shaping the population genetic structure of continental aquatic invertebrates

362 at an intra-continental scale. The patterning of the main phylogenetic lineages, with an

363 Atlantic, Central and Pacific distribution – instead of a North American (Nearctic) *vs*.

364 South American (Neotropical) division reflecting the recognized zoogeographic regions

365 and the long isolation of the continents (Lomolino *et al.* 2010; Holt et al. 2013) - strongly

366 suggests that historical bird migratory flyways, which occur alongside both the major

367 coasts of this continent, determined the historical spread of *A. franciscana* genetic

368 lineages. Bird movements also might have allowed the subsequent persistence of this

369 structure by facilitating colonisation along each migratory flyway, which shaped the main

370 East-West division in mitochondrial lineages. RDA showed that the effect of migratory

371 flyways was highly significant, and accounted for 20% of the genetic variation between

372 populations once geographic distance was taken into account, suggesting that the

373 distribution of genetic lineages in *A. franciscana* is likely to reflect the impact of historical

374 bird flyways on native phylogeographic patterns. In addition, the strong detected IBD

375 pattern suggests that the chances of bird-mediated colonisation are highly distance

376 dependent (see Viana *et al*. 2013), although instances of long-distance dispersal and

377 colonisation, for example from Argentina to Chile, or to Colombia from northern

378 Caribbean populations, are also apparent from our data. A corollary of our results is that

379 bird movements must have shown some stability, forming parallel N-S flyways during the

380 time frame of *A. franciscana* population diversification (i.e., throughout the Pleistocene)

381 extending into new breeding areas becoming available in the north and new wintering

382 sites in the southern extreme of the continent (Buehler *et al.* 2006). These results are no

383   surprise given the high transport rates of *Artemia* cysts by waterbirds (Green *et al*. 2005;

384   Sánchez *et al*. 2007) and the lack of other dispersal vectors. Migratory waterbirds have

385   existed since the Early Cretaceous (Lockley *et al*. 2012). Although now-extinct migratory

386   mammals were once major vectors of plant dispersal in the Americas (Janzen 1984), the

387   hypersaline habitats used by *Artemia* are not conducive to dispersal by large mammals.

388   Our results for *A. franciscana* agree with other studies in N America pointing to an

389   effect of bird movements on the genetic structure of passively dispersed aquatic

390   invertebrates (Figuerola *et al.* 2005), but extends these findings to the whole of the

391   Americas, and emphasizes the major role of bird movements in facilitating colonisation

392   into new suitable habitats – in agreement with the patterns found in *A. franciscana* in the

393   invaded Mediterranean range (J Muñoz, A Gómez, J Figuerola, F Amat, C Rico, AJ

394   Green, 2013 unpublished), and those for *A. salina* in its native range (Muñoz *et al.*

395   2008). Although at a continental scale our results suggest that bird movements do not

396   promote contemporary gene flow between *Artemia* populations, such gene flow may still

397   have a role at a local scale or when new areas suitable for colonization become

398   available.

399

400   *Phylogeographic patterns*

401   The high level of endemism and population structure in native *A. franciscana*

402   populations, with low ongoing gene flow and occasional long-distance migration

403   resembles the patterns found in *Artemia salina* (Muñoz *et al.* 2008), a sexual native

404   Mediterranean brine shrimp species, and in a range of other aquatic passively dispersed

405   taxa including sexual reproduction in their life cycles (Gómez *et al.* 2000; Edmands

406   2001; De Gelas & De Meester 2005; Mills *et al.* 2007; Ketmaier *et al.* 2008; Korn *et al.*

407   2010). Our results also expand and confirm the deep phylogenetic breaks found by

408 Maniatsi *et al*. (2009) in a mtDNA and nDNA study of a more limited number of

409 populations revealing only three lineages. As populations of these passively dispersed

410 organisms can be founded by a small number of propagules, followed by rapid

411 population growth and establishment of large diapausing egg banks, this favours the

412 presence of persistent/long term founder effects, thus reducing the effect of gene flow,

413 possibly reinforced by the development of local adaptation, according to the

414 Monopolisation hypothesis (De Meester *et al.* 2002). Cysts are undoubtedly still

415 regularly dispersed between suitable habitats by waterbirds, but they are unlikely to

416 become established owing to the Monopolisation effects.

417   Given the range of genetic divergence between lineages (from 2 to 6%) the time

418 frame of their fragmentation can be approximated roughly using a COI molecular clock

419 for other shrimp taxa (1.4% sequence divergence per million year; Knowlton & Weigt

420 1998), which translates into 1.4-4.3 million years of divergence (reaching the Pliocene),

421 between *A. franciscana* lineages. Even a faster rate of 2% per million years will result on

422 pre-Pleistocene divergence times between the main lineages. Therefore, a contribution

423 of Pliocene/Pleistocene climatic oscillations to population fragmentations after range

424 expansions across the continent can be inferred from our data, possibly allowing

425 survival of lineages in separate geographical areas including Caribbean islands and

426 areas in N and S America. Mexico has the highest lineage richness, with five out of the

427 12 COI lineages being native to this country - including lineages from both Pacific and

428 Atlantic clades. These findings suggest that this region is likely to have supported

429 separate refugia during climatically adverse periods. The occurrence of a highly

430 divergent Central Canadian prairie lineage was unexpected, as an ice sheet covered

431 this area during the last glacial maximum (Ehlers & Gibbard 2004). However, the 16S

432 data from GenBank suggest that this lineage also occurs in more southern areas in

433 central USA (Nebraska), where it may have survived south of the ice sheets, and then

434 undergone postglacial colonisation of Central Canada.

435   Following climate-driven turnover of hypersaline habitats, migratory shorebirds

436 would be involved in expanding the lineages into newly appearing suitable habitats and

437 the chances of successful spread would be strongly distance dependent. However, long-

438 distance colonisation events must have also occurred. For example, assuming that the

439 ancestor of lineages 9 and 10 inhabited the Caribbean area, the colonisation of

440 Argentina by the ancestor of lineage 12 must have entailed successful transfer of some

441 cysts between these distant areas (see Figure 2). The genetic composition of lineage

442 12, where most of the haplotypes were found in Argentinean populations (see also

443 Figure 3) but four of them (three shared across populations) were found in a Chilean

444 Altiplano population (Salar de Llamara), suggests recent colonisation from Argentina to

445 Chile. Finally, the colonisation of Cape Verde islands, with private haplotypes distantly

446 related to Caribbean lineages, must have involved long-distance transport, possibly from

447 birds accidentally landing there after storms sent them off course, although we cannot

448 rule out the possibility of an unreported human-mediated introduction involving cysts

449 from a Caribbean population not included in our study. Shorebirds are likely vectors for

450 such long-distance dispersal events, and have often been implicated in the dispersal of

451 plants between North and South America, or to oceanic islands (Cruden 1966; Proctor

452 1968).

453   The 16S analysis also shows higher richness of lineage 12 in Argentina with two

454 Chilean populations (Convento and Salar de Llamara), for which evidence of nuclear

455 DNA (nDNA) introgression among lineages exists (Maniatsi *et al.* 2009). Furthermore, as

456 migratory flyways overlap on some areas, this could have occasionally resulted in

457 transfers from the Atlantic to the Pacific coasts, as has been suggested for other

458 passively dispersed aquatic species (Miura *et al.* 2011).

459   Natural spread of lineages from refugial areas is likely to result in contact zones

460 between lineages, which we expect to be sharp, as we found between lineages 1 and 11

461 in Colombia, where despite two lineages being present in the area, there are no sites

462 where both co-occur.

463

464 *Taxonomic considerations*

465 The COI gene is one of the most widely used tools for species delineation (Hebert,

466 Penton, *et al.* 2004; Costa *et al.* 2007; Gómez, Wright, *et al.* 2007). Sequence

467 divergence of 3% have been proposed as a threshold for species delimitation in

468 crustaceans (Costa *et al.* 2007, but see Lefébure *et al.* 2006), but other approaches are

469 also used, such as GMYC (Pons *et al.* 2009) or automatic barcode delimitation, which

470 we used here (Puillandre *et al.* 2012). Our analysis revealed 12 lineages in *A.*

471 *franciscana*, some of them, like lineages 9 (Puerto Rico) plus 10 (Cape Verde)

472 compared to all the others, or lineages 11 (Circum Caribbean) plus 12 (Argentina-Chile)

473 compared to all the others, have genetic divergences of over 5%. Surprisingly,

474 reproductive isolation – due to ecological specialisation and local adaptation – has only

475 been reported between Mono Lake and San Francisco Bay populations (Bowen *et al.*

476 1988) due to the inability of individuals of each of these populations to survive in each

477 others' ecological conditions, but our data show they are very closely related. Indeed

478 cross-fertility has been observed in the laboratory between the San Francisco Bay

479 population and 15 other populations from the whole range of the species, including

480 some populations included in our study that belong to very divergent mtDNA lineages

481 such as Inague Saltern (Bahamas), Little Manitou (Canada, lineage 8), and Puerto Rico

482   (lineage 9) (Clark & Bowen 1976). Therefore, we concur with Bowen *et al.* (1988) in

483   regarding *A. franciscana* as a very diverse "superspecies", where reproductive isolation

484   mediated by habitat adaptation might occur, but populations in intermediate habitats

485   could act as venues for genetic exchange between ecological isolates.

486

487   *Impact of introductions on native population structure and management implications*

488   Given the strong phylogeographic structure of *A. franciscana*, and the high level of

489   private haplotypes found, we used haplotype sharing between the commercialised

490   populations (SFB and GSL) and distant populations as a criterion of recent human

491   mediated introduction. Using this criterion, we identified nine populations where genetic

492   evidence pointed to putative human introductions into Mexico, Cuba, Jamaica, Brazil

493   and Chile. For four of these populations, the occurrence of intentional introductions in

494   the same or nearby sites during the 1970s could be documented. Therefore, our genetic

495   data confirms that the established *A. franciscana* populations in these locations are, at

496   least partially, of introduced origin, and validates our criterion. As for the impact of

497   introductions on native populations, four Mexican populations contained additional

498   private haplotypes not closely related to the introduced ones, which suggests the

499   presence of a pre-existing population before the introduction and the likely introgression

500   of both populations with persistence of native haplotypes, although nuclear loci would be

501   necessary to confirm this. In two populations, we found a discrepancy between our

502   genetic results and previously published ones. The first case is the MEXLC population,

503   which we regard as introduced (lineage 1), whereas in Tizol-Correa *et al*. (2009) the

504   haplotype found is shared with populations from our lineage 11. The second case is

505   CHLC (Laguna Cejas, Chile), which we regard as introduced whereas Maniatsi *et al*.

506   (2009) found that their mtDNA was most closely related to native populations from

507 Central Chile (which belong to our lineage 7). While these authors concluded that the

508 discrepancies between their nDNA and mtDNA data from the latter population pointed to

509 incomplete lineage sorting, an alternative explanation is that they are due to the

510 population being admixed with native and introduced ancestry, as introduced *A.*

511 *franciscana* is known to hybridize even with the genetically divergent *A. persimilis*

512 (Kappas *et al.* 2009). This failure to detect introduced mtDNA lineages in this population

513 might arise from reduced sample sizes, or indicate real temporal changes in these

514 populations reflecting recent introductions, as samples are likely to have been obtained

515 in different years (unfortunately, no collection dates for these population is reported in

516 Maniatsi *et al.* 2009). Despite human introductions, haplotypes presumably from pre-

517 existent populations have survived and coexist with introduced ones, although a

518 possible loss of genetic diversity due to introductions cannot be ruled out, and should be

519 investigated in the future, perhaps using sediment cores (a method widely used in the

520 crustacean model organism *Daphnia*) from populations where commercialised strains

521 were introduced.

522 The high genetic richness found in our COI analyses, and the presence of private

523 haplotypes belonging to lineage 1 in distant populations away from SFB and GSL,

524 including three Colombian and five Mexican populations likely to be native, indicates that

525 the natural distribution of lineage 1 extended further than the two commercialized

526 populations before any human introductions took place. This was also in evidence from

527 our 16S analyses, which included more extended geographic coverage of NW America,

528 and revealed further sites within lineage 1 in the W USA and British Columbia

529 harbouring private haplotypes not found in SFB or GSL. This is also consistent with the

530 presence of private and closely related haplotypes at Mono Lake. The peculiarity and

531 fragmentation of the habitats used by the species, the potential of salinity and varying

532 ionic composition to act as a strong selective agent, the capacity to produce massive

533 quantities of resting eggs that can be readily dispersed by birds, combined with the

534 apparent limitation in modern gene flow, makes this group an ideal system for further

535 studies testing the role of local adaptation and mass effects on reducing gene flow

536 between populations.

537   Given the impact of the invasive *A. franciscana* across the world, and the high

538 genetic and ecological richness of its native populations, further population

539 translocations should be highly discouraged, and the use of native strains as a source of

540 cysts should be encouraged even within the Americas.

541

542 **Conclusions**

543 Our analyses suggest that *A. franciscana* phylogeography in its native range was

544 shaped by (1) Pliocene/Pleistocene climate fluctuations, which contributed to changes in

545 the areas available to the species, (2) historical bird-mediated colonization along

546 migratory flyways, which shaped the East-West population division, (3) strong and

547 persistent founder events, facilitated by high population growth rates and large

548 population sizes, preventing further gene flow despite ongoing bird-mediated dispersal,

549 and (4) human introductions coupled with regional bird dispersal, explaining the large

550 but localised geographic range of the lineages derived from the commercially exploited

551 North American populations. Our findings suggest that, at a continental scale, bird-

552 mediated transport of invertebrate propagules does not result in substantial ongoing

553 gene flow, but instead determines species phylogeography, facilitating the colonisation

554 of newly available aquatic environments along bird flyways.

555

Acknowledgements

We thank G. Van Stappen (Artemia Reference Centre) for providing the Mono Lake cyst

sample and Joachim Mergeay, who read a previous version of this manuscript, and

provided many constructive comments including suggesting using RDA. M.T. Bidwell

helped with sampling in Canada.

References

Abatzopoulos TJ (2002) *Artemia: basic and applied biology*.

Abreu-Grobois FA, Beardmore JA (1982) Genetic differentiation and speciation in the
     brine shrimp *Artemia*. In: *Mechanisms of speciation* (ed. Barigozzi C), pp. 345–376.
     Alan R. Liss, New York.

Alerstam T (1993) *Bird Migration*. Cambridge University Press.

Amat F, Cohen RG, Hontoria F, Navarro C (2004) Further evidence and characterization
     of *Artemia franciscana* (Kellogg, 1906) populations in Argentina. *Journal of
     Biogeography*, 1735–1749.

Amat F, Hontoria F, Navarro JC, Vieira N, Mura G (2007) Biodiversity loss in the genus
     Artemia in the Western Mediterranean Region. *Limnetica*, **26**, 387.

Amat F, Hontoria F, Ruiz O Sanchez MI, Figuerola J, & Hortas F (2005) The American
     brine shrimp as an exotic invasive species in the western Mediterranean.
     Biological Invasions, **7**, 37–47.

Anderson W (1970) A preliminary study of the relationship of salt ponds and wildlife-
     South San Francisco Bay. *Calif Fish Game 56, 240-252. ILLUS. 1970*, **56**.

Bilton DT, Freeland JR, Okamura B (2001) Dispersal in freshwater invertebrates. *Annual
     Review of Ecology and Systematics*, **32**, 159–181.

Boere GC, Stroud DA (2006) The flyway concept: what it is and what it isn't. In:
     Waterbirds around the world. (ed G.C. Boere, C.A. Galbraith & D.A. Stroud), pp.
     40–47. The Stationery Office, Edinburgh, UK.

Boileau MG, Hebert PDN, Schwartz SS (1992) Non-equilibrium gene frequency
     divergence : founder effects in natural populations. *Journal of Evolutionary Biology*,
     5, 25–39.

Borcard D, Legendre P, Drapeau P (1992) Partialling out the spatial component of
     ecological variation. *Ecology*, **73**, 1045-1055.

Bowen ST, Buoncristiani MR, Carl JR (1988) *Artemia* habitats: Ion concentrations
     tolerated by one superspecies. *Hydrobiologia*, **158**, 201–214.

Brochet, AL, M. Gauthier-Clerc, M. Guillemain, H. Fritz, A. Waterkeyn, A. Baltanas, and
     A. J. Green (2010b). Field evidence of dispersal of branchiopods, ostracods and
     bryozoans by teal (*Anas crecca*) in the Camargue (southern France).
     Hydrobiologia **637**:255-261.

Brochet, A. L., M. Guillemain, H. Fritz, M. Gauthier-Clerc, and A. J. Green. 2010a. Plant
     dispersal by teal (*Anas crecca*) in the Camargue: duck guts are more important
     than their feet. Freshwater Biology **55**:1262-1273.

598 Buehler DM, Baker AJ, Piersma T (2006) Reconstructing palaeoflyways of the late
599   Pleistocene and early Holocene Red Knot *Calidris canutus*. *Ardea*, **94**, 485–498.
600 Camara M (2001) Dispersal of *Artemia franciscana* Kellogg (Crustacea; Anostraca)
601   populations in the coastal saltworks of Rio Grande do Norte, northeastern Brazil.
602   *Hydrobiologia*, **466**, 145–148.
603 Carlquist S (1983) Intercontinental dispersal. *Sonderbd. naturwiss. ver Hamburg*, **7**, 37–
604   47.
605 Castro B (1993) Biologia y cultivo de *Artemia franciscana* en el Ex Lago de Texcoco, de
606   Ecatepec, Estado de México. Universidad Nacional Autónoma de México.
607 Castro BT, M.J. Castro, V.R.E. Marín, G.R. Young, D. Jenoure, M.G. Castro, S.A.
608   Malpica, & A.R. De Lara (2000) Calidad de quistes y morfometria de la poblacion
609   de *Artemia franciscana* (Kellogg, 1906) de la laguna pequenya de Yallahs,
610   Jamaica. *Ciencias Marinas*, **26**, 201–214.
611 Castro J, T Castro, J Sánchez, G Castro, A Castro, J Zaragoza, R de Lara, and M. del C
612   Monroy (2006) Cysts and nauplii biometry characteristics of seven *Artemia*
613   *franciscana* (Kellog, 1906) populations from Mexico. *Revista de biología marina y*
614   *oceanografía*, **41**, 187 – 193.
615 Clark LS, Bowen ST (1976) The genetics of *Artemia salina*. VII. Reproductive isolation.
616   *The Journal of heredity*, **67**, 385–8.
617 Clegg JS (2005) Desiccation tolerance in encysted embryos of the animal extremophile,
618   *Artemia*. *Integrative and comparative biology*, **45**, 715–24.
619 Costa FO, deWaard, J R, Boutillier, J, Ratnasingham, S, Dooh, RT, Hajibabaei, M &
620   Hebert, PD (2007) Biological identifications through DNA barcodes: the case of the
621   Crustacea. *Canadian Journal of Fisheries and Aquatic Sciences*, **64**, 272–295.
622 Cruden, RW (1966) Birds as Agents of Long-Distance Dispersal for Disjunct Plant
623   Groups of Temperate Western Hemisphere. Evolution 20:517-532.
624 Darwin CR (1859) On the origin of species by means of natural selection. John Murray,
625   London.
626 Edmands S (2001) Phylogeography of the intertidal copepod *Tigriopus californicus*
627   reveals substantially reduced population differentiation at northern latitudes.
628   *Molecular ecology*, **10**, 1743–50.
629 Ehlers J, Gibbard PL (2004) *Quaternary glaciations: extent and chronology, Part 2*.
630   Elsevier.
631 Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software
632   package for population genetics data analysis. *Evolutionary Bioinformatics Online*,
633   **1**, 47.
634 Figuerola J, Green AJ (2002) Dispersal of aquatic organisms by waterbirds: a review of
635   past research and priorities for future studies. *Freshwater Biology*, **47**, 483–494.
636 Figuerola J, Green AJ, Michot TC (2005) Invertebrate eggs can fly: evidence of
637   waterfowl-mediated gene flow in aquatic invertebrates. *The American Naturalist*,
638   **165**, 274–80.
639 Freeland JR, Romualdi C, Okamura B (2000) Gene flow and genetic diversity: a
640   comparison of freshwater bryozoan populations in Europe and North America.
641   *Heredity*, **85 Pt 5**, 498–508.
642 Frisch D, Green AJ, Figuerola J (2007) High dispersal capacity of a broad spectrum of
643   aquatic invertebrates via waterbirds. *Aquatic Sciences*, **69**, 568–574.
644 Gajardo GGM, Colihueque N, Parraguéz M, Sorgeloos P (1998) International Study on
645   *Artemia* LVIII. Morphologic differentiation and reproductive isolation of *Artemia*

populations from South America. *International Journal of Salt Lake Research*, **7**, 133–151.

Gajardo GGM, Conceicao M, Weber L, Beardmore JA (1995) Genetic variability and interpopulational differentiation of *Artemia* strains from South America. *Hydrobiologia*, **302**, 21–29.

Gelabert F, Solis L (1994) La selección del tamaño de particulas alimenticias por *Artemia* de Guantánamo, Cuba. *Rev. Invest. Mar*, **15**, 141–144.

De Gelas K, De Meester L (2005) Phylogeography of *Daphnia magna* in Europe. *Molecular ecology*, **14**, 753–64.

Gómez A, Carvalho GR, Lunt DH (2000) Phylogeography and regional endemism of a passively dispersing zooplankter: mitochondrial DNA variation in rotifer resting egg banks. *Proceedings of the Royal Society B, Biological sciences*, **267**, 2189–97.

Gómez A, Montero-Pau J, Lunt DH, Serra M, Campillo S (2007) Persistent genetic signatures of colonization in *Brachionus manjavacas* rotifers in the Iberian Peninsula. *Molecular Ecology*, **16**, 3228–3240.

Gómez A, Wright PJ, Lunt DH, Cancino JM, Carvalho GR, & Hughes, RN (2007) Mating trials validate the use of DNA barcoding to reveal cryptic speciation of a marine bryozoan taxon. *Proceedings of the Royal Society B, Biological sciences*, **274**, 199–207.

Green AJ, Figuerola J (2005) Recent advances in the study of long-distance dispersal of aquatic invertebrates via birds. *Diversity and Distributions*, **11**, 149–156.

Green AJ, Sánchez MI, Amat F, Figuerola J, Hontoria F, Ruiz O, & Hortas, F (2005) Dispersal of invasive and native brine shrimps *Artemia* (Anostraca) via waterbirds. *Limnology and Oceanography*, **50**, 737–742.

Green, A. J., D. Frisch, T. C. Michot, L. K. Allain, and W. C. Barrow. 2013. Endozoochory of seeds and invertebrates by migratory waterbirds in Oklahoma, USA. Limnetica **32**:39-46.

Green, A. J., K. M. Jenkins, D. Bell, P. J. Morris, and R. T. Kingsford. 2008. The potential role of waterbirds in dispersing invertebrates and plants in arid Australia. Freshwater Biology **53**:380-392.

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 14812–7.

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004) Identification of birds through DNA Barcodes. *PLoS biology*, **2**, e312.

Hebert PDN, Witt JDS, Adamowicz SJ (2003) Phylogeographical patterning in *Daphnia ambigua*: regional divergence and intercontinental cohesion. *Limnology and Oceanography*, **48**, 261–268.

Holt, B. G. Lessard J-P, Borregaard MK, Fritz SA, Araújo MB, Dimitrov D, Fabre P-H, Graham CH. Graves GR, Jønsson KA, Nogués-Bravo D, Wang Z, Whittaker RJ, Fjeldså J, Carsten Rahbek 2013. An Update of Wallace's Zoogeographic Regions of the World. Science 339: 74-78.

Hontoria F, Amat F (1992) Morphological characterization of adult *Artemia* (Crustacea, Branchiopoda) from different geographical origins. American populations. *Journal of Plankton Research*, **14**, 1461–1471.

Janzen, D. H. 1984. Dispersal of Small Seeds by Big Herbivores - Foliage Is the Fruit. *American Naturalist* 123:338-353.

694 Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance, web service. *BMC*
695   *genetics*, **6**, 13.
696 Kappas I, Baxevanis AD, Maniatsi S, Abatzopoulos TJ (2009) Porous genomes and
697   species integrity in the branchiopod *Artemia*. *Molecular phylogenetics and*
698   *evolution*, **52**, 192–204.
699 Ketmaier V, Pirollo D, Matthaeis E, Tiedemann R, Mura G (2008) Large-scale
700   mitochondrial phylogeography in the halophilic fairy shrimp *Phallocryptus spinosa*
701   (Milne-Edwards, 1840) (Branchiopoda: Anostraca). *Aquatic Sciences*, **70**, 65–76.
702 Knowlton N, Weigt LA (1998) New dates and new rates for divergence across the
703   Isthmus of Panama. *Proceedings of the Royal Society B: Biological Sciences*, **265**,
704   2257–2263.
705 Korn M, Green AJ, Machado M, García-de-Lomas J., Cristo M, Cancela da Fonseca L.,
706   Frisch D., Pérez-Bote JL, & Hundsdoerfer AK (2010) Phylogeny, molecular ecology
707   and taxonomy of southern Iberian lineages of *Triops mauritanicus* (Crustacea:
708   Notostraca). *Organisms Diversity & Evolution*, **10**, 409–440.
709 Lefébure T, Douady CJ, Gouy M, Gibert J (2006) Relationship between morphological
710   taxonomy and molecular divergence within Crustacea: proposal of a molecular
711   threshold to help species delimitation. *Molecular Phylogenetics and evolution*, **40**,
712   435–47.
713 Lockley, M. G., Lim, J. D., Kim, J. Y., Kim, K. S., Huh, M., and Hwang, K. G. 2012.
714   Tracking Korea's Early Birds: A Review of Cretaceous Avian Ichnology and Its
715   Implications for Evolution and Behavior. *Ichnos-an International Journal for Plant*
716   *and Animal Traces* 19:17-27
717 Lomolino MV, Riddle BR, Whittaker R, Brown JH (2010) *Biogeography*. Sinauer
718   Associates.
719 Maniatsi S, Kappas I, Baxevanis AD, Farmaki T, Abatzopoulos TJ (2009) Sharp
720   phylogeographic breaks and patterns of genealogical concordance in the brine
721   shrimp *Artemia franciscana*. *International Journal of Molecular Sciences*, **10**, 5455–
722   5470.
723 De Meester L, Gómez A, Okamura B, Schwenk K (2002) The Monopolization
724   Hypothesis and the dispersal–gene flow paradox in aquatic organisms. *Acta*
725   *Oecologica*, **23**, 121–135.
726 Mills S, Lunt DH, Gómez A (2007) Global isolation by distance despite strong regional
727   phylogeography in a small metazoan. *BMC Evolutionary Biology*, **7**, 225.
728 Miura O, Torchin ME, Bermingham E, Jacobs DK, Hechinger RF (2011) Flying shells:
729   historical dispersal of marine snails across Central America. *Proceedings of the*
730   *Royal Society B: Biological sciences*, **279**, 1061–1067.
731 Montero-Pau J, Gómez A, Muñoz J (2008) Application of an inexpensive and high-
732   throughput genomic DNA extraction method for the molecular ecology of
733   zooplanktonic diapausing eggs. *Limnology and Oceanography- Methods*, **6**, 218–
734   222.
735 Moscatello S, Belmonte G (2009) Egg banks in hypersaline lakes of the South-East
736   Europe. *Saline systems*, **5**, 3.
737 Muñoz J, Gómez A, Green AJ, Figuerola, J, Amat, F. Rico, C (2008) Phylogeography
738   and local endemism of the native Mediterranean brine shrimp *Artemia salina*
739   (Branchiopoda : Anostraca). *Molecular Ecology*, **17**, 3160–3177.
740 Muñoz J, Pacios F (2010) Global Biodiversity and Geographical Distribution of
741   Diapausing Aquatic Invertebrates: The Case of the Cosmopolitan Brine Shrimp,
742   *Artemia* (Branchiopoda, Anostraca). *Crustaceana*, **83**, 465–20.

743 Palumbi S (1996) Nucleic acids II: the polymerase chain reaction. *Molecular*
744   *systematics*, **2**, 205–247.
745 Petit R (1998) Identifying populations for conservation on the basis of genetic markers.
746   *Conservation Biology*, **1**, 844–855.
747 Pons, Joan, Timothy G. Barraclough, Jesus Gomez-Zurita, Anabela Cardoso, Daniel P.
748   Duran, Steaphan Hazell, Sophien Kamoun, William D. Sumlin, and Alfried P. Vogler.
749   (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed
750   insects. *Systematic biology* **55**, 595-609.
751 Proctor, V. W. (1968). Long-distance dispersal of seeds by retention in digestive tract of
752   birds. *Science*, **160**, 321-322.
753 Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, Automatic Barcode
754   Gap Discovery for primary species delimitation. *Molecular Ecology*, *21*, 1864-1877.
755 Sánchez MI, Green AJ, Amat F, Castellanos EM (2007) Transport of brine shrimps via
756   the digestive system of migratory waders: dispersal probabilities depend on diet and
757   season. *Marine biology*, **151**, 1407–1415.
758 Sánchez MI, Green AJ, Castellanos EM (2005) Seasonal variation in the diet of
759   Redshank *Tringa totanus* in the Odiel Marshes, southwest Spain : a comparison of
760   faecal and pellet analysis. *Statistica*, **52**, 210–216.
761 Sánchez MI, Hortas F, Figuerola J, Green AJ (2012) Comparing the dispersal potential
762   of a native and an invasive brine shrimp via waterbirds. *Freshwater Biology*, **50**,
763   737-742.
764 Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses
765   with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688-2690.
766 Sui, L. Y., Q. Zhang, J. Wang, H. He, P. Sorgeloos, and G. Van Stappen. 2012. High-
767   Temperature Adapted *Artemia Franciscana* (Kellogg, 1906) from Vietnam Is a
768   Suitable Strain for Inoculation in Temperate Conditions of Bohai Bay Saltworks,
769   China. Crustaceana, **85,** 1709-1723.
770 van Leeuwen, C. H. A., G. van der Velde, J. M. van Groenendael, and M. Klaassen.
771   2012a. Gut travellers: internal dispersal of aquatic organisms by waterfowl.
772   Journal of Biogeography, **39,** 2031-2040.
773 van Leeuwen, C. H. A., G. van der Velde, B. van Lith, and M. Klaassen. 2012b.
774   Experimental Quantification of Long Distance Dispersal Potential of Aquatic
775   Snails in the Gut of Migratory Birds. Plos One **7,** e32292.
776 Van Stappen G (2002) Zoogeography. In: *Artemia: basic and applied biology* (ed.
777   Abatzopoulos ThJ, Beardmore JA, Clegg JS SP), pp. 171–224. Kluwer Academic
778   Publishing.
779 Tackaert W, Sorgeloos P (1993) Use of brine shrimp to increase salt production.
780   *Seventh Symposium on Salt*, **1**, 617–622.
781 Tamura K, Peterson D, Peterson N (2011) MEGA5: molecular evolutionary genetics
782   analysis using maximum likelihood, evolutionary distance, and maximum parsimony
783   methods. *Molecular Biology and Evolution*, **28**, 2731–273.
784 Taylor DJ, Finston TLTL, Hebert PDN (1998) Biogeography of a widespread freshwater
785   crustacean: pseudocongruence and cryptic endemism in the North American
786   *Daphnia laevis* complex. *Evolution*, **52**, 1648–1670.
787 ter Braak CJF, Šmilauer P (2002) CANOCO Reference Manual and CanoDraw for
788   Windows User's Guide: Software for Canonical Community Ordination (version 4.5).
789   Ithaca, NY, USA (www.canoco.com): Microcomputer Power.
790 Tizol-Correa R (2009) Caracterización morfométrica del camarón de salmuera
791   (Artemia), de Cuba. *Revista cubana de investigaciones pesqueras*, **26**, 25–30.

792 Tizol-Correa R, Maeda-Martínez AM, Weekers PHH, Torrentera L, Murugan G (2009)
793   Biodiversity of the brine shrimp Artemia from tropical salterns in southern México
794   and Cuba. *Current Science*, **96**, 81–87.
795 Vanschoenwinkel, B., S. Gielen, M. Seaman, and L. Brendonck. 2009. Wind mediated
796   dispersal of freshwater invertebrates in a rock pool metacommunity: differences in
797   dispersal capacities and modes. *Hydrobiologia* **635**:363-372.
798 Varo N, Green A, Sánchez M, Ramo C (2011) Behavioural and population responses to
799   changing availability of *Artemia* prey by moulting black-necked grebes, *Podiceps*
800   *nigricollis*. *Hydrobiologia* **664**, 163-171.
801 Vest JL, Conover MR (2011) Food Habits of Wintering Waterfowl on the Great Salt Lake,
802   Utah. *Waterbirds*, **34**, 40–50.
803 Viana DS, Santamaría L, Michot TC, Figuerola J (2013) Migratory strategies of
804   waterbirds shape the continental-scale dispersal of aquatic organisms. *Ecography*,
805   **36**, 430-438.
806

Figure 1

Map of *Artemia franciscana* sampled sites and American bird migratory flyways.

The sampled populations are shown, with indication of the main American migratory flyways following Birdlife International (see text for details). Green shading: Pacific flyway, red shading: Central flyway, blue shading, Atlantic flyway.

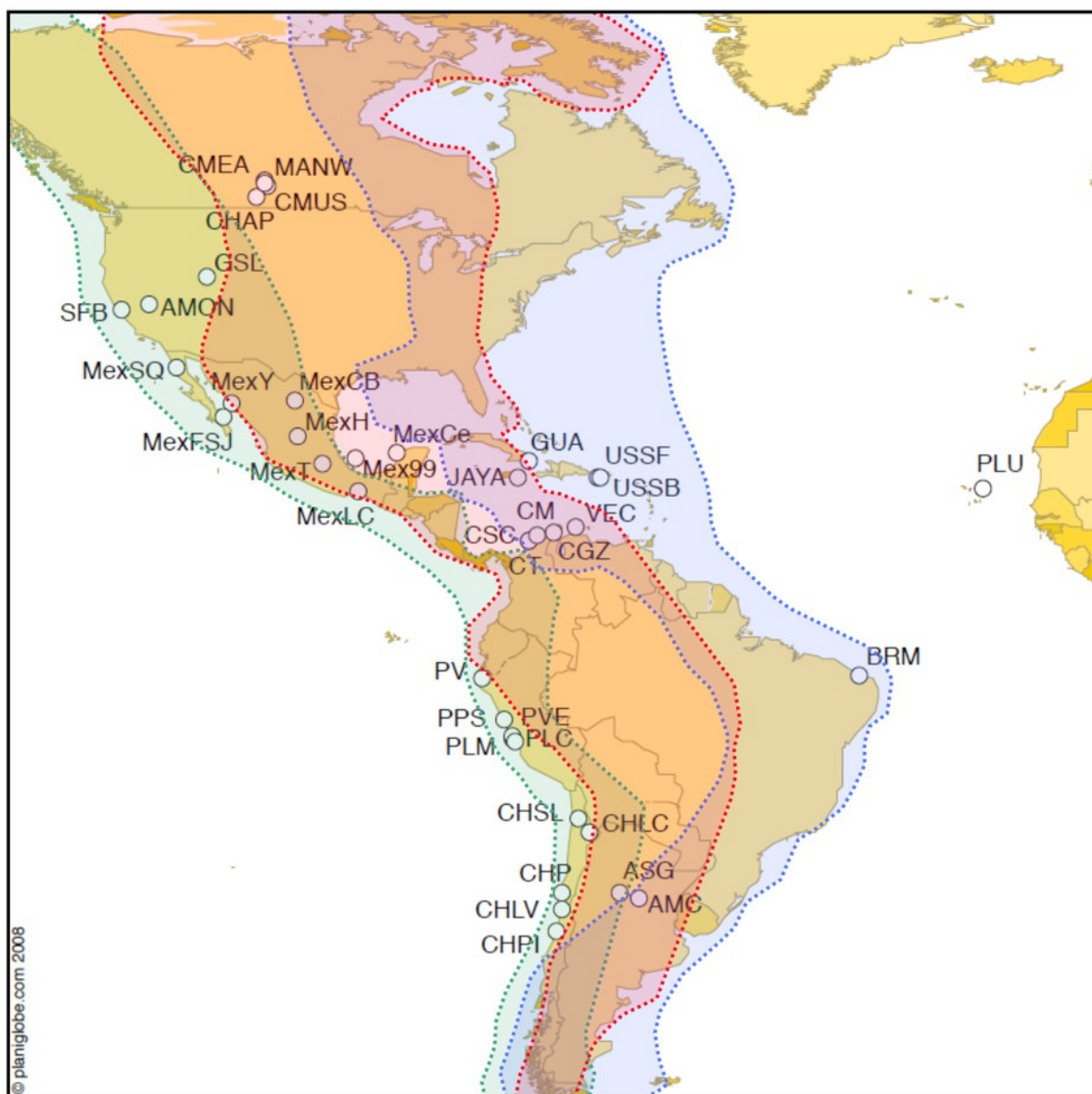



Figure 2

Phylogenetic relationships of native *Artemia franciscana* COI haplotypes

The tree topology is the one obtained in the NJ analysis, with bootstrap values shown for NJ (below branches) and ML (above branches). Haplotypes found in the commercialised populations SFB and GSL are marked in red. Haplotype numbers and populations where these were found are noted at the tips. Each lineage label indicates which countries it is found in and its overlap with the Pacific, Atlantic or Central migratory flyways (P, A or C respectively).

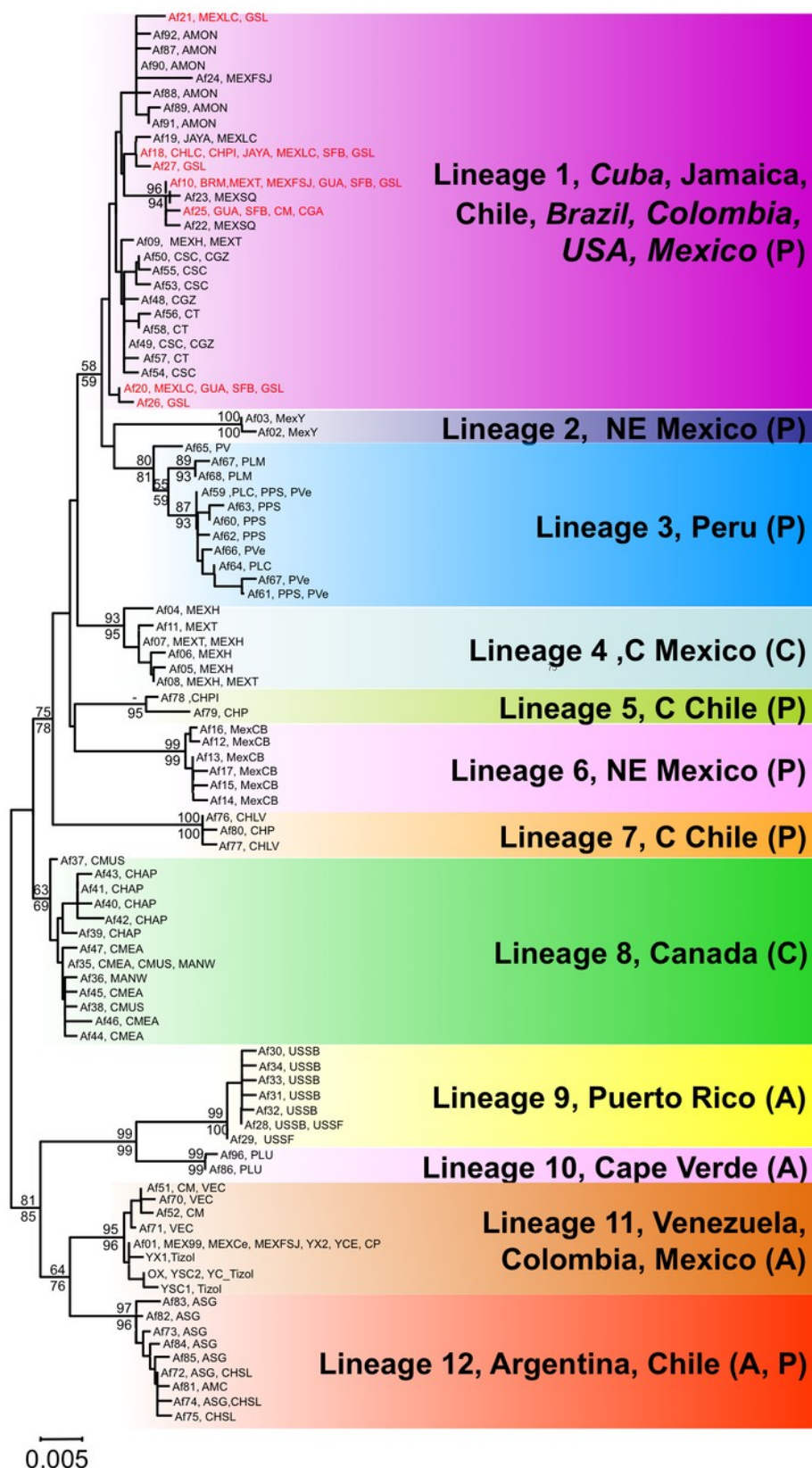



Figure 3

Geographic distribution of *Artemia franciscana* mtDNA lineages.

The distribution of each COI lineage is shown as areas with the same colour coding as in Figure 2. Disjunct areas are linked by lines. Introduced populations are denoted by a grey bucket. Only populations sampled for this study are included. Empty circles denote unsampled *A. franciscana* populations.

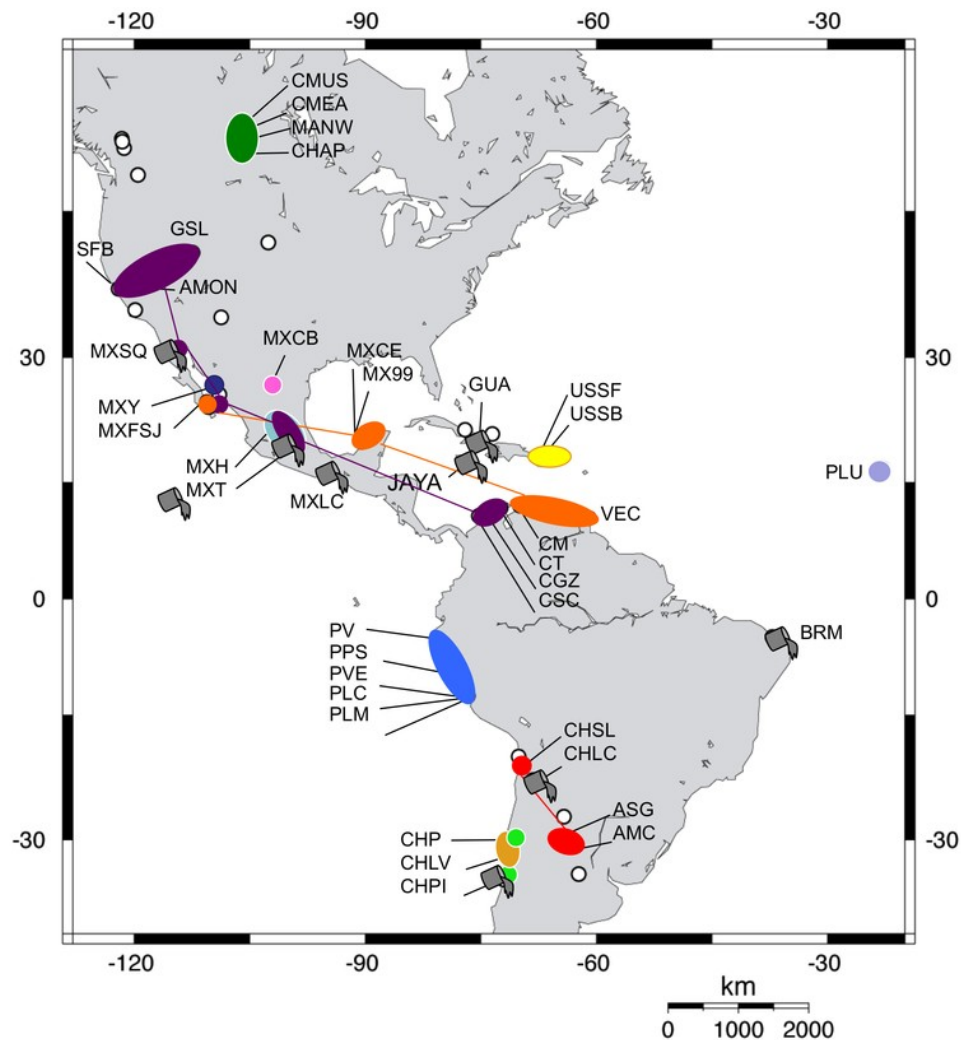



Figure 4

Phylogenetic relationships for native *Artemia franciscana* 16S haplotypes.

The topology obtained in the NJ analysis is shown, with bootstrap values for NJ (below branches) and ML (above branches). Haplotypes found in the commercialised populations SFB and GSL are marked in red. Bold sequences are those produced in this study and italics those from GenBank. The number of individuals sequenced in each location (unless one) is in parenthesis.

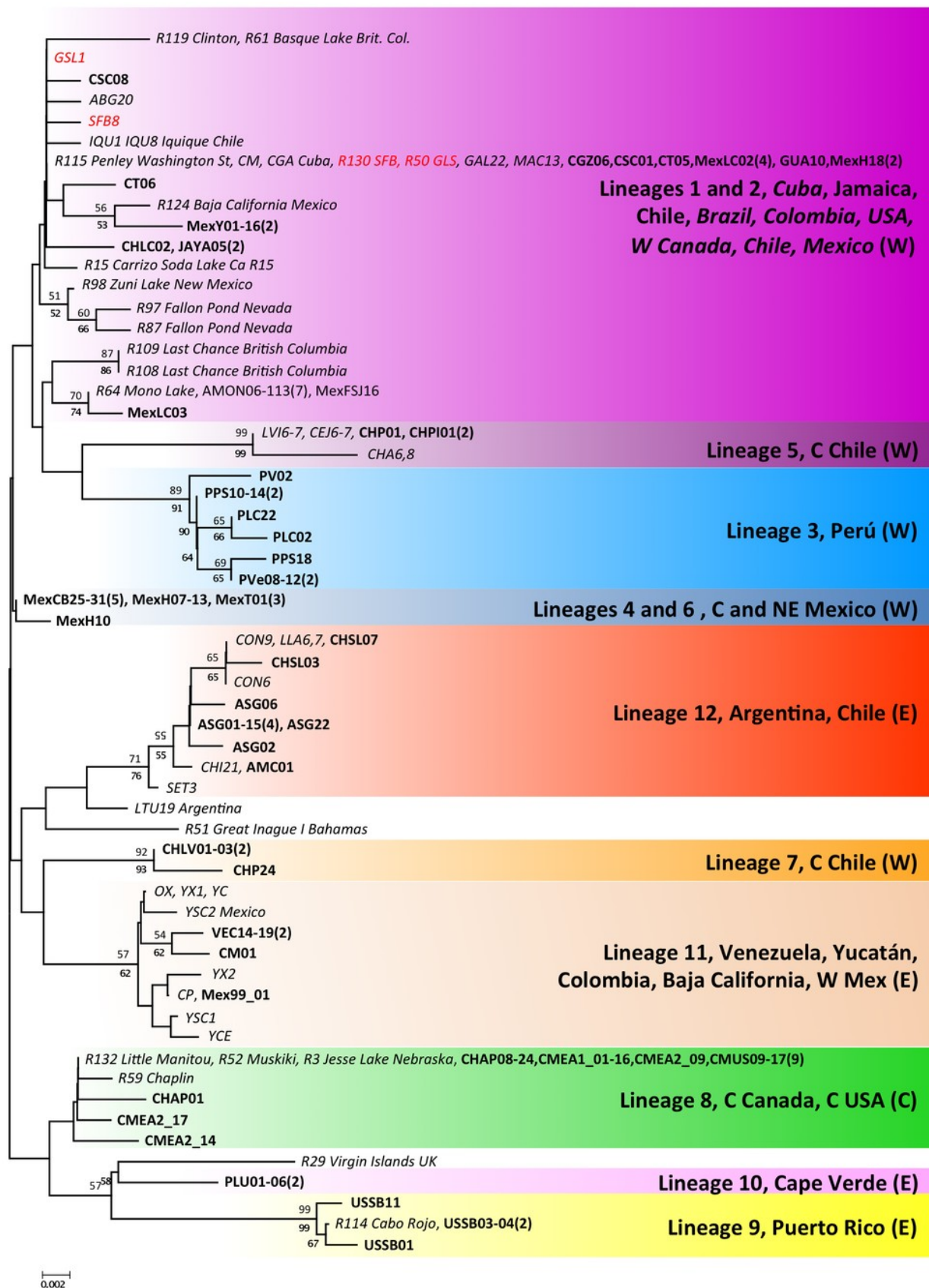
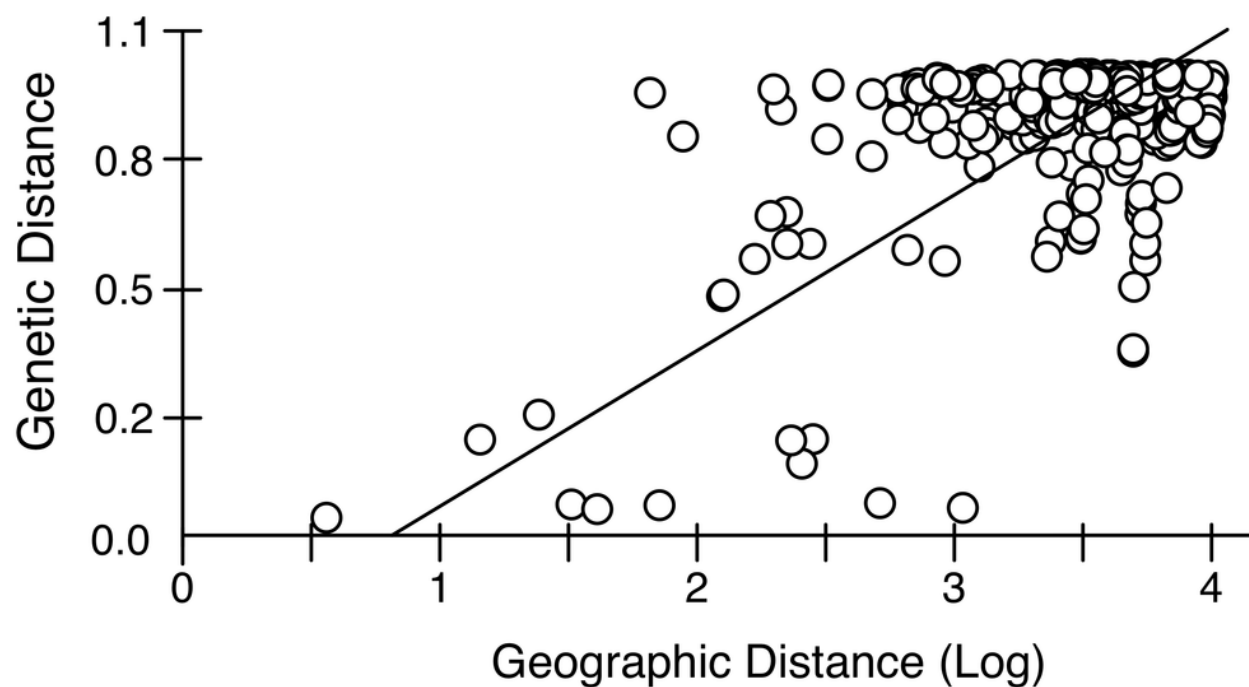



Figure 5

Isolation by distance in native *Artemia franciscana* populations.

Artemia franciscana populations sampled

Population code, year of sampling, COI haplotype code, number of individuals per haplotype per population (N_{HAP}) and number of individuals sequenced per population (N_{TOTAL}), are given for each sampling site. Hs = Standard gene diversity; π = Nucleotide diversity; N.A. = Insufficient data to calculate Hs (the minimum sample size per population performed by RAREFAC was 11). The main commercialised U.S.A. populations (i.e., SFB and GSL) and haplotypes from these populations shared with other populations are shown in red. Populations considered as introduced are indicated in bold. Only populations collected as part of this study are included.

Code	Population (year)	COI haplotypes	N_{HAP}	N_{TOTAL}	*Hs*	*π*	Ref
MexCe	Celestún, Yucatán, México (1984)	Af01	15	15	0.000	0.0000	-
Mex99	Real de las Salinas, Campeche, México (1999)	Af01	14	14	0.000	0.0000	-
MexY	Yavaros, Sonora, México (1993)	Af02; Af03	1; 11	12	0.167	0.0003	-
MexH	Salinas de Hidalgo, San Luis Potosí, México (1989)	Af04; Af05; Af06; Af07; Af08; Af09	1; 1; 1; 2; 3; 3	11	0.873	0.0091	-
MexT	Texcoco, Estado de México, México (1989)	Af07; Af08; Af09; Af10; Af11	1; 2; 7; 1; 2	13	0.705	0.0099	Intr cite
MexCB	Salinas Casa Blanca, Cuatro Ciénagas de Carranza, Coahuila, México (1995)	Af12; Af13; Af14; Af15; Af16; Af17	2; 6; 1; 3; 3; 1	16	0.817	0.0023	-
MexLC	La Colorada lagoon, Oaxaca, México (1993)	Af18; Af19; Af20; Af21	3; 1; 8; 2	14	0.648	0.0041	-
MexSQ	San Quintín, Baja California, México (??)	Af10; Af22; Af23	12; 1; 1	14	0.275	0.0005	-
MexFSJ	Faro San José, Baja California, México (1991)	Af01; Af10; Af24	1; 2; 1	4	N.A.	0.0226	-
GUA	Frank País, Guantánamo, Cuba (1994)	Af10; Af20; Af25	6; 1; 9	16	0.575	0.0019	Intr & S
GSL	Great Salt Lake, Utah, USA (??)	Af10; Af18; Af20; Af21; Af26; Af27	1; 2; 21; 2; 2; 1	29	0.475	0.0028	-
SFB	San Francisco Bay, California,USA (??)	Af10; Af18; Af20; Af25	26; 6; 4; 1	37	0.480	0.0033	-
USSF	Salina Fraternidad, Puerto Rico, Cabo Rojo, USA (2000)	Af28; Af29	12; 4	16	0.400	0.0007	-
USSB	Laguna de las Salinas Bastoncillo, Lajas, Puerto Rico, USA (2000)	Af28; Af30; Af31; Af32; Af33; Af34	6; 4; 1; 1; 1; 1	14	0.769	0.0017	-
MANW	Little Manitou Lake, Saskatchewan, Canada (2009)	Af35; Af36	8; 1	9	N.A.	0.0004	-
CMUS	Muskiki Lake, Saskatchewan, Canada (2009)	Af35; Af37; Af38	12; 1; 1	14	0.275	0.0005	-
CHAP	Chaplin Lake, Saskatchewan, Canada (2009)	Af39; Af40; Af41; Af42; Af43	11; 1; 1; 2; 1	16	0.533	0.0027	-
CMEA	Meacham Lake, Saskatchewan,Canada (2009)	Af35; Af44; Af45; Af46; Af47	26; 1; 1; 1; 1	30	0.253	0.0005	-
BRM	Mossoro, Grossos, Brazil (1994)	Af10	11	11	0.000	0.0000	Intr
CGZ	Salinas de Galerazamba, Colombia (1985)	Af48; Af49; Af50	1; 8; 7	16	0.592	0.0017	-
CM	Salinas de Manaure, Colombia (1999)	Af51; Af52	15; 1	16	0.125	0.0002	-
CSC	Salina Cero, Colombia (1999)	Af49; Af50; Af53; Af54; Af55	8; 4; 1; 1; 1	15	0.676	0.0015	-
CT	Tayrona, Colombia (1999)	Af56; Af57; Af58	9; 4; 2	15	0.590	0.0023	-
PPS	Playa Salinas, Ancash, Perú (1995)	Af59; Af60; Af61; Af62; Af63	6; 2; 1; 3; 1	13	0.756	0.0023	-
PLC	Los Chimus, Perú (1992)	Af59; Af64	14; 1	15	0.133	0.0002	-
PV	Virrilla, Piura, Perú (1996)	Af65	16	16	0.000	0.0000	-
PVe	Humedales de Ventanilla, Callao, Perú (1996)	Af59; Af61; Af66; Af67	8; 1; 6; 1	16	0.642	0.0020	-
PLM	La Milagrosa, Chilca, Perú (1993)	Af68; Af69	14; 2	16	0.233	0.0004	-
VEC	Salinas de Cumaraguas, Venezuela (1994)	Af51; Af70; Af71	10; 4; 1	15	0.514	0.0009	-
JAYA	Yallahs Pond, Jamaica (1998)	Af18; Af19	15; 1	16	0.125	0.0002	Kno sim
CHSL	Salar de Llamará, Chile (1994-lab)	Af72; Af73; Af74; Af75	5; 1; 1; 2	9	N.A.	0.0014	
CHLC	Laguna Cejas, Salar de Atacama, Chile (1995-lab)	Af18	16	16	0.000	0.0000	Ma diff
CHLV	Los Vilos, Poza Palo Colorado, Chile (1997)	Af76; Af77	6; 10	16	0.500	0.0008	
CHPI	Pichilemu Cahuil saltworks, Chile (??-lab)	Af18; Af78	8; 8	16	0.533	0.0103	Rep (Ga
CHP	Poza Pampilla IV Region, Chile (1997)	Af79; Af80	14; 1	15	0.133	0.0048	-
AMC	Mar Chiquita, Córdoba, Argentina (1997)	Af81	16	16	0.000	0.0000	-
ASG	Salinas Grandes, Córdoba, Argentina (2000)	Af72; Af73; Af74; Af82; Af83; Af84; Af85	7; 1; 1; 1; 1; 1; 1	13	0.731	0.0018	-
AMON	Mono Lake (1970s, ARC270)	Af87; Af88; Af89; Af90; Af91;	1; 1; 1;	12	0.682	0.0016	-

| PLU | Pedra de Lume, Sal Island, Cape Verde (??) | Af92 Af86; Af93 | 7; 1; 1 15; 1 | 16 | 0.125 | 0.0002 | - |

Genetic divergence between *Artemia franciscana* mtDNA lineages using COI data

Genetic Distance K2P+G estimated with MEGA between lineages. Genetic distances higher than (or equal to) 4% (0.04) are marked in bold.

Lineage (distribution)	1	2	3	4	5	6	7	8	9	10	11
1 (USA + introduced)	-										
2 (NE Mexico)	0.024 (0.006)	-									
3 (Perú)	0.020 (0.005	0.029 (0.007)	-								
4 (C Mexico)	0.018 (0.005)	0.031 (0.007)	0.027 (0.006)	-							
5 (C Chile)	0.023 (0.005)	0.033 (0.008)	0.031 (0.007)	0.023 (0.006)	-						
6 (NE Mexico)	0.026 (0.006)	0.038 (0.008)	0.038 (0.008)	0.028 (0.007)	0.028 (0.007)	-					
7 (C Chile)	0.032 (0.007)	**0.043** **(0.009)**	0.032 (0.007)	0.032 (0.008)	0.032 (0.007)	0.039 (0.009)	-				
8 (Canada)	0.020 (0.005)	0.034 (0.008)	0.029 (0.007)	0.019 (0.005)	0.025 (0.007)	0.029 (0.007)	0.031 (0.008)	-			
9 (Puerto Rico)	**0.048** **(0.010)**	**0.060** **(0.012)**	**0.058** **(0.011)**	**0.048** **(0.010)**	**0.049** **(0.010)**	**0.055** **(0.011)**	**0.052** **(0.011)**	**0.040** **(0.009)**	-		
10 (Cape Verde)	**0.044** **(0.009)**	**0.059** **(0.012)**	**0.051** **(0.010)**	**0.044** **(0.010)**	**0.047** **(0.010)**	**0.050** **(0.010)**	**0.045** **(0.010)**	0.032 (0.008)	0.023 (0.006)	-	
11 (Yucatán, Colombia…)	0.035 (0.008)	**0.044** **(0.010)**	**0.043** **(0.009)**	0.033 (0.008)	**0.041** **(0.009)**	**0.045** **(0.010)**	**0.043** **(0.010)**	0.023 (0.006)	0.038 (0.008)	0.034 (0.008)	-
12 (Argentina, Chile)	0.037 (0.008)	**0.048** **(0.010)**	**0.046** **(0.009)**	0.037 (0.008)	**0.043** **(0.009)**	**0.047** **(0.010)**	**0.044** **(0.010)**	0.025 (0.006)	**0.044** **(0.009)**	0.036 (0.008)	0.020 (0.005)

Table 3 _(on next page)

Redundancy analyses (RDA) assessing the contribution of spatial (geographical coordinates of populations) and environmental factors to the genetic distance between *Artemia franciscana* populations

Explained variance (%) for three RDAs with different environmental variables is given in separate columns. The first RDA included flyway and human introduction as environmental variables, while the others considered only flyway or introduction. Results are given for flyway assignments made according to Boere & Stroud (2006) and Birdlife international (http://www.birdlife.org/datazone/userfiles/file/sowb/flyways/) (see Table S2 for matrix details).

Source of variation	Flyway + introduction	Flyway	Introduction
Space	8.2 / 8.7	8.3 / 10.2	10.8
Environment	31.2 / 30.6	21.2 / 18.7	15.6
Environment/Space interaction	2.4 / 1.8	2.3 / 0.4	0.0

Figure 6

Artemia franciscana adults and cysts

Possible photo for Journal homepage. Author Andy J Green, released with a CC BY 3.0 licence.