**Original Article**

# Genetic signatures of historical dispersal of fish threatened by biological invasions: the case of galaxiids in South America

Delphine Vanhaecke[1†], Carlos Garcia de Leaniz[2], Gonzalo Gajardo[3], Jason Dunham[4], Guillermo Giannico[5], Sofia Consuegra[1, 2*]

[1]Aberystwyth University, Institute of Biological, Environmental & Rural Sciences, Aberystwyth SY23 3DA, UK

[2]Swansea University, College of Science, Department of BioSciences, Singleton Park, Swansea SA2 8PP, UK

[3]Universidad de Los Lagos, Laboratorio de Genética, Acuicultura y Biodiversidad, Osorno, Chile

[4]U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center, Corvallis, Oregon 97331, USA

[5]Department of Fisheries and Wildlife, Oregon State University, Corvallis, OR, U.S.A. 97330

[†] Present address: Flanders Marine Institute, Wandelaarkaai 7, B-8400 Oostende, Belgium

[*]Corresponding author: Sofia Consuegra

Swansea University, College of Science, Department of BioSciences, Singleton Park, Swansea SA2 8PP, UK ; Tel. +44 (0) 1792 602931 ; Email. S.Consuegra@swansea.ac.uk

Keywords: *Aplochiton,* biological invasions, gene flow, mtDNA, microsatellites, phylogeography, post-glacial colonisation, salmonids.

**Word count: 6,673 (including abstract, main text and references)**

**Figures: 3 (2 pages)**

**Abstract**

**Aim** The ecological effects of biological invasions are well documented, but little is known about the effects of invaders on the genetic structure of native species. We examined the phylogeography, genetic variation and population structuring of two galaxiid fishes, *Aplochiton zebra* and *A. taeniatus* threatened by non-native salmonids, and whose conservation is complicated by misidentification and limited knowledge of their genetic diversity.

**Location** Chile and the Falkland Islands.

**Methods** We combined microsatellite and mitochondrial DNA (16S rDNA and CO-I) markers to compare genetic diversity, effective population size and gene flow of *Aplochiton* spp. populations differentially affected by salmonid presence.

**Results** We identified two 16s rDNA haplotypes among *A. zebra,* one dominant in coastal populations and another dominant in inland populations. Populations living on the Island of Chiloé displayed a mixture of coastal and inland haplotypes, as well as high microsatellite diversity, as one would expect if the island had been a refugium during the Last Glacial Maximum, or a contact zone among populations. Microsatellite data revealed strong population structuring, indicative of current isolation patterns, and a negative correlation between the genetic diversity of *A. zebra* and the relative abundance of invasive salmonids.

**Main conclusion** Our study indicates that population structuring of *A. zebra* reflects the influence of historical patterns of migration, but also current levels of reduced gene flow among watersheds. Invasive salmonids, known to compete with and prey on native galaxiids, may have had negative impacts on *Aplochiton* spp. genetic diversity. The low genetic variation found in some populations, coupled with potential biases in abundance estimates due to species misidentification, highlight the urgent need for more research into the conservation status of the two *Aplochiton* species.

## Introduction

Biological invasions represent a major cause of biodiversity loss (Clavero & García-Berthou, 2005), and although they seldom cause wholesale extinctions (Ricciardi *et al.*, 2011), they can trigger ecological changes which can make native species less resilient to subsequent stressors (Parker *et al.*, 1999). Surprisingly, relatively little is known about the genetic impact of invasions (e.g. Strayer *et al.*, 2006; Carroll, 2011), as most studies tend to focus on genetic changes exhibited by the invaders (Monzón-Argüello *et al.*, 2013; Monzón-Argüello *et al.*, 2014b; Monzón-Argüello *et al.*, 2014a), rather than genetic responses of native species (Vanhaecke *et al.*, 2012a). This is unfortunate because without genetic data it may be difficult to get accurate assessments of the conservation status of threatened populations (Traill *et al.*, 2010) or to quantify the impact of biological invasions. For example, genetic data can be applied to detect changes in effective population size, estimate gene flow, or detect range contractions of native species in relation to the presence of invaders (Arenas *et al.*, 2012). Estimates of genetic diversity are also essential for understanding the long-term evolutionary consequences of biological invasions (Strauss *et al.*, 2006), and could perhaps also be used as an early warning of impending impacts, before range shifts or local extirpations take place.

In this study we employed molecular markers to understand how non-native salmonids may have impacted native galaxiid fishes, taking into account historical biogeographic patterns. We focused on two closely related species, *Aplochiton zebra* (Jenyns, 1842) and *A. taeniatus* (Jenyns 1842) inhabiting rivers of Chilean Patagonia and the Falkland Islands: two areas where native galaxiids are threatened by the introduction of non-native salmonids since the 19th century (Garcia de Leaniz *et al.*, 2010; Schröder & Garcia de Leaniz, 2011; Arismendi *et al.*, 2014). *A. zebra* is considered endangered over its entire range (Lattuca *et al.*, 2008; Arismendi *et al.*, 2009) due to its ecological overlap with non-native salmonids (McDowall, 2006; McDowall, 2010). However, the conservation status of *A.*

*taeniatus* remains unclear (McDowall, 2006) and until recently its range did not include the Falkland Islands, where it had been misidentified as *A. zebra* (Vanhaecke *et al.*, 2012b). This is not surprising because the two species are morphologically very similar, and the limited information available on their ecology and genetic structure makes prioritisation of populations for conservation difficult (Vanhaecke *et al.*, 2012b; Alò *et al.*, 2013). *A. taeniatus* appears to attain larger sizes than *A. zebra* and is considered a specialist that preys mostly on fish, in contrast to *A. zebra* which is considered a generalist that feeds mainly on aquatic invertebrates. The existence of a third species, *A. marinus,* has been suggested, but its taxonomic status remains unclear (Alò *et al.*, 2013).

The distribution of native galaxiids in South America is related to historical pathways of dispersal (Habit *et al.*, 2012). To understand these historical pathways for *Aplochiton* spp. we applied molecular markers with different modes of inheritance, mitochondrial DNA (mtDNA; maternally inherited) and microsatellites (bi-parentally inherited). This combination of markers allowed us to reconstruct the phylogeography of *Aplochiton* spp. and also to provide insights into the historical and current drivers of genetic structure in relation to the presence of non-native salmonids. We hypothesised that ecological impacts of non-native salmonids (e.g., predation, competition, or other processes) could decrease population sizes that would be reflected by reduced microsatellite genetic diversity of *Aplochiton* in rivers and lakes invaded by salmonids, whereas mtDNA genetic variation should reflect the extent of the last glacial maximum (LGM) showing the influence of historical routes of dispersal.

## Materials and methods

### *Sampling*

We collected tissue samples from 456 individual *Aplochiton* spp. from 20 streams in Chile and 15 streams in the Falkland Islands (Fig. 1; Appendix S1). Given the morphological similarity between *A. zebra* and *A. taeniatus*, molecular analysis was needed to discriminate between both species and their hybrids (Vanhaecke *et al.*, 2012b; Alò *et al.*, 2013). After DNA extraction and genetic barcoding using the CO-I gene, samples were classified by species as in Vanhaecke *et al.* (2012b). We found 341 samples were classified as *A. zebra* (317 from Chile and 24 from the Falklands) and 115 to be classified as *A. taeniatus* (61 from Chile and 54 from the Falklands). As a result, sample sizes were unbalanced and reduced in some populations. We analysed the genetic structure of 13 *A. zebra* populations from Chile which had at least 16 individuals per sampling site (Table 1). We used data on catch per unit effort (CPUE, fish/min/$m^2$) from single-pass electro-fishing to derive indices of the relative abundance of *Aplochiton* and salmonids (as in Vanhaecke *et al.*, 2012a), and employed the non-parametric Spearman's rank correlation coefficient (SPSS v19) to examine the relationship between CPUE and altitude. Although CPUE from electrofishing surveys is typically a poor indicator of absolute fish abundance, it can be used as a proxy to compare relative species abundance among sites (Bergman *et al.*, 2011), and use of rank-based methods are more appropriate in these cases.

### *DNA extraction and genotyping*

DNA was extracted using the Wizard® SV 96 DNA Purification. All samples were amplified for 13 microsatellite loci (Aggarwal *et al.*, 2011) and mtDNA CO-I (Vanhaecke *et al.*, 2012b), and those individuals identified as hybrids (Vanhaecke *et al.*, 2012b) were removed from the dataset. To estimate the repeatability of scoring, between 11 and 71 individuals were

98 re-scored for each marker and allele sizes were compared. Potential errors in genotyping were
99 estimated using MICRO-CHECKER (Van Oosterhout *et al.*, 2004), with the exception of
100 microsatellites Aze2, Aze4 and Aze5 that have a complex motif.
101 A region of 479 bp of the 16S rDNA gene of the mtDNA was amplified in 96 *A.*
102 *zebra* samples using the universal primers 16S rRNAr and 16S rRNAbr (Palumbi *et al.*,
103 1991). Both strands were sequenced on an ABI 3100 DNA analyser (Applied Biosystems
104 CA, USA) and sequences were aligned using BioEdit v. 7.0.9 (Hall, 1999).
105
106 *Microsatellite analyses: genetic variation*
107 Linkage disequilibrium between microsatellite loci was computed using GENEPOP
108 (Raymond & Rousset, 1995) and all markers were tested for $F_{ST}$ deviation from neutral
109 expectations using FDIST in ARLEQUIN 3.5.1.2 (Excoffier & Lischer, 2010). Observed and
110 expected heterozygosity (Ho, He) estimates and tests for Hardy Weinberg equilibrium were
111 conducted in TFPGA v. 1.3 (Miller, 1997) and the significance was adjusted by Bonferroni
112 correction for multiple tests (Rice, 1989). Allelic richness (Ar) was calculated by FSTAT
113 v.2.9.3.2 (Goudet, 1995). Comparisons of genetic diversity (Ar, He, Ho) among populations
114 and geographical groups were also performed in FSTAT using 10,000 permutations. All
115 populations were tested for recent bottlenecks using BOTTLENECK v.1.2.02 (Piry *et al.*,
116 1999) under two models; infinite alleles (IAM) and two-phase (TPM), using a Wilcoxon
117 sign-rank test based on a 1000 iterations. For TPM, ps (frequency of single step mutations)
118 was set to 0.90 with a 10% variance of multistep mutations (Piry *et al.*, 1999). Contemporary
119 effective population size ($N_e$) was estimated using COLONY2 (Wang, 2009) from the
120 frequencies of full or half siblings in each population calculated by sibship assignment
121 analysis. $N_e$ was calculated by three runs of medium length with the full likelihood option,
122 only for populations with a minimum sample size of 20. We also used a new implementation

of the LD method included in NeEstimator V2 (Do *et al.*, 2014), using allele frequencies >0.02 to minimize potential bias caused by rare alleles (Waples & Do, 2008).

*Microsatellite analyses: population structuring*

Pairwise genetic differentiation ($F_{ST}$) between populations was computed in TFPGA and significance values were obtained by 10,000 permutations. Pairwise genetic distances ($D_{est}$) (Jost, 2008) were also calculated in SMOGD (Crawford, 2010). Population genetic structuring was examined using analysis of molecular variance (AMOVA) implemented in ARLEQUIN (10,000 permutations). In order to identify the most plausible spatial driver of genetic variation, two analyses were carried out grouping populations by (a) region: West Chiloé Island, North East Chiloé Island (including Reloncavi), Lake Ranco and Valdivia (Fig. 1), or (b) origin: freshwater (Lake Ranco and Valdivia populations) vs coastal (Chiloé and Reloncavi populations).

Population structuring was further analysed taking into account the spatial distribution of genotyped individuals using TESS 2.3.1 (Chen *et al.*, 2007). We ran 100 replicates per K (2-15) using an admixture model, 500 sweeps of burn-in, and a running period of 2,100 sweeps with the interaction parameter ($\Psi$) fixed at 0.6. The maximum number of clusters (Kmax) was inferred from changes in the Deviance Information Criterion (DIC) (Spiegelhalter *et al.*, 2002). The results from the replicates were averaged using the software CLUMP (Jackobsson & Rosenberg, 2007) and the output was represented using DISTRUCT 1.1 (Rosenberg, 2004).

Isolation by distance (IBD) among all populations, and between coastal and freshwater groups, was estimated with a Mantel test (10,000 permutations) on genetic distance measured by $F_{ST}/(1-F_{ST})$ and geographical distance (km) using the Zt statistic (Bonnet & Van de Peer, 2002). In order to identify barriers to gene flow among *A. zebra*

148 populations, we used the Monmonier's (1973) maximum difference algorithm implemented
149 in BARRIER v. 2.2 (Manni *et al.*, 2004). Geographical coordinates for each sampling
150 location were connected by Delauney triangulation and the analysis was conducted using two
151 matrices of genetic distance ($D_{est}$ and $F_{ST}$, described above). To assess the robustness of the
152 barriers, analyses were also run for 100 bootstrapped $F_{ST}$ matrices and for each microsatellite
153 locus separately (Hemmer-Hansen *et al.*, 2007). Barriers supported by high bootstrap values
154 (>65%) were ranked (I to IV) according to the number of loci supporting them, strong
155 support being inferred when more than 10 loci supported the presence of a barrier.
156 We used Approximate Bayesian Computation (ABC) implemented in DIYABC v.
157 2.0.3 (Cornuet *et al.*, 2008) to reconstruct the potential colonization routes of *A. zebra* using
158 microsatellite data. We compared three simple colonisation scenarios based on the mtDNA
159 groupings (Chiloé, Lake Ranco, Valdivia and the Falklands). The potential scenarios
160 involved a two-step colonisation of Lake Ranco/Valdivia and Chiloé/Falklands (Scenario 1),
161 a sequential colonisation of Lake Ranco, Valdivia and the Falklands (Scenario 2) and the
162 possibility of admixture between Valdivia and Chiloé (Scenario 3). Priors were considered
163 uniformly distributed and we used the default settings for mutation rates. We simulated
164 300,000 data sets per scenario and considered the following summary statistics: mean number
165 of alleles per locus, mean gene diversity, mean size variance, mean M ratio within each
166 population and pairwise $F_{ST}$ values between populations. Effective population sizes were
167 considered to vary between 10 and 100,000 for three Chilean populations and between 10 and
168 50,000 for the Falklands. Confidence in the scenario with the highest posterior probability
169 was estimated by comparing simulated and observed summary statistics, and by calculating
170 type I and II error rates (Cornuet *et al.*, 2008).
171
172

*Mitochondrial DNA (mtDNA) analysis: genetic diversity and population structuring*

For mtDNA, the number of haplotypes (h), haplotype diversity ($H_D$) and nucleotide diversity ($\pi$) were calculated in DnaSP v5 (Librado & Rozas, 2009). AMOVA was conducted in ARLEQUIN to estimate population structuring, using the same groups as for the microsatellite analyses.

Non-parametric Spearman's rank correlations between genetic diversity (microsatellite heterozygosity and allelic richness, mtDNA haplotype and nucleotide diversity) of *A. zebra* populations and salmonid abundance (as estimated by CPUE) were carried out using SPSS v19 to test for potential effects of salmonid presence on *A. zebra* genetic diversity.

## Results

*Relative abundance (CPUE) of A. zebra and salmonids*

We found salmonids in all but two of the study rivers (U28 and U29; Table 1). The relative abundance of salmonids (estimated by CPUE, fish/min/$m^2$) ranged from 0 (West Chiloé) to 0.0075 (R. Punahue, Valdivia) and, as expected for cold water fishes, CPUE increased significantly with altitude (Spearman's $r_s = 0.718$; $P = 0.003$). The relative abundance of *A. zebra* was not correlated with altitude and was highest in site U29, on the West coast of the Island of Chiloé, an area free from fish farming and not yet invaded by salmonids (Young *et al.*, 2010).

*Microsatellite genetic diversity and population genetic structuring of A. zebra*

Scoring repeatability was 94% ± 6% on average (per allele). No large allele drop outs were identified, although null alleles were detected in two populations for markers Aze3 and Aze11, and in five populations for Aze14. Aze14 was therefore excluded from further analyses. Microsatellites Aze6 and Aze13 were in linkage disequilibrium in one of the 13 populations analysed (RPU, $P = 0.00045$). None of the markers displayed evidence of being under selection. Significant deviations from Hardy Weinberg equilibrium were found for microsatellite Aze9 in one population (U29, $P = 0.0002$) and Aze14 in two populations (REN, RPU $P = 0.0001$), in both cases due to an excess of homozygotes (Appendix S1). $F_{IS}$ was high for locus Aze4 in all populations (Appendix S1), and was also excluded from analyses.

Inland and coastal populations did not differ in genetic diversity (Appendix S2; coastal: Ar = 9.03; Ho = 0.65; Hs = 0.74 and inland: Ar = 8.93; Ho = 0.72; Hs = 0.76; $P = 0.393$). Effective population size ($N_e$) estimated using COLONY was generally low for *A. zebra* and ranged between 47 (RPU) and 127 (U29). Confidence intervals were very wide,

particularly those obtained by the LD method probably due to small sample sizes (Table 2); we therefore only considered those confidence intervals obtained from COLONY for further analyses. Neither genetic diversity, nor effective population size ($N_e$), were significantly associated with altitude (Spearman rank correlation $P = 0.237$). We found a negative correlation between microsatellite genetic diversity of *A. zebra* and the relative abundance of salmonids (He: $r = -0.48$, $P = 0.048$; Ar: $r = -0.52$, $P = 0.030$). Only the population with the lowest effective population size (RPU) showed significant evidence of genetic bottleneck after strict Bonferroni correction under the IAM model ($P = 0.0012$). The genetic diversity of *A. zebra* in the Falkland Islands, estimated by pooling samples from all populations (due to limited numbers of individuals in local collections) was similar to that in Chile (N=22, Ar=7.5, He=0.72, Ho=0.61).

Pairwise $F_{ST}$ comparisons of population differentiation ranged between 0 and 0.066 (Appendix S3), and $F_{ST}$ values showed a highly significant pattern of differentiation ($P<$ 0.001) among geographical regions (W. Chiloé, E. Chiloé, Reloncavi, Valdivia and Lake Ranco) but non-significant differences within them (Appendix S3). The only significant $F_{ST}$ within a region was between coastal populations U33 and U20 on the East Coast of Chiloé. $D_{est}$ values ranged between 0 and 0.264 and generally showed similar patterns as $F_{ST}$, with low to zero values within regions and high differentiation among regions (Appendix S1b).

AMOVA revealed that 4.6% of genetic variation could be explained by differences among the four geographical regions considered ($F_{CT} = 0.046$; $P < 0.001$), 0.8% of variation was explained by differences among populations within locations ($F_{SC} = 0.008$; $P < 0.001$), and 94.7% of variation was distributed within populations ($F_{ST} = 0.054$; $P < 0.001$). In contrast, dividing the populations into inland or coastal groups only explained 1.97% of genetic variation between groups ($P = 0.01$).

Individual assignment analysis (TESS) supported regional genetic clusters with a most likely K = 4 (Fig. 1). All populations displayed very uniform genotype clustering except for sites U28 and U29 on the West coast of Chiloé, which were the populations with the highest degree of admixture (average Q membership coefficients ranged between 16% and 59%). The Mantel test revealed significant patterns of isolation by distance (IBD) among all 13 *A. zebra* populations in Chile ($r = 0.490$; $P = 0.001$; Fig. 2), and also in coastal ($r = 0.680$; $P = 0.002$) and inland populations ($r = 0.860$; $P = 0.044$) considered separately. Results from BARRIER supported the population structuring identified by TESS. Two strong barriers were identified based on matrices of pairwise $F_{ST}$ and $D_{est}$ differentiation: barrier I, which separated inland from coastal populations and was supported by 10 of the 11 loci; and barrier II, which separated the two freshwater drainages of Lake Ranco and Valdivia and was supported by 9 of the 11 loci (Fig. 1; Appendix S2). Two weaker barriers included barrier III, which separated East and West Chiloé (supported by 7 loci), and barrier IV which separated River Lenca (coastal mainland) and Chiloé Island (supported by 5 loci), but these were not strongly supported by pairwise $F_{ST}$ values based on all loci combined. The most likely colonisation scenario identified by DIYABC was one where there was population mixture between Chiloé and Valdivia, followed by a more recent split of the Falklands populations from Chiloé (Appendix S3, scenario 3).

*mtDNA diversity and population structuring of A. zebra*

The sequencing of 479 bp of the mtDNA 16S rDNA region in 98 *A. zebra* revealed three parsimoniously informative sites, one of which was a non-synonymous substitution (A/G). Based on this difference, we identified four haplotypes that could be included in either haplogroup 1 (base site 148 = G; H1 with n = 51 and H2 with n = 2) or haplogroup 2 (base site 148 = A; H3 with n = 2 and H4 with n= 43; GenBank accession numbers: JF437635 –

JF437642). Haplogroup 1 tended to be more common in inland populations (Valdivia, L. Ranco) than haplogroup 2, which characterized most of the coastal fish in Chiloé and the Falklands individuals (Fig. 3; Table 3). Grouping populations according to regional location indicated that West Chiloé had the highest genetic diversity ($H_D = 0.56$, $\pi = 0.001$) and Valdivia the lowest ($H_D = 0$, $\pi = 0$; Table 3). For CO-I we had previously identified six unique haplotypes defined by five mutations amongst *A. zebra* (Vanhaecke *et al.*, 2012b), the highest variability corresponding to East Chiloé.

AMOVA by geographical region (W. Chiloé, E. Chiloé and Reloncavi, Lake Ranco and Valdivia) using 16S rDNA revealed that most (56.7%) of the genetic variation could be explained by differences among regions ($\Phi_{CT} = 0.567$; $P < 0.001$) , 44.3% was explained by differences within populations ($\Phi_{ST} = 0.557$; $P < 0.001$), while the amount of variation explained by differences among populations within groups was negligible (i.e. high genetic uniformity within regions; $\Phi_{SC} = -0.023$; $P = 0.746$). When populations were divided into coastal and inland groups, genetic variation between groups accounted for 54% ($\Phi_{CT} = 0.539$; $P = 0.001$), differences within populations accounted for 35.9% ($\Phi_{ST} = 0.641$; $P < 0.001$) and 8% was explained by differences among populations within groups ($\Phi_{SC} = 0.219$; $P = 0.002$).

For CO-I, 5.4% of genetic variation could be explained by differences among groups ($P = 0.04$), 89% by differences within populations ($P < 0.001$), and 5.5% by differences among populations within groups ($P = 0.102$). When populations were divided into coastal and inland groups, genetic variation between groups accounted for 6.2% ($P = 0.001$), differences within populations accounted for 87% ($P < 0.001$) and differences among populations within groups accounted for 6.3% ($P = 0.039$).

In contrast to microsatellite diversity, we did not find a significant correlation between relative salmonid abundance and measures of *A. zebra* mtDNA diversity (haplotype diversity: $r$ = -0.54, $P$ = 0.055; nucleotide diversity: $r$ = -0.45, $P$ = 0.127).

*Genetic diversity of Aplochiton taeniatus*

Analysis of genetic diversity was carried out in those populations of *A. taeniatus* with sufficient sample size following DNA barcoding and species identification: River Huicha from Chiloé (N = 28) and North Arms from East Falklands (N = 30). Comparisons were performed using those nine microsatellites that were variable for both species (Vanhaecke *et al.* 2012b). Allelic richness (Ar) of *A. taeniatus* in North Arms (Ar = 2.8) was significantly lower than in Chiloé (Ar = 9.2, $P$ = 0.016) although they did not differ significantly in Ho (0.27 vs. 0.46 $P$ = 0.11). In addition, *A. taeniatus* in the Falklands (but not in Chiloé) displayed evidence of a recent population bottleneck under the two mutation models considered (IAM, $P$ = 0.004; TPM, $P$ = 0.019) and a shifted allelic frequency distribution compared to the expected L-shaped distribution as well as deviation from Hardy-Weinberg equilibrium ($P$<0.001). Estimates of effective population size of *A. taeniatus* were very low in the Falklands ($N_e$ = 26; 95% CI = 19 - 47) and low in Chiloé ($N_e$ = 49; 95% CI = 29 - 100).

## Discussion

Non-native salmonids are known to impact native galaxiids by displacing, out-competing, and preying upon them (Arismendi *et al.*, 2009; Young *et al.*, 2009; Garcia de Leaniz *et al.*, 2010). However, the potential genetic impact of salmonids on galaxiids remains largely unknown (but see Vanhaecke *et al.*, 2012a). Our estimates of relative salmonid abundance in Chilean streams were similar to those of previous studies that found salmonids were absent in some coastal rivers of Chiloé and were most abundant in areas with greater numbers of fish farms that could serve as sources for invaders (Young *et al.*, 2010; Consuegra *et al.*, 2011; Monzón-Argüello *et al.*, 2014b).

We used two types of molecular markers with different rates of evolution and modes of inheritance to investigate the drivers of population structuring in *A. zebra*: maternally inherited mtDNA to provide information on historical phylogeographic events (e.g., post-glacial colonization) and bi-parentally inherited microsatellite loci to infer more recent demographic events influencing patterns of genetic variability (Emerson & Hewitt, 2005). Analyses of the mitochondrial 16S rDNA gene revealed weak population structuring (i.e., high genetic uniformity) within *A. zebra* populations inhabiting freshwater systems, which was unrelated to salmonid presence, and low divergence between inland and coastal populations, including the Falkland Islands. These, along with our analyses of microsatellite variation, suggest that the population structuring of *A. zebra* reflects the signature of historical patterns of colonisation since the last glacial maximum (LGM) and current geographical barriers. We also found a negative association between microsatellite genetic diversity of *A. zebra* and estimates of salmonid abundance, but not between mtDNA diversity and salmonid presence. These results suggest that any genetic impacts of salmonids upon *Aplochiton* must have been relatively recent (indicated by microsatellite diversity), and have yet to be reflected in changes in mtDNA diversity.

Unlike the case of the more abundant *Galaxias maculatus*, whose genetic diversity appears to be unaffected by salmonid abundance (Vanhaecke et al., 2012a), our results suggest that invasive salmonids may have caused a decrease in the nuclear genetic diversity of the endangered *A. zebra*; it is important to note, however, that sample sizes were small for some populations, and that results were merely correlational; genetic inferences, therefore, need to be made with caution. An unknown variable related to both salmonid abundance and microsatellite diversity in *Aplochiton* could also explain the associations we observed.

Intraspecific mtDNA sequence divergence can be low in fishes (Hubert *et al.*, 2008), but in species such as *Aplochiton spp.* with diadromous and resident life histories, long-term isolation can cause substantial divergence between populations (McCusker & Bentzen, 2010; McDowall, 2010). We found that one non-synonymous mutation separated coastal from inland populations, with the exception of populations on the West Coast of Chiloé that displayed a mixture of both. The populations on Chiloé also displayed the highest haplotype diversity for CO-I. Population structuring based on mtDNA may reflect the pattern of colonisation following the LGM, when freshwater populations are thought to have derived from coastal refugia. A similar pattern of low population structuring has been observed for other diadromous and marine species that retreated to marine refugia during the LGM in this region (Fraser *et al.*, 2010; Zemlak *et al.*, 2010). In contrast, freshwater species occupying recently de-glaciated habitats display stronger divergence between separate watersheds, reflecting the recolonization from disconnected freshwater refugia (Ruzzante *et al.*, 2006; Unmack *et al.*, 2009).

Of all the populations examined, *A. zebra* displayed the highest genetic diversity on the Island of Chiloé, suggesting that this could have been a refugium for the species during the LGM (approximately 20,000 to 10,000 yrs BP), as it has also been suggested for other diadromous and marine species (Fraser *et al.*, 2010; Zemlak *et al.*, 2010). Recolonization

350 from coastal refugia and founder effects could explain the lack of diversity in the most
351 isolated population, such as the Valdivian populations and the Falkland Islands. During the
352 LGM, the Patagonian Ice Sheet spread northward from the southern tip of Patagonia (54°S)
353 to 38°S, and westward from the Andes mountains to the Pacific coastline (Cussac *et al.*,
354 2004); therefore, freshwater species could have migrated from the Andes to the coast and
355 from the south to unglaciated coastal regions in the north, resulting in *A. zebra* from coastal
356 rivers in Chile and the Falkland Islands sharing the same 16S rDNA haplotype. Surface
357 marine currents in the southern coast of Chile move southward with the Cape Horn Current,
358 around Tierra del Fuego and reach the Atlantic Ocean where the Antarctic Circumpolar
359 Current and the Malvinas Currents circumvent the Falkland Islands (Kaiser *et al.*, 2005).
360 Such currents have been found to be important dispersal pathways for Chilean seaweed
361 (Fraser *et al.*, 2010), diadromous *G. maculatus* (Zemlak *et al.*, 2010) and Chinook salmon
362 escaping from Chilean net pens and entering the South Atlantic around Cape Horn (Correa &
363 Gross, 2008). Based on these lines of evidence, a common origin for *A. zebra* in Chile and
364 the Falkland Islands is therefore plausible.

365 The low mtDNA diversity detected in *Aplochiton* spp. is common among many
366 freshwater fish, which typically display much lower diversity than marine fish (McCusker &
367 Bentzen, 2010). This would explain the genotypic uniformity observed between the River
368 Lenca (mainland Chile) and East Chiloé, connected by the Chiloé Interior Sea, and also the
369 low genetic diversity observed at West Chiloé. In contrast, the strong genetic differentiation
370 observed among landlocked populations suggests that inland populations are less likely to
371 migrate to sea following their recolonization from marine refugia, as observed in other
372 species (Waters *et al.*, 2010). Microsatellite data also indicated highly significant population
373 structuring with four distinct clusters: (1) Valdivia, (2) Lake Ranco, (3) East Chiloé with
374 River Lenca (i.e., populations inhabiting coastal rivers draining into the Chiloé Interior Sea)

and (4) West Chiloé (i.e., populations inhabiting coastal rivers draining into the Pacific Ocean). This level of regional genetic differentiation at microsatellite loci may reflect currently limited gene flow among watersheds and is supported by genetic barriers identified by BARRIER and IBD patterns.

In summary, our study of *A. zebra* in Chile and the Falklands indicates that the mtDNA diversity of this species probably reflects the legacy of historical routes of recolonization and migration, whereas the pattern of structuring found at microsatellite loci probably reflects the influence of more recent demographic and isolation processes in the region. Our analysis also indicates that the genetic diversity of *Aplochiton* may have decreased in those populations most affected by salmonid encroachment, particularly in the case of *A. taeniatus.* Although the possibility of ascertainment bias cannot be totally ruled out with current molecular markers used to study *Aplochiton* spp., the observed low genetic diversity, and the limited information available on their abundance and distribution calls for a more detailed analysis of the conservation status of *A. zebra* and *A. taeniatus* in the face of salmonid invasions.

**Acknowledgements**

Funding was provided by a DEFRA Darwin Initiative (Grant No. 162/15/020) and post-project award (Grant No. EIDPO-041) to CGL, GG and SC and an IBERS PhD studentship to DV. We are grateful to Kyle Young, Jessica Stephenson, Daniel Fowler and Gabriel Orellana for help with the collection of samples, to Nick Rendell and Brendan Gara for logistic support in the Falklands and to Ed Pope for commenting on the manuscript. Samples were collected under permits No. 958, 17 April 2008, Chilean Subsecretary of Fishing in Chile and licence No. R0221, Falkland Islands Government. Use of trade or firm names in

this document is for reader information only and does not constitute endorsement of any product or service by the U.S. Government.

401 **References**


402 Aggarwal, R., Allainguillaume, J., Bajay, M., Barthwal, S., Bertolino, P., Chauhan, P.,
403 Consuegra, S., Croxford, A., Dalton, D., den Belder, E., Diaz-Ferguson, E., Douglas,
404 M., Drees, M., Elderson, J., Esselink, G., Fernandez-Manjarres, J., Frascaria-Lacoste,
405 N., Gabler-Schwarz, S., Garcia de Leaniz, C., Ginwal, H., Goodisman, M., Guo, B.,
406 Hamilton, M., Hayes, P., Hong, Y., Kajita, T., Kalinowski, S., Keller, L., Koop, B.,
407 Kotze, A., Lalremruata, A., Leese, F., Li, C., Liew, W., Martinelli, S., Matthews, E.,
408 Medlin, L., Messmer, A., Meyer, E., Monteiro, M., Moyer, G., Nelson, R., Nguyen,
409 T., Omoto, C., Ono, J., Pavinato, V., Pearcy, M., Pinheiro, J., Power, L., Rawat, A.,
410 Reusch, T., Sanderson, D., Sannier, J., Sathe, S., Sheridan, C., Smulders, M.,
411 Sukganah, A., Takayama, K., Tamura, M., Tateishi, Y., Vanhaecke, D., Vu, N.,
412 Wickneswari, R., Williams, A., Wimp, G., Witte, V., Zucchi, M. & C, M.E.R.P.D.
413 (2011) Permanent Genetic Resources added to Molecular Ecology Resources
414 Database 1 August 2010-30 September 2010. *Molecular Ecology Resources*, **11**, 219-
415 222.

416 Alò, D., Correa, C., Arias, C. & Cárdenas, L. (2013) Diversity of Aplochiton Fishes
417 (Galaxiidea) and the Taxonomic Resurrection of A. marinus. *PLoS ONE*, **8**, e71577.

418 Arenas, M., Ray, N., Currat, M. & Excoffier, L. (2012) Consequences of Range Contractions
419 and Range Shifts on Molecular Diversity. *Molecular Biology and Evolution*, **29**, 207-
420 218.

421 Arismendi, I., Soto, D., Penaluna, B., Jara, C., Leal, C. & Leon-Munoz, J. (2009)
422 Aquaculture, non-native salmonid invasions and associated declines of native fishes
423 in Northern Patagonian lakes. *Freshwater Biology*, **54**, 1135-1147.

424 Arismendi, I., Penaluna, B.E., Dunham, J.B., de Leaniz, C.G., Soto, D., Fleming, I.A.,
425 Gomez-Uchida, D., Gajardo, G., Vargas, P.V. & León-Muñoz, J. (2014) Differential

426 invasion success of salmonids in southern Chile: patterns and hypotheses. *Reviews in*
427 *Fish Biology and Fisheries*, 1-23.
428 Bergman, P.S., Hansen, M.J. & Nate, N.A. (2011) Relationship between electrofishing catch
429 rate and adult trout abundance in Wisconsin streams. *North American Journal of*
430 *Fisheries Management*, **31**, 952-961.
431 Bonnet, E. & Van de Peer, Y. (2002) zt: a software tool for simple and partial Mantel tests.
432 *Journal of Statistical Software*, **7**, 1-12.
433 Carroll, S.P. (2011) Conciliation biology: the eco-evolutionary management of permanently
434 invaded biotic systems. *Evolutionary Applications*, **4**, 184-199.
435 Chen, C., Durand, E., Forbes, F. & Francois, O. (2007) Bayesian clustering algorithms
436 ascertaining spatial population structure: a new computer program and a comparison
437 study. *Molecular Ecology Notes*, **7**, 747-756.
438 Clavero, M. & García-Berthou, E. (2005) Invasive species are a leading cause of animal
439 extinctions. *Trends in Ecology & Evolution*, **20**, 110-110.
440 Consuegra, S., Phillips, N., Gajardo, G. & Garcia de Leaniz, C. (2011) Winning the invasion
441 roulette: escapes from fish farms increase admixture and facilitate establishment of
442 non-native rainbow trout. *Evolutionary Applications*, **4**, 660-671.
443 Cornuet, J.-M., Santos, F., Beaumont, M.A., Robert, C.P., Marin, J.-M., Balding, D.J.,
444 Guillemaud, T. & Estoup, A. (2008) Inferring population history with DIY ABC: a
445 user-friendly approach to approximate Bayesian computation. *Bioinformatics*, **24**,
446 2713-2719.
447 Correa, C. & Gross, M.R. (2008) Chinook salmon invade southern South America. *Biological*
448 *Invasions*, **10**, 615-639.
449 Crawford, N.G. (2010) smogd: software for the measurement of genetic diversity. *Molecular*
450 *Ecology Resources*, **10**, 556-557.

Cussac, V., Ortubay, S., Iglesias, G., Milano, D., Lattuca, M.E., Barriga, J.P., Battini, M. & Gross, M. (2004) The distribution of South American galaxiid fishes: the role of biological traits and post-glacial history. *Journal of Biogeography*, **31**, 103-121.

Do, C., Waples, R.S., Peel, D., Macbeth, G., Tillett, B.J. & Ovenden, J.R. (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, **14**, 209-214.

Emerson, B.C. & Hewitt, G.M. (2005) Phylogeography. *Current Biology*, **15**, R367-R371.

Excoffier, L. & Lischer, H.E.L. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564-567.

Fraser, C.I., Thiel, M., Spencer, H.G. & Waters, J.M. (2010) Contemporary habitat discontinuity and historic glacial ice drive genetic divergence in Chilean kelp. *Bmc Evolutionary Biology*, **10**

Garcia de Leaniz, C., Gajardo, G. & Consuegra, S. (2010) From best to pest: changing perspectives on the impact of exotic salmonids in the Southern Hemisphere. *Systematics and Biodiversity*, **8**, 447-459.

Goudet, J. (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. *J Hered*, **86**, 485-486.

Hall, T. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95-98.

Hemmer-Hansen, J., Nielsen, E.E., Frydenberg, J. & Loeschcke, V. (2007) Adaptive divergence in a high gene flow environment: Hsc70 variation in the European flounder (Platichthys flesus L.). *Heredity*, **99**, 592-600.

476 Habit, E., Gonzalez, J., Ruzzante, D.E. & Walde, S.J. (2012) Native and introduced fish
477 species richness in Chilean Patagonian lakes: inferences on invasion mechanisms
478 using salmonid-free lakes. *Diversity and Distributions*, **18**, 1153-1165.
479 Hubert, N., Hanner, R., Holm, E., Mandrak, N.E., Taylor, E., Burridge, M., Watkinson, D.,
480 Dumont, P., Curry, A., Bentzen, P., Zhang, J.B., April, J. & Bernatchez, L. (2008)
481 Identifying Canadian Freshwater Fishes through DNA Barcodes. *Plos One*, **3**
482 Jackobsson, M. & Rosenberg, N.A. (2007) CLUMPP: a cluster matching and permutation
483 program for dealing with label switching and multimodality in analysis of population
484 structure *Bioinformatics*, **23**, 1801-1806.
485 Jost, L. (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**,
486 4015-4026.
487 Kaiser, J., Lamy, F. & Hebbeln, D. (2005) A 70-kyr sea surface temperature record off
488 southern Chile (Ocean Drilling Program Site 1233). *Paleoceanography*, **20**, 15.
489 Lattuca, M.E., Battini, M.A. & Macchi, P.J. (2008) Trophic interactions among native and
490 introduced fishes in a northern Patagonian oligotrophic lake. *Journal of Fish Biology*,
491 **72**, 1306-1320.
492 Librado, P. & Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA
493 polymorphism data. *Bioinformatics*, **25**, 1451.
494 Manni, F., Guerard, E. & Heyer, E. (2004) Geographic patterns of (genetic, morphologic,
495 linguistic) variation: How barriers can be detected by using Monmonier's algorithm.
496 *Human Biology*, **76**, 173-190.
497 McCusker, M.R. & Bentzen, P. (2010) Positive relationships between genetic diversity and
498 abundance in fishes. *Molecular Ecology*, **19**, 4852-4862.

McDowall, R.M. (2006) Crying wolf, crying foul, or crying shame: alien salmonids and a biodiversity crisis in the southern cool-temperate galaxioid fishes? *Reviews in Fish Biology and Fisheries*, **16**, 233-422.

McDowall, R.M. (2010) *New Zealand Freshwater Fishes: an Historical and Ecological Biogeography*.

Miller, M.P. (1997) Tools for population genetic analyses (TFPGA) 1.3: A Windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by author.

Monmonier, M. (1973) Maximum difference barriers - Alternative numerical regionalization method. *Geographical Analysis*, **5**, 245-261.

Monzón-Argüello, C., Garcia de Leaniz, C., Gajardo, G. & Consuegra, S. (2013) Less can be more: loss of MHC functional diversity can reflect adaptation to novel conditions during fish invasions. *Ecology and Evolution*, **3**, 3359–3368.

Monzón-Argüello, C., Garcia de Leaniz, C., Gajardo, G. & Consuegra, S. (2014a) Eco-immunology of fish invasions: the role of MHC variation. *Immunogenetics*, **66**, 393-402.

Monzón-Argüello, C., Consuegra, S., Gajardo, G., Marco-Rius, F., Fowler, D.M., DeFaveri, J. & Garcia de Leaniz, C. (2014b) Contrasting patterns of genetic and phenotypic differentiation in two invasive salmonids in the southern hemisphere. *Evolutionary Applications*, **7**, 921–936.

Palumbi, S.R., Martin, A.P., Romano, S., McMillan, W.O., Stice, L. & Grabowski, G. (1991) The simple fool's guide to PCR, ver. 2.0. *Published and distributed by the authors, Honolulu.*,

Parker, I.M., Simberloff, D., Lonsdale, W., Goodell, K., Wonham, M., Kareiva, P., Williamson, M., Von Holle, B., Moyle, P. & Byers, J. (1999) Impact: toward a

framework for understanding the ecological effects of invaders. *Biological Invasions*, **1**, 3-19.

Piry, S., Luikart, G. & Cornuet, J.M. (1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, **90**, 502-503.

Raymond, M. & Rousset, F. (1995) GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *J Hered*, **86**, 248-249.

Ricciardi, A., Palmer, M.E. & Yan, N.D. (2011) Should Biological Invasions Be Managed as Natural Disasters? *Bioscience*, **61**, 312-317.

Rice, W.R. (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223-225.

Rosenberg, N.A. (2004) distruct: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137-138.

Ruzzante, D., Walde, S., Cussac, V., Dalebout, M., Seibert, J., Ortubay, S. & Habit, E. (2006) Phylogeography of the Percichthyidae (Pisces) in Patagonia: roles of orogeny, glaciation, and volcanism. *Molecular Ecology*, **15**, 2949 - 2968.

Schröder, V. & Garcia de Leaniz, C. (2011) Discrimination between farmed and free-living invasive salmonids in Chilean Patagonia using stable isotope analysis. *Biological Invasions* **13**, 203-213.

Spiegelhalter, D.J., Best, N.G., Carlin, B.R. & van der Linde, A. (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, **64**, 583-616.

Strauss, S.Y., Lau, J.A. & Carroll, S.P. (2006) Evolutionary responses of natives to introduced species: what do introductions tell us about natural communities? *Ecology Letters*, **9**, 354-371.

548 Strayer, D.L., Eviner, V.T., Jeschke, J.M. & Pace, M.L. (2006) Understanding the long-term
549 effects of species invasions. *Trends in Ecology & Evolution*, **21**, 645-651.
550 Traill, L.W., Brook, B.W., Frankham, R.R. & Bradshaw, C.J.A. (2010) Pragmatic population
551 viability targets in a rapidly changing world. *Biological Conservation*, **143**, 28-34.
552 Unmack, P.J., Bennin, A.P., Habit, E.M., Victoriano, P.F. & Johnson, J.B. (2009) Impact of
553 ocean barriers, topography, and glaciation on the phylogeography of the catfish
554 Trichomycterus areolatus (Teleostei: Trichomycteridae) in Chile. *Biological Journal*
555 *of the Linnean Society*, **97**, 876-892.
556 Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M. & Shipley, P. (2004) MICRO-
557 CHECKER: software for identifying and correcting genotyping errors in
558 microsatellite data. *Molecular Ecology Notes*, **4**, 535-538.
559 Vanhaecke, D., Garcia de Leaniz, C., Gajardo, G., Thomas, C.J. & Consuegra, S. (2012a)
560 Metapopulation dynamics of a diadromous galaxiid fish and potential effects of
561 salmonid aquaculture. *Freshwater Biology*, **57**, 1241-1252.
562 Vanhaecke, D., Garcia de Leaniz, C., Gajardo, G., Young, K., Sanzana, J., Orellana, G.,
563 Fowler, D., Howes, P., Monzon-Arguello, C. & Consuegra, S. (2012b) DNA
564 Barcoding and Microsatellites Help Species Delimitation and Hybrid Identification in
565 Endangered Galaxiid Fishes. *PLoS ONE*, **7**, e32939.
566 Wang, J.L. (2009) A new method for estimating effective population sizes from a single
567 sample of multilocus genotypes. *Molecular Ecology*, **18**, 2148-2164.
568 Waples, R.S. & Do, C. (2008) ldne: a program for estimating effective population size from
569 data on linkage disequilibrium. *Molecular Ecology Resources*, **8**, 753-756.
570 Waters, J.M., Rowe, D.L., Burridge, C.P. & Wallis, G.P. (2010) Gene Trees versus Species
571 Trees: Reassessing Life-History Evolution in a Freshwater Fish Radiation. *Systematic*
572 *Biology*, **59**, 504-517.

573 Young, K., Dunham, J., Stephenson, J., Terreau, A., Thailly, A., Gajardo, G. & Garcia de
574 Leaniz, C. (2010) A trial of two trouts: comparing the impacts of rainbow and brown
575 trout on a native galaxiid. *Animal Conservation*, 1-10.
576 Young, K.A., Stephenson, J., Terreau, A., Thailly, A.-F., Gajardo, G. & Garcia de Leaniz, C.
577 (2009) The diversity of juvenile salmonids does not affect their competitive impact on
578 a native galaxiid. *Biological Invasions*, **11**, 1955-1961.
579 Zemlak, T., Habit, E., Walde, S., Carrea, C. & Ruzzante, D. (2010) Surviving historical
580 Patagonian landscapes and climate: molecular insights from Galaxias maculatus. *Bmc*
581 *Evolutionary Biology*, **10**, 67.

582 **Data accesibility**

583 Alignments of the 16S rDNA results generated in this study are publicly available in figshare,
584 http://dx.doi.org/10.6084/m9.figshare.1423279.

585 **Biosketch**

586 This work was part of the PhD thesis of DV, supervised by SC and CGL on salmonid
587 invasion biology in the Southern Hemisphere. SC and CGL conceived the work; DV, CGL,
588 GGa, JD and GGi collected the samples, DV and SC analysed the data, all authors
589 contributed to the writing of the manuscript.

590 **Editor:** Robert Bryson Jr.

591 **Appendix S1.** Genetic diversity and differentiation among *Aplochiton zebra* populations.

592 **Appendix S2.** Barriers to gene flow among *Aplochiton zebra* populations in Chile identified
593 by BARRIER v2.2.

594 **Appendix S3.** Alternative scenarios for ABC analysis population colonisation of *Aplochiton*
595 *zebra* in Chile and the Falklands.

596

**Table 1.** Characteristics and location of study rivers, sample size of *Aplochiton zebra* (n) in Chile, presence o abundance (CPUE, fish/min/$m^2$) of native *A. zebra.* and salmonids(*).

| Watershed | Code | River name | n | Latitude | Longitude | Altitude (m) | Origin | Species present |
|---|---|---|---|---|---|---|---|---|
| W. Chiloé | U28 | N/A | 22 | -41.95 | -74.02 | 10 | Coastal | AZ, AT |
| | U29 | N/A | 21 | -41.96 | -74.04 | 18 | Coastal | AZ |
| E. Chiloé | U17 | N/A | 20 | -42.12 | -73.48 | 10 | Coastal | AZ, GM, OM* |
| | U20 | N/A | 26 | -42.21 | -73.40 | 9 | Coastal | AZ, GM, OM* |
| | U33 | N/A | 25 | -42.17 | -73.48 | 4 | Coastal | AZ, AT, GM, OM*, SS* |
| | U34 | N/A | 16 | -42.11 | -73.48 | 11 | Coastal | AZ, AT, BB, GM, OM* |
| Reloncaví | RL | Lenca | 17 | -41.61 | -72.68 | 13 | Pre-Andean | AZ, GM, OM*, ST* |
| Valdivia | REN | Blanco | 30 | -39.91 | -72.15 | 139 | Andean | AZ, OM*, ST* |
| | RPU | Punahue | 21 | -39.83 | -72.04 | 195 | Andean | AZ, OM*, ST* |
| L. Ranco | RPI | Pitreño | 30 | -40.33 | -72.32 | 82 | Central Valley | AZ, GP, PT, TA, BA, C |
| | RQU | Quiman | 27 | -40.11 | -72.34 | 118 | Central Valley | AZ, GP, OM*, SS*, ST* |
| | RFU | Futangue | 30 | -40.33 | -72.27 | 71 | Central Valley | AZ, GP, PT, OM* |
| | RIC | Iculpe | 30 | -40.32 | -72.44 | 89 | Central Valley | AZ, GP, TA, PT, OM*, |

AZ= *Aplochiton. zebra*; AT = *A. taeniatus;* GM = *Galaxias maculatus*; BB = *Brachygalaxias bullocki;* BA=*Basilichth*

TA=*Trichomycterus areolatus;* PT = *Percichtys trucha;* GP = *G. platei;* OM = *Oncorhynchus mykiss*; ST = *Salmo trut*

601 **Table 2.** Effective population size ($N_e$) and 95% confidence intervals estimated with
602 COLONY (full likelihood score method) and NeEstimator (Linkage Disequilibrium, LD
603 method) for populations with sample size (N) of at least 20 individuals and based on 11
604 microsatellites.

| | COLONY | | | | NeESTIMATOR | |
|---|---|---|---|---|---|---|
| | N | $N_e$ | 0.05 | 0.95 | $N_e$ | CI (Jacknife method) |
| U28 | 22 | 66 | 35 | 205 | 40.6 | (22.7, 122.0) |
| U29 | 21 | 127 | 44 | ∞ | ∞ | (78.6, ∞) |
| U17 | 20 | 84 | 37 | ∞ | ∞ | (88.5, ∞) |
| U20 | 26 | 71 | 39 | 190 | 31.3 | (19.3, 63.2) |
| U33 | 25 | 54 | 31 | 115 | 608.4 | (68.5, ∞) |
| REN | 30 | 83 | 47 | 173 | 172.2 | (47.0, ∞) |
| RPU | 21 | 47 | 25 | 129 | 49.3 | (25.5, 223.5) |
| RPI | 30 | 58 | 34 | 101 | 332.6 | (85.3, ∞) |
| RQU | 27 | 71 | 41 | 136 | ∞ | (102.0, ∞) |
| RFU | 30 | 64 | 39 | 115 | 226.5 | (60.7, ∞) |
| RIC | 30 | 72 | 40 | 137 | 194.2 | (64.5, ∞) |

605

606 **Table 3.** Molecular diversity indices per population based on mitochondrial 16S
607 rDNA and CO-I genes calculated in ARLEQUIN (sample size (N), number of
608 haplotypes (h), haplotype diversity ($H_D$) and nucleotide diversity ($\pi$)) for each river
609 separately and pooled by region.

| **River** | **16S rDNA** | | | | **CO-I** | | | |
|---|---|---|---|---|---|---|---|---|
| | **N** | **h** | **$H_D$** | **$\pi$** | **N** | **H** | **$H_D$** | **$\pi$** |
| U28 | 11 | 3 | 0.56 | 0.0013 | 10 | 1 | 0 | 0 |
| U29 | 8 | 2 | 0.57 | 0.0012 | 9 | 1 | 0 | 0 |
| U17 | 6 | 3 | 0.6 | 0.0014 | 15 | 5 | 0.79 | 0.0024 |
| U20 | 5 | 1 | 0 | 0 | 14 | 3 | 0.39 | 0.0009 |
| U33 | 5 | 2 | 0.4 | 0.0008 | 11 | 4 | 0.69 | 0.0018 |
| U34 | 4 | 2 | 0.5 | 0.0010 | 7 | 2 | 0.29 | 0.0006 |
| RL | 10 | 2 | 0.36 | 0.0007 | 16 | 1 | 0 | 0 |
| REN | 10 | 1 | 0 | 0 | 30 | 3 | 0.25 | 0.0006 |
| RPU | 10 | 1 | 0 | 0 | 21 | 1 | 0 | 0 |
| RPI | 5 | 2 | 0.4 | 0.0008 | 24 | 1 | 0 | 0 |
| RQU | 5 | 2 | 0.4 | 0.0008 | 18 | 1 | 0 | 0 |
| RFU | 4 | 1 | 0 | 0 | 27 | 2 | 0.21 | |
| RIC | 5 | 2 | 0.4 | 0.0008 | 19 | 1 | 0 | 0 |
| **REGIONS** | | | | | | | | |
| West Chiloé | 19 | 3 | 0.56 | 0.0013 | 19 | 1 | 0 | 0 |
| East Chiloé and RL | 30 | 3 | 0.35 | 0.0008 | 63 | 5 | 0.49 | 0.0012 |
| Valdivia | 20 | 1 | 0 | 0 | 50 | 3 | 0.15 | 0.0003 |
| Lake Ranco | 19 | 3 | 0.29 | 0.0006 | 88 | 2 | 0.07 | 0.0001 |
| Falkland Islands | 10 | 1 | 0 | 0 | 23 | 1 | 0 | 0 |

610

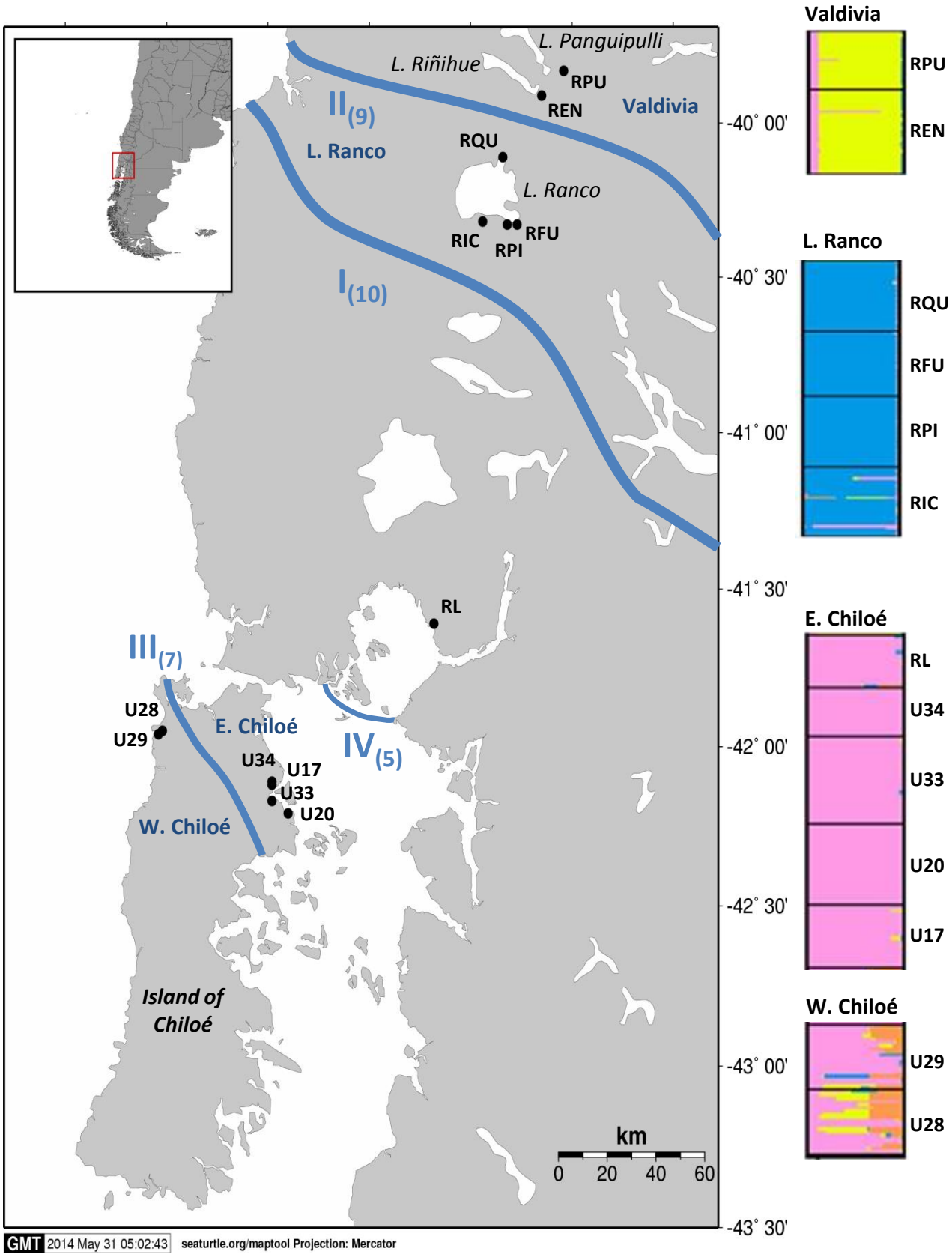
611 **Figure legends**

612 **Figure 1.** Representation of genetic barriers (left) between galaxiid fish populations
613 (*Aplochiton zebra*) in Chile identified using BARRIER v 2.2, based on 11 microsatellites
614 using Monmonier's algorithm. Blue lines represent the main barriers to gene flow, ranked
615 from 1$^{st}$ to 4$^{th}$ (number of loci supporting the barrier is indicated in brackets next to ranking),
616 thickness of the line represents bootstrapping support. Plot of individual assignment (right)
617 based on 11 microsatellites using TESS with K = 4, averaged from 100 runs using CLUMPP
618 and represented using DISTRUCT. Each bar constitutes an individual fish. Y- axis represents
619 the proportion of each individual attributable to each cluster, represented by coloured bars.
620

621 **Figure 2.** Relationship between geographical distance (km) and genetic distance ($F_{ST}$/1- $F_{ST}$)
622 for 13 fish populations of *Aplochiton zebra* in Chile ($y = 0.0157\ln(x) - 0.0164$, $R^2 = 0.462$).

623 **Figure 3.** Distribution of mitochondrial 16S rDNA haplogroups of galaxiid fish (*Aplochiton*
624 *zebra*) in the sampling regions: Valdivia, L. Ranco, West Chiloé, East Chiloé (including Seno
625 Reloncavi) and Falkland Islands .

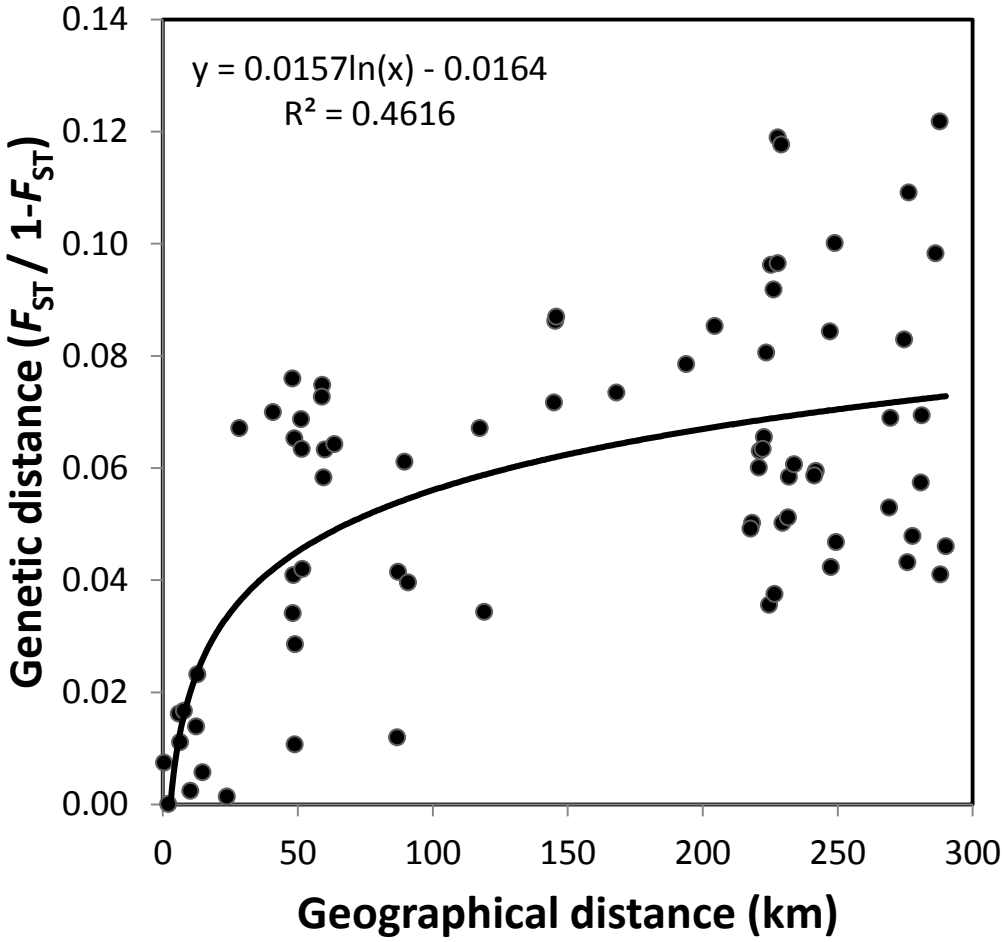
626 **Figure 1.**




627

628

629 **Figure 2.**

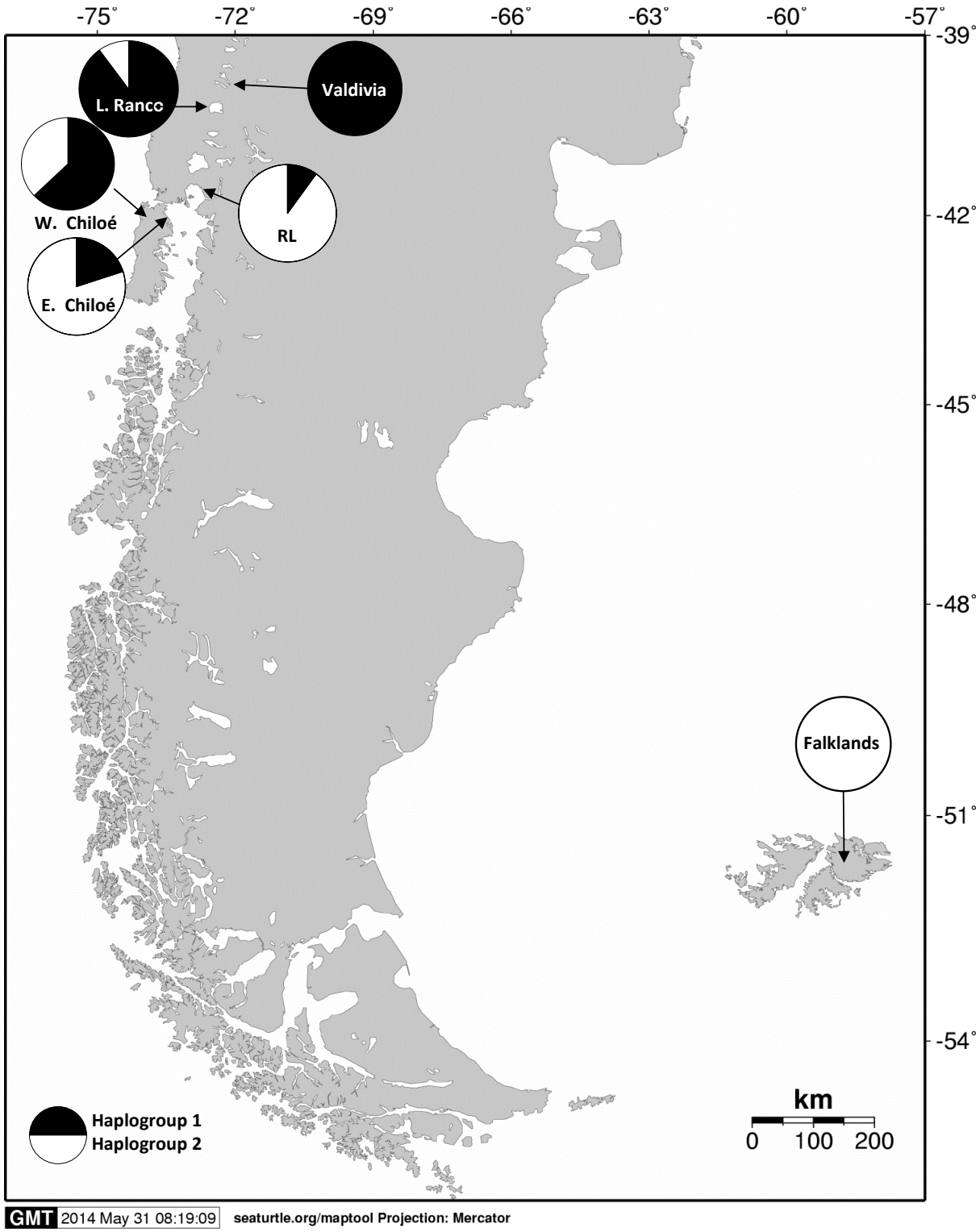



630

631

632 **Figure 3.**

633




634

635

636