Zoological Studies **55:** 16 (2016)
doi:10.6620/ZS.2016.55-16

Zoological Studies

Low Genetic Diversity in *Diplomystes camposensis*, an Endemic and Endangered Catfish from South Chile

Carlos P. Muñoz-Ramírez[1,*], Evelyn Habit[2], Peter J. Unmack[4], Jerald B. Johnson[5], and Pedro F. Victoriano[1,3]

[1]*Departamento de Zoología, Facultad de Ciencias Naturales y Oceanográficas. Universidad de Concepción, Casilla 160-C, Concepción, Chile*

[2]*Facultad de Ciencias Ambientales y Centro EULA-Chile, Universidad de Concepción, Chile*

[3]*Centro de Investigación en Ecosistemas Patagónicos (CIEP), Coyhaique, Chile*

[4]*Institute for Applied Ecology and Collaborative Research Network for Murray-Darling Basin Futures, University of Canberra, ACT 2601, Australia*

[5]*Department of Biology and Monte L. Bean Life Science Museum, Brigham Young University, Provo, UT 84602, USA*

(Received 30 November 2014; Accepted 24 December 2015)

Carlos P. Muñoz-Ramírez, Evelyn Habit, Peter J. Unmack, Jerald B. Johnson, and Pedro F. Victoriano (2016) Despite the fundamental importance of the family Diplomystidae for understanding catfish evolution, its species are poorly known and most of them endangered. *Diplomystes camposensis*, restricted to a single river basin in southern Chile, is perhaps the most vulnerable species due to its small geographic range and imminent habitat alterations by dam constructions. Using mitochondrial DNA sequences, we describe the genetic diversity across its entire distribution in the Valdivia basin and test hypotheses related to the impact of glacial cycles on the genetic diversity and structure. We found that *Diplomystes camposensis* has low genetic diversity and structure across the entire Valdivia basin along with a pattern of decreasing nucleotide and haplotype diversity from West to East. Demographic analyses showed evidence of population expansion in agreement with the glaciated history of the basin. Analyses of population structure showed no evidence of population subdivision. However, coalescent analyses indicated that very recent subdivision (in the last 50 years) cannot be ruled out. Low genetic diversity and genetic structure across the entire basin suggest that the species might be highly vulnerable to habitat fragmentation. Thus, the imminent construction of hydropower dams represents a serious threat to its conservation. Our results suggest that the low genetic diversity can be the product of the glaciated history of the basin, although the influence of species-specific biological traits may also add to this condition. Despite the overall low genetic diversity, higher diversity was found in the central portion of the basin suggesting high priority of conservation for this area as it might be used as a source population in case translocations are required among potential management plans.

Key words: Diplomystidae, Valdivia Basin, Genetic diversity, Genetic structure, Coalescent simulations, Pleistocene Glaciations, Conservation.

BACKGROUND

The family Diplomystidae, endemic to freshwater systems from Southern South America, is one of the earliest branching lineages in the highly diverse order Siluriformes (Lundberg and Baskin 1969; Lundberg and Case 1970; Arratia 1987; Grande 1987; Fink and Fink 1996; Pinna 1998; Sullivan et al. 2006). Unfortunately, despite its importance in terms of phylogenetic biodiversity and for understanding catfish evolution, most species face conservation problems (SERNAPESCA 2008, Bello and Ubeda 1998). Which is worse, although detailed studies on

*Correspondence: Current address: Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA. TEL: +1-225-456-4561. E-mail: carmunoz@umich.edu

comparative anatomy has been conducted in these taxa (Arratia 1987; Arratia 1992; Arratia and Huaquín 1995), there is still a lack of knowledge on basic aspects of their biology (*e.g.* reproductive ecology, population genetics, behaviour; Arratia 1983) which prevents informed conservation strategies. A recent broad-scale phylogeographic analysis has shown that a large portion of the phylogenetic diversity in the family is found at the west of the Andes, with unique genetic lineages in most drainages, suggesting the need for high conservation priorities at broad scale (Muñoz-Ramírez et al. 2014). However, none study has focused on the analysis of genetic diversity within a single river basin, which would be important for local management.

Among all six recognized species within the family, *Diplomystes camposensis* Arratia, 1987 has been recognized as the most geographically restricted, being only recorded in the Valdivia Basin, South Chile (Arratia 1987; Muñoz-Ramírez et al. 2010). A recent phylogeographic study has revealed closely related lineages in the neighbouring basins of Toltén and Imperial (Muñoz-Ramírez et al. 2014). However, until further studies can address whether these lineages belong to *D. camposensis*, we will focus on populations of *D. camposensis* from the Valdivia basin as these populations are genetically well differentiated and isolated geographically. Hereafter we will refer to these populations as *D. camposensis*. Within the Valdivia basin, *D. camposensis* most frequently prefers upstream, fast-flowing waters (Habit et al. 2009), which restricts suitable habitat to the eastern portion of the basin. Ongoing projects to build hydropower dams in areas where the species is more abundant threaten to further reduce habitat availability and cause fragmentation. Previous studies on a reduced number of populations have suggested high levels of gene flow and a large home range (Habit et al. 2009; Victoriano et al. 2012), implying that further habitat alterations might pose a high conservation risk for the species.

The Valdivia basin is located in an area that has been subjected to high environmental instability, mainly caused by volcanic activity and Pleistocene glaciations (*e.g.* Le Roux and Elgueta 2000; Hulton et al. 2002). During the last glacial maximum (LGM; 23000 ybp; Ruzzante et al. 2006), the Valdivia basin was partially covered by ice on its eastern portion, restricting habitat available for freshwater species to the western portions of the basin (Fig. 1).

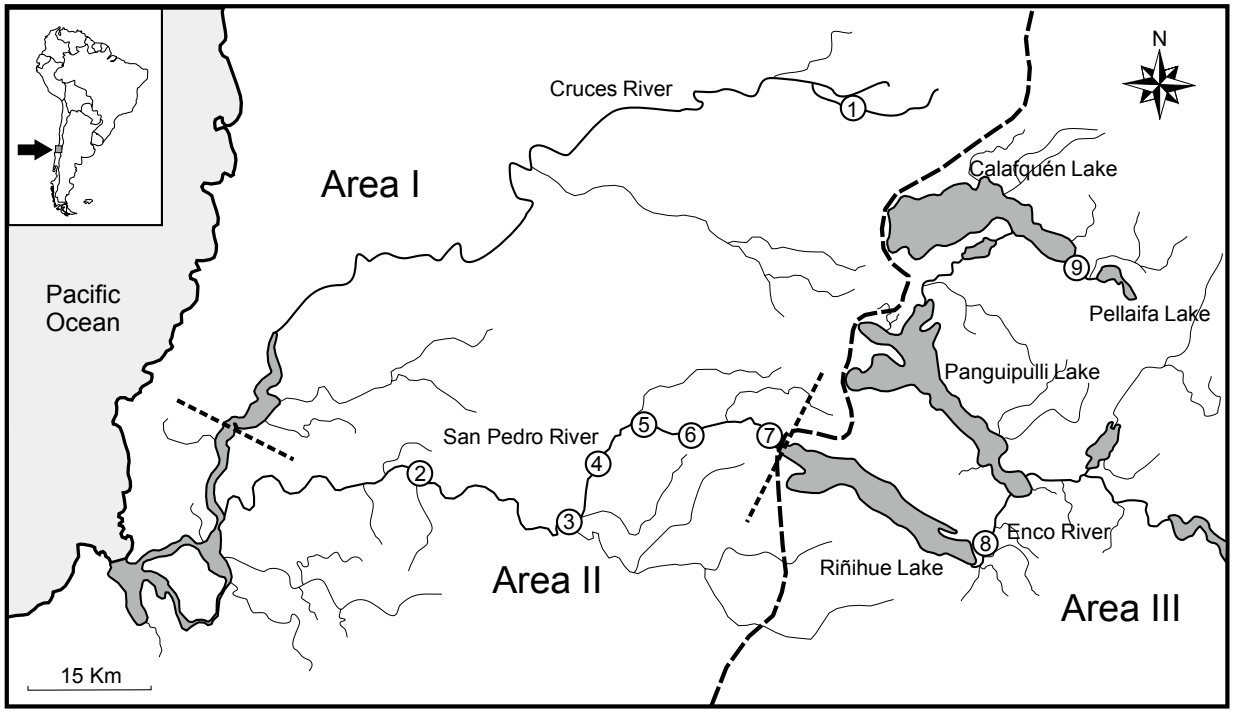

Fig. 1. Map showing location sampling and defined Areas within the Valdivia Basin. Main Areas are separated by short straight dashed lines. Population codes (numbers) are described in Table 1. Long dashed line indicates the western limit of the LGM ice sheet based on Hulton et al. (2002).

In this scenario, populations of *D. camposensis* might have undergone demographic bottlenecks due to a considerable reduction of the drainage area, particularly at the eastern portion of the basin. Previous phylogeographic studies have revealed a role of glaciations in shaping broad scale patterns of genetic diversity in freshwater organisms (*e.g.* Ruzzante et al. 2006; Zemlak et al. 2008; Xu et al. 2009; Muñoz-Ramírez et al. 2014). However, patterns of genetic diversity in relation to glaciations have rarely been analysed at the smaller scale of a single basin (Victoriano et al. 2012), despite its analytical advantages (*i.e.* free from confounding factors derived from inter-basin comparison analyses).

In this study, we analyse sequences of two mitochondrial genes from individuals that encompass an unprecedented geographical coverage in the Valdivia basin to fully describe patterns of genetic diversity in *Diplomystes camposensis*. In addition, we conduct a number of analyses to investigate the role of recent glaciations on patterns of genetic diversity, including a model-based coalescent approach to test for population subdivision and explore the power of the data to distinguish between ancestral polymorphism and current gene flow.

MATERIALS AND METHODS

Sampling

We analysed mitochondrial DNA sequences from 62 individuals of *D. camposensis* collected from all known populations in the Valdivia basin (Fig. 1, Table 1). Thirty of these sequences were also analysed in the context of a broader phylogeographic study of the family Diplomystidae (Muñoz-Ramirez et al. 2014). We did not collect new specimens for the specific purpose of this study given the conservation status of the species. Instead, we analysed specimens already deposited in the collection of the EULA Centre (Universidad de Concepción, Chile) that were previously collected for different ecological studies (*e.g.* Habit et al. 2009; García et al. 2012; Cifuentes et al. 2012; Colin et al. 2012; Montoya et al. 2012; Beltrán-Concha et al. 2012). The inclusion of only one site from the Cruces River is due to the limited distribution of *D. camposensis* in this river. Considerable sampling effort in other sites of the Cruces River was unable to locate additional populations. For the purpose of describing and analysing patterns of genetic diversity across the basin, we divided the basin in three main Areas (see Fig. 1): Area I representing the presumably isolated population from Rio Cruces (population 1); Area II containing populations downstream of the Lago Riñihue, a hypothetically stable area (populations 2-7); and Area III consisting of populations upstream of Riñihue Lake (populations 8 and 9), an area that was glaciated at the LGM. Some potential barriers to dispersal are found between these areas. Between areas I and II, there are wetlands that may represent unsuitable environments for Diplomystids (*e.g.* low flow water, low oxygen concentration, and increased turbidity and water temperature). Between areas II and III there is a large and deep lake from which no captures have been reported (except for captures close to where it discharges), and have led to the idea that this species might not inhabit lakes (Habit et al. 2009). All specimens studied are stored at 95% Ethanol in the Fish Collection of the EULA Center, University of Concepción, Chile.

Table 1. Locality data and voucher codes for the individuals analysed in this study. Site numbers refer to sites in map (Fig. 1)

site	locality	latitude	longitude	individual codes
1	Chesque	-39.3938	-72.3635	Ches06-07,09-10; DCru240-248
2	Antilhue	-39.8028	-72.9609	DcZ416-19, 421-24
3	Cuyincahuin	-39.8532	-72.7524	DMa43,54,60,82; DcZ415
4	Chacaipulli	-39.7881	-72.7111	DMa30,55,56,62,63,66,67,74
5	Malihue Viejo	-39.7450	-72.6577	DMa68
6	La Quinta	-39.7631	-72.5841	DMa29,31,33,34,42,53,61,65
7	Trafún Este	-39.7502	-72.4856	DMa39,40,72,75,77,85
8	Balsa Enco	-39.9012	-72.1520	DEn209-216; DcE01-03
9	Pellaifa River	-39.5875	-72.0163	DCal259-60

Laboratory protocols

We extracted genomic DNA from muscle tissue using the DNeasy Tissue Kit (QIAGEN Inc., Chatsworth CA). Besides including more populations and individuals, we have extended the marker length relative to one previous study (Victoriano et al. 2012), including the genes d-loop and Cyt *b* to increase resolution. Generally, increasing the length of the alignment improve the precision of phylogenetic estimation, a fundamental step for improving parameter estimation (Heled and Drummond 2008). We followed PCR protocols described in Muñoz-Ramírez et al. (2014). PCR products were examined on a 1% agarose gel using SYBR safe DNA gel stain (Invitrogen, Eugene, OR, USA) and purified using a Montage PCR 96 plate (Millipore, Billerica, MA, USA). Sequences were also obtained via cycle sequencing with Big Dye 3.0 dye terminator ready reaction kits using 1/16th reaction size (Applied Biosystems, Foster City, CA). Sequencing reactions were run with an annealing temperature of 52°C following the ABI manufacturer's protocol. We purified sequenced products using sephadex columns. Sequences were obtained using an Applied Biosystems 3730 XL automated sequencer at the Brigham Young University DNA Sequencing Centre. Sequences used in this study are deposited in GenBank (accession numbers: JX648805-JX648834, JX649006-JX649035, and KT886925-KT886988).

Sequence editing and phylogenetic analyses

Chromatograms were edited in CodonCode Aligner 3.0.3 (Dedham, MA, USA). Sequences were imported to BioEdit 7.0.5.2 (Hall 1999) and aligned by eye. Cyt *b* sequences were checked via amino acid coding in MEGA5 (Tamura et al. 2011) for unexpected frame shift errors or stop codons. All downstream analyses were performed with the two genes concatenated since they do not segregate independently and have shown similar levels of genetic diversity in several studies (*e.g.* Burridge et al. 2008; Unmack et al. 2009; Unmack et al. 2013). Phylogenetic relationships were reconstructed by Bayesian inference using MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003). We included each mitochondrial region (cyt *b* and d-loop) as separate partitions in the Bayesian analysis. We identified the best-fitting model of molecular evolution using jModeltest 0.1.1 (Posada 2008). Under the Bayesian Information criterion the model for cyt *b* and d-loop was HKY+I and GTR+I+G, respectively. We conducted two independent runs and its convergence was examined by the average standard deviation of split frequencies (*i.e.* values below 0.01) and the potential scale reduction factor (PSRF; values should approach 1). These two indicators are implemented in MrBayes with the PRSF available through the command *sump*. In addition, we also checked for topological convergence with the program AWTY (Nylander et al. 2008) using the plotting tool *compare*. Topological convergence in this test is confirmed by the observation of values tightly clustered to the diagonal of the plot (see Additional file Figure_S1). Four chains were used for phylogeny estimation, starting with a random tree and running for 5,000,000 generations, sampling every 1000 trees. The initial 10% of the resulting trees was discarded as burn-in using the "sumt" command. Once convergence of the two independent runs was confirmed, parameter estimates from both runs were combined to obtain a total of 9002 trees. A 50% majority rule consensus tree with branch lengths was constructed. Phylogenetic analysis were also conducted by maximum likelihood inference in the program RAxML (Stamatakis 2014). We used the same models used in the Bayesian inference as RAxML allows the use of multiple partitions with different models. Bootstrap values were obtained by the rapid search option with 200 replicates. All other settings were kept as default. In addition to the estimation of trees, we also constructed a haplotype network using the median joining algorithm implemented in Network 4.610 (Bandelt et al. 1999). Haplotypes were inferred from DNASP 5.0 before their use in Network (Librado and Rozas 2009).

Demography, genetic diversity and genetic structure

In order to test for population expansion, we calculated the Tajima's *D* and Fu's *Fs* statistics of neutrality (Fu 1997) in Arlequin 3.5.1.3 (Excoffier et al. 2005). These statistics were only calculated for the entire basin based on the evidence of a single population and the lack of reciprocally monophyletic groups (see the Results section). In addition, we also conducted and plotted a mismatch distribution analysis for which a unimodal distribution is interpreted as evidence of population expansion, while a multimodal distribution supports a constant population size over time.

The raggedness and SSD indexes provide a statistical assessment of the modality of this plot, where significant values reject the hypothesis of expansion. This analysis was also conducted in Arlequin. We used DnaSP to calculate haplotype richness (S), haplotype diversity (H), and nucleotide diversity (π) that were later used to analyse patterns of genetic diversity in regard to the hypothesis of glacial impact. A General Linear Model (GLM) was used to test for correlations between longitude and genetic diversity, since the glacial hypothesis predicts a decrease in diversity in a West-East direction. This analysis was conducted in R (R Development Core Team 2013). In addition, R was also used to test whether the genetic diversity from the glaciated Area III (expected to be lower than other areas of the basin) was lower than expected by chance. This was done by randomly sampling 13 sequences from the total pool of 62 sequences and calculating its nucleotide diversity. This was repeated 10000 times to generate a frequency distribution of random π that was then compared with the observed value of π for the Area III. If the observed value falls within the 5% of the lowest values in the distribution, it is considered significantly lower than expected by chance and consequently as evidence of recent colonization. To investigate patterns of genetic structure between populations, pairwise F_{ST} values were calculated in Arlequin based on *p*-distance, with 1000 randomizations to test for significance. A Bonferroni correction was applied to account for multiple comparisons. Finally, to analyse changes in the effective population size through time and to investigate the timing of a potential past bottleneck, we conducted a Bayesian Skyline plot analysis in Beast 1.8.0 (Drummond and Rambaut 2007). This analysis was conducted by setting different substitution models for each gene region, based on JModeltest (see above in phylogenetic methods). Mutation rate was sampled from a normal distribution with mean set to 1.3 % per million year following recent literature on freshwater fish estimates (Burridge et al. 2008; Unmack et al. 2013), and a standard deviation of 0.2% to account for some degree of uncertainty on this parameter. We set a chain length of 10,000,000, sampling parameter values every 1000 steps to generate 10,000 parameter estimates for 5 independent runs. Likelihood values were visualized in Tracer 1.5 (Rambaut and Drummond 2009), making sure ESS values were all above 200. We also analysed the data under a constant population size model to investigate whether the Bayesian skyline model or the constant size model was a better explanation of the data. Models were compared for support using Bayes factor comparison with 10000 replicates for standard error calculations.

Coalescence simulation analysis

To test whether our data can distinguish between recent gene flow or ancestral polymorphism in cases of recent divergence, we conducted coalescent simulations using the software Mesquite 2.75 (Maddison and Maddison 2008). This approach is based on the simulation of gene trees under different scenarios while accounting for coalescent stochasticity. Briefly, a summary statistic (Slatkin and Maddison's; Slatkin and Maddison 1989) that estimates the degree of incomplete lineage sorting (*i.e.* lack of reciprocally monophyletic clades) is calculated for each gene simulated. In our analyses, a value of 1 means complete lineage sorting (*i.e.* reciprocally monophyletic populations), while values greater than 1 indicate increasing levels of incomplete lineage sorting. Subsequently, the frequency distribution of this statistic, simulated under the two different historical scenarios, can be built and compared with the *s* statistic calculated from our empirical gene tree to assess the statistical support. We simulated population subdivision for two times of divergence (models; Fig. 2): one model with divergence at the end of the last glaciation (11 kybp; AGS model) and the other occurring 50 years ago (RS model). The later model's date of divergence is based on the putative barrier between areas I and II (wetlands) that may have formed 55 years ago with the mega earthquake of Valdivia in 1960 (Reinhardt et al.

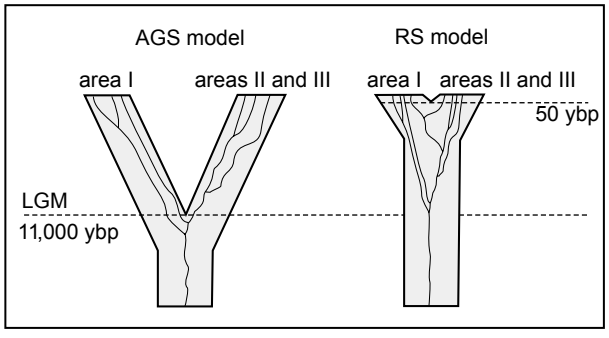

Fig. 2. Models tested in coalescent simulations. Left, the after glaciation subdivision (AGS) model. Right, the recent subdivision (RS) model. Roman numbers on tips indicate main Areas in the basin.

2010). Assuming a generation time of two years (Vila et al. 1996), we built our models as follows: the AGS model was set with a divergence time at 5500 generations ago, whereas RS model was set with a divergence of 25 generations ago. It must be noted that failing to reject the second model should not be taken as evidence for a recent population split, but as evidence that the marker cannot distinguish between ancestral polymorphism (i.e. incomplete lineage sorting) and gene flow at the modelled temporal scale. Given that no empirical estimates of population size are available for *D. camposensis*, we simulated the data setting three different populations sizes for each scenario in order to account for this uncertainty. This totalized six different models combining two divergence times with three different population-size parameter values. We simulated 1000 gene trees for each of the six models, and the effective population size (Ne) parameter was varied to 1000, 10000, and 100,000 individuals.

RESULTS

Genetic diversity and structure

Our results show overall low levels of genetic diversity across the entire basin, as well as for most sample localities (Table 2). The lowest nucleotide diversity was found in population 8, from the glaciated Area III (π = 0.00043), while the highest nucleotide diversity was found in population 4 from Area II (π = 0.0012). Similar patterns were observed for haplotype diversity, with population 4 showing the lowest haplotype diversity (0.44) and population 8 showing the highest (H = 0.93). Overall, nucleotide diversity for the entire basin was 0.00072, while haplotype diversity was 0.68. Despite this low diversity, a decrease in genetic diversity can be observed in a west-east direction.

The GLM analysis showed a negative and significant correlation between H and longitude (R = 0.82; p =0.013), and marginally non-significant (R = 0.7; p = 0.052) negative correlation between π and longitude. At the level of the main areas defined, the highest genetic diversity for both H and π was found in Area II (H = 0.75; π = 0.00082), followed by Area I (H = 0.68; π = 0.00076), and Area III (H = 0.38; π = 0.00038). In the case of the glaciated Area III, the randomization analysis showed that nucleotide diversity was significantly lower than expected by chance (p values equal to 0.027 and 0.016 for sampling with and without replacement, respectively; Fig. 3).

Genetic structure was very low and no significant differentiation was found between most populations and defined areas (Table 3). The only significant comparisons (p < 0.05) were those between La Quinta and Chesque (populations 6 and 1; F_{ST} = 0.246) and between La Quinta and Trafun Este (populations 6 and 7; F_{ST} = 0.305). However, these comparisons became non-significant after Bonferroni correction.

Table 2. Genetic diversity indexes per sampled localities and areas

locality	N	S	H	π
Chesque	13	4	0.68	0.00076
Antilhue	8	5	0.86	0.00080
Cuyincahuin	5	3	0.70	0.00079
Chacaipulli	8	6	0.93	0.00120
Malihue Viejo	1	1	-	-
La Quinta	8	3	0.46	0.00034
Trafún Este	6	3	0.73	0.00076
Balsa Enco	11	2	0.44	0.00043
Río Pellaifa	2	1	0	0
Area				
Area I	13	4	0.68	0.00076
Area II	36	9	0.75	0.00082
Area III	13	2	0.38	0.00038
overall	62	11	0.683	0.00072

N, number of individuals; S, haplotype richness; H, haplotype diversity; π, nucleotide diversity.

Phylogeny and haplotype network

The Bayesian and ML reconstruction showed identical results. Hereafter we only describe and discuss the results from the Bayesian analysis, although ML bootstrap support values were added to the Bayesian tree edges (Fig. 4). The reconstructed gene tree showed a lack of genealogical and geographical structure (Fig. 4). Shallow relationships, indicated by extremely short branch lengths separating most individuals, characterized the genealogy. Individuals do not cluster in groups representing populations or defined areas. On the contrary, individuals from each population and area were interspersed across the entire tree with no pattern of clustering according to geographic origin. The haplotype network mirrored the gene tree regarding the lack of geographic structure (Fig. 5). It was characterized by a central, high-frequency and widely distributed haplotype, connected by only two mutational steps to a second, medium-frequency, and also widely distributed haplotype. Several singletons were connected by one or few mutational steps to the central haplotype, producing a star-like topology. The two most frequent haplotypes were found in all three main areas of the Valdivia basin (haplotypes H1 and

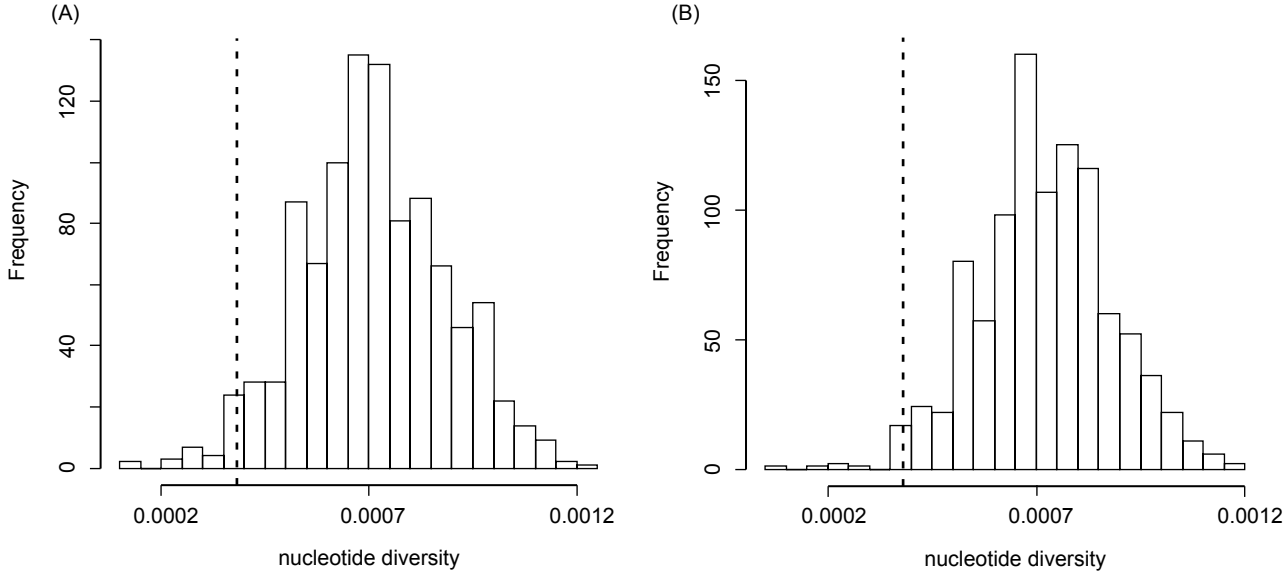

Fig. 3. Frequency distribution of the genetic diversity (π) in the Valdivia basin estimated by bootstrap. Ten thousand samples using 13 randomly chosen individuals each time were used to obtain the distribution. Dashed line indicates the observed nucleotide diversity for Area III. A, bootstrap with replacement. B, bootstrap without replacement. N, number of individuals; S, haplotype richness; H, haplotype diversity; π, nucleotide diversity.

Table 3. Population differentiation values (F_{ST}) between eight localities of *Diplomystes camposensis* from the Valdivia basin based on mitochondrial DNA sequences. Distance matrix was calculated based on pairwise differences and a gamma a value of 0.011. Significance was tested using 1000 permutations. Significant values at the 0.05 level are shown in bold. None value was statistically significant after Bonferroni correction

Locality	1	2	3	4	5	6	7
1. Chesque							
2. Antilhue	-0.015						
3. Cuyincahuin	0.012	-0.125					
4. Chacaipulli	0.000	-0.009	-0.075				
5. La Quinta	**0.246**	0.086	0.068	0.096			
6. Trafún Este	-0.099	-0.020	-0.044	-0.089	**0.305**		
7. Balsa Enco	-0.012	-0.072	-0.066	0.009	0.151	-0.009	
8. Río Pellaifa	0.023	-0.235	-0.290	-0.119	-0.221	0.074	-0.121

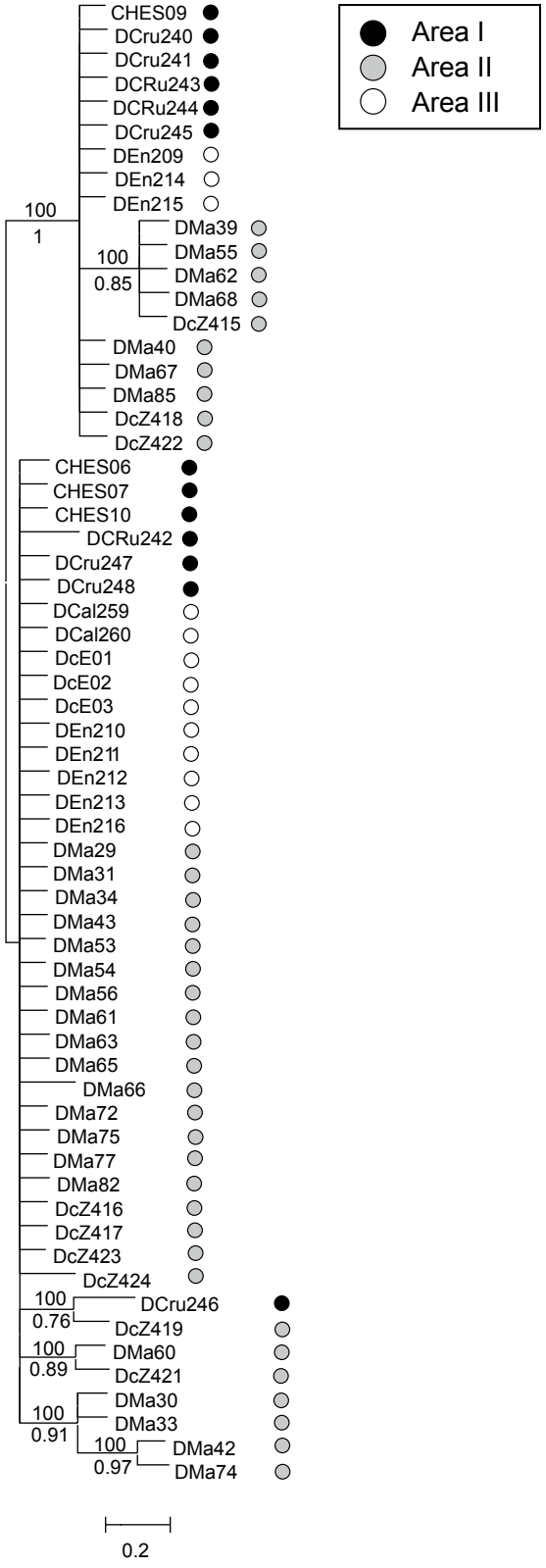

Fig. 4. Bayesian phylogenetic tree based on 62 mitochondrial DNA sequences (Cyt *b* and CR combined). Colored circles in front of tips represent area of origin as shown in figure 1. Tip labels refer to individual codes as in table 1.

H2), while the remaining haplotypes were all private to their corresponding areas. Among the 11 haplotypes found, 9 were present in Area II, 4 in Area I and only 2 in the glaciated Area III. Two haplotypes from Area I were private or exclusive to that area (H3 and H4), whereas the other 2 were shared with areas II and III (H1 and H2). Area II had the higher number of private haplotypes (7 haplotypes: H5-H11). The only 2 haplotypes found in Area III (H1 and H2) were both shared with the other areas and corresponded to those most abundant in the basin.

All the analyses of demographic expansion, except Tajima's *D*, showed evidence of a recent demographic expansion. Mismatch analysis showed a unimodal pattern strongly biased to the left (Fig. 6) indicating that most pairwise

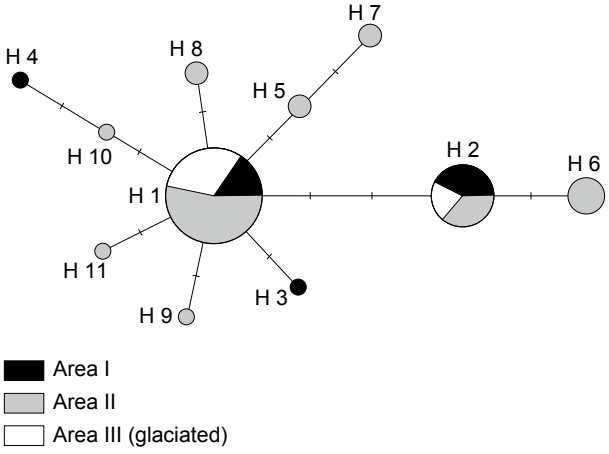

Fig. 5. Haplotype network based on combined Cyt *b* and CR regions. Size of the haplotype represents its frequency. Colors represent sampling Area.

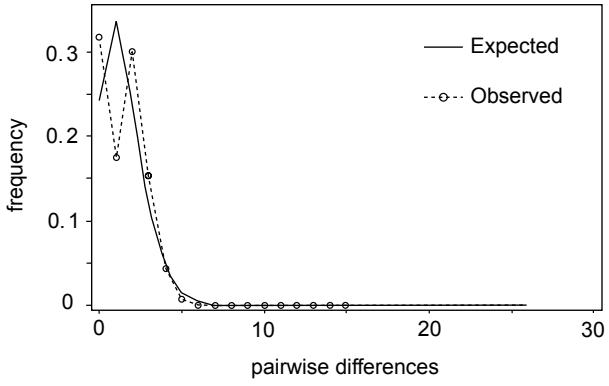

Fig. 6. Mismatch distribution analysis based on pairwise comparisons between 62 mitochondrial DNA sequences.

comparisons involved none very few nucleotide differences. The SSD and raggedness indexes, both of which accompany the mismatch analysis, were both not significantly different from the null expectation of an expansion model (*SSD* = 0018, $p = 0.34$ and *Rgg* = 0.07, $p = 0.56$, respectively). In addition Fu's *Fs* showed a significant negative value (*Fs* = -3.93, $p = 0.03$) which indicates recent demographic expansion. Tajima's *D*, on the other hand, was negative, but not significant (*D* = -1.06; $p = 0.17$). The Bayesian skyline plot analysis showed a slow, but constant increase in population size starting no later than 9,856 ybp (95% lower credibility interval), and no earlier than 80,000 ybp (95% higher credibility interval) (Fig. 7). Although the demographic expansion exhibited by this analysis was not abrupt, it was better supported (likelihood = -2816.357) than a model of constant population size (likelihood = -2868.709) as revealed by the Bayes factor (BF = 22.7; S.E. = 0.297).

Coalescent simulations

Coalescent simulation analyses were conducted to evaluate the ability of our markers to distinguish two hypothetical scenarios of recent population subdivision, one occurring at the end of the last glaciation (11 kybp; AGS model) and the other occurring only 50 years ago (RS model). Our results showed that it is highly unlikely that a population subdivision had occurred and persisted up to date any time since the end of the last glaciation. Slatkin and Maddison's *s* statistic for the observed gene tree was 11, indicating a high level of incomplete lineage sorting between the Cruces River and the rest of the basin. In general, this value was significantly greater than those simulated for the AGS model assuming Ne values of 1000 and 10000, and it was marginally significant for the case with 100000 individuals. For the case of Ne of 1000 individuals, all simulated coalescent gene trees presented an *s*-value of 1 (Fig. 8A). In the case of Ne=10000 individuals, simulated gene trees showed *s*-values that ranged from 1 to 5 with a peak at *s* = 2 (Fig. 8B). For the case of 100000 individuals, simulated gene trees produced *s*-values ranging between 4 and 13, reaching a peak at *s* = 9 (Fig. 8C). For this later value, in which a very large Ne was assumed, *s* was only marginally non-significant. This means that even assuming the very unlikely Ne of 100000

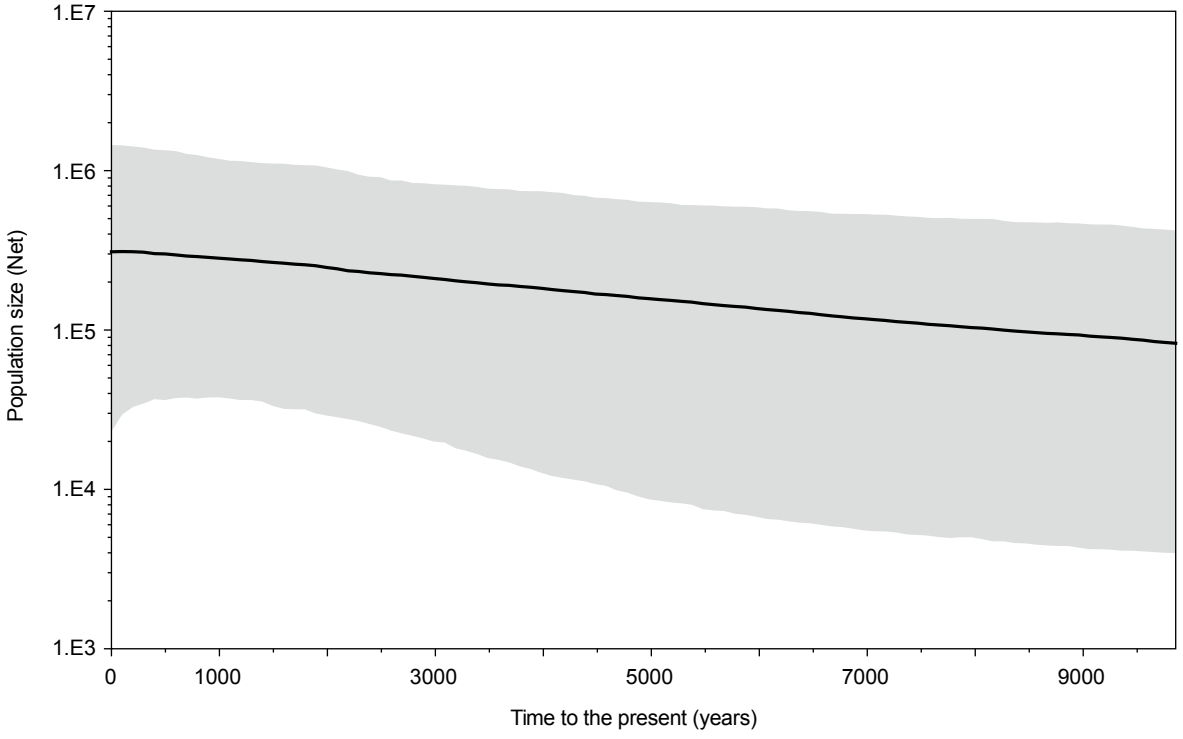

Fig. 7. Bayesian Skyline Plot analysis showing population size over time. The x-axis is the time to the present in years, while the y-axis is the product between the effective population size (Ne) and the generation length (t) in a log scale. The mean estimate (black solid line) and 95% highest probability density limits (grey area) are shown.

individuals, there is only a 5.6% probability of obtaining an *s* value of 11 or higher assuming the AGS model of population subdivision.

Observed *s*-value was not significantly different from those simulated under the recent split model, regardless the Ne assumed (Fig. 8D-F). For all the Ne cases, *s*-values produced by the simulated gene trees were relatively high, yielding the observed *s*-value of 11 or higher in 33%, 90%, and 92% of the times for Ne of 1000, 10000, and 100000 individuals, respectively.

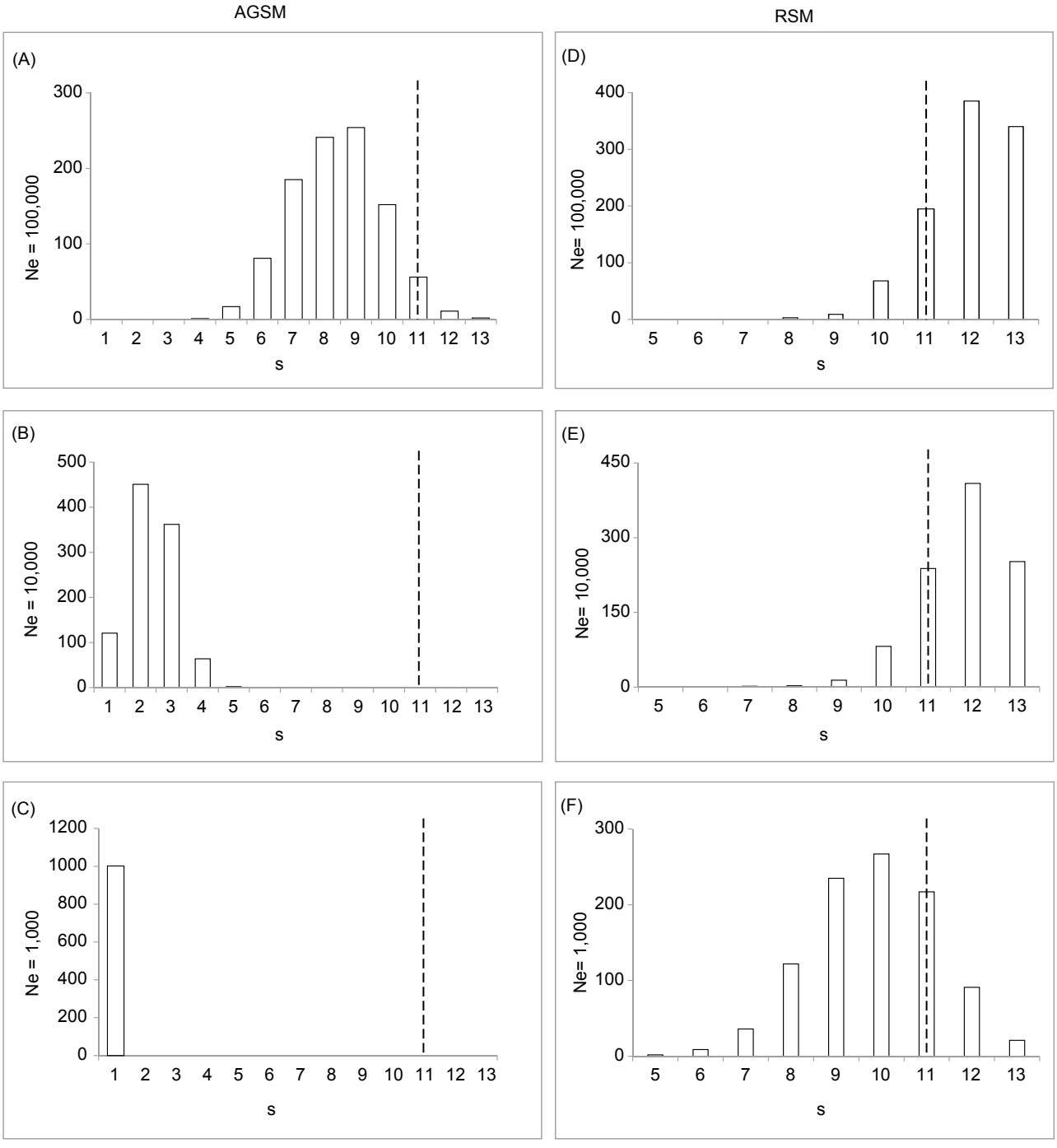

Fig. 8. Frequency distribution of the *s* statistic obtained by simulated gene trees under two different scenarios and values for Ne. Left column, the after glaciation subdivision model (AGSM). Right column, the recent subdivision model (RSM). Segmented line shows where the observed s value (*s* = 11) falls.

DISCUSSION

Although estimates genetic diversity for *Diplomystes camposensis* has been reported previously in the Valdivia basin (Victoriano et al. 2012; Habit et al. 2009), this is the first time that it is analysed for individuals from the entire Valdivia Basin, including populations from the Cruces River, and the glaciated areas beyond the Riñihue Lake. Therefore, our results most likely reflect the true condition of the species in the entire basin. Genetic diversity found in *D. camposensis* was low across the whole basin. Remarkably, it was lower than the genetic diversity found in all other previously studied freshwater fish species cohabiting the same basin (Victoriano et al. 2012), with the solely exception of *Percilia gillisi*, which showed comparable low levels. However, in the case of *Percilia gillisi*, only 20 individuals were analysed from a reduced area within the basin, suggesting that its genetic diversity could have been underestimated. The lower genetic diversity found in *Diplomystes camposensis* relative to the other co-distributed species could reflect naturally lower population sizes. Indeed, *D. camposensis* is generally less abundant than other co-distributed species (Valdovinos et al. 2012). This finding is consistent with genetic evidence of small Ne in *Diplomystes* populations from the Rapel and Mataquito basins in Central Chile (Muñoz-Ramírez et al. 2015). The potential causes of a small Ne in *Diplomystes* have been discussed elsewhere (Muñoz-Ramírez et al. 2014), and they might be related with lower fecundity (Vila et al. 1996).

Despite the globally low genetic diversity in *D. camposensis*, its decrease toward the glaciated area is consistent with a postglacial colonization scenario. Previous phylogeographic studies on fish from this region have failed to find significant correlations between genetic diversity and latitude (*e.g.* Unmack et al. 2009; Muñoz-Ramírez et al. 2014) suggesting a minor role of glaciations on genetic diversity, or a lack of statistical power to detect it. In our case, the finding of this pattern suggests that glaciations may have played a role in the recent microevolutionary history of this species, although more powerful analyses (e.g. model-based demographic analyses) and data (e.g. microsatellites, SNPs) are required to fully test this hypothesis.

Results from the population expansion analyses are consistent with an effect of glaciations on the demographic history of *D. camposensis*. During the LGM, the Valdivia basin was greatly reduced in size due to the advance of glaciers, which may have reduced the population size of the species. After the glacial meltdown (c.a. 11Kybp), more habitat became available, which may have allowed populations expand demographically. Most analyses aiming to test population expansion supported this scenario (Fu's *Fs*, mismatch, BSP), with the only exception of Tajima's *D* test. Tajima's *D* test tends to be more conservative to detect population expansion when the expansion has occurred very recently, or when it has been of low magnitude. This can be seen in the BSP analysis (Fig. 7) that shows a slow population growth over time, suggesting that even though population size was reduced in the recent past, the current population might not be markedly larger. The BSP analysis also placed the bottleneck during the last glaciation (within 9.8 and 80 kybp). Although this time interval is somewhat large, it falls almost completely within the last glaciation period (about 11-110 kybp; Petit et al. 1999) allowing us to rule out some other more recent events (*e.g.* recent volcanism). Altogether, these results are consistent with the hypothesis of an impact of the last glaciation on the demographic history of *D. camposensis*, but contrary to our initial expectations, they also suggest that the impact of glaciations might not have been as dramatic as previously thought.

In natural populations, low levels of genetic diversity may reflect either aspects of the species ecology or recent bottlenecks. The documented impact of the LGM on basins of southern South America (*e.g.* Hulton et al. 2002; Sugden et al. 2005) offers a straightforward explanation for the reduced levels of genetic diversity in this species. However, the finding that the genetic diversity of *D. camposensis* was lower than other co-distributed species (Victoriano et al. 2012) suggests that species-specific biological characteristics may have also contributed to this condition.

Genetic structure

We did not find evidence of population subdivision as revealed by low levels of genetic structure. F_{ST} values and coalescent analyses indicated that the entire basin acts as a single population. In particular, the coalescent simulations indicated that it is highly unlikely that this pattern may have arisen due to ancestral polymorphism, indicating that high gene flow is the most probable explanation for the lack of genetic structure in the basin. This is in agreement with previous

unpublished data (Oyanedel et al. in review) that, using telemetry to study patterns of fish movement, found that *D. camposensis* can move large distances in short time intervals (*ca*. 3 km in one day), suggesting a large home range. Genetic data from other co-distributed freshwater fish species (Victoriano et al. 2012) show diverse levels of genetic structure, suggesting that this pattern may depend on species-specific capabilities to move across the basin. Species like *Galaxias maculatus*, *Aplochiton taeniatus*, and *Percichthys trucha* showed high levels of genetic structure, whereas *Galaxias platei*, *Basilichthys australis*, and *Aplochiton zebra* showed intermediate levels. *Percilia gillissi* showed levels of genetic structure as low as *Diplomystes*, but the sampling for this species was limited to a small portion of the basin. Although the coalescent simulations showed that ancestral polymorphism is unlikely as an alternative explanation, it cannot reject the possibility of current fragmentation due to events occurred in more recent time scales (*e.g.* the formation of wetlands in between areas I and II during the past century; Reinhardt et al. 2010). The use of markers with higher resolution at a more recent temporal scale, such as microsatellites or single nucleotide polymorphism, should be more suitable to investigate genetic structure derived from a recent lack of gene flow.

Conservation considerations

Diplomytes camposensis has been traditionally considered a micro-endemic species of the Valdivia basin (Arratia 1987; Dyer 2000; Habit et al. 2006), and although a recent study has suggested that its distribution could extend further North to two other basins (Tolten and Imperial) based on the presence of shared haplotypes (Muñoz-Ramírez et al. 2014), the high genetic structure found between these basins (*i.e.* F_{ST} values above 0.74) indicates that Valdivian populations are a distinct genetic pool that needs to be treated at least as a different management unit (Moritz 1994). Currently, *D. camposensis* face several potential problems that compromise an adequate conservation of its populations. Current attempts to build several dams with no facilitation for fish pass in rivers that are part of its suitable habitat threaten to reduce and fragment its distribution in the basin. Low genetic diversity found across the entire basin suggests that populations may be prone to high demographic and environmental stochasticity, which could increase the extinction risk if the habitat is further reduced. Fragmentation of an already small population with low genetic diversity (Habit et al. 2009; Victoriano et al. 2012) and large home range carries a number of conservation problems. First, small populations are more exposed to inbreeding and genetic drift which accelerates the loss of genetic diversity and may causes inbreeding depression (Keller 1998; Newman and Pilson 1997; Saccheri et al. 1998). Second, the interruption of migration between areas of the river prevents both the rescue of populations that may go locally extinct (Hanski 1991) and the movement of genes that might be beneficial in scenarios of environmental change (Tallmon et al. 2004).

Populations with the highest levels of genetic diversity were found in Chacaipulli (locality 4), San Pedro River, suggesting that special attention should be given to this area if future conservation goals are the preservation of as much genetic diversity as possible. For example, this population could provide a suitable source of individuals if translocation strategies need to be conducted. On the other hand, fragmentation resulting from dams could also be mitigated by considering the construction of fish pass as evidence suggest that *Diplomystes* seem to disperse well through artificial canals in other rivers (Muñoz-Ramírez et al. 2015). We hope that our study can serve as a starting point in the goal of generating knowledge about the historical dynamics and population genetics of this endangered species. However, more efforts should also be put in increasing the basic understanding of its ecology to improve potential management strategies.

Acknowledgments: This study was funded by Colbún S.A, Fondecyt 1110441 and project GEF/SIRAP (Siempreverde). We thank Alfonso Jara, Alvaro Zuñiga, and Maribel Beltrán for their appreciated help in the field. We also thank Joaquin Ortego for valuable comments to the manuscript.

REFERENCES

Arratia G. 1987. Description of the primitive family Diplomystidae (Siluriformes, Teleostei, Pisces): Morphology, taxonomy, and phylogenetic implications. Bonn Zool Monogr **24:**1-120.

Arratia G. 1992. Development and variation of the suspensorium of primitive catfishes (Teleostei: Ostariophysi) and their phylogenetic relationships. Bonn Zool Monogr **32:**1-140.

Arratia G, Huaquín L. 1995. Morphology of the lateral, line

system and of the skin of diplomystid and systematic and ecological considerations. Bonn Zool Monogr **36:**110.

Arratia GF. 1983. Preferencias de habitat de peces siluriformes de aguas continentales de Chile (Fam. Diplomystidae y Trichomycteridae). Stud Neotrop Fauna Environ **18:**217-237. doi:10.1080/01650528309360637.

Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol **16:**37-48. doi:10.1093/molbev/msl151.

Bello M, Ubeda C. 1998. Estado de conservación de los peces de agua dulce de la Patagonia Argentina. Aplicación de una metodología objetiva. Gayana Zool **62:**45-60.

Beltrán-Concha M, Muñoz-Ramírez C, Ibarra J, Habit E. 2012. Análisis de la dieta de *Diplomystes* (Siluriformes: Diplomystidae) de Chile. Gayana **76:**102-111. doi:10.4067/S0717-65382012000300003.

Burridge CP, Craw D, Fletcher D, Waters JM. 2008. Geological dates and molecular rates: Fish DNA sheds light on time dependency. Mol Biol Evol **25:**624-633. doi:10.1080/10635150701313855.

Cifuentes R, González J, Montoya G, Jara A, Ortíz N, Piedra P, Habit E. 2012. Relación longitud-peso y factor de condición de los peces nativos del río San Pedro (cuenca del río Valdivia, Chile). Gayana **76:**101-110. doi:10.4067/S0717-65382012000100009.

Colin N, Piedra P, Habit E. 2012. Variaciones espaciales y temporales de las comunidades ribereñas de peces en un sistema fluvial no intervenido: río San Pedro, cuenca del río Valdivia (Chile). Gayana **76:**24-35.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol **7:**214. doi:10.1186/1471-2148-7-214.

Dyer BS. 2000. Systematic review and biogeography of the freshwater fishes of Chile. Estud Ocean **19:**77-98.

Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online **1:**47-50. doi:10.1111/j.1755-0998.2010.02847.x.

Fink SV, Fink WL. 1996. Interrelationships of the ostariophysan fishes (Teleostei). *In*: Stiassny MLJ, Parenti LR, Johnson GD (eds) Interrelationships of Fishes. Academic Press, San Diego.

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics **147:**915-925.

García A, González J, Habit E. 2012. Caracterización del hábitat de peces nativos en el río San Pedro (cuenca del río Valdivia, Chile). Gayana **76:**36-44.

Grande L. 1987. Redescription of *Hypsidoris farsonensis* (Teleostei: Siluriformes), with a reassessment of its phylogenetic relationships. J Vertebr Paleontol **7:**24-54.

Habit E, Dyer B, Vila I. 2006. Current state of knowledge of freshwater fishes of Chile. Gayana **70:**100-113.

Habit E, Jara A, Colin N, Oyanedel A, Victoriano P, Gonzales J, Solis-Lufí. 2009. Threatened fishes of the world: *Diplomystes camposensis* Arratia, 1987 (Diplomystidae). Environ Biol Fishes **84:**393-394. doi:10.1007/s10641-008-9434-7.

Hall T. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.Nucl. Acids Symp Ser **41:**95-98.

Hanski I. 1991. Single-species metapopulation dynamics: concepts, models and observations. Biol J Linn Soc **42:**17-38. doi:10.1111/j.1095-8312.1991.tb00549.x.

Heled J, Drummond AJ. 2008 Bayesian inference of population size history from multiple loci. BMC Evol Biol **8:**289. doi:10.1186/1471-2148-8-289.

Hulton NR, Purves R, McCulloch R, Sugden DE, Bentley MJ. 2002. The Last Glacial Maximum and deglaciation in southern South America. Quat Sci Rev **21:**233-241. doi:10.1016/S0277-3791(01)00103-2.

Keller LF. 1998. Inbreeding and Its Fitness Effects in an Insular Population of Song Sparrows (*Melospiza melodia*). Evolution **52:**240-250. doi:10.2307/2410939.

Le Roux J, Elgueta S. 2000. Sedimentologic development of a Late Oligocene-Miocene forearc embayment, Valdivia Basin Complex, southern Chile. Sediment Geol **130:**27-44. doi:10.1016/S0037-0738(99)00096-2.

Librado P, Rozas J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics **25:**1451-1452. doi:10.1093/bioinformatics/btp187.

Lundberg J, Baskin J. 1969. The caudal skeleton of the catfishes, order Siluriformes. Am Museum Novit **2398:**1-50.

Lundberg J, Case G. 1970. A new catfish from the Eocene Green River formation, Wyoming. J Paleontol **44:**451-457.

Maddison WP, Maddison DR. 2008. Mesquite: A modular system for evolutionary analysis. Version 2.1. http://mesquiteproject.org

Montoya G, Jara A, Solis-Lufí K, Colin N, Habit E. 2012. Primeros estadios del ciclo de vida de peces nativos del Río San Pedro (Cuenca del Rio Valdivia, Chile). Gayana **76:**86-100.

Moritz C. 1994. Defining "Evolutionarily Significant Units" for conservation. Trends Ecol Evol pp. 373-375. doi:10.1016/0169-5347(94)90057-4.

Muñoz-Ramírez C, Jara A, Beltrán-Concha M, Zúñiga-Reinoso A, Victoriano P, Habit E. 2010. Distribución de la familia Diplomystidae (Pisces: Siluriformes) en Chile: nuevos registros. Boletín de Biodiversidad de Chile **4:**6-17.

Muñoz-Ramírez CP, Unmack PJ, Habit E, Johnson JB, Cussac VE, Victoriano P. 2014. Phylogeography of the ancient catfish family Diplomystidae: Biogeographic, systematic, and conservation implications. Mol Phylogenet Evol **73:**146-160. doi:10.1016/j.ympev.2014.01.015.

Muñoz-Ramírez CP, Victoriano PF, Habit E. 2015. Inter-basin dispersal through irrigation canals explains low genetic structure in *Diplomystes* cf. *chilensis*, an endangered freshwater catfish from Central Chile. Limnol - Ecol Manag Inl Waters **53:**10-16. doi:10.1016/j.limno.2015.04.004.

Newman D, Pilson D. 1997. Increased Probability of Extinction Due to Decreased Genetic Effective Population Size: Experimental Populations of Clarkia pulchella. Evolution **51:**354-362. doi:10.2307/2411107.

Nylander JAA, Wilgenbusch JC, Warren DL, Swofford DL. 2008. AWTY (are we there yet?): A system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics **24:**581-583. doi:10.1093/bioinformatics/btm388.

Petit RJ, Raynaud D, Basile I, Chappellaz J, Ritz C, Delmotte M, Legrand M, Lorius, C, Pe L. 1999. Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. Nature **399:**429-413. doi:10.1038/20859.

de Pinna, M.C.C. 1998. Phylogenetic relationships of Neotropical Siluriformes (Teleostei: Ostariophysi): historical overview and synthesis of hypotheses. pp. 279-330. *In*: Phylogeny and Classification of Neotropical Fishes. L. R. Malabarba, R. E. Reis, R. P. Vari, Z. M.

Lucena, and C. A. S. Lucena (eds.), EDIPUCRS, Porto Alegre, Brazil 603 pp.

Posada D. 2008. jModelTest: phylogenetic model averaging. Mol Biol Evol **25:**1253-1256. doi:10.1093/molbev/msn083.

R Development Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rambaut A, Drummond AJ. 2009. Tracer V1.5. Available from http//beast.bio.ed.ac.uk/Tracer

Reinhardt EG, Nairn RB, Lopez G. 2010. Recovery estimates for the Río Cruces after the May 1960 Chilean earthquake. Mar Geol **269:**18-33. doi:10.1016/j.margeo.2009.12.003.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19:**1572-1574. doi:10.1093/bioinformatics/btg180.

Ruzzante DE, Walde SJ, Cussac VE, Dalebout, ML, Seibert J, Ortubay S, Habit E. 2006. Phylogeography of the Percichthyidae (Pisces) in Patagonia: roles of orogeny, glaciation, and volcanism. Mol Ecol **15:**2949-68. doi:10.1111/j.1365-294X.2006.03010.x.

Saccheri I, Kuussaari M, Kankare M, Vikman P, Fortelius W, Hanski I. 1998. Inbreeding and extinction in a butterfly metapopulation. Nature **392:**491-494. doi:10.1038/33136.

SERNAPESCA. 2008. Servicio Nacional de Pesca y Acuicultura. http://www.sernapesca.cl/index.php?option=com_remository&Itemid=246&func=fileinfo&id=2143. Accessed 10 January 2014.

Slatkin M, Maddison WP. 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics **123:**603-613.

Stamatakis A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics **30:**1312-1313. doi:10.1093/bioinformatics/btu033.

Sugden DE, Bentley MJ, Fogwill CJ, Hulton NRJ, McCulloch RD, Purves RS. 2005. Late-Glacial Glacier Events in Southernmost South America: a Blend of "Northern" and "Southern" Hemispheric Climatic Signals? Geogr Ann Ser A Phys Geogr **87:**273-288. doi:10.1111/j.0435-3676.2005.00259.x.

Sullivan JP, Lundberg JG, Hardman M. 2006. A phylogenetic analysis of the major groups of catfishes (Teleostei: Siluriformes) using rag1 and rag2 nuclear gene sequences. Mol Phylogenet Evol **41:**636-662. doi:10.1016/j.ympev.2006.05.044.

Tallmon DA, Luikart G, Waples RS. 2004. The alluring simplicity and complex reality of genetic rescue. Trends Ecol Evol **19:**489-96. doi:10.1016/j.tree.2004.07.003.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol **28:**2731-9. doi:10.1093/molbev/msr121.

Unmack PJ, Bennin AP, Habit EM, Victoriano PF, Johnson JB. 2009. Impact of ocean barriers, topography, and glaciation on the phylogeography of the catfish *Trichomycterus areolatus* (Teleostei: Trichomycteridae) in Chile. Biol J Linn Soc **97:**876-892. doi:10.1111/j.1095-8312.2009.01224.x.

Unmack PJ, Hammer MP, Adams M, Johnson JB, Dowling TE. 2013. The role of continental shelf width in determining freshwater phylogeographic patterns in south-eastern Australian pygmy perches (Teleostei: Percichthyidae). Mol Ecol **22:**1683-1699. doi:10.1111/mec.12204.

Valdovinos C, Habit E, Jara A, Piedra P, Gonzalez J, Salvo J. 2012. Dinámica espacio-temporal de 13 especies de peces nativos en un ecotono lacustre-fluvial de la Cuenca del Río Valdivia (Chile). Gayana **76:**45-58.

Victoriano PF, Vera I, Olmos V, Dib M, Insunza B, Muñoz-Ramírez C, Montoya R, Jara A, Habit E. 2012. Patrones idiosincráticos de diversidad genética de peces nativos del Río San Pedro (Cuenca del Río Valdivia), un sistema de la región glaciada del sur de Chile. Gayana **76:**01-09. doi:10.4067/S0717-65382012000100007.

Vila I, Contreras M, Fuentes L. 1996. Reproducción de *Diplomystes nahuelbutaensis* Arratia 1987 (Pises: Diplomystidae). Gayana Oceanol **4:**129-137.

Xu J, Pérez-Losada M, Jara CG, Crandall K. 2009. Pleistocene glaciation leaves deep signature on the freshwater crab *Aegla alacalufi* in Chilean Patagonia. Mol Ecol **18:**904-18. doi:10.1111/j.1365-294X.2008.04070.x.

Zemlak TS, Habit EM, Walde SJ, Battini MA, Adams EDM, Ruzzante DE. 2008. Across the southern Andes on fin: glacial refugia, drainage reversals and a secondary contact zone revealed by the phylogeographical signal of *Galaxias platei* in Patagonia. Mol Ecol **17:**5049-61. doi:10.1111/j.1365-294X.2008.03987.x.

Additional files

File name: Figure_S1. Format: PDF. Title of data: Topological congruence analysis. This file includes a graphical result from the analysis *Compare* (online program AWTY) that checks for topological congruence between trees from the two independent runs of the Bayesian tree analysis.