Biological Journal of the Linnean Society, 2009, **97**, 876–892. With 4 figures

Impact of ocean barriers, topography, and glaciation on the phylogeography of the catfish *Trichomycterus areolatus* (Teleostei: Trichomycteridae) in Chile

PETER J. UNMACK[1]*, ANDRE P. BENNIN[1], EVELYN M. HABIT[2,4], PEDRO F. VICTORIANO[3,4] and JERALD B. JOHNSON[1,5]

[1]*Evolutionary Ecology Laboratories, Department of Biology, Brigham Young University, Provo, UT 84602, USA*
[2]*Unidad de Sistemas Acuáticos, Centro de Ciencias Ambientales EULA-Chile, Universidad de Concepción, Casilla 160-C, Concepción, Chile*
[3]*Centro de Investigaciones en Ecosistema Patagónicos CIEP Coyhaique, Chile*
[4]*Departamento de Zoología, Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepción, Casilla 160-C, Concepción, Chile*
[5]*Monte L. Bean Life Science Museum, Brigham Young University, Provo, UT 84602, USA*

Received 4 September 2008; accepted for publication 9 January 2009

We examined the role of several earth history events on the phylogeographic distribution of the catfish *Trichomycterus areolatus* in Chile using the cytochrome *b* gene. We explored three biogeographic hypotheses: that sea level changes have resulted in the isolation of populations by drainages; that glaciation has impacted genetic diversity; and that ichthyological subprovince boundaries correspond to phylogeographic breaks in our focal species. We found seven well-supported clades within *T. areolatus* with high levels of genetic divergence. The strongest signal in our data was for an important role of sea level changes structuring populations. Five of the seven clades mapped cleanly to the geographic landscape and breaks corresponded closely to areas of narrowest continental shelf. In addition, few haplotypes were shared between rivers within clades, suggesting that only limited local movement of individuals has occurred. There was no relationship between the levels of genetic diversity and the proportion of individual drainages covered by glaciers during the last glacial maximum. Two phylogeographic breaks within *T. areolatus* did match the two previously identified faunal boundaries, but we found three additional breaks, which suggests that faunal breaks have only limited utility in explaining phylogeographic patterns. These results imply that the narrow continental shelf coupled with sea level changes had a strong influence on the obligate freshwater fishes in Chile. © 2009 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2009, **97**, 876–892.

ADDITIONAL KEYWORDS: Andes Mountains – biogeography – community assemblages – continental shelf – cytochrome *b* – Hatcheria macraei – *introgression* – sea level change – South America.

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

INTRODUCTION

A strong geographic bias exists in the global distribution of phylogeographic studies, with the majority of research having focused on taxa from the northern hemisphere (i.e. North America and Europe) (Avise, 2000). This disparity has prompted a call for increased studies in the southern hemisphere (Beheregaray, 2008). Southern South America provides a series of unusual geographic features that make it particularly interesting for studying the phylogeography of the southern hemisphere. Most prominent is the Andes mountain range that runs north to south along the western edge of South America. These mountains formed in the Tertiary and create a barrier to biotic exchange between lowland habitats on either

*Corresponding author. E-mail: peter.pub@unmack.net

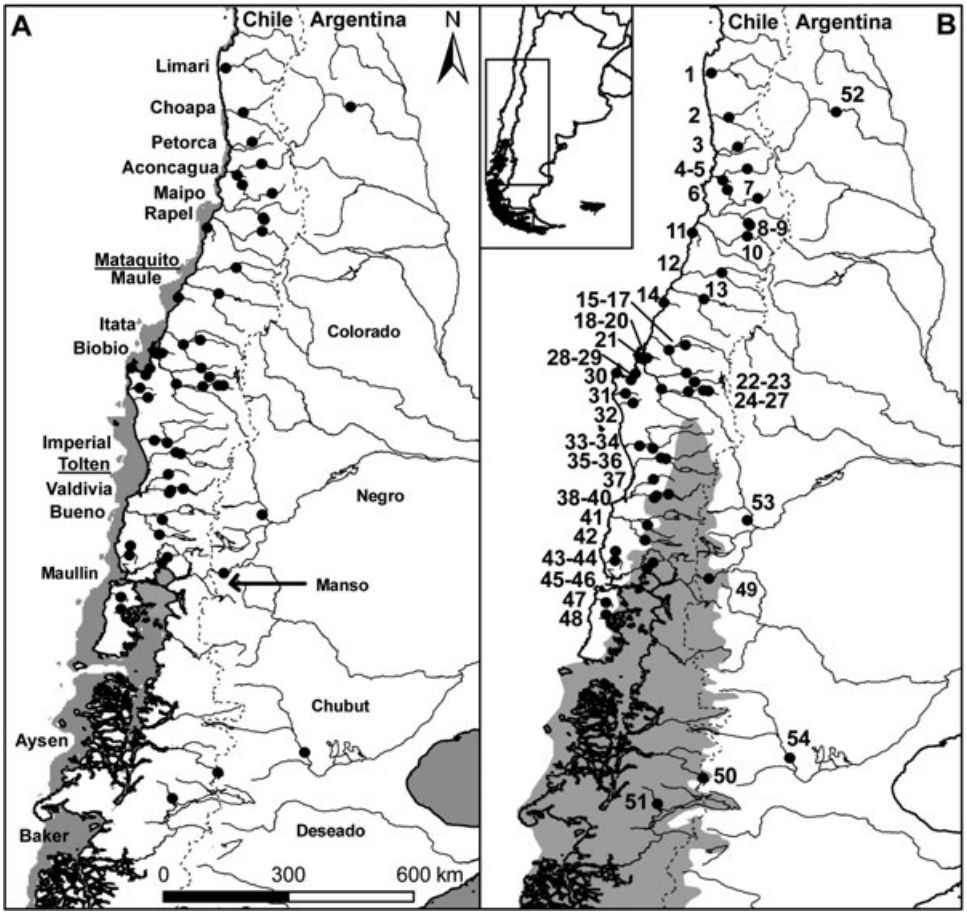

Figure 1. Geographic features of the study area presented in two panels. Dots on both maps represent collection localities and the dotted line represents the political boundary between Chile and Argentina. A, map of central Chile showing rivers and the extent of exposed land when sea level is lowered 150 m during glacial periods, as indicated by the grey-shaded area, comprising information that we used to evaluate hypothesis 1. The Río Mataquito and Río Toltén are underlined to denote the boundaries between the Central, South-central, and Southern fish subprovinces of Dyer (2000), comprising boundaries that we used to evaluate hypothesis 3. B, map of central Chile with site numbers included for our collection localities; details of these sites are found in Table 2. The grey-shaded area represents the estimated maximum extent of the major ice sheet during the last glaciation, comprising information used to evaluate hypothesis 2. The insert map shows the extent of the study area relative to the broader geographic region.

side of the range. In addition, the distance between the high mountains to the east and the ocean to the west is relatively short, with most of central Chile being only 100–220 km wide. This results in an extremely strong elevational gradient from the high Andes down to sea level, with rivers running primarily from the mountains west to the ocean. Such extreme elevation gradients could have a major influence on small-scale climatic patterns in this area, which, in turn, could affect the distribution of some taxa. It is also important to note that the Andes Mountains decrease in elevation from 6960 m in the north near Santiago to below 1000 m in the south in Continental Chiloé. Consequently, for species affected by elevation, we might expect biotic exchange across the Andes to be more common at southern latitudes than at northern latitudes.

Southern South America has also been subject to large-scale climatic fluctuations associated with major glaciations during the Pleistocene. Most notable are repeated cycles where major ice sheets formed and covered vast parts of the southernmost regions of western South America. Current geological hypotheses suggest that glaciers covered most of southern Chile and extended as a single continuous unit north along the higher elevations of the Andes to around 38° south (Fig. 1B) (Hulton *et al.*, 2002). Many smaller isolated glaciers occurred on higher mountain peaks north of this main ice sheet (Grosjean *et al.*, 1998; Harrison, 2004). Hence, glacial events have had

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

an important influence on the distribution of species in southern Chile (Villagrán, 1990; Villagrán & Hinojosa, 2005). Indeed, only taxa that were able to persist in refuges or that could readily migrate along a north to south gradient would have persisted. Those unable to move, such as freshwater aquatic taxa, would have suffered local extirpation. Hence, we expect many species that currently occupy the southernmost regions of Chile to be recent colonists. Combined, the unique geography and climate of southern South America provide a clear framework for exploring the current distribution of species throughout this region.

One group of organisms likely to be strongly impacted by geological and climatic events in southern South America are those restricted to freshwater habitats. Movement for these species is typically limited by connections among riverine systems, with the ocean providing a barrier at each river terminus. As a result of the the topographic nature of this area, most rivers are relatively short and flow from east to west (Fig. 1). Thus, the only escape from unfavorable conditions (e.g. due to climate change or glacial advance) is to move downstream, which would be of limited utility when glaciers are advancing from south to north. Opportunities for south to north movement could occur across drainage divides or if rivers connect along the continental shelf during low sea levels associated with glacial periods (Clark & Mix, 2002). However, Chile has an extremely narrow continental shelf: a sea level fall of 150 m only widens the coastline of Chile by 10–60 km (Fig. 1A). Consequently, few rivers were likely to coalesce during glacial periods, limiting the opportunity for colonization across river systems. Thus, we expect freshwater taxa in southern South America to be strongly impacted by glaciation with limited opportunities for recolonization after glacial events.

The factors responsible for endemism and thus biogeographic province boundaries are also likely to impact widespread species that cross those boundaries (Avise, 2000). Under such circumstances, we might expect to find congruence between biogeographic province boundaries and genetic breaks within widespread species (Burton, 1998). Chile is currently divided into three major ichthyological provinces based on the distribution of endemic freshwater fishes: Titicata in the north; Chilean in the central region; and Patagonian in the south (Dyer, 2000). Most of the obligate freshwater fish species in Chile (18) inhabit the Chilean Province and have relatively restricted ranges that span one or only a few major rivers (Dyer, 2000). In view of these patterns of endemism, Dyer (2000) divided Chilean Province into three subprovinces: Central, South-central and Southern subprovinces, as delimited in Fig. 1A. The boundaries between these subprovinces are delineated by the Río Mataquito and Río Toltén, respectively. It is unclear whether these biogeographic subprovinces are associated with phylogeographic breaks in broadly-distributed species.

The trichomycterid catfish *Trichomycterus areolatus* (Valenciennes) is ideal for understanding Chilean biogeography. It is one of the most widespread obligate freshwater fish in Chile, occurring within Chilean Province in all drainages from the Río Limarí south to Chiloé Island, a distance of approximately 1400 km (Fig. 1). *Trichomycterus areolatus* appears to have broad environmental tolerances and is typically widespread and abundant within most rivers that it occupies (Habit, Victoriano & Campos, 2005; Habit, Belk & Parra, 2007). This species can be found from small headwater streams to low elevation rivers near the coast that have riffles with a variety of substrates, especially gravel and rocks (Arratia, 1983; Garcia-Lancaster, 2006). This makes it an ideal group for exploring the impact of glaciation and climate change on aquatic taxa across central Chile. It is also one of the few species that crosses the two ichthyological subprovince boundaries within the Chilean Province (Fig. 1A; Dyer, 2000).

In the present study, we used *T. areolatus* to explore several biogeographic hypotheses. We generated these hypotheses using three kinds of data: continental shelf width along the coast of Chile; historic climatic patterns and known glacial cycles during Pleistocene; and natural fish community breaks based on existing faunal subprovince boundaries. Using these data, we focused on three general questions. First, we asked what role have sea level changes had in maintaining isolation of major river basins? Given the narrow continental shelf, we expect sea level changes to only have local impacts on *T. areolatus* population connectivity, with the narrowest portions of the continental shelf corresponding to the largest phylogeographic breaks (Fig. 1A). Second, is there a difference in genetic diversity between drainages with significant ice cover during glaciation versus those that were largely unglaciated (Fig. 1B, Table 1)? We might expect to observe a decrease in genetic variation in glaciated basins due to the major decrease in basin size as well as the locally colder climatic conditions and impacts from deglaciation. Finally, do phylogeographic breaks in *T. areolatus* correspond to subprovince biogeographic boundaries identified for fishes (Fig. 1A; Dyer, 2000)?

MATERIAL AND METHODS

Study taxa and sampling

Sample sizes at each collection site varied from one to eleven, but were typically five or eight individuals, for

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

Table 1. Extent of ice cover by river basin at the last glacial maximum (presented from north to south; for reference, see Fig. 1)

River basin	% Glaciated
Imperial	30
Toltén	66
Valdivia	66
Bueno	67
Maullín	90
Ralun	100
Puelo	93
Yelcho	92
Pico-Palena	85
Cisnes	90
Aysén	90
Baker	86

Basins north of the Río Imperial had negligible amounts of glaciation and are not shown. Values are based on the current coastline.

a total of 302 *T. areolatus* from 43 locations spanning the entire distribution of this species in Chile (Fig. 1, Table 2). Our sampling strategy was designed to cover the breadth of their range for phylogeographic analysis. Our sampling scheme has the added advantage in that we are able to comprehensively document within-species diversity and, by including two closely-related species as outgroup taxa (see below), we were able to explore current species boundaries. This approach is particularly appropriate for understudied groups where molecular genetic studies often reveal the presence of cryptic taxa (Johnson, Dowling & Belk, 2004; Wong, Keogh & McGlashan, 2004; Hammer *et al.*, 2007). We initially included a few individuals from two outgroup species: *Hatcheria macraei* (Girard) from Argentina and southern Chile and *Bullockia maldonadoi* (Eigenmann) from central Chile (Fig. 1, Table 2). Our preliminary analyses indicated that *T. areolatus* did not form a monophyletic group, but was paraphyletic with respect to *H. macraei*. Consequently, we expanded our geographic coverage for *H. macraei* to include 12 individuals from six collection sites that span the range of this outgroup species. We also included broad geographic sampling of *B. maldonadoi*, with 26 individuals taken from seven collection sites (Table 2).

Geospatial analysis

We employed a geographic information system approach to quantify several environmental factors needed to evaluate our three general hypotheses. For example, datasets used to generate maps (e.g. Figs 1 and 3) were obtained from the Digital Chart of the World (ESRI, 1993) and manipulated in ArcInfo and ArcMap, version 9.1 (Environmental Systems Research Institute). Bathymetric data (Figs 1A, 3) were obtained via the ETOPO2v2 Global Gridded 2-min Database (NGDC, 2006). The glaciation layer (Fig. 1B) was originally redrawn by J. Milne (*sensu* Clapperton, 1993; Turner *et al.*, 2005; Ruzzante *et al.*, 2006). Drainage basin boundaries were obtained from the HydroSHEDS dataset developed by the World Wildlife Fund (Lehner, Verdin & Jarvis, 2006). Drainage basins were combined with the glaciation layer to determine the percentage cover of glaciers by river basin.

DNA isolation, amplification, and sequencing

We extracted genomic DNA from muscle tissue from each specimen using the DNeasy Tissue Kit (Qiagen Inc.). We amplified the cytochrome *b* (cyt *b*) gene using two primers that flanked this gene: Glu31 5′-GTGACTTGAAAAACCACCGTT-3′ and Cat.Thr29 5′-ACCTTCGATCTCCTGATTACAAGAC-3′. Final concentrations for polymerase chain reaction (PCR) components per 25 μL reaction were: 25 ng of template DNA, 0.25 μM of each primer, 0.625 units of Taq DNA polymerase, 0.1 mM of each dNTP, 2.5 μL of 10 × reaction buffer and 2.5 mM $MgCl_2$. Amplification parameters were: 94 °C for 3 min followed by 35 cycles of 94 °C for 30 s, 48 °C for 30 s, and 72 °C for 90 s, and, finally, 72 °C for 7 min. PCR products were examined on a 1.5% agarose gel using SYBR safe DNA gel stain (Invitrogen). The PCR products were purified using a Montage PCR 96 plate (Millipore). Most sequencing reactions and clean-up were performed using a Parallab 350 robot (Parallabs). Some sequences were also obtained via cycle sequencing with Big Dye 3.0 dye terminator ready reaction kits using 1 : 16th reaction size (Applied Biosystems). Sequencing reactions were run at an annealing temperature of 52 °C in accordance with the ABI manufacturer's protocol. Sequenced products were purified using sephadex columns. Sequences were obtained using an Applied Biosystems 3730 XL automated sequencer at the Brigham Young University DNA Sequencing Center.

Phylogenetic analysis of DNA sequence data

We edited DNA sequences using Chromas Lite, version 2.0 (Technelysium), imported them into BIOEDIT, version 7.0.5.2 (Hall, 1999) and aligned them by eye. Sequences were checked for unexpected frame shift errors or stop codons in MEGA, version 4.0 (Tamura *et al.*, 2007). The editing resulted in a 1076-bp fragment for each individual included in the

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

Table 2. Collection locality data

Population code	Species	Locality	River basin	Subprovince	Latitude			Longitude			*N*
					Degrees	Minutes	Seconds	Degress	Minutes	Seconds	
1	TA	Río Guampulla	Limarí	Central	–30	40	6.5	–71	32	2.6	8
2	TA	Río Choapa	Choapa	Central	–31	37	50.9	–71	9	16.0	8
3	TA	Río Petorca at Zapallar	Petorca	Central	–32	15	46.5	–70	57	36.0	8
4	TA	Estero Limache at Limache	Aconcagua	Central	–32	59	28.9	–71	17	6.7	8
5	TA	Río Aconcagua at San Felipe	Aconcagua	Central	–32	44	25.4	–70	45	21.4	8
6	TA	Estero El Belloto	Maipo	Central	–33	11	49.2	–71	10	57.9	8
7	TA	Estero Arrayan, Santuario Río Arrayan Lo Barnechea	Maipo	Central	–33	22	25.1	–70	31	40.6	11
8	TA	Río en Aguila Sur at Angostura	Maipo	Central	–33	54	40.9	–70	43	48.3	8
9	TA	Río en Angostura del Paine	Maipo	Central	–33	57	16.2	–70	42	8.3	5
10	TA	Río Cachapoal Norte	Rapel	Central	–34	11	37.0	–70	45	13.9	5
11	TA	Estero Topocalma	Topocalma	Central	–34	7	29.6	–71	56	31.7	2
12	TA	Río Mataquito at Lo Valdivia	Mataquito	boundary	–34	59	8.8	–71	18	24.5	5
13	TA	Río Maule	Maule	South-central	–35	33	4.9	–71	41	37.6	5
14	TA	Estero Reloca	Reloca	South-central	–35	37	49.3	–72	33	43.6	8
15	TA	Río Itata at Nueva Aldea	Itata	South-central	–36	39	10.2	–72	27	12.8	5
16	TA	Río Ñuble at el Centro	Itata	South-central	–36	33	2.8	–72	5	31.6	5
19	TA	Estero Nonguén at Puente Carmelitas	Andalién	South-central	–36	50	49.0	–73	0	27.0	37
20	TA	Estero Queule, Aguas de la Gloria	Andalién	South-central	–36	49	52.9	–72	55	12.4	8
21	TA	Río Biobío at Puente Llacolen	Biobío	South-central	–36	49	46.1	–73	4	12.3	5
22	TA	Río Laja	Biobío	South-central	–37	21	6.5	–71	53	54.2	8
23	TA	Río Rucue	Biobío	South-central	–37	21	8.6	–71	53	6.2	3
26	TA	Río Duqueco at Mampil	Biobío	South-central	–37	32	9.0	–71	42	13.5	6
27	TA	Río Duqueco at Peuchen	Biobío	South-central	–37	32	17.4	–71	35	39.8	2
28	TA	Río Laraquete	Lebu	South-central	–37	10	4.8	–73	10	26.3	5
30	TA	Estero Punta Lavapie	Punta Lavapie	South-central	–37	9	29.3	–73	34	43.6	1
31	TA	Río Pilpilco	Lebu 2	South-central	–37	35	40.8	–73	23	29.4	8
32	TA	Río Cayucupil	Cayucupil	South-central	–37	47	51.9	–73	13	17.5	5
33	TA	Río Imperial	Imperial	South-central	–38	43	34.3	–73	5	6.9	8

© 2009 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2009, **97**, 876–892

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

34	TA	Río Cautín	Imperial	South-central	–38	46	40.5	–72	47	51.4	8
35	TA	Río Toltén	Toltén	boundary	–38	59	8.5	–72	37	9.9	8
36	TA	Río Allipén	Toltén	boundary	–39	0	49.2	–72	30	33.0	5
37	TA	Río Cruces Sur	Valdivia	Southern	–39	27	14.5	–72	46	59.7	8
38	TA	Río San Pedro	Valdivia	Southern	–39	51	12.2	–72	45	27.0	6
39	TA	Río San Pedro	Valdivia	Southern	–39	48	8.4	–72	43	27.0	2
40	TA	Río San Pedro	Valdivia	Southern	–39	46	32.7	–72	27	34.0	2
41	TA	Río Pilmahiquen	Bueno	Southern	–40	26	21.8	–72	54	47.1	8
42	TA	Río Rahue	Bueno	Southern	–40	45	45.7	–72	58	14.9	7
43	TA	Río Blanco	Llico	Southern	–41	12	30.6	–73	37	6.7	8
44	TA	Río La Plata	Llico	Southern	–41	0	10.9	–73	36	7.9	8
45	TA	Río Pescado	Maullín	Southern	–41	15	4.2	–72	47	47.5	5
46	TA	Río Alerce	Maullín	Southern	–41	23	8.0	–72	55	22.8	8
47	TA	Río Cudil	Butalcura	Southern	–42	22	28.6	–73	48	22.1	8
48	TA	Río Puntra	Butalcura	Southern	–42	7	11.1	–73	48	39.6	8
17	BM	Río Itata en Liucura	Itata	NA	–37	9	27.8	–72	4	5.6	7
18	BM	Estero Nouguén at Andalién confl.	Andalién	NA	–36	48	57.6	–73	0	55.7	2
19	BM	Estero Nonguén at Puente Carmelitas	Andalién	NA	–36	50	49.0	–73	0	27.0	4
20	BM	Estero Queule, Aguas de la Gloria	Andalién	NA	–36	49	52.9	–72	55	12.4	2
24	BM	Río Biobío at Coigue	Biobío	NA	–37	29	55.8	–72	36	36.6	3
25	BM	Río Duqueco at Villacura	Biobío	NA	–37	33	9.4	–72	1	59.0	4
29	BM	Río Ramadillas en Ramadillas	Ramadillas	NA	–37	18	14.3	–73	16	7.2	4
49	HM	Río Manso	Puelo	NA	–41	35	43.4	–71	34	51.4	2
50	HM	Río Huemules	Aysén	NA	–45	54	22.8	–71	42	42.2	2
51	HM	Lago General Carrera	Baker	NA	–46	28	5.1	–72	41	36.8	2
52	HM	Río San Juan above Ullum Res.	Colorado	NA	–31	30	24.3	–68	50	22.4	2
53	HM	Río Calefu	Negro	NA	–40	20	10.0	–70	45	12.0	2
54	HM	Río Senguer	Chubut	NA	–45	28	12.7	–69	50	7.2	2

Locality codes correspond to the numbers shown in Fig. 1B. BM, *Bullockia maldonadoi*; HM, *Hatcheria macraei*; TA, *Trichomycterus areolatus*. The locality column gives the general location of each sample, followed by the river basin. Biogeographic subprovinces within Chilean Province are based on Dyer (2000). Latitude and longitude are provided based on the precise locality. Sample sizes examined are provided in the last column. NA (not appropriate) indicates the columns relative to the two outgroups that were not calculated as part of any of the specific hypothesis tested.

© 2009 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2009, **97**, 876–892

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

study. Phylogenetic analyses were performed using both parsimony and likelihood approaches using PAUP* 4.0b10 (Swofford, 2003). We conducted our maximum parsimony (MP) analysis using a heuristic search with 1000 random additions and tree bisection and reconnection (TBR) branch swapping. For our maximum likelihood (ML) analysis, we identified the best-fitting model of molecular evolution using the Akaike Information Criterion (AIC) in MODELTEST, version 3.7 (Posada & Crandall, 1998). MODELTEST identified GTR+I+G as the best model with the following parameter estimates: Lset Base = (0.2782 0.2887 0.1424) Nst = 6 Rmat = (8.3552 82.6934 2.3639 5.3730 26.7399) Rates = gamma Shape = 0.8469 Pinvar = 0.5956. Due to the large number of haplotypes in the full dataset (147 haplotypes), we were constrained by computation time in our ability to run bootstrap analyses for both the MP and ML analyses. To overcome this limitation, we created a trimmed dataset of 55 haplotypes that were used to estimate bootstrap values. To construct the trimmed data set, we chose the two most divergent haplotypes plus several intermediate haplotypes from each clade. This gave a total 45 haplotypes for *T. areolatus*, six haplotypes for *H. macraei*, and four haplotypes for *B. maldonadoi*. Most haplotypes within a clade differed only by a few base pairs, thus minimizing the chances of getting significantly different results based on which haplotypes were included. Nodal support on the reduced dataset was estimated with PAUP* by bootstrap with 1000 replicates for MP using a heuristic search with ten random additions of taxa and TBR branch swapping. For ML, we used GARLI 0.951 (Zwickl, 2006) to estimate 1000 bootstrap replicates using the default values recommended for the software. The ML tree topology was estimated with PAUP* via a heuristic search with five random additions of taxa and TBR branch swapping using the GTR+I+G model of evolution with the parameters: Lset Base = (0.2784 0.2859 0.1448) Nst = 6 Rmat = (6.7769 66.9881 1.5695 3.2757 20.9228) Rates = gamma Shape = 0.9155 Pinvar = 0.6229. All tree lengths reported for MP include both informative and uninformative characters. Trees were edited using DENDROSCOPE, version 1.4 (Huson *et al.*, 2007) to arrange haplotypes in a north to south orientation.

Finally, in our discussion, we explore levels of sequence divergence between our focal species and other freshwater taxa that occupy some of the same drainages investigated in the present study. To make these comparisons with freshwater crabs of the genus *Aegla*, we secured published mitochondrial (mt)DNA data directly from M. Perez-Losada (Pérez-Losada *et al.*, 2002). We used crab data from the genes *COI* and *COII* because these typically have closer substitution rates to cyt *b* than do other mitochondrial ribosomal genes that were available. We used model-corrected sequence divergence in the crabs, which required identifying a best-fitting model of molecular evolution using AIC in MODELTEST as outlined above. We found this model to be TrN+I+G with the parameters: Lset Base = (0.3180 0.1390 0.1283) Nst = 6 Rmat = (1.0000 14.4898 1.0000 1.0000 8.9973) Rates = gamma Shape = 0.8667 Pinvar = 0.6814. Pairwise genetic divergence among haplotypes across drainages were then calculated in PAUP* using these parameter estimates.

Phylogeographic analysis

To explore factors that have shaped the distribution of *T. areolatus* through space and time, we employed a set of phylogeographic analytical approaches. We first employed an analysis of molecular variance (AMOVA) to determine how genetic variation was partitioned across the geographic landscape (Excoffier, Smouse & Quattro, 1992). If genetic structuring is driven by isolation among drainages due to the saltwater ocean boundary, then we expect to observe high levels of structuring among drainages. Similarly, if fish community provinces are indicators of historical events, then we expect to see breaks in our *T. areolatus* data set coincident with fish community boundaries. We used the AMOVA framework to create two competing models corresponding to these two hypotheses. Our AMOVA models are distinguished by the way that the 43 *T. areolatus* collecting sites are grouped. Cases where high amounts of variation can be explained between groups suggest an important historic barrier to gene flow exists, coincident with the structure of the model. Our two models were constructed as follows (for geographic reference, see Fig. 1). (1) To evaluate the role of sea level changes in maintaining isolation of river basins, we divided populations into 24 groups that coincide with the 24 river drainages from which *T. areolatus* were collected (Table 2). These groups were delineated by collection sites: 1, 2, 3, 4–5, 6–9, 10, 11, 12, 13, 14, 15–16, 19–20, 21–23/26–27, 28, 30, 31, 32, 33–34, 35–36, 37–40, 41–42, 43–44, and 45–46. (2) To evaluate the association between fish community boundaries and *T. areolatus* phylogeography, we divided populations into three groups based on the three geographic subprovinces defined by Dyer (2000; Fig. 1A, Table 2). These groups were: Central subprovince (sites 1–11); South-central subprovince (sites 13–16, 19–23, 26–28, and 30–34); and Southern subprovince (sites 37–46). We excluded populations from Río Mataquito (site 12) and Río Toltén (sites 35–36) because Dyer (2000) listed these drainages as being the boundary between subprovinces, but did not assign them to a specific sub-

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

province. All of these analyses were executed in ARLEQUIN, version 3.1 (Excoffier, Laval & Schneider, 2006) with significance assessed using 10^4 random permutations of the dataset.

We evaluated the effects of glaciation on *T. areolatus* populations by examining the levels of genetic diversity among drainages across the range of this species. Most drainages in southern Chile that were substantially covered by ice (> 85%) during the last glacial maximum currently do not contain *T. areolatus* (Dyer, 2000), with the exception of Río Maullín of which 90% was covered. However, all other drainages at the southern extent of the range that were partially covered by ice (30–90% cover) do contain *T. areolatus*. We were interested in comparing populations from these sites with those from more northern drainages where glaciers were absent or minor in coverage. If glaciation was important in shaping the genetic diversity in this species, we might expect to find a negative association between the percentage of drainage that was covered with ice and the level of genetic diversity now present in that drainage. To test this hypothesis required calculating the genetic diversity for fishes in each drainage as well as calculating the percentage of each drainage that was historically covered by ice (as described above). Three drainages (i.e. Maipo, Imperial, and Toltén) had individuals from multiple clades: we separated these drainages by clade for calculating genetic diversity because the presence of multiple clades would artificially raise genetic diversity measures relative to surrounding drainages with only single clades. We calculated mtDNA genetic diversity using the measures: number of haplotypes, haplotype diversity (H_D) and nucleotide diversity (π) with the software DnaSP, 4.50.3 (Rozas *et al.*, 2003). We then used linear regression to test for a significant relationship between the percentage of glacial coverage by drainage and each of our three measures of genetic diversity; these analyses were executed in SYSTAT, version 10 (SPSS, 2000).

RESULTS

The distribution of genetic variation in *T. areolatus* allowed us to evaluate each of our three biogeographic hypotheses. We first describe patterns of sequence diversity for the cyt *b* gene within *T. areolatus*; we then present the mtDNA gene tree and show how it maps to the geographic landscape where samples were collected; and, finally, we describe phylogeographic structuring and patterns of genetic diversity among populations.

SEQUENCE DIVERSITY

We found 125 haplotypes of *T. areolatus*, ten haplotypes of *H. macraei*, and 12 haplotypes of *B. maldonadoi* (GenBank accession FJ772091-FJ772237; see Supporting information, Appendix S1). Of the 1076 bp sequenced, 849 characters were invariant, 50 characters were variable but parsimony uninformative, and 175 characters were parsimony informative. Our MP search with all characters weighted equally recovered 15 599 most parsimonious trees with a length of 436 (CI = 0.598, RI = 0.930). ML analysis recovered one tree with a likelihood score of –4310.22826 (Fig. 2). The trimmed dataset contained 55 haplotypes. Of the 1076 bp sequenced, 875 characters were invariant, 55 characters were variable but parsimony uninformative, and 144 characters were parsimony informative. The MP search, with all characters weighted equally, recovered 1637 most parsimonious trees with a length of 355 (CI = 0.634, RI = 0.893). ML analysis recovered three trees with a likelihood score of –3629.74808 (Fig. 3). Overall, the trees were well resolved, with good to strong support (Hillis & Bull, 1993) at most deeper nodes demonstrated by bootstrap values > 70 for both MP and ML analyses (Fig. 3).

PHYLOGENETIC RELATIONSHIPS

We found seven well-supported clades within *T. areolatus* (labelled as clades A through G in Figs 2, 3). Five of the seven clades map cleanly to the geographic landscape with no shared haplotypes between geographic regions. One exception was clade B, which contained haplotypes that were geographically distributed in areas primarily represented by clades C and D. Specifically, two individuals with haplotype 21 occurred at site 7 in the Río Maipo drainage, and eight individuals with haplotype 24 occurred at site 14 in the Estero Reloca drainage (Figs 1, 3). Haplotype 21 was rare in the Río Maipo drainage (two of 32 fish sampled). By contrast, all eight fish sampled from Estero Reloca drainage were from clade B. The second exception was a small region of geographic range overlap between haplotypes found in clades E and F. Clade E haplotypes were geographically distributed from the southern boundary of clade D at the Río Biobio drainage south to Río Toltén; Clade F haplotypes were distributed from Río Imperial south to Río Bueno (Fig. 1A). Thus, there is a region of overlap between clades E and F in the Río Imperial and Río Toltén drainages (sites 33–36 in Fig. 1; drainges described in Fig. 3). Of the 18 individuals sampled in these drainages, 13 were from clade E and five were from clade F (Fig. 2). However, all four localities from these drainages had haplotypes from both clades (see Supporting information, Appendix S1).

Unexpectedly, we also found that the outgroup *H. macraei* nested within the southernmost *T. areolatus* clade G (Figs 2, 3). Moreover, it did not form a single monophyletic group within this clade, but, instead,

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

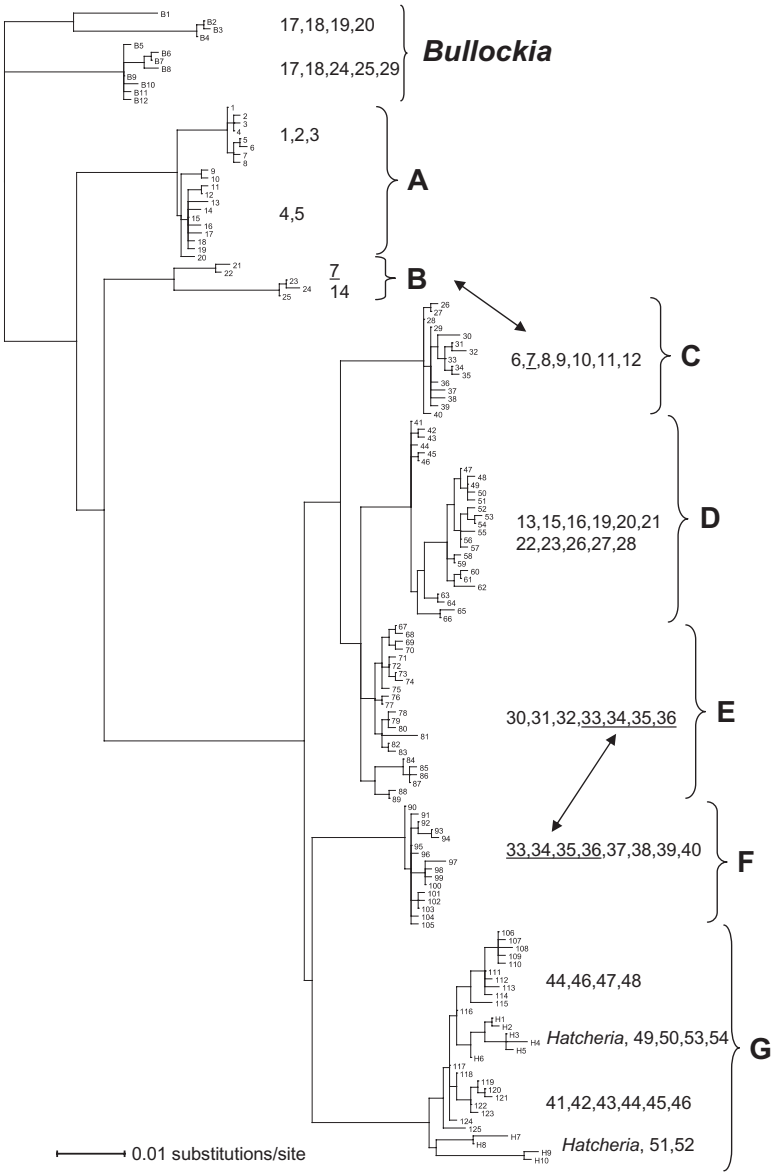

Figure 2. Maximum likelihood (ML) phylogram of *Trichomycterus areolatus* haplotypes generated from 1076 bp of the cytochrome *b* gene. The phylogeny includes 125 haplotypes of *T. areolatus*, 12 haplotypes of the outgroup *Bullockia maldonadoi* (prefixed with B), and ten haplotypes of the outgroup *Hatcheria macraei* (prefixed with H). Branch lengths were estimated via ML under the GTR+I+G model of evolution. The tree was rooted with *B. maldonadoi*. Numbers to the right of the haplotypes indicate the collection localities present in each clade; haplotypes underlined and linked with arrows are shared between clades. Letters to the right of the tree comprise labels for the seven major clades. Details for each haplotype, including locations where haplotypes are found and the number of individuals with that haplotype, are provided in the Supporting information (see Supporting information, Appendix S1).

was represented by two small lineages nested within a set of *T. areolatus* haplotypes. The first *H. macraei* lineage (haplotypes H7 to H10) was composed of individuals collected from the Río Baker drainage at the southern range of *T. areolatus* in Chile and from the Río Colorado drainage far to the north in Argentina (localities 51 and 52 in Fig. 1B). The second *H. macraei* lineage (haplotypes H1 to H6) was composed of individuals from the Manso and Aysén drainages in Chile (sites 49 and 51; Fig. 1B) and from the Negro and Chubut drainages across the Andes mountains in Argentina (sites 53 and 54; Fig. 1B).

GENETIC DIVERSITY

We evaluated patterns of genetic diversity by comparing genetic variation within each of our major clades

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

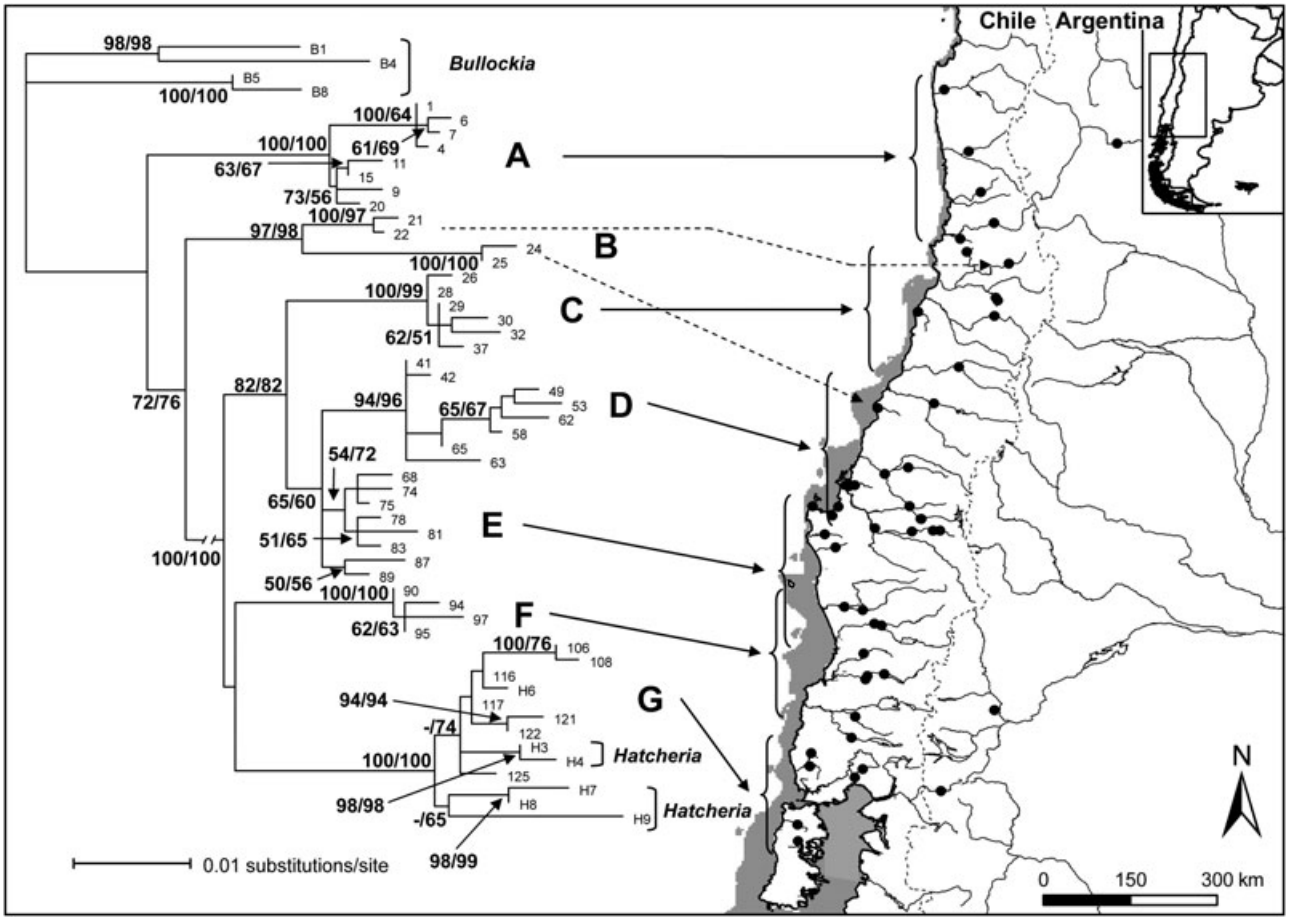

Figure 3. Maximum likelihood (ML) phylogram of *Trichomycterus areolatus* haplotypes generated from 1076 bp of the cytochrome *b* gene from a subset of haplotypes provided in Fig. 2. This tree includes 45 haplotypes of *Trichomycterus areolatus*, four haplotypes of *Bullockia maldonadoi* (prefixed with B), and six haplotypes of *Hatcheria macraei* (prefixed with H). Haplotype designations match those shown in Fig. 2 and the Supporting information (see Supporting information, Appendix S1). Branch lengths were estimated using ML under the GTR+I+G model of evolution. Bootstrap values shown are derived from maximum parsimony and ML analyses (listed in that order) based on 1000 psuedoreplicates each. The dashed bootstrap value indicates support < 50%. The tree was rooted with *B. maldonadoi*. Letters to the right of the haplotypes represent major clade designations. The geographic distribution of each *T. areolatus* clade is indicated by the brackets; the geographic distribution of *H. macraei* is not shown but this species occurs to south and west of the distribution of *T. areolatus*. Dashed arrows indicate the geographic location of those haplotypes within clades that do not map to the geographic regions indicated by the brackets. Overlapping brackets for the Río Imperial and Río Toltén indicate the four populations sampled in these drainages that contain haplotypes found in both clades E and F. The grey-shaded area along the Chilean coast indicates the extent of exposed land when sea level is 150 m lower.

and within each of the 22 river systems. Levels of genetic diversity were generally high within clades: maximum model-corrected sequence divergence within the seven major clades was in the range 0.8–2.6% (Table 3). The maximum percentage of sequence divergence among clades was in the range 2.3–9.1%. Genetic diversity was also high within most river systems, as measured both by haplotype diversity and by nucleotide diversity (Table 4). For example, half of the rivers (12 of 24 sampled) had haplotype diversity measures greater than 0.9 and only seven river systems had hapotype diversity measures less than 0.5. Nucleotide diversity measures showed comparable patterns (Table 4). The two exceptions were the Río Limarí and Río Choapa which showed no genetic variation among individuals for the 16 fish sampled across these two basins (see Supporting information, Appendix S1).

We found no evidence that genetic diversity in *T. areolatus* was affected by Pleistocene glaciation. For

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

Table 3. Genetic distance matrix comparing the seven major clades and two outgroup species shown in Fig. 1

	A	B	C	D	E	F	G	*Hatcheria*	*Bullockia*
A	0.015								
B	0.035–0.057	0.026							
C	0.057–0.072	0.057–0.071	0.008						
D	0.057–0.076	0.058–0.075	0.020–0.033	0.016					
E	0.047–0.069	0.049–0.066	0.016–0.026	0.010–0.023	0.008				
F	0.056–0.071	0.061–0.079	0.030–0.041	0.025–0.040	0.026–0.031	0.008			
G	0.065–0.091	0.066–0.086	0.030–0.045	0.032–0.052	0.031–0.045	0.033–0.048	0.015		
Hat.	0.071–0.094	0.064–0.093	0.034–0.046	0.034–0.053	0.036–0.046	0.036–0.050	0.001–0.026	0.022	
Bul.	0.040–0.072	0.041–0.068	0.070–0.099	0.071–0.091	0.065–0.085	0.070–0.094	0.080–0.102	0.077–0.108	0.054

Trichomycterus areolatus clades are indicated by letters A to G; outgroups are listed by their genera names *Hatcheria* and *Bullockia*. Values on the diagonal show the maximum divergence observed within each clade. Values below the diagonal are the minimum and maximum genetic distances among haplotypes between clades are reported for each comparison. Genetic distances were corrected under the GTR+I+G model of evolution (see text).

Table 4. Genetic diversity measures by river drainage basin for *Trichomycterus areolatus* for the cytochrome *b* gene

River	*N*	Polymorphic sites	# of haplotypes	H_D	π
Limarí	8	0	1	0.000	0.00000
Choapa	8	0	1	0.000	0.00000
Petorca	8	8	7	0.964	0.00263
Aconcagua	16	22	12	0.967	0.00441
Maipo	32	60	10	0.790	0.00729
Maipo clade B	2	3	2	1.000	0.00279
Maipo clade C	30	10	10	0.761	0.00182
Rapel	5	11	5	1.000	0.00410
Topocalma coast	2	1	2	1.000	0.00093
Mataquito	5	2	3	0.700	0.00093
Maule	5	4	4	0.900	0.00168
Reloca coast	8	3	3	0.464	0.00070
Itata	10	11	4	0.778	0.00428
Andalién coast	45	25	15	0.906	0.00444
Biobío	24	20	13	0.942	0.00480
Lebu	5	2	2	0.400	0.00074
Lavapie coast	1	0	1	0.000	0.00000
Lebu 2 coast	8	7	5	0.786	0.00206
Cayucupil coast	5	3	4	0.900	0.00130
Imperial	16	42	12	0.942	0.01477
Imperial clade E	11	22	8	0.891	0.00650
Imperial clade F	5	3	4	0.900	0.00130
Toltén	13	34	7	0.795	0.01409
Toltén clade E	6	11	5	0.933	0.00385
Toltén clade F	7	1	2	0.286	0.00027
Valdivia	18	15	12	0.941	0.00319
Bueno	15	16	7	0.657	0.00253
Llico coast	16	3	4	0.517	0.00054
Maullín	13	13	6	0.782	0.00253
Butalcura coast	16	3	3	0.242	0.00035

Values presented include the number of polymorphic sites, the number of haplotypes, haplotype diversity (H_D), and nucleotide diversity (π) Rivers are listed from north to south. Smaller short coastal rivers are identified by coast after their name. Three rivers (Maipo, Imperial, and Toltén) contained individuals from multiple genetic clades. These were separated into clades, as well as being calculated as a whole.

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

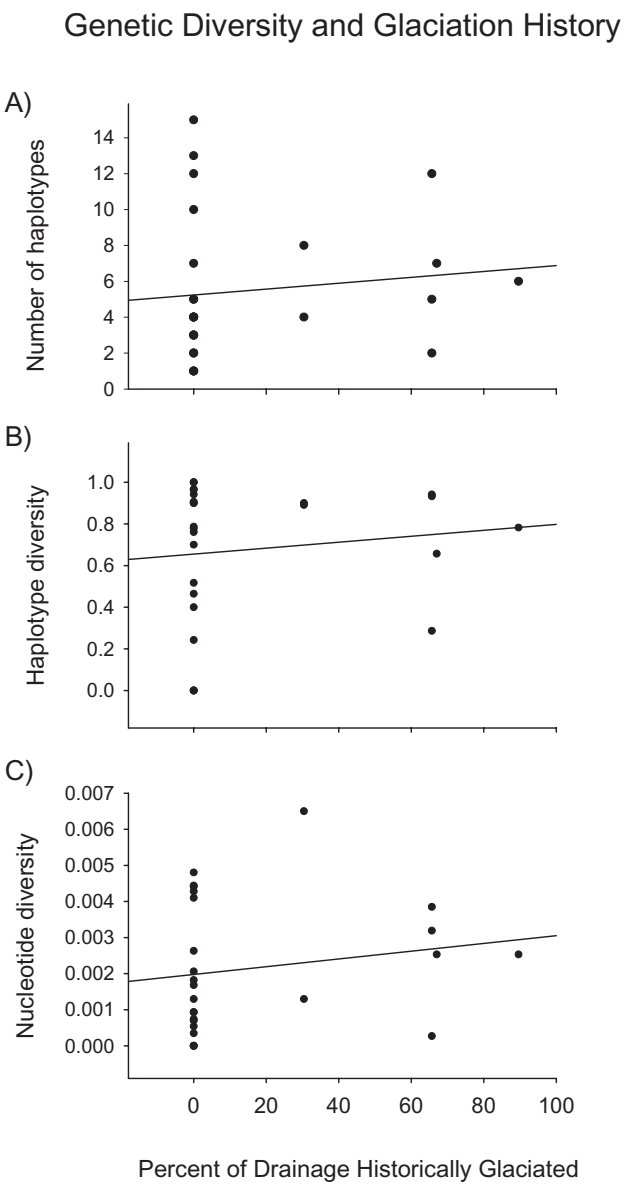

Figure 4. Plots showing relationship between the percentage of drainage basins covered by glaciers during the Pleistocene and (A) number of haplotypes; (B) hapolotype diversity; and (C) π (nucleotide diversity). The line in each plot marks the least square regression line. None of the lines has a slope significantly different from zero, indicating that percentage of glaciation does not have a significant effect on the response variables.

each of the genetic diversity measures that we calculated, there was no relationship between genetic diversity and the percentage of a drainage covered with ice (Fig. 4). The slopes of each of these lines were statistically indistinguishable from zero: number of haplotypes (slope = 0.016, $P = 0.56$); haplotype diversity (slope = 0.001; $P = 0.545$); and nucleotide diversity (slope = 0.000; $P = 0.401$).

PHYLOGEOGRAPHY

Trichomycterus areolatus populations are highly structured. Most of the genetic variation within this species is partitioned among river drainages. That is, there is very little evidence for gene flow between different river basins ($F_{CT} = 0.85$, $P < 0.001$). Our AMOVA under the 'sea level change' hypothesis revealed that 84.9% of the total genetic variation could be explained among river drainages, whereas only 2.1% could be explained among collecting sites within drainages, and 13.0% within collecting sites (Table 5). We found that fish biogeographic regions were comparatively poorer predictors of the partitioning of genetic variation in *T. areolatus* ($F_{CT} = 0.45$, $P < 0.001$). A moderate 45.3% of total genetic variation could be explained among the three fish biogeographic regions, with 44.7% among sites within fish community regions, and 10.0% within sites themselves.

The distribution of haplotypes across the geographic landscape also reveals a limited amount of gene flow among drainage systems. Of the 125 *T. areolatus* haplotypes identified in the present study, only 14 are shared among two or more river systems (see Supporting information, Appendix S1). Those that do share haplotypes tend to drain into the ocean in areas where there is a wide continental shelf that could potentially allow the rivers to connect during low sea level events. For example, the Río Maipo shares one haplotype each with three adjacent river systems: the Río Rapel shares haplotype 28, the Río Estero Topocalma shares haplotype 29, and the Río Mataquito shares haplotype 33. Another example of this can be found in the Itata, Andalién, and Biobío basins, which share nine haplotypes in various combinations. Interestingly, the Río Andalién and Río Biobío have river mouths that almost connect at current sea level, and clearly connect at slightly lower sea levels (Mardones, 2002). Finally, the Imperial, Toltén, and Valdivia drainages all share haplotype 95 and, again, we find a wide continental shelf in this area (Figs 1, 3; see also Supporting information, Appendix S1). Another interesting phylogeographic feature is that at the northernmost range of the *T. areolatus* where the continental shelf is most narrow, we find that two river drainages (i.e. the Río Limarí and Río Choapa) contain only a single haplotype (haplotype 1) for all 16 individuals sequenced.

DISCUSSION

SUPPORT FOR COMPETING HYPOTHESES

The focus of the present study was to describe the phylogeography of *T. areolatus*, aiming to better understand the historical biogeography of Chile.

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

Table 5. Results of the analysis of molecular variance tests evaluating two competing hypotheses to explain the partitioning of genetic variation in *Trichomycterus areolatus*

Hypothesis (number of groups)	Among groups	Among populations within groups	Within populations	F_{CT}
River Basins (24)	84.9	2.1	13.0	0.85*
Subprovinces (3)	45.3	44.7	10.0	0.45*

*$P < 0.001$.

The total genetic variation in the data set is partitioned into three categories: among groups, among populations within groups, and within populations. Groups were defined a priori to evaluate support for each hypothesis. Those hypotheses with more variation explained 'among groups' show greater relative support. Total genetic variation sums to 100%.

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

Freshwater aquatic species are particularly compelling as biogeographic models in this area given their potential susceptibility to geological events, to climate change and glaciation, and to dispersal barriers such as salt water in the ocean and changes in land formations. In view of the geological and climatic history of Chile, we generated three non-exclusive competing hypotheses that could explain the current distribution of genetic variation in *T. areolatus*. We consider support for each of these hypotheses in turn.

Sea level changes

The topography of central Chile and the ocean barrier appear to be the important factors that limit population connectivity between river basins. This is reflected by the low numbers of haplotypes shared between rivers, a pattern indicative of limited movement of *T. areolatus* between basins. In many systems, a decline in sea level allows rivers to coalesce on the exposed continental shelf, especially in areas where the continental shelf is wide. However, in Chile, the continental shelf is very narrow and few rivers are likely to connect during low sea levels (Fig. 1A). Interestingly, all but one of the phylogenetic breaks identified between *T. areolatus* clades correspond to areas with narrow continental shelf widths (Fig. 3). However, some areas with narrow continental shelves also separate several populations within clades. The one case where we found a distinct break between clades but without a narrow continental shelf occurred between clades E and F (Fig. 3). The patterns between these clades are complex, however, because the results of the present study suggest long-term isolation despite this being an area where a wide continental shelf occurs. However, some recent secondary contact has also occurred because the boundary between clades E and F overlaps across the Río Imperial and Río Toltén (Figs 2, 3).

We did find some evidence for shared haplotypes across drainages (e.g. the Limarí and Choapa basins in the north, Fig. 1; see Supporting information, Appendix S1). If the ocean is a barrier to *T. areolatus* movement, cases of shared haplotypes across drainages present a testable hypothesis that faunal exchange between these basins has occurred across drainage divides rather than via coastal connections. An alternative explanation for shared haplotypes across drainages is that of incomplete lineage sorting, rather than recent faunal exchange, where drainages may have historically been interconnected and later became isolated, resulting in shared haplotypes among systems but without any contemporary gene flow.

Glaciation

We found no association between the percentage of glaciation in drainages and contemporary levels of genetic diversity (Fig. 4). These results suggest that populations of *T. areolatus* in partially glaciated basins were not negatively affected by glaciers in a way that resulted in the loss of genetic diversity. Populations forced downstream by headwater glaciers apparently were large enough to maintain typical variation. It is clear, however, that glaciations have had a major effect on the southern distributional limit of *T. areolatus*; this species is not found anywhere on the mainland of Chile south of Río Maullín. All rivers from Río Maullín south experienced extensive glaciation with more than 85% of their catchments being covered by ice during the last glacial maximum (Table 1). Interestingly, there is a difference in the location of the ice within the catchments north and south of Río Maullín. Drainages from Río Maullín north were only glaciated in their headwater reaches and, thus, presumably fish recolonized from downstream reaches and unglaciated tributaries. South of Río Maullín ice covered all but the uppermost reaches of the drainages (Fig. 1B). This presents several potential explanations as to why *T. areolatus* does not occur in drainages south of Río Maullín. It is possible that they simply never occurred there. Alternatively, they did occur at an earlier time but were extirpated and thus unable to recolonize because movement between coastal rivers appears to be difficult (see

above). Although the last glaciation did not completely cover these catchments, an earlier glaciation may have. The largest glacial event in southern South America occurred 1–1.15 Mya (Coronato, Martinez & Rabassa, 2004) and may have more extensively covered these basins, extirpating any *T. areolatus* populations. It is notable that no obligate freshwater fishes found in central Chile, except *Percichthys trucha* (Cuvier & Valenciennes), occur further south on mainland Chile than Río Maullín (Dyer, 2000), which suggests that this glacial boundary is extremely important in the biogeography of Chilean fishes.

Faunal breaks

There is some evidence that phylogenetic differences within *T. areolatus* correspond to the fish community subprovince breaks identified by Dyer (2000). However, it is important to recognize that we found five phylogenetic breaks between *T. areolatus* lineages within the Chilean Province (between clades A, C, D, E, F, and G; Fig. 3), whereas Dyer (2000) only hypothesized two faunal breaks within this province based on patterns of endemism in the fish communities. The northern break is identified to be in the vicinity of the Río Mataquito, and the southern break at the Río Toltén (Fig. 1A). We did find evidence in *T. areolatus* for a separation of clades just south of the Río Mataquito. Samples from the Río Maipo south to Río Mataquito were all recovered as a single closely-related clade (C) with the major separation being between Río Mataquito and Río Maule (Fig. 3). We also found a separation between clades in the vicinity of Río Toltén; however, divergence patterns here were more complex. Clade E contained haplotypes from just south of Río Biobío to the Río Toltén. Clade F contained haplotypes from Río Imperial south to Río Valdivia. Thus, haplotypes from the Río Imperial and Río Toltén were shared between each clade (Fig. 3; see Supporting information, Appendix S1). This implies that a historical barrier may have existed in this area but the break was later crossed by *T. areolatus*, resulting in shared haplotypes between the Río Imperial and Río Toltén. Thus, it appears that the geographic divides among Dyer's fish communities are partially reflected in the phylogeography of *T. areolatus*, but that the overall pattern of geographic structuring in this species is more complex than that predicted by fish community boundaries alone.

Comparative phylogeography

The results obtained for *T. areolatus* are consistent in those found in the other fish species examined in Chile. Ruzzante *et al.* (2006) examined the genera *Percilia* and *Percichthys* from seven and three Chilean rivers, respectively, and found no sharing of haplotypes among rivers (although their sample sizes were often too small to detect rare haplotypes). Populations of *Percilia* from each river were genetically distinct, being separated by between six and sixteen DNA mutational steps. *Percichthys* from Río Andalién and Río Biobío were only one mutational step apart; however, populations in the next river north, Río Itata, differed from these two drainages by at least 30 steps. Such levels of divergence suggest long isolation between most percichthyid populations within these rivers, which is a pattern similar to those found in *T. areolatus*. The only other aquatic group examined genetically across Chile is the freshwater crabs of the genus *Aegla* (Pérez-Losada *et al.*, 2002). Most of these species are restricted to single river basins, with some basins having up to three species present. Only two species occur in more than one river basin: *Aegla abtao* Schmitt and *Aegla laevis* (Latreille). Genetic divergences between geographically isolated *Aegla* species are typically large (1.7–14.1%; mean = 7.1%). Genetic divergence is also substantial among river basins within crab species: *A. abtao* from Río Bueno and Río Valdivia differed by 2.6% (compared to 3.7–4.8% for *T. areolatus* from the same basins) and *A. laevis* from Río Maipo and Maule differed by 2.6–2.8% (compared to 2.0–2.8% for *T. areolatus* from the same basins). Interestingly, levels of genetic divergence between *Aegla* species and *T. areolatus* clades are of a similar magnitude. However, most *Aegla* species are endemic to individual rivers, whereas the major clades within *T. areolatus* occur across multiple river basins (Fig. 3), with considerably lower genetic divergences within clades (Table 3). These results suggest that most freshwater crab species have typically been isolated between individual river basins for considerably longer time frames than have *T. areolatus* if their rates of molecular evolution are similar.

Taxonomic implications

The results obtained in the present study have a number of interesting implications relative to the taxonomy of *T. areolatus*, as well as the status of the monotypic genus *Hatcheria*. There are at least two possible explanations for the placement of *H. macraei* within *T. areolatus*. It could be caused by introgressive hybridization of the mitochondrial genome of *T. areolatus* into *H. macraei* at different times in their history. Some introgression could be relatively recent based on the placement of several *H. macraei* haplotypes (H1–H6) within clade G, as well as the close placement of *H. macraei* individuals (H7–H10) to the remainder of clade G (Figs 2, 3). An alternative hypothesis for the close relationship of *H. macraei* to

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

T. areolatus is that the current taxonomy is incorrect, with *H. macraei* actually representing an individual lineage within *T. areolatus* that has morphologically diverged and specialized from other ancestral *T. areolatus* lineages. That *H. macraei* has been described as a distinct genus indicates that there are clear morphological differences to define these fish as being distinct from other species (Arratia *et al.*, 1978). The next step in the molecular work aiming to distinguish between these competing hypotheses will comprise an examination of DNA from multiple nuclear genes and an expansion of the sample size and geographic sampling within *H. macraei*, as well as the use of additional outgroup taxa. If data from nuclear genes are consistent with our mtDNA results, this would indicate that *H. macraei* evolved from an independent lineage of *T. areolatus*. If, however, the results of the nuclear DNA are consistent with current taxonomy then it provides evidence for introgressive events *of T. areolatus* mtDNA into *H. macraei* populations.

The taxonomy of *T. areolatus* itself might warrant additional consideration. Deep phylogenetic structuring within *T. areolatus* coupled with the observed patterns of geographic isolation among clades suggests that some of these clades are sufficiently distinct to be considered as independent taxonomic entities. Model-corrected sequence divergence among major clades varied in the range 1.0–9.1% (Table 3), with the upper level at a value commonly observed between sister species or even genera in other fishes (Johns & Avise, 1998). Given the observed patterns of phylogenetic divergence, the next step is to determine whether these would also meet the criteria of unique species under morphological and ecological species concepts (Sites & Marshall, 2004). Some preliminary data already exist indicating differences in morphology between populations from clades A and C (Pardo, 2002).

CONCLUSIONS

The paucity of phylogeographic work in the Southern Hemisphere (Beheregaray, 2008) presents an important opportunity: it promises to reveal the role of earth history events in shaping Southern Hemisphere biodiversity. In the present study, we focused on the role of glaciation and ocean barriers in driving the history of the freshwater fish *T. areolatus* in Chile. By contrast to studies conducted on many freshwater aquatic taxa in North America and Europe, we find that colonization of new habitats after glaciation is uncommon and gene flow among drainage systems is low. This can be best explained by the relatively short rivers that run east to west from high elevations in the Andes directly into the Pacific Ocean, and by a narrow continental shelf; when combined, these landscape features create little opportunity for genetic exchange via river capture over land or for river coalescence during low sea levels. In *T. areolatus*, this has resulted in high levels of genetic structuring among drainages. It also creates ideal conditions for local adaptation and genetic differentiation among geographically isolated populations. Hence, the engine creating new biodiversity in Chile for aquatic taxa may largely be controlled by earth history events. How such earth history events affect other taxa in the Southern Hemisphere remains an exciting prospect for future discovery.

ACKNOWLEDGEMENTS

We thank Néstor Ortiz and Waldo San Martín for their help with collecting tissue samples in the field. Benjamin Groves provided assistance in the laboratory. We are also grateful for feedback from Keith Crandall on earlier versions of the analyses and manuscript. This work was funded by a grant from the US National Science Foundation (NSF) PIRE program (OISE 0530267), and projects DIUC-Patagonia 205.310.042-ISP, DIUC 203.113.064-1.0 and DIUC 204.310.041-1.0 from Universidad de Concepción to EMH and PFV. Laboratory work was aided by instrumentation provided by NSF (DBI-0520978). Additional funding to J.B.J. from the BYU Kennedy Center supported A.P.B. in the conduction of field work.

REFERENCES

Arratia G. 1983. Habitat preferences of siluriform fishes of Chilean continental waters (Fam. Diplomystidae and Trichomycteridae). *Studies on Neotropical Fauna & Environment* **18:** 217–237.

Arratia G, Chang A, Menu-Marque S, Rojas G. 1978. About *Bullockia* n. gen. and *Trichomycterus mendozensis* n. sp. and revision of the family Trichomycteridae (Pisces, Siluriformes). *Studies on Neotropical Fauna and Environment* **13:** 157–194.

Avise JC. 2000. *Phylogeography: the history and formation of species*. Cambridge, MA: Harvard University Press.

Beheregaray LB. 2008. Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology* **17:** 3754–3774.

Burton RS. 1998. Intraspecific phylogeography across the Point Conception biogeographic boundary. *Evolution* **52:** 734–745.

Clapperton CC. 1993. *Quaternary geology and geomorphology of South America*. Amsterdam: Elsevier.

Clark PU, Mix AC. 2002. Ice sheets and sea level of the Last Glacial Maximum. *Quaternary Science Reviews* **21:** 1–7.

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

Coronato A, Martinez O, Rabassa J. 2004. Glaciations in Argentine Patagonia, southern South America. In: Ehlers J, Gibbard PL, eds. *Quaternary glaciations – extent and chronology, part III*. Amsterdam: Elsevier, 49–67.

Dyer BS. 2000. Systematic review and biogeography of the freshwater fishes of Chile. *Estudios Oceanológicos* **19:** 77–98.

ESRI [Environmental Systems Research Institute]. 1993. *Digital chart of the world*. Redlands, CA: Environmental Systems Research Institute.

Excoffier L, Laval G, Schneider S. 2006. *ARLEQUIN ver. 3.1: an integrated software package for population genetics data analysis*. Bern: Computational and Molecular Population Genetics Laboratory, University of Berne.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131:** 479–491.

Garcia-Lancaster A. 2006. *Aquatic habitat modeling of Chilean native fish in a reach of the Biobío River*. MSc Thesis, Department of Civil Engineering, University of Idaho.

Grosjean M, Geyh MA, Messerli B, Schreier H, Veit H. 1998. A late-Holocene (< 2600 BP) glacial advance in the south-central Andes (29oS), northern Chile. *The Holocene* **8:** 473–479.

Habit E, Belk M, Parra O. 2007. Response of the riverine fish community to the construction and operation of a diversion hydropower plant in central Chile. *Aquatic Conservation: Marine and Freshwater Ecosystems* **17:** 37–49.

Habit E, Victoriano P, Campos H. 2005. Ecología trófica y aspectos reproductivos de *Trichomycterus areolatus* (Pisces, Trichomycteridae) en ambientes lóticos artificiales. *Revista Biología Tropical* **52:** 195–210.

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

Hammer M, Adams M, Unmack PJ, Walker KF. 2007. *Retropinna* in retrospect: additional taxa and significant genetic sub-structure redefine conservation approaches for Australian smelts (Pisces: Retropinnidae). *Marine and Freshwater Research* **58:** 327–341.

Harrison S. 2004. The Pleistocene glaciations of Chile. In: Ehlers J, Gibbard P, eds. *Pleistocene Glaciations: Extent and Chronology*. Amsterdam: Elsevier, 89–103.

Hillis DM, Bull JJ. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology* **42:** 182–192.

Hulton NRJ, Purves RS, McCulloch RD, Sugden DE, Bentley MJ. 2002. The last glacial maximum and glaciation in southern South America. *Quaternary Science Reviews* **21:** 233–241.

Huson DH, Richter DC, Rausch C, Dezulian T, Franz M, Rupp R. 2007. Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinformatics* **8:** 460.

Johns GC, Avise JC. 1998. A comparative summary of genetic distances in the vertebrates from the mitochondrial Cytochrome *b* gene. *Molecular Biology and Evolution* **15:** 1481–1490.

Johnson JB, Dowling TE, Belk MC. 2004. Neglected taxonomy of rare desert fishes: congruent evidence for two species of leatherside chub. *Systematic Biology* **53:** 841–854.

Lehner B, Verdin K, Jarvis A. 2006. *HydroSHEDS technical documentation*. Washington, DC: World Wildlife Fund US. Available at: http://hydrosheds.cr.usgs.gov

Mardones M. 2002. Evolución morfogenética de la hoya del Río Laja y su incidencia en la geomorfología de la región del Biobío Chile. *Revista Geographyráfica de Chile Terra Australis* **49:** 97–127.

NGDC [US Department of Commerce, National Oceanic and Atmospheric Administration, National Geophysical Data Center]. 2006. 2-minute Gridded Global Relief Data (ETOPO2v2). Available at: http://www.ngdc.noaa.gov/mgg/fliers/06mgg01.html

Pardo R. 2002. Morphologic differentiation of *Trichomycterus areolatus* Valenciennes 1846 (Pisces: Siluriformes: Trichomycteridae) from Chile. *Gayana* **66:** 203–205.

Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* **14:** 817–818.

Pérez-Losada M, Jara CG, Bond-Buckup G, Crandall KA. 2002. Phylogenetic relationships among the species of *Aegla* (Anomura: Aeglidae) freshwater crabs from Chile. *Journal of Crustacean Biology* **22:** 304–313.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19:** 2496–2497.

Ruzzante DE, Walde SJ, Cussac VE, Dalebout ML, Seibert J, Ortubay S, Habit E. 2006. Phylogeography of the Percichthyidae in Patagonia: roles of orogeny, glaciation, and volcanism. *Molecular Ecology* **15:** 2949–2968.

SPSS. 2000. *SYSTAT 10 software*. Chicago, IL: SPSS.

Sites JW Jr, Marshall JC. 2004. Operational criteria for delimiting species. *Annual Review of Ecology, Evolution, and Systematics* **35:** 199–227.

Swofford DL. 2003. *PAUP*. Phylogenetic analysis using parsimony (* and other methods),* Version 4.0b10. Sunderland, MA: Sinauer.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24:** 1596–1599.

Turner KJ, Fogwill CJ, McCulloch RD, Sugden DE. 2005. Deglaciation of the eastern flank of the North Patagonian Icefield and associated continental-scale lake diversions. *Geographyrafiska Annaler* **87A:** 363–374.

Villagrán C. 1990. Glacial climates and their effect on the history of vegetation of Chile. A synthesis based on palynological evidence from Isla de Chiloé. *Review of Paleobotany and Palynology* **65:** 17–24.

Villagrán C, Hinojosa F. 2005. Esquema biogeográfico de Chile. In: Bousquets JL, Morrone JJ, eds. *Regionalización biogeográfica en Iberoámeríca y tópicos afines*. Jiménez Editores: Ediciones de la Universidad Nacional Autónoma de México, 551–577.

Wong BBM, Keogh JS, McGlashan DJ. 2004. Current and

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024

historical patterns of drainage connectivity in eastern Australia inferred from population genetic structuring in a widespread freshwater fish *Pseudomugil signifer* (Pseudomugilidae). *Molecular Ecology* **13:** 391–401.

Zwickl D. 2006. *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion*. PhD Thesis, University of Texas at Austin.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Table of haplotype counts in *Trichomycterus areolatus*, *Hatcheria macraei* (prefixed with H), and *Bullockia maldonadoi* (prefixed with B) based on the cytochrome *b* gene. The collection site is listed across the top, the sample size from each population is given in the last column. Each haplotype is listed along the left side with the total number of individuals with each haplotype listed in the last column.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Downloaded from https://academic.oup.com/biolinnean/article/97/4/876/2448016 by guest on 23 April 2024