DR. LILY C HUGHES (Orcid ID : 0000-0003-4006-4036)

Article type : Original Article

Biogeography, habitat transitions and hybridization in a radiation of South American silverside fishes revealed by mitochondrial and genomic RAD data

Lily C. Hughes[1,2*], Yamila P. Cardoso[3], Julie A. Sommer[4], Roberto Cifuentes[5], Mariela Cuello[6], Gustavo M. Somoza[7], Mariano González-Castro[8], Luiz R. Malabarba[9], Victor Cussac[10], Evelyn M. Habit[5], Ricardo Betancur-R.[2,11], and Guillermo Ortí [1,2]

[1] Department of Biological Sciences, George Washington University, Washington, DC, USA

[2] Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA

[3] Laboratorio de Sistemática y Biología Evolutiva-CONICET, Facultad de Ciencias Naturales y Museo, Universidad Nacional de La Plata, Argentina

[4] Vice Chancellor for Research Office, University of Nebraska Medical Center, Omaha, NE, USA

[5] Departmento de Sistemas Acuáticos, Facultad de Ciencias Ambientales y Centro EULA, Universidad de Concepción y Centro de Investigaciones en Ecosistemas Patagónicos, Concepción, Chile

[6] Facultad de Ciencias Naturales y Museo, Universidad Nacional de La Plata-CONICET, Buenos Aires, Argentina

[7] Instituto Tecnológico de Chascomus (CONICET-UNSAM), Chascomús, Buenos Aires, Argentina

[8] Grupo de Biotaxonomía Morfológica y molecular de peces, IIMyC-CONICET, Universidad Nacional de Mar del Plata, Mar del Plata, Argentina

[9] Departamento de Zoologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

[10] Centro Científico Tecnológico CONICET Patagonia Norte - Universidad Nacional del Comahue, Bariloche, Rio Negro, Argentina

[11] Department of Biology, University of Oklahoma, Norman, OK, USA

*Corresponding author: lilychughes@gmail.com

Abstract

Rivers and lake systems in the southern cone of South America have been widely influenced by historic glaciations events, carrying important implications on the evolution of aquatic organisms including prompting transitions between marine and freshwater habitats and by triggering hybridization among incipient species via waterway connectivity and stream capture events. Silverside fishes (*Odontesthes*) in the region comprise a radiation of 19 marine and freshwater species that have been hypothesized on the basis of morphological or mitochondrial DNA data to have either transitioned repeatedly into continental waters from the sea or colonized marine habitats following freshwater diversification. New ddRAD data presented here provide a robust framework to investigate biogeographic history and habitat transitions in *Odontesthes*. We show that *Odontesthes* silversides originally diversified in the Pacific but independently colonized the Atlantic three times, producing three independent marine-to-freshwater transitions. Our results also indicate recent introgression of marine mitochondrial haplotypes into two freshwater clades, with more recurring instances of hybridization among Atlantic- vs. Pacific-slope species. In Pacific freshwater drainages, hybridization with a marine species appears to be geographically isolated and may be related to glaciation events. Substantial structural differences of estuarine gradients between these two geographic areas may have influenced the frequency, intensity, and evolutionary effects of hybridization events.

Key Words: Single Nucleotide Polymorphism (SNP), Patagonia, glaciation, cytochrome *b*, introgression, *Odontesthes*.

1. Introduction

Compared to the hyper-diverse Neotropics, the temperate freshwater lakes and rivers of southern South America harbor a modest fish fauna, following a latitudinal trend towards lowest species diversity in southern Patagonia. Many fish species in Patagonia are endemic (Abell *et al.* 2008), of local economic and ecological importance (Cussac *et al.* 2004), and both the rapidly changing climate and introductions of invasive species are threats to these unique fishes (Strüssmann *et al.* 2010; Rueda *et al.* 2017; Becker *et al.* 2018). Species inhabiting this region have evolutionary histories heavily influenced by temperature, the rise of the Andes, historical connectivity of rivers and lakes, and glaciation events (Ortubay *et al.* 1997; Unmack *et al.* 2009, 2012; Habit *et al.* 2012; Cussac *et al.* 2016).

Freshwater colonization of the hyper-diverse Amazon and Paraná River Basins by marine lineages like drums, pufferfishes, and anchovies are well documented (Bloom & Lovejoy 2017). However, less is known about the origin, relationships, biogeographic and phylogeographic history of marine-derived lineages in southernmost South America, where repeated glaciation cycles extirpated freshwater fish populations or relegated them to refugia (Cussac *et al.* 2004, 2009; Ruzzante *et al.* 2006, 2008, 2011; Zemlak *et al.* 2008, 2010, 2011), and recent changes in sea level resulted in rapid recolonization (Cussac *et al.* 2004), or speciation (Beheregaray & Sunnucks 2001; Beheregaray *et al.* 2002). A majority of these studies have relied on mitochondrial DNA (mtDNA) or microsatellites for inferences, but applying modern high-throughput sequencing methods to these organisms should aid in elucidating the processes and histories underlying speciation among the unique fish lineages inhabiting this region.

Restriction-enzyme associated DNA (RAD) markers provide genome-wide data with demonstrated utility for resolving phylogenetic relationships among rapidly diverged species (Wagner *et al.* 2013) where a handful of PCR-based markers lack resolution. They also are useful for detecting gene flow and introgression among populations and species, processes that can confound phylogenetic inferences and species delimitation (Eaton & Ree 2013). Though genomic data have revolutionized phylogenetics, mtDNA markers, traditionally popular in the field of phylogeography (Avise *et al.* 1987), are still useful for understanding evolutionary patterns. Unlike nuclear markers, mtDNA does not readily recombine, and in cases of introgression it may be retained intact over generations. There are numerous reports of mtDNA haplotypes crossing species boundaries that unambiguously reveal ancient introgression events that, otherwise, may have left a weak or undetectable nuclear signature (Bryson *et al.* 2014; Willis *et al.* 2014; Good *et al.* 2015). Discordance between mtDNA genealogies and nuclear DNA phylogenies has been extensively debated, but it is generally accepted that incomplete lineage sorting of ancestral polymorphism, introgressive hybridization, and sex-biased dispersal rates are the most common causes of mito-nuclear discordance in various organisms (Toews & Brelsford 2012).

New World silversides (family Atherinopsidae) have repeatedly transitioned between marine and freshwater habitats, but precise knowledge of their phylogenetic history, and therefore the chronology and frequency of inferred habitat transitions, remains contentious (Dyer 2006; Bloom *et al.* 2013; Campanella *et al.* 2015). Some freshwater silverside species can osmoregulate in brackish water (Tsuzuki *et al.* 2000; Hughes *et al.* 2017; Silveira *et al.* 2018), suggesting that reversals from freshwater to marine habitats may be possible, despite this being a relatively rare

phenomenon among ray-finned fishes (Betancur-R *et al.* 2015). With seven marine and twelve freshwater recognized species (Eschmeyer & Fong 2017) distributed across coastal marine habitats and in rivers and lakes exclusively in temperate South America (Figure 1A-B), the atherinopsid genus *Odontesthes* (locally known as pejerrey in Spanish or peixe-rei in Portuguese) provides a unique system to study the history and evolutionary consequences of habitat transitions. A phylogenetic hypothesis based on morphological evidence (Figure 1C) implies that freshwater habitats comprised the ancestral range for this genus. Conversely, a recent molecular phylogeny (Campanella *et al.* 2015) reached the opposite conclusion (Figure 1D), instead requiring several instances of marine dispersal from the Pacific to the Atlantic and subsequent freshwater colonization by silversides. Studies based only on mtDNA sequences either included very few taxa or failed to resolve species boundaries and their relationships (Heras & Roldán 2011; García *et al.* 2014; González-Castro *et al.* 2016, 2019). A series of studies on species of *Odontesthes* along the southwestern Atlantic coastline have shown significant genetic and phenotypic differentiation between marine and estuarine or freshwater populations, as populations adapt to different salinities without physical barriers separating them to restrict gene flow (Endler 1977; Beheregaray *et al.* 2000, 2002; Beheregaray & Sunnucks 2001; González-Castro *et al.* 2016, 2019).

An influential study on the mode of diversification of silversides proposed a hypothesis based on repeated colonization of brackish lagoon systems and estuaries by marine populations, facilitated by adaptive genetic and phenotypic plasticity (Bamber & Henderson 1988), making them a ripe system to study the competing processes of gene flow and selection as populations invade new environments. The replicated nature of these ecological transitions in *Odontesthes* makes them

particularly appealing, but we do not currently have a clear understanding of how many times they have made marine-to-freshwater transitions. Several studies have suggested that a marine lineage currently represented by *O. argentinensis* has invaded estuaries multiple times to establish freshwater populations resulting in speciation, spawning both the phenotypically diverse "*O. perugiae* species-group" (Figure 1C) in the coastal lagoons of southern Brazil (Beheregaray *et al.* 2000, 2002; Beheregaray & Sunnucks 2001) and a unique population in the Mar Chiquita Lagoon in Argentina (González-Castro *et al.* 2016, 2019). However, other work contradicts the hypothesis that the *O. perugiae* species-group is marine-derived. Phylogenetic relationships proposed using morphology fail to place *O. argentinensis* as the sister-group to the *O. perugiae* complex (Dyer 1998, 2006) (Figure 1C), contradicting the hypothesis of a marine-derived origin for this group. Furthermore, published mtDNA genealogies (Heras & Roldán 2011; García *et al.* 2014; González-Castro *et al.* 2016, 2019) support a clade that includes the marine species *O. argentinensis* and all freshwater species endemic to the La Plata River Basin and the coastal lagoons of Uruguay and southern Brazil (e.g., the "*O. perugiae* group" plus *O. bonariensis*, *O. retropinnis* and *O. humensis*), but mtDNA consistently fails to establish species boundaries among these taxa or their phylogenetic relationships, potentially due to a history of gene flow among species (García *et al.* 2014) and limited information content of mitochondrial markers.

Hybridization between marine and incipient or established freshwater species has been proposed to be widespread among *Odontesthes* species in the southwestern Atlantic Basin (García *et al.* 2014; González-Castro *et al.* 2016, 2019). The coastline from northern Argentina to southern Brazil has many large estuaries and flood plains, such as the La Plata estuary, Mar Chiquita

Lagoon, and Dos Patos Lagoon, creating potential secondary contact zones between marine and freshwater species. In contrast, the steep southeastern Pacific coastline is punctuated with rivers draining snowmelt from the Andes mountain range, creating a sharper boundary between marine and freshwater habitats (Griffiths 2018), and making hybridization less likely. Given the relative ease with which mtDNA may cross species boundaries in the event of hybridization, nuclear data also are necessary to elucidate the evolutionary history of *Odontesthes*.

Here we analyze a large dataset of mtDNA sequences for species in *Odontesthes* to generate hypotheses about species boundaries and introgression that we further test with genome-wide RAD markers. We propose a new phylogenetic framework for this group and use it to formulate biogeographic hypotheses and to infer the history of habitat transitions. Finally, we explore a unique pattern of marine introgression detected by mtDNA analysis in a lake in southern Chile with nuclear data. The combination of mtDNA and nuclear RAD markers provides a powerful tool to disentangle species limits and to resolve phylogenetic relationships among species in the face of past or ongoing gene flow across species boundaries.

2. Materials and Methods

2.1 Sampling and DNA Extraction

Silversides were collected via seine or gill nets between 2006 and 2014, and immediately euthanized by an MS-222 overdose, with fin clips preserved in 95% ethanol for DNA extraction. Sampling localities in Argentina, Uruguay, Brazil, and Chile span the range of *Odontesthes* and are listed in Table S1 for cytochrome *b* sequences (cyt*b*), and Table S2 and Figure 1B for ddRAD sequences. *Odontesthes brevianalis* and *O. mauleanum* were difficult to distinguish in

the field, and are named here based on typical habitat, which is estuarine/riverine for *O. brevianalis*, and primarily lacustrine for *O. mauleanum*. These two species are separated by a single morphological character, the presence of "noticeably crenate scales" along the side of *O. mauleanum,* or restricted only to the caudal peduncle for *O. brevianalis* (Dyer 2006). Genomic DNA for cyt*b* sequencing was extracted with a DNeasy Blood & Tissue Kit (Qiagen). Samples for RAD sequencing were extracted in 96-plate format via Autogen automated DNA extraction at the Laboratory of Analytical Biology at the Smithsonian National Museum of Natural History.

2.2 mtDNA analysis

This dataset expands a published set of sequences collected to investigate putative hybridization between *O. hatcheri* and introduced *O. bonariensis* (Conte-Grand *et al.* 2015; Rueda *et al.* 2017). We used the forward primer GLU31 (Unmack *et al.* 2009) and the *Odontesthes*-specific reverse primer Pej15929 (Conte-Grand *et al.* 2015) to amplify the target fragment of the cyt*b* mitochondrial gene, following the same protocol as Rueda *et al*. (2017). PCR products were purified on 96-well Excelapure plates (Edge Biosystems), and cycle sequenced at the Brigham Young University DNA sequencing center. Contigs were assembled from raw chromatograms in Sequencher v. 4.8 (Gene Codes Corp.), and then aligned in MAFFT v1.30b (Katoh & Standley 2013). Our dataset for cyt*b* sequences included 450 individuals representing 14 nominal species of *Odontesthes* (of 19 described), and 4 sequences from its sister genus *Basilichthys*, collapsed into 171 unique haplotypes. The number of individuals per species is shown in parentheses in Figure 2A. New sequences are accessioned on Genbank (MK983245 - MK983379; Table S1).

Sequences were partitioned by codon position and the best substitution model for each partition was evaluated with ModelFinder (Kalyaanamoorthy *et al.* 2017), followed by Maximum Likelihood inference in IQ-Tree 1.6.0 (Nguyen *et al.* 2015), and 1000 bootstrap replicates using the ultra-fast bootstrap method (UFBoot) (Hoang *et al.* 2018). Haplotype networks for major clades were inferred with POPART 1.7 (Leigh & Bryant 2015), using median joining networks (Bandelt *et al.* 1999). Bayesian analysis was conducted in BEAST 2.4.8 (Bouckaert *et al.* 2014), partitioned by codon position with the best-fitting model determined by ModelFinder. The BEAST tree was time-calibrated using a relaxed clock approach based on secondary calibrations from Campanella *et al*. (2015). While fossil *Odontesthes* exist, they are not assigned to any particular species or clade (Dyer 2006). We used the 95% highest posterior density (HPD) for the split between *Odontesthes* and *Basilichthys* as a normal distribution for our secondary calibration (mean = 20.86 , standard deviation = 1.0), as well as the split of *O. nigricans* and *O. incisa* from all other *Odontesthes* species as a calibration (mean = 5.68, standard deviation = 1.0), but enforced no other constraints on the tree. We ran three independent runs for 100,000,000 generations on the Cipres Science Gateway (www.phylo.org), and assessed convergence in Tracer 1.6.0.

2.3 ddRAD sequencing and matrix assembly

We followed the protocol developed by Peterson *et al*. (2012), modified to use the enzymes *Mse*I and *Pst*I and a 350-550 bp size selection. Genomic libraries for 164 *Odontesthes* samples representing 13 species (Table S2) and five *Basilichthys microlepidotus* were prepared at the University of Puerto Rico Sequencing and Genomics Facility. Pooled samples were sequenced on a half-lane of a HiSeq 4000 at the University of Chicago Genomics Facility. Sequences were

demultiplexed, filtered, and assembled into RAD loci using ipyrad v. 0.7.2 (Eaton 2014). We used the default parameters for *de novo* assembly of ddRAD data, with a minimum sequence depth of 6 for base-calling and a cluster threshold of 0.85; reads with more than 5 low-quality base calls were discarded. For our complete dataset, we only retained loci that were present for at least 77 individuals, which included outgroup *Basilichthys microlepidotus* samples. Additional matrices were generated removing the outgroup samples, enforcing a minimum of 75 individuals per locus, and 100 individuals per locus.

2.4 Phylogenomics

We used both concatenation and multi-species coalescent approaches for phylogenetic inference. Concatenated ddRAD loci were analyzed under Maximum Likelihood using IQ-Tree (Nguyen *et al.* 2015) and the GTR+G model, and 1000 UFBoot replicates to assess branch support (Hoang *et al.* 2018). Unlinked SNPs (1 per locus) were analyzed under the multispecies coalescent model in SVDQuartets (Chifman & Kubatko 2014) available in PAUP* 4.0a164 (Swofford 2017) using all possible quartet combinations.

2.5 Biogeography

To reconstruct the biogeographic history of *Odontesthes* with our phylogenetic resolution of the group, we first generated an ultrametric tree using only one tip per species scaled to time using treePL1.0 (Smith & O'Meara 2012), an implementation of the penalized-likelihood approach (Sanderson 2002). We used secondary calibrations from the 95% highest posterior densities from the fossil-calibrated analysis of Campanella *et al*. (2015) for the group at three points: the divergence of *Basilichthys microlepidotus* from all *Odontesthes* 20.0-23.04 Ma, the divergence

of *O. incisa* from all other *Odontesthes* species in our ddRAD phylogeny 3.27-6.93 Ma, and the divergence between *O. argentinensis* and other freshwater lineages in the Paraná and Uruguay rivers 1.81-0.68 Ma. The time-calibrated tree was necessary to reconstruct the biogeography of the group, but we are not testing the effect of any particular major geological event; for example, the sustained uplift of the Andes mountains approximately 12 million years ago (Garzione *et al.* 2008) took place well before the inferred origin of extant *Odontesthes* species.

We defined five areas that *Odontesthes* currently occupies (Figure 1A): eastern Pacific Ocean, eastern Pacific freshwater drainages, Patagonia (restricted to only the Atlantic-draining portion east of the Andes), western Atlantic Ocean, and western Atlantic freshwater drainages (excluding Patagonia, but including the Pampas region, La Plata, Paraná and Uruguay rivers, Patos and Tramandaí lagoon systems in Brazil). Though there is no geographic barrier between Atlantic-draining Patagonia and warmer freshwater habitats to the northeast, it is likely too cold for the freshwater representatives of the *O. argentinensis* species-group, some of which have temperature-dependent sex determination (Strussmann *et al.* 1996; Cussac *et al.* 2009). We restricted dispersal between non-adjacent areas, so that dispersal could not occur directly between the Pacific Ocean and Atlantic freshwater drainages, Pacific and Atlantic freshwater drainages, and Pacific freshwater drainages and Patagonia, where the Andes mountain range creates a physical barrier. Additionally, ancestors could only inhabit two of the five areas at one time, although no modern *Odontesthes* species occupies more than one of these ranges. Reconstruction was performed in the BioGeoBEARS R package (Matzke 2013a) under all available models (Matzke 2013b, 2014), with model selection performed using the Akaike information criterion (AIC). R code is available on Figshare (Hughes *et al.* 2019).

2.7 Admixture Analysis with fastSTRUCTURE and Patterson's D

To examine possible nuclear admixture in our samples with ddRAD-derived SNP data, we ran clustering analyses using fastSTRUCTURE (Raj *et al.* 2014) on individuals from two clades where species were not monophyletic in our cyt*b* dataset (Figure 2A). The first clade was the *O. argentinensis*-species group, represented in our ddRAD dataset by *O. argentinensis* (n=18)*, O. bonariensis* (n=15)*, O. ledae* (n=13)*,* and *O. mirinensis* (n=11) (Figure 2B). The species *O. humensis* and *O. perugiae* also fall in this clade, but since we were only able to obtain one individual for each of these, we excluded them from this analysis. The second group included *O. regia* (n=26), *O. gracilis* (n=3), *O. smitti* (n=7), *O. brevianalis* (n=9)*,* and *O. mauleanum* (n=34). We excluded a few samples with high missing data. Due the hierarchical nature of these groups, we decomposed each of these two initial groups into subgroups until the subgroup contained only individuals of one species, or the optimal value of K was 1 as determined by the chooseK.py script packaged with fastSTRUCTURE. For each dataset and each value of K, we ran 10 replicates, and tested values of K between 1 and 6.

We calculated Patterson's *D* for the two groups above, where mtDNA suggested introgression. This statistic is commonly known as the ABBA-BABA test (Durand *et al.* 2011). This test uses biallelic SNPs based on a four-taxon pectinate tree, including one outgroup taxon to determine which allele is ancestral (the 'A' allele), and which is derived ('B'), with the lineages typically denoted as (Outgroup,(P3,(P2,P1))). The *D*-statistic is calculated based on the number of patterns that conflict with the tree (the 'ABBA' and 'BABA' patterns). These patterns should occur in equal frequencies if incomplete lineage sorting is the only process acting, resulting in *D*=0. But if

introgression is occurring between P3 and either P1 or P2, then there will be an excess of either 'ABBA' or 'BABA' patterns. We used DSuite, a tool that calculates *D*-statistics from VCF files and assesses significance using jackknifing (Malinsky 2019). Under this framework, we calculated *D* for two separate clades, with *Odontesthes hatcheri* samples used as the outgroup population in both analyses. For the clade composed of primarily Pacific-drainage taxa, we treated all samples of marine species (*O. smitti*, *O. gracilis*, and *O. regia*) as a single population, given that fastSTRUCTRE did not differentiate amongst these species, but used the Lake Llanquihue (LLA) *O. mauleanum* population as 'P2' and the Lake Calafquen *O. mauleanum* population as 'P1'. In the *O. argentinensis*-species group, *O. argentinensis* represented 'P3', *O. bonariensis* 'P2', and *O. mirinensis* plus *O. ledae* samples were used for 'P1'.

3. Results

3.1 mtDNA

One representative from each new haplotype sequence has been deposited on GenBank (MK983245-MK983397). Sequence alignments and newick tree files are available on Figshare (DOI: 10.6084/m9.figshare.11413209). ModelFinder supported different substitution models for each of the three partitions: K2P+I+G for the first codon position, HKY+I+G for the second, and TN+G for the third. Maximum likelihood analysis supported the monophyly of the genus *Odontesthes* and distinguished seven well-supported clades (99-100% bootstrap support; Figure 2A), but relationships among clades are not resolved with confidence (<95%, as the recommended interpretation of the UFBoot method). Haplotype networks are shown in Figure S1. Most haplotypes from freshwater *O. brevianalis* or *O. mauleanum* were contained in a single

315 well-supported haplogroup that showed no segregation of haplotypes between species (Figure
316 2A, Figure S1). However, individuals from a few freshwater populations in southern Chile, Lake
317 Llanquihue (LLA) and its tributary the Pescado River (PESC) carry divergent mtDNA
318 haplotypes that are closely related to the marine species *O. regia* and *O. smitti* (Figure 2A,
319 Figure S1). *Odontesthes gracilis* nests within *O. regia* haplotypes (Figure 2A, Figure S1), but *O.*
320 *smitti* haplotypes collected from the southwestern Atlantic (MADR and MDP; Figure 1B), are
321 monophyletic (Figure 2A). The divergence of sequences from Lake Llanquihue and the Pescado
322 River from its sister clade formed by the *O. regia-smitti* haplogroup dated to the Pleistocene
323 (mean age 744,000, 95% HPD 1.2 million-154,700 years ago) (Figure S2). Sequences obtained
324 from Patagonian *O. hatcheri* clustered into a distinct haplogroup, with the exception of some
325 hybrid individuals that carried *O. bonariensis* haplotypes, where this species has been introduced
326 into Patagonia (Conte-Grand *et al.* 2015; Rueda *et al.* 2017). Another clade includes haplotypes
327 from species associated with the Atlantic Basin (marine *O. argentinensis*, and freshwater *O.*
328 *bonariensis*, *O. ledae, O. mirinensis, O. perugiae,* and *O. piquava*) that did not sort according to
329 species boundaries (Figure 2A, Figure S1).
330
331 *3.2 ddRADseq species phylogeny*
332 Using the *de novo* assembly pipeline in ipyrad, we assembled three ddRAD matrices with
333 different amounts of missing data. In matrices without *B. microlepidotus* samples, we rooted the
334 trees on *O. incisa*, which is the first branch of *Odontesthes* in all analyses that include the
335 outgroup (Figure S2). Resulting phylogenies based on the three matrices produced nearly
336 identical results between concatenation and multispecies coalescent SVDQuartets analyses
337 (Figures S3-S7), resolving relationships with confidence for major lineages (Figure 2B),

although relationships among individuals within these groups differed. Not all morphospecies were monophyletic, namely *O. ledae, O. mirinensis, O. mauleanum, O. brevianalis,* and *O. gracilis* (Figure 2B, Figures S3-S7).

In agreement with the mtDNA genealogy, the RAD phylogenies resolve *O. incisa* as the sister group of all other species in the genus (no ddRAD data were obtained for *O. nigricans* due to the poor quality of DNA extractions from our available specimens). The remaining *Odontesthes* species split into two groups, mostly corresponding to geography. One is a primarily Pacific Basin clade from localities west of the Andes that contains the two freshwater Chilean species *O. brevianalis* and *O. mauleanum* and the marine *O. regia* and *O. gracilis*, plus *O. smitti* from Atlantic localities (Figure 1B, 2B). *Odontesthes mauleanum* and *O. brevianalis* do not form well-supported separate clades in either concatenation or multispecies coalescent trees, but some structure grouping individuals from the same collection localities was observed (Figure 4B). Individuals from LLA that carried the marine *O. regia/gracilis* mtDNA haplotype (Figure 2A) were grouped by the RAD data with other freshwater populations of *O. mauleanum* and *O. brevianalis* with high support (Figure 2B). *Odontesthes gracilis* individuals never formed a monophyletic group. The second major clade is in the Atlantic Basin, containing *O. hatcheri* as the sister to the *O. argentinensis* species group (Figure 2B). In stark contrast to the cyt*b* genealogy, which does not differentiate the morphological species (Figure 2A), analyses of genome-wide ddRADseq data clearly delineate *O. argentinensis* and *O. bonariensis* into monophyletic groups, excepting *O. ledae* and *O. mirinensis*, though just one individual each of *O. humensis* and *O. perugiae* was available.

3.3 Biogeography

AIC scores selected the best-fit DIVA-like+J model available in BioGeoBEARS, a likelihood implementation of the parsimony-based DIVA model (Ronquist 1997), with the addition of the 'jump dispersal' parameter 'J' (Matzke 2013b, 2014). All models with the 'J' parameter had similar AIC scores, and resulted in highly similar biogeographic reconstructions, inferring a southeastern Pacific Ocean origin for all extant *Odontesthes*, three separate transitions of the Atlantic Ocean, and three separate transitions into freshwater environs (Figure 3).

3.4 Admixture Analysis of Nuclear SNPs with fastSTRUCTURE and Patterson's D

Among members of the *O. argentinensis*-species group, which share indistinguishable mtDNA haplotypes (Figure 2A), fastSTRUCTURE indicated two genetic clusters, one containing at *O. argentinensis* individuals, and the other cluster contained the three remaining freshwater species, *O. bonariensis, O. mirinensis,* and *O. ledae* (Figure 4A). Analyzing this second cluster separately, *O. bonariensis* formed a separate group from the two remaining Brazilian species, but *O. ledae* and *O. mirinensis* did not form two separate clusters when analyzed alone. There was no evidence of nuclear admixture in individual population assignments.

When all individuals from *O. regia*, *O. smitti, O. gracilis, O. brevianalis,* and *O. mauleanum* were included in our fastSTRUCTURE analysis, the optimal value of K was two (Figure 4B), and individuals from the *O. mauleanum* population Lake Llanquihue (LLA) did not show evidence of nuclear admixture with marine congeners, despite what their mtDNA signal suggests (Figure 2A). Analyzing these two clusters separately, the three marine species *O. regia*, *O. smitti,* and *O. gracilis* did not form separate genetic clusters. Among *O. brevianalis* and *O. mauleanum*

384 individuals, there were three clusters: a population for all Lake Calafquen (LCAL) individuals, a
385 population with all Lake Llanquihue (LLA) individuals, and individuals from all remaining
386 populations formed a third cluster (RMAU, HUI, TOP, REL; Figure 4B).
387
388 The *D*-statistic calculated for the *O. argentinensis*-group was greater than zero, though not
389 substantially so at $D = 0.056$ ($p = 0.0013$), suggesting weak introgression between *O.*
390 *argentinensis* and *O. bonariensis* (Figure 4C). The *D*-statistic using *O. hatcheri* again as an
391 outgroup, *O. regia*, *O. smitti*, and *O. gracilis* individuals as P3, LLA *O. mauleanum* population
392 individuals as P2, and *O. mauleanum* individuals from LCAL as P1 was slightly larger at $D =$
393 0.132 ($p < 0.0000001$), suggesting some introgression between the marine individuals and the
394 LLA population.
395
396 **4. Discussion**
397 *4.1 Resolution of species boundaries and phylogeny*
398 Until now, morphological and mtDNA hypotheses have been the primary resource available to
399 understand the evolution South American silversides. Previous studies (García *et al.* 2014;
400 Campanella *et al.* 2015; González-Castro *et al.* 2016, 2019) failed to separate many nominal
401 species, especially in the Atlantic Basin (Figure 1C, D), and to confidently establish phylogenetic
402 relationships in the *O. argentinensis*-species group. In contrast, genomic ddRAD data show
403 strong support for the current taxonomy with a few exceptions and confidently establish
404 phylogenetic relationships. Most morphologically described species were identified as
405 monophyletic, with a few notable exceptions. The *O. argentinensis*-species group,
406 indistinguishable with mtDNA, mostly sorted into monophyletic groups where multiple

individuals per species could be obtained, except for the two Brazilian species *O. ledae* and *O. mirinensis*, which fastSTRUCTURE also did not identify as separate genetic clusters. Genetic differentiation between *O. regia* and *O. gracilis* is not supported by the RAD data (Figure 2B), suggesting that the species in the Juan Fernandez islands (*O. gracilis*) is not genetically different from the common species distributed along the coast of Chile (*O. regia*). Additional sampling to enable explicit assessment of species boundaries is necessary before recommending a taxonomic decision to classify *O. gracilis* Steindachner 1898 as a junior synonym of *O. regia* Humboldt 1821, but notably fastSTRUCTURE did not distinguish between the three marine species *O. gracilis, O. regia,* and *O. smitti. Odontesthes mauleanum* and *O. brevianalis* are not reciprocally monophyletic, though they were difficult to distinguish in the field. Broader geographic sampling is needed to confidently assess these species boundaries, fastSTRUCTURE suggested some population structure but did not clearly separate *O. brevianalis* and *O. mauleanum* populations into distinct species clusters.

Our results are similar to the molecular phylogeny proposed by Campanella *et al*. (Figure 1D) but differ significantly from the hypothesis based on morphology (Figure 1C), especially regarding the implied transitions between marine and freshwater habitats. The morphological hypothesis placed all marine species except *O. argentinensis* in a derived position, suggesting a recent invasion of the marine environment by freshwater species. The phylogenetic framework supported by the RAD data (Figure 2B), although not complete, sheds new light on the processes underlying the diversification of these fishes and their radiations into freshwater habitats.

4.2 Silverside habitat transitions and biogeography

The phylogenomic hypothesis has significant implications for this group's biogeography. Ancestral area reconstruction (Figure 3) places the origin for all species of *Odontesthes* in the eastern Pacific Basin, unsurprising given that the sister group to *Odontesthes* and *Basilichthys*, Atherinopsini, is a northeastern Pacific marine clade (Dyer 1997; Bloom *et al.* 2012, 2013; Campanella *et al.* 2015). This Pacific origin implies at least three separate events of colonization of the Atlantic, and it supports three independent invasions of freshwater habitats by marine ancestors. One of these inferred invasions triggered a species radiation into the La Plata Basin that currently extends over much of the Pampas region and freshwater habitats ranging through inland and coastal water habitats of Uruguay and southern Brazil (Figure 3), encompassing nine described species. Our results strongly suggest that the early radiation of freshwater species—represented in our study by *O. bonariensis*, *O. humensis*, and *O. perugiae*—transpired along the major Paraná and Uruguay river basins and the shallow lakes in the Pampas region in the Buenos Aires province of Argentina and Uruguay, and only recently reached the coastal lagoon systems in southern Brazil (represented in this study by *O. ledae* and *O. mirinensis*). *Odontesthes hatcheri* populations in Patagonian lakes and rivers east of the Andes are likely from a separate, earlier transition to fresh water. The third freshwater clade (*O. brevianalis* and *O. mauleanum*) is found west of the Andes, in rivers and lakes flowing towards the Pacific Ocean, sister to marine *O. regia* and *O. smitti*.

The pattern of evolution of freshwater species from marine ancestors implied by the RAD phylogeny does not support a previously proposed ecological model of speciation involving divergence with gene flow in southwestern Atlantic populations (Beheregaray & Sunnucks 2001; Beheregaray *et al.* 2002; García *et al.* 2014). According to this model, marine populations of *O.*

453 *argentinensis* entering the coastal lagoons and brackish estuaries along the coastal plain of
454 Uruguay (Mirim), southern Brazil (dos Patos, Tramandaí), and Argentina (Mar Chiquita) gave
455 rise repeatedly to phenotypically variable species flocks, similar to well-documented marine-to-
456 freshwater transitions in marine stickleback fishes in Asia, Europe and North America (Bell &
457 Foster 1994). In sticklebacks, independently derived freshwater populations are nested in
458 divergent marine clades (Fang *et al.* 2018). The RAD phylogeny for *Odontesthes*, in contrast,
459 strongly supports a single group containing freshwater species sister to the marine *O.*
460 *argentinensis* (Figure 2B). Species in the coastal lagoon systems were not directly derived from
461 ancestral *O. argentinensis* stock, but derived from other freshwater populations in the "*O.*
462 *perugiae* group" (Figure 3). This single transition to freshwater from a marine ancestor in a
463 particular geographic area is a repeated pattern in marine-derived freshwater fishes including
464 clingfishes (Conway *et al.* 2017), anchovies (Bloom & Lovejoy 2012), drums and pufferfishes
465 (Bloom & Lovejoy 2017), though silversides at large have colonized freshwater habitats in the
466 Americas many times (Bloom *et al.* 2013; Campanella *et al.* 2015).
467
468 *4.3 mtDNA introgression and hybridization*
469 Secondary contact between the established freshwater species and their marine relatives seems to
470 be frequent in brackish and estuarine habitats where environmental gradients are gradual and
471 form extensive habitats, enabling the formation of hybrids exhibiting intermediate phenotypes.
472 Lack of genetic differentiation among species in mtDNA markers may be the consequence of
473 hybridization, as is likely the case with the *O. argentinensis* species-group (García *et al.* 2014).
474 The numerous estuaries and coastal lagoons that have formed on the southwestern Atlantic
475 coastline from northern Argentina to southern Brazil provide ample contact zones for marine and

476 freshwater species that could homogenize mtDNA, with many of the freshwater species in the *O.*
477 *argentinensis* species-group being sympatric as well. Intermediate phenotypes between marine
478 and freshwater species in the Mar Chiquita coastal lagoon system have been documented, but
479 were interpreted as incipient speciation events (González-Castro *et al.* 2016, 2019). Signature for
480 nuclear introgression among *O. argentinensis, O. bonariensis, O. mirinensis* and *O. ledae* is
481 weak according to the ABBA-BABA test, but broader geographic sampling across the complex
482 coastline they inhabit would likely clarify how these interactions play out in local systems.
483 Additionally, selection may play a significant role in limiting the flow of nuclear alleles between
484 marine and freshwater, where vastly different physiological mechanisms are needed for survival.
485
486 Natural hybrids may occur among broadly sympatric species or can be confined to particular
487 contact areas, and can be detected through the observation of intermediate phenotypes or
488 discovery of introgressed mtDNA haplotypes or nuclear alleles cutting across the presumed
489 contact area (Harrison 1990; Funk & Omland 2003; Willis *et al.* 2014). Possible hybrid
490 silversides have been inferred in estuarine areas where marine and freshwater species come into
491 secondary contact (González-Castro *et al.* 2016), but it is unclear whether these hybrid zones are
492 persistent or ephemeral. The fitness of hybrid individuals in the area of secondary contact is an
493 important factor determining the stability of hybrid zones (Barton & Hewitt 1985). Hybrids may
494 be maintained when they have higher fitness than either parental type under intermediate
495 ecological conditions, but alternatively, they may be ephemeral when the hybrid zone is
496 maintained through a balance between hybrid inferiority and continuous invasion of parental
497 types into the area (Moore & Buchanan 1985). The latter scenario may lead to conditions of
498 "genetic tension" among loci coding for important physiological traits.

A recent study comparing gene expression profiles in gills of wild-caught *O. bonariensis* (fresh water) and *O. argentinensis* (marine), revealed significant differences between these two species for more than 3,000 transcripts, some of them coding for osmoregulatory/ion transport and immune system functions (Hughes *et al.* 2017). Additionally, more than one thousand transcripts with nonsynonymous SNPs in the coding sequences were found to be different between the species. These genetic disparities between species likely reflect adaptive responses to critical challenges in relation to transitions from marine to freshwater conditions. The silverside hybrid zones reported in nature may therefore represent genetic tension zones rather than persistent hybrid swarms where incipient species may be forming, but detailed studies of these hybrid zones are necessary to confirm this hypothesis. If a genetic tension zone limits the flow of nuclear alleles, it does not seem to affect the flow of mtDNA haplotypes among the *O. argentinensis* species group (Figure 2A), perhaps due to their relative neutrality.

Steep gradients between marine and freshwater habitats are more typical in southern Argentina and Chile, where rivers originating in the Andes drain straight into the ocean providing scant opportunities for co-existence and hybridization. Freshwater species in Patagonia (*O. hatcheri* in Argentina and *O. brevianalis* and *O. mauleanum* in Chile) show sharply distinct haplogroups from their marine relatives, a pattern not found in the *O. argentinensis* species-group where secondary contact zones abound (Figure 2A). However, we report an exceptional case suggesting a recent mitochondrial capture of marine haplotypes characteristic of *O. regia* and *O. smitti* by a freshwater population of *O. mauleanum* in Lake Llanquihue. We hypothesize based on these biological results that this pattern could have been caused by mitochondrial capture of the marine

haplotype following a marine incursion into freshwater habitats, creating a temporary secondary contact zone between *O. mauleanum* and *O. regia*. This haplotype appears to be common in Lake Llanquihue, a large lake in the southern Chilean Lake District, only ~22 km away from Reloncaví Sound to the south. Strong selection on nuclear loci and smaller population sizes in freshwater lakes could have erased this signal, though the ABBA-BABA test suggested some nuclear introgression between marine individuals and the Llanquihue population compared to the higher-elevation Lake Calafquen. Lake Llanquihue was likely covered by the Llanquihue glacier at least three times, between 70,000 years ago and finally during the last glacial maximum 14,000 years ago (Clapperton 1994; Harrison & Glasser 2011), but the origin of the introgressed haplotypes appears to be older, suggesting that it might be related to the end of the coldest Patagonian glaciation ~700,000 years ago, or the largest Patagonian glaciation 1.2 million years ago (Ruzzante *et al.* 2008), although we do not know of geological evidence that points to marine incursions during these glaciation cycles. The signal of this secondary contact and introgression would not have been apparent without the mtDNA data, highlighting its continued importance in the genomic era.

5. Conclusions

Our newly generated genome-wide ddRAD data provided remarkable resolution to the phylogenetic relationships of *Odontesthes* silverside fishes (Figure 2B), which was critical to understanding the history of habitat shifts, biogeography, and introgression events in this group. This new hypothesis implies three separate transitions to freshwater habitats (Figure 3), one in southern Chile, one in the Atlantic drainages of Patagonia, and one in the lower Paraná and Uruguay river basins, spawning a species radiation that now occupies coastal fresh waters from

545 the Pampas region in Buenos Aires province, Argentina, across Uruguay, to southern Brazil in
546 the dos Patos and Tramandaí lagoon systems. Mitochondrial cyt*b* data were mostly
547 uninformative for assessing phylogeny and species boundaries (Figure 2A), but uniquely
548 highlight the contrasting histories of introgression between marine and freshwater congeners on
549 both the western and eastern coasts of southern South America. Mitochondrial data support
550 secondary contact between *O. mauleanum* and *O. regia* in or near the Lake Llanquihue locality
551 in southern Chile (Figure 2A). In contrast, the relatively homogenous cyt*b* haplotypes found
552 across species in the *O. argentinensis* species group suggest on-going gene flow and complex
553 scenarios of speciation-with-gene flow across habitats (García *et al.* 2014; González-Castro *et al.*
554 2016, 2019) that should be investigated with deeper population sampling to understand these
555 dynamics more clearly to ascertain if incipient speciation is taking place.
556
557 **References**
558

559 Abell R, Thieme ML, Revenga C *et al.* (2008) Freshwater Ecoregions of the World: A New Map
560 of Biogeographic Units for Freshwater Biodiversity Conservation. *BioScience*, **58**, 403–414.
561 Avise JC, Arnold J, Ball RM *et al.* (1987) Intraspecific Phylogeography : The Mitochondrial
562 DNA Bridge Between Population Genetics and Systematics. *Annual Review of Ecology and*
563 *Systematics*, **18**, 489–522.
564 Bamber RN, Henderson PA (1988) Pre-adaptive plasticity in atherinids and the estuarine seat of
565 teleost evolution. *Journal of Fish Biology*, **33**, 17–23.
566 Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific
567 phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.
568 Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual review of Ecology and*
569 *Systematics*, **16**, 113–148.
570 Becker LA, Crichigno SA, Cussac VE (2018) Climate change impacts on freshwater fishes: a
571 Patagonian perspective. *Hydrobiologia*, **816**, 21–38.
572 Beheregaray LB, Levy JA, Gold JR (2000) Population genetics of the silverside *Odontesthes*
573 *argentinensis* (Teleostei, Atherinopsidae): evidence for speciation in an estuary of southern
574 Brazil. *Copeia*, **2000**, 441–447.
575 Beheregaray LB, Sunnucks P (2001) Fine-scale genetic structure, estuarine colonization and
576 incipient speciation in the marine silverside fish *Odontesthes argentinensis*. *Molecular*

Ecology, **10**, 2849–2866.
Beheregaray LB, Sunnucks P, Briscoe DA (2002) A rapid fish radiation associated with the last sea-level changes in southern Brazil: the silverside *Odontesthes perugiae* complex. *Proceedings of the Royal Society B: Biological Sciences*, **269**, 65–73.
Bell MA, Foster SA (1994) *The evolutionary biology of the threespine stickleback*. Oxford University Press, USA.
Betancur-R R, Ortí G, Pyron RA (2015) Fossil-based comparative analyses reveal ancient marine ancestry erased by extinction in ray-finned fishes. *Ecology Letters*, **18**, 441–450.
Bloom DD, Lovejoy NR (2012) Molecular phylogenetics reveals a pattern of biome conservatism in New World anchovies (family Engraulidae). *Journal of Evolutionary Biology*, **25**, 701–715.
Bloom DD, Lovejoy NR (2017) On the origins of marine-derived freshwater fishes in South America. *Journal of Biogeography*, **44**, 1927–1938.
Bloom DD, Unmack PJ, Gosztonyi AE, Piller KR, Lovejoy NR (2012) It's a family matter: Molecular phylogenetics of Atheriniformes and the polyphyly of the surf silversides (Family: Notocheiridae). *Molecular Phylogenetics and Evolution*, **62**, 1025–1030.
Bloom DD, Weir JT, Piller KR, Lovejoy NR (2013) Do freshwater fishes diversify faster than marine fishes? A test using state-dependent diversification analyses and molecular phylogenetics of new world silversides (Atherinopsidae). *Evolution*, **67**, 2040–2057.
Bouckaert R, Heled J, Kühnert D *et al.* (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS computational biology*, **10**, e1003537.
Bryson RW, Smith BT, Nieto-Montes de Oca A, García-Vázquez UO, Riddle BR (2014) The role of mitochondrial introgression in illuminating the evolutionary history of Nearctic treefrogs. *Zoological Journal of the Linnean Society*, **172**, 103–116.
Campanella D, Hughes LC, Unmack PJ *et al.* (2015) Multi-locus fossil-calibrated phylogeny of Atheriniformes (Teleostei, Ovalentaria). *Molecular Phylogenetics and Evolution*, **86**, 8–23.
Chifman J, Kubatko L (2014) Quartet Inference from SNP Data Under the Coalescent Model. *Bioinformatics*, **30**, 3317–3324.
Clapperton CM (1994) The quaternary glaciation of Chile: a review. *Revista Chilena de Historia Natural*, **67**, 369–383.
Conte-Grand C, Sommer J, Ortí G, Cussac V (2015) Populations of *Odontesthes* (Teleostei: Atheriniformes) in the Andean region of Southern South America: body shape and hybrid individuals. *Neotropical Ichthyology*, **13**, 137–150.
Conway KW, Kim D, Rüber L, Espinosa Pérez HS, Hastings PA (2017) Molecular systematics of the New World clingfish genus *Gobiesox* (Teleostei: Gobiesocidae) and the origin of a freshwater clade. *Molecular Phylogenetics and Evolution*, **112**, 138–147.
Cussac VE, Fernandez DA, Gomez SE, Lopez HL (2009) Fishes of southern South America: a story driven by temperature. *Fish Physiology and Biochemistry*, **35**, 29–42.
Cussac VE, Habit E, Ciancio J *et al.* (2016) Freshwater fishes of Patagonia: conservation and fisheries. *Journal of Fish Biology*, **89**, 1068–1097.
Cussac VE, Ortubay S, Iglesias G *et al.* (2004) The distribution of South American galaxiid fishes: the role of biological traits and post-glacial history. *Journal of Biogeography*, **31**, 103–121.
Durand EY, Patterson N, Reich D, Slatkin M (2011) Testing for Ancient Admixture between Closely Related Populations. *Molecular Biology and Evolution*, **28**, 2239–2252.
Dyer BS (1997) Phylogenetic Revision of Atherinopsinae (TELEOSTEI, ATHERINOPSIDAE),

with comments on the Systematics of the South American Freshwater Fish Genus *Basilichthys* GIRARD. *Miscellaneous Publications Museum of Zoology, University of Michigan*, **185**, 1–64.
Dyer BS (1998) Phylogenetic Systematics and Historical Biogeography of the Neotropical Silverside Family Atherinopsidae (Teleostei: Atheriniformes). In: *Phylogeny and Classification of Neotropical Fishes* (eds Malabarba LR, Reis RE, Vari RP, Lucena ZMS, Lucena CAS), pp. 519–536. Porto Alegre.
Dyer BS (2006) Systematic revision of the South American silversides (Teleostei, Atheriniformes). *Biocell*, **30**, 69–88.
Eaton DAR (2014) PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics*, **30**, 1844–1849.
Eaton DAR, Ree RH (2013) Inferring Phylogeny and Introgression using RADseq Data: An Example from Flowering Plants (Pedicularis: Orobanchaceae). *Systematic Biology*, **62**, 689–706.
Endler JA (1977) *Geographic variation, speciation, and clines.* Princeton University Press, Princeton, NJ, United States.
Eschmeyer WN, Fong JD (2017) Species by Family/Subfamily. *Catalog of Fishes*.
Fang B, Merilä J, Ribeiro F, Alexandre CM, Momigliano P (2018) Worldwide phylogeny of three-spined sticklebacks. *Molecular Phylogenetics and Evolution*, **127**, 613–625.
Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 397–423.
García G, Ríos N, Gutiérrez V *et al.* (2014) Promiscuous Speciation with Gene Flow in Silverside Fish Genus *Odontesthes* (Atheriniformes, Atherinopsidae) from South Western Atlantic Ocean Basins (V Ketmaier, Ed,). *PLoS ONE*, **9**, e104659.
Garzione CN, Hoke GD, Libarkin JC *et al.* (2008) Rise of the Andes. *Science*, **320**, 1304–1307.
González-Castro M, Rosso JJ, Delpiani SM, Mabragaña E, Díaz de Astarloa JM (2019) Inferring boundaries among fish species of the new world silversides (Atherinopsidae; genus *Odontesthes*): new evidences of incipient speciation between marine and brackish populations of *Odontesthes argentinensis*. *Genetica*, **147**, 217–229.
González-Castro M, Rosso JJ, Mabragaña E, Díaz de Astarloa JM (2016) Surfing among species, populations and morphotypes: Inferring boundaries between two species of new world silversides (Atherinopsidae). *Comptes Rendus Biologies*, **339**, 10–23.
Good JM, Vanderpool D, Keeble S, Bi K (2015) Negligible nuclear introgression despite complete mitochondrial capture between two species of chipmunks. *Evolution*, **69**, 1961–1972.
Griffiths D (2018) Why does freshwater fish species richness differ between Pacific and Atlantic drainages of the Americas ? *Journal of Biogeography*, 1–9.
Habit E, Gonzalez J, Ruzzante DE, Walde SJ (2012) Native and introduced fish species richness in Chilean Patagonian lakes: inferences on invasion mechanisms using salmonid-free lakes. *Diversity and Distributions*, **18**, 1153–1165.
Harrison RG (1990) Hybrid zones: windows on evolutionary process. *Oxford surveys in evolutionary biology*, **7**, 69–128.
Harrison S, Glasser NF (2011) The Pleistocene Glaciations of Chile. In: *Quarternary Glaciations - Extent and Chronology* (eds Ehlers J, Gibbard PL, Hughes PD), pp. 739–756. Elsevier B.V., Amsterdam.

669 Heras S, Roldán MI (2011) Phylogenetic inference in *Odontesthes* and *Atherina* (Teleostei:
670 Atheriniformes) with insights into ecological adaptation. *Comptes Rendus Biologies*, **334**,
671 273–281.
672 Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2018) UFBoot2: Improving the
673 Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, **35**, 518–522.
674 Hughes LC, Cardoso YP, Sommer JA *et al.* (2019) Data From: Biogeography, habitat transitions
675 and hybridization in a radiation of South American silverside fishes revealed by
676 mitochondrial and genomic RAD data. *Figshare*, DOI: 10.6084/m9.figshare.11413209.
677 Hughes LC, Somoza GM, Nguyen BN *et al.* (2017) Transcriptomic differentiation underlying
678 marine-to-freshwater transitions in the South American silversides *Odontesthes*
679 *argentinensis* and *O. bonariensis* (Atheriniformes). *Ecology and Evolution*, **7**, 5258–5268.
680 Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017) ModelFinder:
681 fast model selection for accurate phylogenetic estimates. *Nature Methods*, **14**, 587–589.
682 Katoh K, Standley DM (2013) MAFFT Multiple Sequence Alignment Software Version 7:
683 Improvements in Performance and Usability. *Molecular Biology and Evolution*, **30**, 772–
684 780.
685 Leigh JW, Bryant D (2015) POPART: Full-feature software for haplotype network construction.
686 *Methods in Ecology and Evolution*, **6**, 1110–1116.
687 Malinsky M (2019) DSuite - fast D-statistics and related admixture evidence from VCF files.
688 *bioRxiv*, 1–3.
689 Matzke NJ (2013a) BioGeoBEARS: BioGeography with Bayesian (and Likelihood)
690 Evolutionary Analysis in R Scripts. *University of California, Berkeley, Berkeley, CA.*
691 *http://CRAN. R-project. org/package= BioGeoBEARS*.
692 Matzke NJ (2013b) Probabilistic historical biogeography: new models for founder-event
693 speciation, imperfect detection, and fossils allow improved accuracy and model-testing.
694 *Frontiers of Biogeography*, **5**, 242–248.
695 Matzke NJ (2014) Model selection in historical biogeography reveals that founder-event
696 speciation is a crucial process in island clades. *Systematic Biology*, **63**, 951–970.
697 Moore WS, Buchanan DB (1985) Stability of the northern flicker hybrid zone in historical times:
698 implications for adaptive speciation theory. *Evolution*, **39**, 135–151.
699 Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ (2015) IQ-TREE: A fast and effective
700 stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology*
701 *and Evolution*, **32**, 268–274.
702 Ortubay SG, Gómez SE, Cussac VE (1997) Lethal temperatures of a Neotropical fish relic in
703 Patagonia , the scale-less characinid *Gymnocharacinus bergi*. *Environmental Biology of*
704 *Fishes*, **49**, 341–350.
705 Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an
706 inexpensive method for de novo SNP discovery and genotyping in model and non-model
707 species. *PLoS One*, **7**, e37135.
708 Raj A, Stephens M, Pritchard JK (2014) FastSTRUCTURE: Variational inference of population
709 structure in large SNP data sets. *Genetics*, **197**, 573–589.
710 Ronquist F (1997) Dispersal-Vicariance Analysis: A New Approach to the Quantification of
711 Historical Biogeography (D Cannatella, Ed,). *Systematic Biology*, **46**, 195–203.
712 Rueda EC, Mullaney KA, Conte-Grand C *et al.* (2017) Displacement of native Patagonian
713 freshwater silverside populations (*Odontesthes hatcheri*, Atherinopsidae) by introgressive
714 hybridization with introduced O. bonariensis. *Biological Invasions*, **19**, 971–988.

Ruzzante DE, Walde SJ, Cussac VE *et al.* (2006) Phylogeography of the Percichthyidae (Pisces) in Patagonia: roles of orogeny, glaciation, and volcanism. *Molecular Ecology*, **15**, 2949–2968.
Ruzzante DE, Walde SJ, Gosse JC *et al.* (2008) Climate control on ancestral population dynamics: insight from Patagonian fish phylogeography. *Molecular Ecology*, **17**, 2234–2244.
Ruzzante DE, Walde SJ, Macchi PJ, Alonso M, Barriga JP (2011) Phylogeography and phenotypic diversification in the Patagonian fish *Percichthys trucha*: the roles of Quaternary glacial cycles and natural selection. *Biological Journal of the Linnean Society*, **103**, 514–529.
Sanderson MJ (2002) Estimating Absolute Rates of Molecular Evolution and Divergence Times: A Penalized Likelihood Approach. *Molecular Biology and Evolution*, **19**, 101–109.
Silveira TLR, Martins GB, Domingues WB *et al.* (2018) Gene and Blood Analysis Reveal That Transfer from Brackish Water to Freshwater Is Less Stressful to the Silverside *Odontesthes humensis*. *Frontiers in Genetics*, **9**, 1–10.
Smith SA, O'Meara BC (2012) TreePL: Divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics*, **28**, 2689–2690.
Strüssmann CA, Conover DO, Somoza GM, Miranda LA (2010) Implications of climate change for the reproductive capacity and survival of New World silversides (family Atherinopsidae). *Journal of Fish Biology*, **77**, 1818–1834.
Strussmann CA, Moriyama S, Hanke EF, Cota JCC, Takashima F (1996) Evidence of thermolabile sex determination in pejerrey*. *Journal of Fish Biology*, **48**, 643–651.
Swofford DL (2017) PAUP*.
Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**, 3907–3930.
Tsuzuki MY, Aikawá H, Strüssmann CA, Takashima F (2000) Physiological responses to salinity increases in the freshwater silversides *Odontesthes bonariensis* and *O. hatcheri* (Pisces, Atherinidae). *Brazilian Journal of Oceanography*, **48**, 81–85.
Unmack PJ, Barriga JP, Battini MA, Habit EM, Johnson GD (2012) Phylogeography of the catfish *Hatcheria macraei* reveals a negligible role of drainage divides in structuring populations. *Molecular Ecology*, **21**, 942–959.
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. *Biological Journal of the Linnean Society*, **97**, 876–892.
Wagner CE, Keller I, Wittwer S *et al.* (2013) Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, **22**, 787–798.
Willis SC, Farias IP, Orti G (2014) Testing mitochondrial capture and deep coalescence in Amazonian cichlid fishes (Cichlidae: *Cichla*). *Evolution*, **68**, 256–268.
Wingert JM, Ferrer J, Malabarba LR (2017) Review of the *Odontesthes perugiae* species group from Río de La Plata drainage, with the description of a new species (Atherinomorpha: Atherinopsidae). *Zootaxa*, **4250**, 501.
Zemlak TS, Habit EM, Walde SJ *et al.* (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. *Molecular Ecology*, **17**, 5049–5061.

Zemlak TS, Habit EM, Walde SJ, Carrea C, Ruzzante DE (2010) Surviving historical Patagonian landscapes and climate: Molecular insights from *Galaxias maculatus*. *BMC Evolutionary Biology*, **10**, 1–18.
Zemlak TS, Walde SJ, Habit EM, Ruzzante DE (2011) Climate-induced changes to the ancestral population size of two Patagonian galaxiids: the influence of glacial cycling. *Molecular Ecology*, **20**, 5280–5294.

Figure Legends

Figure 1. A) Marine and freshwater areas currently occupied by species of *Odontesthes* in southern South America, and photograph of *O. perugiae* by Y.P. Cardoso. B) Sampling localities for ddRAD-sequenced samples (see Table S2 for description of population codes). C) 2006 Morphological phylogenetic hypothesis for all *Odontesthes* species (Dyer 2006), though lacking the newly described *O. yacuman* (Wingert *et al.* 2017). Marine lineages are shown with gray branches. Shapes with black outlines indicate that cyt*b* or ddRAD data were collected for this study for that species. D) Maximum likelihood molecular hypothesis based on seven nuclear genes and cyt*b* (Campanella *et al.* 2015), including three outgroups that inhabit the northern Pacific Ocean that are not otherwise included in this study. Branch lengths for molecular hypotheses are shown in substitutions per site, and gray branches indicate marine lineages. Species included in this study have black outlines around their shapes.

Figure 2. A) Maximum likelihood Cytochrome *b* phylogeny with ultra-fast bootstrap values written on major nodes inferred in IQ-Tree. Numbers in parentheses indicate the number of individuals sampled. None of the species in the *O. argentinensis*-species group are found to be monophyletic; this group is marked with an asterisk at the base. *Odontesthes brevianalis* and *O. mauleanum* are not reciprocally monophyletic, but break into two distinct clades. Notably, the *O.*

mauleanum populations LLA and PESC cluster with haplotypes of marine congeners. *Odontesthes gracilis* individuals are always interspersed among *O. regia* individuals. B) Maximum likelihood ddRADseq phylogeny based on 151 tips and 2,211,938 bp, with ultra-fast bootstrap values written on the nodes. A minimum of 75 individuals were present for each locus, and the full tree is shown with tip labels in Figure S4. Population codes can be found in Table S2, and on the map in Figure 1B. Branch lengths are shown in substitutions per site. Numbers in parentheses indicate the number of individuals sampled. The *O. argentinensis*-species group is again marked with an asterisk, but most species in this group now form distinct clades in contrast with mtDNA haplotypes. *Odontesthes brevianalis* and *O. mauleanum* do not cluster into distinct species, but individuals from the *O. mauleanum* Lake Llanquihue population (LLA) do not cluster with marine congeners as they do with mtDNA. *Odontesthes gracilis* is still interspersed among *O. regia* individuals.

Figure 3. Biogeographic reconstruction from BioGeoBEARS under the DIVA-like+J model, placing the origin of all modern *Odontesthes* in the Pacific Ocean, with three separate colonization events of the Atlantic Ocean, and three freshwater transitions: one to freshwater Pacific drainages, and two separate invasions of Atlantic freshwater drainages.

Figure 4. A) Distruct plots generated from fastSTRUCTURE results for the *O. argentinensis*-species group. B) Distruct plots generated from fastSTRUCTURE results for the clade of *Odontesthes* that includes *O. brevianalis, O. mauleanum, O. smitti, O. regia* and *O. gracilis*. C) ABBA-BABA tests for introgression for the *O. argentinensis*-species group and the clade containing three marine species (*O. smitti, O. regia* and *O. gracilis*) and two freshwater

811 populations of *O. mauleanum* (LLA and LCAL). LLA individuals possessed marine haplotypes,
812 which suggested introgression between marine individuals and this population. *Odontesthes*
813 *hatcheri* individuals were used as an outgroup in both cases.
814
815 **Funding**
816 This work was funded in part by the National Science Foundation Grant OISE-0530267 to G.O.
817 and NSF-DEB1457426 to G.O. and R.B.R. During her work on this project, L.C.H. was
818 supported by funds from the Harlan Endowment to the Department of Biological Sciences at
819 George Washington University.
820
821 **Acknowledgements**
822 The authors would like to thank Y. Ortiz (Sequencing and Genotyping Facility at University of
823 Puerto Rico-Rio Piedras) for assistance in preparing ddRAD libraries. We are grateful to R. P.
824 Vari for facilitating access to the Smithsonian laboratory facilities, and who is greatly missed
825 since his passing in 2016. We thank C. D. de Santana and L. Plaisance (Smithsonian Institution)
826 for their assistance with the Autogen. J. B. Stiegler (George Washington University) provided
827 assistance using Adobe Illustrator. We thank K. A. Crandall and R. A. Pyron (George
828 Washington University), S. E. Alter (City University of New York), and V. González
829 (Smithsonian Institution) for their helpful comments on earlier drafts of this manuscript.
830 Analyses were conducted on the Colonial One High Performance Computing cluster at George
831 Washington University unless otherwise stated, and we thank Adam Wong for help installing
832 software. We thank Devin Bloom (Western Michigan University) and one anonymous reviewer
833 for their comments that greatly improved this manuscript.

Author Contributions

L.C.H, G.O., R.B-R., E.M.H. and L.R.M. designed the study. Y.P.C., M.C., R.C., L.R.M., M.G-C., V.C., E.M.H, and G.M.S. provided tissues. L.C.H., J.A.S., and R.C. performed laboratory work. L.C.H., J.A.S. and R.C. analyzed the data. L.C.H., Y.P.C., R.B-R., and G.O. wrote the manuscript. All authors contributed to the final draft of the manuscript.

Data Accessibility

Cytochrome *b* sequences are accessioned on Genbank with numbers MK983245-MK983397. Raw reads for ddRAD sequences are archived under NCBI BioProject PRJNA546013. Data matrices, newick tree files, and R code are available for download on Figshare DOI: 10.6084/m9.figshare.11413209.