

Genomic patterns of diversity and divergence of two introduced salmonid species in Patagonia, South America

Shawn R. Narum¹  | Pablo Gallardo² | Cristian Correa^{3,4} | Amanda Matala¹ | Daniel Hasselman¹ | Ben J. G. Sutherland⁵ | Louis Bernatchez⁵

¹Hagerman Fish Culture Experiment Station, Columbia River Inter-Tribal Fish Commission, Hagerman, ID, USA

²Centro de Cultivos Marinos Bahía Laredo, University of Magallanes, Punta Arenas, Chile

³Facultad de Ciencias Forestales y Recursos Naturales, Instituto de Conservación Biodiversidad y Territorio, Universidad Austral de Chile, Valdivia, Chile

⁴Facultad de Ciencias, Instituto de Ciencias Marinas y Limnológicas, Universidad Austral de Chile, Valdivia, Chile

⁵Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, QC, Canada

Correspondence

Shawn Narum, Hagerman Fish Culture Experiment Station, Columbia River Inter-Tribal Fish Commission, Hagerman, ID, USA. Email: nars@critfc.org

Funding information

CONICYT-PAI, Grant/Award Number: 82130009; FONDECYT, Grant/Award Number: 11150990; CONICYT-FONDEF, Grant/Award Number: D10i1258; NSERC-Discovery to Mart Gross; Bonneville Power Administration, Grant/Award Number: 200900500

Abstract

Invasive species have become widespread in aquatic environments throughout the world, yet there are few studies that have examined genomic variation of multiple introduced species in newly colonized environments. In this study, we contrast genomic variation in two salmonid species (anadromous Chinook Salmon, *Oncorhynchus tshawytscha*, 11,579 SNPs and resident Brook Charr *Salvelinus fontinalis*, 13,522 SNPs) with differing invasion success after introduction to new environments in South America relative to populations from their native range in North America. Estimates of genetic diversity were not significantly different between introduced and source populations for either species, indicative of propagule pressure that has been shown to maintain diversity in founding populations relative to their native range. Introduced populations also demonstrated higher connectivity and gene flow than those in their native range. Evidence for candidate loci under divergent selection was observed, but was limited to specific introduced populations and was not widely evident. Patterns of genomic variation were consistent with general dispersal potential of each species and therefore also the notion that life history variation may contribute to both invasion success and subsequent genetic structure of these two salmonids in Patagonia.

KEYWORDS

adaptation, genomics, invasive species, RAD-seq, salmonids

1 | INTRODUCTION

In the last century, non-native aquatic plant and animal species have become widely distributed throughout the world, primarily through human-mediated activities (Cohen & Carlton, 1998; Roman & Darling, 2007). While many species have been dispersed accidentally (e.g., fish that escape from aquaculture or aquatic invertebrates transported in ship ballast tanks), some organisms are intentionally introduced for various objectives such as pest control, recreational fishing, or harvest opportunities. Introduced species that become established in

non-native locations can alter ecosystem dynamics in a manner that negatively impacts native species (Lodge, 1993), and many spread invasively to new locations. This may lead to local extirpation of native species and complications for water resource users, resulting in urgency to balance human activities with requirements for the maintenance of natural ecosystems (Pimentel, Zuniga, & Morrison, 2005).

Successful and rapid colonization of novel habitats by non-native species often occurs despite genetic limitations stemming from founder effects that are typically associated with extinction risk (e.g., low diversity and inbreeding depression; Frankham, 2005). This

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2017 The Authors. *Evolutionary Applications* published by John Wiley & Sons Ltd

“genetic paradox” has largely been resolved through meta-analyses that have identified key components that lead to successful establishment of newly founded populations. In particular, studies have shown that large propagule size and multiple introduction events (i.e., propagule pressure) allow species to maintain and sometimes surpass levels of neutral genetic diversity relative to native sources and expand their introduced range (Lockwood, Cassey, & Blackburn, 2005; Roman & Darling, 2007; Consuegra, Phillips, Gajardo, & de Leaniz, 2011; Kolbe, Leal, Schoener, Spiller, & Losos, 2012). Admixture of genetic lineages within species and introgressive hybridization among species can also increase genetic diversity and produce novel genotypes (e.g., Kelly, Muirhead, Heath, & MacIsaac, 2006). Maintenance of genetic variation is expected to allow for selection to occur in introduced species, and several studies have demonstrated that adaptive genetic change can occur rapidly (Huey, Gilchrist, Carlson, Berrigan, & Serra, 2005; Stockwell et al., 2003; Kinnison, Unwin, & Quinn, 2008). Thus, genetic variation can play a complex role in the establishment of introduced species in new environments (Dlugosch, Anderson, Braasch, Cang, & Gillette, 2015; Roman & Darling, 2007).

In addition to genetic variation, several factors influence the establishment success of organisms including environmental similarity between source and founding locations, life history variation, biotic interactions, and demographics (Arismendi et al., 2014; Bock et al., 2015; Facon et al., 2006). Species that are introduced to areas with similar climate and habitat as their source locations often have high establishment success (Hayes & Barry, 2008; Moyle & Marchetti, 2006), but some species become established in environments that differ from their native range through life history variation (Facon et al., 2006; Sax et al., 2007). Species that become established in dissimilar habitats typically display broad phenotypic plasticity and environmental tolerance (Arismendi et al., 2014) and may experience fewer competitive and biological pressures compared to their source niche (Keane & Crawley, 2002). Demographic factors such as dispersal (natural or human-mediated) contribute to initial colonization of non-native species, while connectivity of established populations may facilitate maintenance and expansion of invasive species (Facon et al., 2006).

Despite the many factors that must be considered for studies of the invasiveness of non-native species, new molecular tools offer the potential to address both demographic and evolutionary processes by surveying neutral and adaptive genetic variation throughout the genome (e.g., Davey et al., 2011; Narum, Buerkle, Davey, Miller, & Hohenlohe, 2013). Neutral loci are effective for evaluating demographic factors such as genetic diversity and gene flow, while adaptive variation can reveal signals of selection when introduced collections are compared to those in a species' native range (e.g., Hamilton, Okada, Korves, & Schmitt, 2015). Further, species that are introduced to novel geographic regions without native congeners provide effective study systems to investigate invasion genetics as there is no confounding genetic background from extant populations or closely related taxa.

This study examines two salmonid species with differing life histories and invasion success after introduction to new environments in Patagonia, South America. In South America, salmonids (Salmonidae) were not historically native but have been widely introduced to rivers

and lakes throughout Patagonia with similar environments as their native ranges (MacCrimmon, 1971; Pascual et al., 2007; Basulto, 2003; Gallardo et al., 2007; Correa & Gross, 2008; Consuegra et al., 2011; Arismendi et al., 2014; Monzón-Argüello, de Leaniz, Gajardo, & Consuegra, 2014). Some species have been more successful than others at establishing viable populations in non-native aquatic systems of South America, with 5 of 12 introduced salmonid species considered established (Arismendi et al., 2014; Monzón-Argüello et al., 2014). Species such as Chinook Salmon (*Oncorhynchus tshawytscha*) are widespread and considered invasive, while others such as Brook Charr (*Salvelinus fontinalis*) are only locally established despite multiple introduction attempts for both species that were initiated in the early 1900's, but more intensively in recent decades (Arismendi et al., 2014). In particular, Chinook Salmon aquaculture programs were initiated in the 1980's, and escapees from ocean net pens spread widely to rivers throughout Patagonia (Arismendi et al., 2014; Correa & Gross, 2008; Di Prinzio, Riva Rossi, Ciancio, Garza, & Casaux, 2015). Brook charr were heavily introduced for recreational fishing opportunities in various locations in Patagonia in the last 30 years, but there is little evidence that this species has become widespread (Arismendi et al., 2014; Gallardo et al., 2007). Here, we investigate Chinook Salmon and Brook Charr populations in South America, and contrast life history and genomic variation of introduced populations with stocks from their native range. Specific objectives included testing for the following: (i) patterns of genetic differentiation and connectivity within introduced populations relative to those from their native range, (ii) evidence of reduced genetic diversity in introduced populations due to founder effects, and (iii) evidence for divergent selection between native and introduced populations in differing environments.

2 | METHODS

2.1 | Samples of introduced and native populations

A major premise of our study design was to compare genomic variation of introduced populations in Patagonia, South America, to collections from their North America native ranges that represented putative source stocks as well as those under natural population dynamics. Samples from introduced populations in South America included four collections of Chinook Salmon and five collections of Brook Charr (Figure 1). As native ranges for each species are extensive spanning thousands of kilometers in North America, it was not feasible to compare native populations from across entire geographic distributions. Thus, we utilized information from previous studies to identify appropriate reference collections from native ranges to address study objectives. Specific stocks used for historical introductions to South America have not been thoroughly documented, but putative source stocks for each species have been identified in previous studies (Correa & Gross, 2008; Di Prinzio et al., 2015; Neville & Bernatchez, 2013) as described below.

Native Chinook Salmon collections were targeted based on information from a previous study (Correa & Moran, 2016) that identified the Pacific Northwest region of North America as the source of

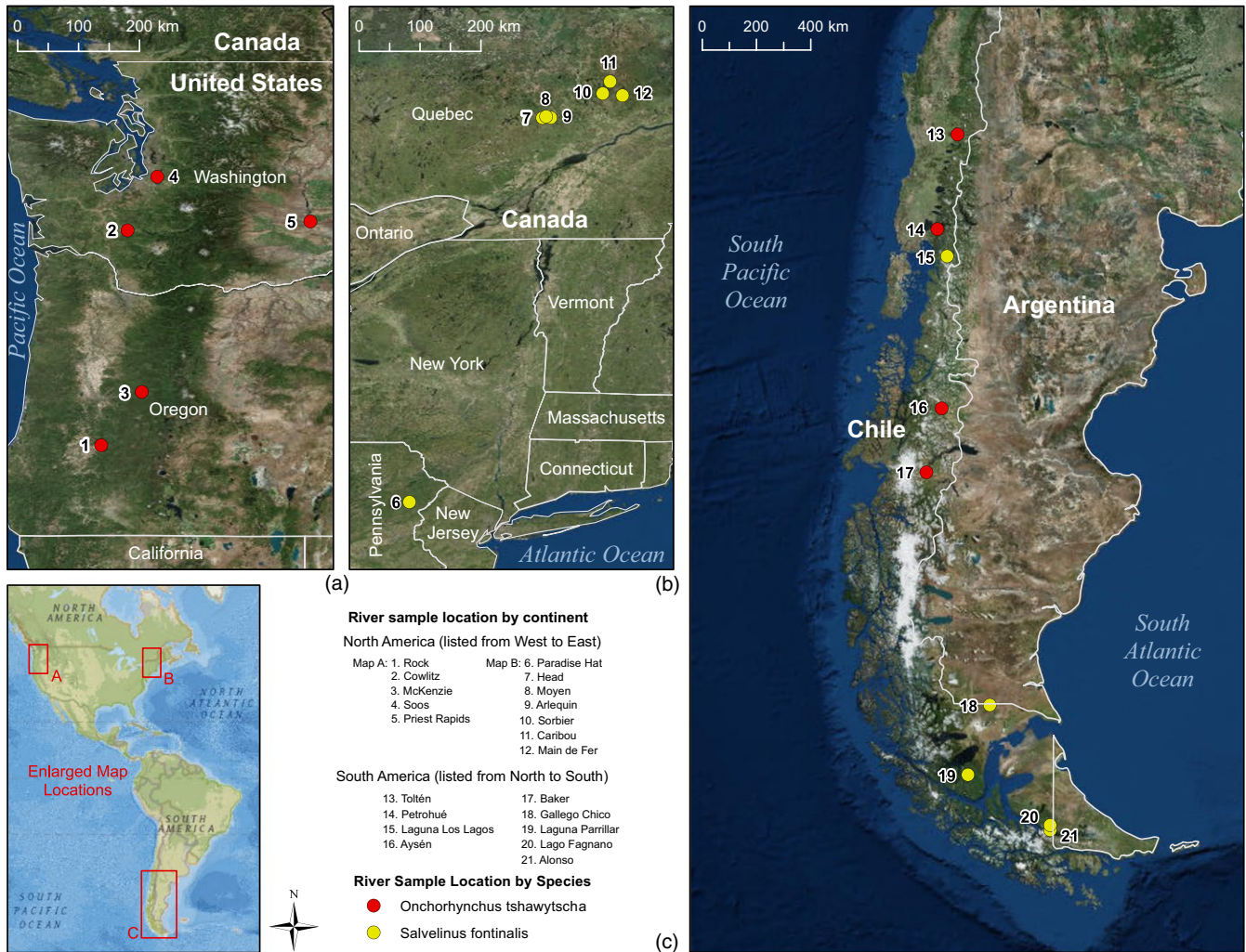


FIGURE 1 Map of sample collections for (a) Chinook Salmon in their native range from North America, (b) Brook Charr in their native range from North America, and (c) introduced populations of both species in South America

the specific introduced populations from Patagonia, South America, examined herein. Probable sources identified with microsatellite markers (Correa & Moran, 2016) included lower Columbia R. spring and fall runs, South Puget Sound fall run/Whidbey Basin, North Oregon Coast, and interior Columbia R. (summer/fall run), which are all stocks from Oregon and Washington, USA. For this study, the five most representative native collections available were (i)—Cowlitz River (akin to lower Columbia R. spring run, Washington); (ii)—Soos Hatchery (akin to Puget Sound fall run, Washington); (iii)—Rock Creek (akin to north Oregon Coast); (iv)—McKenzie River (akin to north Oregon Coast); and (v)—Priest Rapids Hatchery (akin to interior Columbia R. fall/summer-run, Washington; Figure 1a). These five native collections were included to test for contrasting patterns with the four introduced populations from South America for a total of nine Chinook Salmon collections ($n = 364$; Table 1).

For Brook Charr, collections from the native range were targeted to include a hatchery stock (Paradise Hatchery, Pennsylvania, USA; Figure 1b) that has been commonly used for broad introductions over the past century (Neville & Bernatchez, 2013) and was

expected to represent a likely source stock for introductions in Patagonia. However, this was not expected to account for all putative source stocks which were not feasible to include in this study. Additionally, six wild populations from the native range in Quebec, Canada, that have not experienced introgression with stocked fish from other regions (Marie, Bernatchez, & Garant, 2010) were included to represent genetic variation that would be expected in natural populations. Of the six wild populations, three were from the Portneuf system and three were from the Mastigouche system in Quebec, Canada. These seven native collections were included (Figure 1b) from North America along with the five introduced collections from South America for a total of 12 Brook Charr collections ($n = 414$; Table 1).

To contrast general environments of native versus introduced populations for each species, indicators of precipitation, temperature, and elevation were obtained for each sampling location from a database of global climate layers with high-spatial resolution (WorldClim, www.worldclim.org; Hijmans, Cameron, Parra, Jones, & Jarvis, 2005). A total of 20 variables were extracted using ArcGIS software (ESRI; Redlands,

TABLE 1 Genetic diversity and effective population sizes of collections of Chinook Salmon (*O. tshawytscha* = *Ots*; 11,759 SNP markers) and Brook Charr (*S. fontinalis* = *Sfo*; 13,522 SNP markers) from introduced (South America = S.A.) and native (North America = N.A.) locations. H_E = expected heterozygosity; AR = allelic richness; PAR = private allelic richness; N_e = effective population size with 95% confidence intervals (95% CI)

Species	Collection Description & Abbreviation	Sample size (n)	H_E	AR	PAR	Nb ^a (95% CI)
<i>Ots</i>	N.A.—Cowlitz River (COW)	45	0.337	1.21	0.024	927 (669–1498)
<i>Ots</i>	N.A.—McKenzie River (MCK)	48	0.337	1.20	0.026	442 (368–553)
<i>Ots</i>	N.A.—Priest Rapids Hatchery (PRH)	46	0.274	1.20	0.025	2471 (1234–315870)
<i>Ots</i>	N.A.—Rock Creek (ROC)	48	0.292	1.22	0.028	1243 (873–2144)
<i>Ots</i>	N.A.—Soos Creek (SOO)	48	0.265	1.21	0.032	1989 (1178–6265)
<i>Average across N. American populations (Ots)</i>	<i>47</i>	<i>0.301</i>	<i>1.21</i>	<i>0.027</i>	<i>1414</i>	
<i>Ots</i>	S.A.—Aysén Rio (AYS)	30	0.299	1.21	0.025	202 (179–233)
<i>Ots</i>	S.A.—Baker Rio (BAK)	34	0.353	1.21	0.023	356 (292–457)
<i>Ots</i>	S.A.—Petrohué Rio (PET)	36	0.305	1.22	0.027	67 (64–70)
<i>Ots</i>	S.A.—Toltén Rio (TOL)	19	0.283	1.20	0.031	149 (125–183)
<i>Average across S. American populations (Ots)</i>	<i>29.8</i>	<i>0.310</i>	<i>1.21</i>	<i>0.026</i>	<i>194</i>	
<i>Sfo</i>	N.A.—Arlequin Lac (ARL)	38	0.247	1.12	0.017	48 (45–51)
<i>Sfo</i>	N.A.—Caribou Lac (CAR)	30	0.316	1.17	0.01	16 (15–16)
<i>Sfo</i>	N.A.—Paradise Hatchery (PAR)	19	0.367	1.20	0.015	414 (250–1174)
<i>Sfo</i>	N.A.—Head Lac (HEA)	30	0.303	1.19	0.014	136 (123–152)
<i>Sfo</i>	N.A.—Main de Fer Lac (MDF)	29	0.224	1.12	0.012	461 (275–1366)
<i>Sfo</i>	N.A.—Moyen Lac (MOY)	29	0.226	1.08	0.013	133 (99–199)
<i>Sfo</i>	N.A.—Sorbier Lac (SOR)	27	0.333	1.18	0.009	199 (160–262)
<i>Average across N. American populations (Sfo)</i>	<i>28.9</i>	<i>0.288</i>	<i>1.15</i>	<i>0.013</i>	<i>201</i>	
<i>Sfo</i>	S.A. - Alonso Rio (ALO)	39	0.299	1.23	0.019	101 (96–106)
<i>Sfo</i>	S.A. - Gallego Chico Rio (CHI)	47	0.258	1.17	0.014	58 (56–60)
<i>Sfo</i>	S.A. - Lago Fagnano (FGN)	15	0.305	1.23	0.018	601 (327–3441)
<i>Sfo</i>	S.A. - Los Lagos Laguna (LAG)	41	0.289	1.17	0.029	233 (203–273)
<i>Sfo</i>	S.A. - Parrillar Laguna (PAR)	41	0.222	1.17	0.014	112 (105–121)
<i>Average across S. American populations (Sfo)</i>	<i>36.6</i>	<i>0.275</i>	<i>1.19</i>	<i>0.019</i>	<i>221</i>	

^aNb estimates adjusted for downward bias following Waples et al. (2016). Average values are shown in italics.

CA, USA) including elevation and several bioclimatic indicators of temperature and precipitation (“bio01”–“bio19”; Table S1). These data were obtained from the “Current Conditions” dataset, which includes interpretations of observed data, representative of the years 1950–2000, at a resolution of 30 arc-seconds (~1 km). Environmental relationships among collections were evaluated with principal component analysis (PCA) implemented in the R statistical package version 3.3.1 (R Development Core Team 2013). To aid interpretation of resulting canonical variables, vectors representing the original variables were passively correlated with the canonical axes and plotted onto the ordination using the “envfit” function of the R package “vegan” version 2.0-10 (Oksanen et al., 2013).

2.2 | Molecular techniques for RAD sequencing

DNA was extracted from fin tissue using Qiagen DNeasy kits (Qiagen) following the manufacturers’ recommended protocols. Extracted genomic DNA was quantified using Quantit PicoGreen dsDNA Assay Kits (Invitrogen) and a Victor2 microplate fluorometer (Perkin Elmer).

Restriction site-associated DNA (RAD) libraries were prepared for Illumina HiSeq 1500 sequencing using a protocol similar to those previously published (Baird et al., 2008), but modified as described in Hecht, Matala, Hess, and Narum (2015). Libraries were prepared with a starting DNA concentration between 250 and 500 ng per sample, with samples of similar DNA quantity included in each library

to provide equal representation of sequencing reads. Samples were digested individually with the restriction enzyme *Sbf*I-HF (NEB, Ipswich, Massachusetts, USA) and individually barcoded using a six nucleotide barcode adapter sequence. Digested and barcoded samples of the same starting concentration were pooled into libraries that averaged 48 individuals, where no two samples within a library were assigned the same barcode sequence, and each barcode sequence within a library differed by at least two bases from another barcode sequence. Libraries were then mechanically sheared using a Bioruptor 300 sonicator (Diagenode) to generate DNA fragment lengths between 200 and 700 bp, and fragments were isolated using an Agencourt AMPure XP bead purification system (Beckman Coulter). Prior to sequencing, RAD libraries were quantified using real-time PCR on an ABI 7900HT Sequence Detection System (Life Technologies). Each library was sequenced in one lane on a HiSeq 1500 sequencer at a single read length of 100 bp, across a total of 16 lanes.

2.3 | Bioinformatics pipeline and filters

SNP discovery was completed with the *de novo* pipeline in STACKS (Catchen, Amores, Hohenlohe, Cresko, & Postlethwait, 2011; Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013) with sequence data for each species. For Chinook Salmon, this included nine populations (five from North America and four from South America). For North America collections of Chinook Salmon, RAD data were available from a previous study (Hecht et al., 2015) with libraries prepared in an identical manner as those from South America and included 10 double haploid fish to test for paralogous sequence variants (PSVs). For Brook Charr, *de novo* SNP discovery included 12 populations (seven from North America and five from South America).

For the STACKS pipeline of each species, raw Illumina reads were first checked for quality using the program FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). The 25 bases on the 3' end of sequence reads had reduced quality scores relative to the 5' 75 base positions across our sequence data, so reads were trimmed to 75 bp. Reads were quality filtered and de-multiplexed using the "process_radtags" program of STACKS, including options for cleaning the data by discarding any read with an uncalled base (-c), discarding reads with low-quality scores (-q), and rescuing barcodes and partial restriction enzyme recognition sites (-r). All other parameters and options were executed with the default values as outlined in the manual for the program (<http://creskolab.uoregon.edu/stacks>).

After individual sample reads were quality filtered, trimmed, and de-multiplexed, sequences for each sample were taken through the "ustacks" module of STACKS to identify loci. In "ustacks," the deleveraging (-d) and removal (-r) algorithms were applied to filter out those sequences that were likely to be paralogous and highly repetitive. Parameters for STACKS were minimum depth of coverage at a stack ("m" = 2–5), maximum distance between stacks ("M" = 2), distance between secondary reads, and primary stacks ("N" = 4). SNP discovery was carried out using the default SNP model with a chi-square

significance level of 0.05. We created a *de novo* catalog of RAD tag loci using the "cstacks" module by selecting two individuals from each population with at least 2.5 million reads (but no >4 million reads) to represent genetic variation in each species. Individual samples were then aligned to the catalog using the module "sstacks," and genotypes were exported using the "populations" module.

Genotypes were filtered with multiple steps as shown in Table S2 including exclusion of: (i) any RAD tag locus with more than four SNP sites to remove putative PSVs, hypervariable, or poorly sequenced tags, (ii) any RAD tag locus where one of the ten double haploid samples was observed to be heterozygous at any of the SNP positions to remove putative PSVs, (iii) any SNP marker with more than two alleles to remove SNPs with sequencing errors, putative PSVs, or loci that do not fit a bi-allelic statistical model, (iv) any SNP marker missing more than 30% of the genotypes across all of the populations to limit the amount of missing data, (v) any SNP marker failing tests of Hardy-Weinberg Equilibrium in more than one-third of the populations to exclude technical artifacts such as null alleles (heterozygote deficit loci) and putative PSVs (heterozygote excess loci; GENEPOP v.4.0.6, Rousset 2008; false discovery rate corrected critical value BY-FDR, Benjamini and Yekutieli 2001; Narum, 2006), and (vi) any SNP marker with an average minor allele frequency (MAF) across the populations falling below 0.02 to exclude spurious rare SNPs or sequencing errors. Filters for loci that deviated from HWE were implemented to remove false positives due to PSVs or null alleles at the expense of possibly excluding some markers under selection (false negatives). As physically linked SNPs would bias population genetic estimates such as F_{ST} , we only retained the first SNP marker per RAD tag. Individual samples were also filtered from the dataset if they were missing more than 30% of genotypes across all filtered loci.

Sequences from RAD tags for Chinook Salmon were aligned to those from a high-density RAD-based linkage map in this species (Briec, Waters, Seeb, & Naish, 2014) in an effort to determine the relative genetic position and linkage group assignment of loci. Alignment to the RAD database of Briec et al. (2014) was conducted using the end-to-end mode in the short sequence alignment software program *Bowtie2* v.2.2.3 (Langmead, Trapnell, Pop, & Salzberg, 2009) and with a threshold of mapping quality (MAPQ) score ≥ 2 . Alignment of Brook Charr RAD tag sequences against a Brook Charr linkage map (Sutherland et al., 2016) resulted in low numbers of homologous markers, potentially due to the different library preparation method used for the map generation from that used in this study. Therefore, the MapComp approach (described in Sutherland et al., 2016) was applied to pair homologous as well as proximal anonymous markers with mapped markers using a reference genome as an intermediate (here the most recent version of the Atlantic Salmon reference genome, ICSASG_v2; GenBank AGKD00000000.4; Lien et al. 2016). The MapComp algorithm was applied with default settings, except that it was adapted to permit multiple anonymous markers pairing with a single mapped marker to obtain approximate genomic locations. Proximal markers were only considered when an estimated distance between markers on the reference genome was <1 Mb. All nonpositioned markers were considered as "unknown."

2.4 | Statistical analyses

Genetic diversity was estimated for all populations of each species with measures of expected heterozygosity (H_e), allelic richness (AR), and private allelic richness (PAR) in HP-RARE (Kalinowski, 2005). Differences in genetic diversity among collections were tested with analysis of variance (ANOVA). Genetic relationships among populations within each species were estimated by pairwise F_{ST} (Weir & Cockerham, 1984).

To determine clusters of genetically similar collections, we used discriminant analysis of principal components (DAPC) from the “adegenet” package (v1.4-2/ade4 v1.7-2 Jombart, 2008; Jombart & Ahmed, 2011) in the R statistical computing environment (R Development Core Team 2013). The number of clusters was determined by running 10 iterations of the “find.clusters” module of “adegenet” for 1 to 12 possible cluster values (K). The Bayesian information criterion (BIC) values, which provide a metric of model selection similar to likelihood, were averaged across the 10 iterations and standard deviation was estimated for each value of K . The rate of change in BIC value was used to determine the most appropriate value of K (Evanno, Regnaut, & Goudet, 2005).

Estimates of contemporary effective population size (N_b ; following Luikart, Ryman, Tallmon, Schwartz, & Allendorf, 2010) for collections of each species were generated using the linkage disequilibrium (LD) method implemented in the program N_e ESTIMATOR (Do et al., 2014). The number of SNPs available for both species was computationally prohibitive for these analyses, so 10% of the markers were randomly chosen for Chinook Salmon and Brook Charr after removing any putative outlier loci (i.e., 1164 and 1339 SNPs, respectively). Since loci that are physically linked can bias estimates of effective size (Waples, 2006), we first used N_e ESTIMATOR to detect pairs of loci that exhibited strong evidence of physical linkage (mean $R^2 \geq 0.5$) for more than a third of populations. Based on these criteria, we excluded one locus per pair that exhibited signs of physical linkage which resulted in the removal of eight loci for Chinook Salmon and 25 loci for Brook Charr. This revised dataset was used to generate estimates of N_b using the random mating model and exclusion of rare alleles (critical value $P_{crit} = 0.05$) following procedures described in previous studies (Candy, Campbell, Grinnell, Beacham, & Narum, 2015; Gruenthal et al., 2014). Additional estimates of N_b were generated with standardized sample sizes for Chinook Salmon and Brook Charr (i.e., $n = 30$ and 27 , respectively) by randomly selecting individuals for analyses. These sample size criteria resulted in the exclusion of one Chinook Salmon population (i.e., Toltén) and two Brook Charr populations (i.e., Lago Fagnano and Paradise Hatchery) from analyses to estimate N_b . Finally, we incorporated a correction for downward bias in estimates of “naïve” N_b (Waples, Larson, & Waples, 2016) using haploid chromosome number for each species (Chinook Salmon = 34; Brook Charr = 42).

Outlier tests were performed with PCAdapt (Duforet-Frebourg, Bazin, & Blum, 2014), a Bayesian program based on a hierarchical factor model that jointly determines neutral structure and outlier loci using K latent factors. To identify candidate outlier loci, the model

searches for markers that deviate from neutral expectations as measured by latent factors. This approach provides the ability to parse outlier loci into latent factors for specific population clusters rather than over all populations. This was a particularly important feature for our study as one of our objectives was to distinguish signals of selection in introduced populations rather than native populations. This method has been shown to have sufficient power and low false-positive rates in independent evaluations of various outlier methods (Lotterhos & Whitlock, 2015). We ran PCAdapt with a burn-in of 200 steps and 400 steps in the MCMC, with $K = 4$ for Chinook Salmon and $K = 12$ for Brook Charr. Values of K were determined from the mean squared error rate for each species following criteria from Evanno et al. (2005) and a range of K between 1 and 12. The SNP data were also scaled so that all markers had a variance equal to one in order to reduce false positives as previously recommended (Duforet-Frebourg et al., 2014). For all loci considered to be candidates from PCAdapt, we used existing paired end data for each species to assemble longer sequence contigs from overlapping reads with the program Sequencher (v.5.3) to annotate potential genes. We attempted BLAST searches (match threshold of e^{-10} ; Blast2GO 3.1.3, Conesa et al., 2005) for outlier loci of each species using the contig sequences. For further annotation of candidates, we aligned contigs to the Atlantic Salmon genome (NCBI ICSAGS_v2) with Bowtie2 and searched for coding sequences within 5 kb of either direction of the genome (10-kb window).

3 | RESULTS

3.1 | Environmental differentiation between native and introduced locations

The PCA of environmental features for both species indicated that the introduced locations in South America were distinct from those in North America and typically most similar within each continent (i.e., the first two components explained $\geq 80\%$ of the variation for each species; Figure 2a,b). For Chinook Salmon, 13 of the 20 variables were statistically different among locations ($p \leq .01$; Table S3a) and environmental differences between N. and South American locations were primarily driven by higher precipitation and cooler temperatures in the South American collections. Also for Chinook Salmon, there was distinction on PC2 among environments within South America between northern collections from Toltén and Petrohué relative to those from the south (i.e., Aysén and Baker) primarily due to lower precipitation and colder temperatures in the two southern locations. In North America, the Chinook Salmon collection from Priest Rapids in the interior was distinct on PC1 from all the other collections on the coast due to warm and dry conditions and more variable seasonal temperatures. For Brook Charr, 17 of the 20 variables were significantly different ($p \leq .01$; Table S3b1) and collections were closely clustered by continent due to colder winter seasons with more precipitation in North America, with the exception of Los Lagos which was highly distinct in a more northern location with higher precipitation than all the other collections.

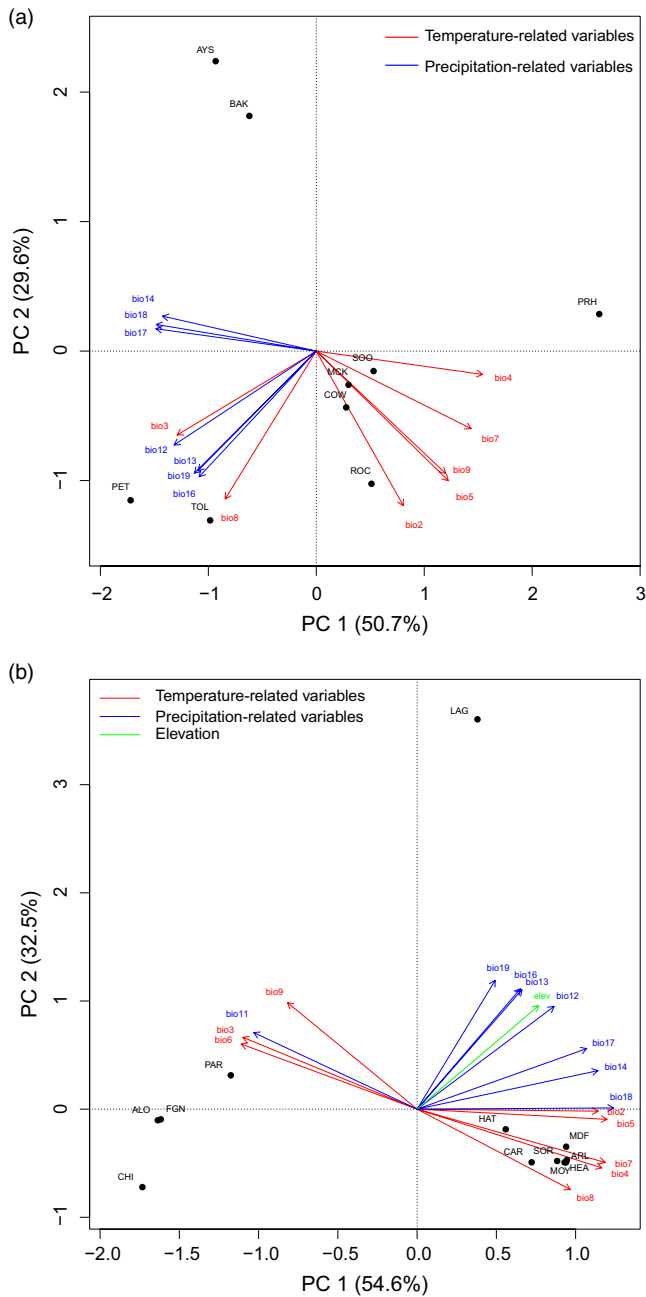


FIGURE 2 Principal component analysis (PCA) of environmental data for (a) Chinook Salmon, and (b) Brook Charr. Vector lines are only shown for environmental variables that accounted for a significant portion of the variation ($p \leq .01$). Individual environmental variables are described in detail in Table S4, but “bio” codes of variables are as follows (WorldClim database): bio1 = Annual mean temp., bio2 = Mean diurnal range, bio3 = Isothermality, bio4 = Temperature seasonality, bio5 = Max temp. of warmest month, bio6 = Min temp. of coldest month, bio7 = Temp. annual range, bio8 = Mean temp. of wettest quarter, bio9 = Mean temp. of driest quarter, bio10 = Mean temp. of warmest quarter, bio11 = Mean temp. of coldest quarter, bio12 = Annual precipitation, bio13 = Precipitation of wettest month, bio14 = Precipitation of driest month, bio15 = Precipitation seasonality, bio16 = Precipitation of wettest quarter, bio17 = Precipitation of driest quarter, bio18 = Precipitation of warmest quarter, elev = elevation. Population abbreviations follow those in Table 1

3.2 | SNP Discovery, filters, and populations

3.2.1 | Chinook Salmon

De novo SNP discovery was completed for Chinook Salmon in STACKS with sequence data from nine populations (five from North America and four from South America). A total of 11,759 SNPs passed filtering criteria and were used for all further analyses. Table S4 includes the RAD tag sequence for all loci included for Chinook Salmon in this study. Of the 11,759 SNPs, 6,020 were aligned to an existing Chinook Salmon linkage map (Briec et al., 2014). Of the initial 364 samples from nine collections, 10 individuals were removed that did not satisfy criteria for missing genotype data (>30%). Mean sample size per population after filtering was 39.3 with a range of 19–48 per location (Table 1). Allele frequencies, individual genotypes, and HWE results are available in Table S5.

3.2.2 | Brook Charr

De novo SNP discovery was completed for Brook Charr with 12 collections representing seven native (North America) and five introduced (South America) populations. Our STACKS pipeline identified 13,522 SNPs that passed filtering criteria, of which 4,911 were assigned positions based on MapComp alignment with the Brook Charr linkage map (Sutherland et al., 2016). Table S6 includes the RAD tag sequence for all loci included for Brook Charr in this study. Of the initial 414 samples, 29 were removed that did not satisfy criteria for missing genotype data (>30%). Mean sample size per population after filtering was 32.1 with a range of 15–47 per location (Table 1). Allele frequencies, individual genotypes, and HWE results are provided in Table S7.

3.3 | Genetic diversity and differentiation

3.3.1 | Chinook Salmon

No significant differences in genetic diversity were observed between native North America and introduced South America populations of Chinook Salmon (mean H_e , South America = 0.310, North America = 0.301, $p = .69$; mean AR, South America = 1.21, North America = 1.21, $p = .70$; mean PAR, South America = 0.027, North America = 0.027, $p = .70$; Table 1). Despite similar levels of diversity, estimates of N_b revealed that introduced collections of Chinook Salmon had significantly lower N_b than native collections (mean N_b , South America = 194, North America = 1414, $p = .022$; Table 1). These estimates of N_b were similar to those with standardized sample sizes ($n = 30$) among collections (mean N_b , South America = 196, North America = 1257, $p = .09$; Table S1).

All populations of Chinook Salmon were significantly different from one another based on allele frequencies. Values of F_{ST} over all loci between North America and South America populations (mean = 0.074; range = 0.023–0.140) were similar to values among populations from the five native groups (mean = 0.088; range = 0.050–0.134), but F_{ST} values were generally lower among introduced populations

(mean = 0.069; range = 0.011–0.106; Table S8). Relationships with DAPC identified four distinct clusters from the nine collections ($K = 4$) and explained much of the variation (DF1 = 47.2%, DF2 = 28.0%; Figure 3). The clusters from DAPC provided evidence for genetic similarity of Soos (North America) with Toltén, and Petrohué (South America) as one cluster, and Cowlitz (North America) with Aysén, and Baker (South America) as another cluster (Figure 3). Rock Cr. and Priest Rapids formed another cluster, while McKenzie R. was a single distinct cluster (Figure 3). Probability of individual membership strongly supported the four clusters as all but four individuals were assigned to their expected cluster with high probability (Fig. S1).

3.3.2 | Brook Charr

Estimates of genetic diversity were similar among populations of Brook Charr from South America and North America collections (mean H_E , South America = 0.275, North America = 0.288; mean AR, South America = 1.19, North America = 1.15; mean PAR, South America = 0.019, North America = 0.013; Table 1). While South America populations did not have significantly different H_E ($p = .65$) or AR ($p = .10$) than those from North America, PAR ($p = .04$) was statistically significant. The difference in private alleles for Brook Charr was largely driven by higher PAR (0.029) in the Los Lagos collection relative to all other collections (South America range from 0.014 to 0.019; North America range from 0.010 to 0.017; Table 1). When Los Lagos was excluded, there were no significant differences in PAR between native North America and introduced South America populations ($p = .07$). Estimates of N_b revealed that both native and introduced collections of Brook Charr had low N_b that were not significantly different with all individuals included (mean N_b , South America = 221, North America = 201, $p = .659$; Table 1) or with standardized sample sizes ($n = 27$) among collections (mean N_b , South America = 129, North America = 162, $p = .659$; Table S1).

All populations of Brook Charr were significantly different from one another based on allele frequencies with the exception of Lago Fagnano and Alonso River that were from the same drainage in South America (Alonso River is a tributary to Lago Fagnano). High values of F_{ST} over all loci indicated strong differentiation among Brook Charr collections. Values of F_{ST} between North America and South America populations (mean = 0.263; range = 0.070–0.429) were similar to values among populations from North America locations (mean = 0.259; range = 0.113–0.440), but F_{ST} values were generally much lower among introduced populations in South America (mean = 0.133; range = 0.013–0.243; Table S9). Mean F_{ST} among South America collections remained low (0.146) even when the nonsignificant comparison between Lago Fagnano and Alonso River was removed. The signal of strong genetic differentiation among Brook Charr collections was also apparent from DAPC results (DF1 = 29.2%, DF2 = 17.1%). Nearly all collections were distinct ($K = 12$), with the exception of Lago Fagnano and Alonso River from the same system that formed a single cluster that was most similar to the putative source stock from Paradise Hatchery (Figure 4). The other three introduced collections from South America were highly distinct from any North America

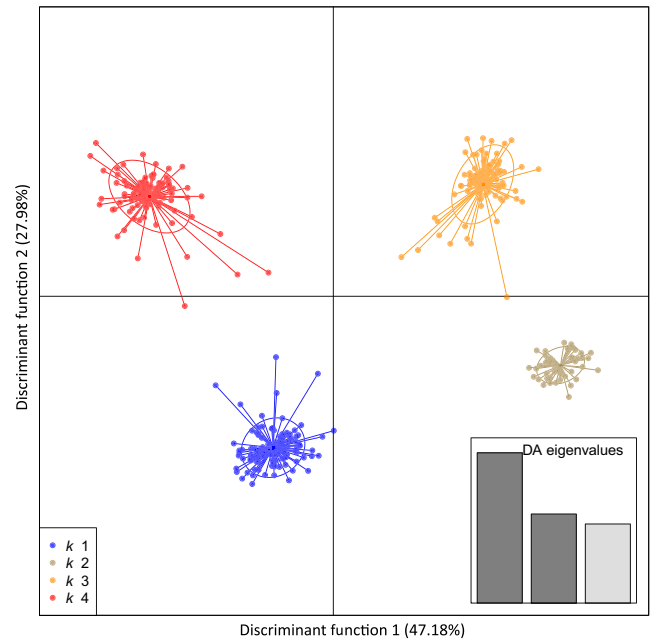


FIGURE 3 Discriminant analysis of principal components (DAPC) of Chinook Salmon (function 1 vs. 2) from 11,759 SNP markers. Cluster membership: k1 = Aysén, Baker, Cowlitz, k2 = McKenzie, k3 = Rock, Priest Rapids, k4 = Toltén, Petrohué, Soos

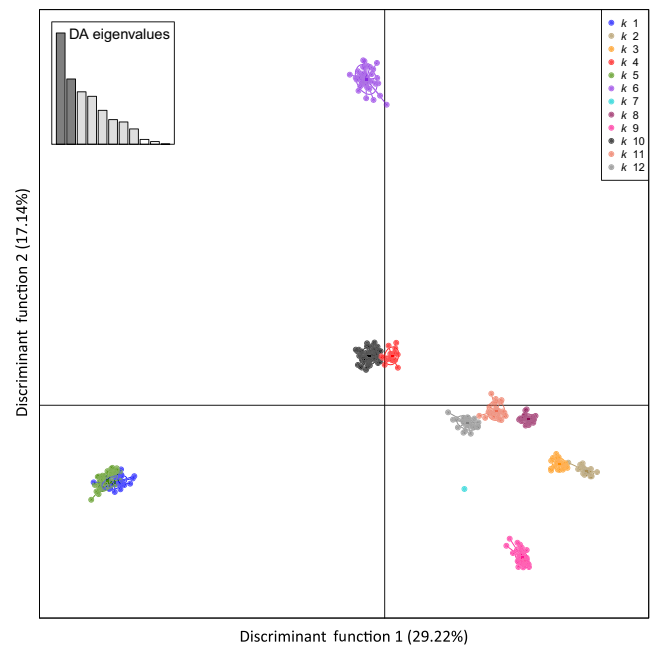


FIGURE 4 Discriminant analysis of principal components (DAPC) of Brook Charr (function 1 vs. 2) from 13,522 SNP markers. Primary cluster membership: k1 = Chico (South America), k2 = Main de Fer (North America), k3 = Sorbier (North America), k4 = Paradise (North America), k5 = Parrillar (South America), k6 = Los Lagos (South America), k7 = Arlequin (North America), k8 = Moyen (North America), k9 = Arlequin (North America), k10 = Alonso and Fagnano (South America), k11 = Head (North America), k12 = Caribou (North America)

collections included in this study. Probability of individual membership strongly supported the 12 clusters with only two individuals mis-assigned at probability >0.90; however, assignment probability was generally lower for individuals between Parrillar Laguna and Gallego Chico clusters relative to other groups (Fig. S2).

3.4 | Outlier tests for divergent selection

3.4.1 | Chinook Salmon

Tests for outlier loci were completed with PCAdapt to jointly account for underlying neutral structure and divergent selection. For Chinook Salmon, PCAdapt identified four significant latent factors ($k = 4$) among the nine collections. Differentiation among populations was evident for specific latent factors (Factor 1 = Toltén/Petrohué/Soos from others, Factor 2 = Soos/McKenzie versus Priest, Factor 3 = Toltén/Rock from all others, Factor 4 = introduced+Cowlitz from others; Figure 5). Latent factors explained variation that allowed examination of outlier loci that were associated with specific population groups (Factor 1 versus Factor 3, Fig. S3a; Factor 2 versus Factor 4, Fig. S3b). Of the top 1% outlier loci (Figure 5; Table S10), the majority were evident for Factor 3 (66.9%) and Factor 2 (23.7%), but a small proportion were also included for Factor 4 (7.6%) and Factor 1 (1.7%). These results suggest the occurrence of adaptive variation in introduced populations to South America, particularly in the Toltén River—which was specifically accounted for in Factor 3. Examination of allele frequencies of the 117 candidate loci for Chinook Salmon (Table S10) indicated that none of the outliers were fixed for alternative alleles among populations, and all outlier loci were polymorphic in one or more population from the native range. We were able to align paired end sequences for 111 of 117 outlier loci to achieve longer search strings (mean = 444 bp; Table S10) for BLAST searches and alignment to the Atlantic Salmon genome. With these 111 loci, 47 aligned in the BLAST search at the e^{-10} threshold. Several gene functions were represented in the BLAST hits and include several pathways (Table S10). Of the 111 sequences, we also aligned 85 of these to the Atlantic Salmon genome with MAPQ scores ≥ 3 (Table S10).

3.4.2 | Brook Charr

Outlier tests for Brook Charr in PCAdapt were completed with 12 latent factors ($K = 12$) among the 12 collections. Differentiation among populations for each factor is shown in Figure 6, but our results focused on only two factors (Factor 9 and Factor 10; Fig. S4) that were represented in the top 1% of outlier loci (Table S11). Of the top outlier loci, 96.3% were associated with Factor 10 that represented Parrillar and Gallego Chico populations relative to all other populations, while the remaining 3.7% were associated with fish from Paradise Hatchery relative to all others in the study (Factor 9). These results may support adaptive hypotheses in these collections. Examination of allele frequencies of the 135 candidate loci for Brook Charr (Table S11) indicated that most outliers (74.8%) were due to differences in allele frequencies among populations rather than fixed differences. However,

some outlier loci (34 of the 135) that were polymorphic in introduced populations were fixed in populations that were included from the native range. We were able to align paired end sequences for 129 of 135 outlier loci to achieve longer search strings (mean = 477 bp; Table S11) for BLAST searches and alignment to the Atlantic Salmon genome. With these 129 loci, 51 aligned in the BLAST search at the e^{-10} threshold (Table S11). Of the 129 sequences, we also aligned 108 of these to the Atlantic Salmon genome with MAPQ scores ≥ 3 (Table S11).

4 | DISCUSSION

Genomic variation plays a potentially complicated role in establishment of introduced species to new environments (Dlugosch et al., 2015). Here, genomic variation of introduced Chinook Salmon and Brook Charr in South America revealed patterns of diversity, population structure, and selection that may be related to invasion success. However, differences in dispersal potential between these two species appeared to be a driving factor of genomic variation and extent of their distribution. Specifically, introduced Chinook Salmon appear to have dispersed broadly in Patagonia by colonizing new freshwater rivers through anadromous migration or escaping from ocean-rearing pens (Correa & Gross, 2008; Di Prinzio et al., 2015), while resident Brook Charr remain relatively isolated in drainages where they were physically stocked by humans with little evidence for natural colonization of new systems (Arismendi et al., 2014; Gallardo et al., 2007). The intensity of stocking efforts likely contributed to the current distribution of each species, but life history diversity and potential for migratory dispersal appear to have been key factors in range expansion.

Each of the introduced species had putative source stocks and natural populations from their native range that provided important references for comparison with introduced populations in South America. For Chinook Salmon, collections from Aysén and Baker rivers in Patagonia clustered with the Cowlitz collection from the lower Columbia River in North America, and the Toltén and Petrohué collections from Patagonia were most similar to the source of Soos Hatchery from Puget Sound in North America. These genetic stocks correspond to a large extent to specific source populations that have been identified in a previous study with very dense representation of source collections (146 native populations from 46 genetic lineages; Correa & Moran, 2016). The other three collections of Chinook Salmon from North America allowed us to account for patterns of natural genomic variation as well as scenarios of colonization from multiple sources or possible admixture (Correa & Gross, 2008; Di Prinzio et al., 2015). Despite the first introduction attempts in the early 1900's, stocking history and genetic analyses indicate that current populations of Chinook Salmon originated from introductions in Chile for ocean ranching in the late 1970s and early 1980s primarily from the lower Columbia River (Cowlitz River Hatchery spring-run; Correa & Gross, 2008; Di Prinzio et al., 2015). In fact, this lineage prevails in our samples south of 43°S (i.e., Aysén and Baker; Correa & Moran, 2016). Additional accidental escapes from net-pen aquaculture in the 1990s

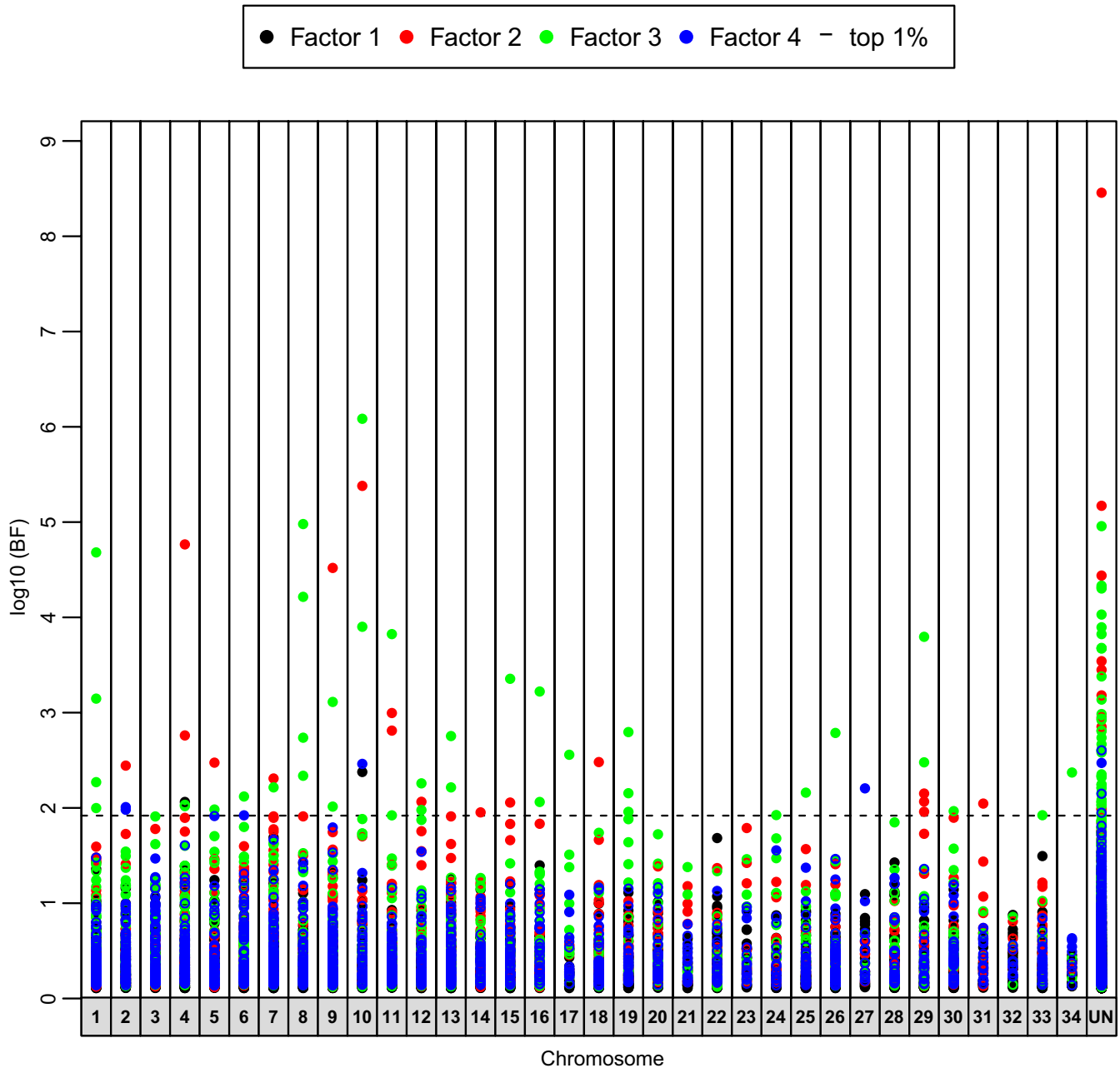


FIGURE 5 Outlier loci for each factor from PCAdapt analyses for Chinook Salmon with 11,759 SNP markers. The top 1% loci are shown above the dotted line based on chromosome position (unmapped markers are in the “UN” bin). Population abbreviations follow those in Table 1. Factors are represented by the following: F1 = Toltén/Petrohué from native, F2 = exotic from native, F3 = Toltén/Rock from all others, F4 = exotic from native

explain a more diverse assembly of lineages in the north (i.e., Toltén and Petrohué) traceable to South Puget Sound (akin to Soos Hatchery), interior Columbia Basin (akin to Hanford Reach and Priest Rapids), and other sources (Correa & Moran, 2016).

For Brook Charr, Paradise Hatchery from North America was a probable source stock representing Lago Fagnano and Alonso River in the extreme southern end of Patagonia. This was consistent with expectations that Paradise Hatchery would be a possible source stock as it has been identified as a common source for broad introductions in other geographic regions outside the species' range

(Neville & Bernatchez, 2013). The other collections of Brook Charr from Patagonia were highly distinct from those in the native range in Quebec, Canada, which is consistent with the high level of genetic divergence commonly observed in this species (e.g., Marie et al., 2010). Results also indicated that additional unidentified source stocks were used for introductions and likely contribute to signals of genetic structure of Brook Charr in Patagonia. Although Brook Charr introductions were attempted in Patagonia in the early 1900's (Arismendi et al., 2014), stocking records of Brook Charr show that there was only one documented introduction event in 1950 to Laguna Parrillar and

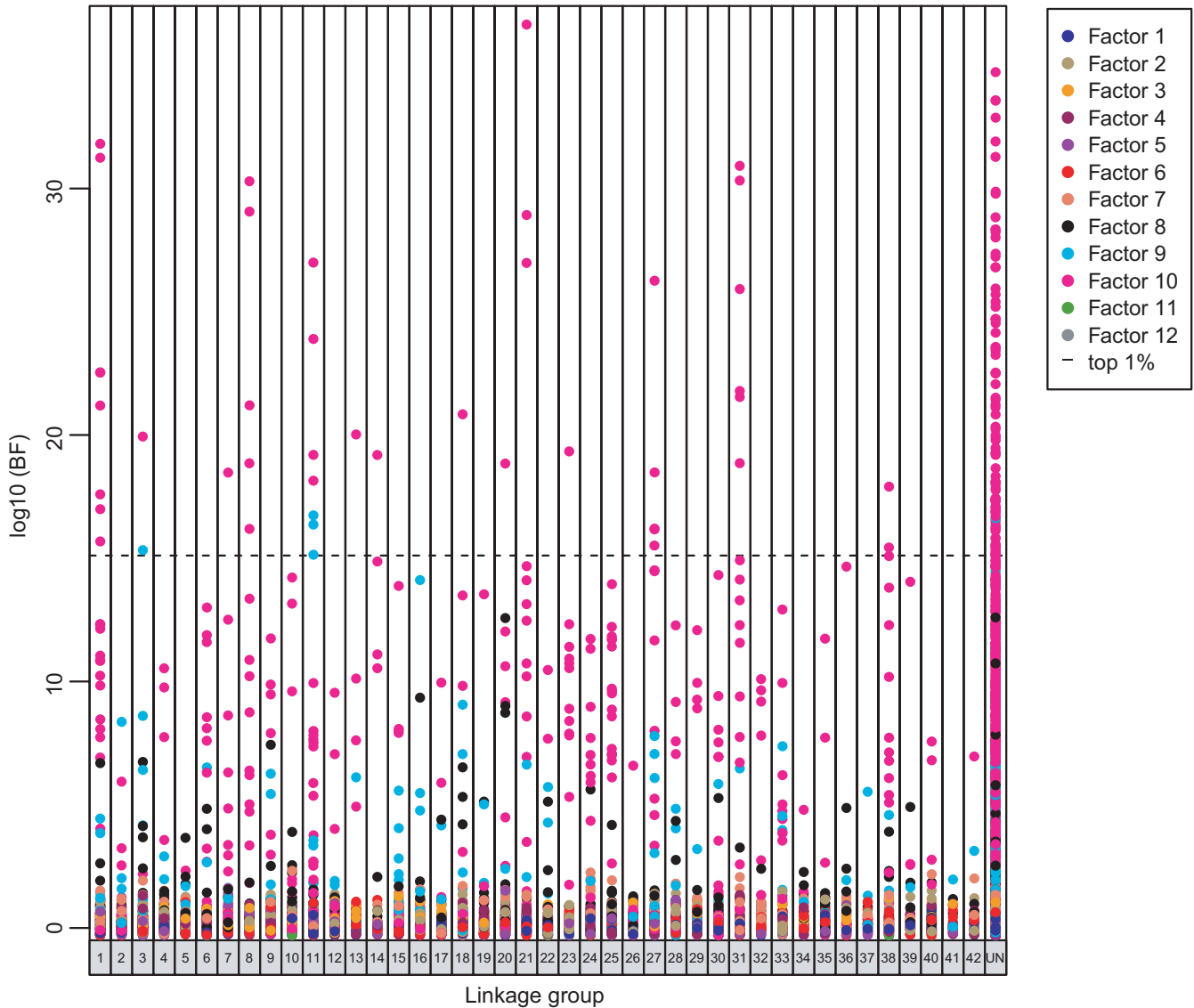


FIGURE 6 Outlier loci for each factor from PCAadapt analyses for Brook Charr with 13,522 SNP markers. The top 1% loci are shown above the dotted line and are represented by Factor 9 (Paradise Hat from all other populations) and Factor 10 (Parrillar and Gallego Chico from all other populations). Population abbreviations follow those in Table 1

Gallego Chico R., but multiple events since 1977 for Lago Fagnano, and between 1931 and 1971 for Los Lagos (Gallardo et al., 2007; Macchi, 2004). Records also indicate that fish were derived from non-native stocks reared in Argentina that had been previously introduced in waters in southeastern Patagonia (Gallardo et al., 2007).

Levels of genetic diversity in the introduced populations of each species were not significantly different than estimates of diversity from native stocks, which reflected evidence of propagule pressure such as multiple introduction attempts and admixture of source stocks. This is consistent with reviews of studies of invasive species that genetic bottlenecks in founding populations are often overcome by propagule pressure (Dlugosch et al., 2015; Lockwood et al., 2005; Roman & Darling, 2007). In Chinook Salmon, genetic diversity of Patagonia stocks has probably been maintained by admixture or multiple colonization events from various stocks that escaped from

ocean ranching (Di Prinzio et al., 2015). Given the significantly lower N_b of Chinook Salmon that we observed in South America collections relative to native stocks, ongoing colonization events and admixture would explain how genetic diversity has been maintained. Admixture and colonization could have caused a downward bias in our estimates of N_b for Patagonia stocks, but these factors would not have a large effect on estimates unless migration is high (i.e., 10 times the equilibrium rate; Waples & England, 2011).

In contrast, limited evidence for genetic bottlenecks in introduced Brook Charr relative to native stocks was rather surprising as this resident species is typically isolated and has had generally lower propagule pressure than many other introduced salmonids in South America (Consuegra et al. 2011; Arismendi et al., 2014). Stocking frequency may have an effect though as the two Brook Charr populations that were founded from a single stocking event (Parillar and Gallego Chico)

had lower diversity and effective sizes than other introduced populations with repeated stocking events. This suggests that even relatively low-intensity propagule pressure may be sufficient to retain genetic diversity for populations that were founded from stocks with low diversity and small effective size. However, examination of additional populations and more information regarding stocking history would be necessary to confirm this hypothesis.

Patterns of genetic structure differed between species introduced to Patagonia, with relatively low genetic differentiation for widespread Chinook Salmon compared to localized populations of Brook Charr with high divergence. These patterns of genetic structure are consistent with the hypothesis that distinct introductions coupled with connectivity and gene flow may be important for establishing populations on a broad geographic scale, consistent with reviews of introduced salmonids (Arismendi et al., 2014) and other invasive aquatic species (Roman & Darling, 2007). Once released or escaped from ocean aquaculture pens, anadromous Chinook Salmon have opportunities for colonization of multiple river systems with suitable habitat (Di Prinzio et al., 2015). Also, these anadromous fish returning from the ocean to freshwater may not always be philopatric, and thus, colonization of new areas may occur (Keefer & Caudill, 2014; Quinn, 1993). This migratory characteristic may allow for existing populations to be bolstered by new migrants and persist through stochastic natural events that might otherwise lead to local extirpation. Diverse life histories for Chinook Salmon have been previously shown to be key factors for colonization success not only in Patagonia (Correa & Gross, 2008; Di Prinzio & Pascual, 2008), but also in other novel environments such as New Zealand (e.g., Quinn, Kinnison, & Unwin, 2001).

In contrast to high dispersal potential in Chinook Salmon, Brook Charr typically remain as residents in freshwater lakes and rivers, although anadromy is also observed in this species (Castric & Bernatchez, 2003; Dauwalter, McGurrin, Gallagher, & Hurley, 2014; Thériault, Bernatchez, & Dodson, 2007). Restricted migration tendency would reduce the potential for Brook Charr to disperse and colonize new areas, thus limiting connectivity of existing populations. In the native range of Brook Charr, differentiation is high and reflects isolation of these resident stocks with little potential for gene flow among populations (i.e., Castric & Bernatchez, 2003; Lamaze, Marie, Garant, & Bernatchez, 2012; Marie et al., 2010). In the introduced stocks, potential for gene flow through migratory dispersal is also limited as these are primarily resident populations, but there is likely gene flow occurring through repeated stocking events (Arismendi et al., 2014; Gallardo et al., 2007) as evidenced by estimates of genetic distance that were approximately twice lower than the native range. In their native range, Brook Charr are threatened by invasive rainbow trout (Larson & Moore, 1985; Rose, 1986) which suggests that biotic resistance is also a plausible scenario in Patagonia given the existence of other salmonids including resident rainbow and brown trout (Monzón-Argüello et al., 2014).

Outlier loci were observed in both species, but only in specific collections and not for all introduced populations. Differences between native and introduced environments were observed for both species, but were within the variation that occurs across each species' native

range (Hecht et al., 2015; Lamaze et al., 2012), suggesting that selection pressure was not likely to be strong relative to genetic drift in the introduced environments. The outlier approach with PCAdapt (Duforet-Frebourg et al., 2014) was chosen specifically for this study as it provided advantages over other popular methods (i.e., Lositan, Antao, Lopes, Lopes, Beja-Pereira, & Luikart, 2008; Beaumont & Nichols, 1996; Foll & Gaggiotti 2008) such as the ability to parse outlier loci into latent factors. It is possible that both false positives and negatives persist in our outlier analyses (e.g., De Mita et al., 2013; Lotterhos & Whitlock, 2014; Narum & Hess, 2011). However, PCAdapt has been demonstrated to be a powerful approach to account for underlying neutral structure in independent simulations to reduce false positives (Lotterhos & Whitlock, 2015). Nevertheless, we interpret our results from PCAdapt with caution as false positives could have occurred due to limited representation of specific source populations and effects of admixture from multiple stocks.

In Chinook Salmon, the collection from Toltén R. had the most outlier loci compared to other introduced collections from South America. These outliers may be explained by unaccounted founder lineages as Toltén and Petrohué populations were founded by multiple sources (Correa & Moran, 2016), but largely concordant identification of source stocks suggests that the current study accounted for much of the genomic variation that contributed to these introduced populations. While multiple ancestry and admixture could affect the performance of outlier tests, this could also lead to enhanced standing genetic variation for selection to act upon in new environments. There was also evidence for selection in native populations such as the Rock and Priest Rapids collections, which have strong differences in migration distance and precipitation that have been found to be significant factors associated with local adaptation in the native range of Chinook Salmon (Hecht et al., 2015). Outlier loci included several markers that were aligned to genes related to immune function, and exposure to novel pathogens could be selective forces for species introduced to new environments (Monzón-Argüello et al., 2014; Roman & Darling, 2007). However, many other pathways were represented by candidate loci and annotation with future resources may provide more extensive biological interpretation.

In Brook Charr, outlier loci were observed in the Parrillar and Gallego Chico populations but not in the other introduced collections. Comparisons to native populations in North America provided critical information regarding relative levels of genomic variation; however, only two of the five introduced Brook Charr collections (Lago Fagnano and Alonso) had a genetically similar native stock (Paradise Hat.) represented in our study. Lack of representative founder stocks for Parrillar, Gallego Chico, and Los Lagos could have affected outlier results, but not consistently as no outliers were detected in the Los Lagos collection. False positives are known to occur in outlier tests due to genetic bottlenecks (Teshima, Coop, & Przeworski, 2006), but estimates of genetic diversity in the Parrillar and Gallego Chico collections were similar to native collections with no significant evidence of bottlenecks. Annotation of outlier genes represented various pathways with no obvious biological signal, but many outliers remained unknown due to limited availability of genomic resources.

The results from this study are consistent with the notion that dispersal potential of aquatic species can influence the extent of invasiveness as long as surrounding environments provide suitable habitat and that genetic diversity may be retained even in species with relatively low propagule pressure and distinct population structure. This information is critical as invasive species become established and cause a myriad of negative impacts that can lead to extirpation or extinction for native species (Crawford & Muir, 2008; Taylor, Courtenay, & McCann, 1984; Townsend, 1996). In the case of salmonids that have been introduced to South America, impacts to native aquatic species have been extensive with severe reductions in indigenous fishes and other aquatic organisms (Arismendi et al., 2014; Correa & Hendry, 2012; Habit et al., 2010). Several species of salmonids are now pervasive in rivers and lakes of southern Chile and Argentina leading to altered ecosystem dynamics (Habit, Gonzalez, Ruzzante, & Walde, 2012), interference with trophic cascades (Elgueta, Gonzalez, Ruzzante, Walde, & Habit, 2013), and reduced native biodiversity (Arismendi et al., 2014; Correa, Bravo, & Hendry, 2012; McDowall, 2006; Vera-Escalona, Habit, & Ruzzante 2015). Despite relatively low estimates of effective population size of invasive species, census sizes can be much higher (Kalinowski & Waples, 2002) and cause substantial impacts on native species.

Despite the negative consequences of introduced salmonids in South America, they continue to be promoted due to high economic value in fisheries and aquaculture production. Thus, there is a need to balance the social, economic, and ecological trade-offs of these introduced species in South America (reviewed in Pascual et al., 2007). Studies that elucidate characteristics that contribute to invasion success and distribution of introduced species may help contribute to successful management scenarios (e.g., Harrison, Pavlova, Telonis-Scott, & Sunnucks, 2014). For example, introduced species such as Chinook Salmon with high propensity to disperse and colonize new areas are likely to impact native aquatic organisms across a broader region than species like Brook Charr that often remain isolated in the systems where they are planted. However, not all anadromous salmonids introduced to Patagonia have become widespread (Arismendi et al., 2014), which emphasizes the point that factors influencing colonization success can be highly complex (Dlugosch et al., 2015). Further studies that examine additional species of introduced salmonids are needed to clarify patterns of invasiveness in Patagonia.

ACKNOWLEDGEMENTS

Chinook Salmon samples from S. America were collected by C. Correa under Chilean permit issued by SUBPESCA (R. Ex. No. 445-2004), and extracted DNA was provided by Paul Moran for these samples. Brook Charr from S. America were collected by several people under permits issued to P. Gallardo in southern Patagonia and Javier Marín in Los Lagos. Denise Kelsey provided the maps in Figure 1. Eric Normandeau provided bioinformatics support for assigning positions to anonymous Brook Charr markers. During the preparation of this article, C. Correa was funded by grants CONICYT-PAI 82130009 and FONDECYT

11150990. Funding for samples for S. America was provided by grants CONICYT-FONDEF D10i1258 to P. Gallardo and NSERC-Discovery to Mart Gross. Funding for samples from N. America was provided by the Bonneville Power Administration 200900500 to S. Narum.

DATA SHARING

Supplementary data of individual genotypes for all SNP loci are stored in a Dryad digital repository under accession ID: <https://doi.org/10.5061/dryad.2f050>

REFERENCES

- Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., & Luikart, G. (2008). LOSITAN: A workbench to detect molecular adaptation based on a Fst-outlier method. *BMC Bioinformatics*, *9*, 323.
- Arismendi, I., Penaluna, B. E., Dunham, J. B., de Leaniz, C. G., Soto, D., Fleming, I. A., & León-Muñoz, J. (2014). Differential invasion success of salmonids in southern Chile: Patterns and hypotheses. *Reviews in Fish Biology and Fisheries*, *24*, 919–941.
- Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., ... Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, *3*, 1–7.
- Basulto, S. (2003). El largo viaje de los salmones. Una crónica olvidada, Maval Editorial, Santiago de Chile.
- Beaumont, M., & Nichols, R. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the National Academy of Sciences of the United States of America*, *263*, 1619–1626.
- Benjamini, Y., & Yekutieli, D. (2001). The control of false discovery rate under dependency. *Annals of Statistics*, *29*, 1165–1188.
- Bock, D. G., Caseys, C., Cousens, R. D., Hahn, M. A., Heredia, S. M., Hübner, S., ... Rieseberg, L. H. (2015). What we still don't know about invasion genetics. *Molecular Ecology*, *24*, 2277–2297.
- Brieuc, M. S. O., Waters, C. D., Seeb, J. E., & Naish, K. A. (2014). A dense linkage map for Chinook Salmon (*Oncorhynchus tshawytscha*) reveals variable chromosomal divergence after an ancestral whole genome duplication event. *G3: Genes Genomes and Genetics*, *4*, 447–460.
- Candy, J. R., Campbell, N. R., Grinnell, M. H., Beacham, T. D., & Narum, S. R. (2015). Population differentiation determined from putative neutral and adaptive genetic markers in Eulachon (*Thaleichthys pacificus*, Osmeridae), and anadromous smelt. *Molecular Ecology Resources*, *15*, 1421–1434.
- Castric, V. L., & Bernatchez, L. (2003). The rise and fall of isolation by distance in the anadromous Brook Charr (*Salvelinus fontinalis* Mitchell). *Genetics*, *163*, 983–996.
- Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping Loci de novo from short-read sequences. *G3: Genes Genomes and Genetics*, *1*, 171–182.
- Catchen, J., Hohenlohe, P., Bassham, S., Amores, A., & Cresko, W. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140.
- Cohen, A. N., & Carlton, J. T. (1998). Accelerating invasion rate in a highly invaded estuary. *Science*, *279*, 555–558.
- Conesa, A., Götz, S., Garcia-Gomez, J. M., Terol, J., Talon, M., & Robles, M. (2005). Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, *21*, 3674–3676.
- Consuegra, S., Phillips, N., Gajardo, G., & de Leaniz, C. G. (2011). Winning the invasion roulette: Escapes from fish farms increase admixture and facilitate establishment of non-native rainbow trout. *Evolutionary Applications*, *4*, 660–671.
- Correa, C., Bravo, A. P., & Hendry, A. P. (2012). Reciprocal trophic niche shifts in native and invasive fish: Salmonids and galaxiids in Patagonian lakes. *Freshwater Biology*, *57*, 1769–1781.

- Correa, C., & Gross, M. R. (2008). Chinook Salmon invade southern South America. *Biological Invasions*, *10*, 615–639.
- Correa, C., & Hendry, A. P. (2012). Invasive salmonids and lake order interact in the decline of Puye Grande Galaxias platei in western Patagonian lakes. *Ecological Applications*, *22*, 828–842.
- Correa, C., & Moran, P. (2016). Polyphyletic ancestry of expanding Patagonian Chinook salmon populations. bioRxiv doi: <https://doi.org/10.1101/079459>.
- Crawford, S. S., & Muir, A. M. (2008). Global introductions of salmon and trout in the genus *Oncorhynchus*: 1870–2007. *Reviews in Fish Biology and Fisheries*, *18*, 313–344.
- Dauwalter, D. C., McGurrin, J., Gallagher, M., & Hurley, S. (2014). Status assessment of coastal and anadromous Brook trout in the United States. Wild Trout Symposium XI.
- Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, *12*, 499–510.
- De Mita, S., Thuillet, A. C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., & Vigouroux, Y. (2013). Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, *22*, 1383–1399.
- Di Prinzio, C. Y., & Pascual, M. A. (2008). The establishment of exotic Chinook Salmon (*Oncorhynchus tshawytscha*) in Pacific rivers of Chubut, Patagonia, Argentina. *International Journal of Limnology*, *1*, 61–68.
- Di Prinzio, C. Y., Riva Rossi, C., Ciancio, J., Garza, J. C., & Casaux, R. (2015). Disentangling the contributions of ocean ranching and net-pen aquaculture in the successful establishment of Chinook Salmon in a Patagonia basin. *Environmental Biology of Fishes*, *98*, 1987–1997.
- Dlugosch, K. M., Anderson, S. R., Braasch, J., Cang, F. A., & Gillette, H. D. (2015). The devil is in the details: Genetic variation in introduced populations and its contributions to invasion. *Molecular Ecology*, *24*, 2095–2111.
- Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Molecular Ecology Resources*, *14*, 209–214.
- Duforet-Frebourg, N., Bazin, E., & Blum, M. G. B. (2014). Genome scans for detecting footprints of local adaptation using a Bayesian factor model. *Molecular Biology and Evolution*, *31*, 2483–2495.
- Elgueta, A., Gonzalez, J., Ruzzante, D., Walde, S. J., & Habit, E. (2013). Trophic interference by *Salmon trutta* on *Aplocheilichthys zebra* and *Aplocheilichthys taeniatus* in southern Patagonian lakes. *Journal of Fish Biology*, *82*, 430–443.
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*, 2611–2620.
- Facon, B., Genton, B. J., Shykoff, J., Jarne, P., Estoup, A., & David, P. (2006). A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology and Evolution*, *21*, 130–135.
- Foll, M., & Gaggiotti, O. E. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, *180*, 977–993.
- Frankham, R. (2005). Genetics and extinction. *Biological Conservation*, *126*, 131–140.
- Gallardo, P., Aguila, R., Roa, R., Cañete, I., Miranda, C., & Olave, C. (2007). Bases técnicas y ambientales para la elaboración de un plan de manejo y reglamentación de la pesca recreativa en la Laguna Parrillar. Fondo de Investigación Pesquera-Chile. Informe Técnico Final FIP 2004-54, 139 pag. Retrieved from <http://www.fip.cl/FIP/Archivos/pdf/informes/inffinal%202004-54.pdf>.
- Gruenthal, K. M., Witting, D. A., Ford, T., Neuman, M. J., Williams, J. P., Pondella, II D., ... Larson, W. A. (2014). Development and application of genomic tools to the restoration of green abalone in southern California. *Conservation Genetics*, *15*, 109–121.
- Habit, E., Gonzalez, J., Ruzzante, D. E., & Walde, S. J. (2012). Native and introduced fish species richness in Chilean Patagonian lakes: Inferences on invasion mechanisms using salmonid-free lakes. *Divers Distributed*, *18*, 1153–1165.
- Habit, E., Piedra, P., Ruzzante, D. E., Walde, S. J., Belk, M. C., Cussac, V. E., ... Colin, N. (2010). Changes in the distribution of native fishes in response to introduced species and other anthropogenic effects. *Global Ecology and Biogeography*, *19*, 697–710.
- Hamilton, J. A., Okada, M., Korves, T., & Schmitt, J. (2015). The role of climate adaptation in colonization success in *Arabidopsis thaliana*. *Molecular Ecology*, *24*, 2253–2263.
- Harrison, K., Pavlova, A., Telonis-Scott, M., & Sunnucks, P. (2014). Using genomics to characterize evolutionary potential for conservation of wild populations. *Evolutionary Applications*, *7*, 1008–1025.
- Hayes, K. R., & Barry, S. C. (2008). Are there any consistent predictors of invasion success? *Biological Invasions*, *10*, 483–506.
- Hecht, B. C., Matala, A. P., Hess, J. E., & Narum, S. R. (2015). Environmental adaptation in Chinook Salmon (*Oncorhynchus tshawytscha*) throughout their North American range. *Molecular Ecology*, *24*, 5573–5595.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*, 1965–1978.
- Huey, R. B., Gilchrist, G. W., Carlson, M. L., Berrigan, D., & Serra, L. (2005). Using invasive species to study evolution: Case studies with *Drosophila* and Salmon. In D. F. Sax et al. (Eds.), *Species invasions: Insights into ecology, evolution and biogeography* (pp. 139–164). Sunderland, MA: Sinauer Associates.
- Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405.
- Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*, 3070–3071.
- Kalinowski, S. T. (2005). HP-Rare: A computer program for performing rarefaction on measures of allelic diversity. *Molecular Ecology Notes*, *5*, 187–189.
- Kalinowski, S. T., & Waples, R. S. (2002). Relationship of effective to census size in fluctuating populations. *Conservation Biology*, *16*, 129–136.
- Keane, R., & Crawley, M. (2002). Exotic plant invasions and the enemy release hypothesis. *Trends in Ecology and Evolution*, *17*, 164–170.
- Keefer, M. L., & Caudill, C. C. (2014). Homing and straying by anadromous salmonids: A review of mechanisms and rates. *Reviews in Fish Biology and Fisheries*, *24*, 333–368.
- Kelly, D. W., Muirhead, J. R., Heath, D. D., & MacIsaac, H. J. (2006). Contrasting patterns in genetic diversity following multiple invasions of fresh and brackish waters. *Molecular Ecology*, *15*, 3641–3653.
- Kinnison, M. T., Unwin, M. J., & Quinn, T. P. (2008). Eco-evolutionary vs. habitat contributions to invasion in salmon: Experimental evaluation in the wild. *Molecular Ecology*, *17*, 405–414.
- Kolbe, J. J., Leal, M., Schoener, T. W., Spiller, D. A., & Losos, J. B. (2012). Founder effects persist despite adaptive differentiation: A field experiment with lizards. *Science*, *335*, 1086–1089.
- Lamaze, F. C., Marie, A., Garant, D., & Bernatchez, L. (2012). Dynamics of introgressive hybridization assessed by coding genes SNP population genomics in stocked brook charr (*Salvelinus fontinalis*). *Molecular Ecology*, *21*, 2877–2895.
- Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*, R25.
- Larson, G., & Moore, S. (1985). Encroachment of exotic rainbow trout into stream populations of native brook trout in the southern Appalachian Mountains. *Transactions of the American Fisheries Society*, *114*, 195.
- Lien, S., Koop, B. F., Sandve, S. R., Miller, J. R., Kent, M. P., Nome, T., ... Davidson, W. S. (2016). The Atlantic salmon genome provides insights into rediploidization. *Nature*, *533*, 200–205.
- Lockwood, J. L., Cassey, P., & Blackburn, T. (2005). The role of propagule pressure in explaining species invasions. *Trends in Ecology and Evolution*, *20*, 223–228.

- Lodge, D. M. (1993). Biological invasions: Lessons for ecology. *Trends in Ecology and Evolution*, 8, 133–137.
- Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, 23, 2178–2192.
- Lotterhos, K. E., & Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, 24, 1031–1046.
- Luikart, G., Ryman, N., Tallmon, D. A., Schwartz, M. K., & Allendorf, F. W. (2010). Estimation of census and effective population sizes: The increasing usefulness of DNA-based approaches. *Conservation Genetics*, 11, 355–373.
- Macchi, P. J. (2004). Respuestas de *Galaxias maculatus* a la depredación por parte de *Percichthys trucha* y los salmónidos introducidos en ambientes lénticos de la Patagonia norte. PhD Thesis, Universidad Nacional del Comahue. 175 pp.
- MacCrimmon, H. R. (1971). World distribution of rainbow trout (*Salmo gairdneri*). *Journal of Fisheries and Research Board of Canada*, 28, 663–704.
- Marie, A., Bernatchez, L., & Garant, D. (2010). Loss of genetic integrity correlates with stocking intensity in brook charr (*Salvelinus fontinalis*). *Molecular Ecology*, 19, 2025–2037.
- McDowall, R. M. (2006). Crying wolf, crying foul, or crying shame: Alien salmonids and a biodiversity crisis in the southern cool-temperate galaxioid fishes? *Reviews in Fish Biology and Fisheries*, 16, 233–422.
- Monzón-Argüello, C., de Leaniz, C. G., Gajardo, G., & Consuegra, S. (2014). Eco-immunology of fish invasions: The role of MHC variation. *Immunogenetics*, 66, 393–402.
- Moyle, P. B., & Marchetti, M. P. (2006). Predicting invasion success: Freshwater fishes in California as a model. *BioScience*, 56, 515–524.
- Narum, S. R. (2006). Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conservation Genetics*, 7, 783–787.
- Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., & Hohenlohe, P. A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, 22, 2841–2847.
- Narum, S. R., & Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, 11, 184–194.
- Neville, H. M., & Bernatchez, L. (2013). Coding gene single nucleotide polymorphism population genetics of nonnative brook trout: The ghost of introductions past. *Transactions of the American Fisheries Society*, 142, 1215–1231.
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., ... Wagner, H. (2013). *vegan*: Community Ecology Package. R package version 2.0-10. Retrieved from <https://CRAN.R-project.org/package=vegan>
- Pascual, M., Cussac, V., Dyer, B., Soto, D., Vigliano, P., Ortubay, S., & Macchi, P. (2007). Freshwater fishes of Patagonia in the 21st Century after a hundred years of human settlement, species introductions, and environmental change. *Aquatic Ecosystem Health Management*, 10, 212–227.
- Pimentel, D., Zuniga, R., & Morrison, D. (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecology and Economics*, 52, 273–288.
- Quinn, T. P. (1993). A review of homing and straying of wild and hatchery-produced salmon. *Fisheries Research*, 18, 29–44.
- Quinn, T. P., Kinnison, M. T., & Unwin, M. J. (2001). Evolution of Chinook Salmon (*Oncorhynchus tshawytscha*) populations in New Zealand: Pattern, rate, and process. *Genetica*, 112, 493–513.
- R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <http://www.R-project.org/>.
- Roman, J., & Darling, J. A. (2007). Paradox lost: Genetic diversity and the success of aquatic invasions. *Trends in Ecology and Evolution*, 22, 454–464.
- Rose, G. A. (1986). Growth decline in subyearling brook trout (*Salvelinus fontinalis*) after emergence of rainbow trout (*Salmo gairdneri*). *Canadian Journal of Fisheries and Aquatic Sciences*, 43, 187–193.
- Rousset, F. (2008). Genepop'007: A complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, 8, 103–106.
- Sax, D. F., Stachowicz, J. J., Brown, J. H., Bruno, J. F., Dawson, M. N., Gaines, S. D., ... Rice, W. R. (2007). Ecological and evolutionary insights from species invasions. *Trends in Ecology and Evolution*, 22, 465–471.
- Stockwell, C. A., Hendry, A. P., & Kinnison, M. T. (2003). Contemporary evolution meets conservation biology. *Trends in Ecology and Evolution*, 18, 94–101.
- Sutherland, B. J. G., Gosselin, T., Normandeau, E., Lamothe, M., Isabel, N., Audet, C., & Bernatchez, L. (2016). Novel method for comparing RADseq linkage maps reveals chromosome evolution in salmonids. *Genome Biology and Evolution*, 8, 3600–3617.
- Taylor, J. N., Courtenay, W. R., & McCann, J. A. (1984). Known impacts of exotic fishes in the continental United States. In W. R. Courtenay, & J. R. Stauffer (Eds.), *Distribution, biology, and management of exotic fishes* (pp. 322–373). Baltimore: Johns Hopkins University Press.
- Teshima, K. M., Coop, G., & Przeworski, M. (2006). How reliable are empirical genomic scans for selective sweeps? *Genome Research*, 16, 702–712.
- Thériault, V., Bernatchez, L., & Dodson, J. (2007). Mating system and individual reproductive success of sympatric anadromous and resident Brook Charr, *Salvelinus fontinalis*, under natural conditions. *Behavioral Ecology and Sociobiology*, 62, 51–65.
- Townsend, C. R. (1996). Invasion biology and ecological impacts of brown trout (*Salmo trutta*) in New Zealand. *Biological Conservation*, 78, 13–22.
- Vera-Escalona, I., Habit, E., & Ruzzante, D. (2015). Echoes of a distant time: Effects of historical processes on contemporary genetic patterns in *Galaxias platei* in Patagonia. *Molecular Ecology*, 16, 4112–4128.
- Waples, R. S. (2006). A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, 7, 167–184.
- Waples, R. S., & England, P. R. (2011). Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. *Genetics*, 189, 633–644.
- Waples, R. K., Larson, W. A., & Waples, R. S. (2016). Estimating contemporary effective population size in non-model species using linkage equilibrium across thousands of loci. *Heredity*, 117, 233–240.
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38, 1358–1370.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Narum SR, Gallardo P, Correa C, et al. Genomic patterns of diversity and divergence of two introduced salmonid species in Patagonia, South America. *Evol Appl*. 2017;10:402–416. <https://doi.org/10.1111/eva.12464>