EVOLUTIONARY GENOMICS OF A PLASTIC LIFE HISTORY TRAIT: *GALAXIAS MACULATUS* AMPHIDROMOUS AND RESIDENT POPULATIONS

by

María Lisette Delgado Aquije

Submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
August 2021

Dalhousie University is located in Mi'kma'ki, the
ancestral and unceded territory of the Mi'kmaq.
We are all Treaty people.

I dedicate this work to my parents, María and José, my brothers JR and Eduardo for their unconditional love and support and for always encouraging me to pursue my dreams, and to my grandparents Victoria, Estela, Jesús, and Pepe whose example of perseverance and hard work allowed me to reach this point.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Understanding the molecular mechanisms that differentiate diadromous from freshwater resident (non-migratory) populations can help understand how species adapt to changing conditions. In this thesis, I first review the literature concerning diadromy in fishes. I discuss how diadromy appears to be the product of independent evolutionary events and how -omics approaches are answering questions regarding the genetic basis, origin, and loss of this life history trait. Using *Galaxias maculatus*, an amphidromous fish from the Southern Hemisphere, as a model and following a RADcap approach, I found that diadromous individuals comprise mainly one large population across the species distribution in Chile, while resident populations, particularly those in the northernmost locations are the product of independent colonization events from a common diadromous source. These geographically close but genetically distinguishable resident populations can thus be considered natural replicates derived from a single diadromous population. A reciprocal transplant experiment consisting of gradual salinity changes with estuarine and resident individuals from two replicate populations, Toltén and Valdivia, revealed that Valdivia residents retained the ability to survive in saltwater environments, but Toltén residents did not. An outlier analysis identified SNPs differentiating diadromy from residency, and the ability, or lack thereof, to survive in salt water. To further understand how diadromous, Toltén resident and Valdivia resident individuals acclimate to salt water and to assess their physiological stress response, I performed an acute salinity change experiment where salinity was changed from 0 ppt to 23-25 ppt. Diadromous and Valdivia resident individuals showed no sign of stress 48 hours post-change, while Toltén residents could not survive the change in salinity. Gill RNAseq analyses revealed key genes related to osmotic adaptations in *G. maculatus* and showed differences between resident populations in the number of genes with retained and lost transcriptional responses. In Toltén residents, key genes including ion transporters (e.g., CFTR) were not upregulated in salt water, suggesting a potential mechanism for the loss of salinity tolerance. Overall, this thesis gives support to the hypothesis that the loss of diadromy can be achieved by several pathways and that drift likely plays an important role in the evolution of resident populations.

LIST OF ABBREVIATION USED

Abbreviation	Description
bp	base pairs
CFTR	Cystic Fibrosis Transmembrane Conductance Regulator
CV	Cross-Validation
DGE	Differential Gene Expression
DNA	Deoxyribonucleic acid
F_{ST}	Fixation index
FW	Fresh water
geno	genotype
He	Expected heterozygosity
Ho	Observed heterozygosity
hr	hour
HWE	Hardy-Weinberg Equilibrium
K	Fulton's condition factor
Km	Kilometer
L	Liter
LD	Linkage disequilibrium
maf	minor allele frequency
minDP	minimum depth
ml	milliliter
mtDNA	mitochondrial DNA
mya	million years ago
N	sample size
Ne	effective size
NKA	Sodium/Potassium ATPase
ORP	Oyster River Protocol
PCA	Principal Component Analysis
ppt	parts per thousand
RADseq	Restriction site-associated DNA sequencing

RNA	Ribonucleic acid
RNAseq	RNA sequencing
RRS	Reduced Representation Sequencing
SNP	Single Nucleotide Polymorphism
SW	Salt water
TolRes	Toltén resident
ValEst	Valdivia estuary
ValRes	Valdivia resident
W1	wet tissue weight
W2	dry tissue weight

ACKNOWLEDGMENTS

I would like to thank my supervisor, Daniel Ruzzante, for giving me the opportunity to do my graduate studies and work in his lab and for his guidance and support during these years. I also would like to thank my committee members Anne Dalziel and Paul Bentzen for their support and helpful feedback which improved this research. Thank you to my co-authors from Chile, Evelyn Habit, Aliro Manosalva, Mauricio Urbina, Oscar Link, and Konrad Górski who welcomed me in their labs. Their help with sample collection and fish rearing was crucial to complete my thesis.

My gratitude to the agencies that funded my studies and lab work, the Nova Scotia Graduate Scholarship, Dalhousie's Faculty of Graduate Studies Scholarship, the Lett fund from the Department of Biology, the NSERC Discovery Grant to Daniel Ruzzante, and to Compute Canada and their support team for providing the frame to run my analyses and helping me with the use of their clusters.

Thank you to Pablo Fierro and Gustavo Díaz for their assistance in the field, Greg McCracken for his support in the molecular lab, Ian Bradbury for providing adapters for the RADseq experiment, René Iribarren for his support rearing the fish, Jon Puritz for his help with the use of dDocent, Angela Fuentes and Tony Einfeldt for answering my bioinformatic questions, and to the editors and anonymous reviewers for their useful comments to improve my manuscripts.

My deepest gratitude to my friends in Halifax who became my family during these years: Angela, Sarah, James, Noor, Francisca, Mailén, Ana E, Ana C, Taylor, and Carlos. Also thank you to my other friends and BOGS members for making graduate studies more fun. Last by definitely not least, I could not have been here without my family and friends from Peru. Even though we are miles away, I always felt your support. To my mom, my dad, and grandparents, thank you because if I am here is because of your hard work and example. To my brothers JR and Eduardo, my sister Ana, and my best friends Zoila, Alejandra, and Alexis thank you for always being in touch and making me feel close to home. And to my aunts, Chacha, Lulú, and Capullo thank you for your support especially for being there for us when my dad was ill.

CHAPTER 1. INTRODUCTION

How species adapt to new environments has long been a focal question in ecology and evolution (Orsini *et al.* 2013) and given the current and forecasted changes in climate, addressing it has acquired a certain degree of urgency (Urban 2015). Our increased understanding of genetics has helped define the evolutionary and molecular mechanisms underlaying adaptation (Orsini *et al.* 2013), yet our knowledge regarding the direction of evolution and the potential and limitations of species to adapt to changes in the environment is still rather limited. In aquatic systems, the transition between marine and freshwater environments has been key in shaping the diversity and distribution of fishes worldwide (Bloom *et al.* 2013). Untangling how this transition occurs is certain to add to our knowledge of how species adapt to contrasting environments. The question of which molecular mechanisms facilitated this transition is thus key to our understanding of the origin of diversity in aquatic environments, and diadromy may hold clues to its answer.

Diadromy has been described as the predictable migration between marine and freshwater environments (McDowall 1997). Diadromous fishes are classified into three categories (Gross 1987): anadromy refers to systems in which individuals spawn in fresh water, then migrate to the sea, and subsequently return to fresh water to reproduce. Catadromy refers to systems in which individuals spawn in the marine environment, then migrate to fresh water, and subsequently return to the sea to reproduce. Lastly, amphidromy refers to systems in which individuals spawn in fresh water, migrate to the sea as larvae and return to fresh water after weeks or months to complete their life cycle (McDowall 1997).

It has been hypothesized that diadromy is an intermediate step in the evolution of fully marine and fully freshwater species. Recent research suggests that in some taxa, diadromy is indeed an intermediate step (Corush 2019), thus the study of diadromous fishes can provide insights into the evolution of freshwater and marine species. The ability to survive in such distinct environments also makes diadromous species good model systems in the quest for understanding the genetic mechanisms underlying the ability to adapt to different environmental conditions.

Non-migratory populations of fish that are derived from diadromous populations, hereby called resident populations, have reduced migratory abilities or have lost them completely (Waters *et al.* 2020). Most resident populations are known to live in fresh water and have recurrently colonized freshwater environments in "landlocking" events (Waters *et al.* 2020). This colonization has led, in some cases, to the extension of the species distribution, however, local adaption has also led to diversification even resulting in speciation.

Diadromous species and their resident counterparts, thus present an opportunity to discuss:

i) The origin of diadromy and the appearance of resident populations, and the commonality of the loss of diadromy across diadromous taxa (chapter 2),

and to investigate:

ii) How genetically distinguishable are diadromous and resident populations (chapter 3)?

iii) Whether replicate resident populations retain certain levels of plasticity to survive in contrasting environments (marine or freshwater) (chapter 4).

iv) Which genes (or expression) are maintained, and which are lost when diadromy is lost (chapter 5)?

Chapter two is a review of diadromy and its loss. This chapter includes an updated list of known diadromous species, a review of the proposed hypotheses regarding the origin of diadromy, and a discussion of how comparative studies between diadromous and resident populations can help answer questions about the origin and molecular bases of diadromy (Chapter 2; Delgado & Ruzzante 2020). This review also highlights how advances in sequencing technology have increased the array of methods in genomics and transcriptomics available to investigate the differences between diadromous and resident populations.

This literature review also raises the point that most research on diadromy has thus far been centered on economically important species, most of which are anadromous from the Northern Hemisphere. Studies in species from the Southern Hemisphere and

classified as catadromous or amphidromous are lacking (Chapter 2; Delgado & Ruzzante 2020). As investigating non-model organisms can only add to our knowledge regarding the genetic architecture of adaptations (Ellegren 2014), this thesis focuses on *Galaxias maculatus*. *Galaxias maculatus* is a non-model organism widely distributed in the Southern Hemisphere (Fig. 1.1). In this thesis, I used it to understand the genetics of diadromous and resident populations, and what molecular mechanisms may occur after the loss of diadromy.

Figure 1.1 Photo of a *Galaxias maculatus* individuals. Source: Paddy Ryan / http://www.ryanphotographic.com/galaxiidae.htm

1.1 *Galaxias maculatus*

Galaxias maculatus occurs in a broad range of locations including Australia, New Zealand, and South America (Barbee *et al.* 2011), fulfilling different roles. From an ecological perspective, it is the main planktivorous fish in Patagonian lakes (Barriga *et al.* 2012). And economically, it is an important resource to local fisheries in many countries (Waters & Burridge 1999; Barile *et al.* 2013). This species can live in a wide variety of environments including salinities from 1 to 62 ppt and temperatures from 11 to 29°C (Laurenson *et al.* 2012). This species is believed to have originated in Australia and to have dispersed by the West Wind Drift to South America (Waters & Burridge 1999; Zemlak *et al.* 2010). Its marine migratory behavior has been hypothesized as the reason for this species' wide distribution, and studies suggest the possibility of an 'ongoing' marine dispersal (Waters *et al.* 2000b).

Due to this migratory behavior, *G. maculatus* is classified as an amphidromous species, which means that individuals migrate from freshwater to marine environments during the larval stage. Larvae remain in the sea for a period of four to six months before returning to the river to complete their life cycle (McDowall *et al.* 1994). Juveniles are transparent and adults are recognized for their pigmentation and their size range between 6 and 10 cm (Rojo *et al.* 2018). Their generation time is short and has been assumed to be between one and two years (Burridge *et al*., 2008). In the case of *G. maculatus*, it has been reported that adults reach between one and up to four years of age (Chapman *et al.* 2006).

Genetic studies in *Galaxias maculatus* are scarce, with most studies focused on either mitochondrial or microsatellite markers (Vera-Escalona *et al.* 2020). *Galaxias maculatus* has 22 chromosomes and a genome size of 1.08 Gb (Jara-Seguel *et al.* 2008), however, thus far there is no reference genome sequence for the species. In Chile, it has been suggested that *G. maculatus* began to diversify in Northern Patagonia and spread south along the coast through the Pacific Ocean (Zemlak *et al.* 2010). Control region analyses found four distinct haplogroups (common matrilineage) with individuals from the same river system not necessarily belonging to the same haplogroup (Zemlak *et al.* 2010). Further studies using one mitochondrial marker found evidence of genetic differentiation between *G. maculatus* populations from 2 marine biogeographic regions in Chile that differ in environmental and oceanographic conditions and in the effect of glacial and habitat continuity: A region in the north referred to as 'Intermediate', and a Magallanic region in the south (González-Wevar *et al*. 2015). This study showed that northern populations have high levels of genetic diversity and structure (González-Wevar *et al*. 2015). While Magallanic populations showed low levels of diversity and no structure even though the Magallanic region presents a more heterogeneous landscape (González-Wevar *et al*. 2015). Phenotypic differences between diadromous or estuarine populations and resident or landlocked populations have also been reported. For example, diadromous populations tend to have larger body sizes (Barriga *et al.* 2012) and higher fecundity (Boy *et al*. 2009) than landlocked populations. Furthermore, diadromous and resident populations appear to exhibit contrasting life history strategies, while diadromous populations move downstream to spawn, resident populations tend to move

upstream (Chapman *et al.* 2006). Studies using bone tissue microchemistry confirmed the existence of diadromous populations across the species distribution in Chile and of allopatric resident populations living in rivers and lakes and sometimes even in sympatric resident populations inhabiting the estuaries (Górski *et al.* 2018). Having identified individuals that have spent time at sea and individuals that have never done so, allowed me to assess genetic differentiation between populations from these two life histories.

1.2 Diadromous vs. resident *G. maculatus* populations

In chapter three, I explore the genetic differences among *G. maculatus* individuals from 20 locations across the Chilean distribution using a genomic approach (i.e., RADcap). In most cases, locations were selected to include a diadromous and resident populations from the same river system, thus for every diadromous population sampled in an estuarine environment, there was a resident counterpart. Most diadromous individuals comprised mainly one large genetic group with high gene flow among sampling locations across the species distribution in Chile. While resident individuals, in particular those collected from the northernmost locations, exhibited significant genetic differentiation with their diadromous counterparts and with each other (Chapter 3; Delgado *et al.* 2019).

This genomic analysis revealed the existence of natural replicates of resident populations that despite being geographically close and derived from a common diadromous source, were phenotypically and genetically distinguishable from each other. They thus represented a good opportunity for an examination of the evolution of residency. I selected two river systems: Toltén and Valdivia for this comparison. The choice of system was made based on their geographic proximity and similar latitude, which allowed me to hypothesize that they appeared at a similar geological time despite exhibiting phenotypic (e.g., Fig. 1.2) and genetic differences (Chapter 4; Delgado *et al.* 2020).

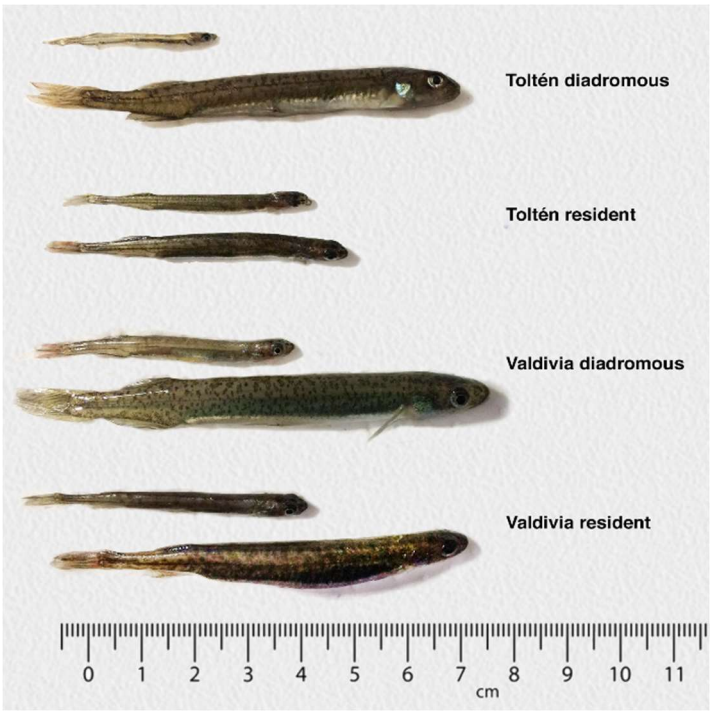


Figure 1.2 Example of size variation of diadromous and resident adults from Toltén and Valdivia River systems

Salinity is a key variable constraining the distribution of fishes (Whitehead 2010; Kultz 2015), with osmoregulatory capacity contributing to the evolution of diadromy. Thus, a main objective of chapter four was to test the ability of diadromous and resident individuals to survive changes in salinity through a reciprocal transplant experiment. Here, I found that Valdivia resident individuals maintained their ability to survive in saltwater environments showing 0% mortality at 25 ppt, while most Toltén resident individuals were not able to acclimate to the gradual changes in salinity. Further genomic analyses were performed to identify SNPs that differentiated diadromous versus resident fish populations and the ability to survive in salt water. Interestingly, genes where some of these outlier SNPs were located have previously been reported in other diadromous species (Chapter 4; Delgado *et al.* 2020).

Knowing that the molecular mechanisms allowing fishes to acclimate or adapt to saltwater environments can occur in regulatory regions (Jones *et al.* 2012) and that the time of colonization of these resident populations might be recent, I chose to analyze the differences in gene expression. To achieve this, I first conducted an abrupt salinity

transfer experiment where diadromous, Toltén resident, and Valdivia resident individuals were first held in tanks with fresh water (0 ppt) for acclimation over eight days and were then exposed to an abrupt change in salinity (23-25 ppt). For this experiment, the percentage of water content in muscle was measured over time (from 0 to 48 hours) to assess osmoregulatory capacity.

The response to the abrupt change in salinity among the Valdivia residents was similar to that exhibited by the diadromous individuals; neither group showed signs of stress 48 hours post salinity change (23-25 ppt). Toltén residents, instead, could not survive the abrupt change in salinity, exhibiting a continuous drop in the percentage of water content in muscle. The percentage of water content analysis detected the greatest differences among populations at 24 hours post-transfer. I thus performed RNAseq of individuals in fresh water (0 hr) and salt water (24hr post-change). Differential gene expression analysis for a total of 50 individuals helped identify key genes associated with osmotic adaptation in diadromous *G. maculatus* and showed that different genes are expressed or lost among Toltén and Valdivia residents as between these and the diadromous individuals.

This thesis integrates the phenotype and physiological responses to salinity variation in diadromous and resident populations of *G. maculatus* with -omics data. I showed that the pathway to the loss of diadromy differed between replicate resident populations at the phenotypic, physiological, genomic, and transcriptomic levels. In chapter six, I highlight the main results of each chapter and discuss the broader implications of my thesis and the limitations of my research. Lastly, I suggest that *Galaxias maculatus* should be considered a model organism for future studies on the evolution of diadromy and its loss.

CHAPTER 2. INVESTIGATING DIADROMY IN FISHES AND ITS LOSS IN AN -OMICS ERA

This chapter has been published as "Delgado, ML & Ruzzante, DE. 2020 Investigating diadromy in fishes and its loss in an -omics era. iScience 23(12): 101837"

2.1 Abstract

Diadromy, the predictable movements of individuals between marine and freshwater environments, is biogeographically and phylogenetically widespread across fishes. Thus, despite the high energetic and potential fitness costs involved in moving between distinct environments, diadromy appears to be an effective life history strategy. Yet, the origin and molecular mechanisms that underpin this migratory behavior are not fully understood. In this review, we aim first to summarize what is known about diadromy in fishes; this includes the phylogenetic relationship among diadromous species, a description of the main hypotheses regarding its origin, and a discussion of the presence of non-migratory populations within diadromous species. Secondly, we discuss how recent research based on -omics approaches (chiefly genomics, transcriptomics, and epigenomics) are beginning to provide answers to questions on the genetic bases and origin(s) of diadromy. Finally, we suggest future directions for -omics research that can help tackle questions on the evolution of diadromy.

2.2 Introduction

Diadromy refers to the predictable migration between marine and freshwater environments that certain species undertake during specific periods in their life (McDowall, 2008a). Although diadromy in fishes is rare (present in less than 1% of all fish species), it is widely distributed both phylogenetically and biogeographically, with many diadromous species known for their evolutionary, historical, cultural, or economic value (McDowall 1999).

The fact that diadromy involves movement between such distinct environments suggests that it requires major physiological and behavioral adaptations. In turn, such movements have important ecological and evolutionary consequences. For instance, diadromy has played a role in the genetic structure of populations (e.g. Chubb *et al*. 1998;

Delgado *et al*. 2019; Taillebois *et al*. 2013), and in postglacial colonization (e.g., Reusch *et al*. 2001; Mateus *et al*. 2016). Despite this relevant influence on species biology, little is known about the potential selective pressures leading to its origin, the molecular mechanisms underlying the capacity for diadromy, and the effects on species evolution. Why has diadromy evolved? Which genes give diadromous individuals the ability to migrate? And why have some diadromous populations stopped migrating? These are questions not yet adequately answered. Important efforts have, however, been made to improve our understanding of diadromy including the formulation of hypotheses about its origins (e.g., Gross 1987; Tsukamoto *et al.* 2009), and the search for genes that differentiate diadromous and non-migratory populations (e.g. Perrier *et al*. 2013; Taugbøl *et al.* 2014).

The development of otolith and bone tissue microchemistry during the last decade has facilitated the description and classification of diadromous fishes as this method traces the presence of individuals to marine or freshwater environments *(e.g., Hale & Swearer 2008; Feutry et al. 2012; Hughes et al. 2014; Warburton et al. 2018; Górski et al.* 2018). Additionally, technological advances in sequencing technologies, specifically in -omics (high-throughput sequencing to study large scale genomes, transcriptomes, epigenomes, etc.) are facilitating significant advances in our understanding of the roles of genetics, the environment, and their interaction in the evolution of life history traits (see reviews on genomics (Orsini *et al*. 2013), transcriptomics (Alvarez *et al*. 2015), and epigenomics (Metzger & Schulte 2016)). Studies based on -omics approaches have been conducted on many diadromous species (see section "2.4 -Omics studies in diadromous fishes" below), although -omics studies analyzing diadromy as a common trait in taxa across the phylogeny of fishes are lacking.

Although migratory behavior is present in all major animal taxa from invertebrates to mammals (Merlin & Liedvogel 2019), the genetic bases and the evolutionary consequences of migratory behavior are not fully understood. The combination of new sequencing technologies and -omics approaches is key for the study of the evolution of life histories including the study of diadromy. Here, we aim to review our knowledge on diadromy across the phylogeny of fishes and how -omics techniques are helping answer questions about the ecology and evolution of diadromous species. We

organized this essay into three sections: 1) What is known about diadromy in fishes? Here, we describe the classification, the distribution of diadromy from a phylogenetic and biogeographic perspective, the main hypotheses proposed to explain its origin, and the presence of non-migratory populations in diadromous species. 2) The contribution of -omics research to our understanding of diadromy. In this section, we discuss how research in genomics, transcriptomics, and epigenomics is providing information about the life history of diadromous species, the facultative nature of this migratory behavior, the molecular bases underpinning this trait, and the origin of diadromy. 3) What questions could future research focus on? In this section, we elaborate on broad questions that future -omics research can help address regarding the evolution, genetic mechanisms, and maintenance of diadromy in fishes.

2.3 Diadromy in fishes – what is known?

The term "diadromy" was first introduced by Myers to describe "truly" migratory fishes (Myers 1949), with "truly" referring to the movement between marine and freshwater environments (McDowall 1993). McDowall (1997) expanded the definition and proposed specific characteristics that all diadromous species must fulfill. These include: migration must be mediated through physiological changes, it must occur at predicted times, and it should involve reciprocal migrations (McDowall 1997). While most diadromous species are known to be euryhaline, some are amphihaline, meaning that they can only adapt to a different salinity at a particular life stage (McDowall *et al.* 2009).

Diadromous species include fishes, gastropod mollusks (family Neritidae, see Abdou *et al.* 2015), and crustaceans (families Atyidae and Palaemonidae) (McDowall 1997); however, in this review, we focus on fishes. More than 440 fish species have been reported to be diadromous (Appendix 1). These species are distributed among 58 of the 482 recognized families of fishes (Nelson 1994); however, almost 62% of all diadromous fishes are concentrated in only seven families (Table 2.1).

Table 2.1 Total number of known diadromous species reported by family and category. The complete list of diadromous species is in Appendix 1.

Order	Family	Diadromous	Anadromous	Catadromous	Amphidromous
Acipenseriformes	Acipenseridae **(R)**	18	18		
	Ambassidae	4		1*	3*
Anguilliformes	Anguillidae **(R)**	16		16	
	Muraenidae	1		1	
	Ophichthidae	1	1*		
Atheriniformes	Atherinidae	1			1*
	Atherinopsidae	2	2*		
Characiformes	Citharinidae	2	2*		
Clupeiformes	Clupeidae **(R)**	31	26	2	3
	Engraulidae **(R)**	11	5	1	5*
	Pristigasteridae	7	4*		3*
Cypriniformes	Cyprinidae **(R)**	6	6		
Elopiformes	Elopidae	1	1*		
	Megalopidae	1		1*	
Gadiformes	Gadidae	1	1		
	Lotidae **(R)**	1	1*		
Galaxiiformes	Galaxiidae **(R)**	11	1		10
Gobiesoformes	Gobiesocidae **(R)**	1			1
Gobiiformes	Eleotridae **(R)**	37		5*	32
	Gobiidae **(R)**	103	2		101
	Lutjanidae	2		2*	
	Moronidae **(R)**	2	2		
Mugiliformes	Mugilidae **(R)**	34	1*	27	6*

Order	Family	Diadromous	Anadromous	Catadromous	Amphidromous
Osmeriformes	Osmeridae **(R)**	10	10		
	Plecoglossidae	1			1
	Retropinnidae	5	1*		4
	Salangidae **(R)**	6	6		
Perciformes	Carangidae	2			2*
	Centropomidae	9		2	7*
	Cheimarrichthyidae	1			1
	Cottidae	8		2	6
	Gasterosteidae **(R)**	2	2		
	Gerreidae	7			7*
	Haemulidae	1			1*
	Kuhliidae	10		5	5
	Lateolabracidae	1		1	
	Latidae	1		1	
	Percichthyidae	1		1	
	Percidae **(R)**	1	1		
	Pseudaphritidae	1		1	
	Rhyacichthyidae	2			2
	Scianidae	3			3*
	Terapontidae	1		1*	
	Tetrarogidae	1		1	
	Toxotidae	3			3*
Petromyzontiformes	Geotriidae	1	1		
	Mordaciidae	2	2		
	Petromyzontidae **(R)**	8	8		
Pleuronectiformes	Pleuronectidae	2		2	
Salmoniformes	Salmonidae **(R)**	35	35		

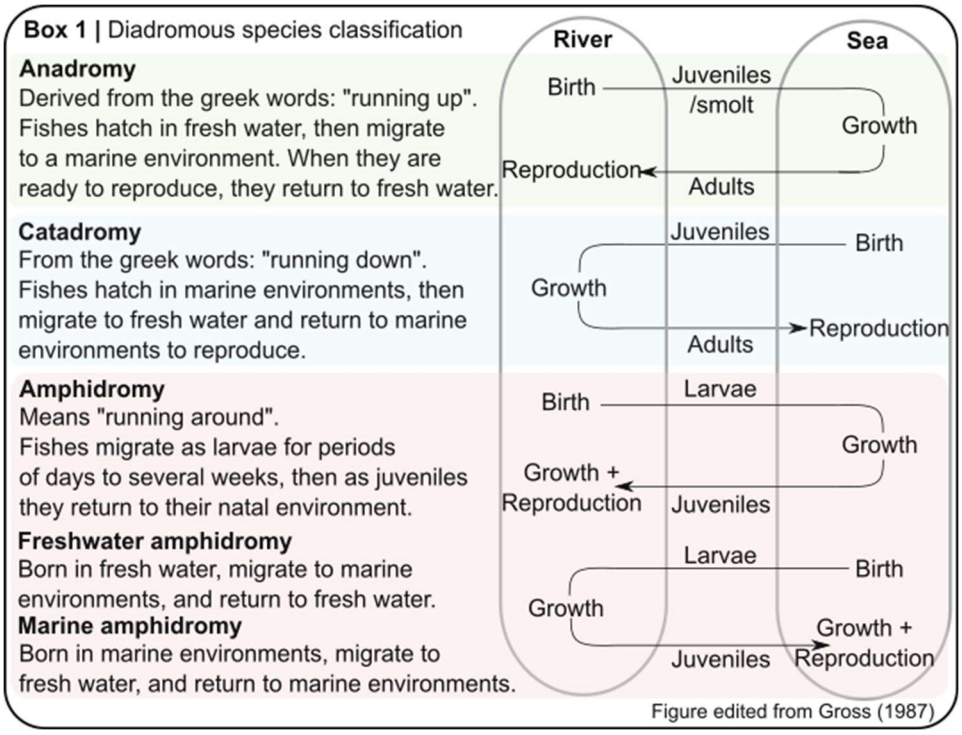
Order	Family	Diadromous	Anadromous	Catadromous	Amphidromous
Siluriformes	Ariidae **(R)**	13	3		10*
	Bagridae	1	1*		
	Claroteidae	1	1*		
	Pangasiidae	1	1		
	Plotosidae	1			1*
	Schilbeidae	1			1
Syngnathiformes	Syngnathidae	5			5
Tetraodontiformes	Tetraodontidae **(R)**	2	2		
Total		**444**	**147**	**73**	**224**

(R) report of resident populations

***** little information available (e.g., no microchemistry analysis)

2.3.1 Classification

There are three categories of diadromy (Gross 1987; McDowall 1997); fishes can be anadromous, catadromous, or amphidromous (Box 1). Categories differ in the direction of the first migration, from rivers to the sea (i.e., anadromy and freshwater amphidromy) or vice-versa (i.e., catadromy, marine amphidromy), the time of migration, particularly the life cycle stage when individuals return to their natal environment (i.e., juveniles vs. adults), and the purpose of the return migration, i.e., if the return to their natal environment is for growing and/or spawning purposes.



Of the 444 species reported here as diadromous, 147 are described as anadromous, 73 as catadromous, and 224 as amphidromous (Table 2.1), with almost all amphidromous species reported as freshwater amphidromous. Appendix 1 lists species that followed McDowall's definition of diadromy and that are referenced in a scientific paper or book, however, for many species particularly amphidromous species, little information is available (Table 2.1 & Appendix 1). Although Riede (2004) reported a higher number of amphidromous species than in this paper, this was due to the use of a broader definition of amphidromy (Chalant *et al.* 2019). We can, however, expect the total number of diadromous (mainly amphidromous) species to increase, with the increased focus on understudied species, for instance, from the tropics.

Anadromy is the most phylogenetically widespread category, present in 29 families, but only two families (Salmonidae and Clupeidae) comprise 41% of all anadromous species. Catadromy is present in 19 families. However, two families (Mugilidae and Anguillidae) represent almost 59% of all catadromous species. Amphidromy holds the highest number of diadromous species and is present in 26 families, two of which (Elotridae and Gobiidae) include 62% of all amphidromous species. Although most species from the same order belong to the same category of diadromy, different categories can be present within an order or even within a family. Four orders (Clupeiformes, Gobiiformes, Mugiliformes, and Perciformes) have species of all three categories (Table 2.1 & Fig. 2.1).

2.3.2 Phylogeny and biogeography

Diadromy is present from agnathans to the most recent bony fishes, indicating that it is an evolutionarily successful strategy (McDowall 1993). Despite the high cost of migration, which includes genetic, morphological, physiological, and behavioral requirements, diadromy is likely to have evolved multiple times (McDowall 1997; Corush 2019). The most recent Actinopterygii phylogeny developed from genomic and transcriptomic data (Hughes *et al.* 2018), confirms that diadromy is widespread across the fish phylogeny (Fig. 2.1).

Anadromy is more phylogenetically widespread than catadromy or amphidromy. While anadromy is found from lampreys, a lineage that appeared before the Actinopterygii to the most recent order, catadromy is present from the Anguilliformes order, a lineage that appeared during the Jurassic Period when marine species reappeared (Fyhn *et al.* 1999), to more recent families (i.e., Lutjanidae). Amphidromous species, on the other hand, are present from the order Clupeiformes to the order Perciformes (Fig 2.1).

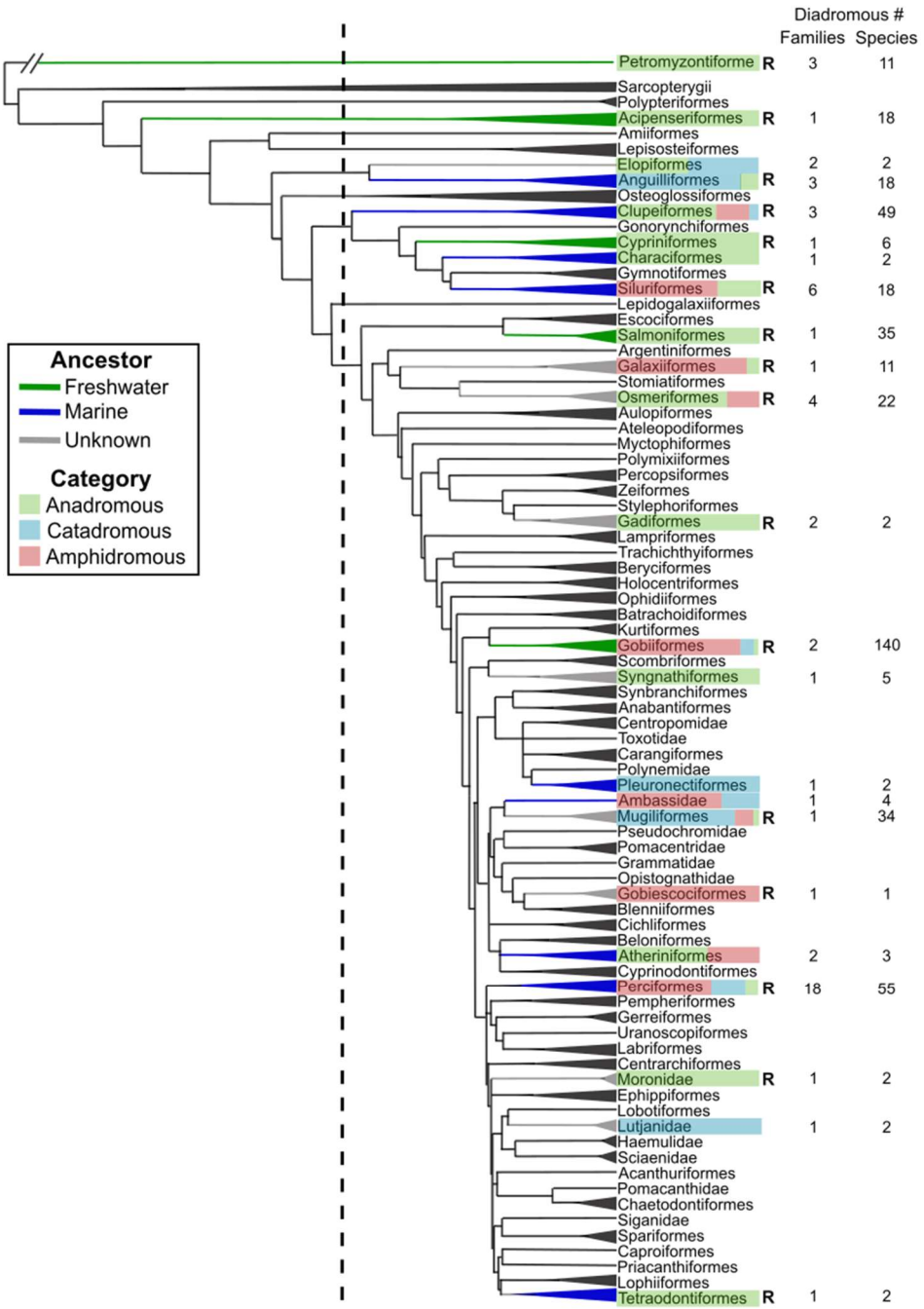


Figure 2.1 Actinopterygians phylogeny adapted from Hughes *et al.* (2018). The colors in each terminal branch have been added to indicate the most likely origin of the ancestor for diadromous taxa (see Table S2.3 for references). Taxa exhibiting diadromy have colored background labels reflecting the category of diadromy (anadromy, catadromy, or amphidromy). Some taxa exhibit more than one form of diadromy, and their proportion is indicated by the different colors in the label backgrounds. The dotted line represents the beginning of the Jurassic Period. The R to the right state the presence of resident populations and the numbers the number of known diadromous families and species for each taxon.

From a biogeographic point of view, diadromy is widely distributed across the globe. Gross (1987) described a latitudinal shift where anadromous species are prevalent at relatively high latitudes, while catadromous species have a relatively high occurrence in the tropics. This pattern led to the productivity hypothesis (see "Hypotheses on the origin" section). Amphidromy also appears to be found predominantly in the tropics (McDowall 2010). However, though categories have higher incidence at certain latitudes, diadromous species of all three categories can be found at both high and low latitudes, providing evidence against the productivity hypothesis (Fig. 2.2). The widespread extent of diadromous species supports its important role in species dispersal, including transoceanic dispersal (Chubb *et al.* 1998; McDowall 1998).

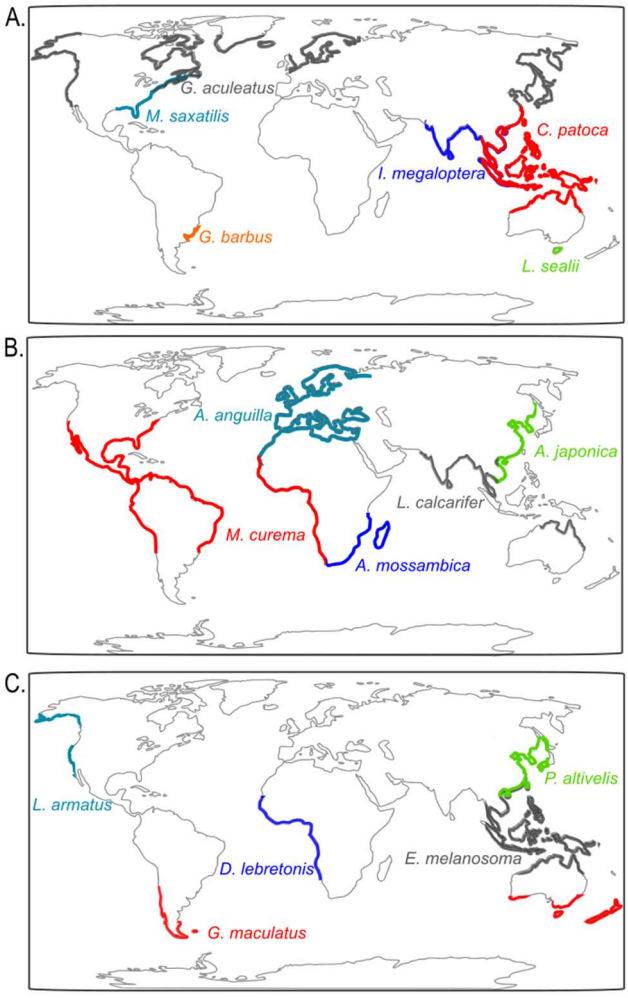


Figure 2.2 World map depicting the distribution of 5- 6 diadromous species present at different latitudes. **A.** Anadromous species. **B.** Catadromous species. **C.** Amphidromous species.

2.3.3 Hypotheses on the origin

While phylogenetic data support the hypothesis that diadromy has evolved multiple times (Fig. 2.1), the origin of diadromy and its evolutionary bases are still under debate (Bloom & Lovejoy 2014). Below, we summarize the main hypotheses for the origin of diadromy.

Productivity or resource availability hypothesis: This is the most accepted hypothesis first proposed by Gross (1987). It is based on the findings of latitudinal trends for anadromous and catadromous species by Baker (1978) and Northcote (1978) and states that anadromous species are more prevalent at high latitudes because productivity in the sea at those latitudes is higher than in freshwater environments. Catadromy instead, is relatively common at tropical latitudes, given that the productivity of fresh water at low latitudes is higher than in the sea.

The presence of all categories of diadromy at different latitudes questions the generality of this hypothesis (McDowall 2008b; Fig. 2.2), for example, temperate eels migrate to freshwater environments that have lower productivity than marine environments (Edeline 2007). Bloom and Lovejoy (2014) tested this hypothesis using the phylogeny of the order Clupeiformes. Their results did not support the productivity hypothesis as the ancestry of diadromy could not be predicted based on latitude (Bloom & Lovejoy 2014).

Historical processes including the expansion and invasion of newly available environments following post-Pleistocene deglaciation could explain the prevalence of anadromous species in northern temperate latitudes (McDowall 2008b). Temperature and particularly temperature fluctuations could also explain the incidence of diadromy. For example, at high latitudes temperature fluctuates more in fresh water than in the sea, a factor that has led to the hypothesis that the invasion of fresh water (i.e., catadromy) is more frequent at low latitudes (Lee & Bell 1999).

Random escapement hypothesis: This hypothesis, proposed by Tsukamoto *et al*. (2009), is based on behavioral models from observations of the amphidromous Ayu (*Plecoglossus altivelis*) and argues that diadromy originated as an escapement behavior of fishes to leave unfavorable environments, instead of as migration to a more nutrient-rich habitat (i.e., productivity hypothesis). A three-step model is used to explain the start

of migration. First, an individual needs to reach a threshold age or size; second, it needs to be physiologically prepared; and third, it needs to receive an endogenous or exogenous cue to initiate migration (Tsukamoto *et al.* 2009).

Similarly, a "safe-site" hypothesis was proposed, where migration to fresh water or a "safe haven" is a consequence of the need to protect early life history stages from marine predators. An example is the early larval migration of osmeroids, which has been hypothesized to maximize their survival, and therefore their fitness, due to the presence of safe sites (Dodson *et al.* 2009).

Ecological opportunity hypothesis: Proposed by Feutry *et al*. (2013), this hypothesis states that diadromy appears as a response to ecological opportunities (Feutry *et al.* 2013). Using the case of the Kuhlia family, within which catadromous species migrate to nutrient-poor environments, the authors proposed that diadromy originated due to the opportunity to colonize insular ecosystems. These isolated habitats would be characterized, for example, by an absence of predators, making them ideal for colonization. This idea to move to an ecological advantageous site is similar to the "safe-site" hypothesis.

Intermediate state hypothesis: This hypothesis states that diadromous fishes have appeared as an intermediate state between fully freshwater and fully marine fishes (Gross 1987). Gross (1987) also proposed that amphidromy is the ancestral state of both anadromy and catadromy. The hypothesis suggested that anadromous species evolved from amphidromous species, which evolved from euryhaline wanderers that evolved from freshwater species. While catadromy evolved from amphidromous species which originated from euryhaline wanderers that evolved from marine species (Gross 1987). The improvement in our understanding regarding the biology of amphidromy, specifically, the short time (i.e. days or weeks) amphidromous species spend in the secondary environment, led to the rejection of the idea that amphidromy was an intermediate step between fully freshwater and marines fishes (Gross 1997).

Recently, Corush (2019) tested this hypothesis by simulating the rate of transitions in and out of diadromy and comparing it between freshwater, marine, and diadromous fishes. Transition rates out of diadromy into strictly marine or strictly freshwater life histories were higher than transitions in the opposite directions (from

marine or freshwater life histories into diadromy), leading to the conclusion that diadromy may sometimes be an intermediate state between freshwater and marine fishes, but not always (Corush 2019).

Conditional evolutionary stable strategy model (CESSM): This model proposed that diadromy is a phenotypically plastic trait in which an individual expresses a migratory phenotype depending on environmental variables and will migrate if this migration leads to higher fitness (Edeline 2007). This model is supported by the fact that the migration in diadromous species reduces inter- and intra-specific competition (Edeline 2007). Although the CESSM model may apply for some species, the presence of sympatric migratory and non-migratory populations that show high levels of genetic differentiation (e.g., Salisbury *et al*. 2018; Delgado *et al.* 2019) suggests that in such cases the decision to migrate does not just depend on environmental variables.

2.3.4 Loss of diadromy

Major reductions in dispersal ability have evolved many times across numerous taxa (Waters *et al.* 2020). Non-migratory (hereafter called resident) populations exist among all three categories of diadromous fish species (Table 2.1). Many resident populations within a species have evolved multiple independent times and derived from a common diadromous ancestor, as the case of the anadromous *Alosa pseudoharengus* in Connecticut, USA (Palkovacs *et al.* 2008) or the amphidromous *Galaxias maculatus* in Chile (Chapter 3; Delgado *et al.* 2019). Resident populations need not be landlocked; they can inhabit environments with access to the sea despite which they do not migrate but remain in their natal habitat. Examples of resident populations have been described for all categories of diadromy as follows:

Anadromous species: Atlantic Salmon (*S. salar*), Brown Trout (*S. trutta*), Brook Trout (*S. fontinalis*), Arctic Charr (*S. alpinus*), White-spotted Charr (*S. leucomaensis*), Dolly Varden (*S. malma*), Rainbow Trout (*O. mykiss*), Masu Salmon (*O. masau*), Coastal Cutthroat Trout (*O. clarki*), Sockeye Salmon (*O. nerka*), Chinook Salmon (*O. tshawytscha*) (e.g., Dodson *et al.* 2013), Three-spined Stickleback (*G. aculeatus*) (e.g., Bell & Foster 1994), and Japanese Smelt (*H. nipponensis*) (e.g., Arai *et al*. 2006b).

Catadromous species: European Eel (*A. Anguilla*) (e.g., Arai *et al.* 2006a), Japanese Eel (*A. japonica*) (e.g., Tsukamoto & Arai 2001), American Eel (*A. rostrata*) (e.g., Lamson *et al*., 2009), and Tupong (*P. urvillii*) (e.g., Crook *et al.* 2010).

Amphidromous species: Common Galaxias (*G. maculatus*) (e.g., Delgado *et al.* 2019), Spotted Galaxias (*G. truttaceus*) (e.g., Waters *et al.* 2001), Big-scaled Redfin (*T. hakonensis*) (e.g., Sakai *et al*. 2002), New Zealand Eleotrid (*G. cotidianus*) (e.g., Michel *et al.* 2008), and *Rhinogobius* sp. (Tsunagawa *et al.* 2010).

The existence of resident populations that can migrate (i.e., inhabiting environments with access to the sea), but do not, suggests that migration may not always be beneficial and that ecological factors likely play an important role in the decision to migrate. Facultative diadromy demonstrates that there is a balance between the benefits and costs of migration and residency tactics (Hogan *et al.* 2014). Ferguson *et al.* (2019) introduced the Threshold-trait model to explain which factors determine or affect the decision of whether or not to migrate. This model involves two components: a genetic and environmental threshold. Individuals remain resident if the energy status is high and exceeds a given threshold. Alternatively, individuals migrate if their energy status is low (i.e., nutritionally deficient).

Gross (1987) proposed a model to explain diadromy, which is simplified in Figure 2.3. In basic terms, the fitness of migrating adding the cost of migration must be higher than the fitness acquired by remaining in the natal environment. Following Gross hypothesis and Ferguson's threshold model, we propose to explain the loss of diadromy from an ecological and evolutionary perspective following one or the combination of 4 scenarios in nature (Fig. 2.3):

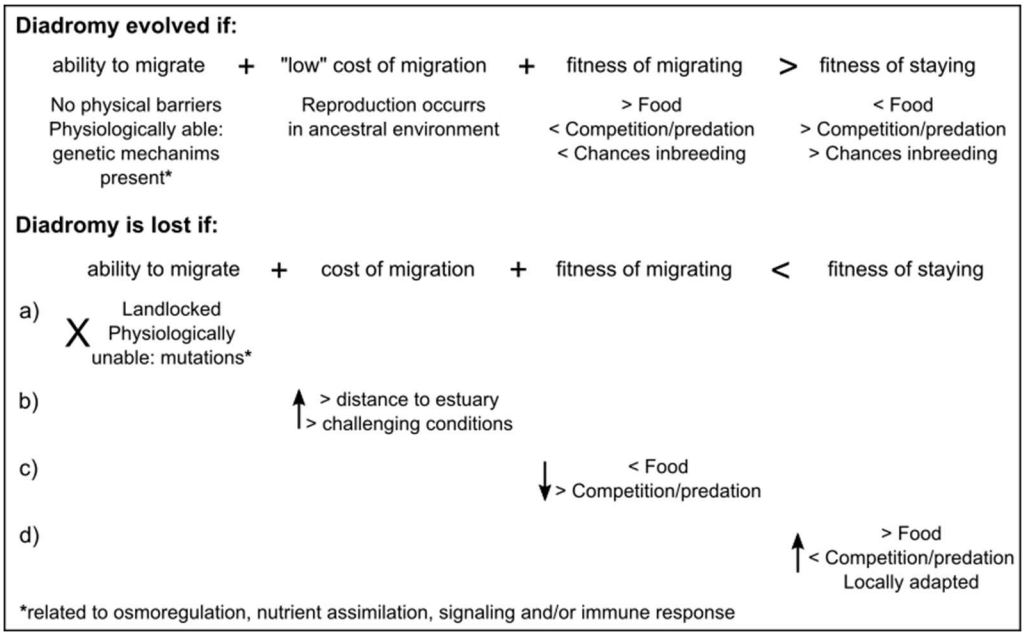


Figure 2.3 A simplified model to explain the "decision" to migrate (Diadromy) or remain in the natal environment (Loss of diadromy) from an Eco-Evo perspective.

No alternative - Fig. 2.3a: There are two possible explanations for the loss of diadromy. Either a population becomes physically landlocked unable to leave its natal environment, or the presence of mutations makes individuals physiologically unable to migrate (i.e., the genetic component of the Threshold-trait model). Mutations that would affect the ability to migrate in diadromous fishes may be linked to osmoregulatory genes, nutrient assimilation genes, and signaling (See "Molecular bases" section below).

Increase in the cost of migration - Fig. 2.3b: This could be due to changes in the river systems; for example, a change in the landscape could lead to a longer distance to reach the estuary increasing the cost of migration. Natural or anthropogenic barriers such as dams would also increase the cost of migration.

Decrease in the fitness of migration - Fig. 2.3c: would be the product of changes in the environmental conditions of the secondary habitat, for example, reduction of food supply, an increase in the number of predators, or an increase in parasite threats. These changes would lead to a reduction in fitness and even the survival of the migrating phenotype.

Increase in the fitness of staying - Fig. 2.3d: Changes in the environmental conditions in the natal habitat that would lead to a higher fitness payoff for a decision to

stay (i.e., the environmental component of the Threshold-trait model). This could result from a decrease in the fitness value of the secondary environment (Fig. 2.3c above) or positive changes in the natal habitat. Examples of the latter are the increment in the quality and quantity of resources and a decrease in competition.

2.4 -Omics studies in diadromous fishes – what current research tells us?

Since the introduction of next-generation sequencing (NGS) and the drop in sequencing cost, the number of studies using DNA, RNA, and methylation to address ecological and evolutionary dynamics questions has increased. Here, we present a list of studies on diadromous species that used -omics techniques, chiefly genomics, transcriptomics, and epigenomics (Table 2.2). This list is based on published articles and excludes books, theses, articles in bioRxiv, and conference abstracts.

Table 2.2 Representation of research in diadromous fishes that have used an -omics approach.

Diadromy category	-omics approach (methods)	Research objective	Species studied	Include residents	References examples
Anadromy	Genomics (GBS, RADseq, ddRADseq, RADcap, SNP array, NextRAD, Pool-seq, whole-genome sequencing)	Development of molecular markers / SNP panels	*Salmo salar; Alosa pseudoharengus; Alosa aestivalis*	No	Houston *et al.* 2014; Yáñez *et al.* 2016; Baetscher *et al.* 2017
		Applicability of SNP array from close-related species	*Salmo trutta*	No	Drywa *et al.* 2013
		Assembly of a reference genome	*Salmo salar; Oncorhynchus tshawytscha; Pungitius pungitius*	No	Davidson *et al.* 2010; Christensen *et al.* 2018; Varadharajan *et al.* 2019
			Salmo salar	Yes	Hauge *et al.* 2016
		Assembly of mitochondrial genome	*Takifugu obscurus; Lethenteron camtschaticum; Coilia nasus*	No, Yes, No	Kim *et al.* 2014; Balakirev *et al.* 2016; Zhang *et al.* 2016
		Development of linkage map/chromosome rearrangements	*Salmonidae*	No	Sutherland *et al.* 2016
		Population diversity and structure	*Leuciscus idus; Oncorhynchus mykiss; Salvelinus fontinalis; Salmo trutta*	Yes	Skovrind *et al.* 2016; Leitwein *et al.* 2017; Elias *et al.* 2018; Lemopoulos *et al.* 2018
			Salmo salar; Salvelinus alpinus; Brachymystax lenok; Tenualosa ilisha; Alosa pseudoharengus; Alosa aestivalis	No	Asaduzzaman *et al.* 2019; Aykanat *et al.* 2015; Madsen *et al.* 2020; Moore *et al.* 2017, 2014; Reid *et al.* 2018; Roman *et al*, 2018
		Genomic divergence/local adaptation	*Thaleichthys pacificus; Oncorhynchus tshawytscha*	No	Candy *et al.* 2015; Narum *et al.* 2018
			Coregonus clupeaformis; Salmo salar; Entosphenus tridentatus; Lampetra fluviatilis; Oncorhynchus nerka; Lampetra planeri; Salvelinus alpinus	Yes	Bourret *et al*. 2013; Hume *et al*. 2018; Mateus *et al*. 2013; Nichols *et al*. 2016; O'Malley *et al*. 2019; Parker *et al.* 2019; Renaut *et al*. 2011; Rougemont *et al.* 2017; Salisbury *et al.* 2020; Veale and Russello, 2017
			Gasterosteus aculeatus	Yes	Hohenlohe *et al.* 2010; Jones *et al.* 2012; Guo *et al.* 2015; Ferchaud & Hansen 2016; Currey *et al.* 2019; Dean *et al.* 2019; Marques *et al.* 2019; Rennison *et al.* 2019; Terekhanova *et al.* 2019

			Oncorhynchus mykiss	Yes	Hale *et al.* 2013; Hecht *et al.* 2013; Pearse *et al.* 2014; Bowersox *et al.* 2016; Matala *et al.* 2017; Arostegui *et al.* 2019
			Salmo salar	Yes	Bourret *et al.* 2011; Culling 2013; Perrier *et al.* 2013
		Introgression/hybridization	*Gasterosteus aculeatus; Leucopsarion petersii; Alosa pseudoharengus; Pungitius pungitius*	Yes	Guo *et al.* 2019; Hirase *et al.* 2020; Reid *et al.* 2020; Yoshida *et al.* 2016
		Genotype - migration associations	*Salmo salar; Oncorhynchus tshawytscha; Oncorhynchus mykiss; Salmo trutta*	No	Brieuc *et al.* 2015; Cauwelier *et al.* 2018; Johnston *et al.* 2014; Lemopoulos *et al.* 2018a; Micheletti *et al.* 2018a; Prince *et al.* 2017; Thompson *et al.* 2020
		Genotype - sex - migration associations	*Oncorhynchus mykiss*	Yes	Kelson *et al.* 2019
		Genotype - environment associations	*Oncorhynchus mykiss; Gasterosteus aculeatus*	Yes	Micheletti *et al.* 2017; Stuart *et al.* 2017; Haenel *et al.* 2019
			Salmo salar; Oncorhynchus mykiss	No	Jeffery *et al.* 2017; Willoughby *et al.* 2018
		Genotype - microbiota associations	*Gasterosteus aculeatus*	Yes	Steury *et al.* 2019
		Sex determination	*Gasterosteus aculeatus*	No	Bissegger *et al.* 2019
	Transcriptomics (RNA-seq, cDNA arrays, microarrays)	Assembly of transcriptomic profiles	*Oncorhynchus mykiss; Salmo salar; Salmo trutta, Salvelinus alpinus, Coregonus lavaretus; Oncorhynchus tshawytscha*	Yes	Salem *et al.* 2015; Carruthers *et al.* 2018; Christensen *et al.* 2018
		Detection of lncRNAs	*Oncorhynchus mykiss*	No	Al-Tobasei *et al.* 2016
		Expression profiles of spermatogenesis	*Coilia nasus*	No	Zhou *et al.* 2015
		Expression profiles before migration	*Salvelinus fontinalis*	Yes	Boulet *et al.* 2012
		Expression profiles of infection response	*Gasterosteus aculeatus; Oncorhynchus mykiss; Salmo salar & Salmo trutta*	Yes	Lenz *et al.* 2013; Sutherland *et al.* 2014
		Expression profiles hatchery vs. wild	*Salmo salar; Oncorhynchus mykiss*	No	Bicskei *et al.* 2014; Fox *et al.* 2014
		Expression profiles juvenile brains	*Oncorhynchus mykiss*	Yes	Hale *et al.* 2016
		Expression profile hybrids	*Salvelinus fontinalis*	Yes	Mavarez *et al.* 2009

		Genomic population/divergence	*Coregonus clupeaformis; Alosa pseudoharengus; Oncorhynchus nerka; Salmo salar*	Yes; No	Jeukens *et al.* 2010; Czesny *et al.* 2012; Lemay *et al.* 2013; Warren *et al.* 2014
		Salinity adaptation	*Alosa pseudoharengus; Gasterosteus aculeatus; Salmo salar*	Yes	Lemmetyinen *et al.* 2013; Gibbons *et al.* 2017; Kusakabe *et al.* 2017; Rastorguev *et al.* 2017; Velotta *et al.* 2017
		Temperature adaptation	*Salmo trutta; Gasterosteus aculeatus*	Yes	Meier *et al.* 2014; Morris *et al.* 2014
		Migratory life history	*Salmo trutta; Salmo Salar*	Yes	Giger *et al.* 2008
		Environmental stress associations	*Takifugu obscurus*	No	Xu *et al.* 2018
		Freshwater colonization	*Gasterosteus aculeatus & non-diadromous sister species*	Yes	Kitano *et al.* 2018; Ishikawa *et al.* 2019
	Proteomics	Salinity adaptation	*Coregonus lavaretus*	Yes	Papakostas *et al.* 2012
	Epigenomics (RRBS, MSAP)	Genotype - environment associations/stressors	*Gasterosteus aculeatus; Salmon trutta*	No	Aniagu *et al.* 2008; Morán *et al.* 2013; Fellous & Shama 2019
		Genotype-phenotype associations	*Salmon trutta; Gasterosteus aculeatus*	No; Yes	Covelo-Soto *et al.* 2015; Smith *et al.* 2015
		Salinity adaptation	*Gasterosteus aculeatus*	Yes	Artemov *et al.* 2017
		Migration effects	*Oncorhynchus mykiss*	Yes	Baerwald *et al.* 2016
		Hatchery effects	*Oncorhynchus kisutch; Oncorhynchus mykiss*	No	Le Luyer *et al.* 2017; Gavery *et al.* 2018
Catadromy	Genomics (genome sequencing, RADseq, Pool-seq)	Molecular markers development	*Anguilla japonica*	No	Sekino *et al.* 2016
		Assembly of mitochondrial genome	*Trachidermus fasciatus*	No	Zhu *et al.* 2018
		Population structure	*Mugil cephalus; Anguilla japónica, Anguilla anguilla & Anguilla rostrata; Trachidermus fasciatus*	No	Krück *et al.* 2013; Igarashi *et al.* 2018; Gong *et al.* 2019; Li *et al.* 2019
		Adaptive divergence	*Anguilla rostrata; Cottus asper*	Yes	Pavey *et al.* 2015; Dennenmoser *et al.* 2017
		Hybridization	*Anguilla anguilla* & *Anguilla rostrata*	No	Pujolar *et al.* 2014; Nikolic *et al.* 2019
	Transcriptomics (RNA-seq, cDNA array)	Transcriptomic profiles	*Anguilla anguilla; Trachidermus fasciatus*	No	Churcher *et al.* 2015; Ma *et al.* 2018
		Genotype-phenotype associations	*Anguilla rostrata*	No	Côté *et al.* 2014
	Proteomics	Salinity acclimation	*Anguilla marmorata; Trachidermus fasciatus*	No	Jia *et al.* 2016; Ma *et al*. 2018

	Epigenomics (MSAP)	Methylation changes between life stages	*Anguilla anguilla*	No	Trautner *et al.* 2017
Amphidromy	Genomics (genome sequencing, RADcap)	Development of molecular markers	*Oncorhynchus clarki lewisi*	No	Campbell *et al.* 2012
		Assembly of mitochondrial genome	*Sicyopterus lagocephalus*	No	Chiang *et al.* 2015
		Phylogeny (mitogenome)	*Sicyopterus genus*	No	Lord et al., 2019
		Populations diversity and structure	*Galaxias maculatus*	Yes	Chapter 3; Delgado *et al.* 2019
		Salinity adaptation	*Galaxias maculatus*	Yes	Chapter 4; Delgado et al., 2020
	Transcriptomics (transcriptome sequencing)	Salinity adaptation	*Plecoglossus altivelis*	Yes	Lu *et al.* 2016

Anadromy is by far the category with the highest number of studies (Table 2.2). From the >120 papers listed, 84% concern anadromous species. Within anadromy, the most intensely studied species are Three-spined Stickleback (*Gasterosteus aculeatus*), Rainbow Trout (*Oncorhynchus mykiss*), and Atlantic Salmon (*Salmo salar*). Catadromy is the second category in terms of the number of studies; however, most research on catadromy has focused on species from the Anguilla family, particularly, the European Eel (*Anguilla anguilla*). Most genetic research on amphidromous species has thus far been based on a few mitochondrial and nuclear markers *(e.g. Crandall et al. 2010; Lord et al*. 2015; Taillebois *et al*. 2013) with studies within this group using -omics approaches being scarce (Table 2.2). These findings are not surprising as anadromous species are more prevalent in the Northern Hemisphere, where most research takes place, and they are of economic importance.

The type of -omics approach used is a function of the research objectives (Table 2.2). Box 2 lists the most common methods with Reduced Representation Sequencing (RRS) being the most widely used. RRS is a cost-effective method that provides a large but limited number of markers across the genome (Wright *et al.* 2019). In some cases, RRS proved to be more powerful than previous genetic markers (e.g., few microsatellites, few mitochondrial genes) in differentiating and assigning individuals to populations (e.g., Moore *et al.* 2014; Candy *et al.* 2015; Yoshida *et al*. 2016), yet this is not necessarily always the case as shown by a study based on a relatively larger number of sequenced microsatellite markers (Layton *et al.* 2020). An alternative to RSS is whole genome sequencing and this can be done by sequencing individuals or pools of individuals (Pool-seq) (Fuentes-Pardo & Ruzzante 2017). However, given its cost, it has been used mainly in economically important species.

Box 2 | Most common methods used in -omics studies about diadromous species.

A. Genomics	Pros/Cons
• RRS: RADseq, ddRAD, RADcap	+ "Thousands" of markers located genome-wide. - Chance of missing informative/ adaptive markers.
• Whole genome Individuals or pools (Pool-Seq)	+ Complete genomic information for the species. - Expensive.
B. Transcriptomics	Pros/Cons
• RRS: RNA-seq	+ Allows *de novo* assembly, great for non-model organims. - Bioinformatically challenging, eg: repetitive sequence, isoforms, etc.
• Microarrays	+ Species-specific arrays provide high reliability. - Require prior knowledge of the genome.
C. Epigenomics	Pros/Cons
• RRS: Bisulfite Sequencing (RRBS)	+ High resolution, single base. - Chance of missing methylation sites.
• Methylation Sensitive Amplified Polymorphism (MSAP)	+ No need for genomic information and inexpensive. - Low resolution.

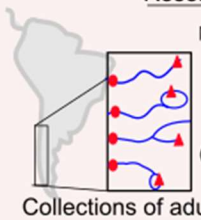
Genomic approaches allow the exploration of both neutral and adaptive markers, thus facilitating the examination of the genetic bases and mechanisms of adaptation (Orsini *et al.* 2013). Markers distributed across the genome have been used to estimate levels of genetic diversity (e.g., Bowersox *et al.* 2016; Gong *et al.* 2019), effective populations size (e.g., Li *et al.* 2019), bottlenecks (e.g., Ferchaud & Hansen 2016), and fine-scale populations structure or lack thereof (e.g., Mateus *et al.* 2013; Aykanat et al. 2015; Skovrind *et al.* 2016) in many diadromous species (Table 2.2). Outlier markers are used for the assessment of genetic divergence among populations (e.g., Box 3A). Estimating this divergence can provide information on the colonization history of diadromous species, the description of glacial lineages, and the effects of secondary contact (e.g., Bourret *et al*. 2013; Dean *et al.* 2019). Genome-wide studies have also revealed the molecular mechanisms (i.e., genes or islands of differentiation) supporting population divergence (e.g., Larson *et al.,* 2016). Genetic markers differentiating phenotypes and populations can be used to assist in the assessment of relevant traits for fisheries/production (e.g., Yáñez *et al*. 2016), in the assignment of regional fisheries stocks (e.g., Baetscher *et al.* 2017), in the detection of introgression (e.g., Bourret *et al.*

2011), in the assessment of hybridization due to secondary contact (e.g., Reid *et al.* 2020), and in the detection of anthropogenic effects (e.g., Leitwein *et al.* 2017); this type of genotype-phenotype association studies have also given insights on which genes (i.e., loci under selection) underpin migratory behavior (e.g., Micheletti *et al*. 2018a).

Box 3| Examples of -omics reasearch in diadromous species.

A. Genomics.- amphidromous *G. maculatus* (Delgado *et al.* 2019)

Research design



Collections of adults from migratory and non-migratory populations pairs (same river)

RADcap: ~4K neutral and adaptive SNPs

Population diversity and divergence

Main findings

Effects
- High gene flow, one genetically distinct diadromous population.
- Non-migratory populations appeared as independent colonization events.
- Cases of sympatry between migratory and non-migratory pop showed little hybridization.

Mechanisms
- Outlier loci that differentiate migratory and non-migratory populations linked to genes involved in processes like: osmoregulation, muscle development, etc.

B. Transcriptomics.- catadromous *A. anguilla* (Côté *et al.* 2014)

Research design

Collection of indviduals arriving to the river, from 2 genetically panmintic locations

River 1 (GRB) River 2 (MR)

Fresh water | Saltwater

Microarrays

Differential Gene Expression

Main findings

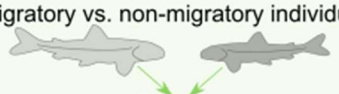
Origin
- Different genes expressed according to the location of origin.
- Confirms the important role of phenotypic plasticity in diadromous species.

Mechanisms
- 883 transcripts showed differential gene expression.
- Different genes involved in process such as: energy metabolism, cell protection/ immunity, growth, membrane transporters, among others.

C. Epigenomics.- anadromous *O. mykiss* (Baerwald *et al.* 2016)

Research design

Common garden experiment of migratory vs. non-migratory individuals



F2 cross

RRBS of fin tissue of smolts and residents

Methylation patterns

Main findings

Mechanisms
- 57 regions showed differential methylation, 62% showed differences between migratory and non-migratory individuals.
- More than half of these regions were located in regulatory regions.
- Examples of genes expressed more in smolts than residents are: AHR2A, known to mediate circadian rhythm; and NHCH2, known to modulate body weight and energy availability.

The facultative nature of diadromy in many species suggests that variation between migratory and resident behavior could happen at a transcriptional (i.e., transcriptome) or post-transcriptional (i.e., epigenome) level, thus the molecular variation would be found in the expression of genes rather than genes themselves. Transcriptomic studies in diadromous species have been used to examine differences in gene expression between populations, revealing critical genes involved in, for example, spermatogenesis (Zhou *et al.* 2015), but also involved in processes related to migratory behavior including osmoregulation (Velotta *et al.* 2017), signaling or sensory perception (how fishes process light – Hale et al., 2016), nutrient assimilation (Ishikawa *et al.* 2019), immune response (Lenz *et al.* 2013), and growth (Box 3B). These studies have also highlighted the importance of regulatory regions and their effects (e.g., Czesny *et al*. 2012). Furthermore, transcriptomics has been used to study salinity and temperature can affect the expression of genes differentially across populations (e.g., Meier *et al.* 2014 and Côté *et al.* 2014: Box 3B). This information is likely to be useful for predicting responses to changes in environmental conditions.

Phenotypic variance in migratory traits can be a product of genetics but may also be solely due to phenotypic plasticity as a response to environmental triggers, suggesting the important role of epigenetics (Merlin & Liedvogel 2019). Epigenetics focuses on modifications of genetic material due to environmental factors (Merlin & Liedvogel 2019). Research in this area is more recent and, therefore, less developed than genomics or transcriptomics. However, research has been conducted in some diadromous species to examine the effects of migration in reared individuals providing clues about the molecular mechanisms that distinguish migratory and resident populations. For instance, the primary location of methylation modifications that distinguish diadromous from resident populations varies in diadromous species, while these modifications are found predominantly within genes in Three-spined Sticklebacks (Smith *et al.* 2015) in Rainbow Trout they are found largely in regulatory regions (Baerwald *et al.* 2016 -Box 3C).

2.4.1 Life history

Genomic analyses have been conducted to assess population structure, and contribute to our understanding of migration in numerous diadromous species including the anadromous Hilsa Shad (*Tenualosa ilisha*) (Asaduzzaman *et al.* 2019), the

catadromous Japanese Eel (*Anguilla japonica*) (Igarashi *et al.* 2018), and the amphidromous Common Galaxias (*Galaxias maculatus*) (Box 3A). For example, while site fidelity led to genetic differentiation in Hilsa Shad (Asaduzzaman *et al.* 2019) and Japanese Eel (Igarashi *et al.* 2018) populations, the presumed absence of site fidelity perhaps combined with a relatively large effective population size in Common Galaxias resulted in a panmictic or nearly panmictic migratory systems among diadromous collections (Chapter 3; Delgado *et al.* 2019).

These genomic results regarding gene flow can also contribute to corroborate the classification of diadromous species. Common Galaxias, for instance, has been classified as a marginal catadromous species as it was hypothesized that larvae only migrate to the estuaries (without reaching the ocean) and after a few weeks migrate back up the river streams (McDowall 2009). The fact that populations across their Chilean distribution showed high levels of gene flow supports the amphidromous nature of the Common Galaxias and is consistent with the hypothesis that larvae do indeed enter the ocean (Chapter 3; Delgado *et al.* 2019).

The analyses of genomic markers have also provided information on species dispersal and reproductive behavior beyond that obtained through other methods such as telemetry. For example, a telemetry study in Arctic Charr showed high dispersal levels, yet this high dispersal did not lead to high gene flow (Moore *et al.* 2017). This result is consistent with the notion that Arctic Charr, an anadromous species, overwinters in non-natal freshwater environments in years when they do not reproduce (Jørgensen & Johnsen 2014). Additionally, information on successful dispersal contributes to our understanding of the balance regarding the costs and benefits of this complex life history trait.

2.4.2 Facultative behavior

While some diadromous species like Atlantic Salmon (*Salmo salar*) and Arctic Cisco (*Coregonus autumnalis*) are considered obligatory diadromous, others like the goby *Awaous stamineus* do not need to visit the marine environment to complete their life cycle (Hogan *et al.* 2014). The importance of migration for an individual's development thus appears to vary depending on the species or family. The presence of sympatric diadromous and resident populations with little genetic differentiation (e.g., Rainbow Trout, Kendall *et al.* 2014) suggests that migrating is not a requisite for the development

of individuals and that diadromy can in some groups be facultative. From an ecological perspective, facultative diadromy may be beneficial as the decision to migrate would depend on environmental pressures.

In some species, resident populations are clearly genetically divergent from their diadromous counterparts (e.g., Common Galaxias; Delgado *et al.* 2019). In Arctic Charr, sympatric anadromous and resident populations were until recently considered genetically indistinguishable, yet recent genomic data revealed genetically differentiated sympatric anadromous and resident populations in Labrador (Salisbury *et al.* 2019, 2020). Sockeye Salmon (*O. nerka*) on the other hand, exhibits examples of both genetically differentiated and non-differentiated sympatric resident and anadromous populations (Nichols *et al.* 2016). These differences in genetic differentiation between diadromous and resident populations are likely the product of local adaptation and/or genetic drift, implying that the ability to migrate can be maintained or lost as a result of selection or random processes.

Common garden studies have been conducted to test individual fitness in different environmental conditions, and the use of reaction norms has helped determine that diadromous species evolved a plastic response to different environmental variables. Examples are the studies on the response to varying salinities in Three-spined Stickleback (McCairns & Bernatchez 2010) and the response to different temperatures in Brown Trout (Meier *et al.* 2014). Both studies showed that diadromous individuals exhibit higher fitness when reared under a variety of environmental conditions than do resident individuals and are thus more plastic than resident individuals. The plastic nature of diadromy has also been demonstrated in Steelhead, where diadromous or resident parents can express alternative offspring (e.g., Zimmerman *et al.* 2009).

Reciprocal transplant experiments with Common Galaxias under laboratory conditions also revealed that resident populations can differ in their response to salinity changes with some populations maintaining their osmoregulatory performance necessary for migration and others not being able to survive such changes (Chapter 4; Delgado *et al.* 2020). Given that these populations were similarly genetically differentiated from their diadromous counterpart, we suspect genetic drift may be one of the factors playing an important role in determining whether the ability to migrate is maintained or lost. Yet,

other forms of relaxed selection could also be playing a role in the differences in persistence and loss of salinity adaptation found between these two populations (Lahti *et al.* 2009). The loss of osmoregulatory capacity can also be partial as revealed by gene expression studies with Alewives (*Alosa pseudoharengus*) resident populations (Velotta *et al.* 2014).

2.4.3 Molecular bases

The physiological adaptations required to survive in both marine and freshwater habitats are extensive (Fig. 2.4). Freshwater species, for example, rely on adaptations related to ions uptake (specific type mitochondria-rich cells, Bartels *et al*. 2017) and visual pigmentations (specific types of chromophores, Toyama *et al*. 2008). And although many genes described as relevant to marine-freshwater adaptations have been reported in diadromous species (Table S2.2), the genetic variation responsible for this migratory behavior is still unknown. This raises the question of whether there is one "diadromous" gene or genes (i.e., islands of differentiation)?

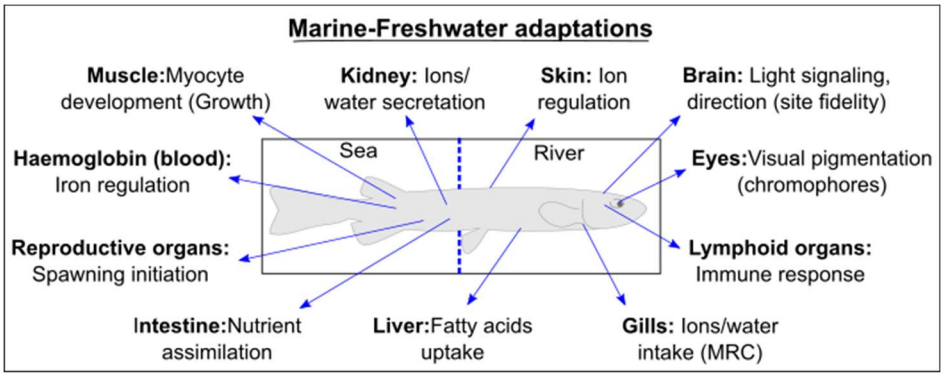


Figure 2.4. Examples of physiological adaptations necessary to survive in marine and freshwater environments.

Research on stickleback sister species provides a clear example of a gene essential for the ability to colonize fresh water: Fads2 (Ishikawa *et al.* 2019). This gene, involved in the assimilation of fatty acids, highlights the importance of food resources in the evolution of diadromous behavior, as well as the significance of one gene (or copy number variant) to promote or constrain the dispersal to a new environment. However, the fact that diadromous species should have evolved all these adaptations (Fig. 2.4)

required to survive in freshwater and marine environments at one point seems unlikely. A more plausible scenario is that adaptations such as wide osmoregulatory capacity may be inherited before a species becomes diadromous and that what makes a species to become diadromous must be a gene or genes related to signaling that would start the migration. In birds, for instance, it has been hypothesized that a gene related to behavior such as circadian behavior or photoreceptors may be responsible for migratory behavior (Lugo Ramos *et al.* 2017).

Genes that influence migration may be linked together in chromosomal rearrangements and genomic islands of differentiation (Wellenreuther & Bernatchez 2018), as it has been suggested that these associations of genes in genomic regions facilitate the selection in favor of or against complex life history traits. Chromosomal rearrangements including inversions and duplications appear to maintain co-adapted alleles facilitating adaptation in many contexts (Sutherland *et al.* 2016; Wellenreuther & Bernatchez 2018; Varadharajan *et al.* 2019). Genomic islands of differentiation also appear to play key roles in linking co-adaptive traits in diadromous species (e.g., Veale & Russello 2017). For example, the Omy5 linkage group, a large region located in chromosome 5 of steelhead shows a strong non-random association with life history differentiation between anadromous and non-migratory populations (Pearse *et al.* 2014). Recently, a study in Chinook Salmon showed that one small "region of strongest association" (RoSA) of ~30Kb was associated with spawning migration time (Thompson *et al.* 2020).

Genomic and transcriptomic analyses have revealed many putative genes that differentiated migratory and resident populations, these genes are related to osmoregulation, muscle contraction, among other processes (Table S2.3). Research in salmonids has shown little parallelism in genes differentiating diadromous and resident populations across species (Schneider *et al.* 2019). Even within the same species, replicate resident populations show that local adaptation and genetic drift can lead to different genes being fixed or lost (Chapter 4; Delgado *et al.* 2020; Salisbury *et al.* 2020). Thus, further research on a variety of species and tissues including the brain is necessary to improve our understanding of the genes that may be common among diadromous

species. Such efforts could help address the overarching question of whether or not there is a "diadromous gene" or gene complex.

2..4.4 Origin and ancestry

The fact that diadromy is present only in a small percentage of species, yet it is widely present across the phylogeny of fishes (Fig. 2.1) would suggest that diadromy could have originated in either of two scenarios (Fig. 2.5). The first hypothesis posits that diadromy appeared multiple independent times across the phylogeny of fishes; the second hypothesis instead, posits that diadromy is ancestral and has been lost on multiple occasions (Fig. 2.5).

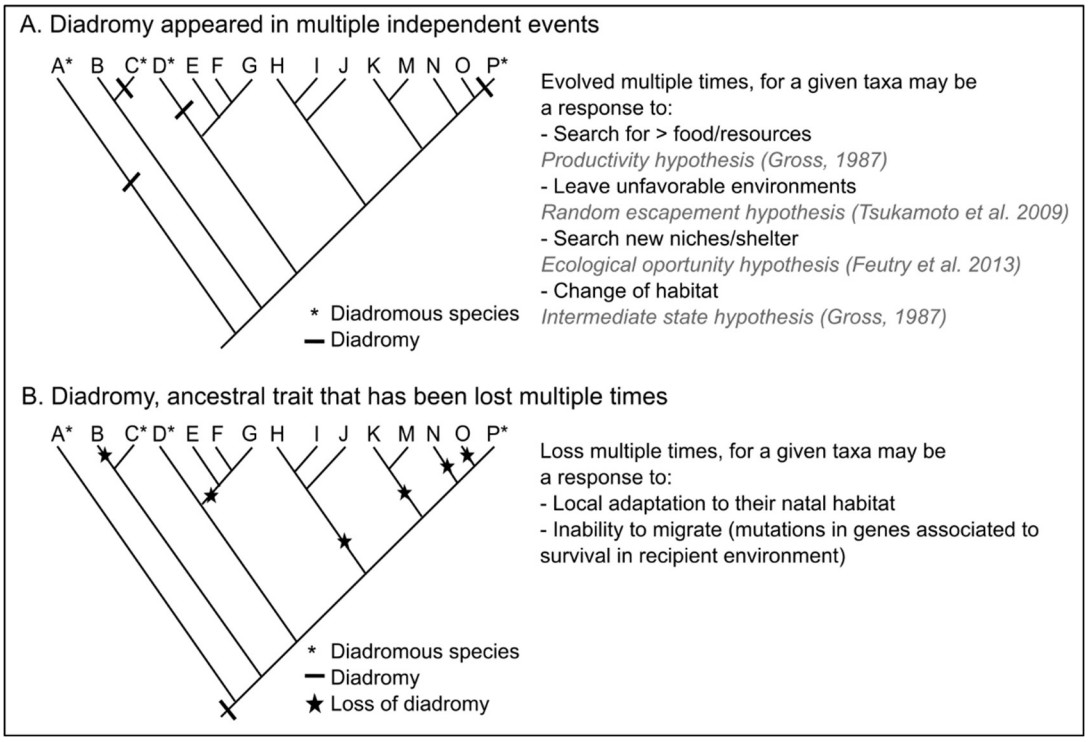


Figure 2.5 Hypotheses regarding the origin of diadromy. **A.** Diadromy appeared in multiple events as a response to different scenarios (one or more of previously proposed hypotheses) **B.** Diadromy is an ancestral trait that has been lost multiple times in fish phylogeny.

The second hypothesis, common to other migratory taxa including birds, assumes that species have an ancestral migratory predisposition, therefore, migration, in theory, could appear in many lineages (Zink 2011). In fishes, the rarity of diadromy (<1%) and

the lack of parallelism in genes associated with a migratory phenotype suggest that diadromy may not be predisposed in all lineages. Although the loss of diadromy does appear to be common in diadromous species, as seen by the existence of resident populations in many diadromous species across the phylogeny (Fig. 2.1), many more instances of loss of diadromy are required under this second hypothesis than are instances of the appearance of diadromy under the first hypothesis. The hypothesis that diadromy appeared multiple independent times across the phylogeny of fishes thus seems more parsimonious.

Under the multiple appearance hypothesis, it can be also assumed that in some lineages/orders where this life history was advantageous, it rapidly expanded, resulting in a relatively large number of related species evolving from a common diadromous ancestor. Examples could include the Gobiiformes and Salmoniformes, where single families within each of these orders exhibit a high number of diadromous species (Fig. 2.1). The first hypothesis also assumes the independence of each diadromy appearance events, suggesting that any of the proposed hypotheses on the origin of diadromy (see "Hypothesis on the origin" section), all of which have evidence for and against, may be true for a given taxon. This is also consistent with the suggestion that no single hypothesis explains the origin of diadromy for all diadromous species adding support to the first hypothesis. Recently, Alò and collaborators tested different hypotheses (e.g., productivity and genetic predisposition) to find one comprehensive migration model. Their results, however, showed that different migratory strategies including the different categories of diadromy, cannot be explained by one model but by different environmental, phylogenetic, and productivity variables (Alò *et al.* 2020).

There is also an ongoing debate about the salinity at which the ancestors of diadromous species lived. The ancestors have been hypothesized to be of freshwater, marine, or diadromous origin (McDowall 1997). The ancestor species can be assumed to be of freshwater or marine environment as a function of the hypothesis of the origin of diadromy. For instance, under the intermediate state hypothesis, anadromy is derived from a freshwater ancestor, but under the safe-site hypothesis, anadromy derives from a marine ancestor. Depending on the taxa, both assumptions can be valid.

The phylogenetic tree suggests that both anadromous and amphidromous species appear to have both marine and freshwater ancestors (Fig. 2.1). Catadromous species, on the other hand, appear to be present mainly in clades where the most recent ancestor was of marine origin (Fig. 2.1). For most clades, however, there is no consensus on the habitat of the ancestral species. Indeed, we were unable to find information on the habitat of the ancestor for many taxa (Table S2.1). Although no single hypothesis explains diadromy for every diadromous species, the presence of a marine ancestor for catadromous species suggests that diadromous species may migrate for ecological reasons (i.e., increase fitness) but return to their natal habitat because they lack adaptations that would allow them to reproduce in the secondary environment. Thus, catadromous and marine amphidromous likely have a marine ancestor that passed on the ability to reproduce in marine environments; and similarly, anadromous and freshwater amphidromous most likely have a freshwater ancestor. However, there are notable exceptions like the order Clupeiformes which held mostly anadromous species yet its most recent ancestor is marine.

The presence of marine ancestry for some anadromous or freshwater amphidromous could be explained by novel mutations that allow these species to reproduce in fresh water or by standing genetic variation of a slightly older freshwater ancestor. The appearance of novel mutations seems unlikely as these mutations would have had to appear multiple times in different lineages. The latter hypothesis (standing genetic variation) seems more plausible, as it is known that most actinopterygians are derived from a common freshwater ancestor (Vega & Wiens 2012). The presence of pre-existing or cryptic genetic variation could explain why some anadromous and freshwater amphidromous species from multiple and independent lineages have a recent marine ancestor. This idea that most actinopterygians have standing genetic variation to reproduce in fresh water may also explain why anadromy and freshwater amphidromy are more prevalent than catadromy (<20% of diadromous species are catadromous – Table 2.1). Although, the higher prevalence of anadromous vs. catadromous species could also be explained by the higher speciation rate of anadromous species as the consequence of more opportunities to isolate and differentiate in freshwater environments than in marine environments.

2.5 Future directions of diadromy in an -omics era – what we can learn?

Research using -omics approaches regarding diadromy is still in its infancy. Further research on different species from all three categories of diadromy would help answer questions on the origin and molecular bases of diadromy. Questions such as whether diadromy evolved multiple independent times or it is an ancestral trait could be addressed by having more genomic data of diadromous species within and among different orders, as thus far, most research has focused on Salmoniformes. Genomic data of species from different categories of diadromy from the same and different orders or even families (e.g., Clupeidae), would also contribute to resolving the question of whether the different categories of diadromy have the same genomic bases. Also, studies focusing on orders that show rapid diversification (i.e., Gobiiformes) could shed light on questions such as why freshwater amphidromy is more prevalent than other categories?

Transcriptomic and epigenomic research with diadromous and resident populations exhibiting little genetic differentiation, and which therefore have not undergone local adaptation, would help assess which differentially expressed genes are responsible for this migratory behavior and if indeed there is a "diadromous" gene(s) or gene complex. Research thus far has focused on osmoregulatory organs (i.e., gills), however, wide resistance to osmoregulatory changes is present too in non-diadromous species as well suggesting that this adaptation is an ancestral trait. Thus, examining other organs including brains where genes related to signaling and photoreceptor are expressed may provide an answer to the question of what gene is responsible for starting the migration of larvae/juveniles.

Research on ecological factors that may influence the decision to migrate, including the presence of predators or the influence of parasites, may also lead to improvements in our understanding of the facultative nature of diadromy which is present in many diadromous species. Finally, understanding the evolutionary consequences of diadromy and its loss as a source of genetic diversity but also considering the increased risk of extinction that this life history trait carries can be important to predict the evolution of these species.

2.6 Concluding remarks

Diadromy is a life history trait that has an important role in the distribution and even diversification of species in aquatic systems, directly influencing the ability of populations to disperse and colonize new habitats and niches. This migratory behavior involved a series of adaptations that grants individuals the ability to survive in such different environments. And although, many questions regarding the origin, genetic bases, and evolutionary consequences of diadromy still need an answer, our understanding of diadromy has improved during the last decade thanks to advances in sequencing technologies. For example, research has revealed key genes and processes responsible for these adaptations. Also, studies suggest that diadromy may have appeared convergently in different taxa, proving to be an important source of genetic variation in fishes. Unfortunately, many diadromous species have experienced population declines throughout the last century (McDowall *et al.* 2009; Righton *et al.* 2012; Duarte 2018), and many diadromous species are in peril due to their migratory nature. Conservation strategies for a diadromous species require the preservation not only of their natal and secondary environment but also the connection between them (McDowall 1999) as well as the consideration of future climate-driven changes in both freshwater and marine environments (Walter *et al.* 2012). Thus, given the conservation status of many diadromous species and the current climate change scenario, understanding the variability and potential to adapt to different environments of diadromous species is crucial for their preservation.

2.7 Acknowledgments

We thank Anne Dalziel, Paul Bentzen, Glenn Crossin, and the two anonymous reviewers for their valuable comments on a previous version of the manuscript. This manuscript is part of MLD's PhD research. MLD has been supported by Nova Scotia Graduate Scholarship (NSGS), a Dalhousie University Faculty of Graduate Studies scholarship, and an NSERC Discovery grant to DER.

2.8 Author contributions

The manuscript was written by MLD with input from DER.

2.9 Supplementary information

Table S2.1 References of diadromous taxa's ancestral environment from Figure 2.1.

Taxa	Origin	Reference
Acipenseriformes	Fresh water	Sulak & Randall 2002
Ambassidae	Marine	Martin 1990
Anguilliformes	Marine	Corush 2019
Atheriniformes	Marine	Campanella *et al.* 2015
Characiformes	Marine	Chen *et al.* 2013
Clupeiformes	Marine	Bloom & Lovejoy 2014
Cypriniformes	Fresh water	Imoto *et al.* 2013
Elopiformes	Unknown	
Gadiformes	Unknown	
Galaxiiformes	Unknown	Vega & Wiens 2012
Gobiesoformes	Unknown	
Gobiiformes	Fresh water	Thacker 2009
Lutjanidae	Unknown	
Moroniformes	Unknown	
Mugiliformes	Unknown	
Osmeriformes	Unknown	Vega & Wiens 2012
Perciformes	Marine/Catadromous	Cottidae (Dickman 1995), Terapontidae (Davis *et al.* 2012)/Kuhliidae (Feutry *et al.* 2013)
Petromyzontiformes	Fresh water	Bartels *et al.* 2017
Pleuronectiformes	Marine	Azevedo *et al.* 2008
Salmoniformes	Fresh water	Wang *et al.* 2011
Siluriformes	Marine	Betancur-R 2010
Syngnathiformes	Unknown	
Tetraodontiformes	Marine	Yamanoue *et al.* 2011

Table S2.2 A representation of putative genes that differentiate migratory vs. non-migratory populations reported by different studies.

Function	Example of putative genes	Reference
Cell junction/adhesion	Tight junction protein ZO-3, Occludin, Protocadherin-18, Cadherin-8	Hale *et al.* 2013; Kozak *et al.* 2014
Cell proliferation	Epidermal growth factor receptor kinase	Kozak *et al.* 2014
Cytoskeletal connections	Obscurin-like 1	Morris *et al.* 2014
Reproduction	Zonadhesin-like, Estrogen receptor, MORC family CW-type zinc finger, Round spermatid basic protein 1-like, RING finger protein 114, life history divergence [9], Gonadotropin-releasing hormone (GnRH), Sperm-associated antigen 16, Gonadotropin subunit beta-2	Mavarez *et al.* 2009; Hale *et al.* 2013; Mateus *et al.* 2013; Kozak *et al.* 2014
Growth/differentiation factor, hormone, FSH inhibitor	Inhibin, alpha, growth hormone 2 (GH2)	Hale *et al.* 2013; Morris *et al.* 2014
Involved in immunity	NOD-like receptor family CARD domain containing 5, Immunoglobulin heavy chain (IgD-A) gene, MHC class I a region	Hale *et al.* 2013; Morris *et al.* 2014
Ion transport	ATPase, Na+/K+ transporting alpha 1, Solute carrier family 9, 10, 12 (Na+/H+ exchanger), ATPase, H+ transporting, Sodium/potassium/calcium exchanger, Potassium voltage-gated channel subfamily H, Na+/Cl- cotransporter	Mavarez *et al.* 2009; Hale *et al.* 2013; Dennenmoser *et al.* 2017; Velotta *et al.* 2017; Brennan *et al.* 2018; Willoughby *et al.* 2018; Delgado *et al.* 2019
Microtubule attachment to the centromere	Bardet-Biedl syndrome 4 protein	Kozak *et al.* 2014
Enzymes	Glutamate dehydrogenase (GDH), d1-pyrroline-5-carboxylase synthase (P5CS), Carbonic anhydrase, Malate dehydrogenase, FMS-related tyrosine kinase 4, Alkaline ceramidase 1, Protein kinase D3	Hale *et al.* 2013; Dennenmoser *et al.* 2017; Debiasse *et al.* 2018; Willoughby *et al.* 2018
Myocyte cytoskeletal development	SPEG, Myosin regulatory light chain 2, Myostatin 2b (MSTN2)	Hale *et al.* 2013; Morris *et al.* 2014
Negative regulator of cell proliferation	Insulin-like growth factor-binding protein 1, 2a, 5	Morris *et al.* 2014; Kusakabe *et al.* 2017; Velotta *et al.* 2017
Osmosensing	Interleukin receptor 17c, 22a, Mitogen-activated protein kinase 1, 8, 13	Velotta *et al.* 2017
Osmotic/salinity stress	Glucocorticoid receptor, Aldehyde dehydrogenase 7, 9, Vasotocin	Mavarez *et al.* 2009; Mateus *et al.* 2013; Kozak *et al.* 2014
Regulation of immune cell proliferation	SAM and SH3 domain-containing protein	Kozak *et al.* 2014; Guo *et al.* 2015
Regulator of fatty acid uptake, intracellular binding	Peroxisome proliferator-activated receptor alpha a, Fatty acid-binding protein, adipocyte (AFABP)	Mavarez *et al.* 2009; Morris *et al.* 2014
Tight junction	Claudin 1, 3, 4, 7, 8, 10, 15	Kozak *et al.* 2014; Dennenmoser *et al.* 2017; Kusakabe *et al.* 2017; Velotta *et al.* 2017
Water transport	Aquaporin 3	Velotta *et al.* 2017

CHAPTER 3. THE EFFECTS OF DIADROMY AND ITS LOSS ON GENOMIC DIVERGENCE: THE CASE OF AMPHIDROMOUS GALAXIAS MACULATUS POPULATIONS

This chapter has been published as "ML Delgado, K Górski, E Habit & DE Ruzzante. 2019. The effects of diadromy and its loss on genomic divergence: the case of amphidromous Galaxias maculatus populations. Molecular Ecology. 28, 5217–523"

3.1 Abstract

Understanding the evolutionary mechanisms that affect the genetic divergence between diadromous and resident populations across heterogeneous environments is a challenging task. While diadromy may promote gene flow leading to a lack of genetic differentiation among populations, resident populations tend to be affected by local adaptation and/or plasticity. Studies on these effects on genomic divergence in non-model amphidromous species are scarce. *Galaxias maculatus*, one of the most widespread fish species in the Southern Hemisphere, exhibits two life histories, an ancestral diadromous, specifically, amphidromous form and a derived freshwater resident form. We examined the genetic diversity and divergence among 20 estuarine and resident populations across the Chilean distribution of *G. maculatus* and assessed the extent to which selection is involved in the differentiation among resident populations. We obtained nearly 4400 SNP markers using a RADcap approach for 224 individuals. As expected, collections from estuarine locations typically consist of diadromous individuals. Diadromous populations are highly differentiated from their resident counterparts by both neutral and putative adaptive markers. While diadromous populations exhibit high gene flow and lack site fidelity, resident populations appear to be the product of different colonization events with relatively low genetic diversity and varying levels of gene flow. In particular, the northernmost resident populations were clearly genetically distinct and reproductively isolated from each other suggesting potential for local adaptation. Our study provides insights into the role of life history differences in the maintenance of genetic diversity and the importance of genetic divergence in species evolution.

3.2 Introduction

The mechanisms that lead to the evolution of alternative life histories and population differentiation are still a challenging and complex theme in evolutionary biology. At the molecular level, it is possible to observe genomic differentiation between alternative life history strategies suggesting a role for selection and local adaptation, a lack of genomic differentiation between forms suggesting instead, an important role for plasticity, or alternatively, a combination of both (Weitere *et al.* 2004; Côté *et al.* 2014; Hendry 2016). The use of molecular tools to distinguish alternative life history forms can help increase our understanding of local adaptation in allopatric and sympatric populations. Diadromy, a life history strategy that indicates an ability to migrate between marine and freshwater environments (McDowall 2008a) and its loss, represent two life histories with contrasting levels of population differentiation. While diadromy has generally been shown to be associated with high gene flow and low genetic divergence among populations (Waters *et al.* 2000a), resident, non-migratory populations are often more closely associated with genetic divergence, local adaptation, and even speciation by facilitating genetic drift, as populations in fresh water become reproductively isolated (McDowall 2001). These opposing responses make *G. maculatus* an appropriate model for the study of the patterns and processes involved in population divergence.

Most studies focusing on the migratory vs. resident life history dichotomy have been conducted with Northern Hemisphere species including Atlantic Salmon (*Salmo salar*) (Perrier *et al.* 2013), Brown Trout (*Salmo trutta*) (Hindar *et al.* 1991; Meier *et al.* 2014), steelhead (*Oncorhynchus mykiss*) (Thrower *et al.* 2004; Matala *et al.* 2017), and Stickleback (*Gasterosteus aculeatus*) (Drevecky *et al.* 2013; Ravinet *et al.* 2015; Ferchaud & Hansen 2016). Most of these model species are economically important and are classified as anadromous, a type of diadromy characterized by the return of mature or maturing individuals to their natal streams for reproduction. Studies focusing on a different type of diadromy, amphidromy where individuals migrate to the sea or streams during the larval stage, are rare (Augspurger *et al.* 2017). The present study focuses on an amphidromous organism from the Southern Hemisphere, *Galaxias maculatus*.

Galaxias maculatus is one of the most widespread freshwater species naturally distributed in the Southern Hemisphere (Chapman *et al.* 2006). Studies examining the

differences between diadromous and resident populations of *G. maculatus* have focused on the phenotypic differences between these two forms. Diadromous populations tend to have larger body size (Barriga *et al.* 2012), higher vertebral count (Campos 1974; McDowall 2003), and fecundity (Boy *et al.* 2009) than landlocked or resident populations. These morphological differences were initially described as being only phenotypic in origin with no genetic component (McDowall 1972; Waters & Burridge 1999 but see McDowall 2001), however, Campos (1974) suggested lacustrine *G. maculatus* populations with a distinct lower vertebral count should be considered as a distinct species, *G. alpinus* (Campos 1974). A subsequent study examining the post-glacial history of *G. maculatus* suggested that landlocked populations originated from different colonization events suggesting some degree of genetic differentiation among ecotypes (Zemlak *et al.* 2010). There is also evidence that resident and diadromous populations do live in sympatry (Carrea *et al.* 2013; Górski *et al.* 2018) suggesting that there likely are genomic regions of differentiation between these life history types.

Genetic studies comparing diadromous and resident *G. maculatus* populations are scarce. Four mitochondrial haplogroups have been described for South American populations, with all those from the Magellanic region (Southern part of the species distribution) belonging to the same haplogroup (Zemlak *et al.* 2010; González-Wevar *et al.* 2015b). A subsequent study comparing a single pair of diadromous and resident populations and using one mitochondrial marker found that the resident population exhibited lower diversity than the diadromous one, yet maintained the salinity tolerance similar to diadromous individuals (Ruiz-Jarabo *et al.* 2016). Advances in genome-wide sequencing for non-model organisms such as RAD-cap (Hoffberg *et al.* 2016) can facilitate the development of SNP markers across the genome at high coverage. This approach can thus help elucidate at a finer scale the genetic differences between diadromous and resident populations.

Our goal in the present study was to examine the genetic diversity and differentiation between and among diadromous and resident populations of *G. maculatus* across its distribution in Chile. First, we identified and characterized SNP markers for *G. maculatus* using a RADcap approach. We hypothesized that limited gene flow between diadromous and resident populations drives significant neutral and adaptive genomic

differences between these two life histories of *G. maculatus*. We expected diadromous and resident populations to be genetically distinguishable at neutral loci as a result of reproductive isolation and genetic drift. We further expected diadromous and resident populations to be distinguishable at a limited number of loci putatively under selection including some reported to be linked to salinity tolerance in other species. This implies a role for selection and local adaptation to different salinity levels. Overall, this study provides insights into the effects of amphidromy and its loss on the genetic differentiation of *G. maculatus* populations across its Chilean distribution.

3.3 Material and Methods

3.3.1 Study area and sample collection

We extracted genomic DNA from *N*=260 *G. maculatus* individuals collected between 2006 and 2017 from a total of 20 populations spanning the range of the species distribution in central-southern Chile and Patagonia (32ºS in central Chile to 52ºS in Tierra de Fuego). Ten of these populations were within a few kilometers from the ocean and thus, presumably harbored diadromous *G. maculatus* populations while the remaining 10 populations targeted resident or landlocked populations (Fig. 3.1, Table 3.1). In seven cases, estuarine and resident collections originated from paired, within river locations (Fig. 3.1). For two populations (ValEst and ValRes), individuals were collected from four different locations, two nearby locations for each population (Table 3.1, Fig. S3.1). Although ValEst[2] is located <35 Km from the ocean, it was considered an estuarine relative to the resident locations (ValRes) that are located >150 Km from the ocean. Stable isotope data, which can be used as an indication of the environment experienced by the individuals, were available for 15 locations (Górski *et al.* 2018). In particular, the Sulphur isotope ratio (34S to ^{32}S) in bone tissue samples, which is expected to differ between the marine and freshwater environments (Peterson & Fry 1987) confirmed that most estuarine collections comprised diadromous individuals, whereas most resident collections comprised individuals that showed no trace of migration to sea water (Górski *et al.* 2018). Five of the 15 collections, four from estuarine locations and one from a presumed resident population contained both diadromous and resident individuals (Table 3.1).

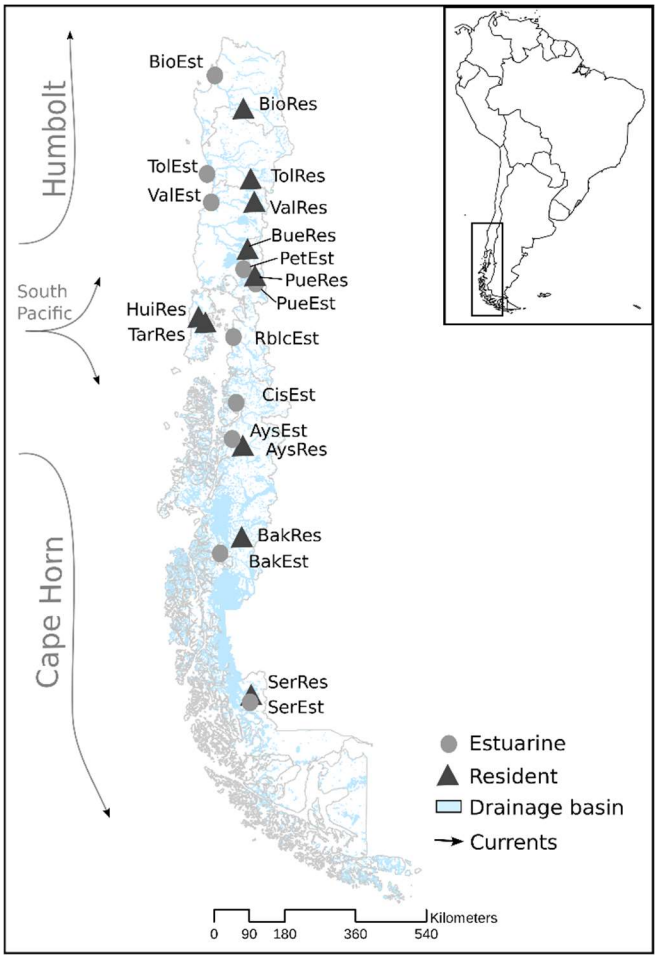


Figure 3.1. Sampling locations of 20 populations of *G. maculatus* across its Chilean distribution. A detailed description of each location is shown in Table 3.1 and Figure S3.1. The direction of the currents is shown with arrows.

Table 3.1. Sampling information of the 20 populations of *G. maculatus* included in the study.

Pop#	Location acronym	River System	Type	Latitude	Longitude	Distance from the ocean (Km)	Tissue sample source	Bone tissue microchemistry results (# sequenced indv w/ confirmed life history) (Gorski *et al.* 2018)
1	BioEst	Biobío river	Estuarine	-36.8067	-73.1698	1	Górski *et al.* 2018	Diadromous (2)
2	BioRes	Biobío river	Resident	-37.5679	-72.3885	115	Górski *et al.* 2018	Resident (9)
3	TolEst	Toltén river	Estuarine	-39.2467	-73.2199	1	Górski *et al*. 2018	Diadromous (2) and Residents
4	TolRes	Toltén river	Resident	-39.2762	-72.2259	117	Górski *et al.* 2018	Resident (2)
5	ValEst[1]	Valdivia river	Estuarine	-39.8546	-73.3325	1	Habit Lab	No data
	ValEst[2]	Valdivia river	Estuarine	-39.7854	-73.0047	<35	Górski *et al*. 2018	Diadromous and Residents
6	ValRes[1]	Valdivia river	Resident	-39.7747	-72.1242	150	Górski *et al.* 2018	Resident (7)
	ValRes[2]	Valdivia river	Resident	-39.7776	-71.9582	155	Habit Lab	No data
7	BueRes	Bueno river	Resident	-40.817	-72.4608	125	Górski *et al.* 2018	Resident (5)
8	PetEst	Petrohué river	Estuarine	-41.3792	-72.3123	1	Górski *et al*. 2018	Diadromous (8)
9	PueEst	Puelo river	Estuarine	-41.6455	-72.2727	1	Górski *et al*. 2018	Diadromous (10) and Residents
10	PueRes	Puelo river	Resident	-41.6381	-72.3194	>4	Górski *et al.* 2018	Diadromous and Residents (4)
11	RblcEst	Blanco river	Estuarine	-42.9345	-72.7233	<20	Zemlak *et al*. 2010	No data
12	HuiRes	Huillinco lake	Resident	-42.6712	-73.9029	<2	Zemlak *et al.* 2010	No data
13	TarRes	Tarahuín lake	Resident	-42.7169	-73.7503	<5	Zemlak *et al.* 2010	No data
14	CisEst	Cisnes river	Estuarine	-44.7442	-72.7014	1	Górski *et al.* 2018	Diadromous (8)
15	AysEst	Aysén river	Estuarine	-45.4172	-72.7401	3	Górski *et al*. 2018	Diadromous (4)
16	AysRes	Aysén river	Resident	-45.4999	-72.6755	>3	Górski *et al*. 2018	Resident (6)
17	BakEst	Baker river	Estuarine	-47.7864	-73.5344	1	Górski *et al*. 2018	Diadromous (6)
18	BakRes	Baker river	Resident	-47.6717	-73.0181	40	Górski *et al*. 2018	Diadromous and Residents (7)
19	SerEst	Serrano river	Estuarine	-51.4133	-73.0929	1	Habit Lab	No data
20	SerRes	Serrano river	Resident	-51.2696	-72.8271	<10	Habit Lab	No data

3.3.2 DNA extraction and RAD-seq pilot

DNA was extracted from a minimum of 10 individuals per population. Whole genomic DNA was extracted using a standard phenol-chloroform protocol. Extracted DNA was quantified using a plate reader, and all samples were normalized to 25ng/µl.

A Restriction site Associated DNA (RAD-seq) pilot was conducted to obtain sequences for reference contigs or "rad-nome". These reference contigs were then used to detect polymorphic sites (loci) and design baits for their "capture". For the pilot, 14 samples including a replicate sample from one individual (i.e., 13 individuals) were selected from six populations (three estuarine and three resident locations). A standard RADseq protocol was performed by a single digest reaction using the restriction enzyme *Sbf*I. Briefly, 1µg of DNA per individuals was digested using SbfI-HF (NEB) for one hour at 37°C. P1 adapters with custom 5 to 8 base pair barcodes were ligated, and samples were then multiplexed. Pooled DNA was sheared on an S220 series ultra-sonicator (Covaris Inc, MA, US) following manufacturers' setting for a mean 500bp fragment size. Sheared fragments were size selected between 300-500bp using 2% cassette Marker B Pippin Prep (Sage Science Inc, MA, US). Ends were then repaired using a blunting kit (NEB), A-addition was done using a dA-Tailing module kit (NEBNext), and finally, P2 adapters were ligated. Libraries were cleaned after every step using the SpeedBeads cleaning protocol (Rohland & Reich 2011). A 16-cycle PCR was performed using a Phusion Master mix (NEB). The library was then sequenced on the Illumina MiSeq platform to obtain paired-end 300bp reads.

Raw reads were demultiplexed to the specific barcode using PROCESS_RADTAGS (Catchen *et al.* 2013) and low quality reads were removed using the software Trimmomatic (Bolger *et al.* 2014). Sequences were analyzed following the dDocent pipeline (Puritz *et al.* 2014). The "ROL" assembly method was used. This method is built for reads obtained by random shearing (reads with different length) and pair-end reads that overlap. The "c" parameter of CD-HIT, which clusters reference sequences by similarity, was set to 0.92 for our specific data. At the end of this pipeline, reference contigs or "rad-ome" were created and contained 30049 contigs. Following filtering, 9536 loci were chosen for further analyses.

3.3.3 Bait design and RADcap

From the 9,536 loci, 12,391 baits were designed and synthesized by Arbor Biosciences. A modification of the RADcap protocol (Hoffberg *et al.* 2016) was used. The standard RADseq protocol described for the pilot was performed and was followed by the "capture" step using the designed baits. The capture was conducted following the manufacturer's protocol. Hybridization temperature was set to 60ºC and 24 hrs were allowed for the hybridization step. Sixteen libraries were created with 260 individuals and sequenced in one lane of Illumina MiSeq and five lanes of Illumina Hiseq (150bp single-end).

3.3.4 Data analyses

All raw reads were demultiplexed with PROCESS_RADTAGS and trimmed with Trimmomatic. Samples with fewer than 100 000 reads were discarded from further analyses. Mapping to the reference "rad-nome" was done using dDocent using the default settings, and the SNP calling was conducted with the ref_map script from STACKS (Catchen *et al.* 2013).

The raw file was filtered using VCFtools (Danecek *et al.* 2011). Different filtering schemes were tested (Table S1). FS3 was chosen as it maximized the number of loci retained while maintaining the patterns of differentiation between populations found by the most conservative schemes. This FS3 scheme was first filtered by a minimum depth of 10 and by removing sites with >15% of missing data. SNPs with a minimum allele frequency (MAF) <0.01 were also removed. Hardy-Weinberg equilibrium (HWE) by population was tested using a perl script (https://github.com/jpuritz/dDocent/blob/master/scripts/filter_hwe_by_pop.pl) with the p-value threshold set to 0.05. Only the first SNP per contig was retained to remove physically linked loci. Tests for linkage were performed for the remaining loci using VCFtools with the default settings (r^2 implied 0.9). Finally, individuals with >20 % missing data were removed. After these filtering processes, we retained 4388 SNPs in 224 individuals. VCFtools, PLINK (Purcell *et al.* 2007) and PGDSpider (Lischer & Excoffier 2012) were used to convert to the appropriate formats.

Principal Component Analyses (PCA) was conducted with PLINK. This analysis was conducted three times, first, considering all 224 individuals and all 4388 loci,

second, considering only neutral loci, and third, considering only outlier loci. PCA was also performed considering only the *N*=156 individuals from 15 populations that cluster together in the first PCA to visually detect the separation between these populations.

Candidate loci under selection were detected using three outlier detection methods that differ in their statistical approaches: *pcadapt* is based on principal component analysis where the markers that correlated to the genetic structure are identified as loci under selection (Luu *et al.* 2017); Bayescan uses differences in allele frequencies between populations to detect outliers (Foll & Gaggiotti 2008); and sNMF is based on population differentiation statistics obtained from ancestry coefficients (Frichot & François 2015). For the *pcadapt* method, the q-value was set to < 0.05. For the Bayescan analyses, we used 50 000 iterations with the 'prior' odd specified to 100. SNPs with a false discovery rate (FDR) < 0.01 were considered outliers. For the sNMF method implemented in the R package LEA, the FDR control was also set to $q < 0.05$. Shared and unique putative outliers found between the three detection methods were identified using bash commands.

VCFtools was used to remove all putative outlier loci, retaining a set of 3516 neutral markers. Diversity indices including observed and expected heterozygosities (Ho and He, respectively) were estimated using the "genind summary" function from *adegenet* (Jombart & Ahmed 2011). Private alleles were found with the "private_allele" function from Poppr (Kamvar *et al.* 2015).

A custom-made script was used to convert a ped file containing only neutral loci to a fasta format. IQ-tree was then used to produce a maximum-likelihood tree performing a bootstrap with 100000 replicates (Nguyen *et al.* 2015). The Poppr R package was used to run a minimum spanning network. Pairwise F_{ST}s (Weir & Cockerham 1984) were estimated using the "genet.dist" function in HIERFSTAT (Goudet 2005).

ADMIXTURE (Alexander *et al.* 2009) was performed using neutral loci. The analyses were run with K=1-20 and recoding the cross-validation (CV) error values. Results with the lowest CV error were plotted using R. Pie charts indicating the genetic group assignments by population were drawn in R. fineRADstructure (Malinsky *et al.*

2018) was run following the software pipeline with the default settings, and using their R script to plot the heatmaps.

The contigs of the top 50 potentially selective loci detected by each of the 3 outlier detection methods were blasted using the BLASTN search tool (NCBI) to identify the genes involved in genetic differences between diadromous and resident individuals. BLAST results were filtered to include only hits that had >50% of identity. The potential role of these genes was searched in the Uniprot protein database (www.uniprot.org).

3.4 Results

RADcap sequencing of 260 *G. maculatus*' individuals from 10 estuarine and 10 resident populations, resulted in approximately half a million polymorphic sites. After filtering for low-confidence SNP calls and missing data, 4388 loci were selected from 224 individuals. Thirty-six individuals were removed during the filtering process. Individuals from the Serrano River system exhibited a high percentage of missing data resulting in low sample sizes from these locations: seven and two individuals were retained for the estuarine and resident locations, respectively. The final percentage of missing data across all 20 populations was <6 % (Table 3.2).

Table 3.2. Sample size (N) after filtering, % of missing data, number of private alleles, observed heterozygosity (Ho), and expected heterozygosity (He) for each of the 20 populations of *G. maculatus.*

#	Locations acronym	N after filtering	% missing data	# private alleles	Ho	He
1	BioEst	14	1.0	1	0.064	0.103
2	BioRes	12	5.9	231	0.042	0.061
3	TolEst	18	1.1	0	0.095	0.111
4	TolRes	10	2.3	64	0.036	0.054
5	ValEst	15	1.1	50	0.073	0.121
6	ValRes	15	1.9	34	0.027	0.039
7	BueRes	11	1.8	43	0.029	0.033
8	PetEst	13	0.9	0	0.094	0.102
9	PueEst	12	0.0	0	0.095	0.107
10	PueRes	12	0.2	1	0.088	0.102
11	RblcEst	10	0.6	0	0.087	0.106
12	HuiRes	9	0.9	1	0.072	0.084

#	Locations acronym	N after filtering	% missing data	# private alleles	Ho	He
13	TarRes	9	1.9	6	0.047	0.056
14	CisEst	9	0.9	1	0.080	0.104
15	AysEst	12	0.9	0	0.093	0.103
16	AysRes	12	1.2	9	0.075	0.081
17	BakEst	11	1.8	0	0.097	0.110
18	BakRes	11	1.6	6	0.071	0.079
19	SerEst	7	1.6	0	0.069	0.091
20	SerRes	2	0.0	0	0.066	0.062

PCA analyses using all 4388 SNPs from 224 individuals showed that resident populations, especially those from the northern locations (Biobío, Toltén, Valdivia, Bueno, and Tarahuín) are genetically distinguishable from estuarine populations (Fig. 3.2). PC1 and PC2 explained most of the differentiation with 24.7 % and 20.1 %, respectively. Resident individuals clustered by populations, while estuarine individuals did not show clear separation among populations. To detect separation at a finer scale, the five most distinct populations were removed, and PCA was performed with only 156 samples from 15 populations (Fig. 3.2). For this subset PCA, the first PC showed the difference between the five remaining resident and the 10 remaining estuarine populations. Estuarine populations were also differentiated from each other along PC1, Aysén estuarine individuals (AysEst) were the most distinguishable from other estuarine individuals, followed by fish from Puelo and Petrohué. PC2 distinguished resident populations, though this axis explained only 6.6 % of the differentiation. It is noteworthy that only half of all Puelo presumed resident individuals clustered together, while the other half clustered with estuarine individuals from Puelo and Petrohué.

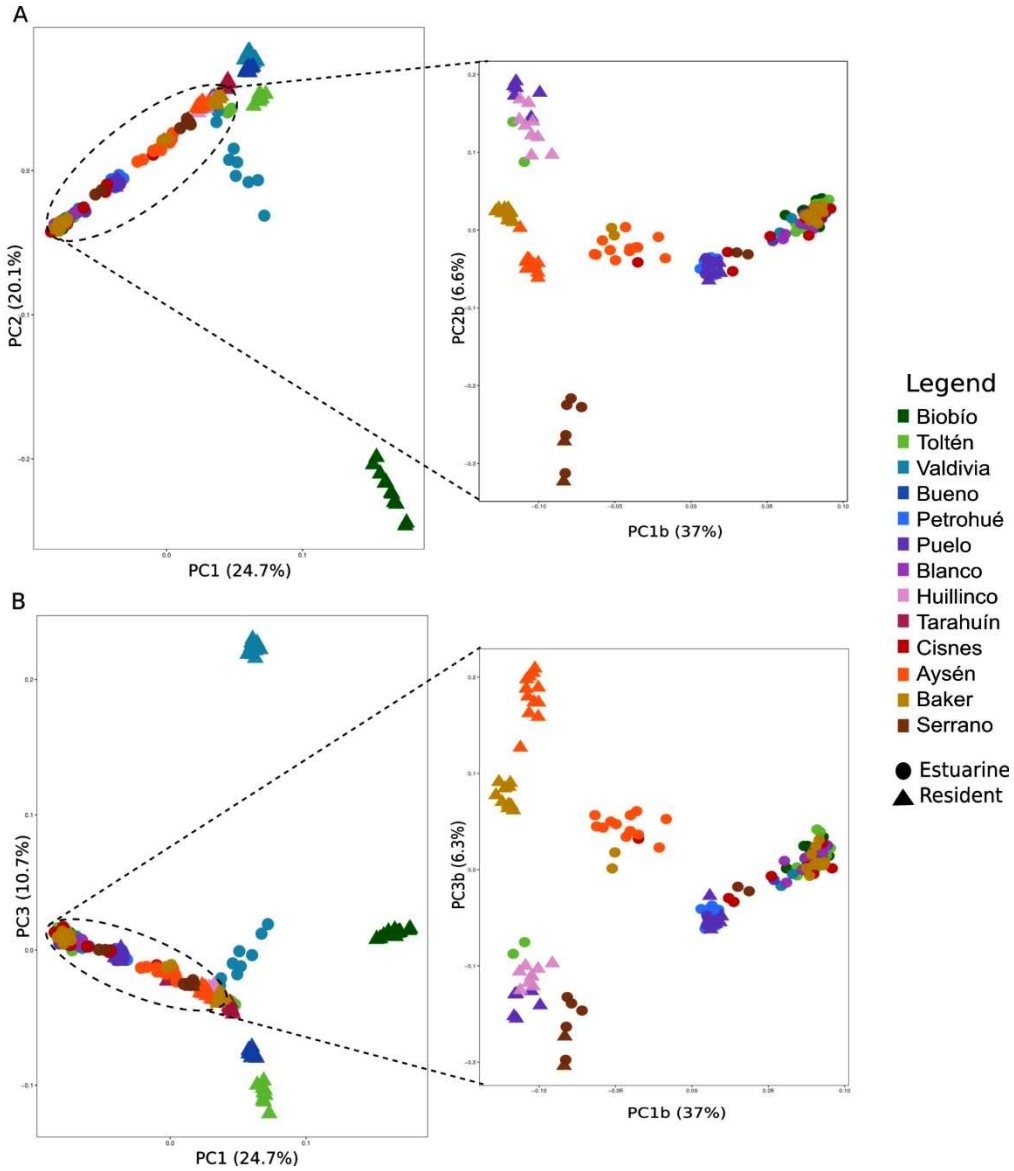


Figure 3.2. PCA performed on 224 individuals from 20 populations of *G. maculatus* using 4388 loci. Each color represents a unique river system. Circles represent estuarine individuals and triangles resident individuals. **A.** First and second PCA axes. **B.** First and third PCA axes. Next, to each graph, there is a PCA performed on the individuals inside the ellipse (156 individuals of 15 populations).

3.4.1 Neutral loci analyses

To follow a conservative approach, all potential outliers found by the three detection methods (872 loci) were removed in the analyses based on neutral markers (See the "Outlier loci analyses" below for details about the outlier detection methods). The neutral loci set thus comprised of 3516 loci. Estuarine populations exhibited higher

heterozygosities than their resident counterparts (Table 3.2). Resident populations, particularly the Biobío resident population exhibited the highest number of loci with private alleles (Table 3.2).

A maximum-likelihood tree based exclusively on the set of neutral loci shows nearly all estuarine individuals clustering together with the exception of some estuarine individuals from Valdivia that cluster with resident populations and some resident individuals from Puelo clustering together with the estuarine group (Fig. 3.3). Estuarine individuals from Valdivia that cluster with resident fish from elsewhere were collected in ValEst^2, a site located >35Km from the sea, suggesting they might be resident fish which nevertheless appear to be distinct from the resident individuals collected upstream in the upper reaches of the Valdivia system. With the exception of half of the Puelo resident individuals, which grouped with estuarine fish from Petrohué (Fig. 3.3), all other resident individuals grouped with individuals from their own location including both ValRes locations, and their relative positions in the tree tend to follow the geographic, latitudinal location. The northernmost resident population (Biobío) is also the most distinct among the resident populations, and the resident populations closest to the estuarine cluster are the southernmost ones (Baker and Serrano). The resident individuals from Aysén appear to derive from the estuarine/diadromous individuals from the same location (Fig. 3.3).

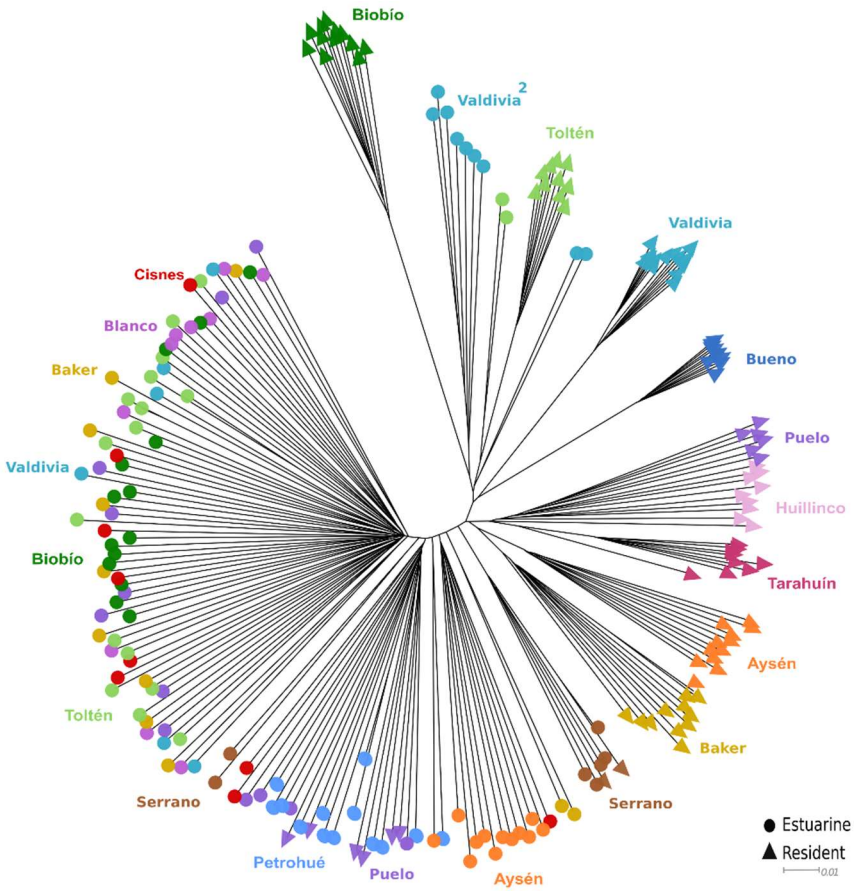


Figure 3.3. Maximum likelihood unrooted tree performed in 224 individuals from 20 populations of *G. maculatus* using 3516 neutral loci. Each color represents a unique river system, the names of which have the same color font. Circles represent estuarine individuals and triangles resident individuals.

Resident populations are more differentiated from each other than are the estuarine populations (Pairwise F_{ST} estimates, Fig. 3.4, Table S3.2), suggesting higher gene flow among estuarine than among resident populations (F_{ST} estuarine: 0-0.1, F_{ST} resident: 0.1-0.7). Further, resident populations appear to be derived from the estuarine populations (Fig. S3.3). Admixture analyses corroborated these results. STRUCTURE analysis indicated the presence of eight genetic groups (Fig. 3.5). Estuarine individuals across the latitudinal landscape appear to share a common genetic group (sea green color). A relatively small proportion of estuarine individuals are assigned to a second genetic group (leaf green color). The four northernmost resident populations formed their own genetic group. In the middle of the species distribution, resident individuals are mostly assigned to one genetic group (blue color), except for Puelo resident (Pop #10)

where some individuals are clearly assigned to the sea-green genetic group (shared among estuarine individuals) and the other half appeared to be residents.

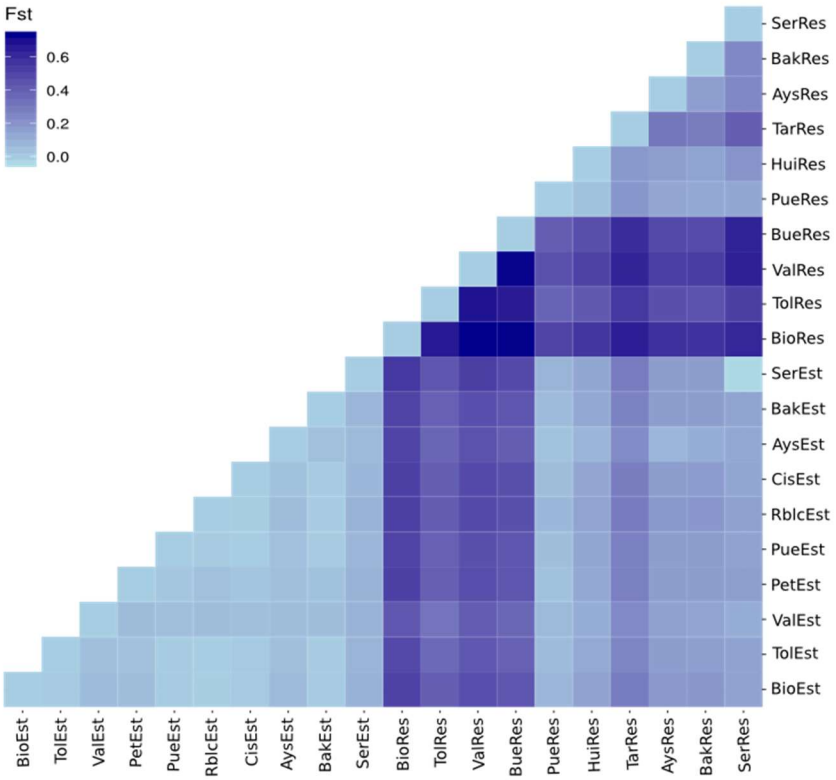


Figure 3.4. Heatmap of $F_{ST\text{-}WC}$ estimates of 20 *G. maculatus* populations. F_{ST} estimates based on 3516 neutral loci.

A population structure analysis based on haplotype coancestry using the neutral set of loci also demonstrated this genetic differentiation between estuarine and resident populations, clearly showing the presence of two genetic groups (Fig. S3.2). One mainly comprised of estuarine populations and the second, comprised of resident populations with the exception of Aysén, Serrano, and Valdivia ($ValEst^2$) estuarine individuals. Among the estuarine collections, a substructure was evident only in the middle of the latitudinal range (Petrohué, Puelo, Cisnes, and Aysén), while nearly all resident collections were distinguishable from each other.

Finally, the minimum spanning network analyses also confirmed these results (Fig. S3.4). Estuarine individuals cluster together and appear to form two genetically distinguishable groups. While resident populations instead, branching out from the

estuarine genetic group with some populations, specifically the northern ones, branching in an independent pattern.

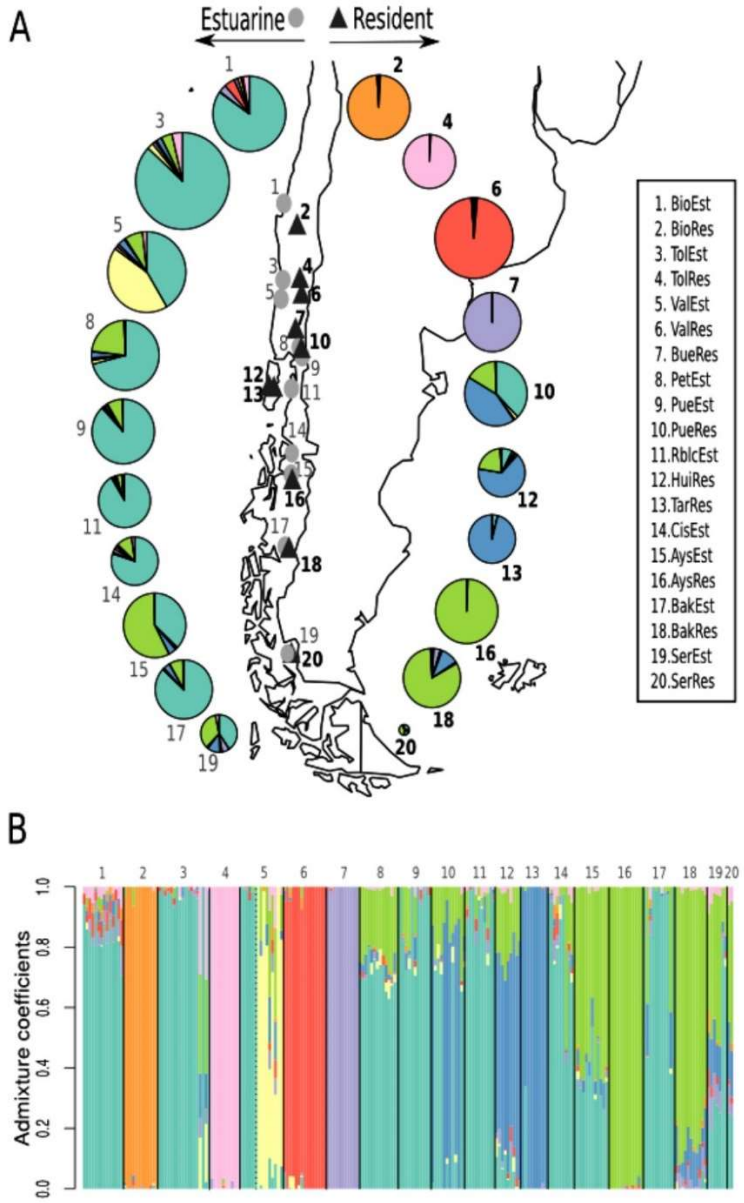


Figure 3.5. Genetic structure of 20 populations of *G. maculatus* in Chile based on 3516 neutral loci. Results are based on the most likely number of genetic groups found by ADMIXTURE **A**. Assignments of populations to 8 genetic groups visualized by pie charts on a map of southern Chile, each color represents one of the 8 distinct genetic groups. Estuarine charts are to the left and their geographic position are depicted with a circle, while resident charts are to the right, and their position is depicted with a triangle. The size of the charts is proportionate to the samples size. **B**. Barplot of admixture analysis of individuals to the 8 admixture groups. Each individual is represented by a bar and each population is separated by a black line. The dotted line separates ValEst[1] from ValEst[2].

3.4.2 Outlier loci analyses

The loci identified as putative outliers differed among the three outlier detection methods used (Fig. 3.6). Bayescan detected 437 outlier loci, while *pcadapt* detected only 366 outlier loci. Bayescan and sNMF shared the highest number of putative loci, 290 common outliers. Over all three methods, there were 872 unique putative outliers but only 13 were shared by all three methods.

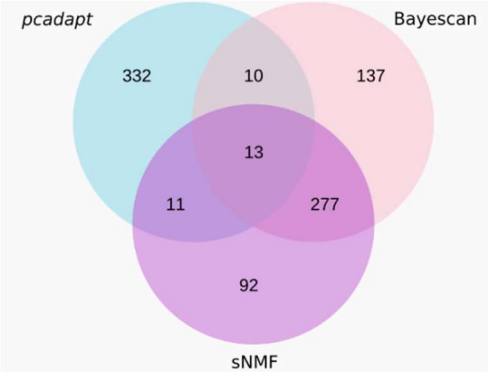


Figure 3.6. Venn diagram showing the different number of outlier loci found by three outlier detection software packages: *pcadapt*, Bayescan, and sNMF.

Using all outlier loci detected by any of the methods (i.e., 872 loci), PCA analyses distinguish the five most distinct resident populations, in agreement with the results using neutral markers (Fig. S3.5). At a finer scale, PC3 distinguishes the other resident populations, though only explaining 3.6 % of the total variance. The coancestry haplotype analysis using these outlier loci clearly distinguishes between estuarine and resident populations, in a similar way as the analyses performed using neutral loci (Fig. S3.6).

Galaxias maculatus does not have a sequenced reference genome yet, which makes it difficult to examine the identity of some of the genes associated with the SNPs. From the top 50 putative outlier loci, half had a function described in other species. We further looked into the top 25 outliers for which a function was found (Table 3.3, Table S3.3). Blast analyses and Uniprot search revealed the presence of some genes related to salinity tolerance and adaptation to marine environments, including activation of calcium channels, muscle contraction, and expression of potassium channels. It is important to note, however, that even though the percentage of identity of the blast results are >67%, the query cover percentages do vary greatly among alignments (13-99%) (Table 3.3).

Table 3.3. Blast annotation results from the top 25 outliers that have a potential function described in TableS3.3.

Contig number	Protein description	% identity	Query cover (%)
Contig_35840	LIM domain-containing protein ajuba-like	75.15	99
Contig_68102	Nuclear receptor subfamily 2 group F member 6	84.98	98
Contig_70236	Protocadherin-8 like	79.94	98
Contig_4613	Voltage-dependent calcium channel gamma-2-subunit	85.02	96
Contig_23889	Tensin 4 (TNS4)	69.44	72
Contig_73552	Myosin-If	83.33	69
Contig_18164	Thrombospondin type-1 domain-containing protein 7A-like	85.19	68
Contig_46934	Kinesin-like protein	90.6	67
Contig_50594	A disintegrin and metalloproteinase with thrombospondin motifs 10-like	67.61	59
Contig_50464	Proteasome subunit alpha type-6-like	87.58	53
Contig_32554	Actin-binding LIM protein 3-like	85.11	51
Contig_55409	Heat-stable enterotoxin receptor-like	87.15	47
Contig_54325	Semaphorin-4B (SEMA4B)	86.09	46
Contig_65480	Synaptotagmin-7	92.65	44
Contig_13463	Cullin-associated NEDD8-dissociated protein 1-like	89.31	43
Contig_29616	Phosphatidylinositol phosphatase PTPRQ	88.37	41
Contig_37112	Sidekick cell adhesion molecule 1 (SDK1)	82.44	37
Contig_19277	Potassium voltage-gated channel subfamily KQT member 5-like	92.86	34
Contig_41866	Inositol 1,4,5-trisphosphate receptor type 1 (Itpr1)	87.39	34
Contig_57173	Inactive dipeptidyl peptidase 10	81.43	34
Contig_71977	Adenylate kinase 9 (AK9)	80.95	33
Contig_39334	PAXIP1 associated glutamate-rich protein 1 (PAGR1)	82.28	25
Contig_7588	Sodium/potassium/calcium exchanger 1-like	84.48	19
Contig_20640	Sodium-independent sulfate anion transporter-like	84.78	13
Contig_69735	High affinity choline transporter 1-like	82.22	13

3.5 Discussion

We have used a suite of nearly 4400 SNP markers identified with the RADcap approach to describe genomic differences between diadromous and resident *G. maculatus* populations across the species' geographic range in Chile. While more than 3500 SNPs were neutral, the remaining 872 markers were identified as potential outliers and, hence, potentially subject to divergent selection, suggesting the genomic differences between diadromous and resident populations are the result of both neutral and adaptive processes.

Estuarine collections exhibited high levels of genetic diversity and low levels of genetic structure as expected from the fact that most individuals collected in the estuaries were indeed diadromous as assessed by analyses of stable isotopes in bone tissue (Górski *et al.* 2018). Resident populations, instead, exhibited relatively low levels of genetic diversity, and the northernmost resident populations were more highly differentiated from each other than the resident populations from locations in the southern part of the species' range in Chile. We discuss these results and their implications in detail below.

3.5.1 Diadromous populations

Estuarine collections, which generally comprised diadromous individuals as assessed by their bone tissue isotopic composition (Górski *et al.* 2018), showed higher levels of genetic diversity than resident populations. Diadromy is the ancestral trait of *G. maculatus*, and individuals have been found up to 700 km from the coast (McDowall *et al.* 1975). It has been hypothesized that Chilean populations originated from New Zealand individuals that arrived via the West Wind Drift (Waters & Burridge 1999), and ongoing migration of diadromous individuals from Australia and New Zealand to South America has also been suggested (Berra *et al.* 1996). A continuous influx of migrants from a genetically distinct location or locations could also contribute to the high levels of genetic diversity found in the diadromous populations.

Oceanic dispersal facilitates gene flow among diadromous populations (McDowall 1998). Many diadromous species including Stickleback (*Gasterosteus aculeatus*) and Atlantic Salmon (*Salmo salar*) are characterized by relatively high gene flow among their diadromous populations (Drevecky *et al.* 2013; Perrier *et al.* 2013; Ferchaud & Hansen 2016). The absence of philopatry or relatively low homing ability that characterizes *G. maculatus* (Barker & Lambert 1988; Waters *et al.* 2000b) can lead

to yet higher gene flow among locations than otherwise expected under philopatry. Here, we have shown that diadromous populations exhibit low levels of differentiation with nearly all estuarine individuals clustering together without necessarily appearing more similar to other individuals from the same location than to individuals from other locations, with the exception of diadromous individuals in the middle of the distribution (Puelo, Petrohué and Aysén) (Fig. 3.3). These results are consistent with the presumed absence of site fidelity to a stream. This high mobility of individuals is expected to contribute to the lack of adaptive differentiation among diadromous populations due to the homogenizing effects of gene flow (Raeymaekers *et al.* 2014), despite the difference in environmental conditions, including temperature, across the species range in Chile (Navarrete *et al.* 2014; Strub *et al.* 2019).

Previous studies regarding the genetic structure of populations of *G. maculatus* focused on the D-loop or control region mitochondrial marker. Zemlak *et al.* (2010) described 4 haplogroups across the distribution of the species in South America (see also González-Wevar *et al.* 2015). In both studies, southern collections (from Chiloe to Tierra de Fuego) exhibited little structure, and these individuals from the South were assigned to one haplogroup. Not surprisingly, the use of genome-wide SNP markers in the present study facilitated the detection of genetic differentiation at a finer geographic scale among the southern populations than did the previous studies based on mtDNA (Zemlak *et al.* 2010; González-Wevar *et al.* 2015b). Although the southern populations are indeed less diverse than the northern populations, there are still detectable differences between the diadromous and resident populations in the southern region.

Although estuarine individuals cluster together, they also exhibit some degree of differentiation. Patterns recovered by admixture and co-ancestry analyses detected one subgroup (Fig. 3.5, Fig. S3.2, Fig. S3.6), comprising mainly individuals from the middle part of the species distribution: Petrohué, Puelo, and Cisnes (Fig. 3.1). At the neutral level, estuarine individuals from Aysén appear to cluster together (Fig. 3.3, Fig. S3.2). Based on their geographic position, the larvae from these locations are expected to migrate into the Chiloé inner sea. The Chiloé island serves as a barrier, limiting the current flow into the Chiloé inner sea (Strub *et al.* 2019). This barrier could explain why

the diadromous individuals collected from streams and estuaries that drain into the Chiloe inner sea are slightly genetically distinguishable from other diadromous individuals.

3.5.2 Resident populations

McDowall suggested that anadromous and amphidromous fish species may be evolving towards an entirely freshwater life cycle (McDowall 1997) with the differences between diadromous and resident individuals being partly genetic in origin, and partly the result of plasticity (McDowall 2001). In sticklebacks, for instance, resident populations appear to be the result of independent colonization events (Defaveri *et al.* 2011). A reduction in genetic diversity is expected if these populations are the product of different founder effects as shown in steelhead trout (Willoughby *et al.* 2018). Under this scenario, local adaptation would be expected to play an important role in the differentiation between ecotypes. The facts that the *G. maculatus* resident populations examined here were highly differentiated from each other and from the estuarine collections at both neutral and adaptive loci, and that resident populations were also less diverse than their estuarine counterparts are consistent with independent colonization events and local adaptation. Different colonization or founding events have been hypothesized for *G. maculatus* Patagonian landlocked populations, suggesting the presence of older and younger populations (Cussac *et al.* 2004). The high gene flow among diadromous collections prevents us from distinguishing independent colonization events in the phylogenetic tree. One exception, however, is the estuarine (diadromous) collection from Aysén, which is somewhat distinguishable from all other diadromous collections. The resident collection from the same system (Aysén) appears to be derived from the Aysén estuarine/diadromous population (Fig. 3.3), supporting the hypothesis of a separate colonization event for at least this system.

Although resident populations were less diverse than their diadromous counterparts, there were also differences among resident populations in their levels of genetic diversity and differentiation: resident populations in the northern part of the distribution were more highly differentiated and exhibited lower genetic diversity than southern resident populations (Table 3.2, Fig. 3.2, Fig. 3.4). Resident populations, especially those in the North, experience low to no gene flow and are most likely characterized by lower effective sizes compared to their diadromous counterparts. Loci

can, therefore, be expected to reach fixation at a relatively fast rate consistent with the high number of private alleles present in these northern resident populations (Table 3.2). Furthermore, as shown in sticklebacks strong environmental heterogeneity among freshwater locations can lead to inconsistency in the outliers found among different freshwater and marine/diadromous populations (Ferchaud & Hansen 2016).

The "weak" differentiation that characterized the southern resident populations is likely a consequence of their relatively recent origin. Most of the river systems in this study (except the 3 northernmost rivers, i.e., Biobío, Toltén, and Valdivia) were covered by ice during the last glacial maximum (LGM) (Zemlak *et al.* 2010, 2011). As shown in other species (e.g., anadromous American shad, *Alosa sapidissima*), relatively low levels of genetic diversity and structure can be found in rivers that were glaciated and were thus colonized postglacially (Hasselman *et al.* 2013). In Patagonia too, lakes that were ice-covered during the Last Glacial Maximum (LGM) were colonized soon after deglaciation (Zemlak *et al.* 2011; Vera-Escalona *et al.* 2015, 2019). An earlier colonization of lakes for the northernmost populations can explain their higher genetic differentiation from their diadromous counterparts. While, in the southern populations, recent colonization can explain the weak differentiation as the standing genetic variation from the ancestral population (diadromous ecotype) will likely be the source of variation (Rivas *et al.* 2018).

All seven individuals sequenced from the second estuarine location of Valdivia ($ValEst^2$) appear to be resident individuals, though stable isotope results suggest that this location harbored both diadromous and resident individuals we did not have stable isotope results for these specific seven individuals. Interestingly, although these individuals from the second estuarine location ($ValEst^2$) cluster with other resident populations, they are genetically distinguishable from the resident population collected further upstream of the Valdivia River system (ValRes). The neutral and adaptive differentiation between these two locations (e.g., between residents from >35 Km and >150 Km upstream) could result from environmental heterogeneity.

3.5.3 Diadromous vs Resident

Based on our results and the stable isotope analyses in Górski *et al.* (2018) of a subset of our samples, we have shown that diadromous and residents individuals can be found in sympatry, the clearest example in our study is the resident population from the

Puelo river (PueRes), where nearly 50 % of the individuals were assigned genetically to the diadromous group. Stable isotope analyses also suggested that Toltén and Valdivia estuarine populations have diadromous and resident individuals, yet our genetic analyses with a limited number of individuals found only diadromous individuals in Toltén and the most estuarine location in the Valdivia River ($ValEst^1$), and only resident individuals in the second (upstream) estuarine location in the Valdivia River ($ValEst^2$).

High gene flow among different populations can constrain selection and have a homogenizing effect on differentiation. When in sympatry, two forms can in principle experience hybridization, and perhaps even introgression and recombination but there could also be strong reproductive isolation between ecotypes. For example, in sticklebacks, the high phenotypic diversity in contact zones reflects reproductive isolation (Ravinet *et al.* 2015). Sticklebacks are rarely sympatric but when they occur in sympatry and hybridize, they rarely show introgression (McPhail 1993), yet freshwater stickleback populations in the North and Baltic seas with access to the sea exhibit little morphological and genomic differentiation from marine stickleback indicating that gene flow is overriding selection (Ferchaud & Hansen 2016). Our collection from the Valdivia River estuary (ValEst) and from the Puelo River resident location (PueRes) comprised both diadromous and resident individuals and our admixture analysis suggests these two groups are at some extent reproductively isolated. Although it has been reported that diadromous and resident populations differ in their spawning time with diadromous individuals spawning in the fall and residents in the spring and summer (Pollard 1971), we have indeed observed diadromous individuals spawning in the spring as has also been reported by Cussac *et al.* (2004). Thus, though complete reproductive isolation is unlikely, we cannot discard temporal restriction to gene flow or "isolation by time" (Hendry & Day 2005) as being responsible, at least in part, for the lack of hybridization between sympatric populations. The presence of distinct genetic groups and high levels of genetic differentiation between sympatric populations suggest a role for divergent selection. Adaptive divergence has also been shown in other species exhibiting diadromous and resident forms such as prickly sculpin (*Cottus asper*) (Dennenmoser *et al.* 2017).

Most of the top 25 loci differentiating diadromous from resident populations are located within or nearby genes related to salinity adaptation including cation channels and solute exchangers (Table S3.3). Some of these genes have been reported through transcriptomic analyses to be potentially linked to osmoregulation in many species including Alewife (*Alosa pseudoharengus*) (Velotta *et al.* 2017) and Killifish (*Lucania parva*) (Kozak *et al.* 2014). However, it is important to point out that the SNP markers used in the present study were not specifically selected to find the adaptive genes that differentiate diadromous and resident individuals. Furthermore, the method used is a reduced representation sequencing approach. Many potential SNPs linked to genes involved in adaptation to the diadromous vs. resident freshwater life histories may have been missed. These BLAST results have a varying percentage of identity and coverage. These limitations suggest that most likely there are other genes not included in this analysis that may be linked to osmoregulation in *G. maculatus*. An annotated genome for this species or a closely related one will be required to specifically detect candidate genes locally adapted with their effects correlating with life history.

3.5.4 Methodological considerations, conservation implications, and future directions

Although useful for population genomics studies, RADseq is known to underestimate genetic diversity (Cariou *et al.* 2016). Further, the choice of filtering criteria when selecting SNPs has the potential to influence results (O'Leary *et al.* 2018). To overcome these difficulties, we tested different filtering schemes as shown in Table S3.1. The chosen filtering scheme based on PCA (i.e., FS3) minimized the percentage of missing data while retaining sufficient informative loci. A more relaxed filtering scheme (e.g. FS1) provided less power for population differentiation with, only one population (i.e., BioRes) clearly differentiated (data not shown). On the other hand, a stricter filtering scheme (e.g. FS5) produced the same PCA pattern as FS3 but retained only 969 loci.

Sample sizes were relatively low, particularly for the southernmost collections in the Serrano river (SerEst and SerRes) for which DNA quality was not optimal resulting in a relatively high number of missing data for these two collections. Although a recent study suggests that a low number of samples (i.e. 2 individuals) is sufficient to accurately

estimate genetic differentiation indices such as F_{ST} (Nazareno *et al.* 2017), increasing the sample size of some of the studied populations and including populations from other locations should lead to an improvement of our understanding of the differences between *G. maculatus* diadromous and resident populations.

Results suggest a clear differentiation among resident populations, particularly the northernmost populations (i.e., F_{ST} between resident populations up to 0.7), as well as between resident and diadromous populations from the same river system, thus we argue strongly for consideration of these differences in conservation initiatives. The range of *G. maculatus* in Chile has declined 26 %, and this decline largely involves the disappearance of populations in the northernmost part of the range (Habit *et al.* 2010). These northern river systems (i.e. Biobío river system) are impacted by hydropower developments that lead to habitat fragmentation and population isolation, affecting the migration of populations to spawning sites (Habit *et al.* 2019). Thus, considerations of the conservation status of northernmost resident populations are therefore strongly warranted.

Although our results suggest a strong role for local adaptation in the genetic differentiation between diadromous and resident populations, a better understanding of the relative role played by plasticity is critical to fully comprehend the extent to which genetics accounts for the presence of these life histories. Results from a reciprocal transplant experiment we recently conducted will likely provide further insight not only into the role of phenotypic plasticity but also the genes that are differentially expressed in diadromous and resident populations of *G. maculatus*.

3.6 Acknowledgements

Fieldwork was funded by a CONICYT Chile Grant to KG (Project 3130690) and a grant from the Committee for Research and Exploration of the National Geographic Society (9247-13) to DER. The genomics work was funded by an NSERC Discovery grant to DER. MLD was supported by the NSERC DG grant to DER, Dalhousie's Faculty of Graduate Studies Scholarship and by Nova Scotia Graduate Scholarship (NSGS). We would like to thank Aliro Manosalva, Jorge González, Nestor Ortíz, Juan Jose Ortíz, Francisco Navarro, and Tyler Zemlak for their assistance in the field, Greg McCracken for his help with the laboratory work, Ian Bradbury for providing NGS adapters, Angela Fuentes-Pardo and Tony Einfeldt for their assistance with bioinformatic

analyses, and Jon Puritz for adapting his dDocent pipeline to accommodate our data ("ROL" method).

3.7 Author contributions

This manuscript is part of MLD's PhD thesis. All authors contributed to the conception and design of the study. Samples were collected between 2006 and 2017 and all authors were involved in the field collections. Genomic data were generated and analyzed by MLD. The manuscript was written by MLD and DER with input from KG and EH.

3.8 Supplementary information

Table S3.1. Filtering schemes applied to raw data. The filters used were: percentage of missing data allowed per loci (geno missing), minimum depth (minDP), minor allele frequency (maf), loci under Hardy Weinberg equilibrium (HWE), loci under linkage disequilibrium (LD), and missingness per individual (indv missing). For each filter the remainder number of loci (SNPs) is presented and the total number of individuals (indv) that passed the filters.

	More relax			---------->									More conserve		
	FS1			**FS2**			**FS3**			**FS4**			**FS5**		
Filter	value	SNPs	indv	value	SNPs	indv	value	SNPs	indv	value	SNPs	indv	value	SNPs	indv
geno missing (%)	30	117109		25	101321		15	77383		15	77383		1	26385	
minDP	6	117109		10	101321		10	77383		10	77383		10	26385	
maf	0.01	49452		0.05	14367		0.01	23545		0.05	10892		0.05	1510	
HWE	yes	49136		yes	13905		yes	23348		yes	10891		yes	1508	
LD	yes	5573		yes	4172		yes	4388		yes	3492		yes	969	
indv missing (%)	20	5573	244	20	4172	225	20	4388	224	20	3492	224	20	969	224

Table S3.2. Pairwise $F_{ST\text{-}WC}$ matrix of 20 populations of *Galaxias maculatus* from Chile. $F_{ST\text{-}WC}$ estimates based on 3516 neutral loci.

	BioEst	BioRes	TolEst	TolRes	ValEst	ValRes	BueRes	PetEst	PueEst	PueRes	RblcEst	HuiRes	TarRes	CisEst	AysEst	AysRes	BakEst	BakRes	SerEst
BioRes	0.514																		
TolEst	0.004	0.485																	
TolRes	0.403	0.662	0.361																
ValEst	0.064	0.427	0.049	0.324															
ValRes	0.465	0.736	0.430	0.685	0.412														
BueRes	0.441	0.733	0.396	0.658	0.372	0.726													
PetEst	0.054	0.516	0.041	0.398	0.063	0.466	0.441												
PueEst	0.006	0.508	0.002	0.396	0.050	0.461	0.439	0.020											
PueRes	0.082	0.511	0.064	0.381	0.066	0.456	0.406	0.039	0.053										
RblcEst	-0.003	0.523	0.000	0.410	0.056	0.483	0.461	0.044	0.002	0.078									
HuiRes	0.152	0.560	0.126	0.425	0.110	0.513	0.460	0.131	0.137	0.039	0.148								
TarRes	0.292	0.649	0.256	0.553	0.241	0.626	0.592	0.278	0.278	0.190	0.296	0.185							
CisEst	0.009	0.519	0.009	0.403	0.048	0.484	0.461	0.026	0.000	0.055	0.001	0.133	0.283						
AysEst	0.065	0.505	0.048	0.376	0.056	0.451	0.406	0.047	0.044	0.036	0.057	0.083	0.232	0.038					
AysRes	0.188	0.583	0.168	0.458	0.156	0.527	0.483	0.167	0.171	0.133	0.187	0.164	0.307	0.170	0.079				
BakEst	0.005	0.507	0.002	0.393	0.051	0.462	0.436	0.039	0.003	0.062	0.002	0.129	0.268	0.004	0.045	0.169			
BakRes	0.194	0.574	0.162	0.451	0.146	0.535	0.479	0.169	0.176	0.132	0.192	0.147	0.288	0.173	0.110	0.165	0.167		
SerEst	0.096	0.552	0.088	0.433	0.091	0.525	0.481	0.091	0.082	0.086	0.092	0.135	0.291	0.078	0.066	0.170	0.080	0.171	
SerRes	0.151	0.618	0.150	0.525	0.112	0.642	0.636	0.165	0.154	0.135	0.157	0.203	0.401	0.137	0.128	0.249	0.141	0.245	-0.039

Table S3.3. Species name and potential functions of the proteins mentioned in Table 3.3.

Contig number	Protein description	Species name	Potential function
Contig_35840	LIM domain-containing protein ajuba-like	*Acanthochromis polyacanthus*	Cell fate determination and cytoskeletal organization
Contig_68102	Nuclear receptor subfamily 2 group F member 6	*Salvelinus alpinus*	Modulation of hormonal responses, and development of forebrain circadian clock
Contig_70236	Protocadherin-8 like	*Seriola dumerili*	Calcium-dependent cell-adhesion protein, may play a role in activity-induced synaptic reorganization
Contig_4613	Voltage-dependent calcium channel gamma-2-subunit	*Clupea harengus*	Stabilize the calcium channel in an inactivated (closed) state
Contig_23889	Tensin 4 (TNS4)	*Xiphophorus maculatus*	Cell migration and link signal transduction pathways to the cytoskeleton
Contig_73552	Myosin-If	*Larimichthys crocea*	Muscle contraction
Contig_18164	Thrombospondin type-1 domain-containing protein 7A-like	*Gadus harengus*	Role in actin cytoskeleton rearrangement
Contig_46934	Kinesin-like protein	*Salmo trutta*	Microtubule binding
Contig_50594	A disintegrin and metalloproteinase with thrombospondin motifs 10-like	*Lates calcarifer*	Metalloendopeptidase activity
Contig_50464	Proteasome subunit alpha type-6-like	*Sinocyclocheilus rhinocerous*	Numerous essential roles within the cell by associating with different regulatory particles
Contig_32554	Actin-binding LIM protein 3-like	*Gadus morhua*	Muscle fiber development
Contig_55409	Heat-stable enterotoxin receptor-like	*Gadus morhua*	Transduction of mitogenic signals from the cell membrane to the nucleus

Contig number	Protein description	Species name	Potential function
Contig_54325	Semaphorin-4B (SEMA4B)	*Oncorhynchus nerka*	Inhibition of axonal extension
Contig_65480	Synaptotagmin-7	*Larimichthys crocea*	Ca^{2+} sensor involved in Ca^{2+}-dependent exocytosis of secretory and synaptic vesicles
Contig_13463	Cullin-associated NEDD8-dissociated protein 1-like	*Lates calcarifer*	Promoter of the exchange of the substrate-recognition F-box subunit in SCF complexes
Contig_29616	Phosphatidylinositol phosphatase PTPRQ	*Salvelinus alpinus*	Protein tyrosine phosphatase activity
Contig_37112	Sidekick cell adhesion molecule 1 (SDK1)	*Paralichthys olivaceus*	Homophilic cell adhesion via plasma membrane adhesion molecules
Contig_19277	Potassium voltage-gated channel subfamily KQT member 5-like	*Gadus morhua*	Voltage-gated potassium channel activity
Contig_41866	Inositol 1,4,5-trisphosphate receptor type 1 (Itpr1)	*Oncorhynchus nerka*	Intracellular channel that mediates calcium release from the endoplasmic reticulum
Contig_57173	Inactive dipeptidyl peptidase 10	*Esox lucius*	Promoter of cell surface expression of the potassium channel KCND2
Contig_71977	Adenylate kinase 9 (AK9)	*Mastacembelus armatus*	Maintenance of the homeostasis of cellular nucleotides
Contig_39334	PAXIP1 associated glutamate-rich protein 1 (PAGR1)	*Anabas testudineus*	Association with the histone methyltransferase complex, role in epigenetic transcriptional activation
Contig_7588	Sodium/potassium/calcium exchanger 1-like	*Notothenia coriiceps*	Transmembrane transport
Contig_20640	Sodium-independent sulfate anion transporter-like	*Myzus persicae*	Secondary active sulfate transmembrane transporter activity
Contig_69735	High affinity choline transporter 1-like	*Seriola lalandi*	Choline: sodium symporter activity

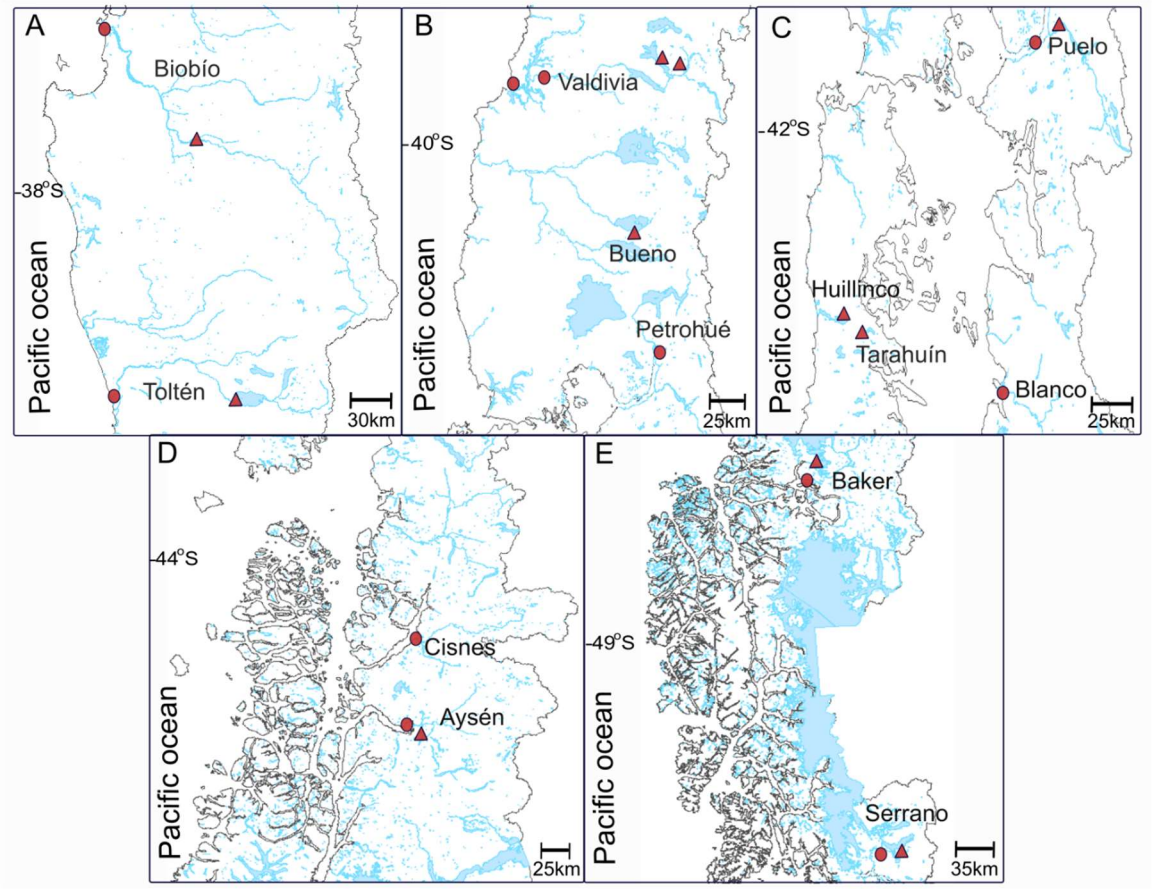


Figure S3.1. Detailed maps that highlight the landscape of the area and closeness to the sea. Circles represent estuarine populations, and triangles represent resident populations. **A.** Biobío and Toltén locations. **B.** Valdivia, Bueno and Petrohué locations. **C.** Puelo, Huillinco and Tarahuín locations. **D.** Cisnes and Aysén locations. **E.** Baker and Serrano locations.

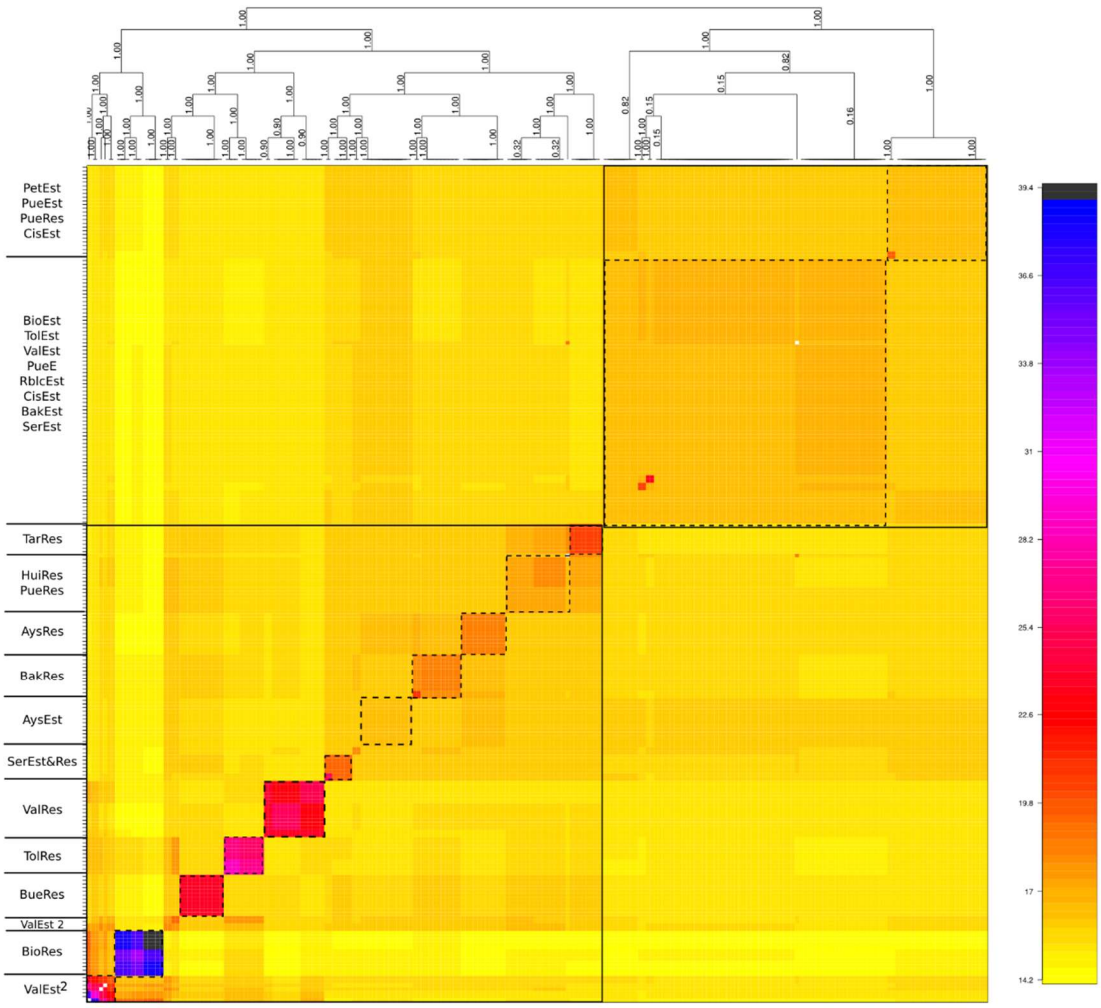


Figure S3.2. Neutral loci - fineRADstructure coancestry matrix using 3516 neutral loci and 224 individuals from 20 populations. Color bar represents the estimates of coancestry. Genetic groups (diadromous and residents) are marked by a solid black line. Substructure is outline by dashed lines. The population's names which individuals belong for each genetic group are in the left side of the matrix.

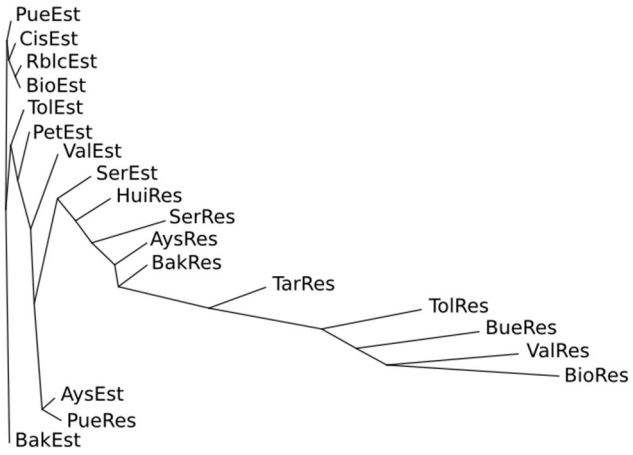


Figure S3.3. Distance tree based on F_{ST} estimates of 20 *G. maculatus* populations. F_{ST} estimates were obtained from the analysis of 3516 neutral loci.

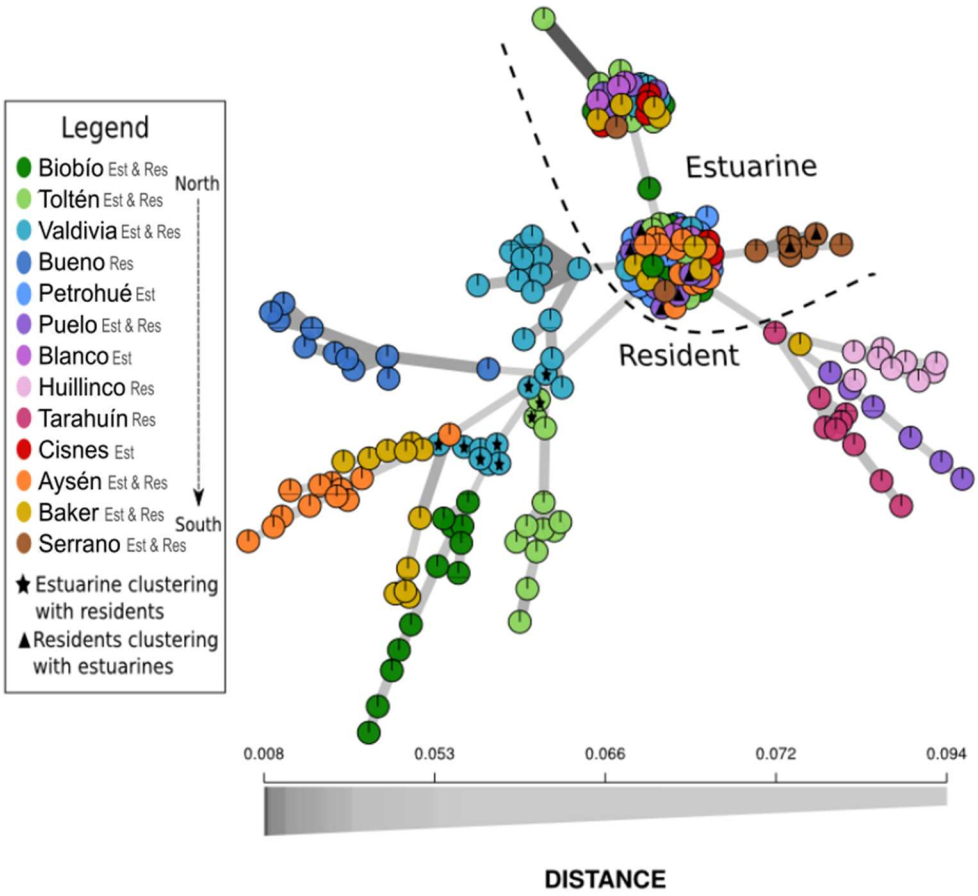


Figure S3.4. Minimum spanning network of 224 *G. maculatus* individuals using 3516 neutral loci. Each circle represents an individual which color represents the river system of origin. Individuals above the dashed line are estuarine and individuals below are resident, except for individuals marked (star or triangle) as explain in the legend.

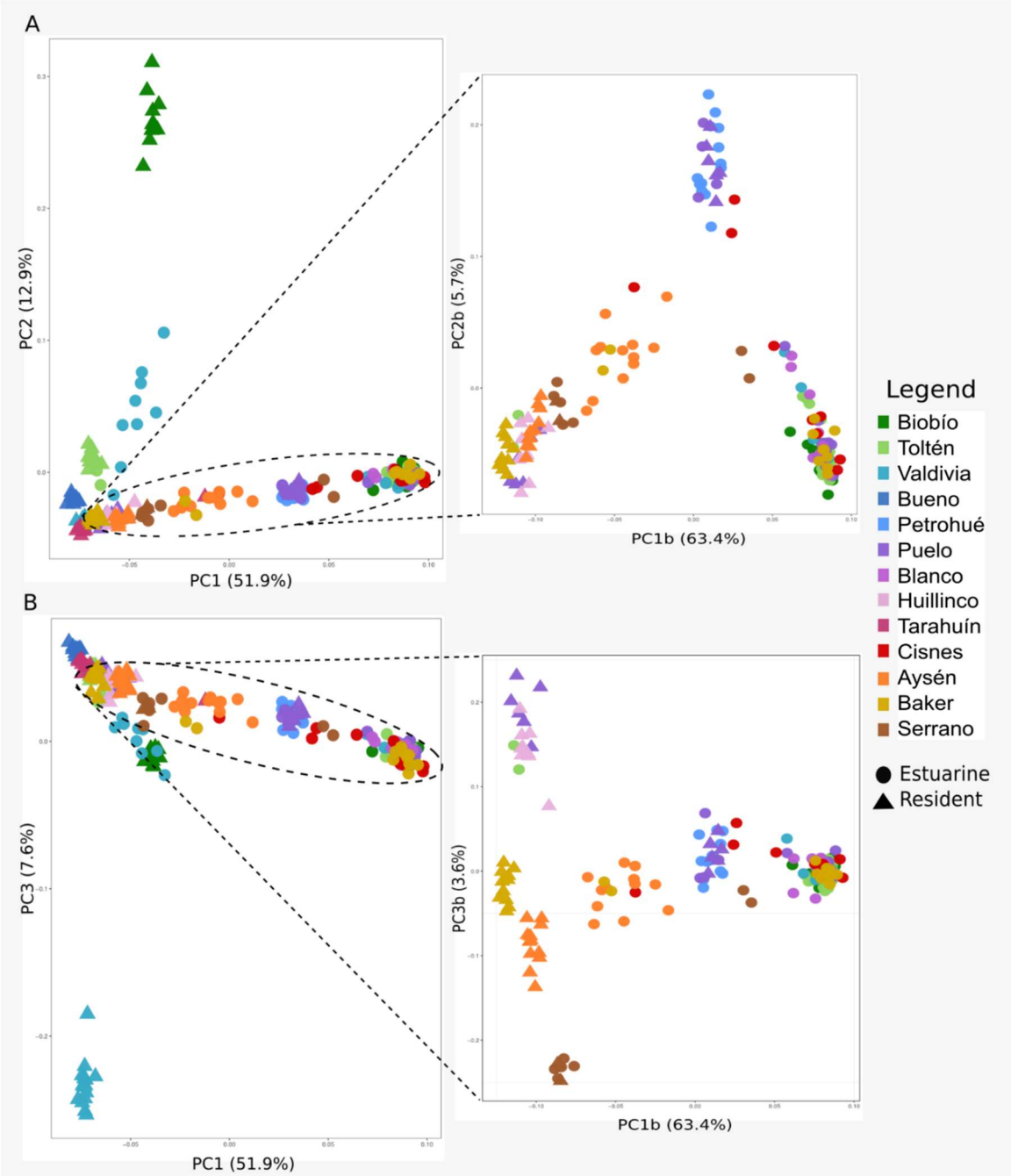


Figure S3.5. PCA performed on 224 individuals from 20 populations of *G. maculatus* using 872 outlier loci. Each color represents a unique river system. Circles represent estuarine individuals, and triangles resident individuals. **A**. The first and second PCA axes. **B**. The first and third PCA axes. Next to each graph, there is a PCA performed on the individuals inside the ellipse (156 individuals of 15 populations).

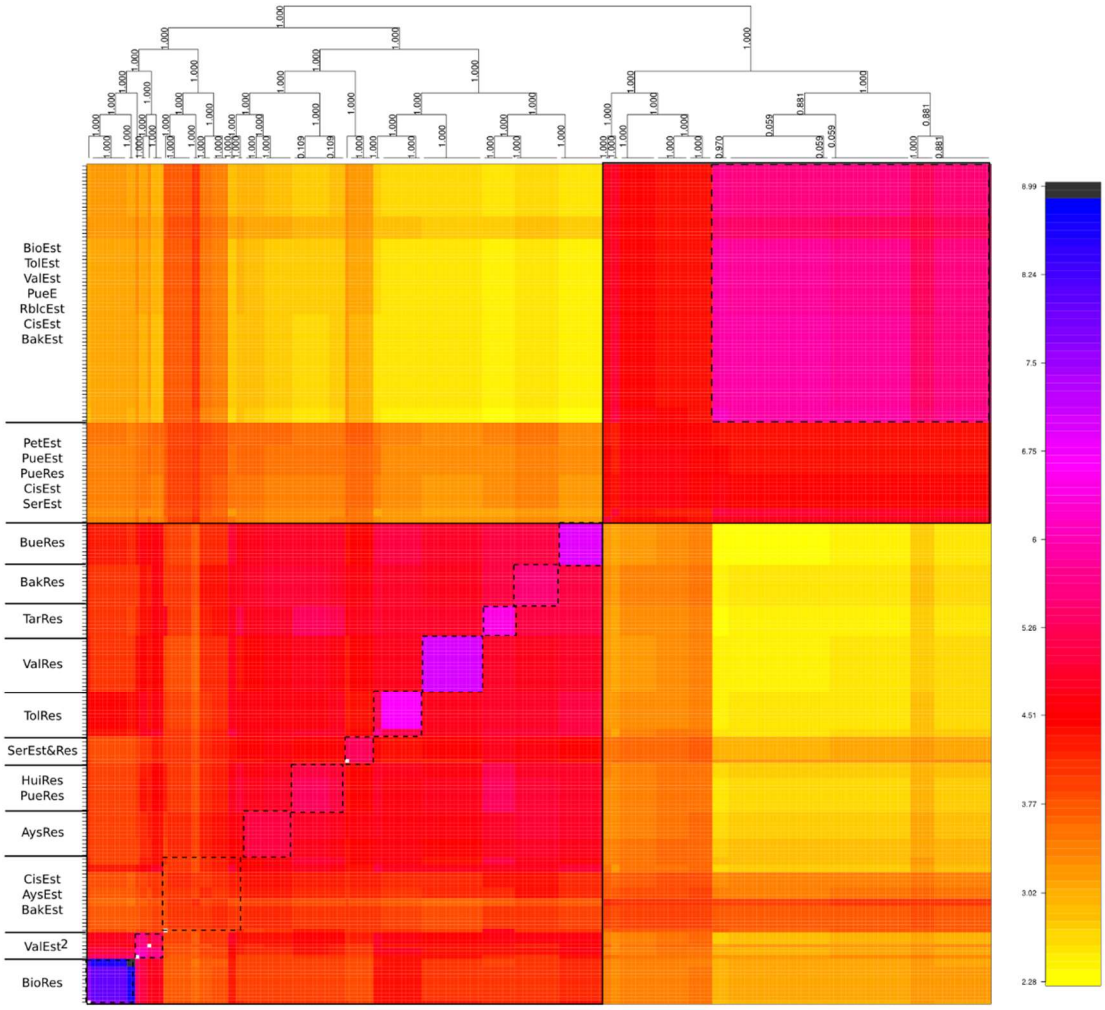


Figure S3.6. Outlier loci - fineRADstructure coancestry matrix using 872 outlier loci and 224 individuals from 20 *G. maculatus* populations. Color bar represents the estimates of coancestry. Genetic structure (diadromous and residents) is marked by a black solid line. Substructure is outline by dashed lines. The population's names which individuals belong for each genetic group are in the left side of the matrix.

CHAPTER 4. GENOMIC BASIS OF THE LOSS OF DIADROMY IN GALAXIAS MACULATUS: INSIGHTS FROM RECIPROCAL TRANSPLANT EXPERIMENTS

This chapter has been published as "ML Delgado, A Manosalva, M Urbina, E Habit, O Link & DE Ruzzante. 2020. Genomic basis of the loss of diadromy in Galaxias maculatus: Insights from reciprocal transplant experiments. Molecular Ecology. 29(24):4857-4870."

4.1 Abstract

Diadromy is known for having major effects on the distribution and richness of aquatic species, and so does its loss. The loss of diadromy has led to the diversification of many species, yet research focusing on understanding its molecular basis and consequences are limited. This is particularly true for amphidromous species despite being the most abundant group of diadromous species. *Galaxias maculatus*, an amphidromous species and one of the most widely distributed fishes in the Southern Hemisphere, exhibits many instances of non-migratory or resident populations. The existence of naturally replicated resident populations in Patagonia can serve as an ideal system for the study of the mechanisms that lead to the loss of the diadromy and its ecological and evolutionary consequences. Here, we studied two adjacent river systems in which resident populations are genetically differentiated yet derived from the same diadromous population. By combining a reciprocal transplant experiment with genomic data, we showed that the two resident populations followed different evolutionary pathways by exhibiting a differential response in their capacity to survive in salt water. While one resident population was able to survive salt water, the other was not. Genomic analyses provided insights into the genes that distinguished 1) migratory from non-migratory populations, 2) populations that can vs. those that cannot survive a saltwater environment, and 3) between these resident populations. This study demonstrates that the loss of diadromy can be achieved by different pathways and that environmental (selection) and random (genetic drift) forces shape this dynamic evolutionary process.

4.2 Introduction

Diadromy, the predictable movements between freshwater and marine environments, is known for affecting the distribution and richness of aquatic species, and so does its loss. Diadromous fishes leave their natal environment (freshwater or marine) for varying periods according to the category of diadromy and return to their natal environment to continue growing and/or to reproduce (McDowall, 2001). This migratory behavior facilitates the colonization of new habitats and its loss is a driver for diversification and even speciation, influencing the biodiversity of many marine and freshwater systems (McDowall 1998; Burridge & Waters 2020).

Diadromy is classified into three types: anadromy, catadromy, and amphidromy; they differ in the environment in which reproduction occurs, the life stage in which migration takes place, and the length of time spent in the alternate environment (McDowall, 1997). Many diadromous species of all three types have non-migratory or resident populations, i.e., populations that remain their entire life cycle and complete their reproductive cycle in their natal environment (McDowall 2001). Although genomic differences have been described between diadromous and resident populations in various species including Three-spined Stickleback (*Gasterosteus aculeatus*) (e.g. Drevecky, Falco, & Aguirre, 2013), Atlantic Salmon (*Salmo salar*) (e.g. Perrier, Bourret, Kent, & Bernatchez, 2013), Rainbow Trout (*Oncorhynchus mykiss*) (e.g. Bowersox, Wickersham, Redfield, & Ackerman, 2016), Brown Trout (*Salmo trutta*) (e.g. Lemopoulos, Uusi-Heikkilä, Huusko, Vasemägi, & Vainikka, 2018), and Arctic Charr (*Salvelinus alpinus*) (e.g. Salisbury *et al.* 2018), the evolutionary basis and consequences of the loss of diadromy are not fully understood. Model species such as the Three-spined Stickleback have, however, provided crucial information on the roles of genetics and the environment in the evolution of resident populations. For example, a common garden experiment with Three-spined Stickleback populations that inhabit environments with different salinity levels revealed signs of local adaptation, where the population native to low salinity performed poorly in high salinity (Defaveri & Merila 2014). Further studies with Three-spined Stickleback resident populations in British Columbia demonstrated that resident populations can exhibit more genes with plastic responses than do anadromous populations (Morris *et al.* 2014), yet the prevalence of plastic traits can vary among

populations (Oke *et al.* 2016). Although some traits such as gill raker number declined in a parallel and predictable way among freshwater Three-spined Stickleback populations (Glazer *et al.* 2014), migratory behavior was not lost in a similar manner across resident populations. Thus, further studies in other diadromous species appear necessary.

Freshwater amphidromy is the most common type of diadromy, and it characterizes fishes born in rivers, which then migrate to the sea as larvae for a short period, usually from several days to a few weeks, before returning to fresh water as juveniles to continue their growth and subsequent reproduction (McDowall 2007). Despite the high number of species in this group, it is the least studied (Augspurger *et al.* 2017). *Galaxias maculatus*, a widespread amphidromous species with a Gondwanan distribution (Berra *et al.* 1996) exhibits migratory as well as resident populations across its distribution. Phylogeographic studies with *G. maculatus* suggest that the diadromous life history is the ancestral state. Diadromy is in fact the life history trait that allowed *G. maculatus* to disperse from New Zealand to Australia, and later to South America via the West Wind Drift (Waters & Burridge 1999; Vera-Escalona *et al.* 2020). Recent genomic data indicate that diadromous *G. maculatus* populations in Chile are highly differentiated from their resident counterparts (i.e. freshwater population inhabiting the same river system) (Fixation index that assesses population differentiation was moderate to high = F_{ST}: 0.2-0.6) (Chapter 3; Delgado *et al.* 2019).

Galaxias maculatus diadromous individuals generally exhibit a high tolerance to gradual changes in salinity (Lethal Dose, LD_{50}=62 ppt) and even to abrupt changes (LD_{50}=45 ppt) (Chessman & Williams 1975) due to its rapid activation of molecular responses (Urbina *et al.* 2013). Studies on the physiology and adaptation to salinity on *G. maculatus* resident populations are scarce. A recent study comparing estuarine and resident individuals in a laboratory setting showed that residents have a lower tolerance to salinity (i.e. 33% survival at 25 ppt) than estuarine population (80% survival at 25 ppt) (Ruiz-Jarabo *et al.* 2016). The authors also found that resident individuals did not increase branchial H^+ -ATPase (HA) activity at any salinity compared to estuarine individuals (Ruiz-Jarabo *et al.* 2016). HA, among other functions, is involved in sodium (Na^+) uptake (Potts 1994). Because freshwater environments are generally poorer in Na^+ than are estuarine environments, freshwater populations are expected to prime their

branchial HA so as to ensure Na^+ uptake at much lower environmental concentrations than their estuarine counterparts.

Multiple studies have demonstrated that common garden and reciprocal transplant experiments between contrasting environments can assist in the study of local adaptation and phenotypic plasticity (de Villemereuil *et al.* 2015). Thus, here we studied two adjacent river systems that have highly differentiated populations of *G. maculatus.* Both systems have lake populations which have been assessed as resident by bone tissue microchemistry (Górski *et al.* 2018). We aimed to answer four questions: 1) Do these resident populations differ in their individuals' ability to acclimate to salt water (25ppt)? 2) Do individuals show similar levels of plasticity across both resident populations? 3) What genes or genomic markers distinguish a) migratory vs. non-migratory individuals, b) individuals that can and cannot acclimate to salt water, and c) individuals from both resident populations? 4) Are the genes that differentiate diadromous and resident individuals common to other migratory species? To answer these questions, we combine the information obtained from a reciprocal transplant experiment with genomic data.

4.3 Materials and methods

4.3.1 Study area

Two river systems, Toltén and Valdivia, from the northern section of the species distribution in Chile were studied (Fig.4.1). Both river systems originate in the Andes (Zemlak *et al.* 2010). The Toltén River system has a length of 123 km from the estuary to Lake Villarica. Lake Villarica is located 230 m above sea level and has a surface area of 173 km^2 (Ministerio de Obras Públicas 2004). The Valdivia River system is larger than the Toltén system and includes a series of lakes. Samples were collected from Lake Neltume, which is located 186 m above sea level and has a surface area of 9.8 km^2 (Ministerio de Obras Públicas 2004b). There are no physical barriers (i.e. hydroelectric dams) that would restrict migration between the fish collection points in either river system.

Figure 4.1 Collection sites (estuary and lake) within the Toltén and Valdivia River systems.

4.3.2 Reciprocal transplant

Adult *G. maculatus* were collected using nets and electrofishing from the estuary (Toltén: Lat. -39.25 Long. -73.22; Valdivia: Lat. -39.85 Long. -73.33) and one lake in each of the two river systems, Lake Villarica (Lat. -39.28 Long. -72.23) in the Toltén river system, and Lake Neltume (Lat. -39.78 Long. -71.96) in the Valdivia River system (Fig. 4.1). Individuals were placed in 80 L containers with air pumps filled with water from their site of collection and brought to the Universidad de Concepción within 2 days of capture. Individuals were collected first from the Valdivia River system, and 4 days later from the Toltén River system. Salinity at the collection sites was measured with a field conductivity meter (ThermoFisher, Orion Star A222). Estuarine locations exhibit a range of salinity levels from close to 0 ppt to 30 ppt depending on the tides, while salinity in the lakes was 0ppt. Estuarine individuals were collected during low tide, thus the salinity at the time was low (~1 ppt). Around 130-140 individuals were collected from each location, mortality during the transfer to the laboratory was <4%.

Upon arrival, individuals were distributed among four 65 L tanks, each tank holding approximately 35 individuals (Fig 4.2). While resident adults were placed in tanks at 0 ppt, estuarine adults were placed in tanks at 25 ppt. Two days after arrival, 20

individuals were chosen randomly from each of the four locations and placed in either control or treatment tanks (Fig. 4.2). Each treatment had three replicate 65 L tanks with a fish density of 0.16 g fish/L. Water temperature, pH, ammonia, and nitrate levels were checked regularly. Thermometers in each tank were used to check the temperature daily. Commercial kits (©API) were used to test pH, ammonia, and nitrate levels weekly. Temperature was maintained at 18 ± 1° C, pH ranged between 7.2 – 7.6, ammonia levels registered were <0.25 (mg/L), and nitrate levels were 0 (mg/L). Tanks were kept under natural light and a feeding regime of brine shrimp and flakes once daily was followed.

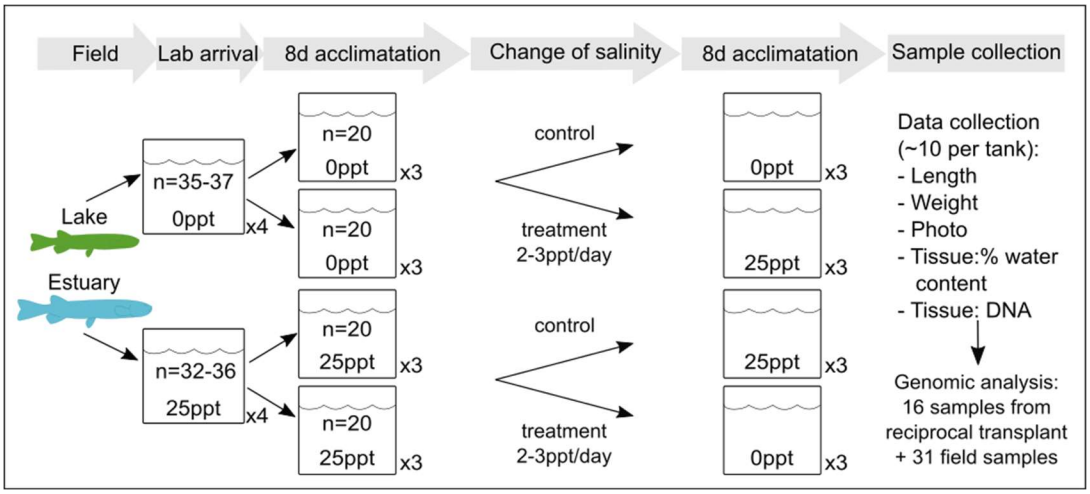


Figure 4.2 Reciprocal transplant experiment design performed for each river system. Upon arrival to the lab, four holding 65L tanks were set up each containing ~35 individuals, the 8 days of acclimation started once individuals were randomly distributed in tanks containing only 20 individuals. The number of samples in the genomic analysis is for the whole experiment (both river systems).

After eight days of acclimation to either 0 or 25 ppt, salinity was gradually changed. Salinity was kept unchanged in the control tanks. Salinity was increased or decreased around 2-3 ppt per day for a total of 8 days (Fig. 4.2). Only 20% of water was removed per day for both the treatments and control. The amount of fresh or salt water replaced was estimated using an online calculator (https://www.hamzasreef.com/Contents/Calculators/TargetSalinity.php). Salt water was prepared by adding synthetic salt (©Instant Ocean) to chlorine-free water, and the salinity was measured for each tank after every change.

Tanks were checked daily, and dead fish were removed. Eight days after the target salinity was reached, ~10 individuals per tank (depending on the number that survived) were sacrificed with an overdose of benzocaine (100mg L^{-1}), measured (i.e. total length), weighed, and photographed. We also took two samples of muscle tissue from each individual. One muscle sample was placed in ethanol for DNA analysis; the second sample was placed on a pre-weighed 1.5 mL microtube for muscular water content calculations. The percentage of water content provides information on whether or not a fish is struggling by experiencing a net loss of water and a gain of ions. This information is relevant when exposing fish to elevated salinities. The wet and dry weights of the tissues were measured and the percentage of water content in muscle was calculated using the following equation.

$$\text{(eq. 4.1)} \qquad \%\text{Water Content} = \frac{W1 - W2}{W1\ x\ 100}$$

Where W1 is the weight of wet tissue and W2 is the weight of dry tissue.

Fulton's condition factor (K) was also calculated, to assess the nutritional condition or "robustness" of the individuals:

$$\text{(eq. 4.2)} \qquad K = W/L^3 \text{ x}100$$

Where W is the wet weight in grams of the individual and L is the total length in cm.

All experimental procedures were performed according to Chilean guidelines on animal care and approved by the Ethics, Bioethics, and Biosafety Committee of the Universidad de Concepción and the University Committee on Laboratory Animals (UCLA) of Dalhousie University.

4.3.3 Genomic data

Our previous study combined bone tissue microchemistry and genomic analysis to confirm that estuarine individuals were diadromous and that lake individuals were residents (Chapter 3; Delgado *et al.* 2019). Genomic data were obtained using a RADcap approach, which is a variant of the reduced representation method RADseq. This method allows the sequencing of thousands of DNA fragments located near restriction enzyme cut sites (Hoffberg *et al.* 2016). RADcap raw sequence data were retrieved from the previous study (N= 47 individuals, 11 and 15 individuals were residents in the Toltén and Valdivia River system, respectively, and 21 were diadromous individuals, 16 and 5 from

the estuaries of Toltén and Valdivia, respectively) (Chapter 3; Delgado *et al.* 2019). Sixteen of these 47 individuals were involved in the reciprocal transplant experiment (Toltén diadromous = 5, Toltén resident = 3, Valdivia diadromous = 5, Valdivia resident = 3).

4.3.4 Data analysis

Analyses regarding size, weight, condition factor (K), and survival data per population were conducted on R (R Core Team 2020). As the data did not follow a normal distribution, Wilcoxon tests followed by Bonferroni correction were performed to compare the median between control and treatment tanks and between estuary and lake collection. Wilcoxon tests and boxplots were conducted with the rstatix (Kassambara 2020a) and ggbur (Kassambara 2020b) packages. Survival differences between control and treatment were assessed using the Kaplan-Meier method, a non-parametric statistic to estimate survival function. The tests and graphs were performed using the survival (Therneau *et al.* 2020) and ggplot2 (Wickham 2016) packages in R. Survival and the percentage of water content at both salinities were plotted using the ggplot2 package (Wickham 2016).

Raw SNP data were filtered using VCFtools (Danecek *et al.* 2011). Filters removed SNPs with >25% missing data per site, a minimum depth <10, and a minimum allele frequency (MAF) <0.01. Hardy-Weinberg equilibrium (HWE) by population was tested using a perl script https://github.com/jpuritz/dDocent/blob/master/scripts/filter_hwe_by_pop.pl) with the p-value threshold set to 0.05. Only the first SNP per contig was retained to remove physically linked loci. Tests for linkage were performed using VCFtools with the default settings (r^2 implied 0.9). Finally, individuals with >20 % missing data were removed. The number of SNPs removed at each step is shown in Table S4.1. VCFtools, PLINK (Purcell *et al.* 2007), and PGDSpider (Lischer & Excoffier 2012) were used to convert to the appropriate formats. Principal Component Analyses (PCA) was conducted with PLINK.

Three pairwise comparisons were performed: 1) between diadromous and Toltén residents, 2) between diadromous and Valdivia residents, and 3) between the two resident populations. Candidate loci under selection for each comparison were detected using

PCAdapt (Luu *et al.* 2017), where the q-value was set to < 0.01, and Bayescan (Foll & Gaggiotti 2008) where the false discovery rate (FDR) was set to < 0.05.

Common and unique SNPs were detected using the "comm" and "uniq" bash commands. All contigs containing outlier SNPs were blasted using the BLASTN search tool (NCBI). BLAST results were filtered to include only hits belonging to an animal species with an e-value <0.01, a query coverage ≥20, a percentage of identity >70%. The biological role of these potential genes under selection was searched in the Uniprot protein database (www.uniprot.org). A literature search was conducted to find genes differentiating diadromous/migratory vs. resident/non-migratory life histories in other fish species.

4.4 Results

4.4.1 Reciprocal transplant

Significant phenotypic differences in size and weight were observed between individuals from the estuaries and lakes (Fig. S4.1). Estuarine adults were heavier and larger than residents from Toltén and Valdivia River systems. The condition factor (K) showed that estuarine individuals were more robust than resident individuals (Fig. S4.1). Valdivia resident individuals were slightly heavier and larger than Toltén resident individuals (Fig. S4.1). There were no significant differences in size between individuals in the control and treatment tanks. Weight, however, differed slightly between control and treatment individuals from the freshwater lake populations, particularly in the Toltén resident population, but not between control and treatment individuals from the estuarine collections (Fig. S4.1).

The response to salinity changes varied between populations. The estuarine individuals exhibited 100% survival in both fresh water (0 ppt) and salt water (25 ppt) (Fig. 4.3). The Toltén and Valdivia resident individuals differed in their survival rate, exhibiting some mortality in fresh water (i.e., 7.5% and 17.5%, respectively). When exposed to salt water, however, only ~23% of Toltén residents survived at 25 ppt, while 100% of Valdivia residents survived the increment in salinity (Fig. 4.3).

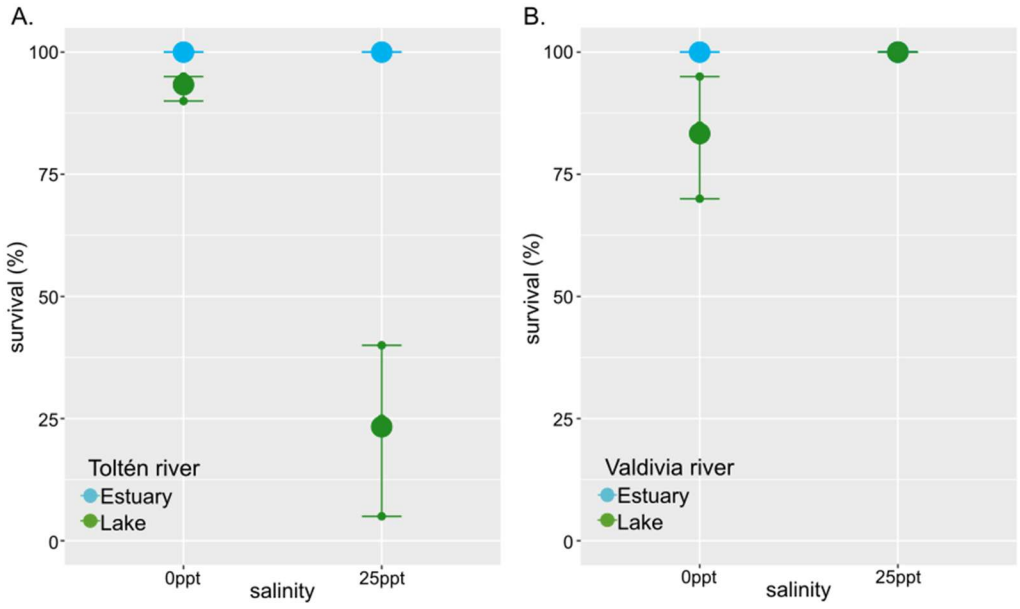


Figure 4.3 Survival at the end of the experiment (8 days after treatment reach their set salinity) at two salinity concentrations (0 and 25ppt) of individuals from collected at the estuary and lake of **A.** Toltén, and **B.** Valdivia River systems.

The Toltén residents started dying when salinity reached 20 ppt (Fig. S4.2). Once salinity reached 25 ppt, approximately 40% of the individuals died within the first two days (Fig. S4.2). Yet, the ~23% resident Toltén individuals that survived exhibited a range of sizes and were generally not the largest individuals. After 8 days at 25 ppt salinity, these Toltén resident survivors showed a slight increase in the percentage of water content in muscle relative to estuarine individuals (Fig. S4.3). Similarly, the percentage of water content among Valdivia residents increased slightly at 25 ppt relative to estuarine individuals (Fig. S4.3).

4.4.2 Genomic analyses

We obtained a total of 5154 SNPs after filtering (see Table S4.1). Two individuals (one Toltén resident and one Valdivia resident, both collected from the field) showing ~30% of missing data were removed; thus, bringing the number of genotyped individuals used in subsequent analyses to 45. Three distinct genetic groups were observed, one comprising all estuarine individuals, and one for each of the resident populations. A neighbor-joining tree using all neutral and outlier markers confirmed that diadromous

individuals clustered into a single group and that the two resident populations are highly distinguishable from this diadromous population (Fig. 4.4).

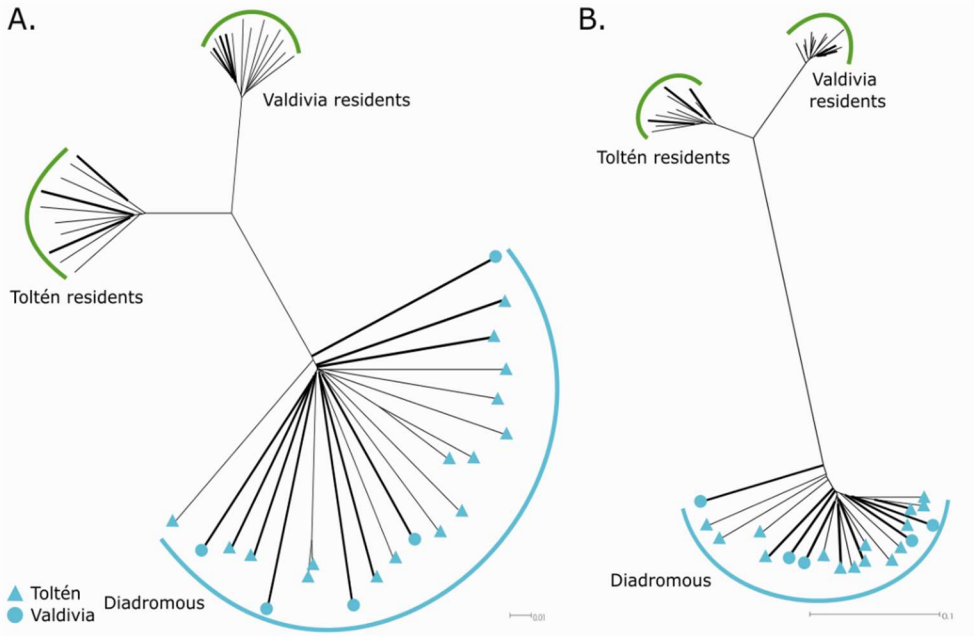


Figure 4.4 Neighbor-joining tree of 45 individuals from 2 river systems (Valdivia and Toltén) using **A.** 5154 neutral and outlier loci and **B.** 523 outlier loci. Both estuarine locations conformed to a single genetic group (diadromous). Bold branches reflect the individuals involved in the reciprocal transplant (N=16).

Overall, 523 SNPs were detected as outliers using two detection methods (PCAdapt and Bayescan), with only 11 SNPs common to both methods. PCA plots using all 523 outlier SNPs showed that diadromous individuals are the most highly differentiated of the three groups. PC1 (59.1% of variance) distinguishes the diadromous individuals from either of the resident populations with Valdivia resident individuals being less differentiated from the diadromous individuals than the Toltén resident individuals. PC2 (11.8% of variance) differentiates the two resident populations (Fig. S4.4). A neighbor-joining tree using only outlier loci showed a similar layout as the tree built with all loci, with the differentiation among the three genetic groups being more marked (Fig. 4.4).

Pairwise, 287 SNPs differentiated the diadromous from the Valdivia resident individuals, while 300 SNPs distinguished the diadromous from the Toltén resident individuals. Only 144 SNPs were common between these two comparisons (Fig. 4.5A).

Forty-three outlier SNPs distinguished the populations that are able to survive in salt water (i.e., diadromous & Valdivia residents) from that which is not able to do so (i.e., Toltén residents) (Fig. 4.5B). Additionally, 99 SNPs distinguished the two resident populations from each other but were not involved in the differentiation with the diadromous population (Fig. 4.5C).

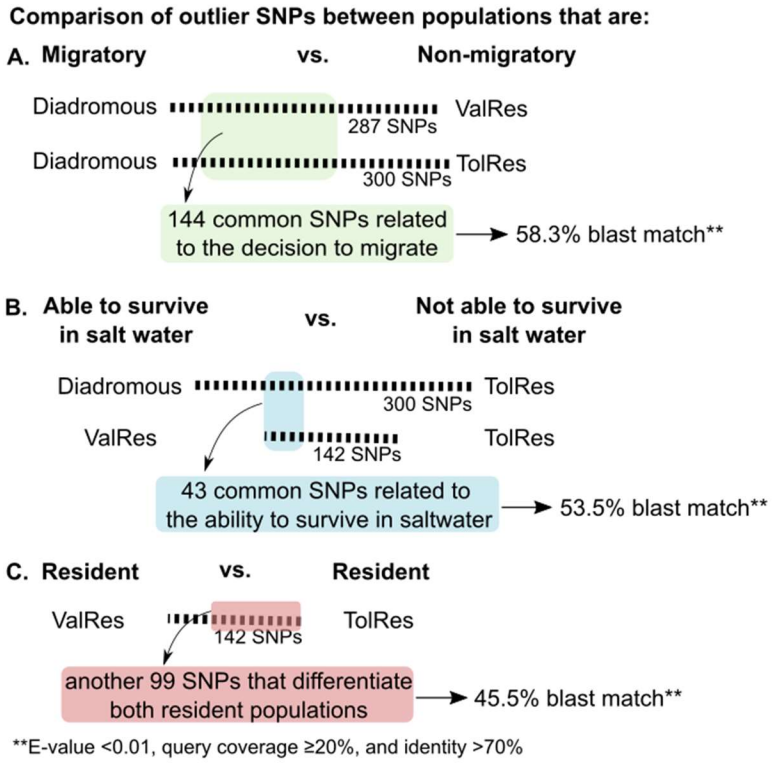


Figure 4.5 Number of outlier loci found when comparing **A.** migratory vs non-migratory populations, **B.** populations that can survive in salt water vs. a population that cannot, and **C.** resident populations.

A BLAST search found that, on average, ~52% matched an annotated gene after selecting a relaxed threshold of E-value <0.01, query coverage ≥20%, and identity >70% (Fig. 4.5). The BLAST analyses generated 84 genes that differentiated migratory from non-migratory populations (Table S4.2), 22 genes that differentiated between individuals that were able and unable to survive in salt water (Table S4.3), and 45 genes that were distinct only between the two resident populations (Table S4.4).

The search in the Uniprot database revealed the different functions these genes are involved in. A list of biological processes associated with these genes is reported for each of the comparisons: migratory vs. non-migratory individuals (Table 4.1A), ability vs. lack

thereof to survive in salt water (Table 4.1B), and resident populations of different lineages (Table 4.1C). These biological processes were classified into cellular activity/processes, energy/metabolism, regulation of gene expression, immune response, osmoregulation, response to stimuli, muscle function, and growth (Table 4.1).

Biological processes found in other comparisons that stand out were different regulators of gene expression and methylation, different genes involved in response to light stimuli, and locomotor behavior. Surprisingly, only one gene (i.e., acid-sensing ion channel subunit gene) found to distinguish individuals that survived in salt water from those that do not, seems to be directly involved in osmoregulation. However, many biological processes related to osmoregulation were found to differentiate migratory and non-migratory populations.

Thirty-one loci of the 84 found to distinguish diadromous from resident individuals have been reported previously in other fish species (Table 4.2). These 31 genes have been associated with differentiation between migratory and non-migratory ecotypes of other diadromous species and have also been linked with salinity tolerance, salinity adaptation, and other processes including nutrient intake and toxicity.

Table 4.1. Biological processes associated with outlier loci found to differentiate **A.** migratory vs non-migratory populations, **B.** populations that can survive in saltwater vs. a population that cannot, and **C.** resident populations (see Fig. 4.5).

A. Loci differentiating migratory and non-migratory populations	B. Loci differentiating populations that can and can't survive at 25ppt	C. Loci differentiating both resident populations
Cellular activity/processes		
Cell-cell adhesion	Cell-cell adhesion	Cell-cell adhesion
Cell differentiation including fatty cells	Cell differentiation	Cell differentiation
Cell migration	Cell migration	Cell migration
Cell-cell signaling	Protein transport	Protein transport
Angiogenesis	Vesicle-mediated transport	Angiogenesis
Regulation of signal transduction	Cellular response to DNA damage stimulus	Ribosome biogenesis
Activation of GTPase activity		Rab protein signal transduction
DNA repair		Signal transduction
		DNA damage checkpoint
Energy/metabolism		
Lipid biosynthesis process	Lipid catabolic process	Catabolic process including lipid
Protein biosynthesis	Glutamate metabolic process	Phosphatidic acid biosynthesis process
Fatty acid metabolic process	Glucose homeostasis	Adipose tissue development
Carbohydrate metabolic process		
Energy homeostasis		
Catabolic process		
Hydrolase activity		
Bile acid and bile salt transport		
Gene expression		
Regulation of transcription by RNA polymerase II	Regulation of transcription by RNA polymerase II	
Regulation of gene expression	Chromatin remodeling	Regulation of gene expression
		Regulation of histone methylation

A. Loci differentiating migratory and non-migratory populations	B. Loci differentiating populations that can and can't survive at 25ppt	C. Loci differentiating both resident populations
Regulation of alternative mRNA splicing	Sequence-specific DNA binding	Activation of MAPK activity
Histone acetylation		
Histone methylation and demethylation		
Immune response		
Innate immune response	Adaptive immune response	Innate immune response
Regulation of defense response to bacteria		Protein ubiquitination
Osmoregulation and intracellular transport		
Ion transport including potassium transport	Ion transport	Ion transport including calcium transport
Regulation of calcium ion transport		Regulation of sodium ion transport
Regulation of sodium ion transport		Symporter activity
Maintenance of tissue homeostasis		Ion homeostasis
Response to stimulus		
Detection of a stimulus (sensory perception)	Sensory perception to light stimuli	Visual perception
Eye photoreceptor cell development	Circadian rhythm	Circadian rhythm
Regulation of circadian rhythm	Response to cold	Cellular response to heat
Regulation of rhodopsin mediated signaling pathway		Neurotransmitter receptor transport
		Microtubule depolymerization
Others		
Locomotor behavior	Regulation of muscle contraction	Regulation of muscle contraction
Muscle fiber development		Regulation of synapse assembly
Regulation of cell growth		Regulation of growth

Table 4.2. Genes found to differentiate diadromous and resident populations of *G. maculatus* that have been previously reported to be important in the migratory behavior of other fish species.

Associated to	Gene	Species	Reference
Fatty acid oxidation during fasting	Probable palmitoyltransferase	*Oncorhynchus mykiss*	(Morash & McClelland 2011)
High-affinity Na+ intake	Acid-sensing ion channel 2	Euryhaline fishes	(Edwards & Marshall 2012)
Maturity and running time	Pecanex 4	*Salmo salar*	(Erkinaro *et al.* 2018)
Migratory vs. non-migratory ecotypes	Dystonin	*Oncorhynchus mykiss*	(Hale *et al.* 2013)
	E3 ubiquitin-protein ligase, Zinc fingers, Lysine-specific demethylase & Trichohyalin	*Oncorhynchus mykiss*	(Arostegui *et al.* 2019)
	Elongation factor (EF)	*Alosa pseudoharengus*	(Czesny *et al.* 2012)
	Homeobox 2, Protocadherin 16-like & Thrombospondin-like	*Oncorhynchus mykiss*	(Hale *et al.* 2013)
	L-Lactate dehydrogenase A chain	*Salvelinus fontinalis*	(Crespel *et al.* 2017)
	Pleckstrin homology domain A7	*Gasterosteus aculeatus*	(Smith *et al.* 2015)
Salinity adaptation	Cadherin 18-like & Collagen type V alpha 2, Nuclear receptor ROR beta-like, Ryanodine receptor, Solute carrier family & Signal-induced proliferation-associated 1-like	*Lateolabrax maculatus*	(Zhang *et al.* 2017)
	Histone-N-methyltransferase	*Fundulus heteroclitus*	(Brennan *et al.* 2018)
	Serine/threonine-protein kinase	*Fundulus heteroclitus*	(Brennan *et al.* 2018)
	Sodium/potassium/calcium exchanger	*Lucania parva*	(Kozak *et al.* 2014)
	Lipoxygenase like	*Trachidermus fasciatus*	(Ma *et al.* 2018)
	Oxysterol binding protein-like	*Fundulus heteroclitus*	(Brennan *et al.* 2018)
	Rho guanine nucleotide exchange factor	*Gasterosteus aculeatus*	(Konijnendijk *et al.* 2015)
	Sh3 domain-containing protein	*Lucania parva*	(Kozak *et al.* 2014)
Salinity tolerance	Brain protein	*Gasterosteus aculeatus*	(Kusakabe *et al.* 2017)
	Centrosomal protein 290	*Cottus asper*	(Dennenmoser *et al.* 2017)
	Cyclin-dependent kinase	*Coregonus lavaretus*	(Papakostas *et al.* 2012)

4.5 Discussion

Galaxias maculatus diadromous and resident populations from two adjacent river systems exhibited clear phenotypic and genetic differences. Phenotypically, diadromous individuals were larger than residents (Fig. S4.1). Genetically, diadromous individuals from the two river systems comprised a single genetic group, and this genetic group was highly genetically distinct from both resident populations (Fig. 4.4, Fig. S4.4). Resident populations from the two river systems (i.e., Toltén and Valdivia) were also distinguishable from each other. Valdivia resident individuals were larger than Toltén residents (Fig. S4.1). Surprisingly, 100% of Valdivia resident individuals survived the transplant experiment to salt water (25ppt), while only ~23 % of Toltén residents survived by the end of the experiment (Fig. 4.3). In fresh water, both resident populations showed some mortality (Fig. 4.3) which we speculate may have been the result of the protozoan *Ichthyophthirius multifilis*, only found in fresh water. No method was used to assess parasite levels though.

The genetic data suggest that the resident populations are the result of independent colonization events, and previous research also showed that they are reproductively isolated from the diadromous population (Delgado *et al.* 2019). The differences found between the resident populations, in their phenotypes, genotypes, and ability to survive in salt water, suggest that the loss of diadromy can lead to different outcomes regarding the retention or loss of plastic and/or adaptive traits. The comparisons of the differences found between diadromous and the resident populations have also revealed important genes associated with migratory behavior, salinity tolerance, and local adaptation. We discuss these results in detail below.

The genetic basis of migratory and resident behavior

Hundred and forty-four outlier SNPs distinguished the diadromous individuals from Toltén and Valdivia resident individuals (Fig. 4.5A). This consistency (nearly half) of common outliers indicate some level of parallelism, which is considerably higher than that observed among replicate populations of Three-spined Stickleback (Ferchaud & Hansen 2016), Atlantic Salmon (Perrier *et al.* 2013), and Arctic Charr (Salisbury *et al.* 2020). The high genomic divergence found between diadromous and resident populations in *G. maculatus* could result from the lack of gene flow between the diadromous and

resident populations; gene flow in other species appears to overwhelm selection (Ferchaud & Hansen 2016). We expected that some of these outlier loci are functionally driven and are related to migratory behavior. The BLAST analyses revealed 84 genes associated with different biological processes (Table 4.1). From these genes, 31 have been previously reported for other diadromous species, many of which are the result of comparing migratory and non-migratory life histories, and most of these genes are associated with osmoregulatory processes (e.g., ion exchangers and solute carriers) (see Table 4.2). Many of these common genes were shared not only with salmonids, but with other phylogenetically distant fish species including *Lateolabrax maculatus* suggesting some degree of genomic parallelism likely driven by common challenges associated with migratory behavior.

Besides genes associated with osmoregulation, other relevant processes that distinguish migratory from non-migratory *G. maculatus* individuals were revealed (Table 4.1):

1) Response to stimulus, eye photoreceptor, and regulation of circadian rhythm: Although these processes have not been extensively studied in diadromous species, there is some evidence of circadian rhythm genes associated with early and late-migrating Pink Salmon (*Oncorhynchus gorbuscha*) (Kovach *et al.* 2013). Visual pigments are also known to change in European Eels (*Anguilla anguilla*) before starting their downstream migration (Kusmic & Gualtieri 2000) and Chum Salmon (*Oncorhynchus keta*) individuals use moonlight to navigate during migration (Hasegawa 2012). Further studies comparing the expression of these genes between diadromous and resident individuals are necessary for further insights on their roles.

2) Lipid biosynthesis and fatty acid metabolic processes: Comparative studies between diadromous and resident individuals reveal that lipid and fatty acid levels may be affected by the nutritional composition of their respective environment and that the different levels of specific lipids can affect the performance or ability of an individual to migrate. For instance, in the anadromous Hilsa Shad (*Tenualosa ilisha*), the difference in nutritional composition between two habitats changed the composition of amino acids (Ganguly *et al.* 2020); whereas in Brown Trout the linoleic fatty acid conversion is higher in young resident individuals than anadromous individuals. This difference was

hypothesized to be associated with migratory behavior, as the synthesis of endohormones demands high levels of arachidonic acid (Murzina *et al.* 2018). Diadromous *G. maculatus* individuals may have physiological advantages when inhabiting brackish waters, reducing environment to fish gradients (vice versa), and therefore expending less energy on osmoregulation (Urbina & Glover 2015).

3) Muscle fiber development: Muscle development can be hypothesized to be responsible for the differences in size found between diadromous and resident individuals. In many diadromous species including Arctic Charr in Canada (Salisbury *et al.* 2018, 2020) and Pond Smelt (*Hypomesus nipponensis*) in Japan (Katayama *et al.* 2000), resident individuals are smaller in size than migratory individuals.

4.5.1 The genetic basis of salinity tolerance

Only 43 outlier SNPs distinguished individuals from populations that could (i.e., diadromous and Valdivia residents) from that which could not (i.e., Toltén residents) survive in salt water (Fig. 4.5B). From the 22 loci that had a blast match, only one gene, acid-sensing ion channel (Table S4.3), has been directly involved in osmoregulation. Acid-sensing ion channels are activated by extracellular protons, modifying their permeability to ions such as sodium, and it has been revealed to play an important role in sodium uptake in adult fishes while in fresh water (Dymowska *et al.* 2015). Regarding the link between Acid-Base regulation and osmoregulation, the pH values measured in the field at the time of collection do not deviate from the expected values for estuaries (e.g., Valdivia estuary water pH = 7.18) and fresh water (e.g., Valdivia lake pH = 6.78), and so they should not represent an extra burden to the salinity challenge used here.

Other genes, however, were involved in the regulation of gene expression (e.g., regulation of transcript by RNA polymerase). This finding suggests that regulation of gene expression could have an essential role in the plastic nature of diadromous species. Yet, it raises the question of why would Valdivia residents maintain this plasticity while most of Toltén residents did not? Under the assumption that colonization of these lakes was possible only following deglaciation during the Late Pleistocene, and given the lakes' geographic proximity to each other and similar altitude, we speculate that over a geological timeframe there are no differences in the time they were colonized by *G. maculatus,* although further research would be necessary to test this assumption.

Assuming a similar colonization time, we suggest three possible explanations to answer the question: 1) Environmental consequences: Selection associated with the environmental differences between the lakes could have indirectly affected genes related to salinity tolerance. Thus, the maintenance of this trait (i.e., ability to survive in salt water) varied between these resident populations due to environmental differences (i.e. relaxed selection), while in one population the trait persisted, in the other population, the trait was partially lost (Lahti *et al.* 2009). 2) Random process: genetic drift could be responsible whereby specific alleles are fixed in some populations but lost in others. Under this scenario, the alleles responsible for surviving in salt water could have been randomly fixed in Valdivia individuals, while in Toltén they may be in the process of being lost. 3) A combination of the two processes. Repeating these experiments with more river systems would be an important contribution to clarify which scenario is the main driver of the loss of this plasticity.

The absence of physical barriers and similar F_{ST} values between the diadromous and resident populations in these river systems suggests that the divergence in the resident populations' tolerance to salinity increases is not due to man-made imposed barriers limiting gene flow. Genetic drift, can however, differ between the resident populations as these populations probably differ in effective size (N_e). Although Lake Neltume (Valdivia resident) is smaller than Lake Villarica (Toltén resident), Lake Neltume is part of a much larger lake system comprising eight lakes. It can be hypothesized that the Valdivia resident population, as part of a larger system, has a larger N_e than the Toltén resident population. If this is the case, the fixation or loss of alleles due to genetic drift would have occurred faster in Toltén. Also, it has been hypothesized that plasticity is retained under more lenient conditions, for example, in a metapopulation (Masel *et al.* 2007). The fact that ~23% of Toltén residents survived the increase in salinity and that they were not experiencing osmoregulatory stress (water content analysis) in a salty environment suggests that the loss of these plastic and adaptive osmoregulation related traits may be incomplete. With time no Toltén resident individual would be expected to survive in salt water, as alleles may be lost in this population. A similar study in a landlocked lake close to Lake Villarica found that only 33% of these resident individuals survived at 25 ppt (Ruiz-Jarabo *et al.* 2016). This result is consistent

with our findings that residency may have different outcomes and that a combination of selective and random processes leads to the maintenance or loss of osmoregulatory capabilities. Further testing would be required to estimate N_e as that task could not be performed given the small sample size.

4.5.2 The genetic basis of local adaptation

Resident populations have been reported in several diadromous species. In some cases, the transition to residency is abrupt and man-made imposed by the construction of physical barriers such as river dams that landlock populations; but in others, resident populations exist even though they have access to the sea. The presence of facultative diadromous species or diadromous and resident populations inhabiting the same locations (sympatry) suggest that the decision to remain in their natal environment can be due to ecological factors, such as lenient environmental conditions and intra-specific competition, that reduced migratory costs and increased survival (Chapter 2; Delgado & Ruzzante 2020).

The phenotypic differences (i.e., mass, length, and condition factor K) observed between diadromous and resident individuals, indicate that these populations are exposed to different environmental conditions. For instance, the differences found in the condition factor (K) between diadromous and resident individuals (Fig. S4.1) demonstrate that these habitats differ in ecological (i.e. food resources) and physical variables (i.e. stream flow and currents). Estuaries are characterized by higher productivity and water flow than lakes, leading individuals to exhibit a higher condition factor K or “robustness” in estuarine than in landlocked populations. And despite the fact that the two lakes are geographically close (Fig. 4.1), they do differ in many environmental variables including species composition, minimum temperature, anthropogenic pressures (Ministerio de Obras Públicas 2004, 2004b). These differences are clearly reflected in the phenotypic differences that are the product of genetic variation.

Forty-five genes (Fig. 4.5C) differentiated the resident populations from each other. This differentiation is likely driven by isolation and associated with each population's environmental conditions. Processes associated with these genes include:

1) Regulation of growth: the fact that Valdivia residents were slightly larger and heavier than Toltén residents can be attributed to this process (Fig. S4.1), however, 2)

Adipose tissue development: could also be attributed to the differences in size between the resident populations, as in Zebrafish (*Danio rerio*), white adipose tissue appearances have been correlated with size (Imrie & Sadler 2010).

4.5.3 Experimental considerations and future directions

It is important to note that only a subset of samples from the reciprocal transplant experiment (N = 16) was used in the genomic analysis of Delgado *et al.* (Chapter 3; 2019; N = 45). However, our population genomic analyses demonstrated that no significant genetic differences were found between individuals tested in the reciprocal transplant and individuals collected in the same location at a different time, as they all grouped in PCA plots according to their respective collection site (See Fig.2 in Chapter 3; Delgado *et al.* 2019).

The amphidromous life history in *G. maculatus* involves migration during the larval and juvenile stages, yet due to logistical constraints, we used adult individuals in our reciprocal transplant experiment. The euryhaline nature of *G. maculatus,* whereby individuals are able to tolerate a wide range of salinities through their life span (see Chessman & Williams, 1975; Urbina & Glover, 2015), allowed us however, to obtain key information about their salinity response using adults. Using adults, we were still able to identify the genes characterizing: diadromous vs. resident, populations that can survive salt water vs. population that cannot, and between resident populations. Future studies that test responses to salinity changes in larvae and juveniles are warranted.

Although RADcap data provide a genome-wide array of neutral and adaptive loci, this is a reduced representation method, it is therefore likely that many other loci under selection across the genome have not been analyzed (Catchen *et al.* 2017). Only 11 SNPs were common to the two outlier detection methods used. Discrepancies between different detection approaches have been documented, as the number of false positives and negatives can vary greatly between approaches (Narum & Hess 2011). Given that we used a reduced representation approach and are thus limited with the number of SNPs, we considered all outlier SNPs whether detected by a single or both detection methods. Furthermore, the lack of a reference genome for *G. maculatus* raises some caveats. The BLAST threshold selected was somewhat lenient, with some genes presenting only 20% of coverage, yet only ~52% of our contigs had a match. The development of a reference

genome with good gene annotation for this species would be advantageous to fully determine which genes or regulatory regions differentiated migratory and non-migratory populations.

There is no reference genome for any species from the Galaxiiformes order, and according to the latest Actinopterygii phylogeny, the closest order would be the Osmeriformes or the Stomiatiformes, which diverged >100mya from the Galaxiiformes (Hughes *et al.* 2018). Furthermore, the (albeit limited) evidence for parallelism in some of the SNPs expressed across species, supports the idea that a more extensive exploration of genes involved or expressed specifically in *G. maculatus* is necessary. With a genome, it would be possible to assess if the migratory genes identified here are linked, as suggested by some studies in fishes (Hale *et al.* 2013) and migratory birds (Ruegg *et al.* 2014), which identified islands of differentiation linking genes associated to migration. Further research on the outlier SNPs, specifically if the variation leads to synonymous or non-synonymous mutations, would also be needed for the improvement of our insight on the genetic basis of migratory behavior and its loss.

Given the number of genes associated with the regulation of gene expression that we observed distinguishing migratory from non-migratory populations, the study of the transcriptomics and differential gene expression is expected to be another step to increase our understanding of the different mechanisms involved in the loss of diadromy. Studies focused on transcriptomics have identified some of the molecular mechanisms underlying gene expression variation between diadromous and resident populations from species like Brook Charr (*Salvelinus fontinalis*) (Boulet *et al.* 2012), Ayu (*Plecoglossus altivelis*) (Lu *et al.* 2016), Rainbow Trout (Hale *et al.* 2016), and Nine-spined Stickleback (*Pungitius pungitius*) (Wang *et al.* 2020).

Questions, such as whether the loss of diadromy leads to the loss of plasticity, have yet to be addressed. In amphidromous species, larvae and juveniles are often the primary migrants and they generally exhibit incomplete site fidelity leading to the colonization of new habitats. In such cases, knowledge of whether or not plasticity is maintained can lead to a better understanding and more accurate predictions regarding the success or failure in the colonization of new habitats. *Galaxias maculatus* exhibits life history characteristics that could make it a good model organism for the study of the

genetic basis of diadromy and the consequences of its loss. Many of these resident populations of *G. maculatus* in Patagonia and most likely many others across the species Gondwanan distribution, derived from a single migratory population and can thus serve as an example for the studies of evolutionary processes such as relaxed selection, parallel evolution, and even speciation.

4.6 Acknowledgements

Sample collection and reciprocal transplant experiments were funded by FONDECYT REGULAR-1150154 grant to EH and CONICYT-FONDECYT-11160019 grant to MAU. The genomics work was funded by an NSERC Discovery grant to DER. MLD was supported by the NSERC DG grant to DER, Dalhousie's Faculty of Graduate Studies Scholarship and by Nova Scotia Graduate Scholarship (NSGS). We would like to thank Anne Dalziel, for helping with the design of the reciprocal transplant. For their assistance in the field, we thank Pablo Fierro and Gustavo Díaz. For assistance in rearing the fish, we thank René Iribarren. For his help with the laboratory work, we thank Greg McCracken. We also appreciate the comments by the Editor and the four anonymous reviewers which helped improve this work.

4.7 Author contributions

This manuscript is part of MLD's PhD thesis research. All authors contributed to the conception and design of the study. Fish were collected by MLD and AM. Data from the reciprocal transplant experiment were collected by MLD, AM, and MAU. MLD generated and analyzed the genomic data. The manuscript was written primarily by MLD with input from DER and contributions by AM, MAU, EH, and OL.

4.8 Supporting information

Table S4.1. Filters applied to raw data: percentage of missing data allowed per locus (geno missing), minimum depth (minDP), minor allele frequency (maf), only first SNP per contig (thin), loci under Hardy Weinberg equilibrium (HWE), loci under linkage disequilibrium (LD) and missingness per individual (indv missing). For each filter the remaining number of loci (SNPs) is presented and the total number of individuals (indv) that passed the filters.

Filter	value	SNPs	indv
geno missing (%)	25	133745	47
minDP	6	133745	
maf	0.01	44086	
thin	yes	5654	
HWE	yes	5154	
LD	yes	5154	
indv missing (%)	20	5154	45

Table S4.2. Common outlier loci between diadromous vs. resident individuals (ValRes + TolRes) (see Fig. 4.5A) that had a Blast match (E-value <0.01, query coverage ≥20% and identity >70%).

Gene	Query Coverage	% Identity
Signal-induced proliferation-associated 1-like	21	85.19
Elongation factor like GTPase 1 (elf1)	21	89.29
Probable palmitoyltransferase	21	82.35
Carcinoembryonic antigen-related cell adhesion molecule 5-like	22	79.49
Calcium voltage-gated channel auxiliary subunit alpha2delta1	22	82.89
Acid-sensing ion channel 2	22	80.9
Organic solute transporter subunit alpha-like	22	90
Elongation factor 1 alpha (EF1- alpha)	22	97.06
E3 ubiquitin-protein ligase RNF213-like	22	85.92
Zinc finger protein 5-like	22	88.57
Proprotein convertase subtilisin/kexin type 5-like	23	78.21
Rho related BTB domain containing 3 (rhobtb3)	23	96
Cilia and flagella associated protein 53 (cfap53)	24	84.93
Sodium/potassium/calcium exchanger	24	78.67
Natterin 3-like	24	83.16
BEN domain containing 5 (bend5)	25	87.18
Lipoxygenase homology domain-containing protein 1-like	25	88.61
IK cytokine (ik)	25	82.11
Prolyl-4-hydroxylase transmembrane	25	91.3
Carcinoembryonic antigen-related cell adhesion molecule 5-like	26	75.95
Brain protein I-3-like	26	87.74
Ubiquitin carboxyl-terminal hydrolase 25-like	26	90.36
Sh3 and Px domain-containing protein 2A-like	26	90.36
Transformation/transcription domain associated protein (TRRAP)	27	93.81
Serine/threonine-protein kinase Nek3-like	27	75.24
Thrombospondin-like 27	27	85.71
Cadherin 18-like	27	90.48
Laminin subunit alpha 4 (lama4)	27	78.18
Oxysterol binding protein like-11	27	81.82
Nucleolar transcription factor 1-A-like	29	76.6
Angiopoietin 2-like	29	74.47
Solute carrier family 25 member 12 (slc25a12)	29	90.8
Serine/threonine-protein kinase (taok2)	29	89.01
Phosphatidylinositol 4-kinase alpha (pi4ka)	29	92.98
Plexin domain-containing protein-1-like	30	85.48
Ryanodine receptor 3-like	30	91.09
Nk2 homeobox 5 (nkx2-5)	30	80.95

Gene	Query Coverage	% Identity
Lysine-specific demethylase 4C-like	31	80.19
Ef-Hand domain-containing protein Dw-like	31	89.22
Glycylpeptide N-tetradecanoyltransferase 2-like	31	79.63
Dystonin (dst)	31	92.92
RNA polymerase II subunit A (polr2a)	32	90
Protocadherin 16-like	33	78.43
BICD family-like cargo adapter 1	33	87.39
Angiopoietin like 6 (angptl6)	33	81.13
D-glutamate cyclase (dglucy)	34	83.04
G protein signaling modulator 1-like	35	84.73
Collagen type V alpha 2 (col5a2)	36	78.26
Histone-N-methyltransferase H3 lysine-36 and H4 lysine-20 specific-like	36	84.24
Nuclear receptor ROR beta-like	37	77.31
Microtubule actin crosslinking factor 1 (macf1)	38	84.43
Glyceronephosphate o-acyltransferase (gnpat)	39	73.65
Mitogen-activated kinase-binding protein 1	40	87.32
Trichohyalin-like	41	69.1
Nuclear receptor subfamily 1 group D member 2-like	42	82.89
Pleckstrin homology domain A7 (plekha7)	42	78.26
Neuroligin 1 (nlgn1)	43	91.08
Pecanex 4 (pcnx4)	43	90
4-Hydroxy-2-Oxoglutarate aldolase 1 (hoga1)	43	84.96
Transmembrane protein 266 (tmem266)	44	92.91
Gamma taxilin-like	45	95.56
Vesicle-trafficking protein SEC22b-B	48	86.61
Cyclin-dependent kinase 6	50.9	76.92
Ubiquitin protein ligase E3 component n-recognin 5 (ubr5)	52	87.73
Vesicle-fusing ATPase	54	87.57
Centrosomal protein 290 (cep290)	56	75.58
RNA-binding protein 10-like	56	91.3
Zinc finger protein 501-like	57	74.47
E3 ubiquitin-protein ligase TRIP-12-like	57	88.54
Rho guanine nucleotide exchange factor 3-like	59	81.2
Vimentin type intermediate filament associated coiled-coil protein (vmac)	60	75
Calcium voltage-gated channel auxiliary subunit alpha2delta (cacna2d2)	60.8	71.93
Fatty acid synthase (fasn)	61	87.05
RNA-binding protein 25-like	62	74.11
L-Lactate dehydrogenase A chain	62	91.3
Protein glucosylgalactosylhydroxylysine glucosidase (pgghg)	72	72.53
RNA polymerase III subunit 3 (polr3e)	73	73.11
Ski oncogene-like	78	87.23

Gene	Query Coverage	% Identity
Aerolysin-like protein	87	70.52
Interferon-inducible GTPase 5-like	90	75.84
Neuroepithelial cell transforming 1 (net1)	90	77.98
Zinc finger and BTB domain containing 4 (zbtb4)	94	78.22
Zinc finger and homeoboxes 2 (zhx2)	94	73.06
Chondroitin sulfate proteoglycan 4 (cspg4)	98	71.29

Table S4.3. Common outlier loci between individuals that can survive in marine water (Diadromous + ValRes) vs. individuals that cannot survive in marine water (TolRes) (see Fig. 4.5B) that had a Blast match (E-value <0.01, query coverage ≥20% and identity >70%).

Gene	Query Coverage	% Identity
Adhesion G protein-coupled receptor V1 (adgrv1)	20	80
Protein HoxD13aa	20	98.36
Trypsin 2-like	23	85.06
Elongator acetyltransferase complex subunit 2 (elp2)	24	83.54
Protein particle complex subunit 8-like	26	93.68
Very long-chain specific acyl-CoA dehydrogenase mitochondrial-like	27	91.49
Rho GTPase-activating protein 42-like	30	89.25
ASPSCR1, UBX domain-containing	34	79.84
Junctional adhesion molecule 3 (jam3)	35	80
Scaffold attachment factor B2-like	36	86.27
Calpain 10 (capn10)	37	76.19
WD repeat domain 90 (wdr90)	37	75.41
Bromodomain adjacent to zinc finger domain 1B (baz 1b)	38	89.08
Cytosolic 5'-nucleotidase 3-like	42	90
Lysine methyltransferase 2C (kmt2c)	47	86.98
Membrane metallo-endopeptidase-like	47	86.81
Adhesion G protein-coupled receptor E2-like	53	80.12
Patatin like phospholipase domain-containing 6 (pnpla6)	61	88.6
Acid-sensing ion channel subunit 2 (asic2)	73	75.1
Nuclear receptor subfamily group F member 6-like	75	88.48
E3 ubiquitin-protein ligase Topors-like	79	68.75
Claudin 4-like	93	80.56

Table S4.4. Other outlier loci found to differentiate resident individuals from both river systems (Valdivia and Toltén) (see Fig. 4.5C) which had a blast match (E-value <0.01, query coverage ≥20% and identity >70%).

Gene	Query Coverage	% Identity
Ribosome binding factor A (rba)	20	77.63
Myb/SANT DNA binding domain-containing 1 (msantd1)	20	92.31
Major facilitator superfamily domain-containing protein 1-like	21	93.24
Cyclin-dependent kinase 4	21	89.71
Protocadherin-10-like	21	90
Sodium channel subunit beta-4-like	21	75.34
Zinc finger protein 609	23	79.45
Semaphorin-3F-like	23	77.67
Vinexin-like	23	87.34
LIM and senescent cell antigen-like-containing domain protein 1	24	94.57
Histidine-rich glycoprotein-like	24	72.17
Stathmin-4	25	74.47
Triple functional domain protein-like	26	93.55
F-box-like/WD repeat-containing protein (tbl1xr1)	28	92.63
Glutamine-rich protein 2-like	28	83.53
Monocarboxylate transporter 13-like	28	73.12
Solute carrier family 22 member 23-like	30	88.71
Phospholipase DDHD1-like	30	91.8
Protein tyrosine kinase 7 (inactive) (ptk7)	30	91.38
Calpain-5-like	31	93.75
E3 ubiquitin-ligase itchy-like	32	91.89
Huntingtin interacting protein 1 (hip1)	32	79.49
RAS-related protein Rab-19-like	33	78.3
Heat shock 70 kDa protein 1B-like	34	74.29
Inhibitor of DNA binding 1, HLH protein (id1)	34	81.2
Cadherin, EGF LAG seven-pass G-type receptor 1	35	85.45
ES1 protein homolog mitochondrial	36	86.36
Kalirin-like	37	88.71
Teneurin-1-like	38	87.7
KIAA0100 ortholog (kiaa0100)	39	87.1
TOR signaling pathway regulator (tiprl)	43	88.81
Zinc finger protein 462-like	43	80.12
Adenylosuccinate synthetase isozyme 2-like	45	83.57
Cerebral endothelial cell adhesion molecule	46	84.96
Anaphase-promoting complex subunit2-like	47	87.58
Tyrosine kinase (egfrb)	55	91.37
COBW domain-containing protein 1-like	56	91.26
TUB-like protein 4 (tulp4)	61	94.05

Gene	Query Coverage	% Identity
Myosin-IIIb-like	68	90.75
LHFPL tetraspan subfamily member 4 (lhfpl4)	70	92.79
WD repeat-containing protein 38-like	72	80.53
Leucine-rich repeat neuronal protein 2	83	75.35
TGF-beta activated kinase 1 (MAP3K7) binding protein 2 (tab2)	84	77.62
Potassium voltage-gated channel protein Shal 22	86	67
C-C chemokine receptor type 1	98.7	81.97

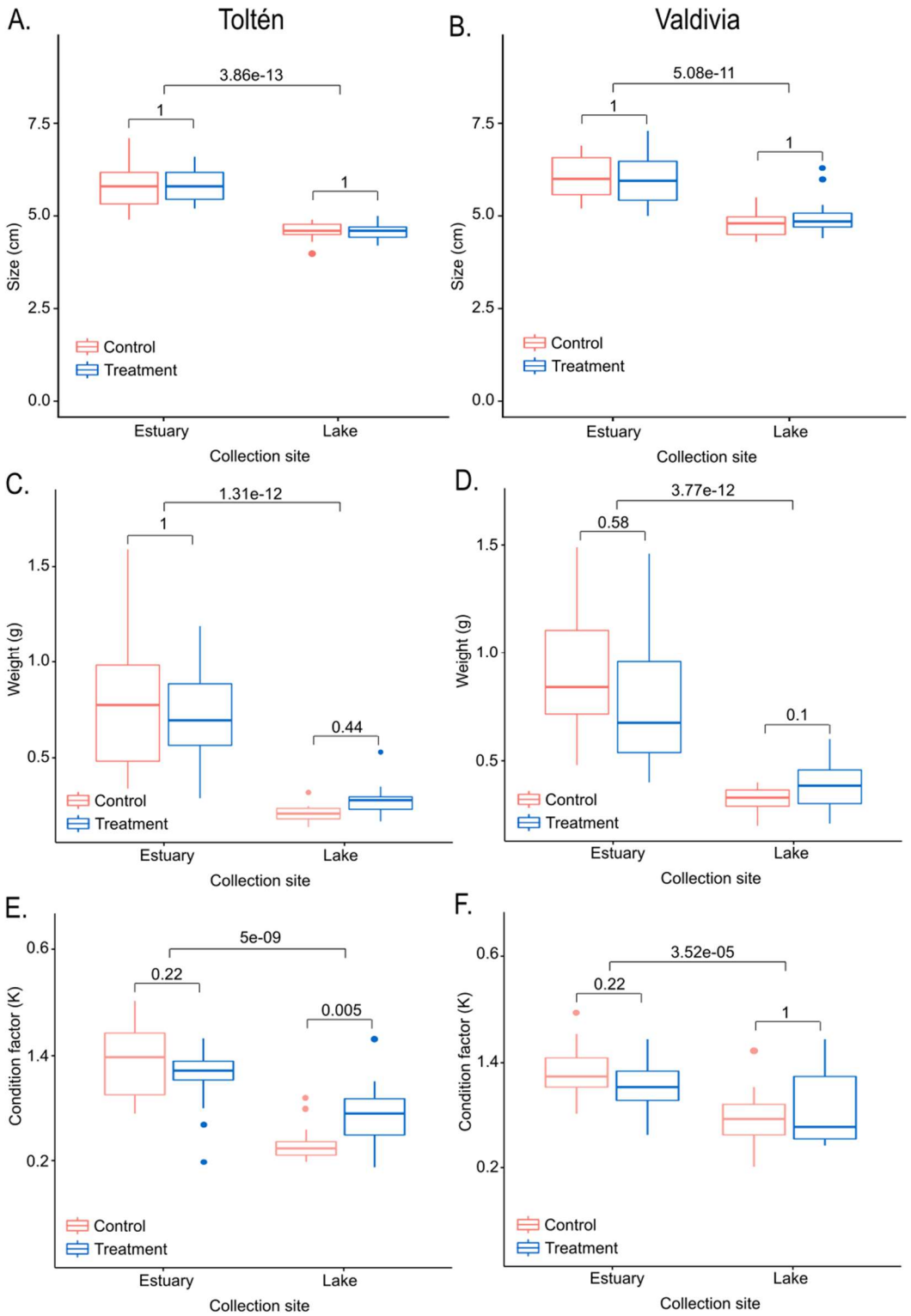


Figure S4.1. Box plots of size (total length) (**A** and **B**), weight (**C** and **D**), and condition factor (**E** and **F**) distributions between individuals in the control vs. treatment tanks collected from the estuary or lake from Toltén (**A, C,** and **E**) and Valdivia (**B, D,** and **F**) river systems. Wilcoxon tests (with Bonferroni correction) revealed significant difference mainly between collection sites, p-values are shown top of the boxplots.

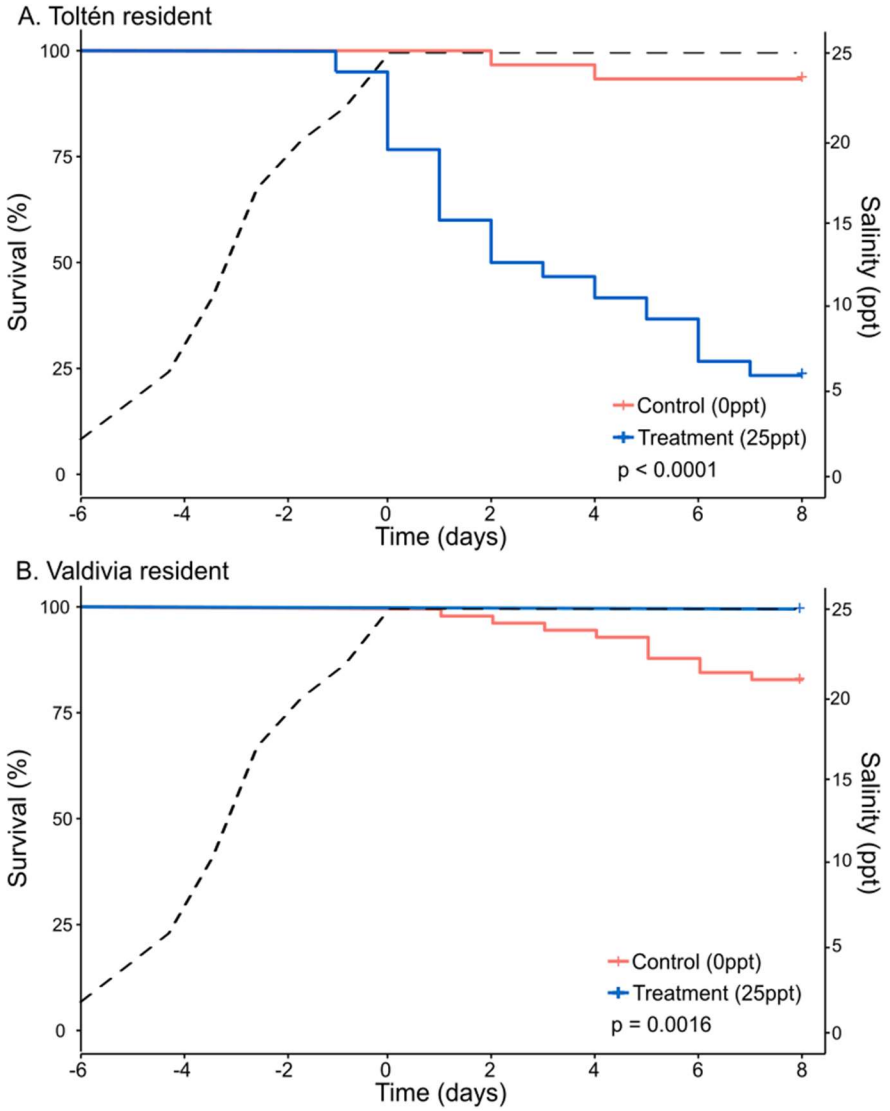


Figure S4.2. Survival percentage of resident adults from **A.** Toltén and **B.** Valdivia as a result of a gradual increment of salinity (0 to 25ppt).

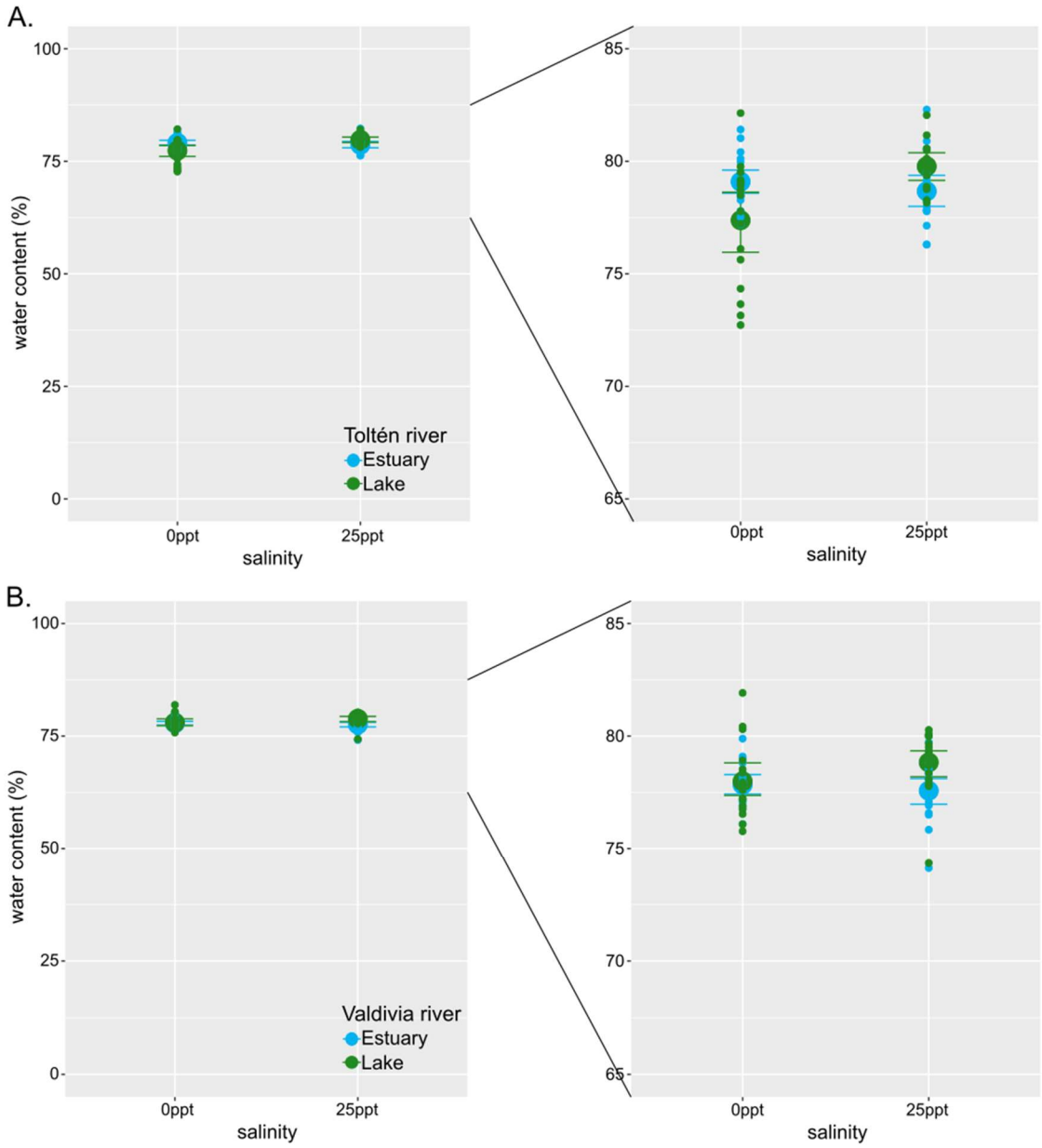


Figure S4.3. Percentage of water content in the muscle at two salinity concentrations (0 ppt and 25 ppt) of **A**. Toltén and **B**. Valdivia individuals that survived at the salinity change. The right panel shows the points at a large scale.

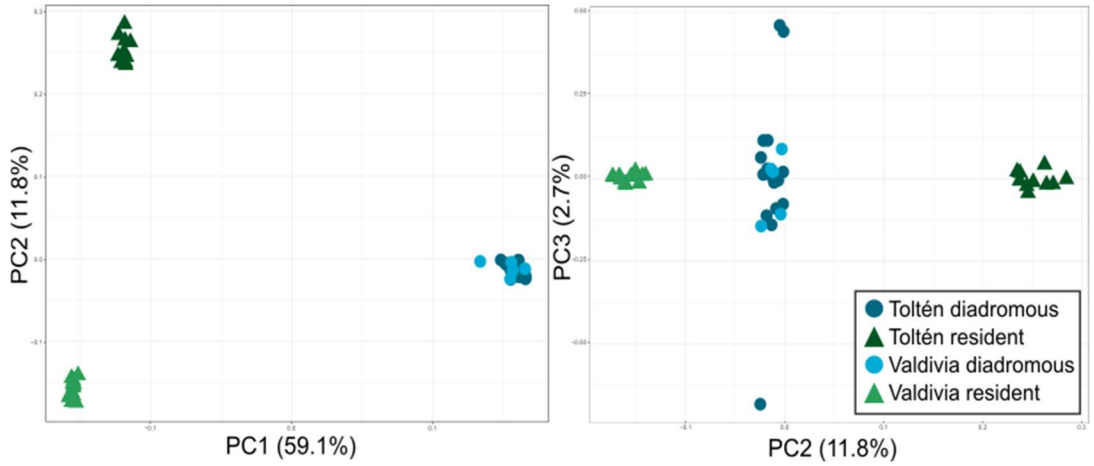


Figure S4.4. PCA performed on 45 *G. maculatus* individuals from 4 locations using 523 outlier loci.

CHAPTER 5. GENE EXPRESSION VARIATION ASSOCIATED WITH THE LOSS OF DIADROMY AND UPPER SALINITY TOLERANCE IN *GALAXIAS MACULATUS*

This chapter will be submitted to Molecular Ecology as "ML Delgado, A Manosalva, M Urbina, Anne C. Dalziel, E Habit, O Link & DE Ruzzante. Gene expression variation associated with loss of diadromy and upper salinity tolerance in *Galaxias maculatus*".

5.1 Abstract

The recurrent colonization of freshwater habitats and subsequent loss of diadromy is a major ecological transition that has been reported in many ancestrally diadromous fishes. Such residency is often accompanied by a loss of tolerance to seawater environment. The amphidromous *Galaxias maculatus* has repeatedly colonized freshwater streams with evidence that freshwater-resident populations exhibit stark differences in their salinity tolerance. Here we used transcriptomics to test the hypothesis that the costs of residency in derived freshwater environments associated with the loss of diadromy would result in reduced salinity tolerance in resident populations. We conducted an acute salinity challenge (0 ppt to 23-25 ppt) and measured osmoregulatory ability (muscle water content) over 48 hours in diadromous, saltwater intolerant (Toltén), and saltwater tolerant (Valdivia) freshwater populations. RNA sequencing of the gills identified genes that were differentially expressed in association with the salinity change and elucidate those genes associated with the loss of saltwater tolerance in the Toltén population. Key genes associated with saltwater acclimation were characterized in diadromous *G. maculatus* individuals, some of which were also expressed in the saltwater tolerant resident population (Valdivia). We found that some of these "saltwater acclimation" genes, including CFTR, were not expressed in individuals of the saltwater intolerant resident population (Toltén), suggesting a potential mechanism for the loss of salinity tolerance. As the suite of differentially expressed genes in the diadromous-resident comparison differed between freshwater populations, we hypothesize that drift may be responsible for the unique evolutionary trajectories of resident populations originating from a common diadromous population.

5.2 Introduction

Our understanding of the genetic architecture of diadromous species which inhabit both seawater and freshwater environments is still limited. Many populations of ancestrally diadromous fishes have evolved changes in their behavior and physiology to fully complete their life cycle in their derived natal freshwater environment. Remarkably, the pace of adaptive divergence in such non-migratory or resident populations is rapid given phenotypic and genetic divergence from their diadromous counterparts (e.g., Three-spined Stickleback (*Gasterosteus aculeatus*) (Liu *et al.* 2018), Rainbow Trout (*Oncorhynchus mykiss*) (Pearse *et al.* 2014), and Arctic Charr (*Salvelinus alpinus*) (Salisbury *et al.* 2018)). Such resident populations can originate allopatrically (O'Malley *et al.* 2019; Härer *et al.* 2021) or sympatrically with diadromous populations (Marques *et al.* 2019; Salisbury *et al.* 2020), and they can do so in a relatively short evolutionary time frame, even within a few generations (e.g., Bell *et al.* 2004).

Diadromous and recently-evolved resident populations represent an ideal system to elucidate the mechanisms contributing to evolutionary variation in a suite of behavioral, physiological, and morphological traits associated with a migratory life history (Seehausen & Wagner 2014). Recent studies have identified genetic changes that underlie phenotypic differences between populations with these two life histories. These genetic differences encompass a range of mutation types, including Single Nucleotide Polymorphisms (SNPs) in genes or regulatory regions as seen in the ectodysplasin (EDA) locus in sticklebacks, where alternate alleles lead to variation in body armor (O'Brown *et al.* 2015), variation in linkage groups, such as the omy5 chromosomal region in steelhead/Rainbow Trout, a region that is associated with variation in life history (Pearse *et al.* 2014), and gene copy number variation, such as the fatty acid desaturase 2 (Fads2) duplications associated with the ability to survive on a freshwater diet in Three-spined Stickleback (Ishikawa *et al.* 2019). Elucidating the loci associated with adaptive phenotypic variation in traits such a salinity tolerance, swimming performance, time of migration, etc. should help us better assess potential mechanisms underlying diadromy and those resulting from this change in life history. By comparing replicate resident populations, we can also assess the extent of genetic and phenotypic parallelism occurring after the evolutionary loss of diadromy. While it might be expected that similar

changes in the selective environment will produce a parallel phenotypic and genetic response (Rivas *et al.* 2018), the evidence for parallelism/convergence among replicate non-migratory populations thus far is scarce, suggesting many different mechanisms can contribute to the loss of migratory phenotypes (Perrier *et al.* 2013; Bolnick *et al.* 2018; Liu *et al.* 2018; Salisbury *et al.* 2020). Yet, the scarcity of annotated genomes and reliable functional genomic information in fishes has hindered our understanding of the molecular mechanisms underlying the divergence of resident and migratory populations (Pavey *et al.* 2012; Todd *et al.* 2016). Thus, the molecular changes contributing to behavioral modifications, including the choice to not migrate (in cases where migration is not physically blocked), as well as the changes occurring following colonization still need to be further investigated.

Given the short time frame over which population differentiation in association with diadromy loss can take place, *de novo* mutations are unlikely, and the molecular mechanisms underlying the transition from diadromy to residency may include changes in gene regulation as opposed to extensive changes in gene coding regions (López-Maury *et al.* 2008) and/or selection on standing genetic variation (Rogers *et al.* 2013). Such regulatory changes in existing genetic diversity (e.g., adaptive phenotypic plasticity) present in the resident populations may facilitate this ecological and environmental transition (Morris *et al.* 2014). In sticklebacks, for example, repeated colonization of freshwater environments is predominantly associated with regulatory changes (Jones *et al.* 2012), which are mainly the result of *cis*-acting variation (Verta *et al.* 2019). Thus, transcriptomic approaches in tissues associated with the many phenotypic changes occurring during the evolution of residency should improve our understanding of the mechanisms associated with the loss of diadromy (Chapter 2; Delgado & Ruzzante 2020). Data obtained from sequencing of RNA (RNA-seq) can be used to construct or assemble transcriptomes in non-model species without the need for a reference genome or transcriptome thus providing the information needed to assess genetic variation among populations (Todd *et al.* 2016).

Amphidromy is a category of diadromy describing species that are born in fresh water and move to the sea as larvae. After a few weeks or months, amphidromous fishes return to fresh water to complete their life cycle (McDowall 1997). *Galaxias maculatus,*

an amphidromous fish widely distributed in the Southern Hemisphere, has several resident populations across its distribution in Chile (Górski *et al.* 2018). Resident populations from the northernmost region originated from independent colonization events from a common diadromous population (Chapter 3; Delgado *et al.* 2019), presenting the opportunity to study natural replicates for the loss of diadromy. Within this region, resident *G. maculatus* populations can have a wide range of tolerance to salinity (Ruiz-Jarabo *et al.* 2016). Delgado *et al.* (2020; Chapter 4) previously showed that resident populations from two neighboring river systems, Toltén and Valdivia, were genetically distinguishable and exhibited stark differences in their ability to acclimate to salt water. Yet, the mechanisms underlying this variation remain unknown.

Delgado *et al.* (2020; Chapter 4) found the saltwater intolerant Toltén and saltwater tolerant Valdivia resident populations differ at many candidate genes containing outlier SNPs providing evidence of genetic adaptation to these environments. However, the lack of an annotated genome limited the ability to predict the effect of these outlier SNPs on gene function or expression. Here, to characterize functional candidate genes associated with the loss of upper (25 ppt) salinity tolerance in the Toltén population of *G. maculatus* we challenged diadromous (saltwater tolerant), Toltén (saltwater intolerant), and Valdivia (saltwater tolerant) fish to an abrupt salinity change from 0 ppt to 23-25 ppt. We measured muscle water content as a proxy for osmoregulatory ability at different time points following a post salt water (23-25 ppt) challenge and examined RNAseq data from whole gill tissue before and at 24-hours post salinity change experiment. Our main objectives were to 1) identify salinity-dependent differentially expressed genes (DGE) within populations by comparing fish pre- and post- salinity change, and 2) compare the transcriptomic response among the three populations using the response of diadromous individuals as a reference. To accomplish these goals, we identified "stress-responsive" genes using the handling controls (fish in tanks where fresh water was replaced with just fresh water) to differentiate genes responding to salinity change and general experimental stress. Given the results of a previous gradual salinity acclimation experiment (Delgado *et al.* 2020), we hypothesized that genes related to ion regulation (e.g., ion transporters found in gill ionocytes; Hwang et al. 2011) may exhibit loss of functional variation in the saltwater intolerant Toltén resident individuals, while the saltwater tolerant Valdivia

resident individuals will express ion regulatory genes at levels and directions similar to those exhibited by diadromous individuals.

5.3 Materials and Methods

5.3.1 Fish collection

Adults were collected from three locations (one estuary and two lakes) and then brought to the Universidad de Concepción to acclimate to common freshwater conditions (0 ppt). Individuals used in this study were different from Delgado *et al.* 2020 (Chapter 4). Diadromous individuals were collected from the estuary of the Valdivia River system (Lat. –39.85 Long. –73.33) and freshwater resident individuals were collected from two genetically distinguishable populations: Lake Villarica (Lat. –39.28 Long. –72.23) in the Toltén River system and Lake Neltume (Lat. –39.78 Long. –71.96) in the Valdivia River system (Fig. 5.1).

All sampling and experimental procedures were performed following the Chilean guidelines on animal care and approved by the Ethics, Bioethics, and Biosafety Committee of the Universidad de Concepción and the University Committee on Laboratory Animals (UCLA) of Dalhousie University.

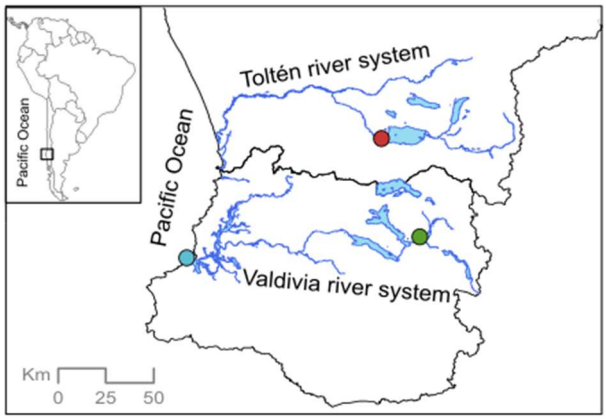


Figure 5.1. Sampling sites for the diadromous (blue circle), Toltén resident (red circle), and Valdivia resident (green circle) individuals located in the Intermediate Area in Chile.

Fish were transported to the Universidad de Concepción in 80L containers with air pumps and filled with water from each site. Salinity (measured with a salinometer ThermoFisher, Orion Star A222) at the time of collection was <1 ppt for all sites, even in the estuary as fish were collected during low tide. In the laboratory, individuals were

placed in 65L tanks (n = 35 fish per tank) for acclimation (Fig. 5.2A). Tanks were filled with chlorine-free water (0 ppt) and temperature (checked daily) was maintained at 18 ± 1°C. Ammonia, nitrate, and pH levels were checked weekly using commercial kits (©API). Ammonia levels were <0.25 (mg/L), nitrate levels were 0 (mg/L), and pH for all tanks ranged between 7.2-7.6. Fish were fed brine shrimp and flakes daily and were kept under natural photoperiod for the season (14 hr light). The water used for replacement was chlorine-free water kept in containers in the same laboratory, thus water temperature when water was replaced for cleaning the tanks and for the experiment was similar to the temperature fish experienced in the tanks.

5.3.2 Lab acclimation (0 ppt) and salt water (23-25) ppt transfer experiment

After two weeks of acclimation, individuals from each population were randomly chosen from within a population and placed in one of four 65 L future treatment (three tanks) or handling control (one tank) tanks. A total of 25 individuals were placed per tank (Fig. 5.2A). Following eight days of further acclimation to fresh water in these tanks, salinity in the treatment tanks was abruptly changed to ~25 ppt by removing approximately half the tank water volume and immediately replacing this volume with salty water. An online calculator (https://www.hamzasreef.com/Contents/Calculators/TargetSalinity.php) was used to estimate the amount of synthetic seawater salt blend (Instant Ocean) needed so the tanks reach 25 ppt after the replacement of half of the tank water. Salinity was checked with a salinity meter (ThermoFisher, Orion Star A222), but given that the amount of water removed and filled was not exactly the same for each replicate tank, the salinity ranged between 23 and 25 ppt. The three handling control tanks (one per population) were subject to the same water removal treatment (half the tank), but this water was immediately replaced with fresh water (0 ppt) to allow us to distinguish the effects of the change in salinity from that of the water replacement stress (Fig. 5.2A).

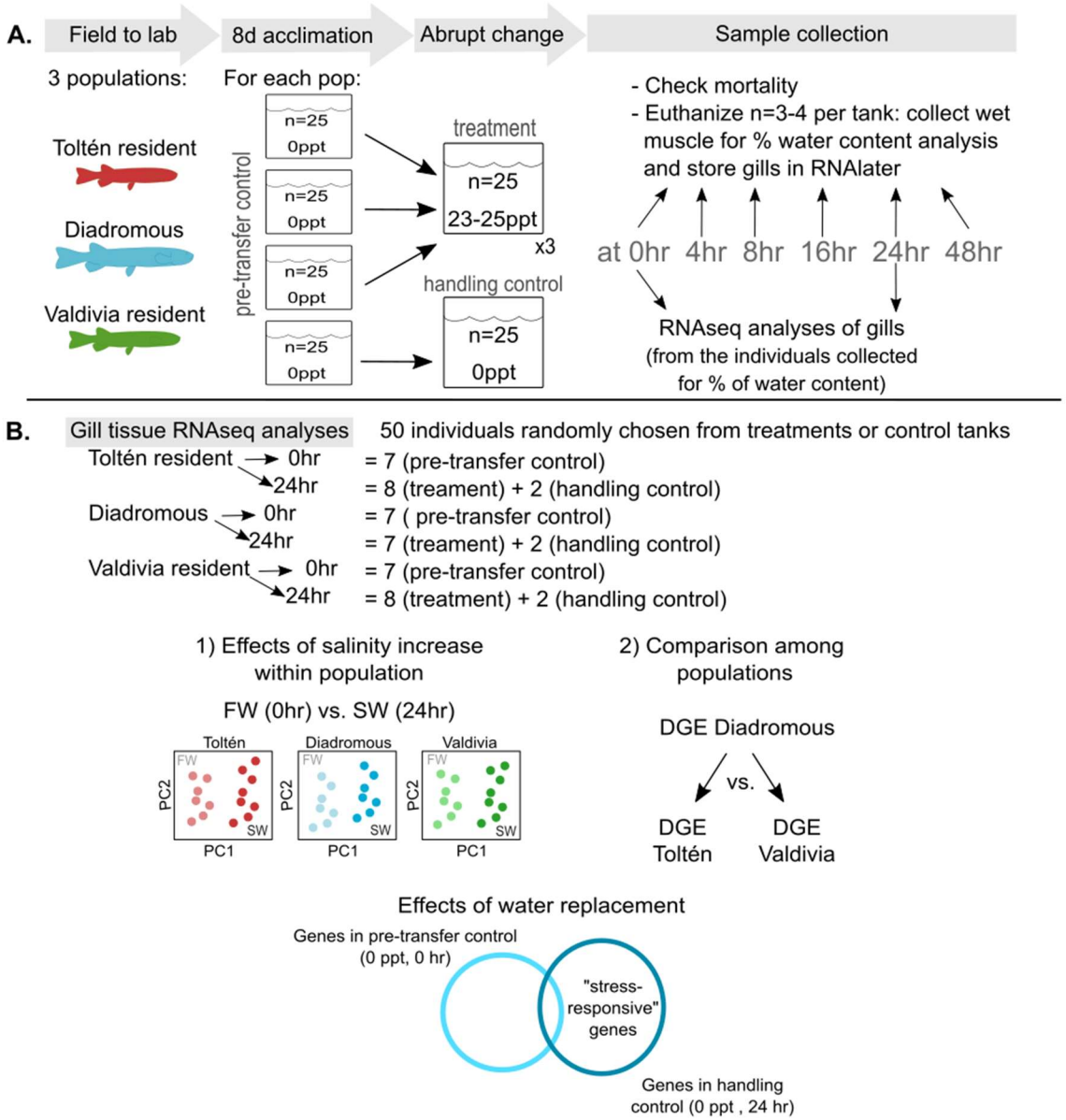


Figure 5.2. Experimental design for this abrupt salinity change experiment. **A.** Timing of experimental salinity transfers, controls for handling stress, and sampling. **B.** Number of individuals per population included in the RNAseq analyses and graphic representation of the two main questions tested in this paper: 1) effects of salinity increase, DGE analysis per population comparing expression between FW (0 hr) and SW (24 hr). 2) Comparison of DGE results among populations using the results of the diadromous population as reference. To control for the effects of stress during water replacement, we identified possible "stress-responsive "genes only found in the assembly built with handling control (0 ppt, 24 hr) and removed these transcripts form the analysis in Figure 4 and 5. "FW" = fresh water, "SW" = salt water, and "DGE" = Differential gene expression.

5.3.3 Sampling and data collection

Muscle and gill tissue were collected from fish at 0 ppt just prior to the experimental water change (0 hr), and at 4, 8-, 16-, 24-, and 48-hours post water changes from the handling control (0 ppt) and experimental (23-25 ppt) tanks. At these times, we checked for mortality and euthanized four randomly chosen individuals from each of the nine treatment tanks (three replicates for each of the populations: diadromous, Toltén, and Valdivia) and three from each of the handling control tanks with an overdose of benzocaine (100 mg/L). Thus, density per tank decreased following each data collection point. Individuals were not showing signs of morbidity (i.e., swimming not visually affected) at the time they were euthanized. A piece of white muscle (~0.5 cm) was removed to estimate % of water content and the heads (including the gills) were placed in 2 mL tubes filled with RNAlater. Gills were stored at -20°C immediately after extraction at the Universidad de Concepción and at -80°C once at the laboratory at Dalhousie University for subsequent transcriptomic analyses (Fig. 5.2A).

Percentage of water content is defined as,

$$\text{(eq. 5.1)} \qquad \%\text{Water Content} = \frac{W1 - W2}{W1\ x\ 100}$$

Where W1 and W2 are wet and dry weights, respectively, provided information on osmoregulatory capacity, as saltwater intolerant fish are expected to passively lose body water to the hypersaline salt water, as water flows down its osmotic gradient increases.

Wet tissue was first weighed (W1), then dried at 60^{o} C in an oven for 24 hours before being weighed again (W2). As the data did not meet assumptions of normality, a Friedman's test was used to test for a significant difference in the % of water content at the different time points for the treatments and handling control samples in each population. All statistical analyses were conducted in R (R Core Team 2020).

5.3.4 RNA extraction and sequencing

The % of water content in muscle informed the selection of the most variable time point for sequencing. We randomly chose a total of 50 individuals (from the set of individuals used to estimate the % of water content) for the RNAseq analysis (see Fig. 5.2B). The 50 samples included: 16 diadromous individuals (7 at 0 hr – pre-transfer control at 0 ppt, 7 at 24 hr at 23-25 ppt, and 2 handling controls taken at 24 hr at 0 ppt), 17 Toltén residents (7 at 0 hr at 0 ppt – pre-transfer control, 8 at 24 hr at 23-25 ppt, and 2

handling controls taken at 24 hr at 0 ppt), and 17 Valdivia residents (7 at 0 hr – pre-transfer control at 0 ppt, 8 at 24hr at 23-25 ppt, and 2 handling controls taken at 24 hr at 0 ppt).

To isolate total RNA, one gill was removed from each individual and placed in tubes for shearing using a beadbeater. Subsequently, the RNeasy Mini Kit (Qiagen©) extraction protocol was followed. Extractions were treated with DNase I to eliminate any DNA. RNA samples were kept at -80^{o} C until sent for sequencing. Quality control of the RNA samples, RNAseq library preparation, and sequencing using the Illumina NovaSeq platform were conducted at Génome Québec Inc. All samples passed the quality control, and their RNA integrity number (RIN) was >6.5. Sequencing was performed in 1 lane of the NovaSeq platform. All samples (n =50) had more than 30 million reads. The number of reads per sample ranged between 30 million and 80 million (Table S5.1). The software packages FastQC (Andrews 2010) and Trimmomatic (Bolger *et al.* 2014) were used to assess read quality, trim adapters, and remove low-quality reads, and SortMeRNA (Kopylova *et al.* 2012) was used to remove rRNA.

5.3.5 *De novo* assembly

The Oyster River Protocol (ORP), a multi-assembler (Trinity, SPAdes, and TransABySS) and k-mer approach for transcriptome assembly (MacManes 2018), was used to create *de novo* assemblies. Multiple *de novo* assemblies were created using different individuals from different populations (see Table 5.1) as the F_{ST} between the diadromous and resident populations are ~0.4 (Delgado *et al.* 2019). The software packages TransRate (Smith-Unna *et al.* 2016) and BUSCO (Seppey *et al.* 2019) were used to evaluate the "correctness" and "completeness" of the assemblies. Assemblies built with just the Trinity assembler were performed, but their TransRate values were lower than the assemblies built with ORP (data not shown). The software package dammit (Scott 2016) was used to annotate the transcripts using the Actinopterygii BUSCO group database. And the package Salmon (Patro *et al.* 2017) was used to pseudo-align and quantify transcripts to the *de novo* transcriptomes.

Table 5.1. Basic statistics and evaluation of *de novo* assemblies obtained using Oyster River Protocol (ORP) based on different numbers and combinations of individuals. The 3 populations are the diadromous, Toltén resident, and Valdivia resident. FW= fresh water and SW=salt water.

ORP assemblies	Includes	N used to build the assembly	Number of transcripts	Mean len	Percent GC	Transrate score	BUSCO Actinop-terygii	n90	n70	n50	n30	n10
Assembly 1	1 from each population in FW (0hr) and SW (24hr)	6	102664	1119.09	0.47	0.45	80.1%	460	1076	1985	3094	5366
Assembly 2	3 diadromous in FW (0hr)	3	85214	1103.85	0.47	0.45	76.8%	462	1098	1853	2773	4568
Assembly 3	1 from each population in FW (0hr)	3	90712	1115.18	0.46	0.46	77.5%	469	1083	1847	2902	4948
Assembly 4	3 Toltén resident in FW (0hr)	3	74495	1526.76	0.47	0.52	84.4%	357	1046	2082	3394	5905
Assembly 5	3 Valdivia resident in FW (0hr)	3	81447	1040	0.45	0.53	84.0%	374	1058	2042	3339	5829
Assembly 6	3 diadromous in SW (24hr)	3	69402	1054.84	0.47	0.51	82.3%	377	1079	2128	3393	6187
Assembly 7	1 from each population in SW (24hr)	3	68188	1648.13	0.48	0.47	84.1%	769	1856	2826	4026	6378
Assembly 8	3 Toltén resident in FW (24hr)	3	64941	1512.2	0.48	0.53	80.7%	681	1707	2654	3842	6168
Assembly 9	3 Valdivia resident in SW (24 hr)	3	63822	1600.07	0.48	0.53	82.5%	737	1777	2737	3974	6427
Assembly 10	handling controls (0 ppt, 24 hr)	6	101516	1414.17	0.46	0.5	86.1%	613	1508	2463	3692	6080

5.3.6 RNAseq analyses

Analyses of RNA-seq data included two main objectives: 1) Identifying differentially expressed genes (DGE) that distinguished the gene expression of individuals in fresh water (0 hr) and salt water (24 hr) within each population (diadromous, Toltén, and Valdivia). And 2) comparing if individuals from both resident populations had a similar transcriptomic response to the salinity increment as diadromous individuals (i.e. if the same genes that were up- or down-regulated in the diadromous population were also up- or down-regulated in the resident populations). To accomplish these objectives, we first needed to distinguish the effects of general handling stress (i.e., water changes) from the effects of salinity (Fig. 5.2B).

The effects of general handling stress were examined in two different ways. First, we identified and removed "stress-responsive" genes found in the assembly built with individuals from the handling control (0 ppt, 24 hr – Table 5.1 assembly 10), but not the assembly built with individuals from the pre-transfer (0 ppt, 0 hr – Table 5.1 assembly 3) using bash commands. Freshwater fish at 0 hr (used in assembly 3) were kept in fresh water for at least eight days without any major stress (tanks were cleaned by removing only 10% of water), while the freshwater fish sampled at 24 hours experienced the same water changes as the experimental fish at 23-25 ppt. Thus, the comparison of assembly 3 to assembly 10 should detect genes responsive to stress, but not salinity (as all are at 0 ppt). Note that this step was conducted because the sample size for freshwater fish at 24 hours was not sufficient to allow for a direct comparison of fish at the two salinities with the same history of stress. By excluding transcripts found only in fish after "handling-stress" we reduced the likelihood that differences in gene expression between fresh and salt water were due to handling stress alone. However, this would not exclude genes that are stress-responsive but detected in both fresh water 0- and 24-hour assembly, so we do note that the effect of salinity and handling cannot be unequivocally separated in this work. Thus, secondly, because some genes are differentially expressed during both salinity and handling stress, we directly compared the gene expression of fish at 0 ppt and 23-25 ppt at 24 hr. Differential gene expression analysis was done using the library DESeq2 (Love *et al.* 2014) and 2FoldChanges were calculated for each comparison and statistical significance was estimated based on false discovery rate (FDR) adjusted P-

values <0.01 (Benjamini and Hochberg1995). We used the UniProt database (www.uniprot.org) to identify the proteins characterized by the affected genes.

To answer our first objective to detect genes that responded to salinity within each population, we compared the DGE of individuals in fresh water (0hr) vs. those in salt water (24hr) for each of the three populations using also the DESeq2 library and 2FoldChanges. This analysis excluded the "stress-responsive" genes identified from the comparison of assemblies from freshwater fish pre- and 24 hr post salinity increase (Fig. 5.2B). The patterns of expression were also assessed using PCAs. Finally, to answer our second objective, using the DGE with known proteins found in the diadromous population as a reference, we identified if the same genes were present in one or both resident populations and if the response (up- or down-regulation) was similar.

5.4 Results

5.4.1 Survival and physiological responses to salinity change

All diadromous and Valdivia residents survived the abrupt change in salinity from 0 ppt to 23-25 ppt (excluding the fish sampled for muscle water content measures). Conversely, some Toltén resident individuals died within 48 hours after exposure to 23-25 ppt. In fact, in one of the three replicate Toltén tanks, all individuals died before the end of the experiment (Fig. 5.3, tank 2), thus no sample was collected at 48 hours for this tank.

No significant change over time (from 0 to 48 hr) in % of water content was found in any of the three handling control tanks kept at 0 ppt (1 tank per population), suggesting that the stress of water replacement alone did not result in mortality (Friedman test $p > 0.05$), nor in any intracellular osmotic stress (Fig. 5.3). Among diadromous individuals, there were tank effects; two replicate tanks did not show a significant change over time in % of water content when exposed to 23-25 ppt water, but one did (tank1, Friedman test $p \leq 0.01$). In general, diadromous individuals tolerated the abrupt salinity increase to 23-25 ppt without experiencing significant osmoregulatory stress. Conversely, Toltén resident individuals exhibited a significant drop in % of water content in all three replicates, indicating they could not effectively osmoregulate at 23-25 ppt (Fig. 5.3B). Valdivia resident individuals showed a small drop, significant in two of the three replicate tanks, in the % of water content between 16 and 24 hours post salinity change

but by 48 hours they were able to return muscle water % to control levels (Fig. 5.3C). These results indicate that Valdivia residents can still activate the physiological response required to survive an abrupt increase in salinity and acclimate to restore homeostasis within 48 hours post-change. Together, these data suggest diadromous and Valdivia resident populations are tolerant to drastic increases in salinity, but Toltén resident fish are not.

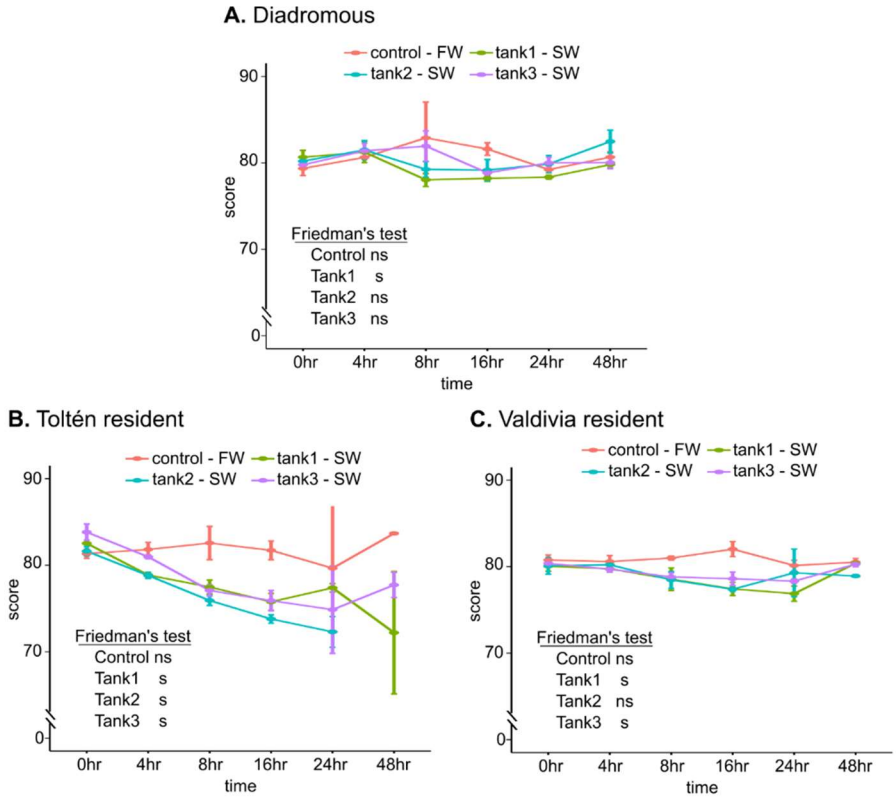


Figure 5.3. Osmoregulatory abilities of diadromous and resident populations as inferred from changes in percentage of muscle water content over time **A.** diadromous **B.** Toltén residents, and **C.** Valdivia residents. "ns" indicates not significant ($p > 0.05$ Friedman's test) and "s" indicates significant ($p < 0.05$). "FW" = fresh water and "SW" = salt water.

5.4.2 RNAseq analyses

Assemblies built with individuals from different populations and exposed to fresh water or salt water showed variation in the number of transcripts detected and the percentage matching Actinopterygii genes (Table 5.1). The assembly of diadromous individuals in fresh water (Table 5.1 - assembly 2) had 85214 transcripts, 76.8% of which matched Actinopterygii genes as assessed by the BUSCO analysis, but the assembly of

diadromous individuals in salt water had 69402 transcripts, 82.3% of which matched Actinopterygii genes (Table 5.1 - assembly 6).

We explored if the transcriptomic response would vary if the *de novo* transcriptome used as reference was built with individuals from different populations. For example, to test if indeed the origin of the samples selected to build the assemblies affected the results, we compared the assembly built with only diadromous individuals at 0 hr (the predicted ancestral population, Table5. 1 – assembly 2) and the assembly built with individuals from each of the three populations at 0 hr (diadromous, Toltén resident, and Valdivia resident, Table 5.1 – assembly 3). Both assemblies showed similar Principal Component Analysis (PCA) patterns of gene expression (see Fig. S5.1) suggesting that despite the differences in the number of transcripts found in the different assemblies the general trend of transcriptomic response is maintained. In both analyses, little differentiation was found between resident individuals in fresh water. Some diadromous individuals, however, could be distinguishable from other diadromous individuals in both analyses. We did not find a relation between batch, number of reads, and % of muscle water content that could explain the difference between these distinguishable diadromous individuals.

Subsequent RNAseq analyses were focused on answering our two main questions: 1) What are the effects of salinity increase within each of the three populations? And 2) To what extent does the transcriptomic response to salinity increase differ between the resident populations and the diadromous population? Specifically, which genes differ in their salinity responsiveness between the saltwater intolerant Toltén fish compared to the other two saltwater tolerant populations? To accomplish these goals, we first identified "stress-responsive" genes associated with water replacement and tested if the expression of these genes also responded to salinity.

5.4.3 "Stress-responsive" genes

To identify genes that were associated with the stress of the water replacement alone, we compared the response between fish in fresh water pre-transfer (0 hr) and post-transfer (24 hr). Since the only difference is the water change stress, these transcriptomic changes should not be related to salinity change. We found a total of 24705 "stress-responsive" transcripts putatively associated with the stress induced by the water

replacement in fish maintained at 0 ppt by comparing the assemblies at 0 hours to the assemblies at 24 hours in fresh water (Fig. 5.4A). To test for the effects of the "stress-responsive" genes on the transcriptomic response of the three populations, we used a reference built with diadromous individuals (the ancestral trait) in salt water (Table 5.1 - assembly 6) but excluded genes associated with the stress of the replacement of water (Fig. 5.4A). A plot of Principal Components 1 and 2 of a PCA shows individuals grouping by population (Fig. 5.4B) with those from the handling control tanks clustering by themselves slightly separated from their own population. This indicates that stress-induced changes in mRNA content as a response to the replacement of water and handling in the experimental tanks are driving a certain level of differentiation. Fig 5.4C which excludes these "stress-responsive" genes, shows that the PCA maintains the differentiation between populations, but in this case, the difference between the two resident populations is not as profound; individuals from one resident population are more similar to individuals from the other resident population than individuals from either population are to diadromous individuals (Fig. 5.4C).

The analysis of the effects of handling stress on the expression of all genes (fresh water at 0 hr vs. fresh water 24 hr) (Fig. 5.4B) further suggests that salinity does not have a major physiological impact on diadromous individuals. PCA of gene expression shows that the diadromous fish in the handling controls (0 ppt, 24 hr) group closely with the diadromous individuals in salt water (23-25 ppt, 24 hr), although one control individual is slightly separated. However, the handling controls (0 ppt, 24 hr) of Valdivia are separated from the Valdivia individuals in salt water (23-25 ppt, 24 hr) and the Toltén freshwater (24 hr) are even more clearly differentiated from Toltén individuals in salt water (23-25 ppt, 24 hr).

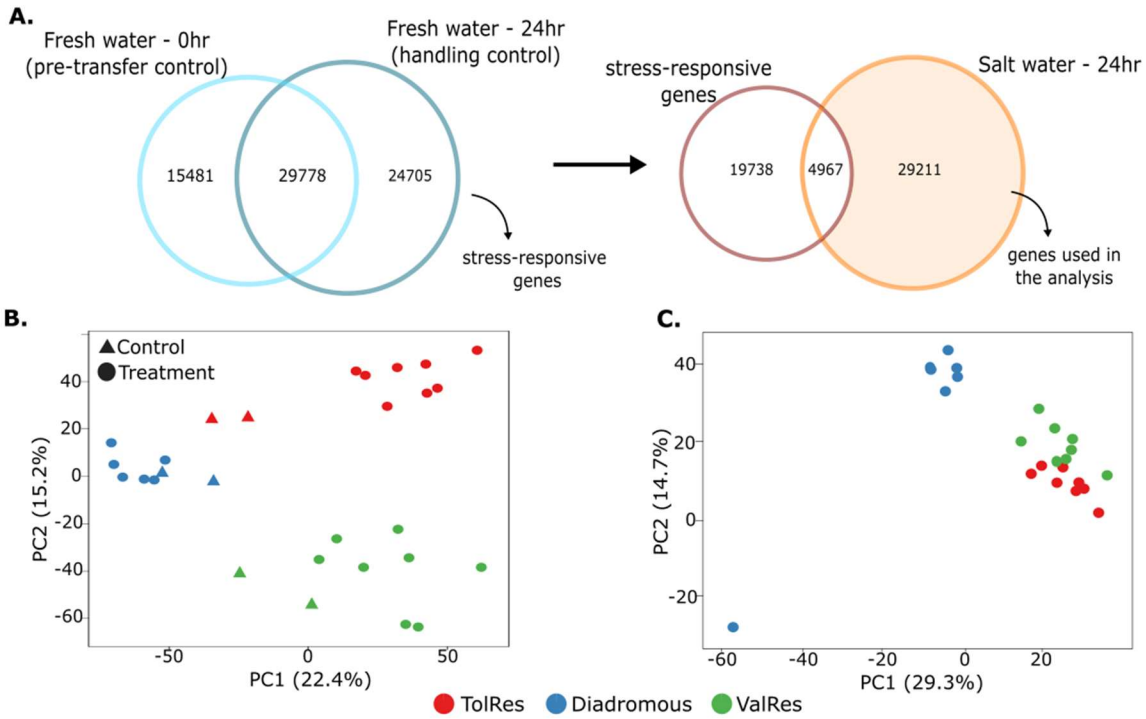


Figure 5.4. A. Transcriptional responses to salinity (0 vs. 23-25 ppt) and handling stress (0 hr vs. 24 hr) among populations **A.** The number of genes that could be attributed to the stress of water replacement termed "stress-responsive genes", and the number of genes used in the analysis that excluded the "stress-responsive genes" detected by comparing genes found in the pre- and post-transfer freshwater assemblies (Table 5.1 – assemblies 3 and 10). Principal Component Analysis (PCA) of gene expression **B.** including stress-responsive genes (analysis considering 34178 genes) and handling control samples (24 hr at 0ppt) and treatment samples (24 hr at 23-25 ppt), and **C**. to clarify the response to salinity among populations while excluding all genes that are considered stress-responsive genes (analysis considering only 29211 genes in (A)).

As some of these "stress-responsive" genes might have a role in salinity tolerance too, we performed a DGE analysis using these "stress-responsive" genes between handling controls (0 ppt, 24 hr) and individuals from the same population in salt water (23-25 ppt, 24 hr). The DGE analysis revealed 144 differentially expressed genes (p_{adj} <0.01) in the diadromous population, 145 in the tolerant Valdivia population, and 789 in the intolerant Toltén population. From the 144 differentially expressed genes in the diadromous population, 87 have characterized proteins, 47 of which were upregulated and 40 were downregulated. From the 145 genes differentiating Valdivia's controls, 78 have characterized proteins, 47 of which are upregulated and 31 are downregulated. And from the 789 genes differentiating Toltén's controls, 444 have characterized proteins, 264

of which are upregulated and 180 are downregulated (Table S5.2). Only two genes were found in common in all three comparisons (individuals in salt water vs. handling controls at 24 hr): Sodium/hydrogen exchanger (NHE) and a transporter (solute carrier family 6). In all three comparisons, these two genes were upregulated, meaning their gene expression increased significantly in salt water. This indicates that the expression of these genes was affected by both the water replacement and salinity.

5.4.4 Objective 1. Effects of salinity increase within populations

To better examine the isolated effects of salinity on gene expression, we removed all "stress-responsive" genes from our analysis; these were genes detected to vary between assemblies from fresh water exposed fish at 0 hr and 24 hr (Table 5.1-assemblies 3 and 10). This would exclude any genes that are responsive to both salinity and stress, so is likely a conservative estimate. Using the assemblies obtained from individuals from each of the three populations in fresh water at 0 hr and assemblies from individuals in salt water (23-25 ppt) at 24 hr (Table 5.1) we identified genes (after excluding the "stress-responsive" genes) that were differentially expressed in salt water (24hr) vs. fresh water (0hr) for each of the three populations. Overall, we observed that the transcriptomic response of individuals, when exposed to salt water is similar among individuals within each of the three populations as individuals in salt water grouped closely together (Fig. 5.5). For diadromous individuals which were able to physiologically cope with the salinity change, DGE analysis revealed only 398 significantly differentially expressed genes ($p_{adj} < 0.01$). This result was obtained after excluding two individuals in fresh water that showed large differences, yet we could not identify any variable that could explain these differences (Fig. S5.2). Both resident populations exhibited a higher number of differentially expressed genes between fresh water (0 ppt, 0 hr) and salt water (23-25 ppt, 24 hr) than the diadromous population, with 2339 DGE in Valdivia, and 2932 in Toltén.

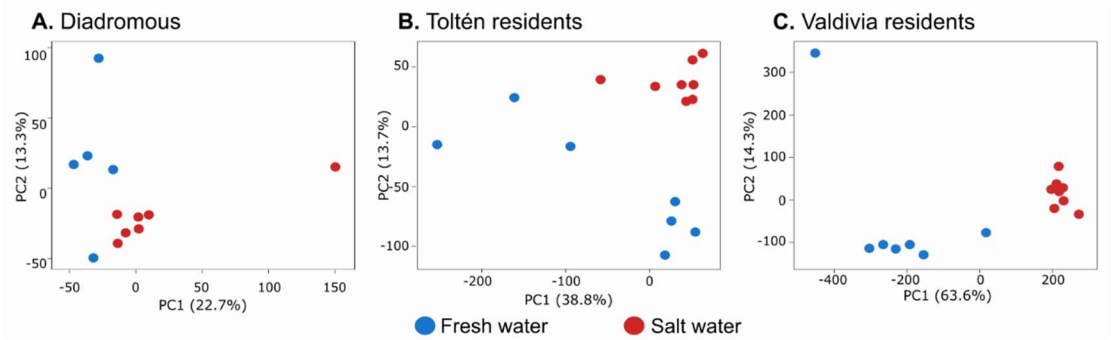


Figure 5.5. The effects of salinity on gene expression are much greater in freshwater than diadromous populations. Principal Component Analysis (PCA) of the transcriptomic response at 0 hours in fresh water (FW) and at 24 hours in salt water (SW) of **A.** diadromous, **B.** Toltén resident, and **C.** Valdivia resident individuals. Note that this figure excludes stress-response genes (see Fig. 5.4).

5.4.5 Objective 2. Comparison of transcriptomic response among populations

As diadromous adult individuals are expected to retain the ability to survive in salt water and showed no or little physiological stress, we examined the 398 DGE that showed to up- or down-regulated when a diadromous individual was transferred from fresh water to salt water (Fig. 5.6). From these 398 genes, 203 had characterized proteins (Table S5.3) and 59 of these genes were also found to be DGE ($p_{adj} < 0.01$) in Toltén and/or Valdivia individuals (Fig. 5.6B, C). Of these 59 genes, 37 were upregulated in diadromous samples (Fig. 5.6B). Of these upregulated 37 genes, 27 were also upregulated in Valdivia individuals, and 22 were upregulated in Toltén individuals. The remaining 22 of the 59 differentially expressed genes with known proteins were downregulated in diadromous samples (Fig. 5.6C). Of these, only 6 were downregulated in Valdivia individuals, and 18 were downregulated in Toltén individuals (Fig. 5.6C).

Ion transport proteins known to play a crucial role in osmoregulation via gill ionocytes in fishes (reviewed by Hwang *et al.* 2011) were detected in our analysis, including cystic fibrosis transmembrane conductance regulator (CFTR) and Na(+)/K(+)-ATPase (NKA). NKA subunit alpha was upregulated after salt water change, respectively, in all populations (Hwang *et al.* 2011). However, CFTR, a key apical membrane transporter responsible for excreting Cl- from ionocytes, and NKA subunit beta were only significantly upregulated in diadromous and Valdivia resident individuals in 23-25 ppt. This suggests that the absence of upregulation of CFTR and NKA subunit

beta among Toltén individuals may be responsible for their inability to acclimate to salt water. Yet, there are other 13 genes upregulated in diadromous and Valdivia resident individuals but not in Toltén resident individuals that could potentially also be critical for acclimation to salt water in *G. maculatus* (Fig 5.6B).

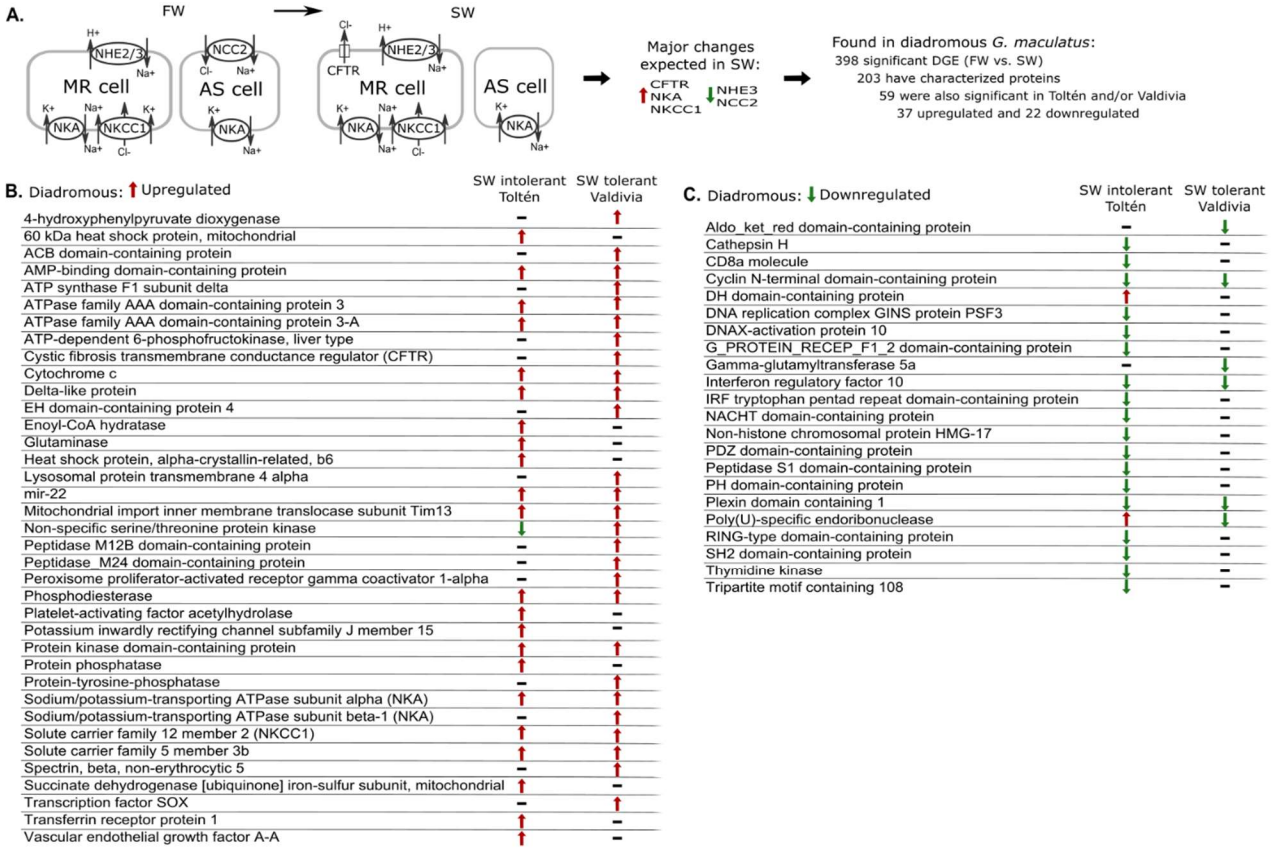


Figure 5.6. A. Variation in gene expression of characterized proteins in response to salinity among populations **A.** Model of some key ion transporters found in teleost gill ionocytes in fresh water (FW) and salt water (SW) that are expected to be upregulated (red arrow) or downregulated (green arrow) when fish are transferred from FW to SW. The number of significant differentially expressed genes (DGE, padj < 0.01) found in diadromous individuals after FW to SW change is noted on the right. **B.** Genes found to be upregulated or **C.** downregulated in diadromous individuals and also found to be significantly differentiated (padj < 0.01) in Toltén resident or Valdivia resident individuals are shown with an arrow (red = upregulated, green = downregulated), and genes not found to be significant differentiated in the resident populations are represented with a dash.

5.5 Discussion

Here we investigated the transcriptomic response to an acute salinity change (0 ppt to 23-25 ppt) in a diadromous population of *Galaxias maculatus* and two resident populations (Toltén and Valdivia) derived from a common diadromous source, and which are considered natural replicates for the loss of diadromy. We focused on salinity adaptation as this variation is key to understanding the evolution of diadromous populations (Nakamura *et al.* 2021). Although Delgado *et al*. (2020; Chapter 4) found that both resident populations exhibited similar levels of genetic differentiation with the diadromous population (i.e., F_{ST} ~0.4), they differed in their ability to gradually acclimate to saltwater environment, such that Toltén fish were less tolerant of salt water than Valdivia fish (Chapter 4; Delgado *et al.* 2020). After lab acclimation to 0 ppt, diadromous and Valdivia resident individuals survived an abrupt increase in salinity (Fig. 5.3A, 5.3C), as would be experienced by larvae that are washed away from the estuary to the sea. Toltén resident individuals, instead, did not survive the abrupt salinity increase and exhibited a continuous drop in the % of water content in muscle just four hours post salinity change, indicating osmoregulatory failure led to their death (Fig. 5.3B). Overall, these results confirm that diadromous *G. maculatus* can quickly acclimate to a wide range of environmental salinities (Urbina & Glover 2015) and indicate that while Valdivia residents have maintained the genetic machinery required to osmoregulate in saltwater environments, Toltén residents have not. This loss of saltwater acclimation makes these *G. maculatus* populations a good comparative system to study the transcriptomic signatures associated with the loss of upper salinity tolerance.

Below we address the following questions: What genes are important for osmoregulatory acclimation to salt water in diadromous and Valdivia *G. maculatus* populations? Have the saltwater tolerant Valdivia residents maintained all the molecular mechanisms required to cope with salt water at the gill? Which transcriptomic responses have been lost in saltwater intolerant Toltén residents? and what processes result in the present-day variation in Toltén and Valdivia residents?

5.5.1 Osmoregulatory gene expression in response to increased salinity

We observed 203 genes with characterized proteins that were differentially expressed as a response to salinity changes (0 to 23-25 ppt) among diadromous

individuals (Table S3). Among these genes, Sodium/potassium ATPase (NKA), Cystic fibrosis transmembrane conductance regulator (CFTR), and solute carrier family 12 member 2 (NKCC1) stand out as they are known to be present in 'saltwater' gill ionocytes and have a key role in osmoregulation in other euryhaline and diadromous species (Hiroi & McCormick 2012). In particular, CFTR, NKCC1, and NKA subunit alpha and beta work together to excrete sodium and chloride ions that diffuse into fish in sea water to help maintain osmoregulatory homeostasis (see Fig. 5.6A). For example, NKA enzyme activity is upregulated in diadromous *G. maculatus* when salinity was changed from fresh water to salt water (Ruiz-Jarabo *et al.* 2016), and its activity increases as early as 8 hours post salinity change and remains high for up to 72 hours (Urbina *et al.* 2013). CFTR has also been found to have an important ion transporter function in gills when exposed to sea water in other species (e.g., anadromous Takifugu species (Nakamura *et al.* 2021) and euryhaline Killifish (Scott *et al.* 2004)). Sodium/hydrogen exchanger (NHE), key in saltwater ionocytes (Hwang *et al.* 2011) was found to be responsive to both salinity and handling stress, but in salt water (24 hr) was significantly upregulated in the three populations. This result is different from what has been reported in Three-spined Stickleback (Gibbons *et al.* 2018) and Killifish (Scott *et al.* 2004) where NHE expression increased but when transferred to fresh water, but not sea water.

Three of five well-known genes involved in the osmoregulatory response, NKA subunit alpha, NKCC1, and NHE, were indeed upregulated in all three populations. The CFTR and NKA subunit beta genes were upregulated only among the diadromous and tolerant Valdivia resident individuals and not among the intolerant Toltén resident individuals. This finding is consistent with the hypothesis that CFTR and NKA subunit beta are important osmoregulatory genes in *G. maculatus* and a loss of transcriptional induction of these genes during transfer to salt water may limit tolerance in the Toltén fish. However, 12 other genes (see Fig. 5.6B) were also upregulated among diadromous and Valdivia resident individuals but not among the Toltén resident individuals in the gill, and one, non-specific serine-threonine protein kinase, was even significantly downregulated in Toltén residents (Fig. 5.6B). These 13 genes could also have an important role in salinity acclimation, yet knowledge of their ecological or functional role is lacking at present.

Analysis of gene expression clearly shows that although Valdivia residents can acclimate to salt water, not all differentially expressed genes found among the diadromous individuals were also found among the Valdivia resident individuals (Fig. 5.6B and 6C). For example, 10 genes were only upregulated in the diadromous and intolerant Toltén resident individuals, suggesting they may not be key for beneficial salinity acclimation. Surprisingly, the tolerant Valdivia residents also exhibited fewer similarities with diadromous individuals among the downregulated genes than intolerant Toltén residents did. While Toltén residents share 18 commonly downregulated genes with the diadromous individuals, Valdivia residents share only six downregulated genes. This finding differed from our initial hypothesis that gene expression of Valdivia resident individuals would be more similar to the diadromous than Toltén residents. Thus, Valdivia residents have not retained the expression of all genes from their diadromous ancestor, but most likely by chance, the ones that it has retained are key for osmoregulation.

The fact that we do find other important ion transporting genes such as solute carrier families and NKA subunit alpha upregulated in Toltén residents, indicates that Toltén residents do retain some of the plastic mechanisms involved in osmoregulation. In fact, ionic homeostasis is always required, perhaps except in the isosmotic point where fish and medium have the same concentration of salts and osmolites (Urbina & Glover 2015). This retention of expression of some osmoregulatory genes may also indicate that genes involved in osmoregulation capacity are regulated by different transcription factors. The partial loss of osmoregulatory capacity has also been reported in other species such as Alewife (Velotta *et al.* 2017), and different patterns of change in gene expression found in these two resident populations, have also been reported in sticklebacks (Gibbons *et al.* 2017).

5.5.2 Evolutionary processes influencing resident populations post-colonization

Given that these populations are expected to have derived from a common diadromous *G. maculatus* population, standing genetic variation would be expected to be similar between the two resident populations. We do observe some level of parallelism in gill mRNA content in response to salinity change post-colonization, as some of the differentially expressed (DGE) are shared between the two resident populations. A small

number of osmoregulatory genes exhibiting parallelism have also been reported for replicate populations of resident Alewife (Velotta *et al.* 2017) and Nine-spined Stickleback (Wang *et al.* 2020).

Yet, the variation in differential gene expression we found when comparing Valdivia residents with diadromous versus what we found when comparing Toltén residents with diadromous, suggests drift may at least be partially responsible for the maintenance or loss of the ability to osmoregulate. Key salinity-responsive transcriptomic responses were maintained among Valdivia residents but not among Toltén residents. Although evidence suggests both resident populations were colonized over similar evolutionary times (Chapter 4; Delgado *et al.* 2020), the number of founders may have differed between populations, affecting genetic diversity. Thus, estimating the time since colonization and isolation from the diadromous population and the effective size may shed light on the issue, as has been suggested that time since the access to standing genetic variation can be a major factor for gene reuse (Liu *et al.* 2018). Environmental similarities are also a good predictor for parallelism (Magalhaes *et al.* 2021), in such cases, the differences in gene expression could also indicate that environmental conditions do differ between these geographically close populations with local adaptation playing a role in the evolution of these resident populations.

5.5.3 Experimental considerations and future directions

Table S5.1, shows that the number of raw reads per sample varied between 30M and 80M, yet bioinformatic pipelines include steps to normalize the data, so this difference should not affect the differential gene expression analysis. In Fig. 5.4A, a percentage of the genes found are likely related to the stress of water replacement and fish handling. This highlights the importance of handling controls, not only in common garden experiments but also in the transcriptomic analysis of acute exposures, as the handling of fishes clearly affected gene expression (Liu *et al.* 2014). Recommendations for RNAseq studies tend to commonly highlight the importance of replicates and sequencing depth (Todd *et al.* 2016), yet handling controls should also be more strongly recommended as a standard practice.

Our results focus on gills as this tissue is the main site of sodium, chloride, and potassium regulation, but osmoregulation is achieved by the coordinated functioning of

several tissues including gills (Urbina *et al.* 2013), esophagus (Brijs *et al.* 2015), kidneys, and intestine (Grosell 2006). As such, our gill transcriptome results only partially represent the full osmoregulatory response at the molecular level. The examination of differential gene expression in these different organs and/or different cell types is expected to lead to an increased understanding of osmoregulatory capacities in *G. maculatus*. Furthermore, to better understand the loss of diadromy is important to look at how fish cope with other environmental stressors that vary in resident and migratory populations and study the mechanism underlying the decision to migrate. In particular, gene expression in the brain may provide important clues regarding the changes in behavior between diadromous and resident populations (Chapter 2; Delgado & Ruzzante 2020).

In this study, we examined the differential gene expression within and among adult individuals of three populations, yet larvae represent the life stage during which *G. maculatus* individuals migrate. Although we believe these results still provide important input regarding the molecular mechanisms differentiating populations of diadromous and resident fishes, future studies with larvae will be necessary to provide an improved understanding of the gene expression changes that take place during exposure to salt water.

Galaxias maculatus has the potential to become a model organism to study the evolution of recurrent colonization and loss of diadromy and mechanisms underlying the evolution of salinity tolerance. The examination of other populations within the Chilean distribution as well as among populations in New Zealand and Tasmania are likely to assist us in answering key questions including how populations evolved? what is their adaptive potential? and to what extent has plasticity been maintained through the species' global distribution?

5.6 Acknowledgments

Sample collection and the challenge experiment were funded by FONDECYT REGULAR- 1150154 grant to EH, CONICYT-FONDECYT-11160019 and FONDECYT REGULAR- 1210071 grants to MAU. The transcriptomic work was funded by an NSERC Discovery grant to DER (RGPIN-2019-04679). MLD was supported by the NSERC DG grant to DER, Dalhousie's Faculty of Graduate Studies Scholarship, and the

Lett fund from the Department of Biology of Dalhousie University. We thank Compute Canada and their support team for providing the frame to run the transcriptomic analyses. We also thank Pablo Fierro and Gustavo Díaz for their assistance in the field and René Iribarren for his assistance in rearing the fish.

5.7 Author contribution

This manuscript is part of MLD's PhD thesis research. All authors contributed to the conception and design of the study. Fish were collected by MLD and AM. Data from the salinity change experiment was collected by MLD, AM, and MAU. MLD generated and analyzed the transcriptomic data. The manuscript was written primarily by MLD with input from DER and ACD, and contributions by MAU, EH, and OL.

5.8 Supplementary information

Table S5.1. Number of reads obtained per sample after sequencing (NovaSeq).

ID	Description	# of reads (Millions)
S01	Dia0hr	80
S61	Dia0hr	56
S62	Dia0hr	61
S63	Dia0hr	50
S64	Dia0hr	56
S67	Dia0hr	32
S82	Dia0hr	80
S23	Dia24hr	62
S24	Dia24hr	30
S25	Dia24hr	56
S27	Dia24hr	80
S28	Dia24hr	80
S68	Dia24hr	72
S87	Dia24hr	51
S29	Dia24hr - Control	80
S30	Dia24hr - Control	45
S16	TolRes0hr	61
S17	TolRes0hr	65
S15	TolRes0hr	81
S18	TolRes0hr	37
S19	TolRes0hr	80
S20	TolRes0hr	30
S21	TolRes0hr	60
S42	TolRes24hr	42
S43	TolRes24hr	44
S44	TolRes24hr	44
S74	TolRes24hr	34
S92	TolRes24hr	49
S40	TolRes24hr	65
S45	TolRes24hr	37
S75	TolRes24hr	59
S47	TolRes24hr - Control	67
S48	TolRes24hr - Control	60
S08	ValRes0hr	51
S11	ValRes0hr	80
S14	ValRes0hr	64
S83	ValRes0hr	65
S84	ValRes0hr	77
S85	ValRes0hr	72

ID	Description	# of reads (Millions)
S86	ValRes0hr	59
S31	ValRes24hr	51
S36	ValRes24hr	80
S73	ValRes24hr	43
S88	ValRes24hr	31
S90	ValRes24hr	57
S69	ValRes24hr	33
S72	ValRes24hr	51
S89	ValRes24hr	45
S39	ValRes24hr - Control	48
S91	ValRes24hr - Control	68

Table S5.2. Genes with known proteins responsive to both water replacement and salinity obtained from comparing expression of individuals in freshwater (0 ppt, 24hr) and in salt water (23-25 ppt, 24 hr).

Diadromous	Valdivia resident	Toltén resident
1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase	3'-phosphoadenosine-5'-phosphosulfate synthase	[histone H4]-N-methyl-L-lysine20 N-methyltransferase KMT5B
3'-phosphoadenosine-5'-phosphosulfate synthase	Aa_trans domain-containing protein	10 kDa heat shock protein, mitochondrial
60 kDa chaperonin	Aamy domain-containing protein	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta-1
Abhydrolase domain containing 14B	Alpha-2C adrenergic receptor	26S proteasome non-ATPase regulatory subunit 13
Actin, cytoplasmic 2	Apolipoprotein D	26S proteasome regulatory subunit 8
ADAMTS-like protein 3	ATPase family AAA domain containing 2	3-hydroxy-3-methylglutarate-CoA lyase
Adenylate kinase 4, mitochondrial	Carn_acyltransf domain-containing protein	3-hydroxy-3-methylglutaryl coenzyme A reductase
Annexin	Carnosine synthase 1	3'-phosphoadenosine-5'-phosphosulfate synthase
ATP-dependent RNA helicase DHX30	Caseinolytic mitochondrial matrix peptidase chaperone subunit	40S ribosomal protein S11
BicC family RNA binding protein 1	CD209 antigen-like protein 2	40S ribosomal protein S12
C2 domain-containing protein	CDC42 effector protein (Rho GTPase binding) 1a	40S ribosomal protein S13
C2 domain-containing protein 5	Ceramidase	40S ribosomal protein S14a
Calcium-activated neutral proteinase 2	Clustered mitochondria protein homolog	40S ribosomal protein S15a
Carboxypeptidase Z	Coiled-coil domain containing 12	40S ribosomal protein S2
Chromobox 8	Collagen alpha-2(I) chain	60 kDa chaperonin
Clustered mitochondria protein homolog	Collagen alpha-6(VI) chain	60S acidic ribosomal protein P0

Diadromous	Valdivia resident	Toltén resident
Collagen alpha-1(XVII) chain	Condensin complex subunit 3	60S ribosomal protein L13a
Collagen alpha-1(XVII) chain	Cyclin dependent kinase 6	60S ribosomal protein L27
Collagen, type XVII, alpha 1b	Cytoglobin-2	60S ribosomal protein L30
Cyclic GMP-AMP synthase	Dimethylargininase	60S ribosomal protein L31
Cytochrome P450, family 1, subfamily C, polypeptide 2	DNA (cytosine-5)-methyltransferase 1	60S ribosomal protein L34
Cytochrome P450, family 2, subfamily N, polypeptide 13	DNA helicase MCM8	60S ribosomal protein L36
Cytochrome P450, family 3, subfamily A, polypeptide 65	DNA polymerase	60S ribosomal protein L7a
Dimethylargininase	DNA polymerase alpha subunit B	AA_permease_C domain-containing protein
Dynamin-binding protein	DNA repair protein RAD51 homolog	Aa_trans domain-containing protein
E3 ubiquitin-protein ligase MYLIP-B	Double-strand break repair protein	AAA_16 domain-containing protein
FBJ murine osteosarcoma viral oncogene homolog B	Double-stranded RNA-specific adenosine deaminase	Aamy domain-containing protein
Frizzled-7	Dynamin-binding protein	AB hydrolase-1 domain-containing protein
Glypican-1	E3 ubiquitin-protein ligase MYLIP-B	Abhydrolase domain containing 14B
Golgin subfamily A member 6-like protein 2	Endopeptidase S2P	Abnormal spindle-like microcephaly-associated protein homolog
Grass carp reovirus (GCRV)-induced gene 2p	Eukaryotic translation initiation factor 3 subunit F	Acetyl-coenzyme A synthetase
Heat shock cognate 70	Family with sequence similarity 13 member A	Acidic leucine-rich nuclear phosphoprotein 32 family member E
Heat-stable enterotoxin receptor	FBJ murine osteosarcoma viral oncogene homolog B	Actin, cytoplasmic 2

Diadromous	Valdivia resident	Toltén resident
Hsd17b3 protein	Fibrillar collagen NC1 domain-containing protein	Activated leukocyte cell adhesion molecule b
Hydroxysteroid 17-beta dehydrogenase 3	Flap endonuclease 1	Activated RNA polymerase II transcriptional coactivator p15
IF rod domain-containing protein	Glycerol-3-phosphate acyltransferase 3	Activator of basal transcription 1
Insulin-like growth factor-binding protein 6b	GRAM domain-containing 2Aa	Acyl-CoA-binding domain-containing protein 7
Integrin beta	Hexosyltransferase	Adhesion G-protein coupled receptor F2
Integrin subunit alpha 2	IF rod domain-containing protein	Adipocyte plasma membrane-associated protein
Interleukin 13 receptor, alpha 2	Inactive carboxypeptidase-like protein X2	ADP/ATP translocase
Intraflagellar transport 140	L-2-hydroxyglutarate dehydrogenase	ADP-ribosylation factor GTPase activating protein 2
Isochorismatase domain-containing protein	L-ornithine N(5)-monooxygenase	AH domain-containing protein
Jacalin 4	Lymphocyte cytosolic protein 1	AIG1-type G domain-containing protein
Kinesin-like protein	MFS domain-containing protein	Akirin-2
Malate dehydrogenase	Occludin	Alkaline ceramidase
Matrix metallopeptidase 13a	Oxidative stress induced growth inhibitor 1	Alpha-2C adrenergic receptor
Matrix metalloproteinase-9	Peptidylprolyl isomerase	Alpha-ketoglutarate-dependent dioxygenase FTO
Membrane-spanning 4-domains subfamily A member 4D-like	Peroxisome proliferator-activated receptor gamma coactivator 1-alpha	Alpha-L-fucosidase
Metallophos domain-containing protein	Phosphate transporter	Aminopeptidase
Metalloproteinase inhibitor 2	Plastin-3	AMP-binding domain-containing protein

Diadromous	Valdivia resident	Toltén resident
Methionine aminopeptidase 2	PNPLA domain-containing protein	Anillin
Mitochondrial Rho GTPase	Polymerase (DNA directed), epsilon 3 (p17 subunit)	ANK_REP_REGION domain-containing protein
Myoglobin	POU domain protein	Ankyrin repeat and EF-hand domain-containing protein 1
Myosin regulatory light chain 2a	Probable E3 ubiquitin-protein ligase makorin-1	Ankyrin repeat domain 22
Myosin-6	Protein AF1q	Ankyrin-1
Myosin-7	Protein timeless homolog	Annexin
N-acetylgalactosaminide beta-1,3-galactosyltransferase	RAB20, member RAS oncogene family	Apolipoprotein D
NEDD4-binding protein 2-like 1	Replication factor C (activator 1) 4	ArfGAP with FG repeats 1a
Neurogenic locus notch homolog protein 1	Reticulon	Argininosuccinate synthase
Nidogen-1	Rho GTPase-activating protein 32b	ATPase family AAA domain containing 2
Nuclear factor of activated T cells 5a	SAM domain-containing protein	ATP-binding cassette, sub-family A (ABC1), member 12
Nucleoporin NSP1	Septin-type G domain-containing protein	ATP-binding cassette, sub-family F (GCN20), member 2a
Peptidase domain-containing-associated with muscle regeneration 1	SH3 domain-containing protein	Atrial natriuretic peptide-converting enzyme
Peroxisome proliferator-activated receptor gamma coactivator 1-alpha	Shootin-1	Bisphosphate nucleotidase 1
Poly(U)-specific endoribonuclease	Small monomeric GTPase	BK channel
Polypeptide N-acetylgalactosaminyltransferase	Sodium- and chloride-dependent taurine transporter	Brefeldin A-inhibited guanine nucleotide-exchange protein 3

Diadromous	Valdivia resident	Toltén resident
Protein kinase domain-containing protein	Sodium/hydrogen exchanger	B-related factor 1
Protein mono-ADP-ribosyltransferase PARP6	Sodium-dependent phosphate transporter 1-B	BRO1 domain-containing protein
Protein phosphatase, Mg2+/Mn2+-dependent, 1Ab	Solute carrier family 2 member 11b	BTB and CNC homology 1, basic leucine zipper transcription factor 1 a
Reverse transcriptase domain-containing protein	Spindle and kinetochore-associated protein 3	BZIP domain-containing protein
RFX-type winged-helix domain-containing protein	SRA1 domain-containing protein	C2 domain-containing protein
Rho GTPase-activating protein 32b	Sulfotransferase	C3H1-type domain-containing protein
Scavenger receptor class A member 5	Thioredoxin	C6 finger domain protein, putative
Secretogranin III	Thread biopolymer filament subunit alpha-like	Ca_chan_IQ domain-containing protein
Secretory carrier-associated membrane protein 5	Transcription factor Sp9	Cadherin EGF LAG seven-pass G-type receptor 1a
Sodium/hydrogen exchanger	Transmembrane protein 245	Calpain 9
Solute carrier family 10 member 1	Transporter (solute carrier family 6 member 6)	Calponin-homology (CH) domain-containing protein
Solute carrier family 25 member 15b	Vacuole membrane protein 1	Carbonic anhydrase
Spondin 2b, extracellular matrix protein		Carn_acyltransf domain-containing protein
Stork_head domain-containing protein		Caseinolytic mitochondrial matrix peptidase chaperone subunit
TIR domain-containing protein		CD209 antigen-like protein 2
TNFR-Cys domain-containing protein		CDC42 effector protein (Rho GTPase binding) 1a

Diadromous	Valdivia resident	Toltén resident
Transporter (solute carrier family 6 member 15)		Cell growth regulator with EF-hand domain 1
Tropomyosin 4		Cellular oncogene fos
von Willebrand factor A domain-containing protein 2-like		Centromere protein F
VWFA domain-containing protein		Centromere protein Q
Zinc finger protein 385B		Ceramidase
		Charged multivesicular body protein 5
		Cleavage and polyadenylation specific factor 1
		CLPTM1 regulator of GABA type A receptor forward trafficking
		Clustered mitochondria protein homolog
		Coiled-coil domain containing 12
		Collagen alpha-2(I) chain
		Collagen alpha-2(VIII) chain
		Collagen alpha-6(VI) chain
		Condensin complex subunit 3
		Condensin-2 complex subunit D3
		Cramped chromatin regulator homolog 1
		CS domain-containing protein
		CXXC-type zinc finger protein 1
		Cyclic AMP-responsive element-binding protein 5
		Cyclic nucleotide-gated channel subunit alpha 3a
		Cyclin dependent kinase 6

Diadromous	Valdivia resident	Toltén resident
		Cyclin-dependent kinase 15
		Cyclin-Y-like protein 1
		Cytochrome b5 heme-binding domain-containing protein
		Cytochrome b-c1 complex subunit 6
		Cytochrome b-c1 complex subunit Rieske, mitochondrial
		Cytochrome P450, family 3, subfamily A, polypeptide 65
		Cytoglobin-2
		Cytosolic Fe-S cluster assembly factor nubp2
		DDHD domain-containing 2
		DENN domain-containing protein 1A
		Derlin
		Diacylglycerol kinase
		Diamine acetyltransferase 1
		Dickkopf_N domain-containing protein
		Diphosphomevalonate decarboxylase
		Dirigent protein 10
		DNA (cytosine-5)-methyltransferase
		DNA helicase
		DNA polymerase
		DNA repair protein RAD51 homolog
		DNA-damage regulated autophagy modulator 2b
		DOCKER domain-containing protein

Diadromous	Valdivia resident	Toltén resident
		Dolichyl-phosphate-mannose--protein mannosyltransferase
		Double-stranded RNA-binding protein Staufen homolog 2
		Double-stranded RNA-specific adenosine deaminase
		Dynein axonemal heavy chain 1
		Dynein axonemal intermediate chain 3
		E3 ubiquitin-protein ligase
		E3 ubiquitin-protein ligase MYLIP-B
		E3 ubiquitin-protein ligase rnf213-beta
		EEF1A lysine methyltransferase 1
		EH domain-containing and endocytosis protein 1
		eIF-2-alpha kinase GCN2
		Elongation factor 1-beta
		Elongation of very long chain fatty acids protein 1
		Elongation of very long chain fatty acids protein 7
		EMI domain-containing protein
		Endopeptidase S2P
		Endoplasmic reticulum junction formation protein lunapark
		Engulfment and cell motility 3
		Envoplakin
		Ephrin type-B receptor 1

Diadromous	Valdivia resident	Toltén resident
		Epidermal growth factor receptor substrate 15-like 1
		Essential meiotic structure-specific endonuclease 1
		ETS variant transcription factor 5a
		Eukaryotic translation initiation factor 3 subunit F
		Exportin 5
		Extended synaptotagmin-like protein 1a
		FAM20A golgi associated secretory pathway pseudokinase
		FAM83 domain-containing protein
		Family with sequence similarity 160 member B2
		Family with sequence similarity 83 member Fa
		Fanconi anemia group I protein
		F-box only protein 45
		Feline leukemia virus subgroup C cellular receptor family, member 2a
		Fes1 domain-containing protein
		Filamin B
		Flap endonuclease 1
		Flotillin-2a
		Forkhead box C1-A
		Forkhead box protein G1
		F-spondin

Diadromous	Valdivia resident	Toltén resident
		G_PROTEIN_RECEP_F1_2 domain-containing protein
		Galectin
		Gap junction protein
		GCS light chain
		Glucosamine 6-phosphate N-acetyltransferase
		Glutamate dehydrogenase (NAD(P)(+))
		Glutamate receptor
		Glutamate--cysteine ligase
		Glutamine synthetase
		Glutathione peroxidase 1
		Glutathione S-transferase P
		Glutathione-disulfide reductase
		Glycerol kinase
		Glycerol-3-phosphate acyltransferase 3
		Glycylpeptide N-tetradecanoyltransferase
		Glypican-1
		GOLD domain-containing protein
		Grainyhead-like protein 1 homolog
		Grass carp reovirus (GCRV)-induced gene 2p
		GTP cyclohydrolase 1
		Guanine nucleotide exchange factor DBS

Diadromous	Valdivia resident	Toltén resident
		Guanine nucleotide-binding protein G(i) subunit alpha-1
		Guided entry of tail-anchored proteins factor 1
		Hcy-binding domain-containing protein
		Heat shock cognate 70-kd protein, tandem duplicate 2
		Heat shock protein 90 alpha
		Heat shock protein 90, alpha (cytosolic), class A member 1, tandem duplicate 2
		Heat shock protein family A (Hsp70) member 8
		HECT domain-containing protein
		Heme oxygenase (biliverdin-producing)
		Hemoglobin subunit beta-2
		HepA-related protein
		Hexosyltransferase
		Histone deacetylase 10
		HMG box domain-containing protein
		HMG box-containing protein 1
		Homeobox protein DLX-5
		HORMA domain-containing protein
		HpcH_HpaI domain-containing protein
		IF rod domain-containing protein
		Ig-like domain-containing protein

Diadromous	Valdivia resident	Toltén resident
		Importin subunit alpha-1
		Inactive C-alpha-formylglycine-generating enzyme 2
		Inactive carboxypeptidase-like protein X2
		Inositol 1,4,5-trisphosphate receptor, type 1a
		Insulin-like growth factor-binding protein 6b
		Intelectin
		Interferon-inducible GTPase 5-like
		Inversin
		Isocitrate dehydrogenase [NAD] subunit, mitochondrial
		J domain-containing protein
		Jacalin 4
		Kelch like family member 34
		Kinetochore-associated protein 1
		Krueppel-like factor 2
		L-2-hydroxyglutarate dehydrogenase
		Leucine rich repeat containing 1
		Leucine-rich repeat extensin-like protein 6
		L-methionine (R)-S-oxide reductase
		L-ornithine N(5)-monooxygenase
		Low density lipoprotein receptor class A domain containing 1

Diadromous	Valdivia resident	Toltén resident
		LRRCT domain-containing protein
		LSM6 homolog, U6 small nuclear RNA and mRNA degradation associated
		Lymphoid-specific helicase
		Lysine (K)-specific demethylase 6B, b
		Lysine-specific demethylase 7B
		Lysophosphatidic acid receptor 6b
		MADS-box domain-containing protein
		Malic enzyme
		Matrix metallopeptidase 25a
		Matrix metallopeptidase 25b
		Matrix metalloproteinase-9
		MCL1 apoptosis regulator, BCL2 family member b
		Mediator of RNA polymerase II transcription subunit 20
		Methyltranfer_dom domain-containing protein
		MFS_1_like domain-containing protein
		Microtubule-associated protein
		Microtubule-associated serine/threonine-protein kinase 2
		Microtubule-associated serine/threonine-protein kinase 3
		MIF4G domain-containing protein B
		Mitochondrial pyruvate carrier
		Mitogen-activated protein kinase

Diadromous	Valdivia resident	Toltén resident
		Mitogen-activated protein kinase kinase kinase
		MRP-S28 domain-containing protein
		Myosin motor domain-containing protein
		N-acetylneuraminic acid phosphatase
		N-acetyltransferase domain-containing protein
		NAD(P)(+)--arginine ADP-ribosyltransferase
		NADH dehydrogenase [ubiquinone] iron-sulfur protein 7, mitochondrial
		NFU1 iron-sulfur cluster scaffold homolog, mitochondrial
		NIPA magnesium transporter 2
		NTP_transferase domain-containing protein
		Nuclear factor of activated T cells 5a
		Nuclear factor of activated T-cells, cytoplasmic 1
		Nucleoporin NSP1
		Nucleoside-diphosphate kinase
		Occludin
		O-GlcNAc transferase subunit p110
		Origin recognition complex subunit 4
		Ornithine decarboxylase antizyme 1
		Oxidative stress induced growth inhibitor 1

Diadromous	Valdivia resident	Toltén resident
		Oxidoreductase-like domain-containing protein
		P13797
		P53 apoptosis effector related to pmp22
		Paired box protein Pax-9
		Patatin like phospholipase domain containing 7
		Peptidyl-prolyl cis-trans isomerase
		Peptidylprolyl isomerase
		Periostin
		Peroxiredoxin
		PHD finger protein 10
		PHD-type domain-containing protein
		Phosphate transporter
		Phosphatidylinositol-3-phosphatase SAC1
		Phosphoinositide 5-phosphatase
		Phospholipid-transporting ATPase
		Phosphoserine aminotransferase
		Pituitary tumor-transforming gene 1 protein-interacting protein
		Plasminogen
		Plastin 3
		Plastin 3 (T isoform)
		Platelet-activating factor acetylhydrolase
		Poly [ADP-ribose] polymerase
		Poly(U)-specific endoribonuclease

Diadromous	Valdivia resident	Toltén resident
		Polymerase (DNA directed), epsilon 3 (p17 subunit)
		Polypeptide N-acetylgalactosaminyltransferase
		Prenylcys_lyase domain-containing protein
		Pribosyltran_N domain-containing protein
		Profilin
		Prospero homeobox 1
		Prostacyclin synthase
		Prostaglandin E2 receptor EP3 subtype
		Protection of telomeres protein 1
		Protein bunched, class 2/F/G isoform
		Protein Churchill
		Protein kinase C alpha type
		Protein kinase domain-containing protein
		Protein kinase, membrane associated tyrosine/threonine 1
		Protein kish-B
		Protein MTSS 1
		Protein O-mannosyl-transferase 1
		Protein O-mannosyl-transferase TMTC2
		Protein phosphatase 1, regulatory subunit 32
		Protein phosphatase, Mg2+/Mn2+ dependent, 1Lb

Diadromous	Valdivia resident	Toltén resident
		Protein RER1
		Protein S100
		Protein timeless homolog
		Protein tyrosine phosphatase non-receptor type 3
		Protein XRP2
		Protein-tyrosine-phosphatase
		Protocadherin-15
		Putative ATP-dependent RNA helicase an3
		Putative ferric-chelate reductase 1
		Putative GPI-anchored protein pfl2
		Putative mediator of RNA polymerase II transcription subunit 21
		Pyridoxal-dependent decarboxylase domain containing 1
		Pyruvate carboxylase
		RAB interacting factor
		RAB20, member RAS oncogene family
		RAB8B, member RAS oncogene family
		Rab-GAP TBC domain-containing protein
		Rab-GAP TBC domain-containing protein
		Ral GEF with PH domain and SH3 binding motif 1
		RAS p21 protein activator 3

Diadromous	Valdivia resident	Toltén resident
		RasGAP-activating-like protein 1
		Ras-related GTP-binding protein
		Receptor protein serine/threonine kinase
		Relaxin family peptide receptor 3.3a3
		Replication protein A subunit
		Reticulon
		Rho guanine nucleotide exchange factor 26
		Rho-GAP domain-containing protein
		Rhomboid domain-containing protein
		Ribosomal protein S6 kinase
		RING-type domain-containing protein
		RNA binding motif, single stranded interacting protein 2b
		RNA helicase Mov10l1
		RNA-binding motif, single stranded-interacting protein 1b
		Roundabout, axon guidance receptor, homolog 1
		RRM domain-containing protein
		Rubicon-like autophagy enhancer
		RWD domain containing 4
		Ryanodine receptor 1
		S-adenosylmethionine synthase
		SAM domain-containing protein
		SCP2 domain-containing protein
		SEA domain-containing protein

Diadromous	Valdivia resident	Toltén resident
		Serine and arginine rich splicing factor 5b
		Serine and arginine rich splicing factor 6a
		Serine/threonine-protein kinase WNK4
		Serine/threonine-protein phosphatase
		SH3_10 domain-containing protein
		Short transmembrane mitochondrial protein 1
		SHSP domain-containing protein
		Shugoshin_C domain-containing protein
		Small monomeric GTPase
		Small nuclear ribonucleoprotein E
		SMC_N domain-containing protein
		Smoothelin, like
		Sodium- and chloride-dependent taurine transporter
		Sodium/hydrogen exchanger
		Solute carrier family 11 member 2
		Solute carrier family 2 member 11b
		Solute carrier family 2, facilitated glucose transporter member 1
		Solute carrier family 25 member 21
		Solute carrier family 35 member A3a
		Sorting nexin 8a
		Sorting nexin-10B
		Spartin a
		Spermatogenesis associated 20

Diadromous	Valdivia resident	Toltén resident
		Spindle and kinetochore-associated protein 3
		SpoU_methylase domain-containing protein
		Squalene synthase
		SRY-box transcription factor 9a
		Ssemaphorin 4F
		ST14 transmembrane serine protease matriptase b
		STAS domain-containing protein
		Stromal interaction molecule 1a
		Structure-specific endonuclease subunit SLX1
		Succinate-semialdehyde dehydrogenase
		Synaptotagmin-like 1
		Tartrate-resistant acid phosphatase type 5
		TEF transcription factor, PAR bZIP family member a
		Testis-specific H1 histone
		Tetraspanin
		THADA armadillo repeat containing
		Thioredoxin
		Thioredoxin reductase-like selenoprotein T1a
		Thioredoxin-disulfide reductase
		Thioredoxin-like protein

Diadromous	Valdivia resident	Toltén resident
		Threonine-rich protein
		TIR domain-containing protein
		TMEM189_B_dmain domain-containing protein
		TNFAIP3 interacting protein 1
		TPR_REGION domain-containing protein
		TPT domain-containing protein
		TPX2 microtubule nucleation factor
		Transcription factor BTF3
		Transcription factor SOX-30
		Transcription factor SOX-5
		Transcription factor Sp9
		Transcription initiation factor IIF subunit alpha
		Transcriptional repressor protein YY1
		Transforming growth factor beta
		Transforming growth factor beta regulator 4
		Transforming growth factor, beta receptor III
		Transforming growth factor-beta receptor-associated protein 1
		Transglutaminase 1-like 1
		Translationally-controlled tumor protein
		Transmembrane 7 superfamily member 3
		Transmembrane protein 107

Diadromous	Valdivia resident	Toltén resident
		Transmembrane protein 120B
		Transmembrane protein 163
		Transmembrane protein 164
		Transmembrane protein 245
		Transmembrane protein 41B
		Transporter (solute carrier family 6 member 6)
		Tripeptidyl-peptidase 1
		Troponin I type 2a (skeletal, fast), tandem duplicate 1
		TTF-type domain-containing protein
		Tubulin beta chain
		Tyrosine-protein kinase
		U4/U6.U5 small nuclear ribonucleoprotein 27 kDa protein
		Ubiquitin-like modifier-activating enzyme 1
		UDP-glucuronate decarboxylase 1
		UDP-glucuronosyltransferase
		Unconventional myosin-6
		Uridine-cytidine kinase
		Vacuole membrane protein 1
		Vav guanine nucleotide exchange factor 1
		Vitamin-K-epoxide reductase
		VRK serine/threonine kinase 1
		WD repeat domain 41
		WD repeat domain 66

Diadromous	Valdivia resident	Toltén resident
		WH2 domain-containing protein
		Xylulose kinase
		Zinc finger protein 572
		Zinc finger protein 703
		ZP domain-containing protein

Table S5.3. Characterized proteins (203) of the 398 DGE found to distinguish diadromous individuals in fresh water vs. saltwater.

	log2FoldChange		
Protein Name	**Diadromous**	**Toltén**	**Valdivia**
4-hydroxyphenylpyruvate dioxygenase	1.369063715		1.631452
60 kDa heat shock protein, mitochondrial	1.151448469	1.125243	
75 kDa glucose-regulated protein	1.355108282		
A disintegrin and metalloproteinase with thrombospondin motifs 1	2.113853177		
AAA domain-containing protein	2.010153752		
ACB domain-containing protein	1.32694132		1.010158
Acetyltransferase component of pyruvate dehydrogenase complex	1.0730131		
Aconitate hydratase, mitochondrial	1.118930006		
Actin cytoskeleton-regulatory complex protein pan1	2.210375054		
Acyltransferase	-3.959255248		
ADP/ATP translocase	2.368646482		
AIF-MLS domain-containing protein	1.444144867		
AIG1-type G domain-containing protein	-2.785209448		
Aldo_ket_red domain-containing protein	-1.712700141		-1.95198
Alpha 2-HS glycoprotein	-3.692905711		
Alpha-1-microglobulin	-4.655786249		
Alpha-2-macroglobulin	-3.919392518		
Ammonium_transp domain-containing protein	-1.244371676		
AMP-binding domain-containing protein	1.489187604	1.555877	1.126648
Androgen-induced gene 1 protein	1.368807558		
Apolipoprotein A-I-1	-3.435618351		
Apolipoprotein Ea	-4.055798551		
Aquaporin	-1.103282148		
ATP synthase F1 subunit delta	1.073259789		1.972771
ATP synthase lipid-binding protein	1.00134796		

Protein Name	Diadromous	Toltén	Valdivia
ATPase family AAA domain-containing protein 3	1.974942602	1.634101	1.289415
ATPase family AAA domain-containing protein 3-A	1.80506599	1.432294	1.512586
ATP-dependent 6-phosphofructokinase, liver type	1.094002188		1.512586
BTB domain-containing protein	2.072869319		
C1q domain-containing protein	-5.178954574		
Calcium and integrin binding family member 2	-1.404801017		
Calcium binding protein 2b	3.275831731		
Calcium uniporter protein	1.473482531		
Calcium-transporting ATPase	2.902970635		
cAMP-specific 3',5'-cyclic phosphodiesterase 4D	1.314335269		
Carbonic anhydrase	2.810871258		
CARD domain-containing protein	3.250691532		
Carnitine O-palmitoyltransferase	1.661378929		
Cathepsin H	-1.034429445	-1.21106	
CD8 beta	-1.327735817		
CD8a molecule	-1.659398688	-2.09266	
Cell adhesion molecule-related/down-regulated by oncogenes	1.241144222		
Centromere protein U	-1.118291701		
CHCH domain-containing protein	1.320617769		
Chemokine (C-C motif) ligand 19a, tandem duplicate 1	-2.175009476		
Chromodomain protein, Y-like	-1.212549311		
Clathrin heavy chain	1.020506924		
Claudin	1.783483435		
Complement C3	-3.37676589		
Complement C5	-4.460054689		
Complement component 1 Q subcomponent-binding protein, mitochondrial	1.346011308		
Complement component 1, r subcomponent	-3.904958214		

Protein Name	Diadromous	Toltén	Valdivia
Creatine kinase	-3.860067722		
Cryptochrome circadian regulator 3a	1.241614459		
Cryptochrome DASH	1.219681081		
Cryptochrome-1	1.38326791		
C-type lectin domain-containing protein	-5.820629728		
Cyclin N-terminal domain-containing protein	-1.13159809	-1.29383	-1.37062
Cystatin domain-containing protein	-4.654096174		
Cystic fibrosis transmembrane conductance regulator	1.912136625		1.639981
Cytochrome c	1.931915051	2.765219	1.943877
Cytochrome P450, family 1, subfamily C, polypeptide 1	1.552800078		
Cytohesin 4a	-1.040509453		
Ddb1 and cul4 associated factor 6	2.763973989		
Dedicator of cytokinesis 2	-2.009445299		
Delta-like protein	1.793162485	1.184782	1.222138
Delta-like protein B	1.75655099		
DH domain-containing protein	-1.598520405	1.069406	
Dipeptidase	2.45584539		
DNA replication complex GINS protein PSF3	-1.501354336	-1.19915	
DNAX-activation protein 10	-1.98728416	-1.78178	
DOP1 leucine zipper like protein B	-2.322088366		
Dynein intermediate chain 3, axonemal	1.122607919		
EH domain-containing protein 4	1.26432289		1.407035
Electron transfer flavoprotein subunit beta	1.075263798		
Electron transfer flavoprotein-ubiquinone oxidoreductase	1.151897473		
Elongation factor G, mitochondrial	1.198534371		
Enoyl-CoA hydratase	1.530872854	1.056157	
Fatty acid-binding protein 10-A, liver basic	-4.363119652		

Protein Name	Diadromous	Toltén	Valdivia
FCH and double SH3 domains 2	1.415923225		
FHA domain-containing protein	3.017062221		
Fibrinogen beta chain	-3.685536284		
Fibrinogen C-terminal domain-containing protein	-3.903114687		
Fibrinopeptide A	-4.069156188		
Formin binding protein 1b	-1.057852721		
Fructose-bisphosphate aldolase C-B	1.294725542		
FYVE, RhoGEF and PH domain containing 5b	1.574139314		
G protein-coupled receptor 55	-1.079516894		
G_PROTEIN_RECEP_F1_2 domain-containing protein	-1.421286025	-1.82032	
Galactose-3-O-sulfotransferase 1b	1.361969842		
Gamma-butyrobetaine hydroxylase	1.410310487		
Gamma-glutamyltransferase 5a	-1.032138326		-1.45368
Glutaminase	1.102620418	1.147593	
GRB2 related adaptor protein 2a	-1.137962532		
Guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1	1.275108323		
HATPase_c domain-containing protein	1.266586392		
Heat shock protein 75 kDa, mitochondrial	1.173197098		
Heat shock protein family A (Hsp70) member 8	1.833150091		
Heat shock protein, alpha-crystallin-related, b6	1.504451813	1.572494	
Heme-binding protein soul5	-2.077594798		
Hemopexin	-3.782529192		
Ig-like domain-containing protein	-3.627566186		
Immunoglobulin heavy variable 5-5	-1.996472843		
Interferon regulatory factor 10	-1.22187347	-1.86327	-1.96376
IRF tryptophan pentad repeat domain-containing protein	-1.301985232	-1.48076	
Isoleucine--tRNA ligase, cytoplasmic	1.336674023		

Protein Name	Diadromous	Toltén	Valdivia
J domain-containing protein	1.445534514		
Jumping translocation breakpoint	1.097060008		
Kinesin-like protein	-2.32472232		
Krueppel-like factor 7	1.272872901		
LIM domain-containing protein	1.456255551		
Lon protease homolog, mitochondrial	1.456156769		
Lysosomal protein transmembrane 4 alpha	2.307574748		1.441242
Meteorin-like protein	2.68344425		
MFS domain-containing protein	1.439203244		
mir-22	2.622383287	2.532456	2.301592
Mitochondrial import inner membrane translocase subunit Tim13	1.186075727	1.11039	1.655304
Mitochondrial import inner membrane translocase subunit Tim50	1.31306089		
Monocarboxylate transporter 2	1.768379943		
MTSS I-BAR domain containing 2a	1.093404673		
Mucin-5AC	-1.283131054		
Myosin light chain kinase family, member 4a	1.486548888		
NACHT domain-containing protein	-1.058377166	-1.23711	
Neural cell adhesion molecule 3	3.227438838		
Non-histone chromosomal protein HMG-17	-1.373818169	-1.23268	
Non-specific serine/threonine protein kinase	2.266921774	-1.89859	1.773672
Oxoglutarate (alpha-ketoglutarate) receptor 1a, tandem duplicate 1	-1.571442866		
PAS domain-containing protein	1.191817823		
PCNA-associated factor	-1.470330167		
PDZ domain-containing protein	-1.440317625	-1.25546	
PE repeat family protein	-4.511259638		
Pentatricopeptide repeat-containing protein 2, mitochondrial	1.734102611		
Peptidase M12B domain-containing protein	2.381310215		1.988146

Protein Name	Diadromous	Toltén	Valdivia
Peptidase S1 domain-containing protein	-1.848002966	-2.22309	
Peptidase_M24 domain-containing protein	1.323899393		1.517552
Peroxisome proliferator-activated receptor gamma coactivator 1-alpha	1.618393592		1.576724
PH domain-containing protein	-1.793557649	-1.63261	
Phosphodiesterase	1.524514392	1.841581	1.674236
Pitrilysin metalloproteinase 1	1.412373097		
Platelet-activating factor acetylhydrolase	1.198716127	1.029912	
Plexin domain containing 1	-1.25482494	-1.782	-1.87453
Poly(U)-specific endoribonuclease	-1.567332302	1.101401	-1.17497
Potassium inwardly rectifying channel subfamily J member 15	2.11141022	1.604996	
Procollagen-proline 4-dioxygenase	2.642403727		
Progestin and adipoQ receptor family member Iva	-1.060453643		
Proline rich coiled-coil 2C	2.220401079		
Protein kinase domain-containing protein	1.587825495	1.454691	1.56393
Protein phosphatase	1.433195755	1.189412	
Protein Wnt	1.422561612		
Protein-tyrosine-phosphatase	1.700361465		2.331015
Prothrombin	-3.216056129		
Proton-translocating NAD(P)(+) transhydrogenase	1.184790232		
Pyruvate dehydrogenase E1 component subunit beta	1.298799019		
Pyruvate kinase	1.039375188		
RAP domain-containing protein	1.194638819		
Receptor protein-tyrosine kinase	1.886886938		
Retinoblastoma binding protein 5	3.27132154		
Ribonuclease_T2	-5.062835619		
Ribosomal protein L39 like	4.415338257		
RING-type domain-containing protein	-1.080466903	-1.10979	

Protein Name	Diadromous	Toltén	Valdivia
RING-type E3 ubiquitin transferase	-4.589574089		
RNA helicase	1.463933644		
RRM domain-containing protein	1.673124713		
Sema domain-containing protein	1.278119104		
SERPIN domain-containing protein	-4.234132127		
Serpin peptidase inhibitor, clade G (C1 inhibitor), member 1	-3.448623417		
Sex hormone-binding globulin	-3.890716052		
SH2 domain-containing protein	-1.238882441	-1.90998	
SH3 domain-binding protein 1	-1.546167988		
Sodium/potassium-transporting ATPase subunit alpha	1.754687239	2.679353	2.98517
Sodium/potassium-transporting ATPase subunit beta	3.898742911		
Sodium/potassium-transporting ATPase subunit beta-1	1.437158516		1.430159
Solute carrier family 12 member 2	2.208432758	2.576543	1.218912
Solute carrier family 5 member 3b	1.333695107	3.236883	3.042832
Spectrin, beta, non-erythrocytic 5	1.408485752		1.32393
Src kinase-associated phosphoprotein 1	5.484299558		
Stress-70 protein, mitochondrial	1.266901094		
Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial	1.932160306	1.707411	
T cell receptor alpha variable 6	-1.443345248		
TAP-binding protein-like	-1.062767715		
T-cell receptor T3 zeta chain	-1.074669946		
Tetraspanin	2.012994036		
Thioesterase superfamily member 6	1.251965084		
Thr_synth_N domain-containing protein	1.055967553		
Thymidine kinase	-1.333432009	-1.13845	
TRAMP-like complex RNA-binding factor ZCCHC8	-1.18413535		
Trans-2,3-enoyl-CoA reductase-like 2b	-2.093590152		

Protein Name	Diadromous	Toltén	Valdivia
Transcription factor SOX	1.05798316		1.084511
Transferrin receptor protein 1	1.70489212	1.217614	
Transmembrane protein 161B	1.039776745		
Tripartite motif containing 108	-1.056716604	-2.16428	
Tripartite motif containing 16	1.668444288		
Tyrosinase-related protein 1a	2.550013937		
Tyrosine-protein kinase	-1.045694497		
Tyrosine-protein kinase receptor TYRO3	1.803973412		
Ubiquitin carboxyl-terminal hydrolase	1.28458883		
Ubiquitin-conjugating enzyme E2T (putative)	-1.355792295		
Vascular endothelial growth factor A-A	1.038322184	1.397017	
Vitronectin b	-5.248523711		
Voltage-dependent anion-selective channel protein 2	1.154882535		
von Willebrand factor A domain containing 2	2.48564601		
Zinc finger protein 536	-1.1395482		
ZnMc domain-containing protein	1.383135429		

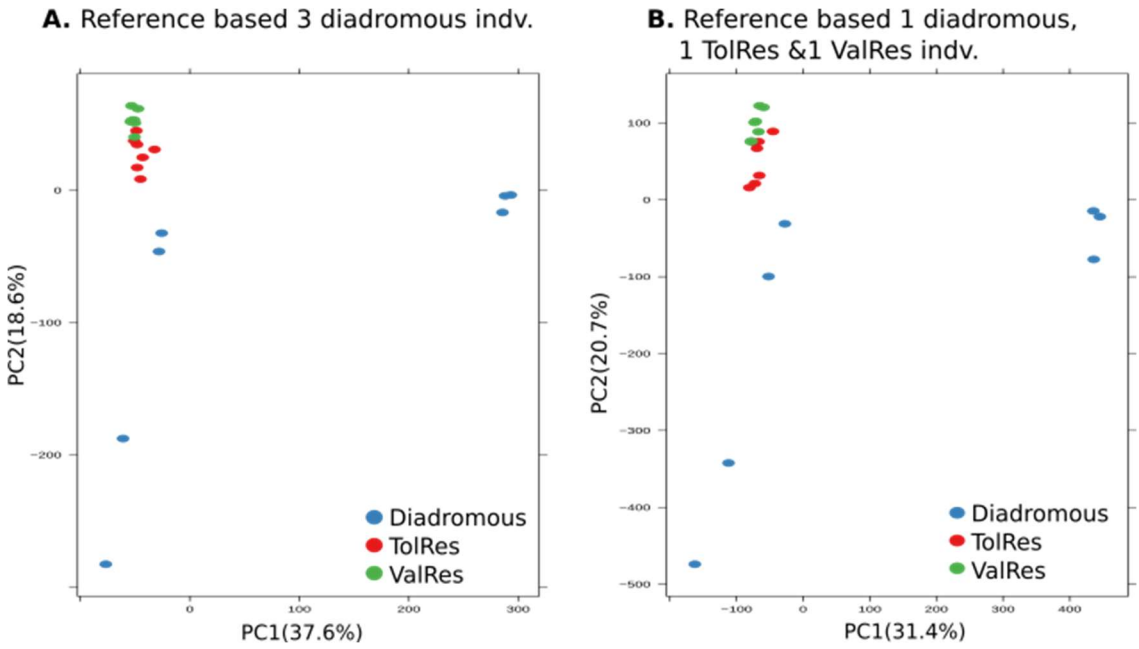


Figure S5.1. Principal Component Analysis (PCA) of samples aligned to a transcriptome build with **A.** 3 diadromous individuals in fresh water and **B.** 1 diadromous, 1 Toltén resident, and 1 Valdivia resident in fresh water.

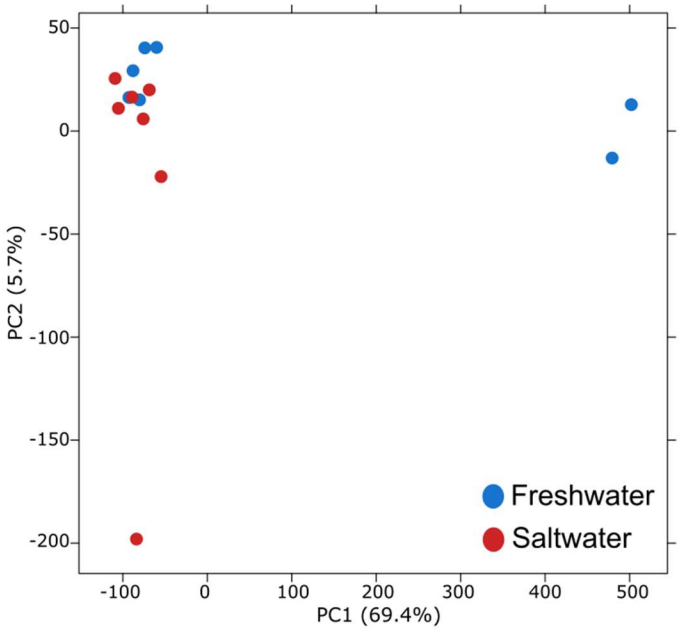


Figure S5.2. Principal Component Analysis (PCA) of gene expression including all seven diadromous individuals in freshwater at 0 hours.

CHAPTER 6. CONCLUSION

6.1 Thesis summary

This thesis focuses on the evolution of diadromy (the movement of aquatic organisms between marine and freshwater environments) and its loss. The various types of diadromy represent different cases of evolution which have had important implications in the current diversity and distribution of fishes (McDowall 2001). Investigating the molecular mechanisms that differentiate diadromous from their derived freshwater resident populations can not only bring insights into the evolutionary transition from the marine into freshwater environments but can also contribute to our understanding of how species adapt to changing conditions. The investigation of the potential evolutionary responses of a life history trait must integrate information about traits from a genetic, phenotypic, and physiological perspective (Stearns 1989). Thus, by combining phenotypic, physiological, genomic, and transcriptomic data of diadromous and resident populations of *Galaxias maculatus,* I bring insights into the degree of genomic divergence between diadromous, specifically amphidromous, and resident populations and on the different evolutionary outcomes or responses associated with the loss of this life history trait.

To start, I reviewed the literature regarding diadromy and its origin. This search led me to hypothesize that diadromy might have appeared in multiple independent events across the fish phylogeny and that migratory behavior is not genetically predisposed meaning that not every taxon has the genetic architecture to move between fresh and saltwater environments (Chapter 2; Delgado & Ruzzante 2020). The idea that diadromy has appeared multiple times emphasizes the importance of studying different diadromous species from the various categories of diadromy and different taxa across the fish phylogeny, as it is likely that similar adaptations we see in multiple diadromous species appeared by convergence. However, studies on diadromy that use -omics data to understand the evolution of the two life histories have thus far focused on a few anadromous species and research on amphidromous species, the most diverse category of diadromy is lacking. Thus, the present study on *Galaxias maculatus* intends to bring attention to the fact that a non-model organism can provide insights into the mechanisms

underlying the evolution of diadromous species and their derived non-migratory populations.

The evolution of diadromous fishes can lead to contradictory outcomes, as diadromy can promote gene flow hindering genetic structure and genetic divergence among populations, but its loss promotes high genetic differences leading to increased genetic divergence among populations and in some cases even speciation (McDowall 2001). Here, I showed how *G. maculatus* Chilean populations exemplified these outcomes. By combining previous research on bone tissue microchemistry (Górski *et al.* 2018) with the genomic data, indeed I found that diadromous and resident populations are genetically distinguishable. While diadromous individuals appear to comprise one large population (with some weak structure) due to the high gene flow, resident populations derived from this single diadromous population. The northernmost resident populations are highly genetically distinguishable from the diadromous populations and from each other with F_{ST} estimates around 0.4-0.5 and can thus provide insights into local adaptation and evolution towards residency (Chapter 3; Delgado *et al.* 2019).

To investigate how resident populations (associated with the loss of diadromy) are evolving, I selected two river systems that can be considered natural replicates: Toltén and Valdivia, as they are geographically close and were colonized from the same diadromous source population. Diadromous and resident individuals from these river systems were tested to compare their osmoregulatory capacity to a gradual change in salinity. While Toltén residents did not tolerate the salinity change (i.e., up to 25 ppt), Valdivia residents tolerated the change showing 100% survival. This result supports the suggestion that alternative genetic routes leading to similar phenotypes are not a rare occurrence (Elmer & Meyer 2011). To further identify genes underlaying adaptive responses, a key step in the study of evolutionary biology (Orsini *et al.* 2013), I used different outlier detection methods to identify genes under selection. Outlier SNPs that differentiated diadromy vs. residency and the tolerance vs. intolerance to salt water were identified. A few of them have previously been reported in other diadromous fishes, but the lack of an annotated genome limited the scope of the analysis. Interestingly, some of these genes such as solute carriers and sodium/potassium exchangers have previously been reported in other migratory fish species from different orders (Chapter 4; Delgado *et*

al. 2020). As expected, many of these genes have roles associated with migration including salinity adaptations and running times. However, some outliers were found in genes not previously reported in diadromous species.

The fact that some genes have not previously been reported in diadromous species could be explained by the limited number of studies and species studied. However, the evolution of closely related lineages can be a complex mosaic of parallel and nonparallel changes (Orsini *et al.* 2013). The high and rapid divergence found between *G. maculatus* diadromous and resident populations indicates that the molecular mechanisms underlying the loss of diadromy may not only occurred in mutations at the coding level but in changes in the regulation of the expression of genes. Indeed in Three Spined Stickleback, recurrent colonization of freshwater environments and the subsequent loss of diadromy have shown the importance of regulatory regions (Jones *et al.* 2012). Thus, I further looked into how these two resident populations (i.e., the salinity intolerant Toltén, and the salinity tolerant Valdivia) responded to an abrupt change in salinity and coupled those results with transcriptomic responses. Given that Toltén resident individuals showed high mortality in salt water and previous research has shown that osmoregulatory mechanisms activate within 24 hours (Handeland *et al.* 1998), osmoregulatory responses were assessed by measuring the percentage of water content in muscle for a period of 48 hours.

Results confirmed that the evolution of these two resident populations followed different genetic routes or pathways. Compared with the gene expression of the diadromous individuals, both Toltén and Valdivia residents showed differences in the genes that were differentially expressed when salinity was abruptly changed from 0 ppt (fresh water) to 23-25 ppt. Two key osmoregulatory genes, Sodium potassium ATPase subunit alpha (NKA) and solute carrier family 12 (NKCC1) were upregulated as expected in both resident populations. In contrast, two important genes involved in ion transportation, NKA subunit beta and Cystic fibrosis transmembrane conductance regulator (CFTR), were not differentially expressed among the saltwater intolerant Toltén individuals when they were exposed to salt water. Contrary to my initial hypothesis, gene expression among Valdivia residents was no more similar to gene expression among the diadromous individuals than was gene expression among the intolerant Toltén residents.

The saltwater tolerant Valdivia residents, however, retained the expression of key osmoregulatory genes (chapter 5).

Overall, my findings provide further evidence that changes at the coding sequence level are decoupled from changes at the gene expression level (Rivas *et al.* 2018) and highlighting the importance of combining genomic with transcriptomics data to bring insights into the evolutionary processes resulting in residency. Finally, given the differences found in coding and regulatory regions between these replicate populations, I conclude that stochastic processes like genetic drift underlay the evolutionary process from a diadromous to resident life history form and that the maintenance of this plastic phenotype in great measure is due to chance.

6.2 Main contributions and implications

This thesis includes an updated list of known diadromous species and their phylogenetic relationship. Knowing the phylogenetic relationship between diadromous species can help us in the debate and discussion of diadromy and its origin. I also proposed that the most parsimonious explanation for the origin of diadromy is that it has appeared in multiple independent events (Chapter 2; Delgado & Ruzzante 2020). This hypothesis is supported by the little parallelism found across species in the genes associated with diadromy, and it also highlights the importance of studying diadromy in different taxa.

This research is the first study on an amphidromous species that combines rearing experiments with -omics data. My work shows that *G. maculatus* has the potential to become a model organism, not only because of its wide distribution but also because of the existence of natural resident replicates that derived from the same diadromous population. These natural replicates that can help elucidate the evolutionary processes involved in the loss of diadromy.

My results show that the loss of diadromy affect replicate resident populations differently. It can lead to differences at the sequence level (*de novo* mutations) most likely as the result of local adaptation and/or genetic drift, and to differences in the retention or loss of the gene expression of different genes. These results indicate that drift likely plays an important role in the maintenance of plasticity bringing its role in the generation of divergence between diadromous and resident populations into sharp focus.

Overall, the results from this thesis highlight the complex processes of evolution. Having a better understanding of these processes can help us predict the potential for adaptation to changing conditions, which is of crucial importance because of climate change. Given their need to depend on distinct environments, diadromous species are vulnerable to changes in both marine and freshwater environments (McDowall *et al.* 2009) and are at risk from anticipated climate change that affects primary productivity in aquatic systems (Chalant *et al.* 2019).

6.3 Limitations and future directions

A major limitation of the study was the use of adults. Although diadromous *G. maculatus* adults are euryhaline, i.e., they are able to survive in high salinity conditions (Chessman & Williams 1975), the life cycle stage during which individuals migrate is the larval stage. Thus, I would recommend the use of larvae in future work. The short life span and small size of *G. maculatus* make the species a good candidate for breeding experiments. With the right conditions, chiefly the maintenance of a relatively low temperature in the lab (between 10 -15 C), it is feasible to reproduce *G. maculatus* (Mitchell 1989). The characteristics mentioned above suggest *G. maculatus* is also a good candidate for conducting crosses within populations (useful to account for maternal effects) as well as between diadromous and resident populations to examine plasticity in hybrids.

The lack an outgroup for the *G. maculatus* Chilean populations that could precisely estimate the branch length of the phylogenetic tree (chapter 3) constrained the estimation of the time since colonization. This estimation will be important to confirm that Toltén and Valdivia's residents were indeed colonized at similar times and that their levels of genetic diversity were similar as well. Furthermore, at the transcriptomic level, Toltén residents were no more different from the diadromous individuals than were Valdivia residents. In fact, Toltén residents exhibited more similarities among downregulated genes with the diadromous individuals than did Valdivia residents. This would suggest that the speculation that the Toltén resident population has a smaller effective population size than the Valdivia resident population is likely wrong or that the colonization times of the resident populations differ from each other more than we hypothesized in this work. Increasing the sample sizes of the genotyped individuals of

both resident populations would allow the estimation of effective population sizes. Such work, would help enrich the discussion regarding the evolutionary forces behind the evolution of these resident populations.

Including more replicate resident populations, particularly those in southern Chile which have lost their migratory trait more recently than the populations I used in my thesis can help provide further evidence regarding the role of drift in the evolution of residency. The fact that *G. maculatus* is also widely distributed means that research can also be performed in other populations in New Zealand or Tasmania, which also comprise diadromous and resident populations. Analyzing genomic data across the entire species distribution can provide insights regarding the species connectivity and if in fact, *G. maculatus* has a continuous migration via the west wind drift (Waters *et al.* 2000a).

Unlike most anadromous fish species, amphidromous organisms spend a relatively short time in the marine environment, and it has been hypothesized that this ability facilitates the colonization of new and changing environments (Hogan *et al.* 2014). Although it is not clear how the category of diadromy may influence the evolution of residency, my results give support to the hypothesis that amphidromous and their lack of site fidelity (contrary to the behavior observed in anadromous fishes) and the subsequent loss of diadromy indeed facilitate the colonization of new environments. Additionally, the fact that there are more amphidromous than anadromous and catadromous species worldwide, suggests further studies in amphidromous species are not just warranted but badly needed. Overall, the importance of studying diadromy in different lineages can help elucidate the remaining questions, particularly concerning their origin and adaptive potential.

Here I suggest that the differences observed in the evolution of the derived resident populations can be ascribed to stochastic processes, but the role of local adaptation and the importance of environmental conditions should also be further investigated. Future research measuring the environmental conditions is warranted. Associating different environmental variables with SNPs and gene expression will be necessary to understand the role of adaptive forces (e.g., redundancy analysis).

Although this work includes the first reference transcriptome of *G. maculatus* in gills proving information for further genetic studies. The development of a reference

genome and whole-genome sequencing can further facilitate the study of diadromy in this species and help determine if there are genomic regions of differentiation associated with migratory behavior as found in other diadromous species like for instance, steelhead/Rainbow Trout (Pearse *et al.* 2014). Knowledge of the genome sequence can also provide clues about the lack of hybridization found between diadromous and resident individuals in sympatry.

The transcriptomic studies can also extend to other osmoregulatory organs including the intestine and kidneys to have a clearer picture of other genes important for salinity acclimation. Yet, to increase our understanding of migratory behavior, transcriptomic studies with brain tissue where signaling must dictate when and if to undergo migration, would also be required. Studies at the cell level (type of cells in the gills or brain) would also be important, this would require the use of -omics advance such single-cell transcriptomics. Having a reference genome can also help elucidate the mechanisms behind the differences in gene expression, specifically if mutations of *cis*- and *trans*-QTLs are responsible for the differences observed.

With the development of -omics methods, the interest in understanding parallel and non-parallel evolution of repetitive adaptations has increased (Waters & Mcculloch 2021). Thus far it has been hypothesized that there must be a limited number of pathways that evolution can follow (Fischer *et al.* 2021). The number of independent resident *G. maculatus* populations that derived from a same diadromous source across the Chilean distribution of the species makes this system ideal to continue to study and bring further insights into the underlying mechanisms that led to similar or different evolutionary trajectories.

BIBLIOGRAPHY

Abdou A, Keith P, Galzin R (2015) Freshwater Neritids (Mollusca : Gastropoda) of Tropical Islands, Amphidromy As a Life Cycle, a Review. *Revue d'Ecologie (Terre et Vie)*, **70**, 387–397.

Acolas ML, Lambert P (2016) Life histories of anadromous fihes. In: *An introduction to fish migration* (eds Morais P, Daverat F), pp. 66–67.

Acolas ML, Rochard E, Le Pichon C, Rouleau E (2012) Downstream migration patterns of one-year-old hatchery-reared European sturgeon (Acipenser sturio). *Journal of Experimental Marine Biology and Ecology*, **430**–**431**, 68–77.

Ahmed Q, Bat L (2016) Levels and Health Risk Assessments of Cd and Pb in Pomadasys maculatus Marketed by Karachi Fish Harbor, Pakistan. *ILMU KELAUTAN: Indonesian Journal of Marine Sciences*, **21**, 53.

Al-Tobasei R, Paneru B, Salem M (2016) Genome-wide discovery of long non-coding RNAs in rainbow trout. *PLoS ONE*, **11**, e0148940.

Albieri RJ, Araújo FG, Ribeiro TP (2010) Gonadal development and spawning season of white mullet Mugil curema (Mugilidae) in a tropical bay: Short communication. *Journal of Applied Ichthyology*, **26**, 105–109.

Alexander DH, Novembre J, Lange K (2009) Fast Model-Based Estimation of Ancestry in Unrelated Individuals. *Genome research*, **19**, 1655–64.

Allen PJ, Cech JJ, Kültz D (2009) Mechanisms of seawater acclimation in a primitive, anadromous fish, the green sturgeon. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, **179**, 903–920.

Allen PJ, Mitchell ZA, Devries RJ *et al.* (2014) Salinity effects on Atlantic sturgeon (Acipenser oxyrinchus oxyrinchus Mitchill, 1815) growth and osmoregulation. *Journal of Applied Ichthyology*, **30**, 1229–1236.

Almeida PR (1996) Estuarine movement patterns of adult thin-lipped grey mullet, Liza ramada (Risso) (Pisces, Mugilidae), observed by ultrasonic tracking. *Journal of Experimental Marine Biology and Ecology*, **202**, 137–150.

Alò D, Correa C, Samaniego H, Krabbenhoft CA, Turner TF (2019) Otolith microchemistry and diadromy in Patagonian river fishes. *PeerJ*, **2019**, 1–25.

Alò D, Lacy SN, Castillo A, Samaniego HA, Marquet PA (2020) The macroecology of fish migration. *Global Ecology and Biogeography*, 1–18.

Alvarez M, Schrey A, Richards C (2015) Ten years of transcriptomics in wild populations: what have we learned about their ecology and evolution? *Molecular ecology*, **24**, 710–725.

Andrews S (2010) FastQC: A Quality Control Tool for High Throughput Sequence Data.

Aniagu SO, Williams TD, Allen Y, Katsiadaki I, Chipman JK (2008) Global genomic methylation levels in the liver and gonads of the three-spine stickleback (Gasterosteus aculeatus) after exposure to hexabromocyclododecane and 17-β

oestradiol. *Environment International*, **34**, 310–317.

Aprahamian MW, Baglinière JL, Sabatié MR *et al.* (2003) Biology, status, and conservation of the anadromous Atlantic twaite shad Alosa fallax fallax. *American Fisheries Society Symposium*, **2003**, 103–124.

Arai T (2006) Comparison of habitat use during early life stage between ayu Plecoglossus altivelis and ice goby Leucopsarion petersi along the Sanriku Coast of Japan, as determined from otolith microchemistry. *Fisheries Science*, **72**, 382–387.

Arai T, Chino N (2012) Diverse migration strategy between freshwater and seawater habitats in the freshwater eel genus Anguilla. *Journal of Fish Biology*, **81**, 442–455.

Arai T, Chino N (2019) Variations in the migratory history of the tropical catadromous eels Anguilla bicolor bicolor and A. bicolor pacifica in south-east Asian waters. *Journal of Fish Biology*, **94**, 752–758.

Arai T, Chino N, Le DQ (2013) Migration and habitat use of the tropical eels Anguilla marmorata and A. bicolor pacifica in Vietnam. *Aquatic Ecology*, **47**, 57–65.

Arai T, Goto A (2008) Diverse migratory histories in a brackish water type of the ninespine stickleback, Pungitius pungitius. *Environmental Biology of Fishes*, **83**, 349–353.

Arai T, Goto A, Miyazaki N (2003) Use of otolith microchemistry to estimate the migratory history of the threespine stickleback, Gasterosteus aculeatus. *Journal of the Marine Biological Association of the United Kingdom*, **83**, 223–230.

Arai T, Kotake A, McCarthy TK (2006a) Habitat use by the European eel Anguilla anguilla in Irish waters. *Estuarine, Coastal and Shelf Science*, **67**, 569–578.

Arai T, Miyazaki N (2001) Use of otolith microchemistry to estimate the migratory history of the Russian sturgeon, Acipenser guldenstadti. *Journal of the Marine Biological Association of the United Kingdom*, **81**, 709–710.

Arai T, Taha H, Amalina R, Iizuka Y, Chang CW (2019) Anadromy and heterogenous population of a tropical shad Tenualosa ilisha in Malaysia, as revealed by otolith microchemistry and molecular evidence. *Journal of Fish Biology*, **95**, 1506–1511.

Arai T, Yang J, Miyazaki N (2006b) Migration flexibility between freshwater and marine habitats of the pond smelt Hypomesus nipponensis. *Journal of Fish Biology*, **68**, 1388–1398.

Arostegui MC, McKinney GJ, Quinn TP *et al.* (2019) Retention of a chromosomal inversion from an anadromous ancestor provides the genetic basis for alternative freshwater ecotypes in rainbow trout. *Molecular Ecology*, **28**, 1412–1427.

Artemov A V, Mugue NS, Rastorguev SM *et al.* (2017) Genome-Wide DNA Methylation Profiling Reveals Epigenetic Adaptation of Stickleback to Marine and Freshwater Conditions. *Molecular Biology and Evolution*, **34**, 2203–2213.

Asaduzzaman M, Wahab MA, Rahman MJ *et al.* (2019) Fine-scale population structure and ecotypes of anadromous Hilsa shad (Tenualosa ilisha) across complex aquatic ecosystems revealed by NextRAD genotyping. *Scientific Reports*, **9**, 1–15.

Augspurger JM, Warburton M, Closs GP (2017) Life-history plasticity in amphidromous and catadromous fishes: a continuum of strategies. *Reviews in Fish Biology and Fisheries*, **27**, 177–192.

Austin CS, Bond MH, Smith JM, Lowery ED, Quinn TP (2019) Otolith microchemistry reveals partial migration and life history variation in a facultatively anadromous, iteroparous salmonid, bull trout (Salvelinus confluentus). *Environmental Biology of Fishes*, **102**, 95–104.

Avigliano E, Carvalho B, Velasco G *et al.* (2017) Nursery areas and connectivity of the adults anadromous catfish (Genidens barbus) revealed by otolith-core microchemistry in the south-western Atlantic Ocean. *Marine and Freshwater Research*, **68**, 931–940.

Aykanat T, Johnston SE, Orell P *et al.* (2015) Low but significant genetic differentiation underlies biologically meaningful phenotypic divergence in a large Atlantic salmon population. *Molecular Ecology*, **24**, 5158–5174.

Azevedo MFC, Oliveira C, Pardo BG, Martínez P, Foresti F (2008) Phylogenetic analysis of the order Pleuronectiformes (Teleostei) based on sequences of 12S and 16S mitochondrial genes. *Genetics and Molecular Biology*, **31**, 284–292.

Baerwald MR, Meek MH, Stephens MR *et al.* (2016) Migration-related phenotypic divergence is associated with epigenetic modifications in rainbow trout. *Molecular Ecology*, **25**, 1785–1800.

Baetscher DS, Hasselman DJ, Reid K, Palkovacs EP, Garza JC (2017) Discovery and characterization of single nucleotide polymorphisms in two anadromous alosine fishes of conservation concern. *Ecology and Evolution*, **7**, 6638–6648.

Baglinière JL, Sabatié MR, Rochard E, Alexandrino P, Aprahamian MW (2003) The allis shad Alosa alosa: Biology, ecology, range, and status of populations. *American Fisheries Society Symposium*, **2003**, 85–102.

Balakirev ES, Parensky VA, Ayala FJ (2016) Complete mitochondrial genomes of the anadromous and resident forms of the lamprey Lethenteron camtschaticum. *Mitochondrial DNA. Part A, DNA mapping, sequencing, and analysis*, **27**, 1730–1731.

Barbee NC, Hale R, Morrongiello J *et al.* (2011) Large-scale variation in life history traits of the widespread diadromous fish, Galaxias maculatus, reflects geographic differences in local environmental conditions. *Marine and Freshwater Research*, **62**, 790–800.

Barile J, Escudero M, Carreño E (2013) Efecto combinado de la temperatura y salinidad sobre la supervivencia embrionaria de Galaxias maculatus. *Revista de biología marina y oceanografía*, **48**, 641–645.

Barker JR, Lambert DM (1988) A genetic analysis of populations of Galaxias maculatus from the Bay of Plenty : implications for natal river return. *New Zealand Journal of Marine and Freshwater Research*, **22**, 321–326.

Barriga J, Battini M, García-Asorey M *et al.* (2012) Intraspecific variation in diet,

growth, and morphology of landlocked Galaxias maculatus during its larval period: the role of food availability and predation risk. *Hydrobiologia*, **679**, 27–41.

Bartels H, Wrede C, Przybylski M, Potter IC, Docker MF (2017) Implications of absence of seawater-type mitochondria-rich cells and results of molecular analyses for derivation of the non-parasitic Ukrainian brook lamprey Eudontomyzon mariae. *Environmental Biology of Fishes*, **100**, 509–518.

Beamish RJ, Levings CD (1991) Abundance and freshwater migrations of the anadromous parasitic lamprey, Lampetra tridentata, in a tributary of the Fraser River, British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences*, **48**, 1250–1263.

Bell MA, Aguirre WE, Buck NJ (2004) Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution*, **58**, 814–824.

Bell MA, Foster SA (1994) The Evolutionary Biology of the Threespine Stickleback. In: *The Journal of Animal Ecology*, pp. 1–27.

Bell KNI, Pepin P, Brown JA (1995) Seasonal, inverse cycling of length- and age-at-recruitment in the diadromous gobies Sicydium punctatum and Sicydium antillarum in Dominica, West Indies. *Canadian Journal of Fisheries and Aquatic Sciences*, **52**, 1535–1545.

Benson LK, Michael Fitzsimons J (2002) Life history of the Hawaiian fish Kuhlia sandvicensis as inferred from daily growth rings of otoliths. *Environmental Biology of Fishes*, **65**, 131–137.

Berra TM, Crowley L, Ivantsoff W, Fuerst PA (1996) Galaxias maculatus: an explanation of its biogeography. *Marine and Freshwater Research*, **47**, 845–849.

Betancur-R R (2010) Molecular phylogenetics supports multiple evolutionary transitions from marine to freshwater habitats in ariid catfishes. *Molecular Phylogenetics and Evolution*, **55**, 249–258.

Bicskei B, Bron JE, Glover KA, Taggart JB (2014) A comparison of gene transcription profiles of domesticated and wild Atlantic salmon (Salmo salar L.) at early life stages, reared under controlled conditions. *BMC genomics*, **15**, 884.

Bijoy Nandan S, Jayachandran PR, Sreedevi OK (2012) Temporal pattern of fish production in a microtidal tropical estuary in the south-west coast of India. *Indian Journal of Fisheries*, **59**, 17–26.

Bissegger M, Laurentino TG, Roesti M, Berner D (2019) Widespread intersex differentiation across the stickleback genome – The signature of sexually antagonistic selection? *Molecular Ecology*, **29**, 262–271.

Blaber SJM, Milton DA, Brewer DT, Salini JP (2003) Biology, fisheries, and status of tropical shads Tenualosa spp. in south and southeast Asia. *American Fisheries Society Symposium*, **2003**, 49–58.

Bloom DD, Lovejoy NR (2014) The evolutionary origins of diadromy inferred from a time-calibrated phylogeny for Clupeiformes (herring and allies). *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20132081.

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.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.

Borcherding J, Breukelaar AW, Winter H V., König U (2014) Spawning migration and larval drift of anadromous North Sea houting (Coregonus oxyrinchus) in the River IJssel, the Netherlands. *Ecology of Freshwater Fish*, **23**, 161–170.

Boulet M, Normandeau É, Bougas B, Audet C, Bernatchez L (2012) Comparative transcriptomics of anadromous and resident brook charr Salvelinus fontinalis before their first salt water transition. *Current Zoology*, **58**, 158–170.

Bourret V, Kent MP, Primmer CR *et al.* (2013) SNP-array reveals genome-wide patterns of geographical and potential adaptive divergence across the natural range of Atlantic salmon (Salmo salar). *Molecular Ecology*, **22**, 532–551.

Bourret V, O'Reilly PT, Carr JW, Berg PR, Bernatchez L (2011) Temporal change in genetic integrity suggests loss of local adaptation in a wild Atlantic salmon (Salmo salar) population following introgression by farmed escapees. *Heredity*, **106**, 500–510.

Bowersox B, Wickersham T, Redfield L, Ackerman MW (2016) The Genetic Relationship between Anadromous and Resident Oncorhynchus mykiss at a Putative Barrier with Implications for Habitat Improvement. *Transactions of the American Fisheries Society*, **145**, 305–318.

Boy CC, Pérez AF, Lattuca ME, Calvo J, Morriconi E (2009) Reproductive biology of Galaxias maculatus (Jenyns 1842) in the Río Ovando, a high-latitude environment in southernmost Patagonia, Argentina. *Journal of Applied Ichthyology*, **25**, 661–668.

Bradbury IR, Campana SE, Bentzen P (2008) Otolith elemental composition and adult tagging reveal spawning site fidelity and estuarine dependency in rainbow smelt. *Marine Ecology Progress Series*, **368**, 255–268.

Brennan RS, Healy TM, Bryant HJ *et al.* (2018) Integrative Population and Physiological Genomics Reveals Mechanisms of Adaptation in Killifish. *Molecular biology and evolution*, **35**, 2639–2653.

Brieuc MSO, Ono K, Drinan DP, Naish KA (2015) Integration of Random Forest with population-based outlier analyses provides insight on the genomic basis and evolution of run timing in Chinook salmon (Oncorhynchus tshawytscha). *Molecular Ecology*, **24**, 2729–2746.

Brown RJ, Daum DW, Zuray SJ, Carter WK (2008) Documentation of annual spawning migrations of anadromous coregonid fishes in a large river using maturity indices, length and age analyses, and CPUE. *Advances in Limnology*, **63**, 101–116.

Burridge CP, Craw D, Fletcher D, Waters JM (2008) Geological dates and molecular rates: Fish DNA sheds light on time dependency. *Molecular Biology and Evolution*,

25, 624–633.

Burridge CP, Waters JM (2020) Does migration promote or inhibit diversification? A case study involving the dominant radiation of temperate Southern Hemisphere freshwater fishes. *Evolution*, **74**, 1954–1965.

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.

Campbell NR, Amish SJ, Pritchard VL *et al.* (2012) Development and evaluation of 200 novel SNP assays for population genetic studies of westslope cutthroat trout and genetic identification of related taxa. *Molecular Ecology Resources*, **12**, 942–949.

Campos H (1974) Population studies of Galaxias maculatus (Jenyns) (Osteichthys: Galaxiidae) in Chile with reference to the number of vertebrae. *Studies on the Neotropical Fauna*, **9**, 55–76.

Candy JR, Campbell NR, Grinnell MH *et al.* (2015) Population differentiation determined from putative neutral and divergent adaptive genetic markers in Eulachon (Thaleichthys pacificus, Osmeridae), an anadromous Pacific smelt. *Molecular Ecology Resources*, **15**, 1421–1434.

Cariou M, Duret L, Charlat S (2016) How and how much does RAD-seq bias genetic diversity estimates? *BMC Evolutionary Biology*, **16**, 1–8.

Carrea C, Cussac VE, Ruzzante DE (2013) Genetic and phenotypic variation among Galaxias maculatus populations reflects contrasting landscape effects between northern and southern Patagonia. *Freshwater Biology*, **58**, 36–49.

Carruthers M, Yurchenko AA, Augley JJ *et al.* (2018) De novo transcriptome assembly, annotation and comparison of four ecological and evolutionary model salmonid fish species. *BMC Genomics*, **19**, 448.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Molecular ecology*, **22**, 3124–40.

Catchen JM, Hohenlohe PA, Bernatchez L *et al.* (2017) Unbroken: RADseq remains a powerful tool for understanding the genetics of adaptation in natural populations. *Molecular Ecology Resources*, **17**, 362–365.

Cauwelier E, Gilbey J, Sampayo J, Stradmeyer L, Middlemas SJ (2018) Identification of a single genomic region associated with seasonal river return timing in adult Scottish Atlantic salmon (Salmo salar), using a genome-wide association study. *Canadian Journal of Fisheries and Aquatic Sciences*, **75**, 1427–1435.

Chabarria R, Furiness S, Patterson L *et al.* (2014) Genetic Structure and Demographic History of Endemic Micronesian Blue Riffle Goby, Stiphodon caeruleus (Gobiidae) Inferred from Mitochondrial DNA Sequence Analysis. *Copeia*, **2014**, 23–37.

Chalant A, Jézéquel C, Keith P, Hugueny B (2019) The global geography of fish diadromy modes. *Global Ecology and Biogeography*, **28**, 1272–1282.

Chang CW, Iizuka Y (2012) Estuarine use and movement patterns of seven sympatric

Mugilidae fishes: The Tatu Creek estuary, central western Taiwan. *Estuarine, Coastal and Shelf Science*, **106**, 121–126.

Chapman A, Morgan DL, Beatty SJ, Gill HS (2006) Variation in life history of land-locked lacustrine and riverine populations of Galaxias maculatus (Jenyns 1842) in Western Australia. *Environmental Biology of Fishes*, **77**, 21–37.

Chen WJ, Lavoué S, Mayden RL (2013) Evolutionary Origin And Early Biogeography Of Otophysan Fishes (Ostariophysi: Teleostei). *Evolution*, **67**, 2218–2239.

Chenoweth SF, Hughes JM (1997) Genetic population structure of the catadromous perciform: Macquaria novemaculeata (Percichthyidae). *Journal of Fish Biology*, **50**, 721–733.

Chessman BC, Williams WD (1975) Salinity tolerance and osmoregulatory ability of Galaxias maculatus (Jengns) (Pisces, Salmoniforormes Galaxiidae). *Freshwater Biology*, **5**, 135–140.

Chiang TY, Chen IS, Lin H Du, Hsiao ST, Ju YM (2015) Complete mitochondrial genome of the amphidromous, red-tailed goby Sicyopterus lagocephalus (Pallas) (Teleostei, Gobiidae). *Mitochondrial DNA*, **26**, 670–671.

Christensen KA, Leong JS, Sakhrani D *et al.* (2018) Chinook salmon (Oncorhynchus tshawytscha) genome and transcriptome. *PLoS ONE*, **13**, 1–15.

Chubb AL, Zink RM, Fitzsimons JM (1998) Patterns of mtDNA variation in Hawaiian freshwater fishes: The phylogeographic consequences of amphidromy. *Journal of Heredity*, **89**, 8–16.

Churcher AM, Pujolar JM, Milan M *et al.* (2015) Transcriptomic profiling of male European eel (Anguilla anguilla) livers at sexual maturity. *Comparative Biochemistry and Physiology - Part D: Genomics and Proteomics*, **16**, 28–35.

Clarke AD, Lewis A, Telmer KH, Shrimpton JM (2007) Life history and age at maturity of an anadromous smelt, the eulachon Thaleichthys pacificus (Richardson). *Journal of Fish Biology*, **71**, 1479–1493.

Clemens BJ, van de Wetering S, Sower SA, Schreck CB (2013) Maturation characteristics and life-history strategies of the Pacific lamprey, Entosphenus tridentatus. *Canadian Journal of Zoology*, **91**, 775–778.

Como S, Carpentier A, Rossi F *et al.* (2018) Stable isotopes as tracers can reveal resource allocation in juvenile golden gray mullets (Liza aurata, Risso, 1810). *Journal of Experimental Marine Biology and Ecology*, **503**, 72–79.

Conover CA, Bale LK, Overgaard MT *et al.* (2004) Metalloproteinase pregnancy-associated plasma protein A is a critical growth regulatory factor during fetal development. *Development*, **131**, 1187–1194.

Corush JB (2019) Evolutionary patterns of diadromy in fishes: more than a transitional state between marine and freshwater. *BMC Evolutionary Biology*, **19**, 168.

Côté CL, Castonguay M, Kalujnaia MS, Cramb G, Bernatchez L (2014) In absence of local adaptation, plasticity and spatially varying selection rule: a view from genomic

reaction norms in a panmictic species (Anguilla rostrata). *BMC genomics*, **15**, 403.

Couillard MA, Cabana G, Dery JF, Daigle G, Dodson JJ (2011) Ontogenetic Habitat Shifts of the Atlantic Tomcod (Microgadus tomcod) Across an Estuarine Transition Zone. *Estuaries and Coasts*, **34**, 1234–1245.

Covelo-Soto L, Leunda PM, Pérez-Figueroa A, Morán P (2015) Genome-wide methylation study of diploid and triploid brown trout (Salmo trutta L.). *Animal Genetics*, **46**, 280–288.

Crandall ED, Taffel JR, Barber PH (2010) High gene flow due to pelagic larval dispersal among South Pacific archipelagos in two amphidromous gastropods (Neritomorpha: Neritidae). *Heredity*, **104**, 563–572.

Crespel A, Dupont-Prinet A, Bernatchez L *et al.* (2017) Divergence in physiological factors affecting swimming performance between anadromous and resident populations of brook charr Salvelinus fontinalis. *Journal of Fish Biology*, **90**, 2170–2193.

Crook DA, Koster WM, Macdonald JI *et al.* (2010) Catadromous migrations by female tupong (Pseudaphritis urvillii) in coastal streams in VICtoria, Australia. *Marine and Freshwater Research*, **61**, 474–483.

Culling M (2013) Signatures of Selection on Growth, Shape, Parr Marks, and SNPs Among Seven Canadian Atlantic Salmon (Salmo Salar) Populations. *The Open Evolution Journal*, **7**, 1–16.

Currey MC, Bassham SL, Cresko WA (2019) Genetic divergence outpaces phenotypic evolution among threespine stickleback populations in old freshwater habitats. *Biological Journal of the Linnean Society*, **128**, 415–434.

Cussac VE, Ortubay S, Iglesias G *et al.* (2004) The distribution of South American galaxiid shes: the role of biological traits and post-glacial history. *Journal of Biogeography*, **31**, 103–121.

Czesny S, Epifanio J, Michalak P (2012) Genetic divergence between freshwater and marine morphs of alewife (Alosa pseudoharengus): a “next-generation” sequencing analysis. *PloS one*, **7**, e31803.

Danecek P, Auton A, Abecasis G *et al.* (2011) The variant call format and VCFtools. *Bioinformatics*, **27**, 2156–2158.

Davidson WS, Koop BF, Jones SJM *et al.* (2010) Sequencing the genome of the Atlantic salmon (Salmo salar). *Genome Biology*, **11**, 403.

Davis AM, Unmack PJ, Pusey BJ, Johnson JB, Pearson RG (2012) Marine-freshwater transitions are associated with the evolution of dietary diversification in terapontid grunters (Teleostei: Terapontidae). *Journal of Evolutionary Biology*, **25**, 1163–1179.

Dean LL, Magalhaes IS, Foote A *et al.* (2019) Admixture between Ancient Lineages, Selection, and the Formation of Sympatric Stickleback Species-Pairs. *Molecular Biology and Evolution*, **36**, 2481–2497.

Debiasse MB, Kawji Y, Kelly MW (2018) Phenotypic and transcriptomic responses to

salinity stress across genetically and geographically divergent Tigriopus californicus populations. *Molecular ecology*, **27**, 1621–1632.

Defaveri J, Merila J (2014) Local adaptation to salinity in the three-spined stickleback? *Journal of Evolutionary Biology*, **27**, 290–302.

Defaveri J, Shikano T, Shimada Y, Goto A, Merilä J (2011) Global analysis of genes involved in freshwater adaptation in threespine sticklebacks (Gasterosteus aculeatus). *Evolution*, **65**, 1800–1807.

Delgado ML, Górski K, Habit E, Ruzzante DE (2019) The effects of diadromy and its loss on genomic divergence: The case of amphidromous Galaxias maculatus populations. *Molecular Ecology*, **28**, 5217–5231.

Delgado ML, Manosalva A, Urbina MA *et al.* (2020) *Genomic basis of the loss of diadromy in Galaxias maculatus : insights from reciprocal transplant experiments* .

Delgado ML, Ruzzante DE (2020) Investigating Diadromy in Fishes and Its Loss in an -Omics Era. *iScience*, **23**, 101837.

Dennenmoser S, Vamosi SM, Nolte AW, Rogers SM (2017) Adaptive genomic divergence under high gene flow between freshwater and brackish-water ecotypes of prickly sculpin (Cottus asper) revealed by Pool-Seq. *Molecular Ecology*, **26**, 25–42.

Dickman M (1995) An isolated population of fourhorn sculpins (Myoxocephalus quadricornis, family Cottidae) in a hypersaline high arctic Canadian lake. *Hydrobiologia*, **312**, 27–35.

Dodson JJ, Aubin-Horth N, Thériault V, Páez DJ (2013) The evolutionary ecology of alternative migratory tactics in salmonid fishes. *Biological Reviews*, **88**, 602–625.

Dodson JJ, Laroche J, Lecomte F (2009) Contrasting evolutionary pathways of anadromy in euteleostean fishes. *American Fisheries Society Symposium*, **69**, 63–77.

Donaldson TJ, Myers RF (2002) Insular freshwater fish faunas of Micronesia: Patterns of species richness and similarity. *Environmental Biology of Fishes*, **65**, 139–149.

Dou SZ, Yokouchi K, Yu X *et al.* (2012) The migratory history of anadromous and non-anadromous tapertail anchovy Coilia nasus in the Yangtze River Estuary revealed by the otolith Sr:Ca ratio. *Environmental Biology of Fishes*, **95**, 481–490.

Drevecky CJ, Falco R, Aguirre WE (2013) Genetic divergence of a sympatric lake-resident-anadromous three-spined stickleback Gasterosteus aculeatus species pair. *Journal of Fish Biology*, **83**, 111–132.

Drywa A, Poćwierz-Kotus A, Was A *et al.* (2013) Genotyping of two populations of Southern Baltic Sea trout Salmo trutta m. trutta using an Atlantic salmon derived SNP-array. *Marine Genomics*, **9**, 25–32.

Duan J, Zhang H, Liu K *et al.* (2012) An overview of Coilia ectenes in Jiangsu Section of the Yangtze River. *Agricultural Science & Technology*, **13**, 1950–1954.

Duarte GFF (2018) Ghost of Diadromous Fish Past: Streamlining Research on Diadromous Fish Species Using Historical Data At the European Scale. Universidade de Lisboa.

Dymowska AK, Boyle D, Schultz AG, Goss GG (2015) The role of acid-sensing ion channels in epithelial Na+ uptake in adult zebrafish (Danio rerio). *Journal of Experimental Biology*, **218**, 1244–1251.

Ebner BC, Donaldson JA, Allen GR, Keith P (2017) Visual census, photographic records and the trial of a video network provide first evidence of the elusive Sicyopterus cynocephalus in Australia. *Cybium*, **41**, 117–125.

Ebner BC, Thuesen P (2011) Discovery of stream-cling-goby assemblages (Stiphodon species) in the Australian Wet Tropics. *Australian Journal of Zoology*, **58**, 331–340.

Edeline E (2007) Adaptive phenotypic plasticity of eel diadromy. *Marine Ecology Progress Series*, **341**, 229–232.

Edo K, Kawaguchi Y, Nunokawa M, Kawamula H, Higashi S (2005) Morphology, stomach contents and growth of the endangered salmonid, Sakhalin taimen Hucho perryi, captured in the Sea of Okhotsk, northern Japan: Evidence of an anadromous form. *Environmental Biology of Fishes*, **74**, 1–7.

Edwards SL, Marshall WS (2012) *Principles and Patterns of Osmoregulation and Euryhalinity in Fishes*.

Elahi N, Yousuf F, Tabassum S *et al.* (2017) Life-history Traits of the Blacktrip sardinella, Sardinella melanura (Clupeidae) in the Gwadar, Balochistan Coast, Pakistan. *Indian Journal of Geo-Marine Sciences (IJMS)*, **46**, 397–404.

Elias A, McLaughlin R, Mackereth R, Wilson C, Nichols KM (2018) Population structure and genomic variation of ecological life history diversity in wild-caught Lake Superior brook trout, Salvelinus fontinalis. *Journal of Great Lakes Research*, **44**, 1373–1382.

Ellegren H (2014) Genome sequencing and population genomics in non-model organisms. *Trends in Ecology and Evolution*, **29**, 51–63.

Elmer KR, Meyer A (2011) Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends in ecology & evolution*, **26**, 298–306.

Erkinaro J, Pritchard VL, Primmer CR *et al.* (2018) Genomic signatures of fine-scale local selection in Atlantic salmon suggest involvement of sexual maturation, energy homeostasis and immune defence-related genes. *Molecular Ecology*, **27**, 2560–2575.

Fellous A, Shama LNS (2019) Genome Survey of Chromatin-Modifying Enzymes in Threespine Stickleback: A Crucial Epigenetic Toolkit for Adaptation? *Frontiers in Marine Science*, **6**, 1–16.

Ferchaud AL, Hansen MM (2016) The impact of selection, gene flow and demographic history on heterogeneous genomic divergence: Three-spine sticklebacks in divergent environments. *Molecular Ecology*, **25**, 238–259.

Ferguson A, Reed TE, Cross TF, McGinnity P, Prodöhl PA (2019) Anadromy, potamodromy and residency in brown trout Salmo trutta: the role of genes and the environment. *Journal of Fish Biology*, **95**, 692–718.

Feutry P, Castelin M, Ovenden JR *et al.* (2013) Evolution of diadromy in fish: insights

from a tropical genus (Kuhlia species). *The American naturalist*, **181**, 52–63.

Feutry P, Tabouret H, Maeda K, Pécheyran C, Keith P (2012) Diadromous life cycle and behavioural plasticity in freshwater and estuarine Kuhliidae species (Teleostei) revealed by otolith microchemistry. *Aquatic Biology*, **15**, 195–204.

Filipe AF, Araújo MB, Doadrio I, Angermeier PL, Collares-Pereira MJ (2009) Biogeography of Iberian freshwater fishes revisited: The roles of historical versus contemporary constraints. *Journal of Biogeography*, **36**, 2096–2110.

Fischer EK, Song Y, Hughes KA, Zhou W, Hoke KL (2021) Non-parallel transcriptional divergence during parallel adaptation. *Molecular Ecology*.

Flura -, Alam MA, Hossain MRA (2018) A review on Silonia silondia (Hamilton, 1822) threatened fish of the world: (Siluriformes: Schilbeidae). *Research in Agriculture Livestock and Fisheries*, **5**, 235–240.

Foll M, Gaggiotti O (2008) A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics*, **180**, 977–993.

Fox SE, Christie MR, Marine M *et al.* (2014) Marine Genomics Sequencing and characterization of the anadromous steelhead (Oncorhynchus mykiss) transcriptome Frequency Biological process. *Marine Genomics*, **15**, 13–15.

Franklin P, Gee E (2019) Living in an amphidromous world: Perspectives on the management of fish passage from an island nation. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **29**, 1424–1437.

Frichot E, François O (2015) LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, **6**, 925–929.

Frotté L, Ringelstein J, Monti D *et al.* (2019) Detection of full and limited amphidromous migratory dynamics of fish in Caribbean rivers. *Ecology of Freshwater Fish*, eff.12501.

Fuentes-Pardo AP, Ruzzante DE (2017) Whole-genome sequencing approaches for conservation biology: Advantages, limitations and practical recommendations. *Molecular Ecology*, **26**, 5369–5406.

Fuji T, Kasai A, Yamashita Y (2018) Upstream migration mechanisms of juvenile temperate sea bass Lateolabrax japonicus in the stratified Yura River estuary. *Fisheries Science*, **84**, 163–172.

Fyhn HJ, Finn RN, Reith M (1999) Yolk protein hydrolysis and oocyte free amino acids as key features in the adaptive evolution of teleost fishes to seawater. *Sarsia*, **84**, 451–456.

Gallagher ZS, Bystriansky JS, Farrell AP, Brauner CJ (2013) A novel pattern of smoltification in the most anadromous salmonid: Pink salmon (Oncorhynchus gorbuscha). *Canadian Journal of Fisheries and Aquatic Sciences*, **70**, 349–357.

Ganguly S, Mitra T, Mahanty A, Mohanty S, Mohanty BP (2020) A comparative metabolomics study on anadromous clupeid Tenualosa ilisha for better

understanding the influence of habitat on nutritional composition. *Metabolomics*, **16**, 1–14.

Garbin T, Castello JP, Kinas PG (2014) Age, growth, and mortality of the mullet Mugil liza in Brazil's southern and southeastern coastal regions. *Fisheries Research*, **149**, 61–68.

Gavery MR, Nichols KM, Goetz GW, Middleton MA, Swanson P (2018) Characterization of genetic and epigenetic variation in sperm and red blood cells from adult hatchery and natural-origin steelhead, Oncorhynchus mykiss. *G3: Genes, Genomes, Genetics*, **8**, 3723–3736.

Gibbons TC, Metzger DCH, Healy TM, Schulte PM (2017) Gene expression plasticity in response to salinity acclimation in threespine stickleback ecotypes from different salinity habitats. *Molecular Ecology*, **26**, 2711–2725.

Giger T, Excoffier L, Amstutz U *et al.* (2008) Population transcriptomics of life-history variation in the genus Salmo. *Molecular Ecology*, **17**, 3095–3108.

Glazer AM, Cleves PA, Erickson PA, Lam AY, Miller CT (2014) Parallel developmental genetic features underlie stickleback gill raker evolution. *EvoDevo*, **5**, 1–16.

Gong X, Davenport ER, Wang D, Clark AG (2019) Lack of spatial and temporal genetic structure of Japanese eel (Anguilla japonica) populations. *Conservation Genetics*, **20**, 467–475.

González-Murcia S, Álvarez FS (2018) Your place, my place…, distribution of Agonostomus monticola and Sicydium multipunctatum in the Acahuapa Watershed. *Revista Mexicana de Biodiversidad*, **89**, 854–864.

González-Wevar C, Salinas P, Hüne M *et al.* (2015a) Contrasting Genetic Structure and Diversity of *Galaxias maculatus* (Jenyns, 1848) Along the Chilean Coast: Stock Identification for Fishery Management. *Journal of Heredity*, **106**, 439–447.

González-Wevar C a., Salinas P, Hüne M *et al.* (2015b) Phylogeography in Galaxias maculatus (Jenyns, 1848) along Two Biogeographical Provinces in the Chilean Coast. *Plos One*, **10**, e0131289.

Gordoa A (2009) Characterization of the infralittoral system along the north-east Spanish coast based on sport shore-based fishing tournament catches. *Estuarine, Coastal and Shelf Science*, **82**, 41–49.

Górski K, Habit EM, Pingram MA, Manosalva AJ (2018) Variation of the use of marine resources by Galaxias maculatus in large Chilean rivers. *Hydrobiologia*, **814**, 61–73.

Goto A, Arai T (2003) Migratory histories of three types of Cottus pollux (small-egg, middle-egg, and large-egg types) as revealed by otolith microchemistry. *Ichthyological Research*, **50**, 67–72.

Goudet J (2005) HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, **5**, 184–186.

Gross M (1987) Evolution of Diadromy in Fishes. *American Fisheries Society Symposium*, **1**, 14–25.

Gross M (1997) Evolution of life history and migration in fish. *Memoirs of the faculty of fisheries Hokkaido University*, **44**, 17–21.

Guo B, DeFaveri J, Sotelo G, Nair A, Merilä J (2015) Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. *BMC Biology*, **13**, 19.

Guo B, Fang B, Shikano T *et al.* (2019) A phylogenomic perspective on diversity, hybridization and evolutionary affinities in the stickleback genus Pungitius. *Molecular Ecology*, **28**, 4046–4064.

Habit E, García A, Díaz G *et al.* (2019) River science and management issues in Chile: Hydropower development and native fish communities. *River Research and Applications*, **35**, 489–499.

Habit E, Piedra P, Ruzzante DE *et al.* (2010) Changes in the distribution of native fishes in response to introduced species and other anthropogenic effects. *Global Ecology and Biogeography*, **19**, 697–710.

Haenel Q, Roesti M, Moser D, MacColl ADC, Berner D (2019) Predictable genome-wide sorting of standing genetic variation during parallel adaptation to basic versus acidic environments in stickleback fish. *Evolution Letters*, **3**, 28–42.

Hale MC, McKinney GJ, Thrower FP, Nichols KM (2016) RNA-seq reveals differential gene expression in the brains of juvenile resident and migratory smolt rainbow trout (Oncorhynchus mykiss). *Comparative Biochemistry and Physiology - Part D: Genomics and Proteomics*, **20**, 136–150.

Hale R, Swearer SE (2008) Otolith microstructural and microchemical changes associated with settlement in the diadromous fish Galaxias maculatus. *Marine Ecology Progress Series*, **354**, 229–234.

Hale MC, Thrower FP, Berntson EA, Miller MR, Nichols KM (2013) Evaluating adaptive divergence between migratory and nonmigratory ecotypes of a salmonid fish, Oncorhynchus mykiss. *G3: Genes, Genomes, Genetics*, **3**, 1273–1285.

Handeland SO, Berge Å, Björnsson BT, Stefansson SO (1998) Effects of temperature and salinity on osmoregulation and growth of Atlantic salmon (Salmo salar L.) smolts in seawater. In: *Aquaculture*, pp. 289–302.

Härer A, Bolnick DI, Rennison DJ (2021) The genomic signature of ecological divergence along the benthic-limnetic axis in allopatric and sympatric threespine stickleback. *Molecular Ecology*, **30**, 451–463.

Hasegawa EI (2012) Chum salmon Oncorhynchus keta respond to moonlight during homeward migrations. *Journal of Fish Biology*, **81**, 632–641.

Hashemi SA, Taghavimotlagh SA, Hedayati A, Ghorbani R (2013) Original article Some biological aspect of thinspine seacatfish (plicofollis tenuispinis) in northwest of persian gulf (Khuzestan coastal waters , Iran). , **2**, 176–183.

Hasselman DJ, Ricard D, Bentzen P (2013) Genetic diversity and differentiation in a wide ranging anadromous fish, American shad (Alosa sapidissima), is correlated with latitude. *Molecular Ecology*, **22**, 1558–1573.

Hauge H, Dahle MK, Kristoffersen AB *et al.* (2016) The genome of a landlocked Atlantic salmon Salmo salar characterized through high-throughput sequencing. *Journal of Evolutionary Biology Research*, **8**, 1–14.

Hecht BC, Campbell NR, Holecek DE, Narum SR (2013) Genome-wide association reveals genetic basis for the propensity to migrate in wild populations of rainbow and steelhead trout. *Molecular Ecology*, **22**, 3061–3076.

Heim-Ballew H, Moody KN, Blum MJ, McIntyre PB, Hogan JD (2020) Migratory flexibility in native Hawai'ian amphidromous fishes. *Journal of Fish Biology*, **96**, 456–468.

Hendry AP (2016) Key questions on the role of phenotypic plasticity in eco-evolutionary dynamics. *Journal of Heredity*, **107**, 25–41.

Hendry AP, Day T (2005) Population structure attributable to reproductive time : isolation by time and adaptation by time. *Molecular ecology*, **14**, 901–916.

Hess JE, Campbell NR, Close D a., Docker MF, Narum SR (2013) Population genomics of Pacific lamprey: adaptive variation in a highly dispersive species. *Molecular Ecology*, **22**, 2898–2916.

Hickford MJH, Schiel DR (2016) Otolith microchemistry of the amphidromous Galaxias maculatus shows recruitment to coastal rivers from unstructured larval pools. *Marine Ecology Progress Series*, **548**, 197–207.

Hindar K, Jonsson B, Ryman N, Ståhl G (1991) Genetic relationships among landlocked, resident, and anadromous brown trout, *Salmo trutta* L. *Heredity*, **66**, 83–91.

Hirase S, Kokita T, Nagano AJ, Kikuchi K (2020) Genomic and phenotypic consequences of two independent secondary contact zones between allopatric lineages of the anadromous ice goby Leucopsarion petersii. *Heredity*, **124**, 223–235.

Hiroi J, McCormick SD (2012) New insights into gill ionocyte and ion transporter function in euryhaline and diadromous fish. *Respiratory Physiology and Neurobiology*, **184**, 257–268.

Hoffberg SL, Kieran TJ, Catchen JM *et al.* (2016) RADcap: sequence capture of dual-digest RADseq libraries with identifiable duplicates and reduced missing data. *Molecular Ecology Resources*, **16**, 1264–1278.

Hogan Z, Baird IG, Radtke R, Vander Zanden MJ (2007) Long distance migration and marine habitation in the tropical Asian catfish, Pangasius krempfi. *Journal of Fish Biology*, **71**, 818–832.

Hogan JD, Blum MJ, Gilliam JF, Bickford N, McIntyre PB (2014) Consequences of alternative dispersal strategies in a putatively amphidromous fish. *Ecology*, **95**, 2397–2408.

Hogan JD, Kozdon R, Blum MJ *et al.* (2017) Reconstructing larval growth and habitat use in an amphidromous goby using otolith increments and microchemistry. *Journal of Fish Biology*, **90**, 1338–1355.

Hohenlohe PA, Bassham S, Etter PD *et al.* (2010) Population genomics of parallel

adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics*, **6**, e1000862.

Honţ Ștefan, Paraschiv M, Ion Iani M *et al.* (2019) Detailed analysis of beluga sturgeon (Huso huso) and stellate sturgeon (Acipenser stellatus) migration in the Lower Danube River. *Turkish Journal of Zoology*, **43**, 457–464.

Houston RD, Taggart JB, Cézard T *et al.* (2014) Development and validation of a high density SNP genotyping array for Atlantic salmon (Salmo salar). *BMC Genomics*, **15**, 1–13.

Hughes LC, Ortí G, Huang Y *et al.* (2018) Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. *Proceedings of the National Academy of Sciences of the United States of America*, **115**, 6249–6254.

Hughes JM, Schmidt DJ, Macdonald JI, Huey JA, Crook DA (2014) Low interbasin connectivity in a facultatively diadromous fish: Evidence from genetics and otolith chemistry. *Molecular Ecology*, **23**, 1000–1013.

Hume JB, Recknagel H, Bean CW, Adams CE, Mable BK (2018) RADseq and mate choice assays reveal unidirectional gene flow among three lamprey ecotypes despite weak assortative mating: Insights into the formation and stability of multiple ecotypes in sympatry. *Molecular Ecology*, **27**, 4572–4590.

Hwang P-P, Lee T-H, Lin L-Y (2011) Ion regulation in fish gills: recent progress in the cellular and molecular mechanisms. *AJP: Regulatory, Integrative and Comparative Physiology*, **301**, R28–R47.

Igarashi Y, Zhang H, Tan E *et al.* (2018) Whole-genome sequencing of 84 Japanese eels reveals evidence against panmixia and support for sympatric speciation. *Genes*, **9**, 474.

Iguchi K, Mizuno N (1999) Early starvation limits survival in amphidromous fishes. *Journal of Fish Biology*, **54**, 705–712.

Iida M, Kondo M, Tabouret H *et al.* (2017) Specific gravity and migratory patterns of amphidromous gobioid fish from Okinawa Island, Japan. *Journal of Experimental Marine Biology and Ecology*, **486**, 160–169.

Imoto JM, Saitoh K, Sasaki T *et al.* (2013) Phylogeny and biogeography of highly diverged freshwater fish species (Leuciscinae, Cyprinidae, Teleostei) inferred from mitochondrial genome analysis. *Gene*, **514**, 112–124.

Imrie D, Sadler KC (2010) White adipose tissue development in zebrafish is regulated by both developmental time and fish size. *Developmental Dynamics*, **239**, 3013–3023.

Ishihara T, Tachihara K (2008) Reproduction and early development of a freshwater pipefish Microphis leiaspis in Okinawa-jima Island, Japan. *Ichthyological Research*, **55**, 349–355.

Ishikawa A, Kabeya N, Ikeya K *et al.* (2019) A key metabolic gene for recurrent freshwater colonization and radiation in fishes. *Science*, **364**, 886–889.

Jara-Seguel P, Valdebenito I, Palma-Rojas C, Rebolledo C (2008) Nuclear DNA content

in Galaxias maculatus (Teleostei: Osmeriformes: Galaxiidae). *Latin American Journal of Aquatic Research*, **36**, 87–91.

Jarvis MG, Harland HA, Warburton ML, Closs GP (2018) The spawning and early life-history of a New Zealand endemic amphidromous eleotrid, bluegill bully (Gobiomorphus hubbsi). *New Zealand Journal of Marine and Freshwater Research*, **52**, 55–68.

Jeffery NW, Stanley RRE, Wringe BF *et al.* (2017) Range-wide parallel climate-associated genomic clines in Atlantic salmon. *Royal Society Open Science*, **4**, 171394.

Jeong SY, Kim JH, Lee WO, Dahms HU, Han KN (2014) Salinity changes in the anadromous river pufferfish, Takifugu obscurus, mediate gene regulation. *Fish Physiology and Biochemistry*, **40**, 205–219.

Jessop BM, Cairns DK, Thibault I, Tzeng WN (2007) Life history of American eel Anguilla rostrata: New insights from otolith microchemistry. *Aquatic Biology*, **1**, 205–216.

Jeukens J, Renaut S, St-Cyr J, Nolte AW, Bernatchez L (2010) The transcriptomics of sympatric dwarf and normal lake whitefish (Coregonus clupeaformis spp., Salmonidae) divergence as revealed by next-generation sequencing. *Molecular Ecology*, **19**, 5389–5403.

Jia Y, Yin S, Li L *et al.* (2016) iTRAQ proteomic analysis of salinity acclimation proteins in the gill of tropical marbled eel (Anguilla marmorata). *Fish Physiology and Biochemistry*, **42**, 935–946.

Johnston SE, Orell P, Pritchard VL *et al.* (2014) Genome-wide SNP analysis reveals a genetic basis for sea-age variation in a wild population of Atlantic salmon (Salmo salar). *Molecular Ecology*, **23**, 3452–3468.

Jombart T, Ahmed I (2011) adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, **27**, 3070–3071.

Jones F, Grabherr M, Chan Y *et al.* (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, **484**, 55–61.

Jørgensen EH, Johnsen HK (2014) Rhythmic life of the Arctic charr: Adaptations to life at the edge. *Marine Genomics*, **14**, 71–81.

Kamvar ZN, Brooks JC, Grünwald NJ (2015) Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics*, **6**, 1–10.

Kassambara A (2020a) rstatix: Pipe-friendly framework for basic statistical test.

Kassambara A (2020b) ggbur: “ggplot2” based publication ready plots.

Katayama S, Radtke RL, Omori M, Shafer DJ (2000) Coexistence of anadromous and resident life history styles of pond smelt, Hypomesus nipponensis, in Lake Ogawara, Japan, as determined by analyses of otolith structure and strontium: Calcium ratios. *Environmental Biology of Fishes*, **58**, 195–201.

Keith P (2003) Biology and ecology of amphidromous Gobiidae of the Indo-Pacific and

the Caribbean regions. *Journal of Fish Biology*, **63**, 831–847.

Kelson SJ, Miller MR, Thompson TQ, O'Rourke SM, Carlson SM (2019) Do genomics and sex predict migration in a partially migratory salmonid fish, Oncorhynchus mykiss? *Canadian Journal of Fisheries and Aquatic Sciences*, **76**, 2080–2088.

Kendall NW, McMillan JR, Sloat MR *et al.* (2014) Anadromy and residency in steelhead and rainbow trout (Oncorhynchus mykiss): a review of the processes and patterns. *Canadian Journal of Fisheries and Aquatic Sciences*, **72**, 319–342.

Kim JH, Han KN, Lee JS (2014) Complete mitochondrial genome of the anadromous river pufferfish, takifugu obscurus (tetraodontiformes: Tetraodontidae). *Mitochondrial DNA*, **25**, 46–47.

Kissinger BC, Gantner N, Anderson WG *et al.* (2016) Brackish-water residency and semi-anadromy in Arctic lake trout (Salvelinus namaycush) inferred from otolith microchemistry. *Journal of Great Lakes Research*, **42**, 267–275.

Kitano J, Ishikawa A, Kusakabe M (2019) Parallel transcriptome evolution in stream threespine sticklebacks. *Development Growth and Differentiation*, **61**, 104–113.

Kohestan-Eskandari S, Anvarifar H, Mousavi-Sabet H, Yousefi M, Khanzade M (2014) A morphology-based hypothesis for homeward migration success and population differentiation in the anadromous kutum Rutilus kutum (Pisces: Cyprinidae) along the southern Caspian Sea, Iran. *Folia Zoologica*, **63**, 151–160.

Kokita T, Nohara K (2011) Phylogeography and historical demography of the anadromous fish Leucopsarion petersii in relation to geological history and oceanography around the Japanese Archipelago. *Molecular Ecology*, **20**, 143–164.

Konijnendijk N, Shikano T, Daneels D, Volckaert FAM, Raeymaekers JAM (2015) Signatures of selection in the three-spined stickleback along a small-scale brackish water - freshwater transition zone. *Ecology and Evolution*, **5**, 4174–4186.

Kopylova E, Noé L, Touzet H (2012) SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*.

Koshelev VN, Mikodina E V., Mironova TN, Presnyakov A V., Novosadov AG (2012) New data on biology and distribution of Sakhalin sturgeon Acipenser mikadoi. *Journal of Ichthyology*, **52**, 619–627.

Koshelev VN, Ruban G, Shmigirilov A (2014) Spawning migrations and reproductive parameters of the kaluga sturgeon, Huso dauricus (Georgi, 1775), and Amur sturgeon, Acipenser schrenckii (Brandt, 1869). *Journal of Applied Ichthyology*, **30**, 1125–1132.

Kovach RP, Gharrett AJ, Tallmon DA (2013) Temporal patterns of genetic variation in a salmon population undergoing rapid change in migration timing. *Evolutionary Applications*, **6**, 795–807.

Kozak GM, Brennan RS, Berdan EL, Fuller RC, Whitehead A (2014) Functional and population genomic divergence within and between two species of killifish adapted to different osmotic niches. *Evolution*, **68**, 63–80.

Krück NC, Innes DI, Ovenden JR (2013) New SNPs for population genetic analysis reveal possible cryptic speciation of eastern Australian sea mullet (Mugil cephalus). *Molecular Ecology Resources*, **13**, 715–725.

Kultz D (2015) Physiological mechanisms used by fish to cope with salinity stress. *Journal of Experimental Biology*, **218**, 1907–1914.

Kusakabe M, Ishikawa A, Ravinet M *et al.* (2017) Genetic basis for variation in salinity tolerance between stickleback ecotypes. *Molecular Ecology*, **26**, 304–319.

Kusmic C, Gualtieri P (2000) Morphology and spectral sensitivities of retinal and extraretinal photoreceptors in freshwater teleosts. *Micron*, **31**, 183–200.

Kuzishchin K V., Gruzdeva MA, Filenko VA, Pavlov DS (2020) Caspian Anadromous Shad Alosa kessleri kessleri (Grimm, 1887) from the Akhtuba River, Lower Volga River Basin: Biological and Morphological Features. *Inland Water Biology*, **13**, 79–87.

Kynard B (1997) Life history, latitudinal patterns, and status of the shortnose sturgeon, Acipenser brevirostrum. *Environmental Biology of Fishes*, **48**, 319–334.

Kynard B, Zhuang P, Zhang T, Zhang L (2003) Ontogenetic behavior and migration of Dabry's sturgeon, Acipenser dabryanus, from the Yangtze River, with notes on body color and development rate. *Environmental Biology of Fishes*, **66**, 27–36.

Lahti DC, Johnson NA, Ajie BC *et al.* (2009) Relaxed selection in the wild. *Trends in Ecology and Evolution*, **24**, 487–496.

Lamson HM, Cairns DK, Shiao JC, Iizuka Y, Tzeng WN (2009) American eel, Anguilla rostrata, growth in fresh and salt water: Implications for conservation and aquaculture. *Fisheries Management and Ecology*, **16**, 306–314.

Larson WA, Limborg MT, McKinney GJ *et al.* (2016) Genomic islands of divergence linked to ecotypic variation in sockeye salmon. *Molecular Ecology*, **26**, 554–570.

Laurenson LJB, French RP, Jones P *et al.* (2012) Aspects of the biology of Galaxias maculatus. *Journal of fish biology*, **81**, 1085–100.

Layton K, Dempson J, Snelgrove P *et al.* (2020) Resolving fine-scale population structure and fishery exploitation using sequenced microsatellites in a northern fish. *Evolutionary Applications*, 1–14.

Lee CE, Bell MA (1999) Causes and consequences of recent freshwater. *Tree*, **14**, 284–288.

Lehtonen H, Nyberg K, Vuoronen PJ, Leskelä A (1992) Radioactive strontium (85Sr) in marking whitefish [Coregonus lavaretus (L.)] larvae and the dispersal of larvae from river to sea. *Journal of Fish Biology*, **41**, 417–423.

Leitwein M, Garza JC, Pearse DE (2017) Ancestry and adaptive evolution of anadromous, resident, and adfluvial rainbow trout (Oncorhynchus mykiss) in the San Francisco bay area: application of adaptive genomic variation to conservation in a highly impacted landscape. *Evolutionary Applications*, **10**, 56–67.

Lemay MA, Donnelly DJ, Russello MA (2013) Transcriptome-wide comparison of

sequence variation in divergent ecotypes of kokanee salmon. *BMC Genomics*, **14**, 308.

Lemmetyinen J, Piironen J, Kiiskinen P, Hassinen M, Vornanen M (2013) Comparison of gene expression in the gill of salmon (Salmo salar) smolts from anadromous and landlocked populations. *Ann Zool Fennici*, **50**, 16–35.

Lemopoulos A, Uusi-Heikkilä S, Huusko A, Vasemägi A, Vainikka A (2018a) Comparison of migratory and resident populations of brown trout reveals candidate genes for migration tendency. *Genome Biology and Evolution*, **10**, 1493–1503.

Lemopoulos A, Uusi-heikkilä S, Vasemägi A, Huusko A (2018b) Genome-wide divergence patterns support fine-scaled genetic structuring associated with migration tendency in brown trout. *Canadian Journal of Fisheries and Aquatic Sciences*, **999**, 1–13.

Lenz TL, Eizaguirre C, Rotter B, Kalbe M, Milinski M (2013) Exploring local immunological adaptation of two stickleback ecotypes by experimental infection and transcriptome-wide digital gene expression analysis. *Molecular Ecology*, **22**, 774–786.

Li YL, Xue DX, Zhang BD, Liu JX (2019) Population genomic signatures of genetic structure and environmental selection in the catadromous roughskin sculpin trachidermus fasciatus. *Genome Biology and Evolution*, **11**, 1751–1764.

Limburg KE (2001) Otolith microchemistry indicates unexpected patterns of residency and anadromy in blueback herring, alosa aestivalis, in the hudson and mohawk rivers. *BFPP - Bulletin Francais de la Peche et de la Protection des Milieux Aquatiques*, 931–938.

Lischer HEL, Excoffier L (2012) PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, **28**, 298–299.

Liu S, Ferchaud AL, Grønkjær P, Nygaard R, Hansen MM (2018) Genomic parallelism and lack thereof in contrasting systems of three-spined sticklebacks. *Molecular Ecology*, **27**, 4725–4743.

Liu S, Gao G, Palti Y *et al.* (2014) RNA-seq analysis of early hepatic response to handling and confinement stress in rainbow trout. *PLoS ONE*, **9**, 1–8.

Lord C, Bellec L, Dettaï A, Bonillo C, Keith P (2019) Does your lip stick? Evolutionary aspects of the mouth morphology of the Indo-Pacific clinging goby of the Sicyopterus genus (Teleostei: Gobioidei: Sicydiinae) based on mitogenome phylogeny. *Journal of Zoological Systematics and Evolutionary Research*, **57**, 910–925.

Lord C, Maeda K, Keith P, Watanabe S (2015) Population structure of the asian amphidromous Sicydiinae goby, Stiphodon percnopterygionus, inferred from mitochondrial COI sequences, with comments on larval dispersal in the northwest pacific ocean. *Vie et Milieu*, **65**, 63–71.

Love M, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion

for RNA-seq data with DESeq2. *Genome Biology*, **15**, 550.

Lowerre-Barbieri S, Villegas-Rís D, Walters S *et al.* (2014) Spawning site selection and contingent behavior in common snook, Centropomus undecimalis. *PLoS ONE*, **9**.

Lu XJ, Zhang H, Yang GJ, Li MY, Chen J (2016) Comparative transcriptome analysis on the alteration of gene expression in ayu (Plecoglossus altivelis) larvae associated with salinity change. *Zoological research*, **37**, 126–135.

Lugo Ramos JS, Delmore KE, Liedvogel M (2017) Candidate genes for migration do not distinguish migratory and non-migratory birds. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, **203**, 383–397.

Łuszczek-Trojnar E, Drąg-Kozak E, Kleszcz M, Popek W, Epler P (2008) Gonadal maturity in vimba (Vimba vimba L.) raised in carp ponds. *Journal of Applied Ichthyology*, **24**, 316–320.

Luu K, Bazin E, Blum MGB (2017) pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, **17**, 67–77.

Le Luyer J, Laporte M, Beacham TD *et al.* (2017) Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon. *Proceedings of the National Academy of Sciences*, **114**, 12964–12969.

Lyle AA, Maitland PS (1997) The spawning migration and conservation of smelt Osmerus eperlanus in the River Cree, Southwest Scotland. *Biological Conservation*, **80**, 303–311.

Lyons J, Schneider DW (1990) Factors influencing fish distribution and community structure in a small coastal river in southwestern Costa Rica. *Hydrobiologia*, **203**, 1–14.

Ma Q, Liu X, Feng W, Liu S, Zhuang Z (2018) Analyses of the molecular mechanisms associated with salinity adaption of Trachidermus fasciatus through combined iTRAQ-based proteomics and RNA sequencing-based transcriptomics. *Progress in Biophysics and Molecular Biology*, **136**, 40–53.

MacManes MD (2018) The Oyster River Protocol: A multi-assembler and kmer approach for de novo transcriptome assembly. *PeerJ*, 6:e5428.

Madsen RPA, Jacobsen MW, O'Malley KG *et al.* (2020) Genetic population structure and variation at phenology-related loci in anadromous Arctic char (Salvelinus alpinus). *Ecology of Freshwater Fish*, **29**, 170–183.

Maeda K (2014) Stiphodon niraikanaiensis, a new species of sicydiine goby from Okinawa Island (Gobiidae: Sicydiinae). *Ichthyological Research*, **61**, 99–107.

Maeda K, Mukai T, Tachihara K (2011) A new species of amphidromous goby, Stiphodon alcedo, from the ryukyu archipelago (Gobiidae: Sicydiinae). *Cybium*, **35**, 285–298.

Maeda K, Tachihara K (2005) Recruitment of amphidromous sleepers Eleotris acanthopoma, Eleotris melanosoma, and Eleotris fusca into the Teima River,

Okinawa Island. *Ichthyological Research*, **52**, 325–335.

Magalhaes IS, Whiting JR, D'Agostino D *et al.* (2021) Intercontinental genomic parallelism in multiple three-spined stickleback adaptive radiations. *Nature Ecology and Evolution*, **5**, 251–261.

Mai ACG, dos Santos ML, Lemos VM, Vieira JP (2018) Discrimination of habitat use between two sympatric species of mullets, Mugil curema and Mugil liza (Mugiliformes: Mugilidae) in the rio Tramandaí Estuary, determined by otolith chemistry. *Neotropical Ichthyology*, **16**, 1–8.

Mai ACG, Vieira JP (2013) Revisão e considerações sobre o uso do habitat, distribuição e história de vida de Lycengraulis grossidens (Agassiz, 1829) (Actinopterygii, Clupeiformes, Engraulididae). *Biota Neotropica*, **13**, 121–130.

Malinsky M, Trucchi E, Lawson DJ, Falush D (2018) RADpainter and fineRADstructure: Population Inference from RADseq Data. *Molecular Biology and Evolution*, **35**, 1284–1290.

Mark Shrimpton J (2012) *Seawater to Freshwater Transitions in Diadromous Fishes*. Elsevier Inc.

Marques DA, Lucek K, Sousa VC, Excoffier L, Seehausen O (2019) Admixture between old lineages facilitated contemporary ecological speciation in Lake Constance stickleback. *Nature Communications*, **10**, 4240.

Martin TJ (1990) Osmoregulation in three species of Ambassidae (Osteichthyes: Perciformes) from estuaries in Natal. *South African Journal of Zoology*, **25**, 229–234.

Martínez-Álvarez RM, Sanz A, García-Gallego M *et al.* (2005) Adaptive branchial mechanisms in the sturgeon Acipenser naccarii during acclimation to saltwater. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, **141**, 183–190.

Masel J, King OD, Maughan H (2007) The Loss of Adaptive Plasticity during Long Periods of Environmental Stasis. *The American Naturalist*, **169**, 38–46.

Matala AP, Allen B, Narum SR, Harvey E (2017) Restricted gene flow between resident Oncorhynchus mykiss and an admixed population of anadromous steelhead. *Ecology and Evolution*, **7**, 8349–8362.

Mateus CS, Almeida PR, Mesquita N, Quintella BR, Alves MJ (2016) European lampreys: New insights on postglacial colonization, gene flow and speciation. *PLoS ONE*, **11**, e0148107.

Mateus CS, Stange M, Berner D *et al.* (2013) Strong genome-wide divergence between sympatric European river and brook lampreys. *Current Biology*, **23**, R649–R650.

Mavarez J, Audet C, Bernatchez L (2009) Major disruption of gene expression in hybrids between young sympatric anadromous and resident populations of brook charr (Salvelinus fontinalis Mitchill). *Journal of Evolutionary Biology*, **22**, 1708–1720.

McBride RS, Holder JC (2008) A Review and Updated Assessment of Florida's

Anadromous Shads: American Shad and Hickory Shad. *North American Journal of Fisheries Management*, **28**, 1668–1686.

McBride RS, Matheson RE (2011) FLORIDA ' S DIADROMOUS FISHES : BIOLOGY , ECOLOGY , CONSERVATION , AND MANAGEMENT. *Florida Scientist*, **74**, 187–213.

McCairns RJS, Bernatchez L (2010) Adaptive divergence between freshwater and marine sticklebacks: insights into the role of phenotypic plasticity from an integrated analysis of candidate gene expression. *Evolution*, **64**, 1029–1047.

McDowall R (1972) The species problem in freshwater fishes and the taxonomy of diadromous and lacustrine populations of Galaxias maculatus (Jenyns). *Journal of the Royal Society of New Zealand*, **2**, 325–367.

McDowall R (1993) A recent marine ancestry for diadromous fishes? Sometimes yes, but mostly no! *Environmental Biology of Fishes*, **37**, 329–335.

McDowall R (1997) The evolution of diadromy in fishes (revisited) and its place in phylogenetic analysis. *Reviews in Fish Biology and Fisheries*, **7**, 443–462.

McDowall R (1998) Driven by diadromy: Its role in the historical and ecological biogeography of the New Zealand freshwater fish fauna. *Italian Journal of Zoology*, **65**, 73–85.

McDowall R (1999) Different kinds of diadromy: Different kinds of conservation problems. *ICES Journal of Marine Science*, **56**, 410–413.

McDowall RM (2000) Biogeography of the New Zealand torrentfish, Cheimarrichthys fosteri (Teleostei: Pinguipedidae): A distribution driven mostly by ecology and behaviour. *Environmental Biology of Fishes*, **58**, 119–131.

McDowall R (2001) Diadromy, diversity and divergence: implications for speciation processes in fishes. *Fish and Fisheries*, **2**, 278–285.

McDowall R (2003) Variation in vertebral number in galaxiid fishes (Teleostei: Galaxiidae): A legacy of life history, latitude and length. *Environmental Biology of Fishes*, **66**, 361–381.

McDowall R (2004) Ancestry and amphidromy in island freshwater fish faunas. *Fish and Fisheries*, **5**, 75–85.

McDowall R (2007) On amphidromy, a distinct form of diadromy in aquatic organisms. *Fish and Fisheries*, **8**, 1–13.

McDowall R (2008a) Diadromy, history and ecology: A question of scale. *Hydrobiologia*, **602**, 5–14.

McDowall R (2008b) Why are so many boreal freshwater fishes anadromous? Confronting "conventional wisdom." *Fish and Fisheries*, **9**, 208–213.

McDowall RM (2009) Early hatch: A strategy for safe downstream larval transport in amphidromous gobies. *Reviews in Fish Biology and Fisheries*, **19**, 1–8.

McDowall R (2010) Why be amphidromous : expatrial dispersal and the place of source and sink population dynamics ? *Reviews in Fish Biology and Fisheries*, **20**, 87–100.

McDowall R, Haro A, Smith KL *et al.* (2009) Making the best of two worlds: diadromy in the evolution, ecology, and conservation of aquatic organisms. In: *Challenges for diadromous fishes in a dynamic global environment. American Fisheries Society, Symposium*, pp. 1–22.

McDowall R, Mitchell C, Brothers E (1994) Age at migration from the sea of juveniles Galaxias in New Zealand (Pisces: Galaxiidae). *Bulletin of Marince Science*, **54**, 385–402.

McDowall R, Robertson DA, Saito R (1975) Occurrence of Galaxiid Larvae and Juveniles in the Sea. *Journal of Marine and Freshwater Research*, **9**, 1–9.

McEnroe M, Cech JJ (1985) Osmoregulation in juvenile and adult white sturgeon, Acipenser transmontanus. *Environmental Biology of Fishes*, **14**, 23–30.

McPhail J (1993) Ecology and evolution of sympatric sticklebacks (Gasterosteus): origin and evolution of the species pair. *Canadian Journal of Zoology*, **71**, 515–523.

Meier K, Hansen MM, Normandeau E *et al.* (2014) Local adaptation at the transcriptome level in brown trout: evidence from early life history temperature genomic reaction norms. *PloS one*, **9**, e85171.

Mennesson MI, Tabouret H, Pécheyran C, Feunteun E, Keith P (2015) Amphidromous life cycle of Eleotris fusca (Gobioidei: Eleotridae), a widespread species from the Indo-Pacific studied by otolith analyses. *Cybium*, **39**, 249–260.

Merlin C, Liedvogel M (2019) The genetics and epigenetics of animal migration and orientation: Birds, butterflies and beyond. *Journal of Experimental Biology*, **222**, jeb191890.

Metzger DCH, Schulte PM (2016) Epigenomics in marine fishes. *Marine Genomics*, **30**, 43–54.

Michel C, Hicks BJ, Stölting KN *et al.* (2008) Distinct migratory and non-migratory ecotypes of an endemic New Zealand eleotrid (Gobiomorphus cotidianus) - Implications for incipient speciation in island freshwater fish species. *BMC Evolutionary Biology*, **8**, 1–14.

Micheletti SJ, Hess JE, Zendt JS, Narum SR (2018a) Selection at a genomic region of major effect is responsible for evolution of complex life histories in anadromous steelhead. *BMC Evolutionary Biology*, **18**, 140.

Micheletti SJ, Matala AR, Matala AP, Narum SR (2018b) Landscape features along migratory routes influence adaptive genomic variation in anadromous steelhead (Oncorhynchus mykiss). *Molecular Ecology*, **27**, 128–145.

Miles NG, Butler GL, Diamond SL *et al.* (2018) Combining otolith chemistry and telemetry to assess diadromous migration in pinkeye mullet, Trachystoma petardi (Actinopterygii, Mugiliformes). *Hydrobiologia*, **808**, 265–281.

Miles NG, Walsh CT, Butler G, Ueda H, West RJ (2014) Australian diadromous fishes - Challenges and solutions for understanding migrations in the 21st century. *Marine and Freshwater Research*, **65**, 12–24.

Milton D (2009) *Living in Two Worlds: Diadromous Fishes, and Factors Affecting Populaiton Connectivity Between Tropical Rivers and Coasts* (I Nagelkerken, Ed,).

Milton DA, Chenery SR, Farmer MJ, Blaber SJM (1997) Identifying the spawning estuaries of the tropical shad, terubok Tenualosa toli, using otolith microchemistry. *Marine Ecology Progress Series*, **153**, 283–291.

Ministerio de Obras Públicas (2004) *Diagnóstico y clasificación de los cursos y cuerpos de agua según objetivos de calidad, cuenca del río Toltén.* Gobierno de Chile

Ministerio de Obras Públicas (2004b) *Diagnóstico y clasificación de los cursos y cuerpos de agua según objetivos de calidad, cuenca del Rio Valdivia* . Gobierno de Chile

Mitchell CP (1989) Laboratory culture of Galaxias maculatus and potential applications. *New Zealand Journal of Marine and Freshwater Research*, **23**, 325–336.

Miyazaki Y, Terui A (2016) Temporal dynamics of fluvial fish community caused by marine amphidromous species in the Shubuto River, southwestern Hokkaido, Japan. *Ichthyological Research*, **63**, 173–179.

Moore JS, Bourret V, Dionne M *et al.* (2014) Conservation genomics of anadromous Atlantic salmon across its North American range: Outlier loci identify the same patterns of population structure as neutral loci. *Molecular Ecology*, **23**, 5680–5697.

Moore JS, Harris LN, Le Luyer J *et al.* (2017) Genomics and telemetry suggest a role for migration harshness in determining overwintering habitat choice, but not gene flow, in anadromous Arctic Char. *Molecular Ecology*, **26**, 6784–6800.

Morán P, Marco-Rius F, Megías M, Covelo-Soto L, Pérez-Figueroa A (2013) Environmental induced methylation changes associated with seawater adaptation in brown trout. *Aquaculture*, **392**–**395**, 77–83.

Morash AJ, McClelland GB (2011) Regulation of carnitine palmitoyltransferase (CPT) I during Fasting in rainbow trout (Oncorhynchus mykiss) promotes increased mitochondrial fatty acid oxidation. *Physiological and Biochemical Zoology*, **84**, 625–633.

Morin R, Dodson JJ, Power G (1982) Life history of anadromous cisco (Coregonus artedii), lake whitefish (Coregonus clupeaformis), and round whitefish (Prosopium cylindraceum) populations of eastern James-Hudson bay. *Canadian Journal of Aquatic Sciences*, **39**, 958–967.

Morris R, Pickering AD (1976) Changes in the ultrastructure of the gills of the river lamprey, Lampetra fluviatilis (L.), during the anadromous spawning migration. *Cell And Tissue Research*, **173**, 271–277.

Morris MRJ, Richard R, Leder EH *et al.* (2014) Gene expression plasticity evolves in response to colonization of freshwater lakes in threespine stickleback. *Molecular Ecology*, **23**, 3226–3240.

Moyers BT, Rieseberg LH (2013) Divergence in gene expression is uncoupled from divergence in coding sequence in a secondarily woody sunflower. *International Journal of Plant Sciences*, **174**, 1079–1089.

Murase I, Iguchi K (2019) Facultative amphidromy involving estuaries in an annual amphidromous fish from a subtropical marginal range. *Journal of Fish Biology*, **95**, 1391–1398.

Murzina SA, Nefedova ZA, Pekkoeva SN *et al.* (2018) Fatty Acid Status of Freshwater Resident and Anadromous Forms of Young Brown Trout (Salmo trutta L.). *Contemporary Problems of Ecology*, **11**, 320–324.

Musarrat-ul-Ain, Farooq RY, Masood Z (2015) Gonadosomatic Index of Terapontid species. *International Journal of Biology and Biotechnology*, **12**, 575–578.

Myers GS (1949) Usage of Anadromous, Catadromous and Allied Terms for Migratory Fishes. *Copeia*, **1949**, 89–97.

Nakamura M, Jiang T, Xu G *et al.* (2021) Capacity for freshwater acclimation and differences in the transcription of ion transporter genes underlying different migratory life histories of Takifugu fish. *Gene*, **767**, 145285.

Narum SR, Genova A Di, Micheletti SJ, Maass A (2018) Genomic variation underlying complex life-history traits revealed by genome sequencing in Chinook salmon. *Proceedings of the Royal Society B: Biological Sciences*, **285**, 20180935.

Narum SR, Hess JE (2011) Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, **11**, 184–194.

Navarrete AH, Lagos NA, Ojeda FP (2014) Latitudinal diversity patterns of Chilean coastal fishes : searching for causal processes. *Revista Chilena de Historia Natural*, **87**.

Nazareno AG, Bemmels JB, Dick CW, Lohmann LG (2017) Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Molecular Ecology Resources*, **17**, 1136–1147.

Nelson JS (1994) *Fishes of the World (3rd ed.)*. Wiley, New York.

Nesbø CL, Magnhagen C, Jakobsen KS (1998) Genetic differentiation among stationary and anadromous perch (Perca fluviatilis) in the Baltic Sea. *Hereditas*, **129**, 241–249.

Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, **32**, 268–274.

Nichols KM, Kozfkay CC, Narum SR (2016) Genomic signatures among Oncorhynchus nerka ecotypes to inform conservation and management of endangered Sockeye Salmon. *Evolutionary Applications*, **9**, 1285–1300.

Nikolic N, Liu S, Jacobsen MW *et al.* (2020) Speciation history of European (Anguilla anguilla) and American eel (A. rostrata), analysed using genomic data. *Molecular Ecology*, **29**, 565–577.

Nordlie FG (2012) Life-history characteristics of eleotrid fishes of the western hemisphere, and perils of life in a vanishing environment. *Reviews in Fish Biology and Fisheries*, **22**, 189–224.

O'Brown NM, Summers BR, Jones FC, Brady SD, Kingsley DM (2015) A recurrent

regulatory change underlying altered expression and Wnt response of the stickleback armor plates gene EDA. *eLife*, **2015**, 1–17.

O'Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS (2018) These aren't the loci you'e looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology*, **27**, 3193–3206.

O'Malley KG, Vaux F, Black AN (2019) Characterizing neutral and adaptive genomic differentiation in a changing climate: The most northerly freshwater fish as a model. *Ecology and Evolution*, **9**, 2004–2017.

Oke KB, Bukhari M, Kaeuffer R *et al.* (2016) Does plasticity enhance or dampen phenotypic parallelism? A test with three lake-stream stickleback pairs. *Journal of Evolutionary Biology*, **29**, 126–143.

Orsini L, Andrew R, Eizaguirre C (2013) Evolutionary ecological genomics. *Molecular Ecology*, **22**, 527–531.

Oto Y (2019) Upstream migration without complete osmoregulatory switching in an amphidromous goby: Estimated by body condition changes in different salinity environment. *Journal of Experimental Marine Biology and Ecology*, **514**–**515**, 67–75.

Palavai V (2009) Diversity and Distribution of Freshwater Fishes of the Andaman.

Palkovacs EP, Dion KB, Post DM, Caccone A (2008) Independent evolutionary origins of landlocked alewife populations and rapid parallel evolution of phenotypic traits. *Molecular Ecology*, **17**, 582–597.

Papakostas S, Vasemägi A, Vähä JP *et al.* (2012) A proteomics approach reveals divergent molecular responses to salinity in populations of European whitefish (Coregonus lavaretus). *Molecular Ecology*, **21**, 3516–3530.

Parker KA, Hess JE, Narum SR, Kinziger AP (2019) Evidence for the genetic basis and epistatic interactions underlying ocean- and river-maturing ecotypes of Pacific Lamprey (Entosphenus tridentatus) returning to the Klamath River, California. *Molecular Ecology*, **28**, 3171–3185.

Patro R, Duggal G, Love M, Irizarry R, Kingsford C (2017) Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference. *Nat Methods*, **14**, 417–419.

Pavey SA, Gaudin J, Normandeau E *et al.* (2015) RAD Sequencing Highlights Polygenic Discrimination of Habitat Ecotypes in the Panmictic American Eel. *Current Biology*, **25**, 1666–1671.

Pearse DE, Miller MR, Abadía-Cardoso A, Garza JC, Abadı A (2014) Rapid parallel evolution of standing variation in a single, complex, genomic region is associated with life history in steelhead/rainbow trout. *Proceedings of the Royal Society B Biological Sciences*, **281**, 565–577.

Perrier C, Bourret V, Kent MP, Bernatchez L (2013) Parallel and nonparallel genome-wide divergence among replicate population pairs of freshwater and anadromous Atlantic salmon. *Molecular Ecology*, **22**, 5577–5593.

Peterson BJ, Fry B (1987) Stable isotopes in ecosystem studies. *Annual review of ecology and systematics*, **18**, 293–320.

Pollard DA (1971) The biology of a landlocked form of the normally catadromous salmoniform fish Galaxias maculatus.pdf. *Aust. J. mar. Freshwat. Res.*, **22**, 91–123.

Potts WTW (1994) Kinetics of sodium uptake in freshwater animals: A comparison of ion- exchange and proton pump hypotheses. *American Journal of Physiology - Regulatory Integrative and Comparative Physiology*, **266**, 315–320.

Prince DJ, O'Rourke SM, Thompson TQ *et al.* (2017) The evolutionary basis of premature migration in Pacific salmon highlights the utility of genomics for informing conservation. *Science Advances*, **3**, e1603198.

Pujolar JM, Jacobsen MW, Als TD *et al.* (2014) Assessing patterns of hybridization between North Atlantic eels using diagnostic single-nucleotide polymorphisms. *Heredity*, **112**, 627–637.

Purcell S, Neale B, Todd-Brown K *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics*, **81**, 559–75.

Puritz JB, Hollenbeck CM, Gold JR (2014) dDocent : a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ*, **2**, e431.

R Core Team (2020) R: A language and environment for statistical computing. *Foundation for Statistical Computing, Vienna, Austria*.

Raeymaekers JAM, Konijnendijk N, Larmuseau MHD *et al.* (2014) A gene with major phenotypic effects as a target for selection vs. homogenizing gene flow. *Molecular Ecology*, **23**, 162–181.

Rahman MM, Rahman M, Parvez S, Mallik N, Nabi MR (2015) Population dynamics of mudskipper *Periophthalmus novemradiatus* from Bakkhali river estuary , Cox ' s Bazar, Bangladesh. *Agricultural Science Research Journal*, **5**, 118–123.

Rastorguev SM, Nedoluzhko A V., Gruzdeva NM *et al.* (2017) Differential miRNA expression in the three-spined stickleback, response to environmental changes. *Scientific Reports*, **7**, 18089.

Ravinet M, Hynes R, Poole R *et al.* (2015) Where the lake meets the sea: Strong reproductive isolation is associated with adaptive divergence between lake resident and anadromous three-spined sticklebacks. *PLoS ONE*, **10**, 1–23.

Reid K, Carlos Garza J, Gephard SR *et al.* (2020) Restoration-mediated secondary contact leads to introgression of alewife ecotypes separated by a colonial-era dam. *Evolutionary Applications*, **13**, 652–664.

Reid K, Palkovacs EP, Hasselman DJ *et al.* (2018) Comprehensive evaluation of genetic population structure for anadromous river herring with single nucleotide polymorphism data. *Fisheries Research*, **206**, 247–258.

Renaut S, Nolte AW, Rogers SM, Derome N, Bernatchez L (2011) SNP signatures of selection on standing genetic variation and their association with adaptive

phenotypes along gradients of ecological speciation in lake whitefish species pairs (Coregonus spp.). *Molecular Ecology*, **20**, 545–559.

Rennison DJ, Stuart YE, Bolnick DI, Peirch CL (2019) Identification of ecological and morphological traits associated with regions of repeated genomic differentiation between lake and stream stickleback. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **374**, 20180241.

Reusch TBH, Wegner KM, Kalbe M (2001) Rapid genetic divergence in postglacial populations of threespine stickleback (Gasterosteus aculeatus): The role of habitat type, drainage and geographical proximity. *Molecular Ecology*, **10**, 2435–2445.

Riede K (2004) Global register of migratory species - from global to regional scales. *Final Report of the R&D-Projekt 808 05 081. Federal Agency for Nature Conservation*, 329 pp.

Righton D, Aarestrup K, Jellyman D *et al.* (2012) The Anguilla spp. migration problem: 40 million years of evolution and two millennia of speculation. *Journal of Fish Biology*, **81**, 365–386.

Rivas MJ, Saura M, Pérez-Figueroa A *et al.* (2018) Population genomics of parallel evolution in gene expression and gene sequence during ecological adaptation. *Scientific Reports*, **8**.

Rodríguez A, Gallardo MA, Gisbert E *et al.* (2002) Osmoregulation in juvenile Siberian sturgeon (Acipenser baerrii). *Fish Physiology and Biochemistry*, **26**, 345–354.

Rohland N, Reich D (2011) Cost-effective, high-throughput DNA sequencing. *Genome Research*, 939–946.

Rohtla M, Vetemaa M, Taal I *et al.* (2014) Life history of anadromous burbot (Lota lota, Linneaus) in the brackish Baltic Sea inferred from otolith microchemistry. *Ecology of Freshwater Fish*, **23**, 141–148.

Rojo JH, Figueroa DE, Boy CC (2018) Age and growth of diadromous Galaxias maculatus (Jenyns, 1842) in southernmost South America (54° S) including contribution of age classes to reproduction. *Environmental Biology of Fishes*, **101**, 1149–1160.

Roman I, Bourgeois Y, Reyes-Velasco J *et al.* (2018) Contrasted patterns of divergence and gene flow among five fish species in a Mongolian rift lake following glaciation. *Biological Journal of the Linnean Society*, **125**, 115–125.

Rougemont Q, Gagnaire PA, Perrier C *et al.* (2017) Inferring the demographic history underlying parallel genomic divergence among pairs of parasitic and nonparasitic lamprey ecotypes. *Molecular Ecology*, **26**, 142–162.

Ruegg K, Anderson EC, Boone J, Pouls J, Smith TB (2014) A role for migration-linked genes and genomic islands in divergence of a songbird. *Molecular Ecology*, **23**, 4757–4769.

Ruiz-Jarabo I, González-Wevar CA, Oyarzún R *et al.* (2016) Isolation Driven Divergence in Osmoregulation in Galaxias maculatus (Jenyns, 1848) (Actinopterygii: Osmeriformes). *Plos One*, **11**, e0154766.

Sakai H, Goto A, Jeon S-R (2002) Speciation and Dispersal of Tribolodon Species (Pisces, Cyprinidae) around the Sea of Japan. *Zoological Science*, **19**, 1291–1303.

Sakai H, Imai C (2005) Otolith Sr:Ca ratios of the freshwater and anadromous cyprinid genus Tribolodon. *Ichthyological Research*, **52**, 182–184.

Sakar UK, Naskar M, Roy K *et al.* (2018) Baseline information of reproduction parameters of an amphidromous croaker Johnius coitor (Hamilton, 1822) from Ganga river basin, India with special reference to potential influence of climatic variability. *Aquat. Living Resour.*, **31**.

Salem M, Paneru B, Al-Tobasei R *et al.* (2015) Transcriptome assembly, gene annotation and tissue gene expression atlas of the rainbow trout. *PLoS ONE*, **10**, e0121778.

Salisbury SJ, Booker C, McCracken GR *et al.* (2018) Genetic divergence among and within arctic char (Salvelinus alpinus) populations inhabiting landlocked and sea-accessible sites in Labrador, Canada. *Canadian Journal of Fisheries and Aquatic Sciences*, **75**, 1256–1269.

Salisbury SJ, McCracken GR, Keefe D, Perry R, Ruzzante DE (2019) Extensive secondary contact among three glacial lineages of Arctic Char (Salvelinus alpinus) in Labrador and Newfoundland. *Ecology and Evolution*, **9**, 2031–2045.

Salisbury SJ, Mccracken GR, Perry R *et al.* (2020) Limited genetic parallelism underlies recent, repeated incipient speciation in geographically proximate populations of an Arctic fish (Salvelinus alpinus). *Molecular Ecology*, **29**, 4280–4294.

Samani NK, Esa Y, Amin SMN, Ikhsan NFM (2016) Phylogenetics and population genetics of Plotosus canius (Siluriformes: Plotosidae) from Malaysian coastal waters. *PeerJ*, **2016**.

Schaffler JJ, Young SP, Herrington S, Ingram T, Tannehill J (2015) Otolith Chemistry to Determine Within-River Origins of Alabama Shad in the Apalachicola–Chattahoochee–Flint River Basin. *Transactions of the American Fisheries Society*, **144**, 1–10.

Schliewen UK (2012) Diversity and Distribution of Marine, Euryhaline and Amphidromous Gobies from Western, Central and Southern Africa. In: *The Biology of Gobies* (eds Patzner RA, Van Tassell JL, Kovacic M, Kapoor B), pp. 207–229.

Schmidt DJ, Crook DA, Macdonald JI *et al.* (2014) Migration history and stock structure of two putatively diadromous teleost fishes, as determined by genetic and otolith chemistry analyses. *Freshwater Science*, **33**.

Schneider K, Adams CE, Elmer KR (2019) Parallel selection on ecologically relevant gene functions in the transcriptomes of highly diversifying salmonids. *BMC Genomics*, **20**, 1010.

Scott C (2016) dammit: an open and accessible de novo transcriptome annotator.

Scott G, Richards J, Forbush B, Isenring P, Schulte P (2004) Changes in gene expression in gills of the euryhaline killifish Fundulus heteroclitus after abrupt salinity transfer. *Am J Physiol Cell Physiol*, **287**, C300–C309.

Secor DH, Henderson-Arzapalo A, Piccoli PM (1995) Can otolith microchemistry chart patterns of migration and habitat utilization in anadromous fishes? *Journal of Experimental Marine Biology and Ecology*, **192**, 15–33.

Seehausen O, Wagner CE (2014) Speciation in freshwater fishes. *Annual Review of Ecology, Evolution, and Systematics*, **45**, 621–651.

Sekino M, Nakamichi R, Iwasaki Y *et al.* (2016) A new resource of single nucleotide polymorphisms in the Japanese eel Anguilla japonica derived from restriction site-associated DNA. *Ichthyological Research*, **63**, 496–504.

Seppey M, Manni M, Zdobnov E (2019) BUSCO: Assessing Genome Assembly and Annotation Completeness. *Methods Mol Biol.*, **1962**, 227–245.

Shen KN, Lee YC, Tzeng WN (1998) Use of otolith microchemistry to investigate the life history pattern of gobies in a Taiwanese stream. *Zoological Studies*, **37**, 322–329.

Shiao JC, Chen CY, Zhang J, Iizuka Y (2016) Habitat use and migratory life history of Salangid icefish (salangidae) revealed by otolith Sr/Ca ratios. *Zoological Studies*, **55**, 1–10.

Shiao JC, Tzeng CS, Li PC, Bell KNI (2015) Upstream migration and marine early life history of amphidromous gobies inferred from otolith increments and microchemistry. *Environmental Biology of Fishes*, **98**, 933–950.

Skovrind M, Olsen MT, Vieira FG *et al.* (2016) Genomic population structure of freshwater-resident and anadromous ide (Leuciscus idus) in north-western Europe. *Ecology and Evolution*, **6**, 1064–1074.

Smith-Unna R, Boursnell C, Patro R, Hibberd J, Kelly S (2016) TransRate: reference free quality assessment of de-novo transcriptome assemblies. *Genome Research*.

Smith WE, Kwak TJ (2014) Otolith microchemistry of tropical diadromous fishes: Spatial and migratory dynamics. *Journal of Fish Biology*, **84**, 913–928.

Smith G, Smith C, Kenny JG, Chaudhuri RR, Ritchie MG (2015) Genome-wide DNA methylation patterns in wild samples of two morphotypes of threespine stickleback (Gasterosteus aculeatus). *Molecular Biology and Evolution*, **32**, 888–895.

Stearns ASC (1989) Trade-Offs in Life-History. *Functional Ecology*, **3**, 259–268.

Steury RA, Currey MC, Cresko WA, Bohannan BJM (2019) Population genetic divergence and environment influence the gut microbiome in Oregon threespine stickleback. *Genes*, **10**, 484.

Strub PT, James C, Montecino V, Rutllant JA, Blanco JL (2019) Ocean circulation along the southern Chile transition region (38°–46°S): Mean, seasonal and interannual variability, with a focus on 2014–2016. *Progress in Oceanography*, **172**, 159–198.

Strydom NA (2003) Occurrence of larval and early juvenile fishes in the surf zone adjacent to two intermittently open estuaries, South Africa. *Environmental Biology of Fishes*, **66**, 349–359.

Stuart YE, Veen T, Weber JN *et al.* (2017) Contrasting effects of environment and

genetics generate a continuum of parallel evolution. *Nature Ecology and Evolution*, **1**, 0158.

Sulak KJ, Randall M (2002) Understanding sturgeon life history: Enigmas, myths, and insights from scientific studies. *Journal of Applied Ichthyology*, **18**, 519–528.

Sutherland BJG, Gosselin T, Normandeau E *et al.* (2016) Salmonid chromosome evolution as revealed by a novel method for comparing radseq linkagemaps. *Genome Biology and Evolution*, **8**, 3600–3617.

Sutherland BJG, Koczka KW, Yasuike M *et al.* (2014) Comparative transcriptomics of Atlantic Salmo salar, chum Oncorhynchus keta and pink salmon O. gorbuscha during infections with salmon lice Lepeophtheirus salmonis. *BMC Genomics*, **15**, 200.

Tabouret H, Tomadin M, Taillebois L *et al.* (2014) Amphidromy and marine larval phase of ancestral gobioids Rhyacichthys guilberti and Protogobius attiti (Teleostei: Rhyacichthyidae). *Marine and Freshwater Research*, **65**, 776–783.

Taillebois L, Castelin M, Ovenden JR, Bonillo C, Keith P (2013) Contrasting Genetic Structure among Populations of Two Amphidromous Fish Species (Sicydiinae) in the Central West Pacific. *PLoS ONE*, **8**.

Taillebois L, Tabouret H, Pecheyran C, Keith P (2015) Inputs from microchemistry to the understanding of three Sicydiinae species' life cycle. *Vie et Milieu*, **65**, 73–84.

Taugbøl A, Arntsen T, Østbye K, Vøllestad LA (2014) Small changes in gene expression of targeted osmoregulatory genes when exposing marine and freshwater threespine stickleback (Gasterosteus aculeatus) to abrupt salinity transfers. *PLoS ONE*, **9**, e106894.

Teichert N, Valade P, Lim P *et al.* (2014) Habitat selection in amphidromous Gobiidae of Reunion Island: Sicyopterus lagocephalus (Pallas, 1770) and Cotylopus acutipinnis (Guichenot, 1863). *Environmental Biology of Fishes*, **97**, 255–266.

Terekhanova N V, Barmintseva AE, Kondrashov AS, Bazykin GA, Mugue NS (2019) Architecture of Parallel Adaptation in Ten Lacustrine Threespine Stickleback Populations from the White Sea Area. *Genome Biology and Evolution*, **11**, 2605–2618.

Thacker C (2009) Phylogeny of Gobioidei and Placement within Acanthomorpha, with a New Classification and Investigation of Diversification and Character Evolution. *Copeia*, **2009**, 93–104.

Therneau T, Lumley T, Atkinson E, Crowson C (2020) Survival analysis.

Thompson NF, Anderson EC, Clemento AJ *et al.* (2020) A complex phenotype in salmon controlled by a simple change in migratory timing. *Science*, **370**, 609–613.

Thrower FP, Hard JJ, Joyce JE (2004) Genetic architecture of growth and early life-history transitions in anadromous and derived freshwater populations of steelhead. *Journal of Fish Biology*, **65**, 286–307.

Todd E V., Black MA, Gemmell NJ (2016) The power and promise of RNA-seq in

ecology and evolution. *Molecular Ecology*, **25**, 1224–1241.

Toyama M, Hironaka M, Yamahama Y *et al.* (2008) Presence of rhodopsin and porphyropsin in the eyes of 164 fishes, representing marine, diadromous, coastal and freshwater species - a qualitative and comparative study. *Photochemistry and Photobiology*, **84**, 996–1002.

Trancart T, Lambert P, Rochard E *et al.* (2012) Alternative flood tide transport tactics in catadromous species: Anguilla anguilla, Liza ramada and Platichthys flesus. *Estuarine, Coastal and Shelf Science*, **99**, 191–198.

Trautner JH, Reiser S, Blancke T, Unger K, Wysujack K (2017) Metamorphosis and transition between developmental stages in European eel (Anguilla anguilla, L.) involve epigenetic changes in DNA methylation patterns. *Comparative Biochemistry and Physiology - Part D: Genomics and Proteomics*, **22**, 139–145.

Trevisan dos Santos J. (2016) Revisão taxonômica e anatomia de Awaous tajasica (Lichtenstein, 1822) (Gobiiformes: Gobiidae).

Tsukamoto K, Arai T (2001) Facultative catadromy of the eel Anguilla japonica between freshwater and seawater habitats. *Marine Ecology Progress Series*, **220**, 265–276.

Tsukamoto K, Miller MJ, Kotake A, Aoyama J, Uchida K (2009) The Origin of Fish Migration: The Random Escapement Hypothesis. *Challenges for Diadromous Fishes in a Dynamic Global Environment*, **69**, 45–61.

Tsukamoto K, Watanabe S, Kuroki M, Aoyama J, Miller MJ (2014) Freshwater habitat use by a moray eel species, Gymnothorax polyuranodon, in Fiji shown by otolith microchemistry. *Environmental Biology of Fishes*, **97**, 1377–1385.

Tsunagawa T, Arai T (2008) Flexible migration of Japanese freshwater gobies Rhinogobius spp. as revealed by otolith Sr:Ca ratios. *Journal of Fish Biology*, **73**, 2421–2433.

Tsunagawa T, Suzuki T, Arai T (2010) Migratory history of Rhinogobius sp. OR morphotype “Shimahire” as revealed by otolith Sr:Ca ratios. *Ichthyological Research*, **57**, 10–15.

Tulkani R (2017) Population Biology of Two Species of Grey Mullet, Liza abu in Central Iraq (Heckel, 1843) and Chelon labrosus (Risso, 1827) in North West Wales.

Urban MC (2015) Accelerating extinction risk from climate change. *Science*, **348**, 571–573.

Urbina MA, Glover CN (2015) Effect of salinity on osmoregulation, metabolism and nitrogen excretion in the amphidromous fish, inanga (Galaxias maculatus). *Journal of Experimental Marine Biology and Ecology*, **473**, 7–15.

Urbina MA, Schulte PM, Bystriansky JS, Glover CN (2013) Differential expression of Na(+), K (+)-ATPase α-1 isoforms during seawater acclimation in the amphidromous galaxiid fish Galaxias maculatus. *Journal of comparative physiology. B, Biochemical, systemic, and environmental physiology*, **183**, 345–357.

Varadharajan S, Rastas P, Löytynoja A *et al.* (2019) A High-Quality Assembly of the

Nine-Spined Stickleback (Pungitius pungitius) Genome. *Genome biology and evolution*, **11**, 3291–3308.

Veale AJ, Russello MA (2017) Genomic changes associated with reproductive and migratory ecotypes in sockeye salmon (Oncorhynchus nerka). *Genome Biology and Evolution*, **9**, 2921–2939.

Vega GC, Wiens JJ (2012) Why are there so few fish in the sea? *Proceedings of the Royal Society B: Biological Sciences*, **279**, 2323–2329.

Velotta JP, McCormick SD, O'Neill RJ, Schultz ET (2014) Relaxed selection causes microevolution of seawater osmoregulation and gene expression in landlocked Alewives. *Oecologia*, **175**, 1081–1092.

Velotta JP, Wegrzyn JL, Ginzburg S *et al.* (2017) Transcriptomic imprints of adaptation to fresh water: parallel evolution of osmoregulatory gene expression in the Alewife. *Molecular Ecology*, **26**, 831–848.

Vera-Escalona I, Delgado ML, Habit E, Ruzzante D (2020) Historical and Contemporary Diversity of Galaxiids in South America: Biogeographic and Phylogenetic Perspectives. *Diversity*, **12**, 304.

Vera-Escalona I, Habit E, Ruzzante DE (2015) Echoes of a distant time: effects of historical processes on contemporary genetic patterns in Galaxias platei in Patagonia. *Molecular Ecology*, **24**, 4112–4128.

Vera-Escalona I, Habit E, Ruzzante DE (2019) Invasive species and postglacial colonization : their effects on the genetic diversity of a Patagonian fish. *Proceedings of the Royal Society B Biological Sciences*, **286**, 20182567.

Verta J, Jones FC, Miescher F, Society P (2019) Predominance of cis -regulatory changes in parallel expression divergence of sticklebacks. , 1–30.

Villamarín F, Magnusson WE, Jardine TD *et al.* (2016) Temporal uncoupling between energy acquisition and allocation to reproduction in a herbivorous-detritivorous fish. *PLoS ONE*, **11**, 1–17.

de Villemereuil P, Gaggiotti OE, Mouterde M, Till-Bottraud I (2015) Common garden experiments in the genomic era: new perspectives and opportunities. *Heredity*, **116**, 249–254.

Waldman J, Grunwald C, Wirgin I (2008) Sea lamprey Petromyzon marinus: An exception to the rule of homing in anadromous fishes. *Biology Letters*, **4**, 659–662.

Walter RP, Hogan JD, Blum MJ *et al.* (2012) Climate change and conservation of endemic amphidromous fishes in Hawaiian streams. *Endangered Species Research*, **16**, 261–272.

Walters AW, Barnes RT, Post DM (2009) Anadromous alewives (Alosa pseudoharengus) contribute marine-derived nutrients to coastal stream food webs. *Canadian Journal of Fisheries and Aquatic Sciences*, **66**, 439–448.

Wang Y, Guo R, Li H *et al.* (2011) The complete mitochondrial genome of the Sichuan taimen (Hucho bleekeri): Repetitive sequences in the control region and

phylogenetic implications for Salmonidae. *Marine Genomics*, **4**, 221–228.

Wang Y, Zhao Y, Wang Y *et al.* (2020) Population transcriptomics reveals weak parallel genetic basis in repeated marine and freshwater divergence in nine-spined sticklebacks. *Molecular Ecology*.

Warburton ML, Jarvis MG, Closs GP (2018) Otolith microchemistry indicates regional phylopatry in the larval phase of an amphidromous fish (Gobiomorphus hubbsi). *New Zealand Journal of Marine and Freshwater Research*, **52**, 398–408.

Warren IA, Ciborowski KL, Casadei E *et al.* (2014) Extensive local gene duplication and functional divergence among paralogs in Atlantic Salmon. *Genome Biology and Evolution*, **6**, 1790–1805.

Waters JM, Burridge CP (1999) Extreme Intraspecific Mitochondrial DNA Sequence Divergence in Galaxias maculatus (Osteichthys : Galaxiidae), One of the World 's Most Widespread Freshwater Fish. *Molecular Phylogenetics and Evolution*, **11**, 1–12.

Waters JM, Dijkstra LH, Wallis GP (2000a) Biogeography of a southern hemisphere freshwater fish : how important is marine dispersal? *Molecular Ecology*, 1815–1821.

Waters JM, Emerson BC, Arribas P, McCulloch GA (2020) Dispersal Reduction: Causes, Genomic Mechanisms, and Evolutionary Consequences. *Trends in Ecology and Evolution*, **35**, 512–522.

Waters JM, López J a, Wallis GP (2000b) Molecular phylogenetics and biogeography of galaxiid fishes (Osteichthyes: Galaxiidae): dispersal, vicariance, and the position of Lepidogalaxias salamandroides. *Systematic biology*, **49**, 777–795.

Waters JM, Mcculloch GA (2021) Reinventing the wheel? Reassessing the roles of gene flow, sorting and convergence in repeated evolution. *Molecular Ecology*, 0–3.

Waters JM, Wallis GP (2001) Cladogenesis and loss of the marine life-history phase in freshwater Galaxiid fishes (Osmeriformes: Galaxiidae). *Evolution*, **55**, 587–597.

Weitere M, Tautz D, Neumann D, Steinfartz S (2004) Adaptive divergence vs. environmental plasticity: Tracing local genetic adaptation of metamorphosis traits in salamanders. *Molecular Ecology*, **13**, 1665–1677.

Wellenreuther M, Bernatchez L (2018) Eco-Evolutionary Genomics of Chromosomal Inversions. *Trends in Ecology and Evolution*, **33**, 427–440.

Whitehead A (2010) The evolutionary radiation of diverse osmotolerant physiologies in killifish (Fundulus sp.). *Evolution*, **64**, 2070–2085.

Whitfield AK (2005) Preliminary documentation and assessment of fish diversity in sub-Saharan African estuaries. *African Journal of Marine Science*, **27**, 307–324.

Whittamore JM (2012) Osmoregulation and epithelial water transport: Lessons from the intestine of marine teleost fish. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, **182**, 1–39.

Wickham H (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag.

Willoughby JR, Harder AM, Tennessen JA, Scribner KT, Christie MR (2018) Rapid

genetic adaptation to a novel environment despite a genome-wide reduction in genetic diversity. *Molecular Ecology*, **27**, 4041–4051.

Wilson JPF (1984) A Review of the Biology of the Pollan Coregonus autumnalis pollan Thompson : An Endemic Irish Salmonid Subspecies. *Proceedings of the Royal Irish Academy. Section B: Biological, Geological and Chemical Science*, **84B**, 123–127.

Wood CC, Foote CJ (1996) Evidence for sympatric genetic divergence of anadromous and nonanadromous morphs of sockeye salmon (Oncorhynchus nerka). *Evolution*, **50**, 1265–1279.

Wright B, Farquharson KA, McLennan EA *et al.* (2019) From reference genomes to population genomics: Comparing three reference-aligned reduced-representation sequencing pipelines in two wildlife species. *BMC Genomics*, **20**, 453.

Xia A, Zhong L, Chen X *et al.* (2015) Complete mitochondrial genome of spined sleeper Eleotris oxycephala (Perciformes, Eleotridae) and phylogenetic consideration. *Biochemical Systematics and Ecology*, **62**, 11–19.

Xu DP, Jiang SL, Zhao CS, Fang DA, Hu HY (2018) Comparative transcriptomics analysis of the river pufferfish (Takifugu obscurus) by tributyltin exposure: Clues for revealing its toxic injury mechanism. *Fish and Shellfish Immunology*, **82**, 536–543.

Yamaguchi M, Katayama S, Omori M (2004) Migration pattern of shirauo Salangichthys microdon Bleeker, in the Ishikari River system and adjacent nearshore sea area, Japan, as estimated by otolith microchemistry analysis. *Fisheries Science*, **70**, 546–552.

Yamanoue Y, Miya M, Doi H *et al.* (2011) Multiple invasions into freshwater by pufferfishes (Teleostei: Tetraodontidae): A mitogenomic perspective. *PLoS ONE*, **6**, e17410.

Yáñez JM, Naswa S, López ME *et al.* (2016) Genomewide single nucleotide polymorphism discovery in Atlantic salmon (Salmo salar): validation in wild and farmed American and European populations. *Molecular ecology resources*, **16**, 1002–1011.

Yang Z, Chen YF (2008) Differences in reproductive strategies between obscure puffer Takifugu obscurus and ocellated puffer Takifugu ocellatus during their spawning migration. *Journal of Applied Ichthyology*, **24**, 569–573.

Yatsuyanagi T, Ishida R, Sakata MK *et al.* (2020) Environmental DNA monitoring for short-term reproductive migration of endemic anadromous species, Shishamo smelt (Spirinchus lanceolatus) . *Environmental DNA*, **2**, 130–139.

Yoshida K, Miyagi R, Mori S *et al.* (2016) Whole-genome sequencing reveals small genomic regions of introgression in an introduced crater lake population of threespine stickleback. *Ecology and Evolution*, **6**, 2190–2204.

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**, 67.

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.

Zhang N, Song N, Han Z, Gao T (2016) The complete mitochondrial genome of Coilia nasus (Clupeiformes: Engraulidae) from the coast of Ningbo in China. *Mitochondrial DNA. Part A, DNA mapping, sequencing, and analysis*, **27**, 1660–1661.

Zhang J, Wei F, Li M, Xu M (2007) Threatened fishes of the world: Hemisalanx prognathus (Regan 1908) (Salangidae). *Environmental Biology of Fishes*, **78**, 209–210.

Zhang X, Wen H, Wang H *et al.* (2017) RNA-Seq analysis of salinity stress-responsive transcriptome in the liver of spotted sea bass (Lateolabrax maculatus). *PLoS ONE*, **12**, e0173238.

Zhou YF, Duan JR, Liu K *et al.* (2015) Testes transcriptome profiles of the anadromous fish Coilia nasus during the onset of spermatogenesis. *Marine Genomics*, **24**, 241–243.

Zhu K, Lü Z, Liu L, Gong L, Liu B (2018) The complete mitochondrial genome of Trachidermus fasciatus (Scorpaeniformes: Cottidae) and phylogenetic studies of Cottidae. *Mitochondrial DNA Part B: Resources*, **3**, 301–302.

Zhuang P, Kynard B, Zhang L, Zhang T, Cao W (2002) Ontogenetic behavior and migration of Chinese sturgeon, Acipenser sinensis. *Environmental Biology of Fishes*, **65**, 83–97.

Zimmerman CE, Edwards GW, Perry K (2009) Maternal Origin and Migratory History of Steelhead and Rainbow Trout Captured in Rivers of the Central Valley, California. *Transactions of the American Fisheries Society*, **138**, 280–291.

Zink RM (2011) The evolution of avian migration. *Biological Journal of the Linnean Society*, **104**, 237–250.

Appendix 1. List of known diadromous species. Related to Table 2.1.

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Acipenser baerii	Acipenseridae	anadromous		(Rodríguez *et al.* 2002)	yes
Acipenser brevirostrum	Acipenseridae	anadromous		(Kynard 1997)	
Acipenser dabryanus	Acipenseridae	anadromous		(Kynard *et al.* 2003)	
Acipenser gueldenstaedtii	Acipenseridae	semi-anadromous		(Arai & Miyazaki 2001)	yes
Acipenser medirostris	Acipenseridae	anadromous		(Allen *et al.* 2009)	
Acipenser mikadoi	Acipenseridae	anadromous		(Koshelev *et al.* 2012)	
Acipenser naccarii	Acipenseridae	semi-anadromous		(Martínez-Álvarez *et al.* 2005)	
Acipenser nudiventris	Acipenseridae	anadromous		(Acolas & Lambert 2016)	
Acipenser oxyrinchus	Acipenseridae	anadromous		(Allen *et al.* 2014)	
Acipenser persicus	Acipenseridae	anadromous		(Acolas & Lambert 2016)	
Acipenser schrenckii	Acipenseridae	anadromous		(Koshelev *et al.* 2014)	
Acipenser sinensis	Acipenseridae	anadromous		(Zhuang *et al.* 2002)	
Acipenser stellatus	Acipenseridae	anadromous		(Honţ *et al.* 2019)	
Acipenser sturio	Acipenseridae	anadromous		(Acolas *et al.* 2012)	
Acipenser transmontanus	Acipenseridae	anadromous		(McEnroe & Cech 1985)	
Huso dauricus	Acipenseridae	anadromous		(Koshelev *et al.* 2014)	
Huso huso	Acipenseridae	anadromous		(Honţ *et al.* 2019)	
Scaphirhynchus suttkusi	Acipenseridae	anadromous		(Acolas & Lambert 2016)	
Arius madagascariensis	Ariidae	anadromous	X	(Milton 2009; Acolas & Lambert 2016)	
Genidens barbus	Ariidae	anadromous		(Avigliano *et al.* 2017)	yes
Neoarius graeffei	Ariidae	anadromous	X	(Milton 2009)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Atherinella chagresi	Atherinopsidae	anadromous	X	(Milton 2009)	
Atherinella guatemalensis	Atherinopsidae	anadromous	X	(Milton 2009)	
Mystus gulio	Bagridae	anadromous	X	(Bijoy Nandan *et al.* 2012; Acolas & Lambert 2016)	
Citharinus citharus	Citharinidae	anadromous	X	(Riede 2004)	
Citharinus eburneensis	Citharinidae	anadromous	X	(Acolas & Lambert 2016)	
Clarotes laticeps	Claroteidae	anadromous	X	(Acolas & Lambert 2016)	
Alosa aestivalis	Clupeidae	anadromous		(Limburg 2001)	yes
Alosa alabamae	Clupeidae	anadromous		(Schaffler *et al.* 2015)	
Alosa alosa	Clupeidae	anadromous		(Baglinière *et al.* 2003)	yes
Alosa fallax	Clupeidae	anadromous		(Aprahamian *et al.* 2003)	
Alosa immaculata	Clupeidae	anadromous		(Acolas & Lambert 2016)	
Alosa kessleri	Clupeidae	anadromous		(Kuzishchin *et al.* 2020)	
Alosa mediocris	Clupeidae	anadromous		(McBride & Holder 2008)	
Alosa pseudoharengus	Clupeidae	anadromous		(Walters *et al.* 2009)	
Alosa sapidissima	Clupeidae	anadromous		(McBride & Holder 2008)	
Alosa tanaica	Clupeidae	anadromous		(Acolas & Lambert 2016)	
Alosa volgensis	Clupeidae	anadromous		(Acolas & Lambert 2016)	
Anodontostoma chacunda	Clupeidae	anadromous	X	(Milton 2009)	
Anodontostoma thailandiae	Clupeidae	anadromous	X	(Milton 2009; Acolas & Lambert 2016)	
Clupanodon thrissa	Clupeidae	anadromous	X	(Riede 2004)	
Clupeonella cultriventris	Clupeidae	anadromous		(Bloom & Lovejoy 2014)	
Dorosoma cepedianum	Clupeidae	anadromous	X	(Acolas & Lambert 2016)	
Dorosoma petenense	Clupeidae	anadromous	X	(Acolas & Lambert 2016)	
Herklotsichthys gotoi	Clupeidae	anadromous	X	(Milton 2009)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Hilsa kelee	Clupeidae	anadromous	X	(Milton 2009)	
Nematalosa galatheae	Clupeidae	anadromous	X	(Milton 2009; Acolas & Lambert 2016)	
Nematalosa nasus	Clupeidae	anadromous	X	(Milton 2009; Acolas & Lambert 2016)	
Pellonula leonensis	Clupeidae	anadromous	X	(Milton 2009)	
Pellonula vorax	Clupeidae	anadromous	X	(Milton 2009; Acolas & Lambert 2016)	
Tenualosa ilisha	Clupeidae	anadromous		(Arai *et al.* 2019)	yes
Tenualosa reevesii	Clupeidae	anadromous		(Blaber *et al.* 2003)	
Tenualosa toli	Clupeidae	anadromous		(Milton *et al.* 1997)	
Leuciscus idus	Cyprinidae	semi-anadromous		(Skovrind *et al.* 2016)	yes
Pelecus cultratus	Cyprinidae	anadromous		(Acolas & Lambert 2016)	
Rutilus frisii	Cyprinidae	anadromous		(Kohestan-Eskandari *et al.* 2014)	
Tribolodon brandtii	Cyprinidae	anadromous		(Sakai & Imai 2005)	
Tribolodon hakonensis	Cyprinidae	anadromous		(Sakai *et al.* 2002)	yes
Vimba vimba	Cyprinidae	anadromous		(Łuszczek-Trojnar *et al.* 2008)	
Elops hawaiensis	Elopidae	anadromous	X	(Milton 2009)	
Anchoviella lepidentostole	Engraulidae	anadromous		(Milton 2009)	
Colia ectenes	Engraulidae	anadromous		(Duan *et al.* 2012)	
Coilia nasus	Engraulidae	anadromous		(Dou *et al.* 2012)	yes
Lycengraulis grossidens	Engraulidae	anadromous		(Mai & Vieira 2013)	
Stolephorus commersonnii	Engraulidae	anadromous	X	(Bijoy Nandan *et al.* 2012)	
Microgadus tomcod	Gadidae	anadromous		(Couillard *et al.* 2011)	
Lovettia sealii	Galaxiidae	semi-anadromous		(Schmidt *et al.* 2014)	
Gasterosteus aculeatus	Gasterosteidae	anadromous		(Arai *et al.* 2003)	yes

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Pungitius pungitius	Gasterosteidae	anadromous		(Arai & Goto 2008)	yes
Geotria australis	Geotriidae	anadromous		(Miles *et al.* 2014)	
Leucopsarion petersii	Gobiidae	anadromous		(Kokita & Nohara 2011)	
Luciogobius guttatus	Gobiidae	anadromous		(Riede 2004; Miyazaki & Terui 2016)	
Lota lota	Lotidae	anadromous	X	(Rohtla *et al.* 2014)	yes
Mordacia lapicida	Mordaciidae	anadromous	X	(McDowall 1999)	
Mordacia mordax	Mordaciidae	anadromous		(Miles *et al.* 2014)	
Morone americana	Moronidae	anadromous		(Acolas & Lambert 2016)	
Morone saxatilis	Moronidae	anadromous		(Secor *et al.* 1995)	yes
Rhinomugil corsula	Mugilidae	anadromous	X	(Acolas & Lambert 2016)	
Pisodonophis boro	Ophichthidae	anadromous	X	(Acolas & Lambert 2016)	
Hypomesus japonicus	Osmeridae	anadromous		(Dodson *et al.* 2009)	
Hypomesus nipponensis	Osmeridae	anadromous		(Katayama *et al.* 2000)	yes
Hypomesus olidus	Osmeridae	anadromous		(Acolas & Lambert 2016)	yes
Hypomesus transpacificus	Osmeridae	anadromous		(Acolas & Lambert 2016)	
Osmerus dentex	Osmeridae	anadromous		(Dodson *et al.* 2009)	
Osmerus eperlanus	Osmeridae	anadromous		(Lyle & Maitland 1997)	
Osmerus mordax	Osmeridae	anadromous		(Bradbury *et al.* 2008)	yes
Spirinchus lanceolatus	Osmeridae	anadromous		(Yatsuyanagi *et al.* 2020)	
Spirinchus thaleichthys	Osmeridae	anadromous		(Acolas & Lambert 2016)	yes
Thaleichthys pacificus	Osmeridae	anadromous		(Clarke *et al.* 2007)	
Pangasius krempfi	Pangasiidae	anadromous		(Hogan *et al.* 2007)	
Perca fluviatilis	Percidae	semi-anadromous		(Nesbø *et al.* 1998)	yes
Caspiomyzon wagneri	Petromyzontidae	anadromous		(Mark Shrimpton 2012)	
Entosphenus tridentatus	Petromyzontidae	anadromous		(Clemens *et al.* 2013)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Lampetra ayresii	Petromyzontidae	anadromous		(Acolas & Lambert 2016)	
Lampetra fluviatilis	Petromyzontidae	anadromous		(Morris & Pickering 1976)	
Lampetra tridentata	Petromyzontidae	anadromous		(Beamish & Levings 1991)	
Lethenteron camtschaticum	Petromyzontidae	anadromous		(Acolas & Lambert 2016)	yes
Lethenteron reissneri	Petromyzontidae	anadromous		(Acolas & Lambert 2016)	
Petromyzon marinus	Petromyzontidae	anadromous		(Waldman *et al.* 2008)	
Ilisha filigera	Pristigasteridae	anadromous	X	(Milton 2009)	
Ilisha megaloptera	Pristigasteridae	anadromous	X	(Milton 2009)	
Ilisha sirishai	Pristigasteridae	anadromous	X	(Milton 2009)	
Pellona ditchela	Pristigasteridae	anadromous	X	(Milton 2009)	
Retropinna tasmanica	Retropinnidae	anadromous	X	(Miles *et al.* 2014)	
Hemisalanx prognathus	Salangidae	anadromous		(Zhang *et al.* 2007)	
Neosalanx jordani	Salangidae	anadromous		(Dodson *et al.* 2009)	
Neosalanx reganius	Salangidae	anadromous		(Acolas & Lambert 2016)	
Salangichthys microdon	Salangidae	anadromous		(Yamaguchi *et al.* 2004)	yes
Salanx ariakensis	Salangidae	anadromous		(Shiao *et al.* 2016)	yes
Salanx cuvieri	Salangidae	anadromous		(Riede 2004; Dodson *et al.* 2009)	
Brachymystax lenok	Salmonidae	anadromous	X	(Riede 2004)	
Coregonus albula	Salmonidae	anadromous		(Acolas & Lambert 2016)	
Coregonus artedii	Salmonidae	anadromous		(Morin *et al.* 1982)	
Coregonus autumnalis	Salmonidae	anadromous		(Wilson 1984)	
Coregonus clupeaformis	Salmonidae	anadromous		(Morin *et al.* 1982)	
Coregonus huntsmani	Salmonidae	anadromous		(Acolas & Lambert 2016)	
Coregonus laurettae	Salmonidae	anadromous		(Brown *et al.* 2008)	
Coregonus lavaretus	Salmonidae	anadromous		(Lehtonen *et al.* 1992)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Coregonus muksun	Salmonidae	anadromous		(Acolas & Lambert 2016)	
Coregonus nasus	Salmonidae	anadromous		(Brown *et al.* 2008)	yes
Coregonus oxyrinchus	Salmonidae	anadromous		(Borcherding *et al.* 2014)	
Coregonus pallasii	Salmonidae	anadromous		(Acolas & Lambert 2016)	
Coregonus peled	Salmonidae	anadromous		(Acolas & Lambert 2016)	
Coregonus pidschian	Salmonidae	anadromous		(Acolas & Lambert 2016)	
Coregonus sardinella	Salmonidae	anadromous		(Brown *et al.* 2008)	
Hucho perryi	Salmonidae	anadromous		(Edo *et al.* 2005)	
Oncorhynchus clarkii	Salmonidae	anadromous		(Dodson *et al.* 2013)	yes
Oncorhynchus gorbuscha	Salmonidae	anadromous		(Gallagher *et al.* 2013)	
Oncorhynchus keta	Salmonidae	anadromous		(Wood & Foote 1996)	yes
Oncorhynchus kisutch	Salmonidae	anadromous		(Dodson *et al.* 2013)	
Oncorhynchus masou	Salmonidae	anadromous		(Dodson *et al.* 2013)	yes
Oncorhynchus mykiss	Salmonidae	anadromous		(Dodson *et al.* 2013)	yes
Oncorhynchus nerka	Salmonidae	anadromous		(Dodson *et al.* 2013)	yes
Oncorhynchus tshawytscha	Salmonidae	anadromous		(Dodson *et al.* 2013)	
Salmo labrax	Salmonidae	anadromous		(Acolas & Lambert 2016)	
Salmo marmoratus	Salmonidae	anadromous		(Acolas & Lambert 2016)	
Salmo salar	Salmonidae	anadromous		(Dodson *et al.* 2013)	yes
Salmo trutta	Salmonidae	anadromous		(Dodson *et al.* 2013)	yes
Salvelinus alpinus	Salmonidae	anadromous		(Dodson *et al.* 2013)	yes
Salvelinus confluentus	Salmonidae	anadromous		(Dodson *et al.* 2013; Austin *et al.* 2019)	yes
Salvelinus fontinalis	Salmonidae	anadromous		(Dodson *et al.* 2013)	yes
Salvelinus leucomaenis	Salmonidae	anadromous		(Dodson *et al.* 2013)	
Salvelinus malma	Salmonidae	anadromous		(Dodson *et al.* 2013)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Salvelinus namaycush	Salmonidae	semi-anadromous		(Kissinger *et al.* 2016)	yes
Stenodus leucichthys	Salmonidae	anadromous		(Brown *et al.* 2008)	
Takifugu obscurus	Tetraodontidae	anadromous		(Jeong *et al.* 2014)	yes
Takifugu ocellatus	Tetraodontidae	anadromous		(Yang & Chen 2008)	
Ambassis interrupta	Ambassidae	catadromous	X	(Milton 2009)	
Anguilla anguilla	Anguillidae	catadromous		(Arai *et al.* 2006a)	yes
Anguilla australis	Anguillidae	catadromous		(Miles *et al.* 2014)	
Anguilla bengalensis	Anguillidae	catadromous		(Milton 2009; Bijoy Nandan *et al.* 2012)	
Anguilla bicolor	Anguillidae	catadromous		(Arai & Chino 2019)	yes
Anguilla celebesensis	Anguillidae	catadromous		(Milton 2009)	
Anguilla dieffenbachii	Anguillidae	catadromous		(Arai *et al.* 2003)	
Anguilla interioris	Anguillidae	catadromous		(Arai & Chino 2012)	
Anguilla japonica	Anguillidae	catadromous		(Tsukamoto & Arai 2001)	yes
Anguilla malgumora	Anguillidae	catadromous		(Arai & Chino 2012)	
Anguilla marmorata	Anguillidae	catadromous		(Arai *et al.* 2013)	yes
Anguilla megastoma	Anguillidae	catadromous		(Arai & Chino 2012)	
Anguilla mossambica	Anguillidae	catadromous		(Whitfield 2005)	
Anguilla nebulosa	Anguillidae	catadromous		(Arai & Chino 2012)	
Anguilla obscura	Anguillidae	catadromous		(Miles *et al.* 2014)	
Anguilla rheinhardtii	Anguillidae	catadromous		(Miles *et al.* 2014)	
Anguilla rostrata	Anguillidae	catadromous		(Jessop *et al.* 2007)	yes
Centropomus undecimalis	Centropomidae	catadromous		(Lowerre-Barbieri *et al.* 2014)	
Centropomus pectinatus	Centropomidae	catadromous	X	(Milton 2009)	
Ethmalosa fimbriata	Clupeidae	catadromous		(Bloom & Lovejoy 2014)	
Potamalosa richmondia	Clupeidae	catadromous		(Miles *et al.* 2014)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Cottus kazika	Cottidae	catadromous		(Augspurger *et al.* 2017)	
Trachidermus fasciatus	Cottidae	catadromous	X	(Milton 2009)	
Eleotris annobonensis	Eleotridae	catadromous	X	(Milton 2009)	
Eleotris balia	Eleotridae	catadromous	X	(Milton 2009)	
Eleotris pisonis	Eleotridae	catadromous	X	(Milton 2009)	
Eleotris senegalensis	Eleotridae	catadromous	X	(Milton 2009)	
Eleotris vittata	Eleotridae	catadromous	X	(Milton 2009)	
Thryssa scratchleyi	Engraulidae	catadromous		(Miles *et al.* 2014)	
Kuhlia marginata	Kuhliidae	catadromous		(Feutry *et al.* 2013)	
Kuhlia malo	Kuhliidae	catadromous		(Feutry *et al.* 2013)	
Kuhlia rupestris	Kuhliidae	catadromous		(Augspurger *et al.* 2017)	
Khulia salelea	Kuhliidae	catadromous		(Feutry *et al.* 2013)	
Khulia sauvagii	Kuhliidae	catadromous		(Feutry *et al.* 2013)	
Lateolabrax japonicus	Lateolabracidae	catadromous		(Fuji *et al.* 2018)	
Lates calcarifer	Latidae	catadromous		(Miles *et al.* 2014)	yes
Lutjanus goldiei	Lutjanidae	catadromous	X	(Milton 2009)	
Lutjanus maxweberi	Lutjanidae	catadromous	X	(Milton 2009)	
Megalops cyprinoides	Megalopidae	catadromous	X	(Miles *et al.* 2014)	
Agonostomus monticola	Mugilidae	catadromous	X	(Tulkani 2017)	
Agonostomus telfairii	Mugilidae	catadromous	X	(Milton 2009)	
Aldrichetta forsteri	Mugilidae	catadromous		(Chang & Iizuka 2012)	
Chelon labrosus	Mugilidae	catadromous		(Gordoa 2009)	
Crenimugil heterocheilos	Mugilidae	catadromous	X	(Milton 2009)	
Ellochelon vaigiensis	Mugilidae	catadromous	X	(Milton 2009)	
Joturus pichardi	Mugilidae	catadromous	X	(Tulkani 2017)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Liza alata	Mugilidae	catadromous		(Villamarín *et al.* 2016)	
Liza aurata	Mugilidae	catadromous		(Como *et al.* 2018)	
Liza falcipinnis	Mugilidae	catadromous		(Milton 2009)	
Liza grandisquamis	Mugilidae	catadromous		(Milton 2009)	
Liza haematocheila	Mugilidae	catadromous		(Chang & Iizuka 2012)	
Liza macrolepsis	Mugilidae	semi-catadromous		(Bijoy Nandan *et al.* 2012; Chang & Iizuka 2012)	
Liza parsia	Mugilidae	catadromous		(Bijoy Nandan *et al.* 2012)	
Liza ramada	Mugilidae	catadromous		(Filipe *et al.* 2009)	
Liza richardsonii	Mugilidae	catadromous		(Chang & Iizuka 2012)	
Liza rumadu	Mugilidae	catadromous		(Almeida 1996)	
Liza subviridis	Mugilidae	catadromous		(Chang & Iizuka 2012)	
Mugil cephalus	Mugilidae	catadromous		(Bijoy Nandan *et al.* 2012)	yes
Mugil curema	Mugilidae	catadromous		(Albieri *et al.* 2010)	
Mugil liza	Mugilidae	catadromous		(Garbin *et al.* 2014)	
Mugil soiuy	Mugilidae	catadromous		(McDowall 1997)	
Mugil trichodon	Mugilidae	catadromous		(Mai *et al.* 2018)	
Myxus capensis	Mugilidae	catadromous		(Strydom 2003)	
Trachystoma petardi	Mugilidae	catadromous		(Miles *et al.* 2018)	yes
Valamugil cunnesius	Mugilidae	catadromous		(Bijoy Nandan *et al.* 2012)	
Valamugil speigleri	Mugilidae	catadromous		(Bijoy Nandan *et al.* 2012)	
Gymnothorax polyuranodon	Muraenidae	catadromous		(Tsukamoto *et al.* 2014)	
Macquaria novemaculeata	Percichthyidae	catadromous		(Chenoweth & Hughes 1997)	
Platichthys flesus	Pleuronectidae	catadromous		(Trancart *et al.* 2012)	
Rhombosolea retiaria	Pleuronectidae	catadromous		(McDowall 2000)	
Pseudaphritis urvillii	Pseudaphritidae	catadromous - female		(Crook *et al.* 2010)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Terapon jarbua	Terapontidae	catadromous	X	(Riede 2004; Musarrat-ul-Ain *et al.* 2015)	
Notesthes robusta	Tetrarogidae	catadromous		(Milton 2009)	
Ambassis miops	Ambassidae	amphidromous	X	(Milton 2009; Miles *et al.* 2014)	
Ambassis gymnocephalus	Ambassidae	amphidromous	X	(Milton 2009)	
Ambassis kopsii	Ambassidae	amphidromous	X	(Milton 2009)	
Ameiurus melas	Ariidae	amphidromous	X	(Milton 2009)	
Arius jella	Ariidae	amphidromous	X	(Milton 2009)	
Cephalocassia jatia	Ariidae	amphidromous	X	(Milton 2009)	
Cochlefelis burmanica	Ariidae	amphidromous	X	(Milton 2009)	
Hemiarius sona	Ariidae	amphidromous	X	(Milton 2009)	
Hexanematichthys sagor	Ariidae	amphidromous	X	(Milton 2009)	
Nemapteryx caelata	Ariidae	amphidromous	X	(Milton 2009)	
Netuma thalassina	Ariidae	amphidromous	X	(Milton 2009)	
Plicofollis platystomus	Ariidae	amphidromous	X	(Milton 2009)	
Plicofollis tenuispinis	Ariidae	amphidromous		(Milton 2009; Hashemi *et al.* 2013)	
Atherina boyeri	Atherinidae	amphidromous	X	(Filipe *et al.* 2009)	
Carangaoides malabaricus	Carangidae	amphidromous	X	(Bijoy Nandan *et al.* 2012)	
Caranx sexfasciatus	Carangidae	amphidromous	X	(Bijoy Nandan *et al.* 2012)	
Centropomus ensiferus	Centropomidae	amphidromous		(Milton 2009; McBride & Matheson 2011)	
Centropomus medius	Centropomidae	amphidromous	X	(Milton 2009)	
Centropomus nigrescens	Centropomidae	amphidromous	X	(Milton 2009)	
Centropomus parallelus	Centropomidae	amphidromous		(Milton 2009; McBride & Matheson 2011)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Centropomus pectinatus	Centropomidae	amphidromous		(McBride & Matheson 2011)	
Centropomus robalito	Centropomidae	amphidromous	X	(Milton 2009)	
Centropomus undecimalis	Centropomidae	amphidromous	X	(Milton 2009)	
Cheimarrichthys fosteri	Cheimarrichthyidae	amphidromous		(McDowall 2000; Augspurger *et al.* 2017)	
Clupea harangus	Clupeidae	amphidromous		(Augspurger *et al.* 2017)	
Sardinella melanura	Clupeidae	amphidromous		(Milton 2009; Elahi *et al.* 2017)	
Sprattus sprattus	Clupeidae	amphidromous		(Augspurger *et al.* 2017)	
Cottus aleuticus	Cottidae	amphidromous		(Augspurger *et al.* 2017)	
Cottus amblystomopsis	Cottidae	amphidromous		(Augspurger *et al.* 2017)	
Cottus asper	Cottidae	amphidromous		(Augspurger *et al.* 2017)	
Cottus hangiongensis	Cottidae	amphidromous		(Miyazaki & Terui 2016; Augspurger *et al.* 2017)	
Cottus pollux	Cottidae	amphidromous		(Goto & Arai 2003; Augspurger *et al.* 2017)	yes
Leptocottus armatus	Cottidae	amphidromous		(McDowall 1997)	
Bostrychus africanus	Eleotridae	amphidromous	X	(Milton 2009)	
Bostrychus sinensis	Eleotridae	amphidromous	X	(Milton 2009)	
Bunaka gyrinoides	Eleotridae	amphidromous	X	(Milton 2009; Miles *et al.* 2014)	
Bunaka pinguis	Eleotridae	amphidromous		(Donaldson & Myers 2002; Milton 2009)	
Butis amboinensis	Eleotridae	amphidromous		(Donaldson & Myers 2002; Milton 2009)	
Butis butis	Eleotridae	amphidromous		(Milton 2009; Bijoy Nandan *et al.* 2012)	
Butis humeralis	Eleotridae	amphidromous	X	(Milton 2009)	
Butis koilomatodon	Eleotridae	amphidromous	X	(Milton 2009)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Butis melanostigma	Eleotridae	amphidromous	X	(Milton 2009)	
Dormitator latifrons	Eleotridae	amphidromous		(McDowall 2009; Augspurger *et al.* 2017)	
Dormitator maculatus	Eleotridae	amphidromous		(Milton 2009; Augspurger *et al.* 2017)	
Eleotris acanthopoma	Eleotridae	amphidromous		(Shen *et al.* 1998; Milton 2009)	
Eleotris amblyopsis	Eleotridae	amphidromous		(Nordlie 2012; Augspurger *et al.* 2017)	
Eleotris fusca	Eleotridae	amphidromous		(Bijoy Nandan *et al.* 2012; Mennesson *et al.* 2015)	
Eleotris melanosoma	Eleotridae	amphidromous		(Maeda & Tachihara 2005)	
Eleotris oxycephala	Eleotridae	amphidromous	X	(Xia *et al.* 2015)	
Eleotris perniger	Eleotridae	amphidromous		(Frotté *et al.* 2019)	
Eleotris picta	Eleotridae	amphidromous		(Augspurger *et al.* 2017)	
Eleotris sandwicensis	Eleotridae	amphidromous		(Heim-Ballew *et al.* 2020)	yes
Giuris margaritacea	Eleotridae	amphidromous	X	(Miles *et al.* 2014)	
Gobiomorphus australis	Eleotridae	amphidromous	X	(Miles *et al.* 2014)	
Gobiomorphus cotidianus	Eleotridae	amphidromous		(Augspurger *et al.* 2017)	yes
Gobiomorphus gobioides	Eleotridae	amphidromous		(Augspurger *et al.* 2017)	
Gobiomorphus hubbsi	Eleotridae	amphidromous		(Augspurger *et al.* 2017; Jarvis *et al.* 2018)	
Gobiomorphus huttoni	Eleotridae	amphidromous		(Augspurger *et al.* 2017)	
Gobiomorus dormitor	Eleotridae	amphidromous		(Smith & Kwak 2014; Augspurger *et al.* 2017)	
Gobiomorus maculatus	Eleotridae	amphidromous		(Augspurger *et al.* 2017)	
Guavina guavina	Eleotridae	amphidromous		(Milton 2009; Augspurger *et al.* 2017)	
Hypseleotris cyprinoides	Eleotridae	amphidromous		(Donaldson & Myers 2002)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Hypseleotris guentheri	Eleotridae	amphidromous		(Donaldson & Myers 2002)	
Ophieleotris aporos	Eleotridae	amphidromous		(Donaldson & Myers 2002; Milton 2009)	
Ophiocara porocephala	Eleotridae	amphidromous		(Donaldson & Myers 2002; Milton 2009)	
Thryssa dussumieri	Engraulidae	amphidromous	X	(Milton 2009; Bijoy Nandan *et al.* 2012)	
Thryssa gautamiensis	Engraulidae	amphidromous	X	(Milton 2009)	
Thryssa hamaltonii	Engraulidae	amphidromous	X	(Milton 2009)	
Thryssa kammalensoides	Engraulidae	amphidromous	X	(Milton 2009)	
Thryssa malabarica	Engraulidae	amphidromous	X	(Bijoy Nandan *et al.* 2012)	
Aplochiton taeniatus	Galaxiidae	amphidromous		(Augspurger *et al.* 2017; Alò *et al.* 2019)	
Aplochiton marinus	Galaxiidae	amphidromous		(Alò *et al.* 2019)	
Aplochiton zebra	Galaxiidae	amphidromous		(Augspurger *et al.* 2017)	
Galaxias argenteus	Galaxiidae	amphidromous		(Augspurger *et al.* 2017)	
Galaxias brevipinnis	Galaxiidae	amphidromous		(Augspurger *et al.* 2017)	
Galaxias fasciatus	Galaxiidae	amphidromous		(Augspurger *et al.* 2017)	
Galaxias maculatus	Galaxiidae	amphidromous		(Hickford & Schiel 2016; Augspurger *et al.* 2017)	yes
Galaxias postvectis	Galaxiidae	amphidromous		(Franklin & Gee 2019)	
Galaxias truttaceus	Galaxiidae	amphidromous		(Augspurger *et al.* 2017)	yes
Neochanna cleaveri	Galaxiidae	amphidromous		(McDowall 2004; Miles *et al.* 2014)	
Eucinostomus melanopterus	Gerreidae	amphidromous	X	(Milton 2009)	
Gerres cinereus	Gerreidae	amphidromous	X	(Milton 2009)	
Gerres erythrourus	Gerreidae	amphidromous	X	(Bijoy Nandan *et al.* 2012)	
Gerres filamentosus	Gerreidae	amphidromous	X	(Milton 2009)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Gerres limbatus	Gerreidae	amphidromous	X	(Milton 2009)	
Gerres longirostris	Gerreidae	amphidromous	X	(Milton 2009)	
Gerres seifer	Gerreidae	amphidromous	X	(Milton 2009; Bijoy Nandan *et al.* 2012)	
Gobiesox cephalus	Gobiesocidae	semi-amphidromous		(Frotté *et al.* 2019)	yes
Acantrogobius caninus	Gobiidae	amphidromous		(Palavai 2009)	
Acanthogobius lactipes	Gobiidae	amphidromous		(Miyazaki & Terui 2016)	
Awaous acritosus	Gobiidae	amphidromous		(Augspurger *et al.* 2017)	yes
Awaous banana	Gobiidae	amphidromous		(Smith & Kwak 2014; Augspurger *et al.* 2017)	yes
Awaous bustamantei	Gobiidae	amphidromous	X	(Schliewen 2012)	
Awaous grammepomus	Gobiidae	amphidromous	X	(Milton 2009)	
Awaous guamensis	Gobiidae	amphidromous		(Augspurger *et al.* 2017)	
Awaous lateristriga	Gobiidae	amphidromous	X	(Schliewen 2012)	
Awaous melanocephalus	Gobiidae	amphidromous		(Shen *et al.* 1998; Shiao *et al.* 2015)	
Awaous ocellaris	Gobiidae	amphidromous	X	(Milton 2009)	
Awaous stamineus	Gobiidae	amphidromous		(Augspurger *et al.* 2017; Hogan *et al.* 2017)	yes
Awaous tajasica	Gobiidae	amphidromous		(Trevisan dos Santos 2016)	
Awaous transandeanus	Gobiidae	amphidromous	X	(Lyons & Schneider 1990)	
Cotylopus acutipinnis	Gobiidae	amphidromous		(Milton 2009; Teichert *et al.* 2014)	
Glossogobius aureus	Gobiidae	amphidromous		(Miles *et al.* 2014; Shiao *et al.* 2015)	
Glossogobius celebius	Gobiidae	amphidromous		(Shen *et al.* 1998; Milton 2009)	
Glossogobius giuris	Gobiidae	amphidromous		(Milton 2009; Miles *et al.* 2014)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Gobioides broussonnetii	Gobiidae	amphidromous	X	(Milton 2009)	
Gobioides sagitta	Gobiidae	amphidromous	X	(Milton 2009)	
Gobionellus occidentalis	Gobiidae	amphidromous	X	(Milton 2009)	
Gobionellus oceanicus	Gobiidae	amphidromous	X	(Milton 2009)	
Gobionellus thoropsis	Gobiidae	amphidromous	X	(Milton 2009)	
Gymnogobius petschiliensis	Gobiidae	amphidromous		(Oto 2019)	
Gymnogobius opperiens	Gobiidae	amphidromous		(Miyazaki & Terui 2016)	
Gymnogobius urotaenia	Gobiidae	amphidromous		(Miyazaki & Terui 2016)	
Lentipes armatus	Gobiidae	amphidromous	X	(Milton 2009)	
Lentipes concolor	Gobiidae	amphidromous		(Augspurger *et al.* 2017; Heim-Ballew *et al.* 2020)	yes
Lentipes whittenorum	Gobiidae	amphidromous	X	(Milton 2009)	
Oligolepis acutipennis	Gobiidae	amphidromous		(Shen *et al.* 1998)	
Parasicydium bandama	Gobiidae	amphidromous	X	(Schliewen 2012)	
Periophthalmus argentilineatus	Gobiidae	amphidromous	X	(Milton 2009)	
Periophthalmus barbarus	Gobiidae	amphidromous	X	(Milton 2009)	
Periophthalmus malaccensis	Gobiidae	amphidromous	X	(Milton 2009)	
Periophthalmus modestus	Gobiidae	amphidromous	X	(Milton 2009)	
Periophthalmus novemradiatus	Gobiidae	amphidromous		(Milton 2009; Rahman *et al.* 2015)	
Periophthalmus weberi	Gobiidae	amphidromous	X	(Milton 2009)	
Periophthalmodon schlosseri	Gobiidae	amphidromous	X	(Milton 2009)	
Periophthalmodon septemradiatus	Gobiidae	amphidromous	X	(Milton 2009)	
Porogobius schlegelii	Gobiidae	amphidromous	X	(Milton 2009)	
Pseudapocryptes elongatus	Gobiidae	amphidromous	X	(Milton 2009)	
Pseudogobius javanicus	Gobiidae	amphidromous	X	(Milton 2009)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Pseudogobius melanostictus	Gobiidae	amphidromous	X	(Milton 2009)	
Pseudogobius poicilosoma	Gobiidae	amphidromous	X	(Milton 2009)	
Redigobius balteatus	Gobiidae	amphidromous		(Donaldson & Myers 2002; Milton 2009)	
Redigobius bikolanus	Gobiidae	amphidromous		(Shen *et al.* 1998)	
Redigobius dispar	Gobiidae	amphidromous	X	(Milton 2009)	
Redigobius horiae	Gobiidae	amphidromous		(Donaldson & Myers 2002)	
Redigobius macrostoma	Gobiidae	amphidromous	X	(Milton 2009)	
Redigobius roemeri	Gobiidae	amphidromous		(Donaldson & Myers 2002; Milton 2009)	
Redigobius sapangus	Gobiidae	amphidromous		(Donaldson & Myers 2002; Milton 2009)	
Rhinogobius brunneus	Gobiidae	amphidromous		(Iguchi & Mizuno 1999; Augspurger *et al.* 2017)	
Rhinogobius giurinus	Gobiidae	amphidromous		(Shiao *et al.* 2015)	
Rhinogobius similis	Gobiidae	amphidromous		(Iida *et al.* 2017)	
Rhinogobius sp.	Gobiidae	amphidromous		(Tsunagawa & Arai 2008; Augspurger *et al.* 2017)	yes
Schismatogobius sp.	Gobiidae	amphidromous		(Keith 2003)	
Schismatogobius roxasi	Gobiidae	amphidromous	X	(Milton 2009)	
Sicydium brevifile	Gobiidae	amphidromous	X	(Schliewen 2012)	
Sicydium bustamantei	Gobiidae	amphidromous	X	(Schliewen 2012)	
Sicydium crenilabrum	Gobiidae	amphidromous	X	(Schliewen 2012)	
Sicydium multipunctatum	Gobiidae	amphidromous	X	(González-Murcia & Álvarez 2018)	
Sicydium plumieri	Gobiidae	amphidromous		(Milton 2009; Frotté *et al.* 2019)	
Sicydium punctatum	Gobiidae	amphidromous		(Bell *et al.* 1995; Augspurger *et al.* 2017)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Sicydium salvini	Gobiidae	amphidromous		(Lyons & Schneider 1990)	
Sicyopterus aiensis	Gobiidae	amphidromous		(Augspurger *et al.* 2017)	
Sicyopterus cynocephalus	Gobiidae	amphidromous	X	(Ebner *et al.* 2017)	
Sicyopterus fuliag	Gobiidae	amphidromous	X	(Milton 2009)	
Sicyopterus griseus	Gobiidae	amphidromous	X	(Milton 2009)	
Sicyopterus japonicus	Gobiidae	amphidromous		(Shen *et al.* 1998; Augspurger *et al.* 2017)	
Sicyopterus lacrymosus	Gobiidae	amphidromous	X	(Milton 2009)	
Sicyopterus lagocephalus	Gobiidae	amphidromous		(Augspurger *et al.* 2017)	
Sicyopterus macrostetholepis	Gobiidae	amphidromous	X	(Milton 2009)	
Sicyopterus micrurus	Gobiidae	amphidromous	X	(Milton 2009)	
Sicyopterus rapa	Gobiidae	amphidromous	X	(Milton 2009)	
Sicyopterus sarasini	Gobiidae	amphidromous		(Augspurger *et al.* 2017)	
Sicyopterus stimpsoni	Gobiidae	amphidromous		(Heim-Ballew *et al.* 2020)	
Sicyopus auxilimentus	Gobiidae	amphidromous	X	(Milton 2009)	
Sicyopus jonklaasi	Gobiidae	amphidromous	X	(Milton 2009)	
Sicyopus leprurus	Gobiidae	amphidromous	X	(Milton 2009)	
Sicyopus zosterophorum	Gobiidae	amphidromous		(Taillebois *et al.* 2015; Augspurger *et al.* 2017)	
Smilosicyopus chloe	Gobiidae	amphidromous		(Taillebois *et al.* 2015)	
Stenogobius blokzeyli	Gobiidae	amphidromous	X	(Milton 2009)	
Stenogobius fasciatus	Gobiidae	amphidromous	X	(McBride & Matheson 2011)	
Stenogobius fehlmanni	Gobiidae	amphidromous		(Donaldson & Myers 2002)	
Stenogobius genivittatus	Gobiidae	amphidromous		(Shen *et al.* 1998; Shiao *et al.* 2015)	
Stenogobius gramnepomus	Gobiidae	amphidromous		(Palavai 2009)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Stenogobius hawaiiensis	Gobiidae	amphidromous		(Heim-Ballew *et al.* 2020)	yes
Stiphodon alcedo	Gobiidae	amphidromous		(Maeda *et al.* 2011)	
Stiphodon aureorostrum	Gobiidae	amphidromous	X	(Milton 2009)	
Stiphodon caeruleus	Gobiidae	amphidromous		(Chabarria *et al.* 2014)	
Stiphodon elegans	Gobiidae	amphidromous		(Milton 2009; Shiao *et al.* 2015)	
Stiphodon larson	Gobiidae	amphidromous		(McDowall 2010)	
Stiphodon niraikanaiensis	Gobiidae	amphidromous		(Maeda 2014)	
Stiphodon percnopterygionus	Gobiidae	amphidromous		(McDowall 2009; Iida *et al.* 2017)	
Stiphodon rutilaureus	Gobiidae	amphidromous	X	(Ebner & Thuesen 2011)	
Stiphodon semoni	Gobiidae	amphidromous		(Keith 2003)	
Stiphodon surrufus	Gobiidae	amphidromous	X	(Milton 2009)	
Taenoides cirratus	Gobiidae	amphidromous		(Bijoy Nandan *et al.* 2012)	
Taenoides buchanani	Gobiidae	amphidromous		(Bijoy Nandan *et al.* 2012)	
Tridentiger brevispinis	Gobiidae	amphidromous		(Miyazaki & Terui 2016)M	
Tridentiger kuroiwae	Gobiidae	amphidromous		(Iida *et al.* 2017)	
Zappa confluentus	Gobiidae	amphidromous	X	(Milton 2009)	
Pomadasys maculatus	Haemulidae	amphidromous	X	(Riede 2004; Ahmed & Bat 2016)	
Kuhlia caudavittata	Kuhliidae	amphidromous		(Augspurger *et al.* 2017)	
Kuhlia mugil	Kuhliidae	amphidromous		(Augspurger *et al.* 2017)	
Kuhlia petiti	Kuhliidae	amphidromous		(Augspurger *et al.* 2017)	
Kuhlia sandvicensis	Kuhliidae	amphidromous		(Benson & Michael Fitzsimons 2002; Milton 2009)	
Kuhlia xenura	Kuhliidae	amphidromous		(Augspurger *et al.* 2017)	
Agonostomus monticola	Mugilidae	amphidromous		(Frotté *et al.* 2019)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Liza macrolepis	Mugilidae	amphidromous	X	(Milton 2009)	
Liza melinoptera	Mugilidae	amphidromous	X	(Milton 2009)	
Liza subviridis	Mugilidae	amphidromous	X	(Milton 2009)	
Liza vaigiensis	Mugilidae	amphidromous	X	(Milton 2009)	
Valamugil buchanani	Mugilidae	amphidromous	X	(Milton 2009)	
Plecoglossus altivelis	Plecoglossidae	amphidromous		(Arai 2006; Murase & Iguchi 2019)	
Plotosus canius	Plotosidae	amphidromous	X	(Samani *et al.* 2016)	
Ilisha kampeni	Pristigasteridae	amphidromous	X	(Milton 2009)	
Ilisha melastoma	Pristigasteridae	amphidromous	X	(Milton 2009)	
Ilisha novacula	Pristigasteridae	amphidromous	X	(Milton 2009)	
Prototroctes maraena	Retropinnidae	amphidromous		(Augspurger *et al.* 2017)	
Prototroctes oxyrhynchus	Retropinnidae	amphidromous		(Augspurger *et al.* 2017)	
Retropinna retropinna	Retropinnidae	amphidromous		(Augspurger *et al.* 2017)	
Retropinna semoni	Retropinnidae	amphidromous		(Augspurger *et al.* 2017)	
Rhyacichthys aspro	Rhyacichthyidae	amphidromous		(Donaldson & Myers 2002; Milton 2009)	
Rhyacichthys guilberti	Rhyacichthyidae	amphidromous		(Milton 2009; Tabouret *et al.* 2014)	
Johnius belangerii	Sciaenidae	amphidromous	X	(Bijoy Nandan *et al.* 2012)	
Johnius coitor	Sciaenidae	amphidromous	X	(Sakar *et al.* 2018)	
Otolithoides biauritus	Sciaenidae	amphidromous	X	(Bijoy Nandan *et al.* 2012)	
Silonia silondia	Schilbeidae	amphidromous		(Flura *et al.* 2018)	
Hippichthys cyanospilus	Syngnathidae	amphidromous		(Donaldson & Myers 2002; Milton 2009)	
Hippichthys spicifer	Syngnathidae	amphidromous		(Donaldson & Myers 2002; Milton 2009)	

Species	Family	Category	Little information available (e.g. no microchemistry analysis)	References	Resident populations
Microphis brachyurus	Syngnathidae	amphidromous		(McBride & Matheson 2011)	
Microphis leiaspis	Syngnathidae	amphidromous		(Ishihara & Tachihara 2008; Milton 2009)	
Syngnathus abaster	Syngnathidae	amphidromous		(Filipe *et al.* 2009)	
Toxotes blythii	Toxotidae	amphidromous	X	(Milton 2009)	
Toxotes chatareus	Toxotidae	amphidromous	X	(Milton 2009)	
Toxotes jaculatrix	Toxotidae	amphidromous	X	(Milton 2009)	