The genetic basis of phenotypic differentiation in *Python regius* and *Gasterosteus aculeatus*

Alan Garcia-Elfring

Department of Biology

McGill University, Montréal

August 2023

A thesis submitted to

McGill University in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

TABLE OF CONTENTS

Appendix C

Appendix D

THESIS ABSTRACT

Knowledge of the underlying genetic mechanisms responsible for phenotypic evolution is central for understanding the process of adaptation. In my dissertation I use two complementary study systems to generate insight into distinct parts of this process. I first use captive ball pythons (*Python regius*) from the pet trade to identify causal links between genetic variation and phenotypic diversity. Next, to understand real-world fitness consequences of genetic variation, I use natural populations of threespine stickleback (*Gasterosteus aculeatus*) experiencing rapid environmental change. The first section of my research takes advantage of artificial selection on colour and patterning imposed in captive breeding programs to understand genotype-phenotype connections. Most pigmentation studies lack the functional validation needed to make a causal link between genotype and phenotype. Those that do are usually based on a few model species, like the mouse and zebrafish. This raises the question of whether the knowledge gained from these classic model species is generalizable across vertebrates. Furthermore, by far the most intensely studied colour trait is melanin pigmentation, with relatively little known of the genetics of pteridine pigmentation and iridophore structural colouration in non-mammal vertebrates, particularly reptiles. Captive ball pythons display an extraordinary degree of colour variation, making them an excellent model species for the study of the genetics of phenotypic diversification. I use whole-genome sequencing, population genetics, gene-editing, and electron microscopy methods to uncover the genetic basis of a recessive colour phenotype characterized by blotches of white skin. This research led to the discovery of a transcription factor not previously linked to reptile colouration or white spotting in general. Functional validation confirmed the role of this

transcription factor in reptile pigmentation and showed it is required for iridophore development in a lizard model. A genomic analysis of additional Mendelian colour morphs identified genes not only in the melanin pathway but also pteridine pigmentation.

I next used threespine stickleback fish to study the effects of selection acting on genetic variation within natural populations. Stickleback are a classic system in evolutionary genetics for showing evidence of natural selection through parallel evolution of freshwater-adapted ecomorphs from marine ancestors. However, studies on parallel adaptation in stickleback tend to be restricted in time and space. Most have been focused on populations in which the ecological shift (e.g., colonization of freshwater habitats by marine populations), and thus natural selection, occurred thousands of years prior, and they have been confined to a few geographic regions where certain derived phenotypes are repeatedly observed. This leaves open questions of how quickly genomic responses to selection can be detected – months, years, thousands of years - and what alternative evolutionary pathways to freshwater adaptation have been taken in populations outside of the extensively studied locations. My research shows that parallel genomic changes in estuary stickleback can be detected within a single year near to genes linked to osmoregulation, largely mirroring the longer-term patterns observed in post-glacial populations. In addition, lake populations of stickleback from eastern Canada show different genotypic targets of selection to those that have been repeatedly identified in the more well-studied regions on the Pacific coast of North America, suggesting alternative pathways can be used for adaptation. This is likely due to differences in standing genetic variation among populations from different geographic regions as a result of range expansion. Collectively, this research is helping expand our knowledge of the functional

connections between genotype, phenotype, and fitness, and the ways in which they interact to govern the trajectory of evolutionary change.

RÉSUMÉ DE THÈSE

La connaissance des mécanismes génétiques sous-jacents de l'évolution phénotypique est essentielle pour comprendre le processus d'adaptation. Dans ma thèse, j'utilise deux systèmes d'étude complémentaires pour comprendre des aspects distincts de ce processus. J'utilise tout d'abord des pythons royaux (Python regius) captifs issus du commerce d'animaux de compagnie pour identifier les liens de causalité entre la variation génétique et la diversité phénotypique. Ensuite, pour comprendre les conséquences de la variation génétique sur la valeur sélective dans le milieu naturel, j'utilise des populations sauvages d'épinoches à trois épines (Gasterosteus aculeatus) soumises à des changements environnementaux rapides. La première partie de ma recherche s'appuie sur la sélection artificielle de la couleur et les motifs imposée par les programmes d'élevage en captivité pour comprendre les liens entre le génotype et le phénotype. La plupart des études sur la pigmentation n'inclue pas la validation fonctionnelle nécessaire pour établir un lien de causalité entre le génotype et le phénotype. Celles qui établissent ce lien sont généralement basées sur quelques espèces modèles, comme la souris et le poisson-zèbre. On peut donc se demander si les connaissances acquises à partir de ces espèces modèles classiques sont généralisables à l'ensemble des vertébrés. En outre, la pigmentation de la mélanine est de loin le trait de couleur le plus étudié, alors que les connaissances sont relativement limitées sur la génétique de la pigmentation de la ptéridine et de la coloration structurelle de l'iridophore chez les vertébrés non-mammifères, en particulier les reptiles. Les pythons royaux en captivité présentent un degré extraordinaire de variation de couleur, ce qui en fait une excellente espèce modèle pour l'étude de la génétique de la diversification phénotypique. J'utilise le séquençage du génome entier, la

génétique des populations, l'édition de gènes et des méthodes de microscopie électronique pour révéler la base génétique d'un phénotype de couleur récessif caractérisé par des taches de peau blanche. Cette recherche a conduit à la découverte d'un facteur de transcription qui n'avait jamais été associé à la coloration des reptiles ou aux taches blanches en général. La validation fonctionnelle a confirmé le rôle de ce facteur de transcription dans la pigmentation des reptiles et a montré qu'il est nécessaire au développement de l'iridophore dans un modèle de lézard. Une analyse génomique d'autres morphes de couleur mendéliens a de plus identifié des gènes non seulement dans la voie de la production de mélanine, mais aussi dans la pigmentation de la ptéridine.

J'ai ensuite utilisé l'épinoche à trois épines pour étudier les effets de la sélection sur la variation génétique au sein des populations naturelles. L'épinoche est un système classique en génétique évolutive car elle montre des preuves de la sélection naturelle via l'évolution parallèle d'écomorphes adaptés à l'eau douce à partir d'ancêtres marins. Toutefois, les études sur l'adaptation parallèle chez l'épinoche ont tendance à être limitées dans le temps et l'espace. Elles se concentrent sur des populations dans lesquelles le changement écologique (par exemple, la colonisation des habitats d'eaux douces par des populations marines), et donc la sélection naturelle, s'est produit sur des milliers d'années et sur des régions géographiques particulières où certains phénotypes dérivés sont observés de manière répétée. Des questions restent donc ouvertes concernant la rapidité avec laquelle les réponses génomiques à la sélection peuvent être détectées - mois, années, milliers d'années - et les voies évolutives alternatives vers l'adaptation à l'eau douce qui ont été empruntées dans les populations en dehors de ces quelques régions largement étudiées. Mes recherches montrent que des changements génomiques

parallèles chez l'épinoche présent dans des estuaires peuvent être détectés en une seule année près des gènes liés à l'osmorégulation, reflétant largement les schémas à plus long terme observés dans les populations post-glaciaires. En outre, les populations lacustres d'épinoches de l'Est du Canada présentent des cibles génotypiques de la sélection différentes de celles qui ont été identifiées à plusieurs reprises dans les régions les plus étudiées de la côte pacifique de l'Amérique du Nord, suggérant que d'autres voies peuvent être utilisées pour l'adaptation, probablement dû à des différences dans la variation génétique stockée au sein des populations de différentes régions géographiques à la suite de l'expansion de l'aire de répartition. Collectivement, cette recherche contribue à élargir nos connaissances des liens fonctionnels entre le génotype, le phénotype et la valeur sélection, et de la manière dont ils interagissent pour déterminer la trajectoire des changements évolutifs.

ACKNOWLEDGEMENTS

I wish to extend my heartfelt appreciation to my Ph.D. supervisor, Professor Rowan D. H. Barrett, for his invaluable guidance, unwavering encouragement, and constructive feedback. I am grateful for his patience, empathy, and his persistent challenge to foster my growth as a scientist. Financial support from the Fessenden Innovation Prize and the Dr. Lawrence Light Graduate Fellowships in Sustainability is deeply acknowledged. I am indebted to Andrew Hendry for his valuable insights and writing advice. A special note of thanks to Heather Roffey for generously sharing her expertise in the ball python industry and facilitating connections with commercial breeders, significantly enhancing the content of this dissertation.

I thank the Menke lab at the University of Georgia and the Palkovacs lab at the University of California Santa Cruz for their contributions in refining the work presented in chapters 1 and 4. I am appreciative of the ball python breeders (Mutation Creation, T. Exotics, Designing Morphs, The Ball Room) for their cooperation in providing shed skins. I am also thankful to my colleagues, including Tim Thurman, Charles (Cong) Xu, Mathilde Salamon, Janay Fox, Ananda Martin, Antoine Paccard, and Marc-Olivier Beausoleil, for their insightful discussions and camaraderie, which enriched my academic journey.

I extend my gratitude to Professor Virginie Millien, my M.Sc. supervisor, for offering me the opportunity to embark on my academic pursuit at McGill University. Special thanks to Anne-Marie L'Heureux and the students for whom I served as a teaching assistant, as this experience honed my teaching and communication skills.

Of course, I am profoundly thankful to my family. Especially my mother Rosalia Elfring, for her unwavering support throughout my educational endeavor. I appreciate Aracely Salas Jaramillo for her encouragement, and Damien McGuire for the moments of laughter. Lastly, I thank my late father, Mario Cesar Garcia Cocco, for being my inspiration in academic pursuits.

CONTRIBUTION TO ORIGINAL KNOWLEDGE

Chapter 1

I present the findings of research that led to the discovery of a novel mutation affecting patterning and structural colouration in reptiles. We integrated modern population genomic methods and gene-editing to show that a transcription factor is involved in snake patterning and iridophore development. This work constitutes a significant contribution to the limited literature on the genetics of reptile colouring. At the time that our preprint was published (Garcia-Elfring et al. 2020), we were the first to use ball pythons in collaboration with ball python breeders to discover the genetic basis of a Mendelian phenotype affecting colouration. This work is published as Garcia-Elfring et al. *Current Biology* 2022 doi.org/10.1016/j.cub.2023.01.004

Chapter 2

In this chapter I investigate the genetic basis of naturally occurring mutations (i.e., not induced via gene editing) affecting melanin and pteridine pigmentation. Here, we use whole-genome sequencing, population genomics, and bioinformatics to discover a missense mutation in in the gene *mc1r* as the likely causal variant for the ‘clown’ colour morph. This gene has been implicated in ecologically relevant phenotypes in a broad range of vertebrates, although studies on reptiles are limited to a few lizard species (e.g., Rosenblum et al. 2004). This work constitutes an original contribution to pigmentation research as variation in this gene had not been studied in snakes. We showed that

variation in *mc1r* also contributes to pattern formation and not only the level of light or dark melanin pigment as it does in other vertebrate taxa.

In this chapter we also present the discovery of a mutation affecting pteridine pigmentation. Axanthism is characterized by the absence or deficiency of yellow pigment. Axanthic animals are rare in nature and this pigmentation anomaly has been little studied compared to others, like albino. To our knowledge, we present the first genetic basis of a natural case of axanthism. The discovery of a gene variant likely controlling the production of a whole class of pigments, pteridines, and leading to an axanthic phenotype constitutes a significant and novel advance in our understanding of pigmentation. Apart from a few amphibian species, like the axolotl, model organisms for the study of axanthism in vertebrates are few and are lacking in reptiles. Our work also presents the ball python as a viable model organism for the study of axanthism.

Chapter 3

This chapter sheds light on the evolutionary history of freshwater stickleback populations in Eastern Canada, which have been far less studied than populations in Western Canada and Europe. We show that among post-glacial lake populations there is large variation in the degree of differentiation relative to marine populations. This result hints at a complex evolutionary history, with either (or a combination of) multiple colonization events, ongoing gene flow, drift, and selection playing a role in this pattern. We also discovered evidence of parallel evolution at loci near dopamine receptors, suggesting selection on behaviour or osmoregulation (through effects on prolactin). This work sheds light into the

evolutionary history of stickleback from eastern Canada and provides novel loci contributing to freshwater adaptation. This chapter is in preparation for submission to *Ecology and Evolution*.

Chapter 4

Here, we studied estuary populations of stickleback to understand how natural selection may affect genetic variation over very short periods of time - less than a year. We show that seasonal fluctuations in allele frequency contain parallel changes at loci linked to freshwater adaptation in post-glacial populations. Thus, this work constitutes original knowledge by showing in a vertebrate species that evolutionary patterns that had previously only been observed retrospectively after thousands of years can be glimpsed as they begin to take form, over seasonal timescales. This chapter was published in *Molecular Ecology* 2021 doi.org/10.1111/mec.15879.

CONTRIBUTION OF AUTHORS

Chapter 1

AGE, APH, and RDHB conceived the study and its design, with contributions from JDL and DBM. HLR collected and catalogued shed skin samples. AGE performed DNA extractions and bioinformatics to analyze whole-genome data. DBM and JDL carried-out CRISPR/Cas9 project oversight. CES performed in vitro test of tfec CRISPR gRNA, preparation of *tfec* RNP, *tfec* surgeries/microinjection, breeding of *tfec*-/- lizards, eye & skin dissections and stereomicroscope images, preparation of skin samples for TEM imaging and working with TEM microscopy technician. ALI performed egg collection, egg care, screening hatchlings for phenotypes, documentation, and initial analysis of *tfec* phenotypes, genotyping, raising hatchlings, and breeding *tfec*-/- lizards. SPS performed *tfec* surgeries/microinjections. AJA documented *tfec* phenotypes. RSO was instrumental in the creation of the *tyrosinase* mutant line. JDL contributed with the analysis of *tfec* phenotypes, project oversight, project funding (NSF EDGE grant). DBM performed *tfec* gene annotation, gRNA design, genotyping design, analysis of *tfec* phenotypes, project oversight, project funding (NSF EDGE grant). AGE wrote the original draft with all authors contributing to review and editing.

Chapter 2

AGE and RDHB conceived and designed the study. HR collected samples. AGE performed the molecular work and bioinformatics. AGE wrote the original draft with all authors contributing to review and editing.

Chapter 3

AGE, AP, RDHB conceived and designed the study. AP collected samples and performed the molecular work. AGE performed the bioinformatics. AGE wrote the original draft with all authors contributing to review and editing.

Chapter 4

AP, EPP, APH, and RDHB conceived the study. AGE, AP, APH, and RDHB designed the methodological approach. BAW collected samples. AP performed the molecular work. AGE and AP performed the bioinformatics. AGE analyzed the data and created figures with assistance from TJT. AGE wrote the original draft with all authors contributing to review and editing.

INTRODUCTION

My thesis is composed of two parts involving different, but complimentary, lines of research. The first two chapters (Section One) use artificial selection on Mendelian phenotypes to link phenotypic variation to genotypic variation. In the last two chapters (Section Two), I study how natural selection shapes genetic variation from a geologic timescale (thousands of years) to a seasonal one (less than a year).

In Section One, I use a common model system, animal colouration, to study the genetic basis of phenotypic differentiation in an uncommon study organism, the ball python (*Python regius*). Here, I study Mendelian phenotypes found among captive-bred ball pythons using modern genomic and bioinformatic methods. The aim of Section One is to gain insights into the genetic basis of pigmentation, structural colouration, and pattern formation in a broader range of vertebrates than the traditional models of mouse for melanin pigmentation and zebrafish (and to a lesser extent medaka fish) as the standard for non-melanin pigmentation in poikilothermic vertebrates. I take advantage of the extensive variation of Mendelian phenotypes found in captive-bred ball pythons, bridging a gap between two distinct groups interested in the genetics of colouration – reptile breeders and academic geneticists. In Chapter 1, the aim is to investigate the genetic basis of patterning using an aberration found across vertebrates, piebaldism. Piebald individuals exhibit white blotching on otherwise pigmented skin. We use population genomics, gene-editing, and electron microscopy to advance our understanding of pattern formation and structural colouration (i.e., iridophores).

Following Chapter 1 on structural colouration, the aim of Chapter 2 is to advance our understanding of the genetic basis of melanophore and xanthophore colouration in

reptiles. To do this, I use crowdsourcing of shed skin from ball python breeders, apply pooled whole-genome DNA sequencing, population genetic analysis, and variant annotation. I study melanin and pteridine pigmentation in Mendelian phenotypes characterized by changes to melanin (melanophore) and pteridine (xanthophore) pigmentation, the 'clown' and 'VPI axanthic' colour morphs, respectively.

In Section Two, I study how patterns of genomic variation are influenced by selection across different timescales, from months to thousands of years. In this section I use the threespine stickleback (*Gasterosteus aculeatus*) system to test for patterns of parallel evolution as evidence for the action of natural selection. In western Canada, threespine stickleback have independently evolved reduced body armour in freshwater. Freshwater populations from Eastern Canada are known to exhibit the full set of lateral bony plates which is characteristic of marine individuals. However, genomic data for populations from eastern North America are scarce relative to the wealth of genomic data available for populations from western Canada, where parallel evolution of reduced armour traits is widespread.

For Chapter 3, I first characterize levels of genetic differentiation among marine and freshwater populations in Nova Scotia and Newfoundland. Next, I look for evidence of consistent differentiation between marine and freshwater populations, that is, parallel evolution. The aim of this chapter is to gain insight into the evolutionary trajectories of stickleback populations in Eastern North America, which have enabled adaptation and persistence in fresh water for thousands of years. Here, I sample genome-wide variation using a reduced representation approach (RAD-seq) pool-seq approach.

In Chapter 4, I investigate whether the action of selection can be observed over a seasonal timescale using an estuary stickleback system. The aim of this chapter is to advance our understanding of the pace of adaptive evolution during shifts in environmental conditions. This work is also an investigation on whether the adaptive evolution observed after thousands of years can be glimpsed during seasonal fluctuations in selection. The changes we can detect retrospectively in postglacial populations now might not be representative of the changes that took place during the initial stages of adaptation, the signals eroded by recombination and processes (stochastic or deterministic) happening later and not directly related to freshwater adaptation. Therefore, the work presented in this chapter attempts to shed light on the first steps of an evolutionary trajectory that started thousands of years ago in postglacial populations.

LITERATURE REVIEW

PART I

Genetics of pigmentation: linking genotype to phenotype

Colouration in genetics and evolution

The study of organismal colouration has played an important role in the study of genetics and evolution. It was in large part variation in colour of *Pisum sativum* (three of seven traits: flower, cotyledon, and pod) that led Gregor Mendel to the laws of particulate inheritance, the observation that phenotypic traits can be passed through the generations as discrete 'factors', founding the field of classical genetics (Reid and Ross 2011). After Mendel's Laws of inheritance were rediscovered in 1900, a mutant white-eye fly arose in T. H. Morgan's collection of *Drosophila melanogaster*. Noticing that the inheritance of the white eye phenotype followed the normal segregation of sex chromosomes, Morgan provided direct evidence that Mendel's hereditary factors, which Wilhelm Johanssen rebranded as 'genes', were carried on chromosomes (Morgan 1910; Green 2010). In subsequent years additional mutations affecting eye colour and wing morphology were used by Morgan and his students in the 'fly room' at Columbia University to make important advancements in genetics, making the first genetic map (Sturtevant 1913) and establishing chromosomes as hereditary vehicles (Bridges 1916).

The study of colour has played an important role in evolutionary thought. Before Darwin, it was common to regard animal colouration with teleological significance. The purpose of pigments was to provide colour for the esthetic enjoyment of humans (Sumner 1937). The variability of colour within species (e.g., the occurrence of a white blackbird

or a white peacock) also led early naturalists to conclude that colour was an unstable and thus unimportant trait (as opposed to traits associated with "form and structure", Wallace 1877). Darwin, and particularly Wallace, highlighted the benefits that colouration brings the organism itself through camouflage, warning signals, or by appealing to the opposite sex (Wallace 1877; Caro 2017). While corresponding with Alfred Russel Wallace and thinking about how sexual selection could not apply to the colouration of caterpillars, Darwin asked: "My difficulty is, why are caterpillars sometimes so beautifully and artistically coloured?" (Darwin 1887). Wallace wrote that he was preparing a paper on "Mimicry and Protective Colouring" (Wallace 1867), the latter phenomenon subsequently named aposematism (Poulton 1890; Rowe and Halpin 2013; Smith and Bannerman 2016).

It is debated whether Darwin came across Mendel's work. Some researchers suggest that Darwin read Mendel's work but skipped the results on *Pisum* (Bizzo and El-Hani 2009). Darwin's theory of evolution by small incremental steps, continuous evolution, was influential in the rise of the biometry school of thought, which was founded by Darwin's cousin, Francis Galton, the inventor statistical concepts like regression, correlation, and discoverer of regression toward the mean. Scientists who followed Darwin, like Galton, were interested in continuous traits and believed in the blended inheritance. Whereas Darwin proposed the theory of pangenesis, Galton proposed the "The Law of Ancestral Heredity" (Bulmer 2003). The biometricians, which included Karl Pearson and Raphael Weldon, attempted to corroborate Darwin's theory of evolution by quantitative analysis of continuous traits in biological populations. After the rediscovery of Mendel's work on *Pisum*, scientists from the Mendelian or 'genetics'

school of thought, most influential being William Bateson, argued that evolution was discontinuous, disagreeing with the Darwinian maxim *Natura non facit saltum* (Latin for "nature does not make jumps"). The proportion of white in piebald hooded rats were used by Castle (1919) as an example for continuous traits being governed by genetics rather than environment. Castle's student, Sewall Wright, who would later go on to be one of the scientists who helped integrate genetics (Mendelism) with Darwinism (biometry), studied piebald guinea pigs to shed light on the mechanism of colour formation (Wright 1917; 1920). The debate on the nature of heredity, whether it is discrete or continuous, continued until Ronald Fisher gave a mathematical account reconciling biometry with genetics. Fisher (2018) showed that continuous variability can be explained by the cumulative effects of discretely inherited Mendelian factors (Piegorsch 1990).

The mouse in early genetics and pigmentation research

The mouse has played a central role in genetics (reviewed by Russel 1985; Bennett and Lamoreux 2003; Steingrímsson et al. 2006). Following the rediscovery of Mendelian genetics in 1900, scientists rushed to investigate whether mammals conformed to Mendel's Laws. Many used colour variations in mice (*Mus musculus*) as a model system. Mice are suitable models for genetic research because of the convenience of maintaining them in the lab, their relatively short life span, and similarity to humans in terms of genetics and physiology. At Harvard, Castle (1903) studied the inheritance of colour variations: albino vs non-albino, also spotted vs solid coloured, brown vs black, and yellow vs non-yellow. Similarly, research from France showed independent segregation of albino vs non-albino and yellow vs black (Cuenot 1905). Shortly later, Clarence Cook

Little, an undergraduate in Castle's lab, added nine genetic loci that affect pigmentation (Castle and Little 1909, 1910). The quantitative analysis by Dunn and Charles of types of white spotting showed that phenotype (e.g., amount of white) can be controlled by genetic and nongenetic factors (Dunn 1920; Dunn 1937; Dunn and Charles 1937; Charles 1938). Organismal colouration and patterning therefore played a central role in the origin and maturation of genetics and our understanding of heredity.

Melanin pigmentation

All major classes of vertebrates – fish, amphibians, (non-avian) reptiles, birds, and mammals- produce melanin (reviewed by Wakamatsu and Ito 2021). Melanin in the skin of animals comes in two general types (Ito and Wakamatsu 2011), black to dark-brown eumelanin (EM) and an alkali-soluble reddish to brown pheomelanin (PM). In animals, melanins can be classified as those pigments derived from the oxidation and polymerization of a tyrosine precursor; in 'lower' organisms, of phenolic compounds (d'Ischia et al 2013). Melanins are synthesized in melanocytes inside membrane-bound organelles called melanosomes. Melanosomes are then transferred from epidermal melanocytes to keratinocytes (Sturm 2009). In addition to EM and PM found in the skin, neuromelanin can be found in dopaminergic and norepinephrinergic neurons of certain mammals, including humans.

EM is an insoluble (black to dark brown) pigment synthesized from the oxidative polymerization of L-3,4-dihidroxyphenylalanine (DOPA) via 5,6-dihydroxyindole (DHI) intermediates (also DHICA). PM is (a red to light brown) alkali-soluble and derived from

the oxidation of cysteinyldopa (CD) via benzothiazine (BT) and benzothiazole (BZ) intermediates (Simon and Peles 2010). The enzymes involved in the synthesis of EM include tyrosinase, Tyrp1, and Dct/Tyrp2 (Ito and Wakamatsu 2011). Copper ions can also act as a catalyst to take the role of Dct/Tyrp2 (Ito et al. 2013). Synthesis of PM, on the other hand, only requires the tyrosinase (and sulfur-containing cysteine). Additional proteins with roles in melanin production are P protein (*OCA2* gene), MATP protein (*SLC45A2* gene), α-MSH, agouti signaling protein (ASIP), its receptor MC1R (also the receptor for α-MSH), and SLC7A11, a cystine transporter (Ito and Wakamatsu 2011). The melanogenesis pathway has been described in detail by Simon et al (2009).

Mouse coat colour, like human hair, contains EM and PM, which can be analyzed using alkaline hydrogen peroxide oxidation (Ito et al. 2011). Hair colour in mammals is closely linked to the amount and relative proportions of EM and PM. In human hair, black, dark brown, brown, light brown, and blonde hair have similar amounts of PM and decreasing EM. In red hair, there is a greater amount of PM, and it is found at similar levels as EM. Thus, the 'chemical phenotype' correlates with the 'visual phenotype' (Ito et al. 2011).

Mouse pigmentation in the genomics age

The study of mouse coat colour variation has played a central role in the study of pigmentation, summarized by Silvers (Silvers 1979; also reviewed by Russell 1985; Bennett and Lamoreux 2003; Hirobe 2011; Ito and Wakamatsu 2011). With the advent of DNA sequencing, sequence variation and genes underlying pigmentation phenotypes

have been uncovered. In a review of pigmentation genes by Baxter et al, they found more than 250 genes linked to pigmentation or patterning, primarily in mouse, human, and zebrafish models (Baxter et al. 2019). Since then, many more genes associated with some aspect of pigmentation/colouring and patterning have been identified (https://www.ifpcs.org/colorgenes/). However, only ~170 have been cloned and fewer than 20 genes have been shown to play a direct role in the production of melanin and regulation of EM and PM (Ito and Wakamatsu 2011). These include *tyr*, *tyrp*, *agouti*, *mc1r*, *pmel*, *oca2,* and *slc5a2*. Also important for melanin pigmentation is *Mitf* (microphthalmia-associated transcription factor), coding for a transcription factor and master regulator of melanogenesis (Levy et al. 2006). *Mitf* is linked to the differentiation, proliferation, and migration of melanocytes. Among its targets is *Slc7a5*, a transporter of the melanin precursor tyrosine (Gaudel et al. 2020).

Early on, the *agouti* and *extension* loci were shown to play a key role in the production and distribution of PM in mice (Silvers 1979). With the advent of molecular techniques, these loci were identified as genes coding for the agouti signaling protein and the melanocortin 1 receptor, respectively (Hirobe 2011). Animals with mutations to the *agouti* locus are known to produce only black EM (Silvers 1979). Controlling the switch between EM and PM is α-melanocyte-stimulating hormone (α-MSH) and the agouti signaling protein (ASIP). The concentration of cysteine inside the melanosome is also important for regulating the production of EM and PM, with lower levels of the sulphur-containing amino acid promoting the production of EM (del Marmol et al. 1996). For example, mutations to a gene coding a membrane cysteine/glutamate exchanger result in low intracellular cysteine and decreased PM synthesis. The wild-type allele for *Tyrp1* is

linked to black EM whereas the recessive allele produces brown EM, with tyrosinase activity being higher in brown mice (Tamate et al. 1989).

Variation in *Mc1r* (the *extension* locus) and *agouti* have been linked to light PM phenotypes (Silvers 199; Hirobe et al. 2007; Miller et al. 1993). Albino mice lack melanin not because of an absence of melanocytes, as in the leucistic phenotype, but due to a lack of tyrosinase activity, resulting in no pigment in their coat or eyes (Silvers 199; Yamamoto et al. 1989; Tanaka et al. 1990). Variation in the gene *Oca2* produces *pink-eyed dilution* mice, characterized by reduced melanin pigmentation of the coat and eyes. In contrast to albino mice, *pink-eyed dilution* mice have some pigmentation in the eyes (Silvers 1979). Premelanosome protein (PMEL) is expressed in melanocytes where it forms a scaffold (intralumenal fibrils) for EM deposits. Mutations to *Pmel* alter melanosome shape and result in a significant decrease in EM content in hair (Hellström et al. 2011). Similarly, hair graying with age is associated with alterations to melanosome shape (Itou et al. 2018). Mutations to *Oca2*, a gene thought to regulate melanosome pH (Bellono et al. 2014), result in decreased production of EM but the synthesis of PM is not altered (Hirobe et al. 2011a). The effects of genes affecting mouse coat EM and PM have been summarized in Table 2.1 of Wakamatsu and Ito (2021).

In humans, much of the diversity in normal skin colour is based on allelic variants in a few genes, including *TYR*, *OCA2*, *MATP/SLC45A2* and *SLC24A5* (Lamason et al. 2005; Lao et al. 2007; Norton et al. 2007; Pavan and Sturm 2019). Melanosome pH is critical for controlling the relative amounts of EM and PM (Ancans et al. 2001; Zhou et al. 2018). It is known that melanosomes derived from fair skin are acidic, while melanosomes derived from black/dark skin are near neutral (Smith et al. 2004). At pH

5.8, the activity of the Tyr enzyme is 20% of what it is at pH 6.8 (Fuller et al. 2001). Ancans et al (2001) showed that neutralizing melanosome pH increases the ratio of EM to PM.

Bird feathers

It is now known that feathers, scales, and hairs are homologous structures (Di-Poï and Milinkovitch 2016). Birds are among the most varied animals with respect to colouration and apart from mammals, the most studied (McGraw et al. 2004). Bird feathers contain pigments like carotenoids, flavins, porphyrins, psittacofulvins, pterins, purines, and turacin. However, in contrast to poikilothermic vertebrates like reptiles, birds and mammals only have one pigment cell type, the melanocyte. Avian species also have additional melanins called trichochromes and erythromelanins (Hudon 2005). Birds and mammals share the melanogenesis pathway producing EM and PM, although PM is more extensively distributed in birds and likely more chemically diverse (Wakamatsu and Ito 2021).

An important model in avian pigmentation is the domesticated rock pigeon (*Columbia livia*), which varies extensively in ornamental traits, like feather colour and patterns (Guernsey et al. 2013; Vickrey et al. 2918; Bruders et al. 2020). Plumage colouration is regulated by the type and quantity of melanin (EM and PM) deposited in the feathers. Pigeons homozygous for the "recessive red" allele, which causes a down-regulation of the gene *Sox10*, switch to PM production and display red plumage rather than the wild-type black colouration (Domyan et al. 2019). In certain galliform species

variation in *Mc1r* has been to higher levels of PM (e.g., Nam et al. 2020). The Eurasian Scops-Owl (*Otus scops*) exhibits variation in the degree of redness. Levels of PM correlate with redness, although not variation around *Mc1r* (Avilés et a. 2020).

Fish, amphibians, and reptiles

The genetic basis of melanin pigmentation in mammals and birds has been extensively studied. Yet, melanin colouration produced by poikilothermic vertebrates is not as well understood. In fish, like the red seabream (*Pagrus major*), production of EM is linked to androgen levels (Adachi et al. 2010), and is influenced by many factors including temperature, diet, background colour, feeding density, and stress (e.g., Kawauchi et al. 1983; Adachi et al. 2006). Unsaturated acids have been shown to decrease melanin synthesis and tyrosine activity (Ando et al. 1998). Thus, aqua-cultured fish fed a formulated diet decreases melanogenesis relative to controls (reviewed in Wakamatsu and Ito 2021). In reptiles, colouration is in large part the result of melanin, pteridine pigments, carotenoids, and structural colouration (Olsson et al. 2007). It was initially thought that reptiles only contained EM. However, Roulin et al. (2013) discovered that reptiles do indeed produce EM, showing that the EM/PM ratio in the Eastern Hermann tortoise (*Eurotestudo Boettgeri*) correlates well with a gradient of black/yellow colour on the shell. However, in the amphisbaenian (Squamata) *Trogonophis wiegmanni,* yellow scales lack PM, while EM is present in black and yellow scales, suggesting the yellow colour is caused by a reduction of EM or by other pigments (Recio et al. 2022).

In poikilothermic vertebrates, hormonal and neural control of melanisation plays an important in pigmentation. Hormonal regulation of melanin production in specific regions of the skin via the melanin-concentrating hormone plays a larger role in fish and reptile colouration than in mammals (reviewed by Baker 1993; Goda and Kuriyama 2021). In amphibians, melanin production of the skin and liver is controlled by similar genes as in mammals, including tyrosinase (Gallone et al. 2007), as shown by the albino phenotype produced by *Tyr* (tyrosinase) mutants in the Mexican Axolotl, *Ambystoma mexicanum* (Woodcock et al. 2017). In these mutants, melanin is absent but other types of pigments and iridescence remain. In contrast, an all-white phenotype (absence of pigments and iridescence) is produced by mutations *Edn3*, a ligand for the endothelin receptor.

The zebrafish model and chromatophore biology

Unlike birds and mammals - which only have melanin-producing cells - fish, amphibians, and reptiles have three general types of colour-producing cells: melanophores which produce melanin pigments; xanthophores which produce yellow to red pteridine-based pigments (termed erythrophores if they contain red pteridine or carotenoid pigments), and iridophores which produce structural colouration via scattering of light on stacked guanine crystals (Gur et al. 2017; Reviewed by Goda and Kuriyama 2021). Many fish have white leucophores. These cells produce structural colouration via uric acid crystals, although they resemble xanthophores in the genetics of their specification during development (Lewis et al. 2019). All colour producing cells of the integument (skin) are derived from neural crest cells, which generate multiple cell types including neurogenic

and skeletogenic derivatives (reviewed by Hall 2009; Arnheiter and Debbache 2021; Owen et al. 2021). The mechanism of cell fate specification is debated, with the cyclical fate restriction model recently being proposed as a replacement of the progressive fate restriction model (see Thomas and Erickson 2008; Nikaido et al. 2021; Kelsh et al. 2021; Dawes and Kelsh 2021; Subkhankulova et al. 2023). Under the new cyclical fate restriction model, neural crest stem cells transition into a multipotent progenitor with a highly dynamic transcriptome which transitions between states. This multipotent progenitor expresses key markers of most or all fates (i.e., chromatophore types), although different states have biases towards derivative cell types (Dawes and Kelsh 2021).

A great extent of our knowledge of the genetics of pigment cell development and pigment synthesis has been derived from mammals, especially the mouse model. Phenotypes like white spotting and piebald have led to the discovery of genes affecting melanocyte development, including *Kit*, *Mitf*, and genes of the endothelin pathway (Baxter et al. 2004; Wen et al. 2010; Arnheiter and Debache 2021). Pattern formation in fish has been studied far more than in reptiles (or amphibians), although most of what we know comes from one species, the zebrafish (*Dario rerio*) (Singh and Nüsslein-Volhard 2015; Baxter et al. 2019; Owen et al. 2021).

In zebrafish, larval pigment cells are derived directly from neural crest cells. In the adult, larval chromatophores are replaced by postembryonic neural crest-derived stem cells (Parichy et al. 1999; Owen et al. 2021). Zebrafish patterning results from Turing-like interactions among chromatophores (Nakamasu et al. 2009; Kondo et al. 2021). Mutations that affect chromatophore development can alter adult patterning. In this

regard, zebrafish has been and is a powerful system for advancing our understanding of non-melanin colouration and pattern formation in poikilothermic vertebrates.

Pigment cell development

All chromatophores of the integument are ultimately derived from neural crest cells, which themselves are specified by the expression of certain genes, including *foxd3*, *sox9a*, *sox9b*, *sox10* and *tfap2a* (Simões-Costa and Bronner 2015; reviewed by Hashimoto et al. 2021). Expressed in late chromatoblasts is *tfec*, a member of the mitf-family of transcription factors (Petratou et al. 2018). Late chromatoblasts begin expressing *ltk* and *mitf*, signalling the start of iridophore and melanophore specification, respectively (Hashimoto et al. 2021).

Melanophores

Melanophore differentiation begins with the expression of *mitf* and downregulation of *sox10*. *mitf* regulates many genes, including ones with key roles in melanin production. Melanophore development in zebrafish, like in mammals, is regulated by *mitf* (Lister et al. 1999). *mitf* mutants show no melanophores and spots of xanthophores with S-iridophores. Since xanthophores localize with iridophores, this suggest there is an attraction between xanthophores and iridophores. Double mutants lacking melanophores and xanthophores contain only S-iridophores, whereas single mutants have both iridophore types. These results suggest that cellular cues from both melanophores and xanthophores regulate the transition between L and S iridophores, and that iridophores attract xanthophores, whereas xanthophores repel melanophores (Frohnhöfer et al. 2013).

Cell-cell communication is thought to be regulated by filopodia and dendrites for short range interactions and airinemes for long range interactions. (Eom et al. 2015; Eom 2020). Thyroid hormone, known for its role in mammalian metabolism and amphibian metamorphosis, is also critical for pigmentation in zebrafish. Loss-of-function alleles of the thyroid-stimulating hormone receptor results in mutants with twice as many melanophores relative to wildtype and no xanthophores (McMenamin et al. 2014; Owen et al. 2021).

Iridophores

The specification of iridophores begins with the upregulation and sustained expression of *tfec*. In contrast to the melanophore lineage, where *sox10* expression is reduced, *sox10* expression is maintained and required for *tfec* expression in the iridophore lineage (Petratou et al. 2018). There is also a feedback loop between *tfec* and *ltk* required to sustain *tfec* expression (and *Sox10*), leading to the expression of other genes involved in differentiation, such as *pnp4a* (Hashimoto et al. 2021).

For example, most zebrafish embryos with a mutated *leukocyte receptor tyrosine kinase* (*Ltk*) die in the larval stage. Those that make it to the adult stage lack iridohores. Adult *Ltk* mutants display two dark stripes and have reduced melanophores, suggesting iridophores may play a role in melanophore differentiation (Lopes et al. 2008). More recently, the gene *tfec* was linked to iridophore development in zebrafish (Petratou et al. 2021). The key genes involved in iridophore specification and development are transcription factors *Sox10, Tfec* and *Mitf,* and the leukocyte receptor tyrosine kinase *ltk* (Petratou et al. 2018). When iridophores are differentiated, they express *pnp4a* (Petratou et al. 2018).

Xanthophores

Expression of *pax3* and *pax7* specify the xanthophore or erythrophore lineage (Minchin and Hughes 2008; Nord et al. 2016). Mutations that result in the absence of xanthophores show spots of melanophores (Hafter et al. 1996; Odenthal et al. 1996; Patterson and Parichy 2013). Thus, xanthophores are important for stripe formation, at least in zebrafish. Given that the teleost lineage underwent a whole-genome duplication event 320-350 million years ago, and that divergent evolution of paralogous genes via subfunctionalization or neofunctionalization has occurred (e.g., *asip* and *agrp2*), mechanisms for zebrafish pigmentation may not apply in other vertebrates like reptiles. It is therefore important to study the genetics of pigmentation, structural colouration, and pattern formation in a broad range of vertebrates.

Pigmentation and rapid evolution

Evolutionary change, according to Darwin, occurred over long periods of time, too slow to observe and study. “We see nothing of these slow changes in progress, until the hand of time has marked the long lapses of ages, and then so imperfect is our view into long past geological ages” (Henson et al. 2015). Work in guppies by Endler and others and other species helped to demonstrate that evolution can occur over more observable timescales (Reviewed by Thompson 1998). However, few studies outside microorganisms have documented evolutionary change in action, that is, made the link between genotype, phenotype, and fitness effect. Barrett et al (2019) studied the genetic and ecological basis of cryptic coat colour variation of deer mice (*Peromyscus*

maniculatus) that live in different coloured backgrounds, in light-coloured sand and darker vegetation. Prior work showed that the light coat phenotype was linked to a mutation in the *agouti* gene (Linnen et al. 2009). Light and dark coloured deer mice were then housed in enclosures open to avian predators on both dark and light backgrounds. The results showed the *agouti* variant was selected on light coloured backgrounds, with significant changes in allele frequency occurring within a single generation (Barrett et al. 2019). Rapid adaptive evolution of colour has been documented in Soay Sheep (Maloney et al. 2009), Timema stick insects (Soria-Carrasco et al. 2014)

Reptile colouration

The detailed knowledge of the melanogenesis pathway obtained from mice have been applied to other species, including reptiles (Saenko et al. 2015; Iwanishi et al. 2018; Ullate-Agote et al. 2021; Brown et al. 2022). For example, studies on lizards have linked ecologically relevant blanch phenotypes to signatures of selection (i.e., derived alleles at frequencies significantly above general population structuring) in *mc1r* (Rosenblum et al. 2004; Rosenblum et al. 2009; Des Roches et al. 2017; Jin et al. 2020). Most of the work done on the genetics of reptile colouration are studies on melanin-based pigmentation (reviewed by Hoekstra et al. 2006; Olsson et al. 2013). Despite the great variation in non-melanin pigmentation, much of the basic biology and genetics of iridophore (Teyssier et al. 2015; Tang et al. 2023) and xanthophore/erythrophores (McLean et al. 2019) colouration remains to be characterized in most species of poikilothermic vertebrates - except zebrafish.

The use of snakes as model organisms has provided insights into the evolution and development of limbs and fangs (Kvon et al. 2016; Leal and Cohn 2016; Vonk et al. 2008), venom (Vonk et al. 2013), and extreme physiological responses, for example, to digestion (Wang and Rindom 2021). The use of captive-bred reptiles, including snakes, in cooperation with commercial breeders can lead to new discoveries (Guo et al. 2021). Natural colour polymorphisms in nature (Pizzatto and Dubey 2012) and aberrant colourations in captivity (Bechtel et al. 1978) can be used to advance our knowledge of the genetics of reptile colouration and chromatophore biology in areas where it is deficient (e.g., Nicolaï et al. 2016; Tang et al. 2023). Two emerging reptilian models in the study of pigmentation are the corn snake (*Pantherophis guttatus*) and ball python (*Python regius*). Snakes exhibit vast diversity in colours and patterns, including as ringed, striped, uniform, speckled, and blotched (Kuriyama et al. 2013). Colour variation in corn snakes, for example, has been used to show that all chromatophores are lysosome-derived organelles (Ullate-Agote et al. 2015; Ullate-Agote et al. 2020).

The ball python exhibits an extraordinary amount of colour variation in captivity, with hundreds of reported Mendelian phenotypes segregating in the collections of hobbyist and commercial breeders. With few recent exceptions (Irizarry and Bryden 2016; Brown et al. 2022; Dao et al. 2022; Garcia-Elfring et al. 2023), the ball python has been overlooked in pigmentation research. In this thesis, I use the ball python phenotypic diversity found in captivity to make links between genotype and phenotype. This thesis represents a small contribution to our understanding of the genetics of pattern formation and colouration in reptiles.

PART II

Genetics of freshwater adaptation

Parallel evolution across time scales

Section 1 consisted of leveraging a system in which artificial selection is prevalent to facilitate making the link between phenotype and genotype. Although artificial selection can tell us a great deal about how phenotypic traits can evolve, artificial selection simplifies the process by focusing on a single trait in isolation. The advantage of using natural selection is that it shows how adaptation occurs in the real world. The threespine stickleback (*Gasterosteus aculeatus*) has played an important role in biological research, particularly evolutionary genetics. Stickleback evolved in the Eastern Pacific where they have persisted for approximately 26 million years, only colonizing the Western Pacific and Atlantic Ocean approximately 37–347 thousand years ago (Orti et al. 1994; Fang et al. 2018; Fang et al. 2020a). After the last glacial maximum, marine threespine stickleback colonized many newly formed freshwater habitats across the northern hemisphere where they independently evolved similar phenotypes. In freshwater, threespine stickleback have repeatedly evolved reduced armour plating and pelvic spines. Plates and spines help protect against large piscivorous fish in a marine habitat. In freshwater, however, where levels of ions can be lower (e.g., calcium), bony plating is thought to be too costly. Also, the presence of spines is hypothesized to be maladaptive against invertebrate (Odonata) predators which can latch on to these structures. The repeated reduction in pelvic spines and bony plating in threespine stickleback is considered one of the prime examples of parallel evolution and evidence for natural selection (Colosimo et al. 2005).

DNA sequencing and population genetics has provided insights into the genetic basis of freshwater adaptation and parallel evolution in threespine stickleback (Barrett et al. 2008; Schluter et al. 2010; Jones et al. 2012; Morris et al. 2014; Shanfelter et al. 2019; Mack et al. 2023). Comparative genomic studies between marine and freshwater populations, mostly along Western North America, have shown that the *eda* gene controls armour plating (Hohenlohe et al. 2010; Jones et al. 2012), whereas *pitx1* controls pelvic spines (Chan et al. 2010). Mutations to the protein-coding region of *eda* and deletion of the *pitx1* enhancer result in reduced armour plating and pelvic spines, respectively. However, parallel evolution of the freshwater ecotype has been mostly observed in freshwater habitats along Western North America and Europe (e.g., Schluter et al. 2010; Guo et al. 2015). Notably, many freshwater populations along eastern North America exhibit the full bony plate phenotype associated with marine populations (Hagen and Moodie 1982; Haines et al. 2022). Work by Fang et al (2020c) suggests the lack of reduced plated phenotype is due to stochastic loss of adaptive *eda* alleles during range expansion from the Eastern Pacific to the Atlantic Ocean Basin (Fang et al. 2020b, Fang et al. 2020c). The presence of full-plated stickleback in lakes along Eastern North America suggests persistence in freshwater can be attained through the evolution of physiology, like osmoregulation (Divino et al. 2016; Rudman et al. 2019). However, comparative genomic studies of marine and freshwater populations along Eastern North America are lacking (but see Fang et al. 2020).

Stickleback have been used as a model system in evolutionary genomics, ecological speciation, and eco-evolutionary dynamics (McKinnon and Rundle 2002; Hendry et al. 2013; Reid et al. 2021). The adaptive evolution in freshwater over the last

approximately 12-15 thousand years (Liu et al. 2016) seems consistent with Darwin's view of slow evolution, proceeding by steps imperceptible during a human lifetime. However, the threespine stickleback system has been used to show contemporary evolution, perceptible over a timescale of years or decades (Hendry and Kinnison 1999; Bell and Aguirre 2013) rather than millennia. For example, marine stickleback transplanted to freshwater lakes evolved the same degree of cold tolerance as natural freshwater populations within three generations (Barrett et al. 2011). After a 19-year-old selection experiment on the Haida Gwaii archipelago in Western Canada, stickleback showed 75% of the genetic divergence associated with differences between 12-thousand-year-old ecotype (Marques et al. 2018) Russia, 35-year-old lake populations also showed evidence of rapid adaptation, with significant differentiation relative to marine populations and selection coefficients of up to 0.27 (Terekhanova et al. 2014). Similarly, stickleback living in pools formed during an uplift caused by an earthquake in Alaska show evidence of adaptation in less than 50 generations/years (Lescak et al. 2015).

Standing genetic variation in marine populations of alleles that previously experienced positive selection in fresh water, rather than de novo mutations, has facilitated the contemporary evolution documented in stickleback fish (Schluter and Conte 2009; Kingman et al. 2020). Simulations indicate that only a few marine stickleback fish are sufficient to transport freshwater adaptive alleles and result in adaptation, with no need for continued gene flow (Galloway et al. 2020). One experiment shows that in fresh water beneficial alleles, like the one associated with the low-plated phenotype (*eda* chrIV:12,823,875 T>G), can quickly increase in frequency from less than 0.01 to over 0.50 within eight years, and near fixation in 15 years (Kingman et al. 2020).

For Darwin, evolution by natural selection could only be observed retrospectively, after the passage of many thousands of years. Studies on *Drosophila* have been important for showing the seasonal effects of selection (Bergland et al. 2014; Johnson et al. 2023). Whether seasonal shifts in selection result in correlated shifts in allele frequencies in larger species, like vertebrates, is less clear.

This thesis uses artificial selection on ball python pigmentation to make genotype-phenotype links, advance our knowledge of the genetics of structural colouration (Chapter 1) and pigmentation (Chapter 2). I study populations experiencing natural selection to shed light into freshwater adaptation (Chapter 3) and the pace and initial steps of evolution (Chapter 4). I finish by discussing the findings, their limitations, and making recommendations for future works.

CHAPTER 1

Piebaldism and chromatophore development in reptiles are linked to the *tfec* gene

Authors: Alan Garcia-Elfring[1], Christina E. Sabin[2,3], Anna L. Iouchmanov[2], Heather L. Roffey[4], Sukhada P. Samudra[2], Aaron J. Alcala[2], Rida S. Osman[2], James D. Lauderdale[3,5], Andrew P. Hendry[1], Douglas B. Menke[2], and Rowan D. H. Barrett[1].

[1]Department of Biology, Redpath Museum, McGill University, Montreal, QC, H3A 0G4, Canada

[2]Department of Genetics, University of Georgia, Athens, GA 30602, USA

[3]Neuroscience Division of the Biomedical and Translational Sciences Institute, University of Georgia, Athens, GA 30602, USA

[4]Biology Department, Vanier College, Montreal, QC, H4L 3X9, Canada

[5]Department of Cellular Biology, University of Georgia, Athens, GA 30602, USA

Abstract

Reptiles display great diversity in colour and pattern; yet much of what we know about vertebrate colouration comes from classic model species such as the mouse and zebrafish. Captive-bred ball pythons (*Python regius*) exhibit a remarkable degree of colour and pattern variation. Despite the wide range of Mendelian colour phenotypes available in the pet trade, ball pythons remain an overlooked species in pigmentation research. Here, we investigate the genetic basis of the recessive piebald phenotype, a pattern defect characterized by patches of unpigmented skin (leucoderma). We performed whole-genome sequencing and used a case-control approach to discover a nonsense mutation in the gene encoding the transcription factor *tfec*, implicating this gene in the leucodermic patches in ball pythons. We functionally validated *tfec* in a lizard model (*Anolis sagrei*) using the gene editing CRISPR/Cas9 system and TEM imaging of skin. Our findings show that reading frame mutations in *tfec* affect colouration and lead to a loss of iridophores in Anolis, indicating that *tfec* is required for chromatophore development. This study highlights the value of captive-bred ball pythons as a model species for accelerating discoveries on the genetic basis of vertebrate colouration.

Introduction

Colour variation is one of the most visually striking forms of biodiversity and has a long history of study in evolutionary biology, as it is easily observed and is often important for survival (Caro 2017; Endler and mappes 2017; Davison et al. 2019). Vertebrate colour arises from pigments, structural colouration, and cell-cell interactions of three types of cells called chromatophores (Patterson and Parichy 2013; Jahanbakhsh and Milinkovitch 2022) Mammals and birds have only a single type of chromatophore, the melanocyte, which produces the brown pigment melanin. In contrast, reptiles and other poikilothermic vertebrates have melanophores which produce melanin, but also xanthophores and iridophores. Xanthophores contain yellow to orange pteridine pigments (Ziegler 2003; Andrade and Carneiro 2021) and are called erythrophores if they contain red carotenoid pigments (Fang et al. 2022; Huang et al 2021). Iridophores do not contain pigment, but instead have guanine crystals that act as reflective platelets to produce structural colouration (Nicolaï et al. 2021). To date, the study of melanin-based pigmentation pathways has contributed the most to our understanding of pigmentation evolution and development in vertebrates (McNamara et al 2021). Moreover, a limited number of classic model species like the mouse and zebrafish dominate the literature on pigmentation biology (Ziegler 2003; Ishikawa et al. 2015; Jackson 1997; Logan et al. 2006; Sturm 2006; Sturm 2009; Ito and Wakamatsu 2011; Tsetskhladze et al. 2012; Adhikari et al. 2019; Baxter et al. 2019; Feng et al 2021; Seruggia et al. 2021; Irionl and Nüsslein-Volhard 2022; Neuffer and Cooper 2022; Phelps et al. 2022). Importantly, however, the knowledge gained from these models might not translate to other vertebrate groups like reptiles, which remain less studied (Kuriyama and Hasegawa 2017; reviewed

by Hasegawa et al. 2020). Ball pythons (*Python regius*), native to western sub-Saharan Africa and a popular snake in the international pet trade, present an excellent opportunity to study the genetic basis of vertebrate colouration in an emerging reptile model (Iziarry and Bryden 2016; Brown et al. 2022; Dao et al. 2022). Many Mendelian phenotypes (‘base morphs’), representing rare, aberrant colourations (Iwanishi et al. 2018; Borteiro et al. 2021: Ullate-Agote and Tzika 2021), have been discovered in nature and propagated in captivity. Ball python breeders have crossed these (inferred) single-gene colour morphs to produce many more (inferred) multi-locus phenotypes ('designer morphs’, Figure 1). However, the actual genetic basis of these phenotypes remains largely unknown, with a few recent exceptions (Brown et al. 2022; Dao et al. 2022).

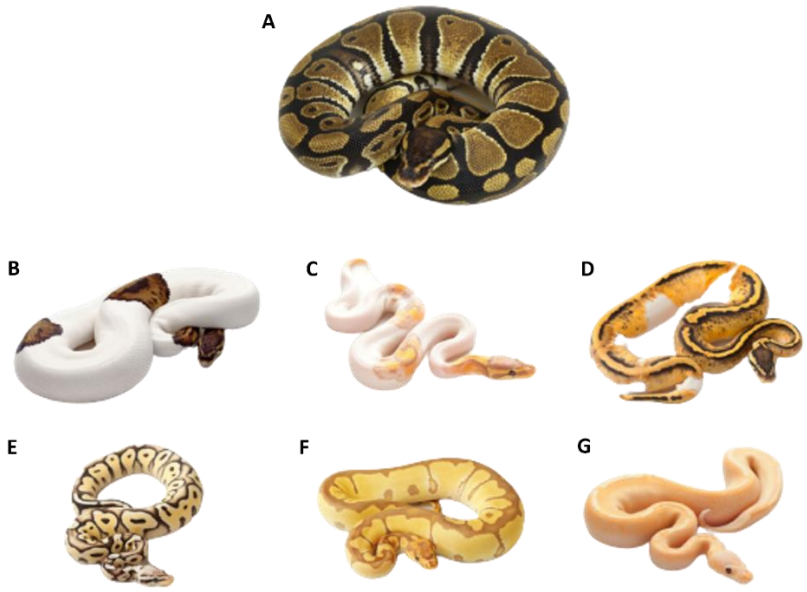


Figure 1. A small sample of the phenotypic variation found in captive-bred ball pythons (*Python regius*). (A) wild type, (B) piebald, (C) banana piebald, (D) pastel piebald, (E) pastel HRA enhancer, (F) ultramel clown, (G) banana champagne. Photo credit: pethelpful.com (A) and *Designing Morphs* (B-G).

We investigated the genetic basis for a classic colour morph found in the pet trade and common across a wide range of vertebrate taxa, the piebald. This phenotype is characterized by leucodermic patches and has been described by commercial breeders as recessive (Barker and Barker 2006). Here, we analyze publicly available clutch data to investigate the mode of inheritance of the piebald phenotype in ball pythons and use whole-genome sequencing and population genomics to identify the genomic region likely containing the causal mutation. Through the annotation of genetic variants (SNPs and indels), we identified a candidate causal mutation in a gene coding for a transcription factor. We functionally validated this locus in a squamate model using CRISPR/Cas9 gene editing and confirmed an effect on chromatophore development by TEM imaging.

Methods and Materials

Experimental model and subject details

Shed skin from ball pythons (*Python regius*) was provided by commercial breeders. We designated shed skin as being 'piebald' if it originated from an individual with white patches hatched from a cross expected to produce the piebald phenotype (i.e., both parents having the allele, as inferred from pedigree). Short read sequencing data was acquired from pooled DNA of snakes with the piebald phenotype (n = 47) and snakes inferred by commercial breeders through pedigree analysis to not have the piebald mutation or mutations (n = 52).

Analysis of clutch data, sample collection, DNA extraction and sequencing

To test whether the piebald phenotype segregates as a simple Mendelian factor, we compiled 10 years' worth of clutch data available online data from a commercial breeder (*KINOVA*). We included data from piebald relevant crosses (https://kinovareptiles.com/incubator/?clutch_id=piebald): piebald vs. inferred non piebald, piebald vs. inferred heterozygotes, and crosses between inferred heterozygotes. We obtained ball python samples (shed skin) by appealing to commercial breeders from Canada (Mutation Creation, T. Dot Exotics, The Ball Room Canada, Designing Morphs). We used a case-control approach, using shed skin samples from 47 piebald individuals (inferred to be homozygous for the piebald variant; Data S1C) and 52 non-piebald individuals (inferred to be homozygous wild-type from pedigrees; Data S1D). Although individuals from both sets of samples contained additional mutations (i.e., other base morphs), the only consistent difference between the two pools was the piebald versus non-piebald phenotype difference. We attempted to maximize the number of individuals that came from different families to minimize the effects of population structure, although there were some exceptions (Data S1D). From each sample, we used approximately 0.1 g of shed skin, cut to small pieces using scissors, for DNA extraction. We extracted DNA following a standard phenol-chloroform procedure, with the modification of a 24-hour proteinase-K incubation time at 37 °C. Piebald and non-piebald samples were prepared on different working days to avoid contamination. We quantified all samples using a Picogreen® ds DNA assay (Thermo Fisher Scientific, Waltham, USA) on an Infinite® 200 Nanoquant (Tecan Group Ltd. Männedorf, Switzerland). After DNA extraction, we mixed DNA of individuals (according to phenotype) in equimolar amounts to obtain a single pool for each phenotype, 'piebald' and 'non-piebald.' Because

extracted DNA from shed skin was degraded, we used PCR-based whole-genome libraries for both pools. We sequenced 150 bp pair-end reads on two lanes of Illumina HiSeqX. Library preparation and DNA sequencing were done at the McGill University and Genome Quebec Innovation Center in Montreal, Canada. The locus of interest (*tfec* exon 5) was validated with PCR and Sanger sequencing.

Bioinformatics

We processed raw reads by filtering for read quality and length with the program *Popoolation* (Kofler et al. 2011). We kept reads with a minimum quality of 20 (--quality-threshold 20) and a length of 50 bp (--min-length 50). We then aligned processed reads to the Burmese python (*Python bivittatus*) draft assembly Pmo2.0 (Castoe et al. 2013) using the program *NextGenMap* (Sedlazeck et al 2013). *NextGenMap* was designed for aligning reads to highly polymorphic genomes or genomes of closely related species. We used *SAMtools* (Li et al. 2009) to convert SAM files to BAM format and remove reads with mapping quality below 20 (samtools view -q 20). PCR duplicates were removed with the program *MarkDuplicates* of *Picard Tools* (Wysoker et al. 2013). We used the *Popoolation2* (Kofler et al. 2011b) protocol to produce a sync file, which contains read counts for all nucleotides sequenced in the genome and used this for subsequent downstream analyses (e.g., F_{ST} scan). In a separate analysis, we applied the same protocol as above but instead aligned reads to the chromosome-length Burmese python reference genome, Python_molurus_bivittatus-5.0.2_HiC.assembly (Dudchenko et al. 2017; Dudchenko et al. 2018).

We applied a genome-wide F_{ST} scan to search for SNPs showing high differentiation between the two pools. For this procedure, we used the *fst-sliding.pl* script

of *Popoolation2* (--min-count 10, --min-coverage 20, --max-coverage 500, --min-covered-fraction 0, --window-size 1, --step-size 1, --pool-size 47:52, --suppress-noninformative). We then identified SNPs with high F_{ST} estimates (F_{ST} = 0.9-1.0) and mapped them to genes. We used a custom script to map SNPs with high differentiation to genes in the gene annotation file using the scaffold name and SNP position. Because the draft assembly of the Burmese python is highly fragmented, we also applied the same F_{ST} scan on data aligned to the chromosome-length genome assembly – thus obtaining better delineation of the genomic region of interest. However, this latter assembly is not annotated with genetic features, hence necessitating the use of both assemblies.

Mendelian phenotypes arise predominately due to mutations to the protein-coding sequences of genes (Chong et al. 2015). We thus annotated variants (SNPs and indels) with the software *snpEff* (Cingolani et al. 2012) to aid in identifying the putative causal mutation for the piebald phenotype within protein-coding genes. *SnpEff* was designed for annotating and predicting loss or reduced function effects of variants on gene protein-products, such as amino acid changes. This program provides an assessment of the impact of a variant, including 'HIGH' (e.g., stop codon), 'MODERATE' (e.g., non-synonymous change), 'LOW' (e.g., synonymous change), or 'MODIFIER' (change in an intergenic area). We verified candidate variants by PCR amplification of target locus and Sanger sequencing of piebald, inferred heterozygotes and inferred non-piebald (i.e., homozygous for reference allel). We used the following primers to amplify exon 5 of *tfec*: Forward: 5'-AACTCAGAGCACTCCATGACC-3'; Reverse: 5'-AGGTGTGCCCCTTTCATAA-3'

Functional validation of the putative piebald mutation in *Anolis sagrei*

Gene editing was performed on wild-caught brown anole females under the approval and oversight of the University of Georgia Institutional Animal Care and Use Committee (A2019 07-016-Y3-A3). All experiments followed the National Research Council's Guide for the Care and Use of Laboratory Animals. CRISPR/Cas9 genome editing was carried out as previously reported [35] with the following modifications: For analgesia, rimadyl (4μg/g) was substituted for meloxicam, and the Cas9 RNP concentration was increased to 10 μM. Cas9 RNP was produced by mixing SpCas9 2NLS with sgRNA (Synthego Corp, Menlo Park, CA) in 10 mM Tris-HCl, pH 7.4. In addition, Cas9 RNP was injected into a maximum of three follicles per ovary, prioritizing the largest follicles. The size of follicles injected ranged from 1mm to 10mm in diameter, and included both previtellogenic and large, yolky follicles. Potential guide sites were obtained using *tfec* coding exon 5 from the *A. sagrei* AnoSag2.1 assembly (Geneva et al. 2021), and targets were chosen using CRISPOR 4.4 (Concordet et al. 2018), selecting targets with Fusi-Scores of 50% or greater. Before performing oocyte injections, we tested the ability of the Cas9 RNP to digest a PCR product than spans the target site. An equal mixture of two sgRNA was used to create *tfec* Cas9 RNP: Targets sites 5' AGAAACAGATACACGAGCAA 3' and 5' AGATACACGAGCAATGGCAA 3'. A total of 44 follicles in 12 adult females were injected to generate four *tfec* mutants. For the production of the *tyr* mutant line, a single sgRNA directed against *tyr* exon 2 was used to create *tyr* Cas9 RNP: Target site 5'ATGATAAAGGGAGGACACCT.

Eggs from CRISPR injected females were collected and incubated at 29°C. Upon hatching, lizard tail clips were collected, and genomic DNA prepared. Hatchlings were screened for mutations in *tfec* coding exon 5 by performing Sanger Sequencing on two different PCR amplicons: 466bp *tfec* amplicon (Tfec-F3: 5'-AAGGGCACATGGCTTGGAAG-3' and Tfec-R3: 5'-CAGTGGGTCTATACTAAACCTGA-3'); 1595bp *tfec* amplicon (Tfec-468-F: 5'-CCATGTACCATTTATCAATGCTATGC-3' and Tfec-1121-R: 5'-CATCGAATTGTTGCCAATCTGTG-3'). Sanger sequencing revealed mutations in two male (*Mutant 1* and *Mutant 3*) and two female (*Mutant 2* and *Mutant 4*) hatchlings. *Mutant 1* and *mutant 2* carried mutant alleles with large size differences that allowed us to gel purify two distinct PCR bands of different sizes from each lizard. We sequenced gel purified bands to obtain clean chromatograms and verify the sequence of the mutant alleles. All mutations shifted the *tfec* reading frame. No evidence of wild-type alleles was detected in the mutants. Only wild-type alleles were detected in lizards with normal pigmentation. To test for germline transmission, *Mutant 1* was crossed with *Mutant 2* and *Mutant 4.* Mutations in *tyr* were identified as previously described (Rasys et al. 2019)[35]. We note that in many vertebrates, F0 genome edited individuals are highly mosaic (Mehravar et al. 2019). Injecting Cas9 RNP into immature lizard oocytes that are not fertilized for days to weeks allows for an extended period for Cas9 RNP to enter the nucleus and cut the target site on the maternal allele (and upon fertilization) the paternal allele. We speculate that may account for the low mosaicism that we observed in the F0 mutants.

Dissection of eyes and TEM imaging of skin

F0 *tfec* mutants were crossed to generate F1 *tfec*-/- progeny that we used for more detailed analyses of eyes and skin. We chose to analyze the eyes and skin as these tissues presented alterations to pigmentation/colouration in *tfec* -/- mutants. An F0 *tyr* mutant male heterozygous for an 8bp deletion in *tyr* exon 2 was crossed to produce heterozygous F1 lizards; F1 *tyr*-/+ lizards were then intercrossed to produce F2 *tyr*-/- lizards. Wild-type, *tyr*-/-, and *tfec* -/- hatchlings were euthanized, and their eyes and skin from the trunk were collected immediately. The freshly dissected tissue was imaged using a ZEISS Discovery V12 SteREO microscope, AxioCam (MRc5), and Axio Vision 4.8.2 (release 06-2010). Electron microscopy was performed following the protocol of Lewis et al. (2017) with modification. Samples were fixed in 2.5% glutaraldehyde in phosphate-buffered saline (PBS) overnight at room temperature. Fixed tissue samples were rinsed three times in PBS for 10 min each, before being dehydrated in increasing concentrations of ethanol consisting of 25%, 50%, 70%, 80%, 90%, 100%, and 100% anhydrous ethanol for 60 min each. Following dehydration, the cells were infiltrated with increasing concentrations of LR White resin in ethanol consisting of 25%, 50%, 75%, and 100% resin for 6 hr each step. After a second change of 100% resin, the samples were embedded in fresh resin in gelatin capsules. The gelatin capsules were capped to exclude air and the resin polymerized in an oven at 60°C for 24 h. The embedded tissues in resin blocks were sectioned with a diamond knife on a Leica Ultracut S microtome and ultrathin sections (60-70 nm) were collected onto formvar-coated 100 mesh hexagonal copper grids. The sections on grids were sequentially stained with 2% aqueous uranyl acetate for 30 min and Reynolds Lead Citrate for 8 min (Reynolds 1963) and viewed in

JEOL JEM-1011 transmission electron microscope at 80-100 kV. Images were captured with an AMT XR80M Wide-Angle Multi-Discipline Mid-Mount CCD digital camera, at a resolution of 3296 x 2460 pixels.

Gene nomenclature

Throughout this article, we follow gene nomenclature established in humans and zebrafish. In humans, gene names are capitalized (e.g., *TFEC* and *MITF*), whereas when referring to genes in other model organisms (e.g., zebrafish, mouse, and reptiles) the gene names are presented in lowercase letters (e.g., *tfec* and *mitf*) for simplicity.

Quantification and statistical analysis

F_{ST} was estimated using the *fst-sliding.pl* script of *Popoolation2*. F_{ST} estimates were visualized using R version 4.2.1. All additional information can be found in METHOD DETAILS.

Results

Mode of inheritance and delineation of genomic region of interest

To test whether the piebald phenotype segregates as a simple Mendelian factor, we compiled 10 years' worth of clutch data from a commercial breeder (*KINOVA*). Consistent with the knowledge among commercial breeders, the proportion of piebald hatchlings indicates the piebald phenotype is inherited as a recessive Mendelian factor (Figure 2A). We applied whole-genome pool-seq to two sets of individuals, one set

having the piebald phenotype and another set inferred by commercial breeders through pedigree analysis to not have the piebald mutation or mutations. Our whole-genome sequencing effort resulted in 977,907,754 reads for the piebald (case) pool and 964,346,462 reads for the non-piebald (control) pool. We obtained an average read coverage of 50.5 and 52.6 for the piebald and non-piebald pools, respectively. To map SNPs showing high differentiation between pools to genes, we aligned reads to the annotated Burmese python (*Python bivittatus*) draft genome (Pmo2.0), from which we obtained 3,095,304 SNPs after filtering. Across all SNPs, we found an average F_{ST} of 0.03456 – indicating that population structure was successfully minimized. To delineate the genomic region of interest, we also mapped reads to a chromosome-length assembly. Using the draft assembly, we identified 129 fixed SNPs ($F_{ST} = 1.0$) and 369 SNPs with $F_{ST} > 0.9$ (Data S1A). Indeed, the chromosome-length assembly shows a single 8 Mb region of high differentiation on scaffold seven (7: 49526089– 57612101), clearly delineating a genomic region of interest (Figure 2B).

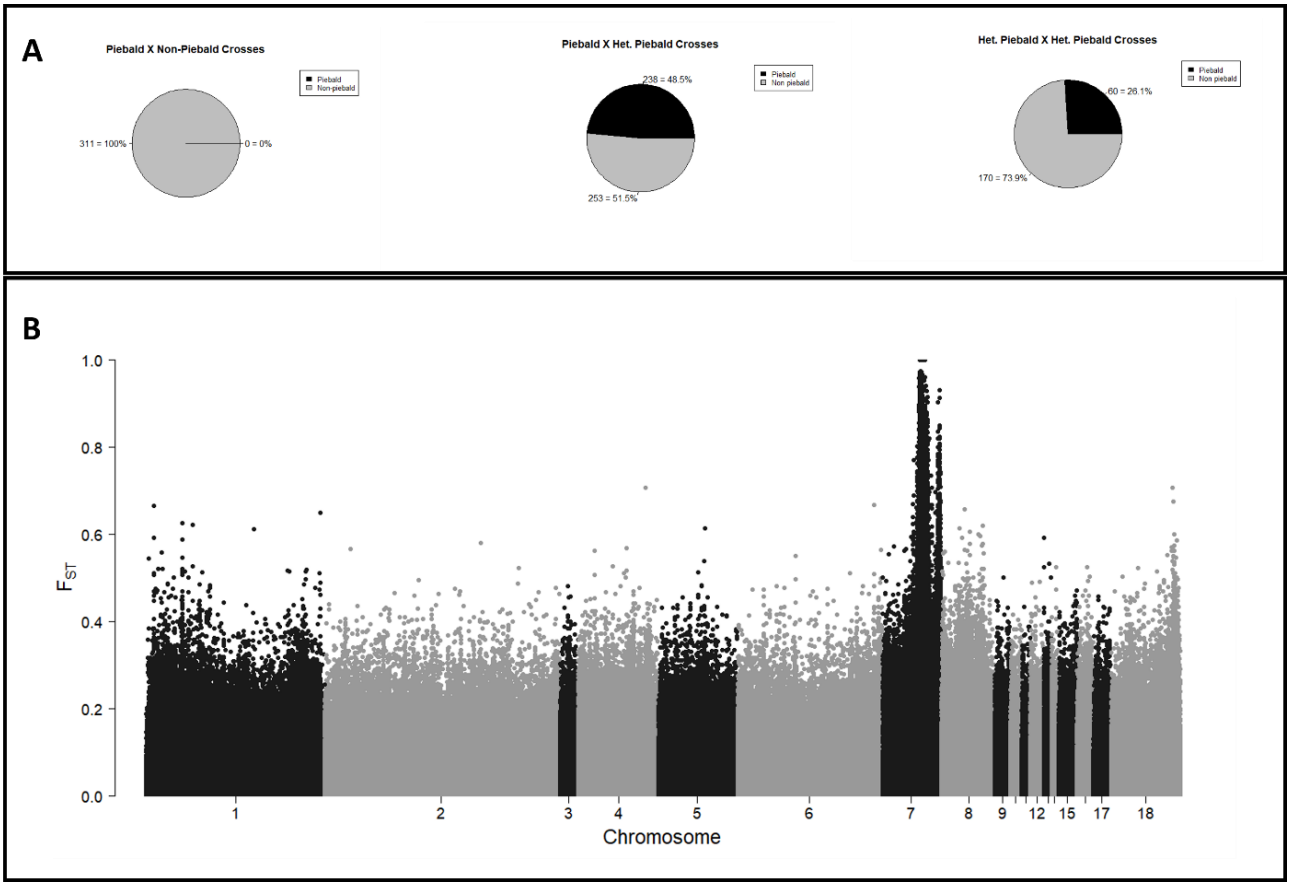


Figure 2. Inheritance patterns and genomic differentiation. (A) Clutch records (2008-2018) from a commercial breeder (*KINOVA*) indicates piebald has a recessive mode of inheritance. Out of 311 hatchlings from piebald to non-piebald crossings, 0% showed the piebald phenotype. Out of 491 hatchlings from piebald to heterozygote crosses, 48.5% were piebald. 26.1% of 230 hatchlings from heterozygote crosses showed the piebald phenotype. (B) F_{ST} plot between piebald and non-piebald samples using a chromosome-length genome assembly. The F_{ST} peak on chromosome 7 delineates the region of interest containing the putative causal gene for the piebald phenotype.

Candidate genes and causal mutation

To obtain a list of candidate genes, we determined the gene annotations of variants with $F_{ST} > 0.90$ (Data S1B). We used $F_{ST} > 0.90$ (rather than $F_{ST} = 1$) to account for factors that might preclude finding a fixed causal mutation (e.g., sequencing error, misidentification of a sample, or minor sample contamination, as multiple snakes are often housed together by commercial breeders during breeding). We annotated SNPs for predicted loss-of-function to identify candidate causal mutations for the piebald phenotype. We found variants that mapped to the protein-coding sequences of 32 different genes. Most of the variants do not have a predicted effect on proteins, instead mapping to intronic and intergenic regions (344 'modifier' variants) and including one synonymous SNP (one 'low' impact mutation). The sole exception was a nonsense SNP (i.e., stop-gained mutation, 'high' impact) with $F_{ST} = 0.96$ located within the fifth coding exon of the *tfec* gene (NW_006534020.1 160458). On the chromosome length assembly, exon five spans 7: 52856864-52856924. This variant consists of a c.493C>T (p.Arg165*) mutation, resulting in a premature opal termination codon. This mutation was validated by Sanger sequencing of *tfec* exon 5 and is expected to result in a truncated protein with functional domains missing (e.g., basic helix-loop-helix on exon 7). The coverage of the reference and alternative alleles are 1X and 47X in the piebald pool and 46X and 0X non-piebald pool, respectively. The single read for the reference allele sequenced in the piebald pool resulted in a F_{ST} below 1.00, potentially due to sample misidentification or minor contamination from co-housed animals. Among non-piebald samples, the reference allele is fixed. We verified the candidate variant by genotyping known piebald, heterozygotes, and non-piebald individuals (Figure A1)

Deletion of a splice acceptor site in snakes

To examine sequence conservation around the candidate variant, we generated a multispecies sequence alignment of the *tfec* coding exon five and flanking intronic sequence. This alignment revealed the presence of a 4 bp deletion in snakes at an intron-exon junction relative to other vertebrates (Figure A3). Our analyses of RNA-seq data from the brown anole lizard (*Anolis sagrei*) demonstrate that this snake-specific deletion removes one of two alternative splice acceptor sites (i.e., 3' splice sites) that appear to be used in other squamates (Figure A3). We further note that certain other vertebrate species have single base pair changes that remove either splice site acceptor 1 (seen in some turtles) or splice site acceptor 2 (seen in some mammals, including humans). The use of splice acceptor 1 results in the inclusion of two additional codons relative to transcripts generated using splice acceptor 2. The functional differences, if any, between *tfec* proteins generated by the two different splice acceptors are unknown. However, we infer that ball pythons likely use the second acceptor site, which is intact in ball pythons and other snakes. The stop codon mutation identified in piebald ball pythons occurs 6 bp from splice acceptor 2.

Targeted mutation of *tfec* in *Anolis* lizards

Protocols for genome editing in reptiles have been slow to develop because microinjection of single-cell embryos (zygotes) is difficult. To date, the brown anole lizard is the only squamate in which CRISPR/Cas9 has been successfully applied (Rasys et al. 2019). Therefore, to functionally validate *tfec* as a gene with a role in reptile colouration, we generated lizards with reading frame disrupting mutations in *tfec* coding exon 5, successfully producing four F0 mutant individuals. *Mutant 1* carried one allele

with a 56 bp deletion and a second allele with a 1 bp deletion, *Mutant 2* carried a 190 bp inversion and a 295 bp deletion, *Mutant 3* had a 4 bp deletion, and *Mutant 4* had a 13 bp insertion. These lizards all exhibited altered pigmentation phenotypes (Figure A4 A-D).

Relative to lizards with normal pigmentation (Figure 3A), the four F0 *tfec* mutants showed reduced colouration, particularly in the snout, arms, and legs (Figure 3B; Figure A4 A-D). In this respect, the mutant phenotype is like the reduced pigmentation observed in piebald ball pythons. However, in contrast to ball pythons, the anole mutants have black eyes and lack the leucodermic patches characteristic of the piebald phenotype. Lateral and ventral views also revealed that the skin of *tfec* mutants is translucent, allowing the internal organs and ribs to become more visible. We noted that Mutant 1 displayed small patches of skin on its head that were wild type in appearance, suggesting the possibility of mosaicism in this gene edited animal (Figure A4A). Therefore, we generated F1 lizards to examine pigmentation patterns in the offspring of *tfec* mutants. Crossbreeding mutant F0s together demonstrated that, just as in ball pythons, *tfec* is not required for viability or fertility in brown anoles in captivity. All F1 progeny (n=33) recapitulated the pigmentation phenotypes observed in the original F0 *tfec* mutants with no evidence of skin patches with wild type pigmentation (Figure A4E).

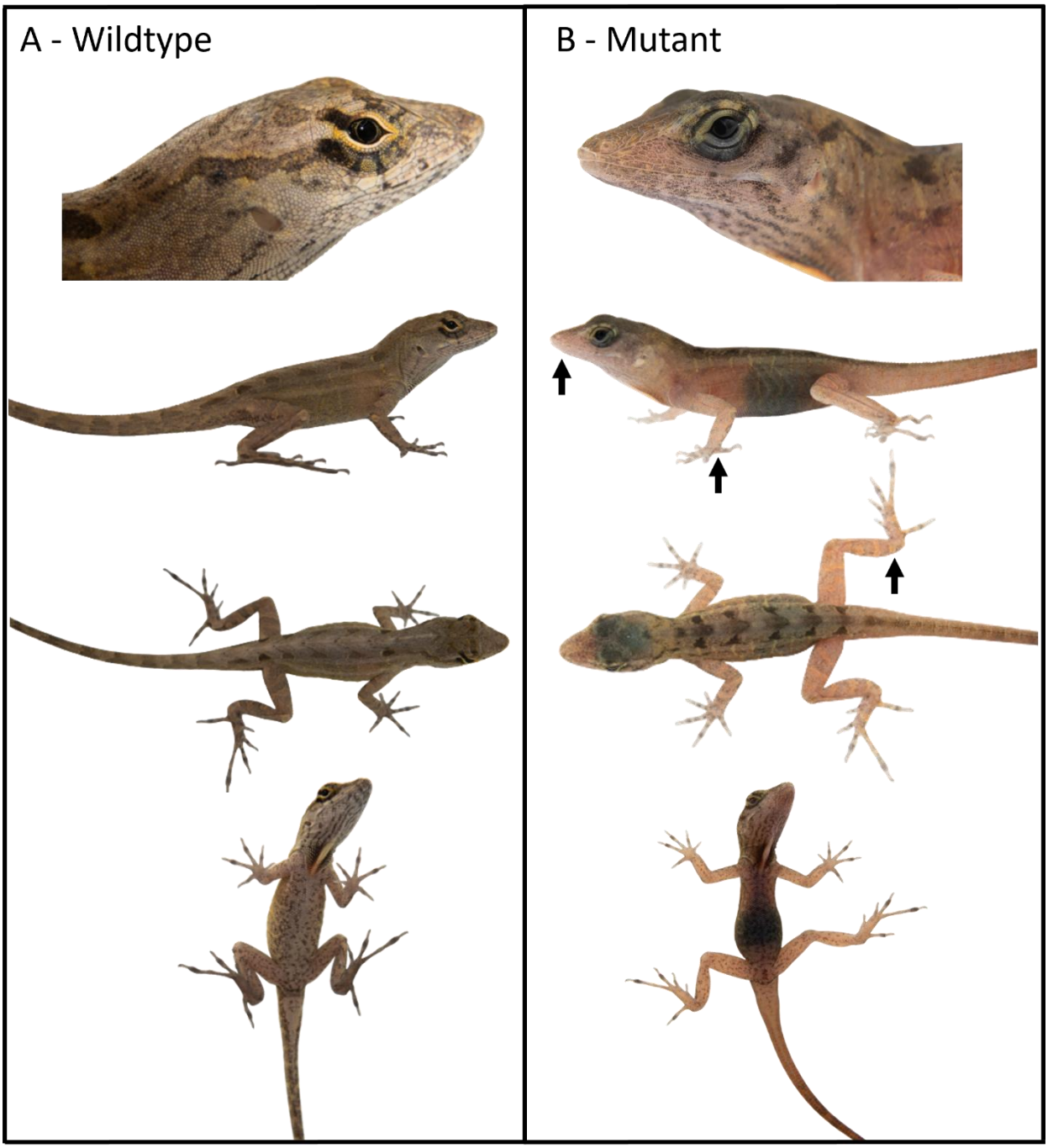


Figure 3. Phenotypic comparisons of *Anolis sagrei*. Wild type (A) and F0 *tfec* mutant (B). The mutant showed reduced body colouration, particularly in the snout, forelimbs and hindlimbs (arrows).

To further understand the phenotype caused by our induced mutations, we examined the eyes and skin of *tfec* and *tyrosinase* (*tyr*) brown anole mutants and compared them to wild-type individuals. The gene *tyrosinase* was chosen as an additional control for these comparisons because of its role in melanin production. Through CRISPR/Cas9 editing we targeted *tyr* and generated a line of *tyr* mutants that carry an 8bp deletion in exon 2 of this gene. We dissected the eyes and skin from hatchlings homozygous for reading frame disrupting mutations in *tfec* or *tyr* (Figure 4, Figure A4). External examination of the eyes and skin revealed a loss of iridophores and presence of melanophores in $tfec^{-/-}$ F1 hatchlings. The loss of the iridescent iridophores make the eyes appear dark, much like the eyes of zebrafish that carry *tfec* mutations (Petratou et al. 2021). In contrast, $tyr^{-/-}$ hatchlings retained iridophores but have an absence of melanophores. These changes in pigmentation were confirmed by transmission electron microscopy (TEM) on skin samples. In wild-type skin, TEM readily detected melanosomes and guanine crystals, which are characteristic features of melanophores and iridophores, respectively. In contrast, melanosomes were absent *from* $tyr^{-/-}$ skin and guanine crystals were absent from the skin of $tfec^{-/-}$ hatchlings.

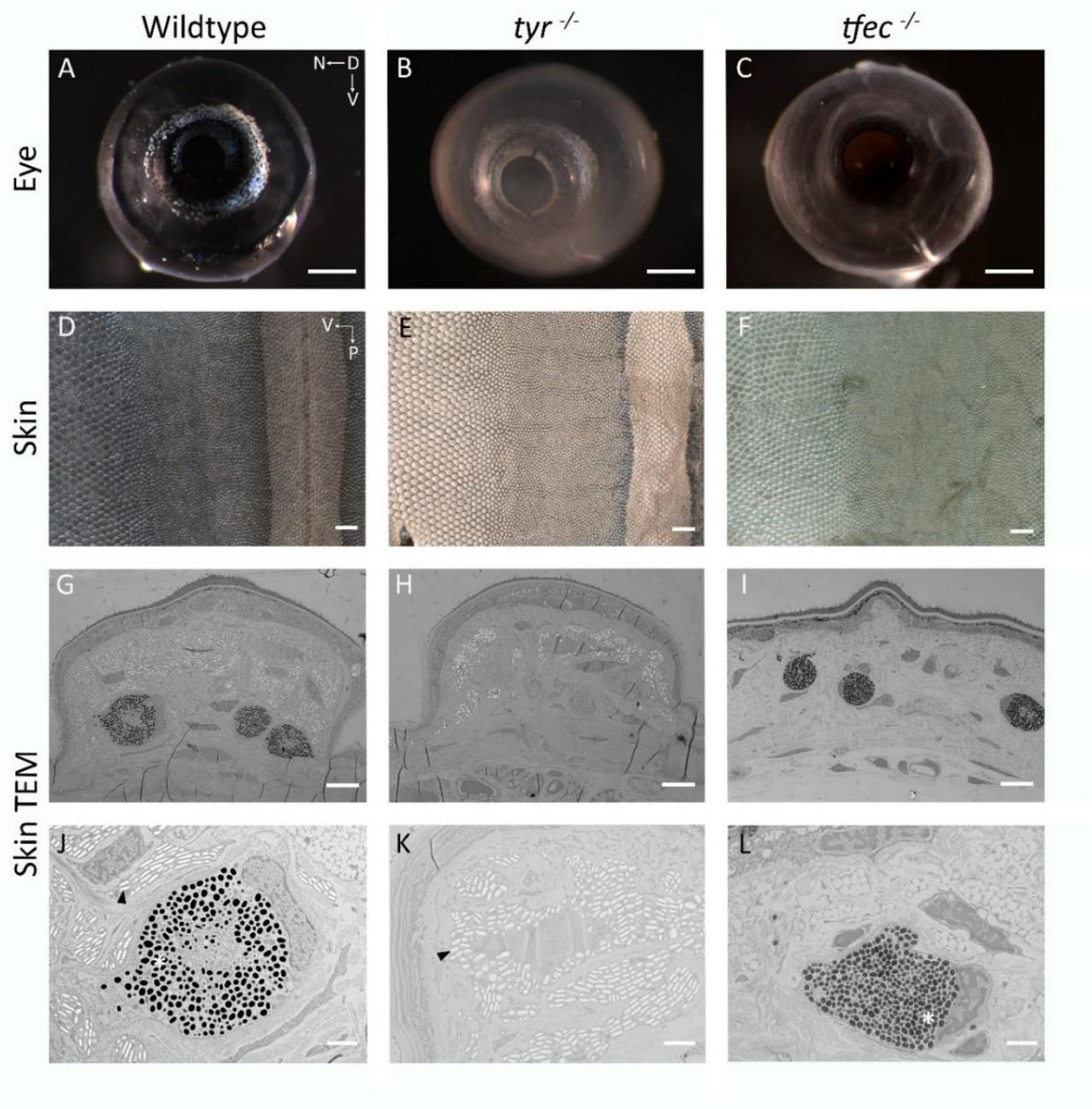


Figure 4. *tfec* is required for iridophore development in *Anolis sagrei*. Presented are eye and skin samples from wild type (A, D, G, J) and mutants with reading frame mutations in *tyr* (B, E, H, K) and *tfec* (C, F, I, L). (A-C) Anterior view of hatchling eyes. (D-F) Dissected skin from the trunk of hatchlings. For these panels, anterior surface is up, and the posterior surface is down. Ventral surface is on the left side of the image and the

dorsal surface is on the right. The dorsal stripe can be seen on panels (D) and (E) while (F) exhibits a lack of this back pattern. (G-L) Transmission electron microscopy images of individual dorsal scales (G-I) and higher magnification images of melanophores and iridophores (J-L). Melanophores hold pigmented melanosomes while iridophore reflectiveness arises from guanine crystals. For *tyr*$^{-/-}$ note the absence of melanosomes and the presence of guanine crystals. For *tfec* $^{-/-}$ note the presence of melanosomes and the absence of guanine crystals. *tyr* samples are from F2 lizards; *tfec* samples are from F1 lizards. Asterisks show melanosomes while arrowheads point to guanine crystals. Scale bars: (A-F) 500 μm, (G-I) 6 μm, (J-L) 2 μm.

Discussion

A MiTF/TFE transcription factor linked to reptile colouration

The *tfec* gene encodes a transcription factor from the MiT-family of genes, which includes *mitf, tfe3, tfeb,* and *tfec*. These genes encode transcription factors that have basic helix-loop-helix and leucine zipper functional domains with important roles in lysosomal signaling, metabolism, and pigmentation (Slade and Pulinilkunnil 2017; Hejna et al. 2019). *TFE3* and *TFEB* have pivotal roles lysosomal acidification and autophagy (Settembre et al. 2011; Martina and Puertollano 2017), while *MITF*, *TFE3*, and *TFEB*, have all been linked to the development of cancer (Levy et al. 2006; Slade and Pulinilkunnil 2017; Goding and Arnheiter 2019). *MITF* is also considered a master regulator of melanocyte development and was first discovered through its association with Waardenburg syndrome type II (Tassabehji et al. 1994), which is characterised by

deafness, hypopigmentation, and microphthalmia (Steingrímsson et al. 1994). Mutations to *mitf* in mammals have been shown to affect melanocyte differentiation, resulting in apoptosis (Hu et al. 2021) and leucodermic patches (Baranowska et al. 2014; Hauswirth et al. 2019). One of the few studies investigating snake pigmentation identified a mutation in *mitf* in leucistic Texas rat snakes (Ullate-Agote and Tzika 2021); this mutation, which results in an all-white phenotype, causes the loss of melanophores and xanthophores, but not iridophores.

Of the genes in the MiT-family, the function of *tfec* is the least well understood (Lister et al. 2011; Agostini et al. 2022), but studies have shown that it is expressed, like *mitf*, in neural crest cells and retinal pigment epithelium of fish and mammals (Rowan et al. 2004; Bharti et al. 2012; George et al. 2016). In mouse and zebrafish models, *mitf* and *tfec* are required for normal eye development. Both *tfec* and *mitf* encode proteins with very similar helix-loop-helix domains, and it has been proposed that these two transcription factors regulate gene expression together as heterodimers (Bharti et al. 2012). Kuiper et al. (2004) studied the expression patterns of MiT genes in human tissues and showed that *tfec* and *mitf* have multiple promoter regions and alternative splicing of functional domains, which may modulate target gene regulation (Lister et al. 2011). Interestingly, they found that *tfec* displays the broadest variety of functionally distinct isoforms, with differential spatiotemporal tissue distribution (e.g., spleen, kidney, bone marrow and small intestine). Although this study did not investigate the expression pattern in skin, the premature stop codon in the fifth exon of *tfec* found in piebald ball pythons is expected to result in a protein with missing basic helix-loop-helix and leucine zipper functional domains, likely disrupting target gene regulation (Kuiper et al. 2004).

Our study adds *tfec* to the list of genes implicated in white spotting and pattern formation (Baxter et al. 2004; Ahi and Sefc 2017).

tfec phenotypes in reptiles and other vertebrates

Identifying genes that affect colour across a wide range of vertebrate species can lead to a deeper understanding of the mechanisms that underlie variation in colour and pattern. Given the absence of iridophore and xanthophore cell types in mammals, it is particularly important to expand functional genetic studies of pigmentation beyond mice to better understand the biology of these chromatophore cell types. For example, while we found that *tfec* affects reptile colour, a mouse study showed that *tfec* mutants have normal coat pigmentation (Steingrímsson et al. 2002), highlighting the need to study a wider range of taxa than traditional model organisms. Indeed, *tfec* was not included in a recent curated list of genes known to affect pigmentation (Baxter et al. 2019). However, more recent work on zebrafish has shown that *tfec* is required for iridophore cell fate specification (Petratou et al. 2021). Zebrafish *tfec* mutants also display delayed development of melanophores and xanthophores, but these chromatophores recover by day 4 post-fertilization. The *tfec*-associated phenotypes in reptiles and zebrafish contrast with reported *mitf* phenotypes in these species. In zebrafish, mutations to *mitf* result in a loss of melanophores, a reduction in xanthophores, and an increase in iridophore density (Lister et al. 1999). In contrast, Texas rat snakes with a mutated *mitf* gene are leucistic (i.e., all-white), lacking melanophores and xanthophores but showing no difference in iridophore density relative to the wild type (Ullate-Agote and Tzika 2021). Therefore, mutations in *tfec* and *mitf* produce distinct pigmentation phenotypes, with *mitf* playing a

key role in melanophore development across vertebrates and *tfec* in iridophore development in fish and lizard models.

Our results support the conclusion that mutations to *tfec* in ball pythons cause piebaldism or white spotting, whereas in the brown anole they result in hypopigmentation and lack of iridophores. However, three main points remain unresolved. First is the question of what accounts for the species-specific differences in pigmentation phenotypes. It is known that ball pythons and lizards likely acquire their adult colour pattern by different mechanisms (Jahanbakhsh et al. 2022). Ball pythons have a fixed pattern specified in the embryo, prior to hatching and scale development, and as adults they do not show scale-by-scale colouration. In contrast, many lizards exhibit scale-by-scale colouration that is specified between the juvenile and adult stage. Thus, differences in the timing of gene expression of chromatophores may play a role in the species-specific differences in pigmentation phenotypes. Also unresolved is whether piebald ball pythons have iridophores in either pigmented or white skin. In the Texas rat snake, white colouration arises with iridophores present (Ullate-Agote and Tzika 2021). In the leopard gecko, skin from the ventral side is white but features a complete absence of all chromatophores (Szydłowski et al. 2016). Since *tfec* is required for iridophore development in both the brown anole and zebrafish, the white patches in piebald ball pythons may lack all chromatophores. However, TEM imaging will be needed to confirm the chromatophore content of piebald skin in ball pythons. Between the initial and final submissions of this dissertation, Tzika (2023, in press) demonstrated that ball pythons, both wild type or piebald, lack iridophores in their skin. Furthermore, Tzika confirmed the absence of all chromatophores from leucodermic patches of piebald ball pythons. A

third point that requires further study is the role of the splice site deletion we detected in snakes and the function of different TFEC protein isoforms across reptile taxa. The splice site deletion itself does not cause piebaldism, since it is present in wild type ball pythons and other snake species. However, two distinct splice acceptor sites are conserved across many squamate reptiles, and our data demonstrate that both acceptor sites are used in anoles. Whether the ability to produce different TFEC isoforms contributes to species-specific differences in *tfec* function remains to be tested.

Conclusion

In summary, the finding of a nonsense mutation associated with the piebald phenotype in ball pythons in combination with targeted mutation and TEM imaging in a brown anole model shows *tfec* has an important role in reptile colouration. Mutations to *tfec* lead to hypopigmentation and a loss of iridophores in the skin and eyes of brown anoles. In snakes, *tfec* is likely to be required for the development of chromatophores migrating to body regions that correspond to leucodermic patches observed in piebald ball pythons. Our work highlights the advantages of using ball pythons as a model organism and working with non-academic communities like reptile breeders to accelerate discoveries in pigmentation research in an under-studied class of vertebrates.

Bridging text

In poikilothermic vertebrates like reptiles, colour is produced by pigments and structural colouration. In Chapter 1, we studied the genetic basis of a Mendelian phenotype affecting the normal pattern of ball pythons and characterized by white blotching (piebald). We associated phenotypic variation (white blotching vs normal) to variation in a gene coding for a transcription factor (*tfec*). Chapter 1 provided a new gene associated with piebaldism and required for structural colouration.

Some ball python colour morphs are characterized by changes to pattern but not pigment (e.g., white spotting phenotypes), some have changes to colour but not pattern (e.g., axanthic and albino phenotypes), whereas others have changes to both pattern and colour. In Chapter 1 showed the genetic basis of structural colouration was uncoverd. to complement our findings on structural colouration and patterning in reptiles, we investigate the genetic basis of pigmentation. Here, we focus on the 'clown' phenotype, which is characterized by altered melanisation (reddish brown) and patterning. We also focus on an overlooked phenotype in the literature, axanthism. This aberration is characterized by the absence of yellow (pteridine) pigments on a normal pattern. By studying the genetic basis of the 'VPI axanthic' colour morph, we promote the ball python as a model organism for the study of axanthism.

CHAPTER 2

The genetic basis of pigment pattern formation and axanthism in ball pythons (*Python regius*)

*Alan Garcia-Elfring, °Heather L. Roffey, *Andrew P. Hendry, and *Rowan D. H. Barrett

*Redpath Museum and Department of Biology, McGill University, Montreal, QC, Canada

°Biology Department, Vanier College, Montreal, QC, Canada

Abstract:

Much of our knowledge of the genetic basis for vertebrate colouration comes from classic model species like the mouse, with relatively few studies on reptiles. In this study, we advance captive-bred ball pythons (*Python regius*) as a vertebrate model for identifying the genetic basis of colour in reptiles. We obtained ball python shed skin samples from commercial breeders to study the genetic basis of melanin and pteridine pigmentation. The clown colour morph exhibits altered melanin pigmentation and reduced patterning, displaying a reddish/golden-brown colour and dorsal stripe. We also studied the genetic basis of an overlooked colour anomaly, axanthism, focusing on the 'VPI axanthic' ball python colour morph. Using whole-genome pool sequencing (pool-seq) and variant annotation, we found a missense mutation in *mc1r* associated with the clown phenotype. This finding is consistent with the idea that melanophores play an important role in pattern formation in reptiles, as iridophores do in zebrafish. For the first time, we identify the genetic basis for the lack of any yellow pigments, showing this recessive phenotype is associated with a premature stop codon in *gtp cyclohydrolase 1* (*gtpch*). Our findings indicate that in snakes pattern formation affected by *mc1r* function but not the production of pteridine pigments.

Introduction

The study of animal colouration has a long history in evolutionary biology (Wallace 1882; Cott 1940; reviewed by Caro 2017) and genetics (Lyon 1963) because it is ecologically important and easily observable (Endler and Mappes 2017). Animal colouration is the end-product of processes that begin in the neural crest during embryonic development and include cell-fate specification, proliferation, migration, cell-cell interactions, and pigment synthesis (Kelsh et al. 1996). Animals show variation in colour (e.g., brown vs. yellow) and pattern (e.g., solid brown vs. white spotting). Aberrations to pattern, like white spotting, are often caused by mutations in genes expressed in neural crest stem cells (Dawes and Kelsh 2021). These genes can affect cell-fate specification and result in changes to pattern. Phenotypes characterized by changes in colour alone are usually caused by mutated genes involved in pigment synthesis (Mills and Patterson 2009).

Within vertebrates, mammals and birds have only one type of pigment-producing cell, the melanocyte. Melanocytes produce melanin, the most widespread pigment in animals and the best studied (Mills and Patterson 2009; Baxter et al. 2019). Melanin occurs in brownish-black (eumelanin) and reddish-yellow (pheomelanin) forms and functions in visual signalling, UV protection, homeostasis, host-parasite interactions, and immunity (Ducrest et al. 2008; Côte et al. 2018; McNamara et al. 2021). In contrast, reptiles and other poikilothermic vertebrates have three general types of colour-producing cells, known as chromatophores (Bechtel 1991; Olsson et al. 2013). Like mammals and birds, reptiles produce melanin in cells called melanophores (Megía-Palma et al. 2018). In addition, however, reptiles produce yellow to red pteridine pigments in cells called

xanthophores (Andrade and Carneiro 2021; Kimura 2021). These cells can also be red from carotenoids, in which case they are referred to as erythrophores. In addition to pigmentation, there is structural colouration in poikilothermic vertebrates. For example, scattering of light in guanine crystals in cells called iridophores. Leucophores, like iridophores, produce structural colouration but from uric acid crystals. However, the genetics of leucophore specification more closely resembles that of xanthophores and melanophores than iridophores (Lewis et al. 2019).

Many studies have documented genes and cellular processes associated with colouration and patterning (Frohnhöfer et al. 2013; Singh et al. 2014; Haupaix et al. 2018; Luo et al. 2021). Yet, the identification of causal mutations has been mostly limited to 'classic' model organisms and melanin-based pigmentation, which has been the focus of most work to date (Steingrímsson et al. 2004; Levy et al. 2006; Lin and Fisher 2007; Pavan and Sturm 2019; Feng et al. 2021; McNamara et al. 2021). In pigmentation research, these models include primarily zebrafish (Quigley et al. 2002; Baxter et al. 2019), the mouse (Baxter et al. 2019) and humans (Feng et al. 2021). These species have advanced our understanding of the genetic basis of vertebrate colour production. However, differences in the cellular mechanisms between these model species and other vertebrate taxa could lead to a biased and limited understanding.

Snake colouration and pattern provide a valuable opportunity to study adaptive evolution, including the evolutionary genetics of warning colouration, its mimicry, convergent evolution of stripes, cryptic colouration, and motion dazzle colouration (Kuriyama et al. 2020). Recent work on reptiles has shed light on the origin of chromatophores, indicating chromatophore vesicles are lysosome-related organelles

(Ullage-Agote et al. 2020). However, compared to other vertebrates like fish (Kimura 2021; Luo et al. 2021), mammals (Baxter et al. 2019; Le et al. 2020) and birds (e.g., Haupaix et al. 2018; Bailleul et al. 2019; Yoshioka et al. 2021), little is known of the genetic basis of pigmentation and patterning in reptiles (Olsson et al. 2013). The dynamic process of cell-cell interactions leading to pattern formation may differ between zebrafish and reptiles. There is evidence to suggest that in reptiles it is melanophores, not iridophores (Singh et al. 2014), that are key in pattern formation (Kuriyama and Hasegawa 2017; Kuriyama et al. 2020).

Ball pythons (*Python regius*) represent an exceptional opportunity to study the genetic basis of colouration in reptiles (Irizarry and Bryden 2016). Over the last 30 years, ball pythons have emerged as one of the most popular snakes in the reptile trade (D'Cruze et al. 2020). Numerous Mendelian phenotypes affecting pigmentation and patterning ('colour morphs') have been discovered and are propagated in captivity (Figure 1). Commercial ball python breeders have used morphs thought to be monogenic ('base morphs') to selectively breed a much larger number of unique phenotypes ('designer morphs') using different combinations of base morphs. The Mendelian phenotypes in captivity can be characterized by changes to pattern, melanin-based pigmentation, non-melanin pigmentation (including structural colouration), or a combination of the former.

Figure 1. Mendelian phenotypes in ball pythons. The wild type (A) differs from the clown morph (B) in colour and pattern. The VPI axanthic retains wild type pattern but lacks pteridine pigments.

In human, mouse, and zebrafish models (George et al. 2016; Cooper et al. 2017; Brooks et al. 2018), mutations to genes that affect colour and patterning often show pleiotropy in other neural-crest-derived tissues and can cause disorders like COMMAD (coloboma, osteopetrosis, microphthalmia, macrocephaly, albinism, and deafness). Similarly, some ball python colour morphs are associated with neurological and morphological defects (Rose and Williams 2014; Hedley et al. 2018; Lambert et al. 2019; D'Cruze et al. 2020). Thus, ball pythons represent a promising model organism for advancing our knowledge of non-melanin pigmentation for the genetic study of epistasis, pleiotropy, and human pigmentation disorders (Baxter and Pavan 2013).

Here, we advance ball python colour morphs as a model system for reptile colouration, as it remains largely untapped by the scientific community (for recent exceptions, see Brown et al. 2022; Dao et al. 2022; Garcia-Elfring et al. 2023). The 'clown' colour morph is characterized as recessive and by changes to melanin pigmentation and patterning (Figure 1). The clown colour anomaly was discovered segregating in Africa and introduced into captivity on two occasions, in 1996 and 1997 (www.morphmarket.com/morphpedia/ball-pythons/clown/). In contrast to the piebald phenotype (Garcia-Elfring et al. 2023), characterized by white spotting on otherwise normal colour, the clown colour morph exhibits a reddish/golden brown colour with reduced lateral 'saddles.' They also display a dorsal stripe not found in the wild type. Thus, we use the clown phenotype to investigate genes not linked to white spotting but nonetheless important for pattern formation. We also study the 'VPI axanthic' colour morph in the descendants of a ball python discovered in Africa in 1990 (www.morphmarket.com/morphpedia/ball-pythons/axanthic-vpi/). Axanthic animals lack yellow or red xanthophore pigments (Cataneo 2015; Jablonski et al. 2014; Kolenda et al. 2017; Bortero et al. 2021) and have been overlooked in pigmentation research relative to more common colour aberrations like albino. Axanthic ball pythons can therefore be used to fill the gap in the literature and advance our understanding of pteridine pigmentation.

Here, we use whole-genome sequencing, population genetics, and annotation of DNA variants to identify candidate mutations for the clown and VPI axanthic colour morphs, using samples collected from the community of Canadian ball python breeders. Because Mendelian phenotypes are thought to be primarily caused by mutations to the protein-coding sequence of genes (Chong et al. 2015; Bamshad et al. 2019), we identified

variants predicted to affect gene product. For the clown colour morph, we found a candidate causal mutation in a gene regulating the production of melanin. For the axanthic phenotype, we found a mutation in a gene regulating the first step of pteridine pigment synthesis.

Methods and Materials

Sample collection, DNA extraction, and sequencing

We appealed to commercial breeders for ball python shed skin samples. These efforts led to the collection of 283 shed skin samples, including 38 samples for the clown colour morph and five for VPI axanthic (Table 1). We extracted and sequenced DNA from shed skins by following a standard three-day phenol-chloroform procedure. Extracted DNA was quantified and quality-checked on an Infinite® 200 Nanoquant (Tecan Group Ltd. Männedorf, Switzerland), and was normalized prior to combining into pools in equimolar amounts, including three control pools corresponding to samples with genetic stripe, champagne, ivory, and black pastel colour morphs. Whole-genome PCR-based libraries were prepared and 150 bp paired-end reads were sequenced using one lane of Illumina NovaSeq 6000. Library preparation and whole-genome sequencing was carried out at the McGill Genome Centre.

Bioinformatics

We first filtered raw sequence reads based on read quality (--quality-threshold 20) and length (--min-length 50) using the *trim-fastq.pl* program of *Popoolation* (Kofler et al. 2011). This process trims bases with a quality score below 20 and discards reads if the

final length is less than 50 bp. We then aligned processed reads to the Burmese python (*Python bivittatus*) reference genome Pmo2.0 (Castoe et al. 2013) using the program *bwa-mem* with default parameters (Li and Durbin 2009). The Burmese python is closest relative of the ball python for which an annotated draft genome sequence is available (Castoe et al. 2013). We converted SAM files to BAM format with *SAMtools* (Li et al. 2009) and removed reads with mapping quality below 20 (-q 20). *SAMtools* and *Popoolation2* (Kofler et al. 2011b) was subsequently used to produce mpileup and sync formats, respectively. We generated VCF files for case and control pools by using *bcftools* to call variants and filter (*vcfutils.pl varFilter*) for a minimum depth of five (-d 5) and mapping quality of 20 (-Q 20). The program *fst-sliding.pl* of *Popoolation2* was used to get SNP-specific F_{ST} estimates. We generated genome-wide F_{ST} estimates for the clown pool relative to two reference pools (n = 16, 3) known by pedigree analysis to not have the focal trait. Similarly, the VPI axanthic (n=5) pool was compared to three reference pools (n = 5, 15, and 26) known to not carry the allele for axanthism. Because we only analyzed DNA from snakes exhibiting the recessive clown and VPI axanthic traits, the causal variant for each phenotype is expected to be fixed in its respective pool and have a high F_{ST}. We therefore identified SNPs with $F_{ST} > 0.90$ across comparisons. We chose a threshold of 0.90 to allow for sample misidentification. We also generated F_{ST} estimates using the chromosome-length Burmese python reference assembly *Python_molurus_bivittatus-5.0.2_HiC* (Dudchenko et al. 2017; Dudchenko et al. 2018). We validated candidate loci by Sanger sequencing known heterozygotes, known homozygotes for the derived allele (snakes of the clown and VPI axanthic colour morph), and known homozygotes for the ancestral allele. Primers to amplify region with

candidate loci were designed using *Primer3* (Kõressaar and Remm 2007; Kõressaar et al. 2018).

Variant annotation and identification of polymorphisms and fixed differences

We annotated variants with the software *SnpEff* (Cingolani et al. 2012) using the annotated Burmese python reference genome (Python_molurus_bivittatus-5.0.2, assembly: GCF_000186305.1). This program predicts the effect size of mutations (SNPs and small indels) on a protein. Variants with an F_{ST} greater than 0.90 across comparisons mapping to a gene and affecting gene product (e.g., nonsense and missense mutations, indels) were considered candidate causal mutations for ball python colour morphs.

Results

Whole-genome sequencing and variant calling found 26,449,779 SNPs in 283 ball python samples. We obtained SNP-specific estimates of allele frequency difference between case and control pools, followed by variant annotation. In the clown colour morph samples, we found one missense variant (i.e., nonsynonymous SNP) in *melanocortin 1 receptor* (*mc1r*) showing high levels of differentiation (F_{ST} = 1.0) between case and control comparisons. This candidate mutation for the clown colour morph consists of a G to A change (c.149G>A), resulting in a serine being replaced by an asparagine on residue 50 (p.Ser50Asn; Table 1) of the Burmese python ortholog (Figure 2; Table 1). In the axanthic samples, we found a candidate SNP on chromosome 2 with F_{ST} = 1.0 between case and control comparisons and mapping to a protein-coding region. Variant annotation revealed a nonsense mutation in exon 5 of the gene *gtp cyclohydrolase 1* (*gtpch*) or

LOC103050242 (Figure 2; Table 1). The candidate SNP consists of a C to T nucleotide change (c.520C>T), resulting in a codon change from a CGA arginine to a TGA stop codon (p.Arg174*). Both mutations were confirmed by PCR amplification and Sanger sequencing.

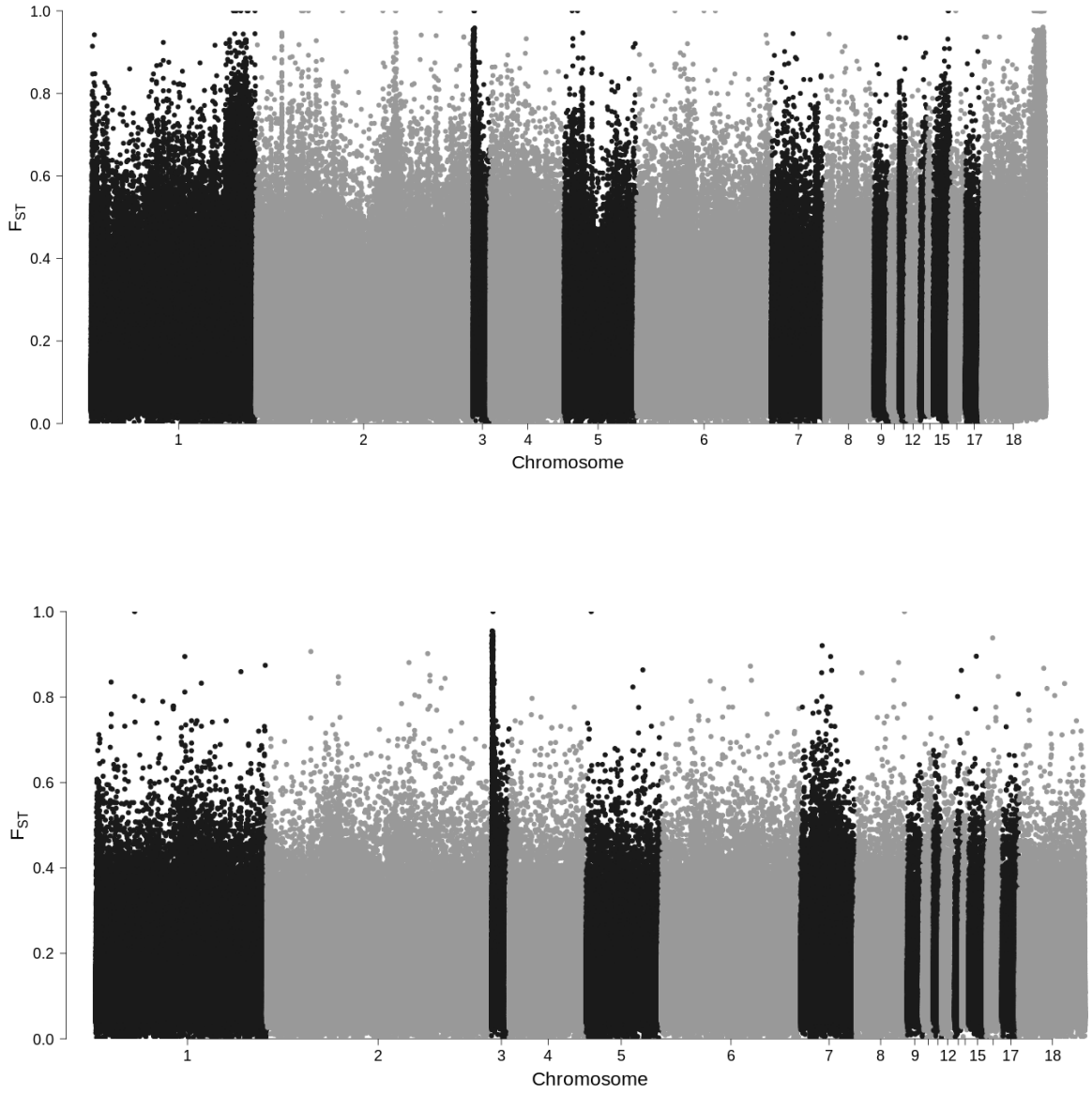


Figure 2. Manhattan plot of F_{ST} values for clown comparisons against reference pools (genetic stripe, top; champagne, bottom). A fixed SNP in the protein coding region of *mc1r* on chromosome 3 is the likely cause for the clown colour morph.

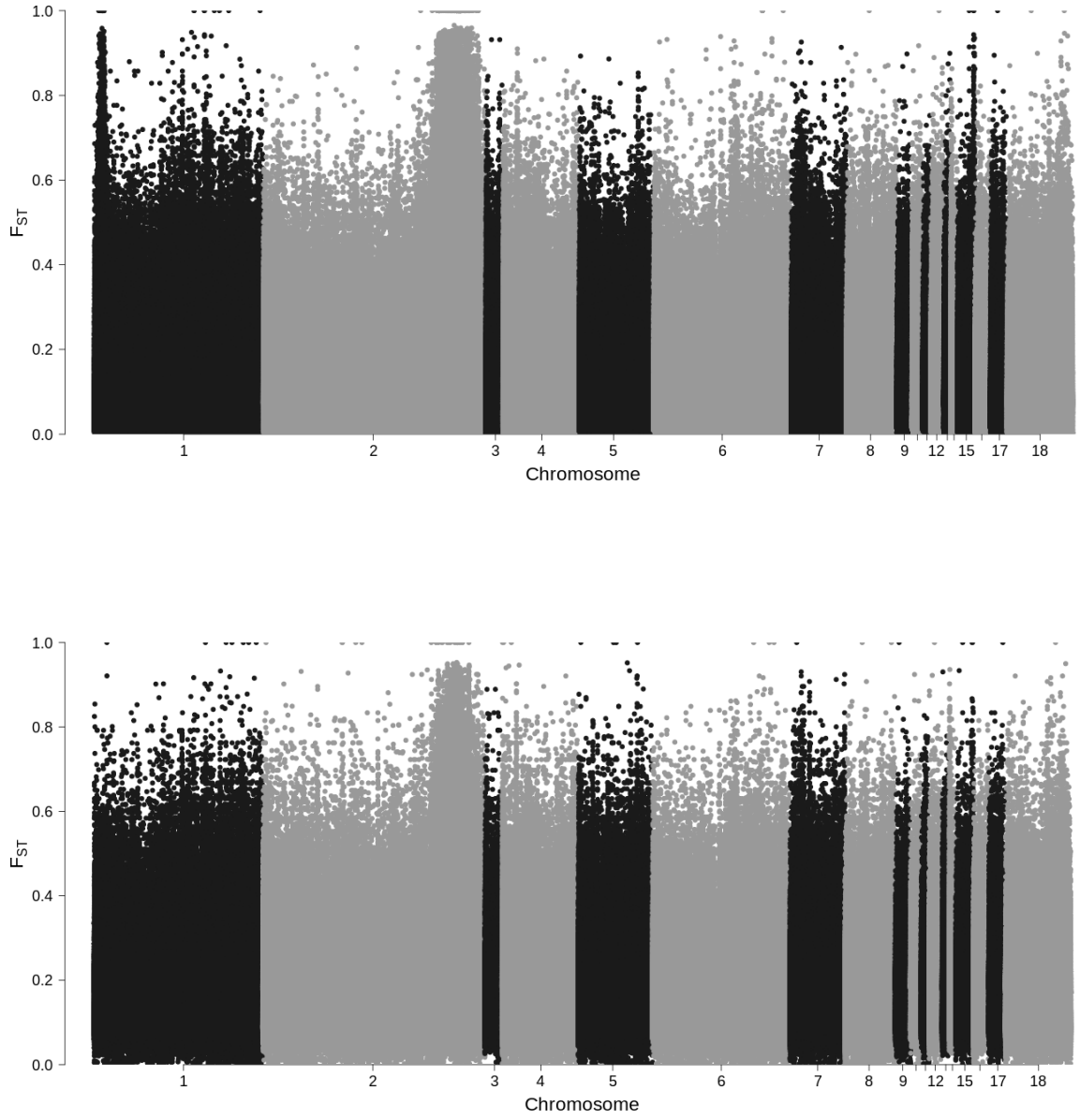


Figure 3. Manhattan plot of F_{ST} values for VPI axanthic comparisons against reference pools (ivory, top; black pastel, bottom). A nonsense mutation on crhomosome 2 (large peak) in the *gtp cyclohydrolase 1* (*gtpch*) is the likely cause for this axanothic phenotype.

Table 1. Candidate variants for the 'clown' and 'VPI axanothic' colour morphs.

Morph	Scaffold	Position	Nucleotide change	Amino acid change	Variant annotation	Gene
Clown	NW_006536545.1	62955	c.149G>A	p.Ser50Asn	missense	*mc1r*
VPI axanthic	NW_006535317.1	19442	c.520C>T	p.Arg174*	nonsense	*LOC103050242 (gtpch)*

Discussion

Ball pythons display a wide range of color and pattern variations or colour variations. Captive breeding programs have led to the development of many different morphs. This makes ball pythons a promising model species for studying the genetic basis of pigmentation, as researchers can compare the genomes of individuals with different colour variations to identify causal genes and mutations. Overall, the use of captive-bred ball pythons in pigmentation research is in its infancy, although it has potential to provide valuable insights into the genetic and developmental processes that control pigmentation.

Associated with the clown colour morph we found a missense mutation in *mc1r*, resulting in a serine to arginine change in the first (of seven) intramembrane domain (residue 50 of the Burmese python ortholog). Transmembrane domains maintain the structural integrity of *mc1r*, are important for signalling and are highly conserved (Rosenblum et al. 2004; Rosenblum et al. 2010). Mutations to the intramembrane domains have been associated with changes to colour in mammals (Matsumoto et al. 2020), birds (Theron et al. 2001), fish (Gross et al. 2009), and reptiles (e.g., Rosenblum et al. 2004; Nunes et al. 2011; Hauser et al. 2022), although not in snakes. In humans, variation in *mc1r* is associated red hair (Zorina-Lichtenwalter et al. 2019), whereas other variants affecting skin tone and hair colour have introgressed from neanderthals into

Eurasians (Ding et al 2014). More than 20 variants on *mc1r* that affect pigmentation have been discovered in humans (reviewed by Dannemann and Kelso 2017), some of which are strongly associated with cancer risk (Bastiaens et al. 2001; Smit et al. 2020). Overall, variation in *mc1r* tends to be associated with changes along the light-to-dark pigmentation continuum and not patterning.

However, the clown colour morph exhibits altered patterning, with the dark lateral 'saddles' being reduced to a dorsal stripe. Most of what is known about the genetics and development of pattern comes from the zebrafish where stripe formation has been studied in detail (Singh and Nüsslein-Volhard 2015). In zebrafish, iridophores and xanthophores arise first and their interactions with melanophores are key for defining the yellow interstripe and blue stripe regions (Frohnhöfer et al. 2013; Singh et al. 2014; Singh and Nüsslein-Volhard 2015; Gur et al. 2020). If zebrafish lack any of the three main types of chromatophores, an unstriped or irregularly striped pattern emerges. The dynamics of stripe formation is not well understood in reptiles, but in *Plestiodon* lizards melanophores appear first, before xanthophores and iridophores, to form a stripe pattern. Work on *Elaphe* snakes also shows that spatial concentration of melanophores distinguishes the striped from nonstriped morph and is laid during embryonic development, suggesting melanophores are the key chromatophore in pattern formation (Kuriyama et al. 2013; Murakami et al. 2016). Kuriyama et al (2013) hypothesize that variants controlling stripe formation also control the distribution and density of melanophores. The density and distribution of melanophores may be regulated by *mc1r*, whose expression improves sensitivity to a-MSH, which promotes the differentiation of melanoblasts into melanophores (Matsuda et al. 2018). Loss-of-function variants to *mc1r* resulted in a loss

of melanophores in common carp (Mandal et al. 2020). Consistent with this hypothesis, our results showed that a missense variant (p.Ser50Asn) in *mc1r* is linked to melanin pigmentation and pattern formation in ball pythons.

Associated with the VPI axanthic colour morph we found a mutation in the gene coding for the first and rate-limiting enzyme involved in the biosynthesis of tetrahydrobiopterin (BH4). The axanthic phenotype is rare in nature and has been documented only a few times, usually amphibian and snake species (e.g., Jablonski et al. 2014; Borteiro et al. 2021; Schluckebier et al. 2022). As such, the genetic basis of axanthism has been much less featured in pigmentation studies. In addition to being rare in nature, model systems for the study of axanthism are limited. The axolotl has served as a model organism, although not ideal as axanthic axolotls have poor survival rates (Lyerla et al. 1971; Frost et al. 1986; Masselinket and Tanaka 2021).

In general, the pteridine pathway has been little studied in vertebrates outside cell cultures and zebrafish (Ziegler 2003). We found that variation in *gtp cyclohydrolase 1* (*gtpch*) results in axanthism in ball pythons. *gtpch* codes for the first enzyme involved in the conversion of GTP to tetrahydrobiopterin (BH4) and other pteridine pigments, including sepiapterin, 7-oxobiopterin, and 2,4,7-trioxopteridine (Masada et al. 1990; Ziegler 2003). Tetrahydrobiopterin is also an essential cofactor in the synthesis of tyrosine, dopamine, and serotonin. Mutations to this gene in humans are known to cause neurological conditions like autosomal dominant dystonia and recessive hyperphenylalaninemia (Nagatsu and Ichinose 1996; Hirano and Ueno 1999)

All captive ball pythons exhibiting the VPI axanthic colour morph are thought to be descendants of a single axanthic individual found in Africa in the early 1990s (Ref).

We found that a change in exon 5 of *gtpch* from an arginine to a stop codon is the likely causal basis for the VPI axanthic colour morph. Nonsense mutations play an important role in Mendelian phenotypes, accounting for approximately 11% of mutations (i.e., gene lesions) causing human inherited diseases and approximately 20% of disease-associated SNPs in gene coding regions (http://www.hgmd.org). CGA codons for arginine can be 'hazardous' for genetic function, and thus fitness, given they are easily turned into TGA stop codons (UGA in mRNA) via epigenetic-mediated mutagenesis or RNA cytosine deamination (Romanov and Sukhoverov 2017).

In contrast to our findings of the clown colour morph, where variation in a gene regulating melanin production (*mc1r*) alters pattern, mutations to *gtpch* that regulate pteridine pigments do not have an effect on pattern. Wu et al (2021) recently found a correlation between increased expression of *gtpch* and yellow colouration in fish, and Rodríguez et al (2020) found differentially expressed pteridine genes associated with red and green colouration in frogs. However, to our knowledge, the nonsense variant in *gtpch* represents the first natural axanthism-causing mutation in vertebrates. We thus provide a new model system for the study of axanthism and help fill a gap in the pigmentation literature on this colour anomaly.

Conclusion

Sequence variation in *melanocortin 1 receptor* (*mc1r)* in ball pythons is associated with melanin and, notably, pattern changes. This result is consistent with the view that melanophores have an important role in pattern formation in snakes, which contrasts with traditional model systems like zebrafish. Mutations to *gtp cyclohydrolase 1 (gtpch)* cause axanthism but do not seem to affect neurological function in snakes as it does in humans.

Our work provides a new reptile model system for the study of pattern formation and pteridine pigmentation.

Bridging text

In Chapter 1 and 2 we took advantage of artificial selection imposed on reptiles in captivity to make a link between phenotype and genotype. In the following two chapters we study how natural selection affects patterns of genetic variation across geologic (Chapter 3) and seasonal timescales (Chapter 4).

In Chapter 3, we investigate the evolutionary history of post-glacial populations of threespine stickleback in Eastern Canada and look for evidence of selection. These populations are of particular interest to the evolutionary biologist as they lack the typical reduced lateral plate phenotype observed in freshwater populations in the west coast.

CHAPTER 3

Genomics of freshwater adaptation in three-spine stickleback from eastern Canada

*Alan Garcia-Elfring, *¶Antoine Paccard, and *Rowan D. H. Barrett

*Redpath Museum and Department of Biology, McGill University, Montreal, QC, Canada

¶McGill University Genome Center, McGill University, Montreal, QC, Canada

Abstract

The threespine stickleback is an important species in the study of parallel evolution. However, the evidence for parallel evolution, at both genetic and phenotypic levels, has largely come from the west coast of North America and parts of southern Europe. Here, we use RAD-sequencing of pooled samples to study stickleback marine-freshwater differentiation in Atlantic Canada and test for evidence of parallel evolution. Our results show considerable heterogeneity in differentiation among populations, with some freshwater populations showing high degree of allele frequency divergence relative to marine ones. The strongest candidates for selection were two SNPs near dopamine receptors (*Drd24a* and *Drd2l*). Gene ontology (GO) analysis of candidate genes shows dopamine neurotransmitter receptor activity as one of the most enriched gene functions, along with nervous system development. Our study suggests a more complex evolutionary history of stickleback populations in the Atlantic than previously thought, with multiple colonization events or ongoing gene flow. In contrast to western populations where selection on genes for morphological traits (spines and plates) has been most noticeable, our results suggest adaptation to freshwater has occurred in eastern populations through changes in behaviour or osmoregulation.

Introduction

Knowledge of the genetic basis of adaptation is central to understanding the mechanisms driving the diversification of species. Threespine stickleback fish (*Gasterosteus aculeatus*) are native to the oceans of the northern hemisphere and have been an important species in studies of ecology and evolution for almost a century (Bertin 1925; Heuts 1947a, b; Moodie and Reimchen 1976). Over the last 12,000 years, with the melting of glaciers after the last glacial maximum, marine populations of threespine stickleback have colonized freshwater lakes and rivers across the northern hemisphere. Many freshwater populations evolved similar low-plated morph phenotypes, particularly in western North America (Baumgartner and Bell 1984; Colosimo et al. 2005; Hohenlohe et al. 2010). It is thought that alleles that confer higher fitness in freshwater are found at low frequencies in marine populations, with bidirectional gene flow maintaining the genetic variation needed to seed new freshwater habitats (the 'transporter hypothesis', Schluter and Conte 2009). Today, the threespine stickleback is considered an exemplar of parallel evolution.

However, even early in the development of stickleback as a model species, it was noted that the remarkable parallel evolution among freshwater populations might be a local phenomenon and limited to western North America (i.e., eastern Pacific) and southern Europe (Hagen and Moodie 1982). Taking armour phenotype as an example, in regions of North America, Northern Europe, the Baltic Sea, and eastern Asia the completely plated morph that is common in marine environments dominates in local lakes (Münzing 1963; Penczak 1965; Hagen and Moodie 1982; Mäkinen et al. 2008; Raeymaerker et al. 2014; Ferchaud and Hansen 2016; Yamasaki et al. 2019). This lack of

differentiation between marine and freshwater populations was also noted at the genetic level using allozymes (Rafiński et al. 1989). More recently, a genomic study (Fang et al. 2020a) that investigated stickleback across their distribution showed that high marine-freshwater (M-FW) differentiation is largely confined to the eastern Pacific, with Atlantic populations showing lower levels of genomic differentiation. In contrast to populations in British Columbia, many freshwater populations in Eastern Canada are dominated by full-plated ecotypes (Hagen and Moodie 1982; Haines et al. 2022), with certain populations showing distinct phenotypes such as … (Scott et al. 2023). However, genomic studies in this region are few (e.g., Fang et al. 2020).

One explanation for why parallel evolution and M-FW differentiation appears to be lower outside of the eastern Pacific is stochastic loss of freshwater-adapted alleles. Threespine stickleback have persisted the Pacific Ocean for 26 million years (Matschiner et al. 2011; Betancur-R et al. 2015). Only much more recently, in the late Pleistocene (36.9–346.5 thousand years ago), did stickleback colonize the western Pacific and Atlantic basin (Orti et al. 1994; Fang et al. 2018; Fang et al 2020). Atlantic populations also show less genetic diversity than those in the Pacific (Fang et al. 2020), a consequence of range expansion. However, many Atlantic populations analyzed thus far are located along the coast of northern Europe (Liu et al. 2016).

We study marine and freshwater threespine stickleback from eastern Canada, in the provinces of Nova Scotia and Newfoundland where the full-plated phenotypes can be found at high frequencies in freshwater habitats (Hagen and Moodie 1982). We use genome-wide SNPs to characterize M-FW differentiation and look for evidence of parallel natural selection. We seek to gain insight into the evolutionary history and the

genetic basis of freshwater adaptation of stickleback populations from eastern North America.

Methods and Materials

Field sampling and DNA extraction

We used minnow traps and beach seines to collect 30 adult stickleback (> 30 mm in length) from nine populations in Nova Scotia and Newfoundland, Canada (Figure 1). In Nova Scotia, we sampled three freshwater sites, including two lakes (Pomquet Lake Lake Ainslie) and a stream (Black River), and two marine populations (Antigonish Landing and Porper Pond). In Newfoundland, we sampled two freshwater (Pinchgut Lake and Blue Pond) and two marine populations (Cooks Brook and Humber Arm). We followed a phenol-chloroform procedure for DNA extraction.

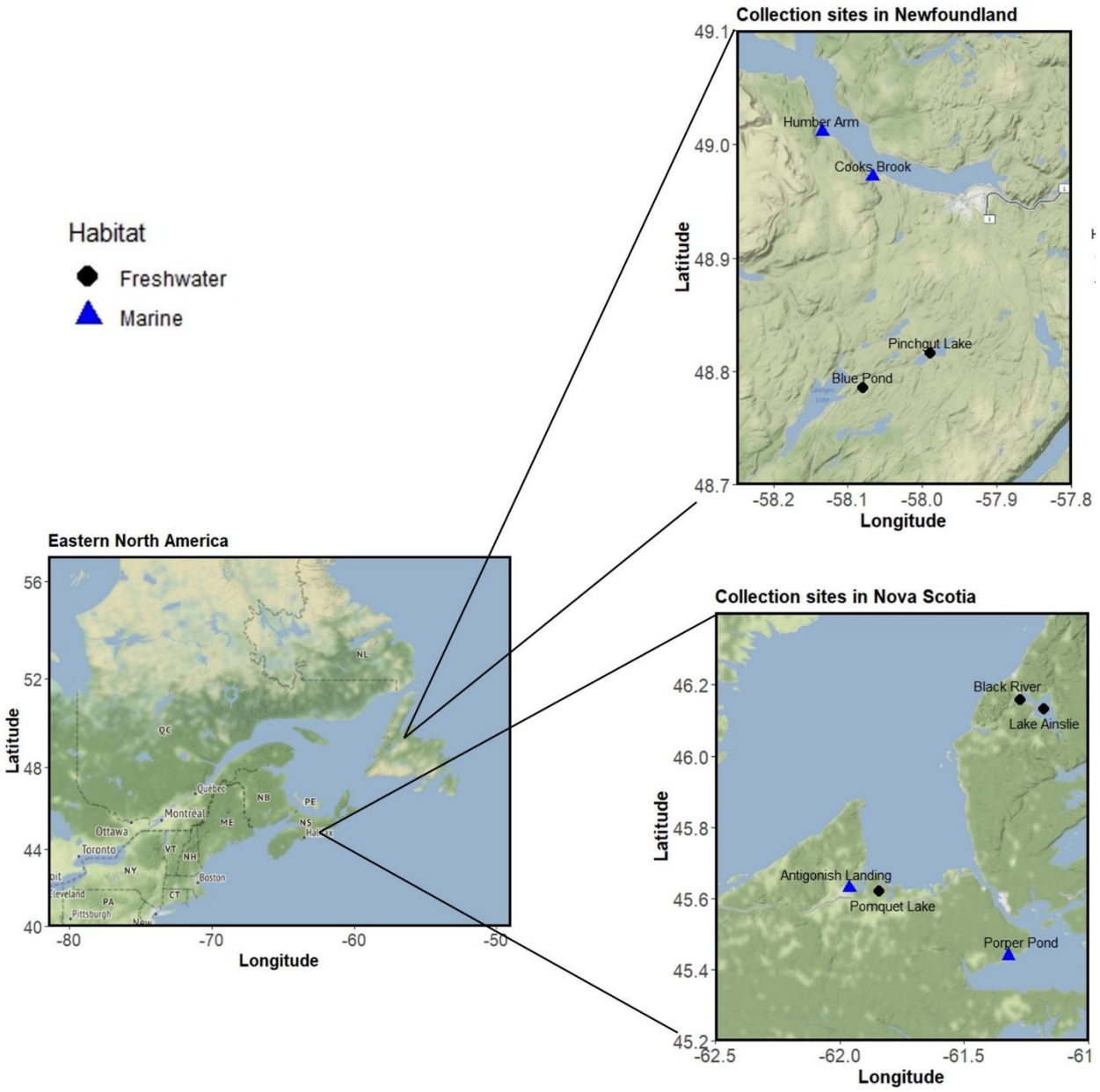


Figure 1. Nine collection sites in Eastern Canada. Stickleback were collected (n=30) from freshwater (black diamond) and marine sites (blue triangle) in Nova Scotia and Newfoundland.

DNA sequencing of pooled samples

For DNA quantification, we used a Picogreen® ds DNA assay (Thermo Fisher Scientific, Waltham, USA) and an Infinite® 200 Nanoquant (Tecan Group Ltd. Männedorf, Switzerland). We then normalized the concentration of dsDNA and pooled according to sampling location. Thus, we created nine pools consisting of 30 individuals each. RAD-seq libraries were prepared and sent for sequencing the McGill University and Genome Quebec Innovation Center, Montreal, Canada. Sequencing was carried out on one lane of Illumina HiSeq2500 which produced 125bp paired-end reads.

Bioinformatics

We used the *trim-fastq.pl* script of *Popoolation* (Kofler et al. 2011a) to process raw reads, filtering based on read quality (--quality-threshold 20) and length (--min-length 50). We then mapped the processed reads to the stickleback reference genome (BROADS S1) with the program *Bowtie2* (Langmead and Salzberg 2012), using the --end-to-end mapping option. We then used *SAMtools* (Li et al. 2009) to convert the output SAM files to BAM format and subsequently removed reads with mapping quality below 20 (samtools view -q 20). We generated an mpileup file (samtools mpilep -B) and converted the mpileup file to the synchronized (sync) format using *Popoolation2* (Kofler et al. 2011b), which was used for downstream analyses.

Estimates allele frequency differentiation

We estimated population differentiation using F_{ST}, as formulated by Hartl and Clark (1997), using the *fst-sliding.pl* script of *Popoolation2*. We applied the following parameters: --min-count 2, --min-coverage 5, --max-coverage 500, --min-covered-

fraction 0, --window-size 1, --step-size 1, --pool-size 30:30:30:30:30:30:30:30:30:30, --suppress-noninformative. We only analyzed genomic regions assembled at the chromosome level (i.e., scaffolds excluded). We categorized F_{ST} outliers as SNPS in the top 5% of the F_{ST} distribution of a given M-FW comparison. We then looked for overlap of F_{ST} outliers among M-FW population pairs. We quantified F_{ST} outlier overlap among M-FW population pairs using a custom bash script. Loci showing relatively large changes in allele frequency (top 5%) in two or more M-FW comparisons within provinces and across Nova Scotia and Newfoundland were taken as three subsets. We then considered 'F_{ST} parallel outliers' SNPs that overlapped among the three sets.

Identification of candidate genes and analysis of molecular function

We applied a custom bash script to map outliers to or near protein-coding genes in the reference genome. We limited our search to a set of 14,252 protein-coding gene annotations with attributes "ID=gene" and "biotype=protein_coding". To gain insights into the traits under selection, we analyzed candidate genes for enrichment of molecular functions using the software *ShinyGO* 0.77 (Ge et al. 2020). This program, which is implemented as a graphical web application, uses a large annotation database derived from Ensembl and STRING-db to link lists of genes to functional categories, like gene ontology (GO). We analyzed whether candidate genes were enriched for GO terms (biological process, cellular component, and molecular function), using the FDR threshold of 10% and keeping the 20 most significant hits. Because our data is based on a reduced representation of the genome, there are likely cases where a causal mutation is not sampled but rather a SNP in linkage disequilibrium to it. We thus analyzed significant enrichment of GO terms for genes with outlier loci or within 5kb of one. We compared

our list of candidate genes to the set of 14,252 genes from which our candidate genes were sampled and used the 'best matching species' option.

Results

Data sets and Genomic differentiation among populations

After processing raw RAD-seq data, we identified 18,582 SNPs across all 36 population comparisons. We find that median genome-wide F_{ST} varies widely, ranging from 0.0118 to 0.0824 (Table S1; Figure 2). Considerable genomic heterogeneity is found among M-FW comparisons (Figure 3) and between populations from the similar environments (Figure 4). Notably, both the highest and lowest levels of genomic differentiation are between freshwater populations. The least differentiated populations are Pomquet Lake (Nova Scotia) and Pinchgut Lake (Newfoundland), while the most differentiated correspond to Lake Ainslie (Nova Scotia) and Blue Pond (Newfoundland; Figure 2, Table C1).

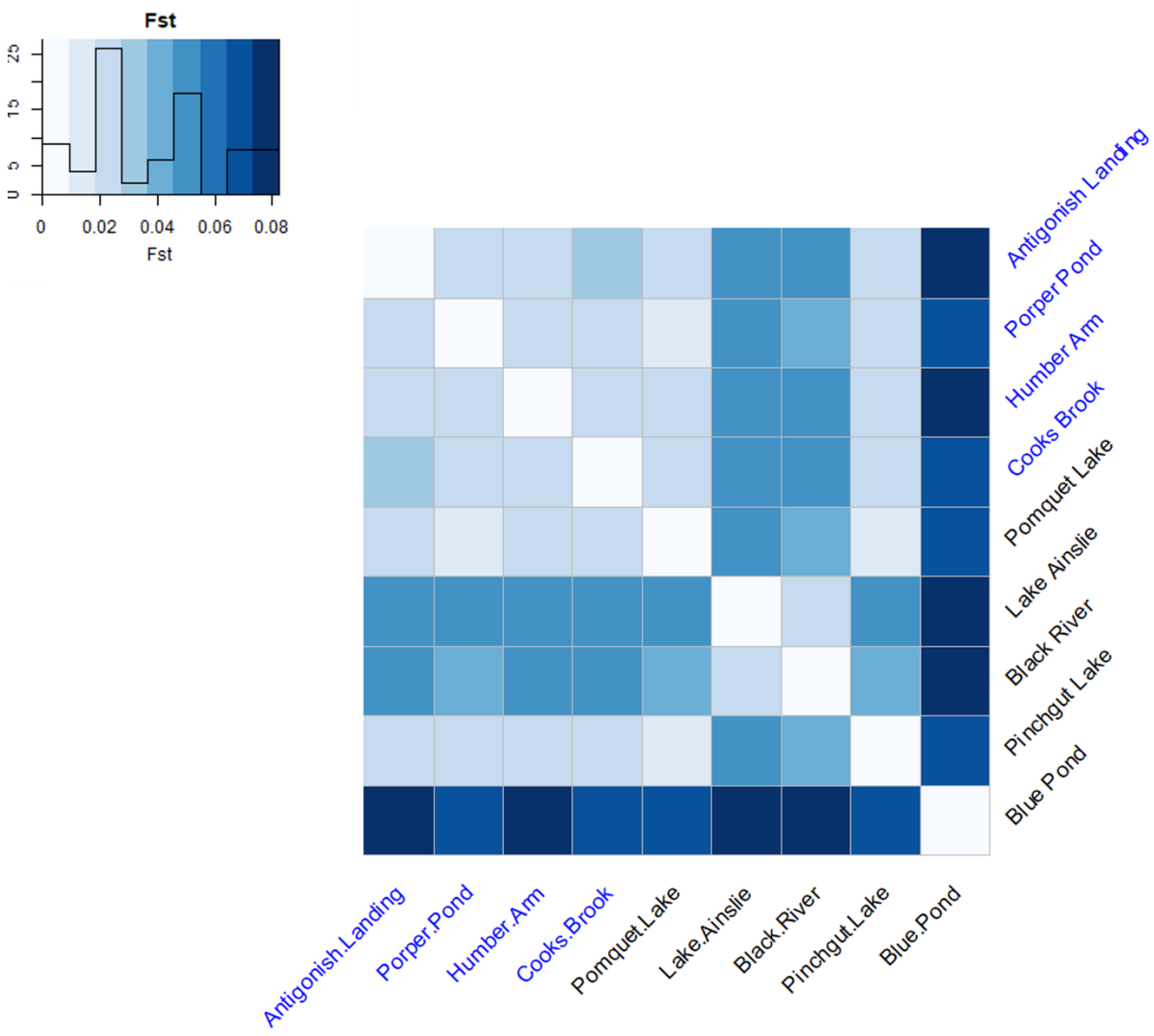
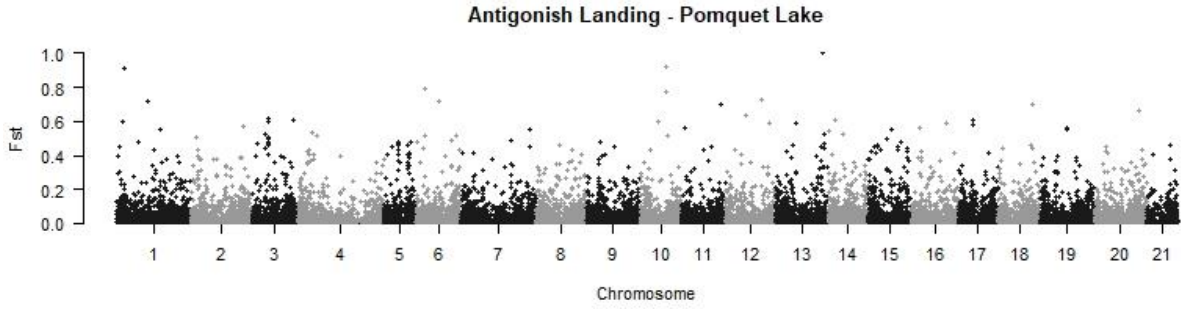
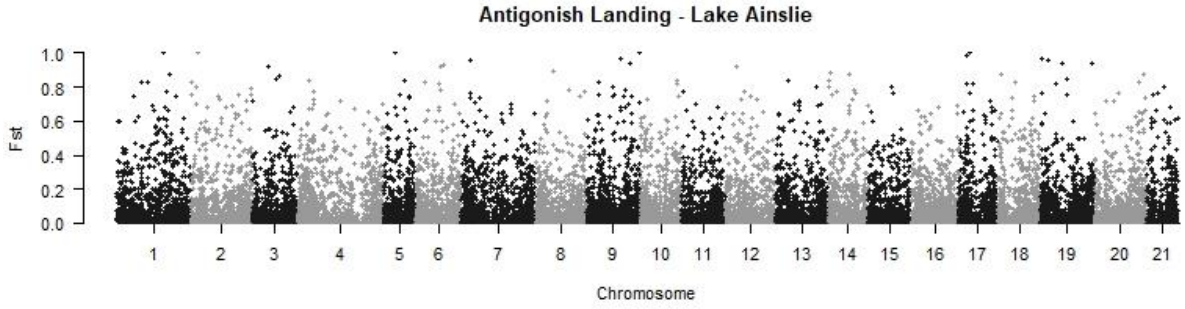


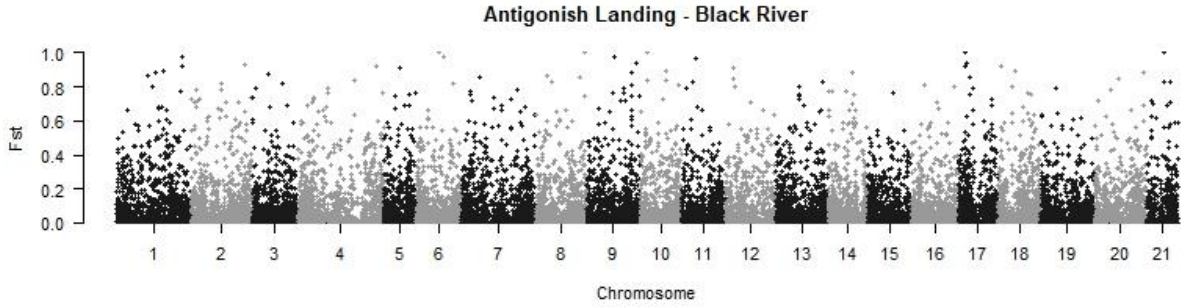
Figure 2. Heat plot depicting average (median) pairwise F_{ST} with histogram of F_{ST} values. Freshwater sites are labelled black, marine are labelled blue.

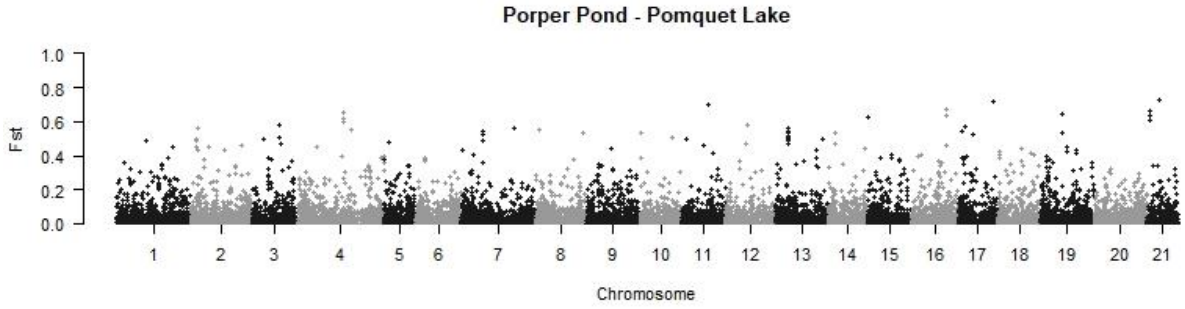
After excluding M-M and FW-FW comparisons, we obtained 25,236 informative SNPs, variable across all 20 M-FW population comparisons. We used this SNP set to look for genomic regions potentially involved in parallel adaptation to freshwater and found 2,100 SNPs as F_{ST} outliers in two or more M-FW comparisons within Nova Scotia (19 SNPs across all six comparisons; Figure C2a). In the four M-FW comparisons within

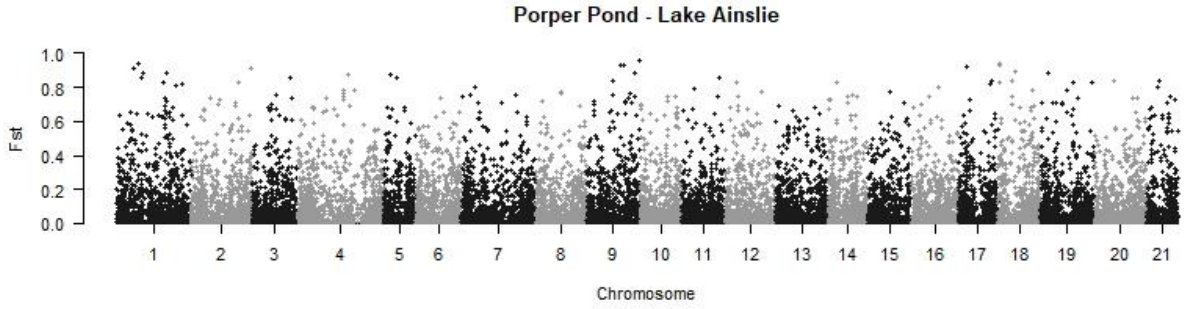
Newfoundland there are 1,306 SNPs in two or more comparisons (Figure C2b). When comparing interprovincial differentiation among 10 M-FW population comparisons, our results show 3,422 SNPs are F_{ST} outliers in at least two comparisons; three SNPs in nine of the 10 comparisons (Figure C2c). In the top 5% of the F_{ST} distribution we find an overlap of 279 SNPs (referred to as 'F_{ST} outliers' hereon) among the three sets M-FW comparisons (NS, NL, and NS-NF), meaning these SNPs are F_{ST} outliers in at least 6/20 M-FW comparisons, representing 1.1 % of SNP data set. We found 133 different genes contain or are within 5 kb of any of the 279 F_{ST} outliers (Table C2), with 68 genes containing at least one outlier. Notably, two SNPs (chromosome VI position 16449327 and chromosome XX position 14009648) fall in the top 5% of the F_{ST} distribution in 19/20 M-FW comparisons: 6/6 M-FW comparisons in Nova Scotia, 4/4 comparisons in Newfoundland, and 9/10 interprovincial M-FW comparisons (not an outlier in the Porper Pond and Pomquet Lake comparison). The latter of these two SNPs is 920 bp downstream from a gene coding for a dopamine receptor, *drd2l* (chromosome XX 14002115-14008728).

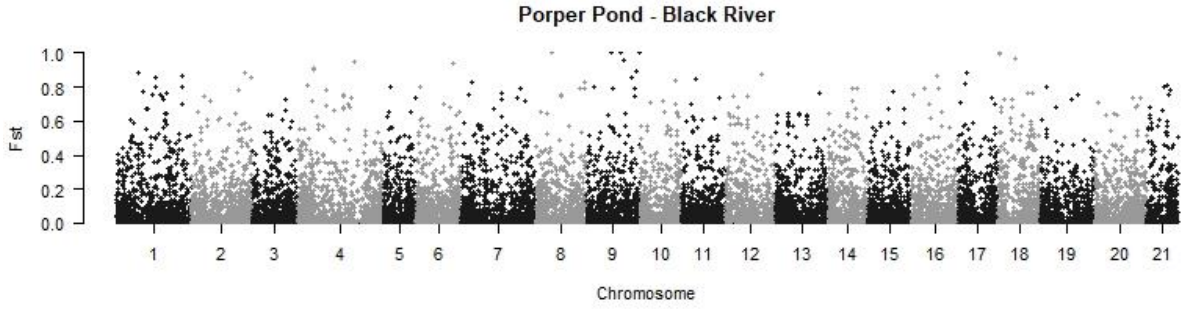

Antigonish Landing - Pomquet Lake
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

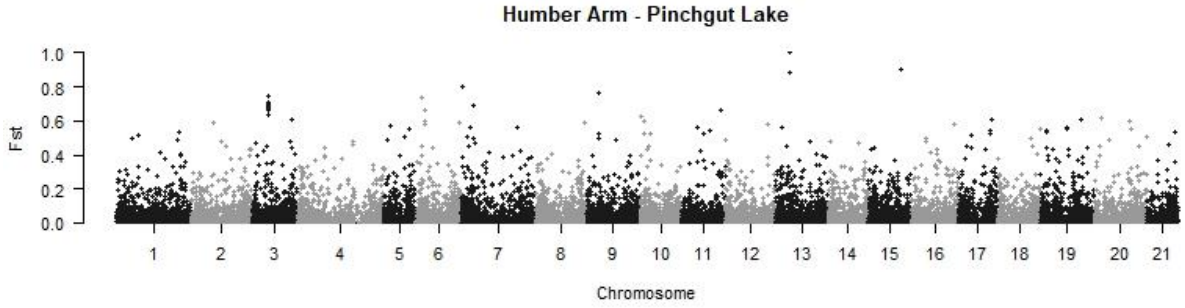

Antigonish Landing - Lake Ainslie
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

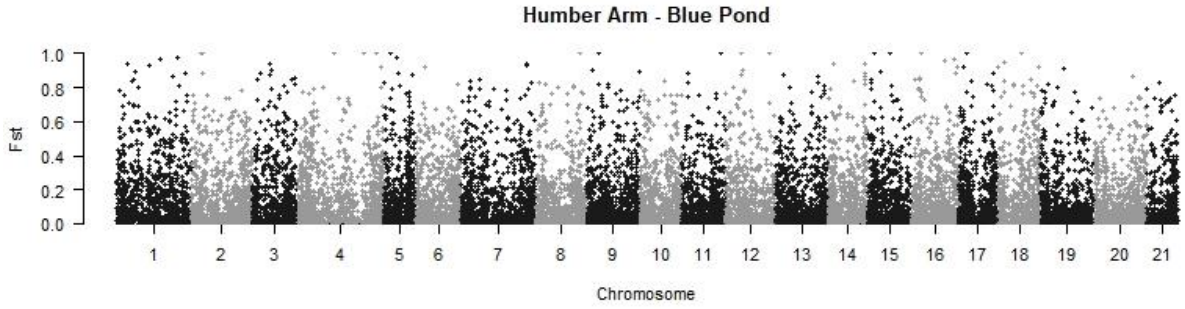

Antigonish Landing - Black River
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

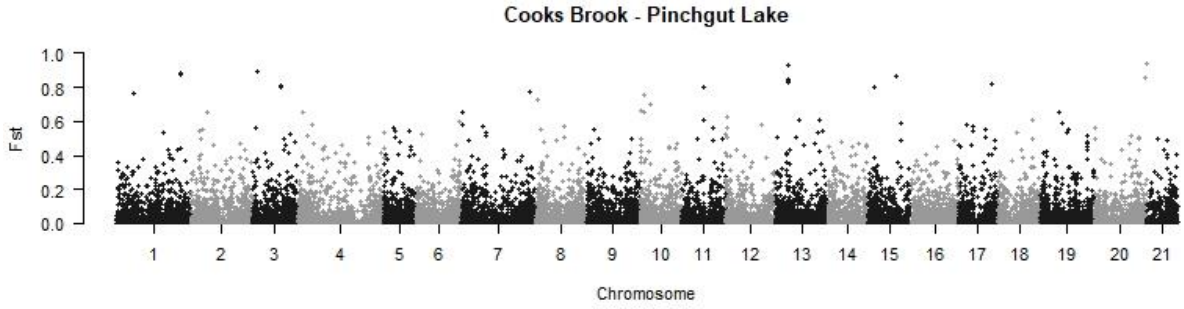

Porper Pond - Pomquet Lake
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

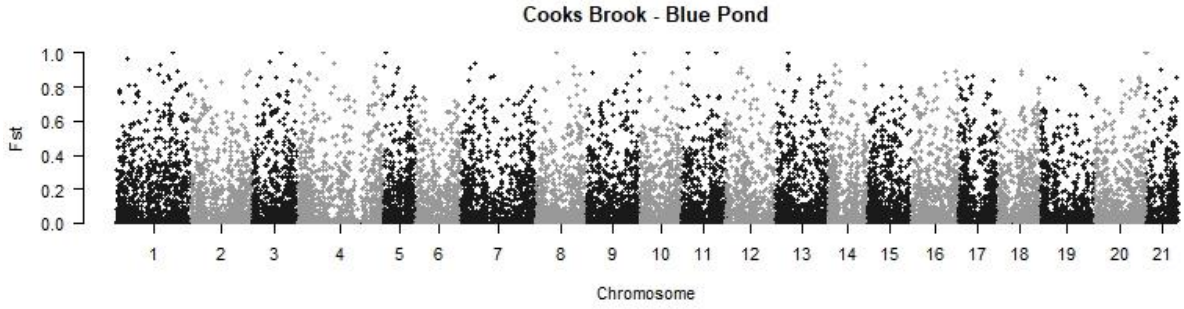

Porper Pond - Lake Ainslie
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

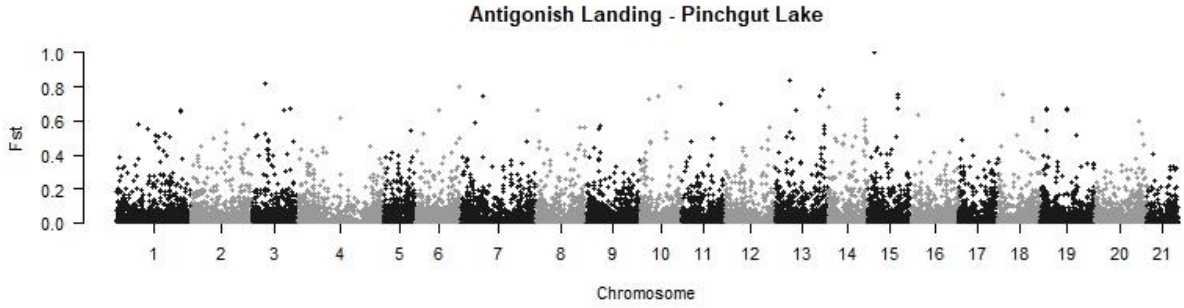

Porper Pond - Black River
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

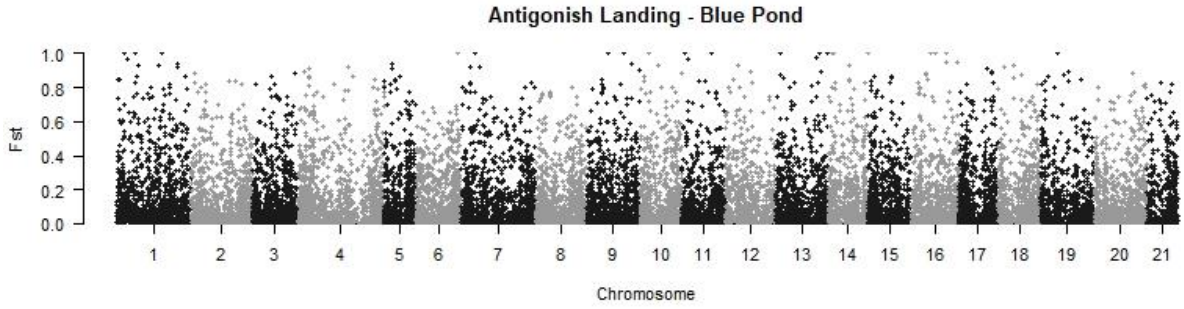

Humber Arm - Pinchgut Lake
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

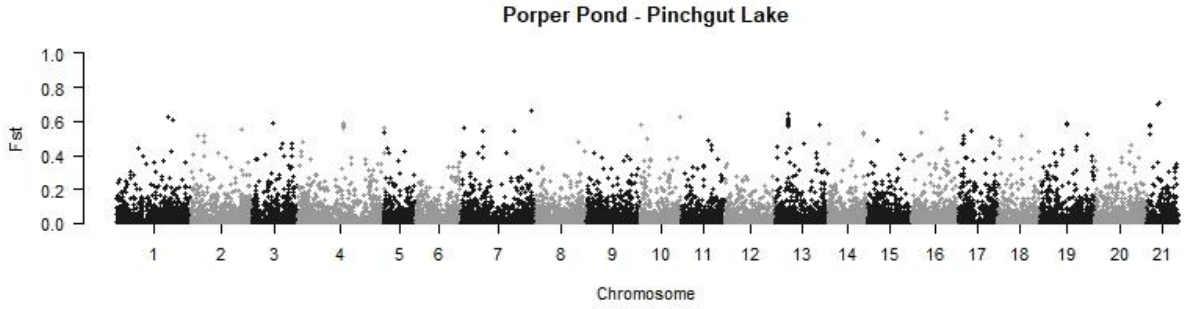

Humber Arm - Blue Pond
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

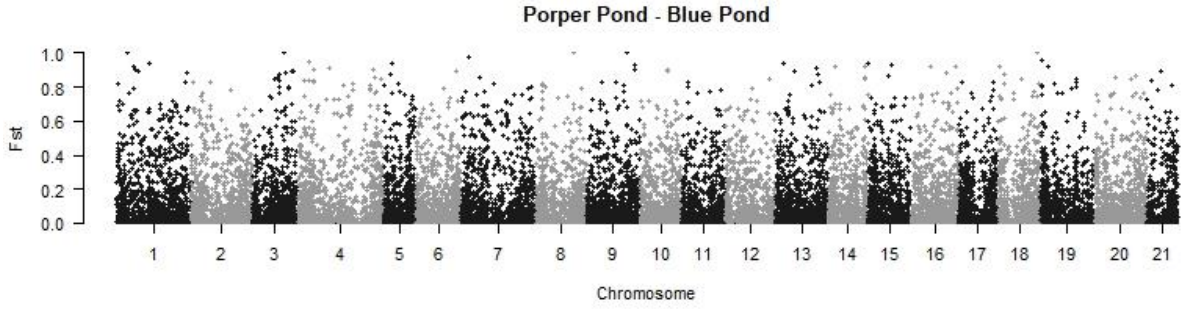

Cooks Brook - Pinchgut Lake
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

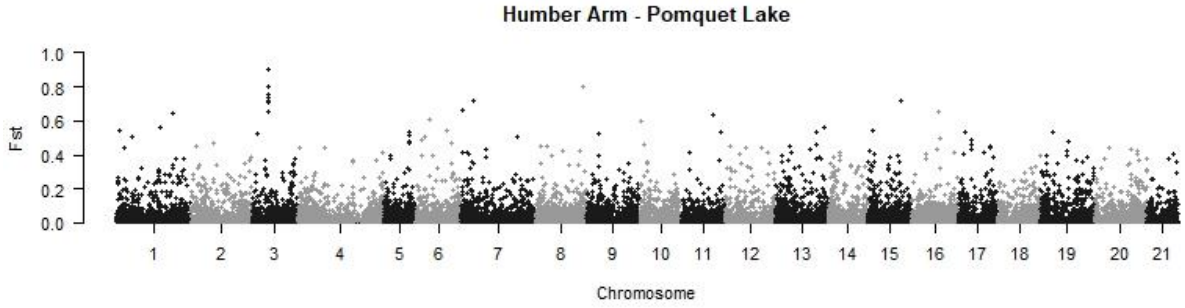

Cooks Brook - Blue Pond
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

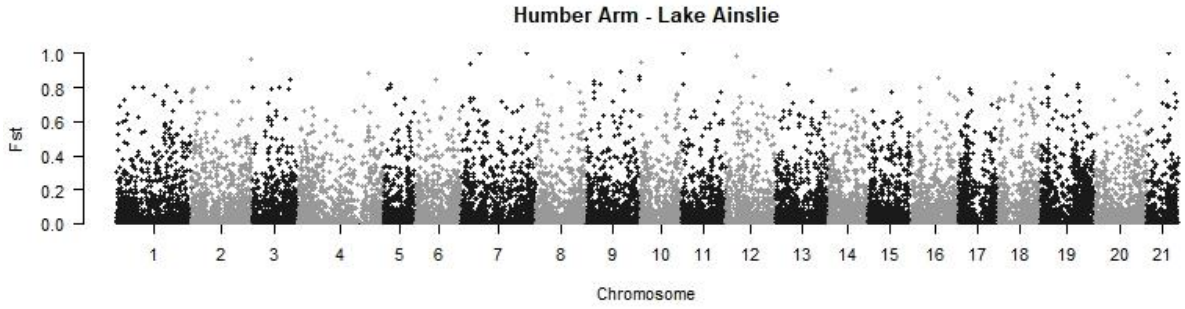

Antigonish Landing - Pinchgut Lake
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome


Antigonish Landing - Blue Pond
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome


Porper Pond - Pinchgut Lake
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome


Porper Pond - Blue Pond
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome


Humber Arm - Pomquet Lake
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome


Humber Arm - Lake Ainslie
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

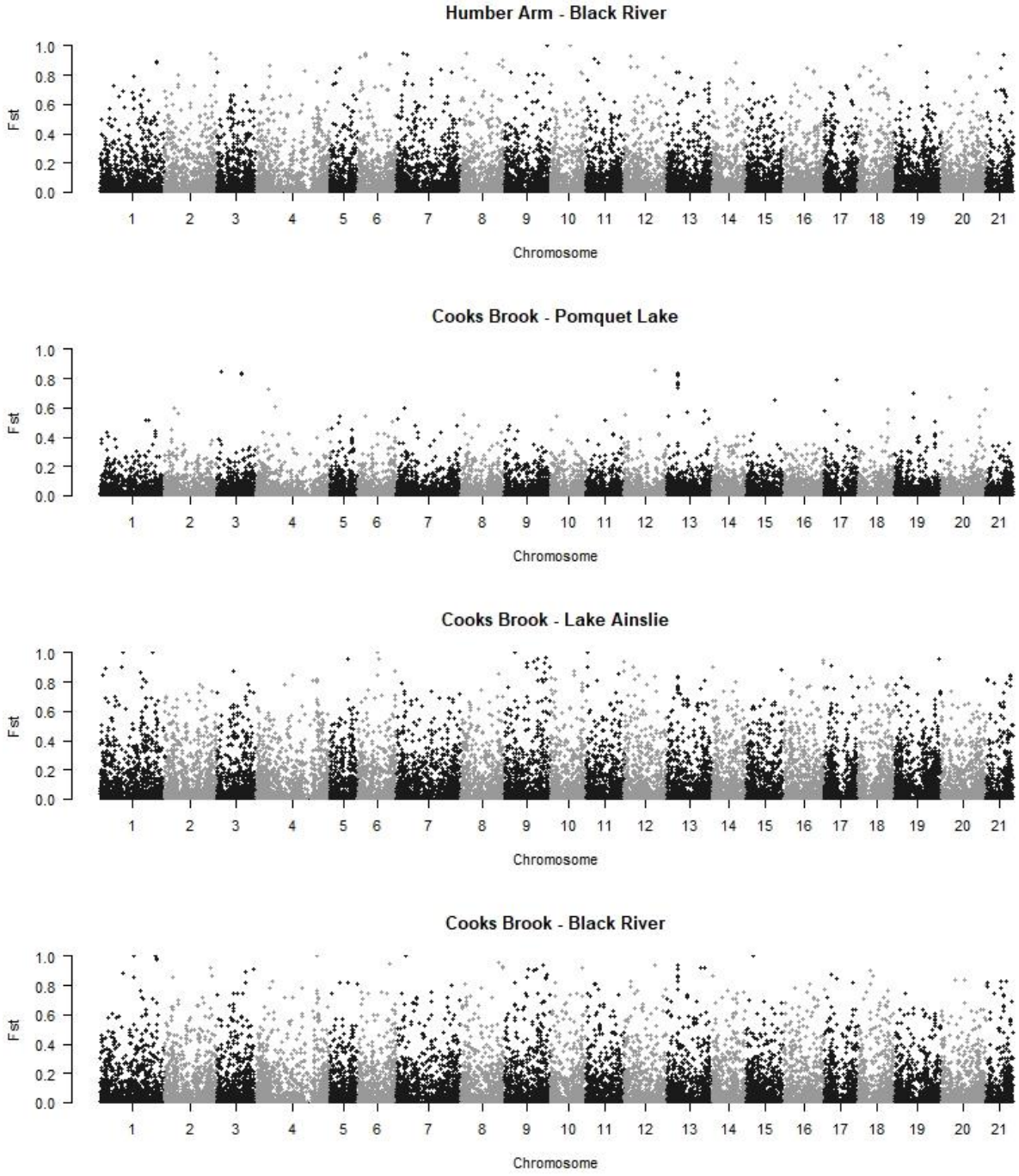
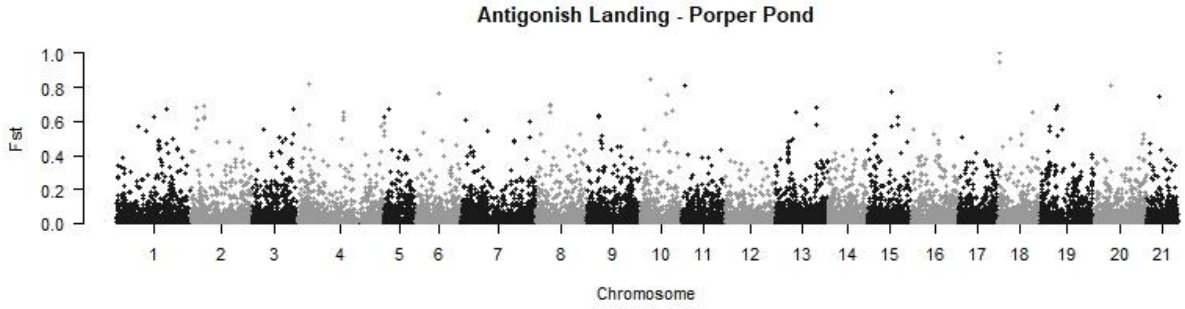
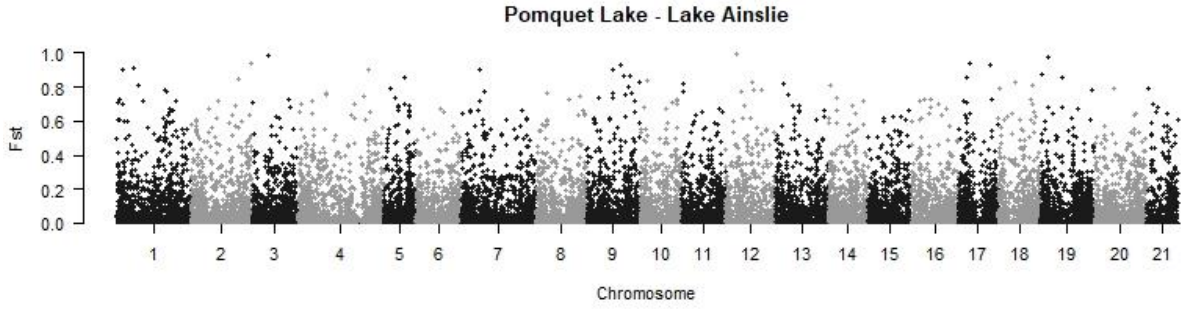
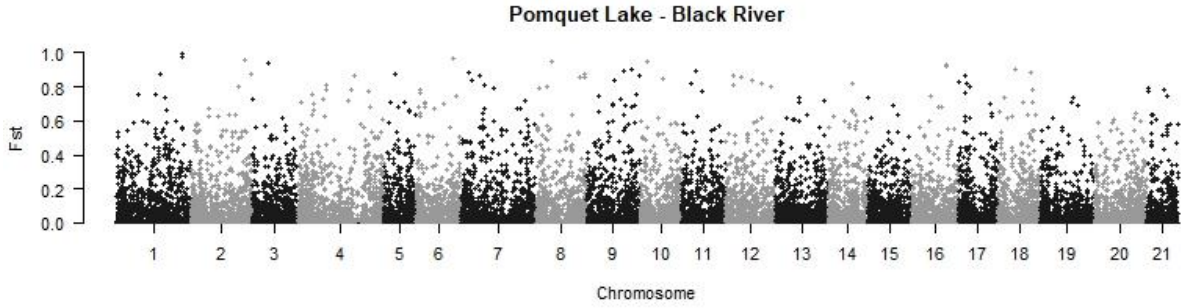
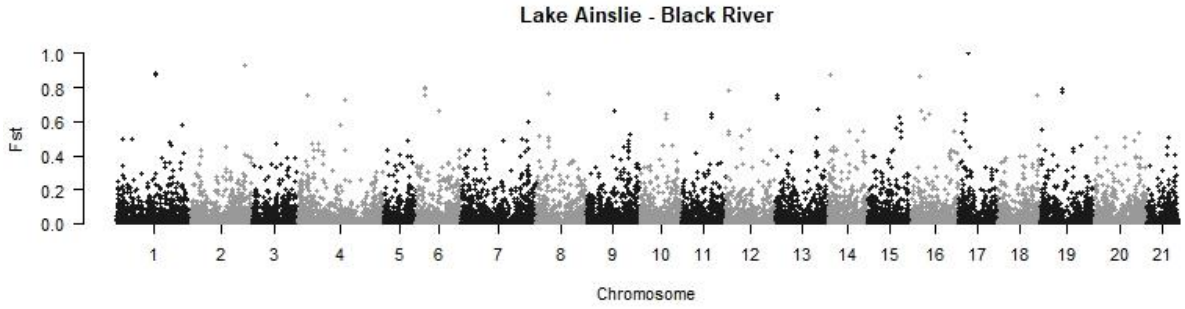


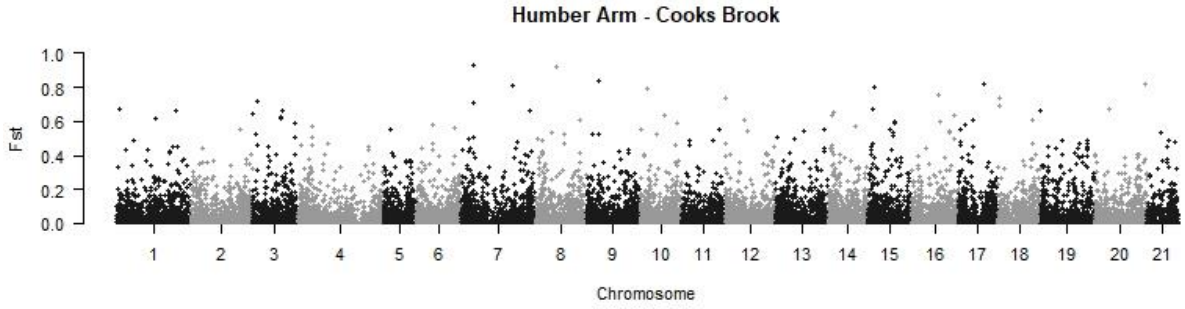
Figure 3. Manhattan plot of F_{ST} values showing heterogeneity in 20 marine-freshwater (M-FW) comparisons.

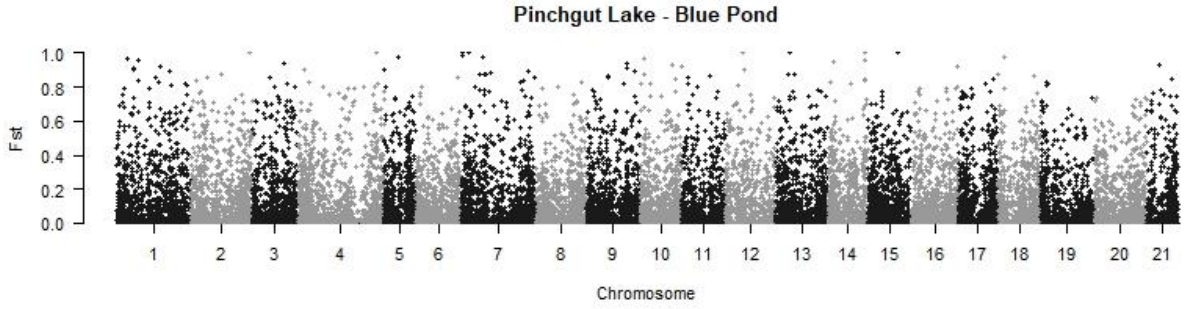

Antigonish Landing - Porper Pond
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

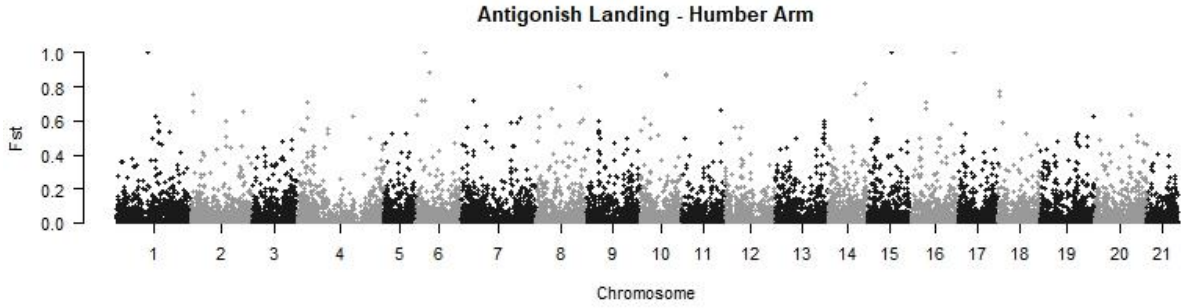

Pomquet Lake - Lake Ainslie
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

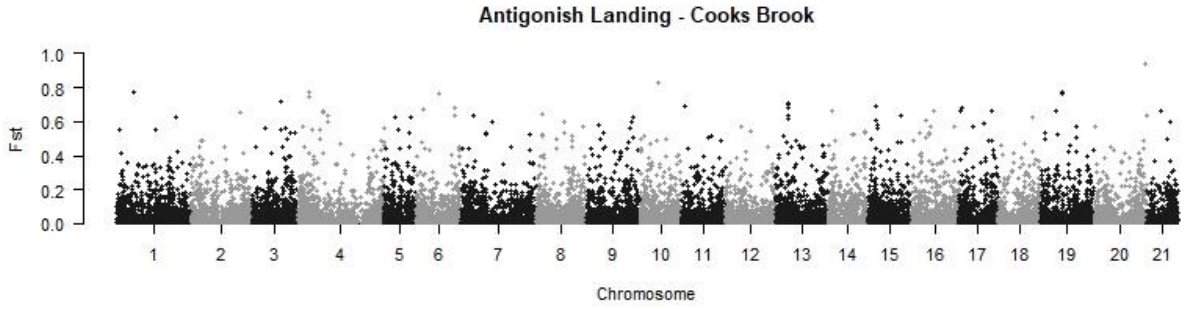

Pomquet Lake - Black River
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

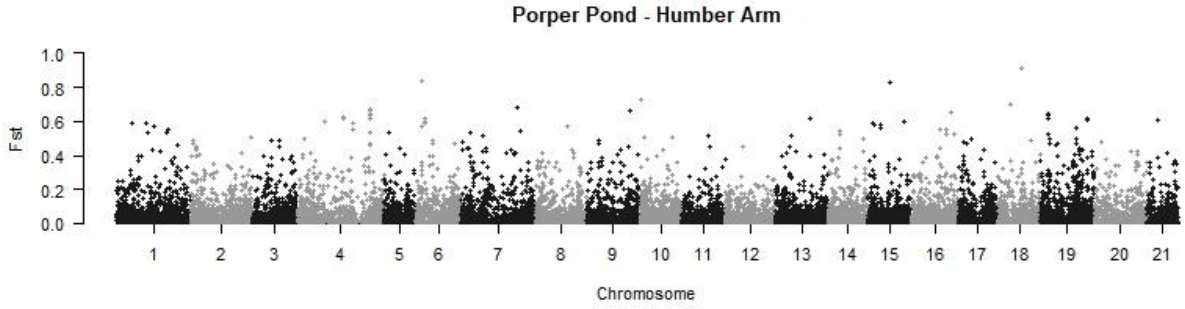

Lake Ainslie - Black River
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

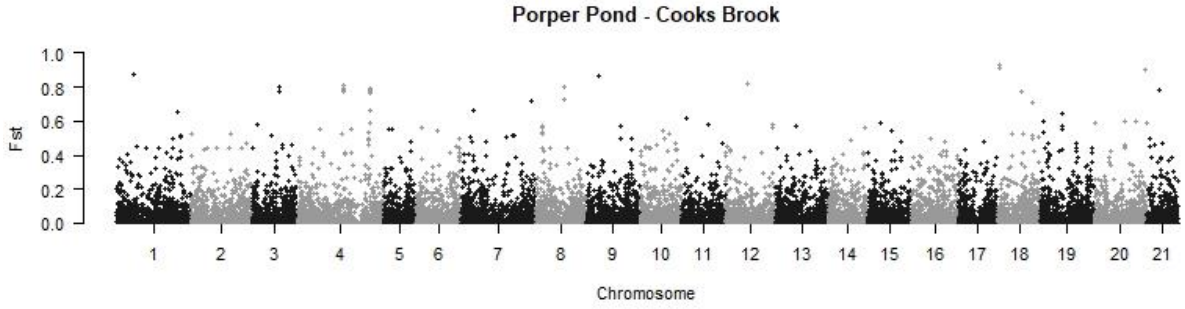

Humber Arm - Cooks Brook
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

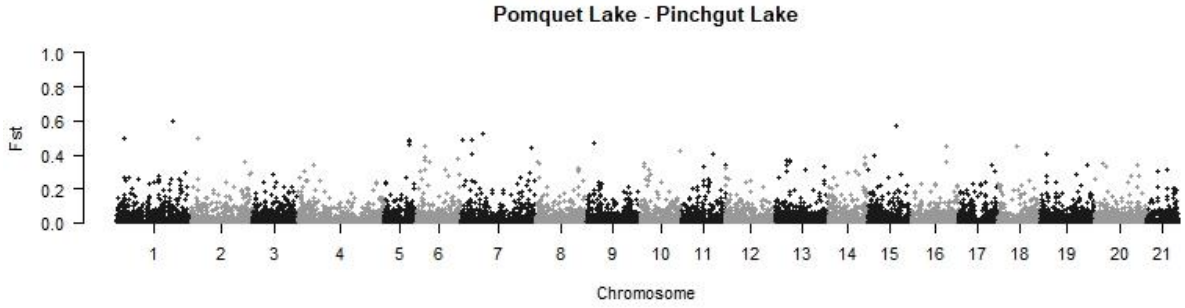

Pinchgut Lake - Blue Pond
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

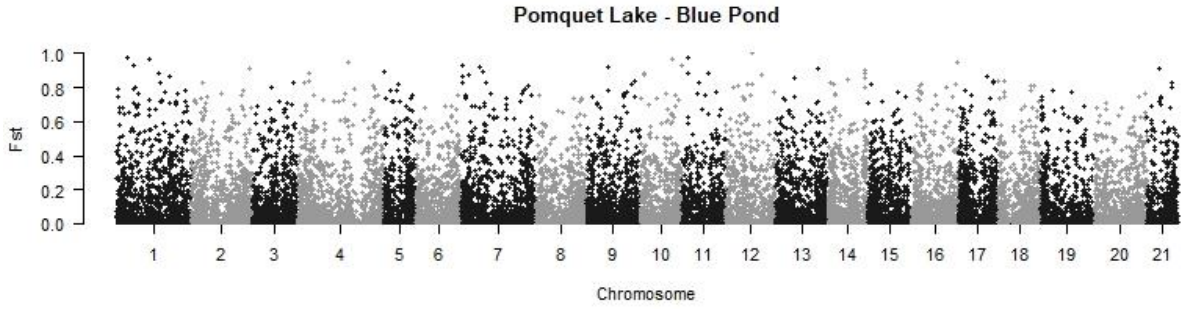

Antigonish Landing - Humber Arm
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome


Antigonish Landing - Cooks Brook
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome


Porper Pond - Humber Arm
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome


Porper Pond - Cooks Brook
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome


Pomquet Lake - Pinchgut Lake
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome


Pomquet Lake - Blue Pond
Fst
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Chromosome

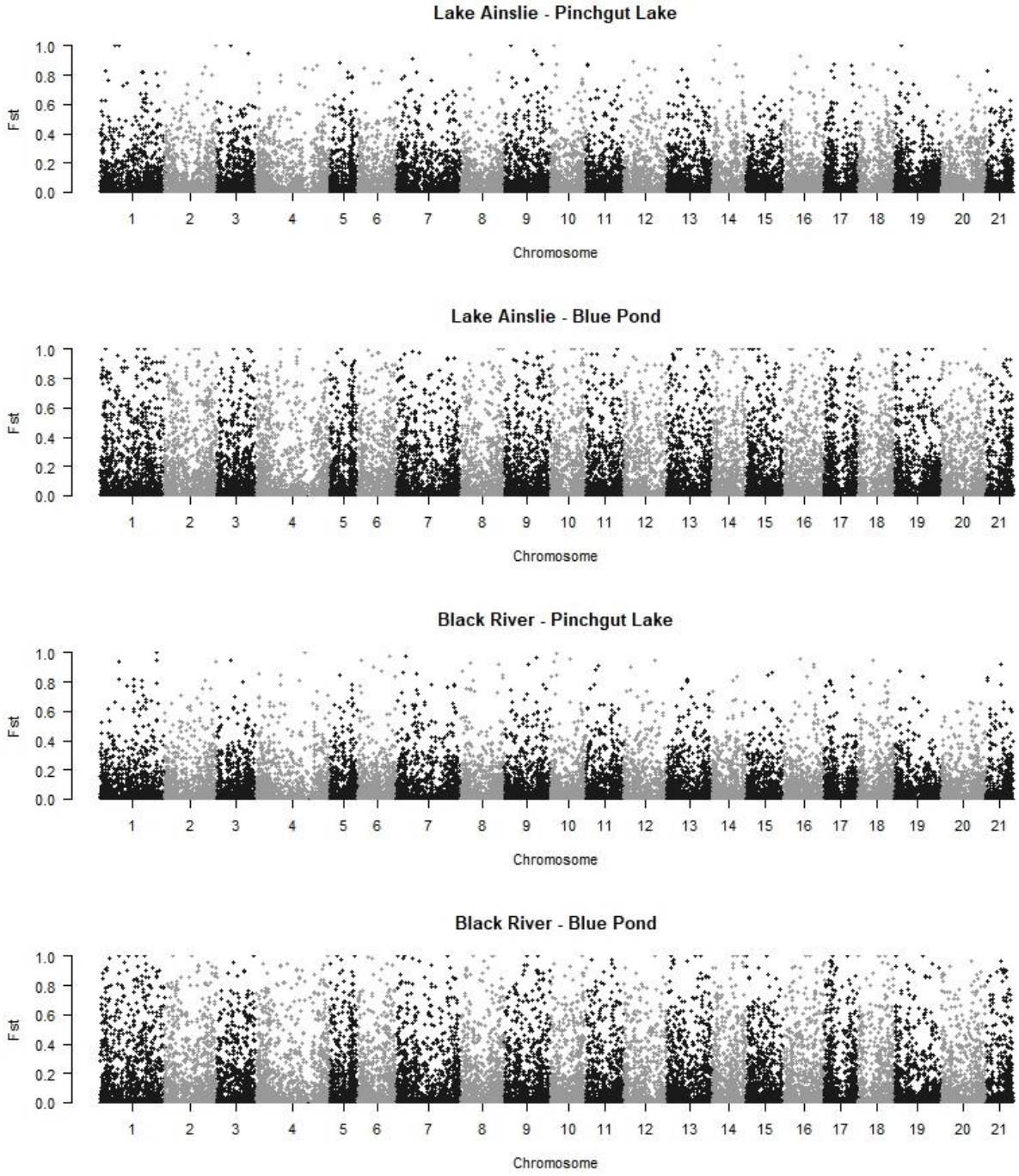


Figure 4. Manhattan plots of F_{ST} populations from similar environments (M-M and FW-FW comparisons).

Gene Ontology

Among 133 genes within 5 kb of an F_{ST} outlier, we found enrichment (FDR = 5%) of biological processes associated with regulation of glutamatergic synaptic transmission (FDR = 0. 0.043, fold-enrichment = 52), catecholamine secretion, including dopamine (FDR = 0.038, fold-enrichment = 22) among our candidate genes under selection (Table C3). Intrinsic component of synaptic membrane was the most enriched GO function related to cellular components (Table C4).

Clustering of marine and freshwater populations

An unrooted tree clarifies the relationships among marine and freshwater populations (Figure 5). Pinchgut and Pomquet Lakes, which are located ~500 km apart, show close relationship, as well as Black River and Lake Ainslie. Blue Pond, which sits on top of its watershed, shows genetic differentiation relative to freshwater and marine sites

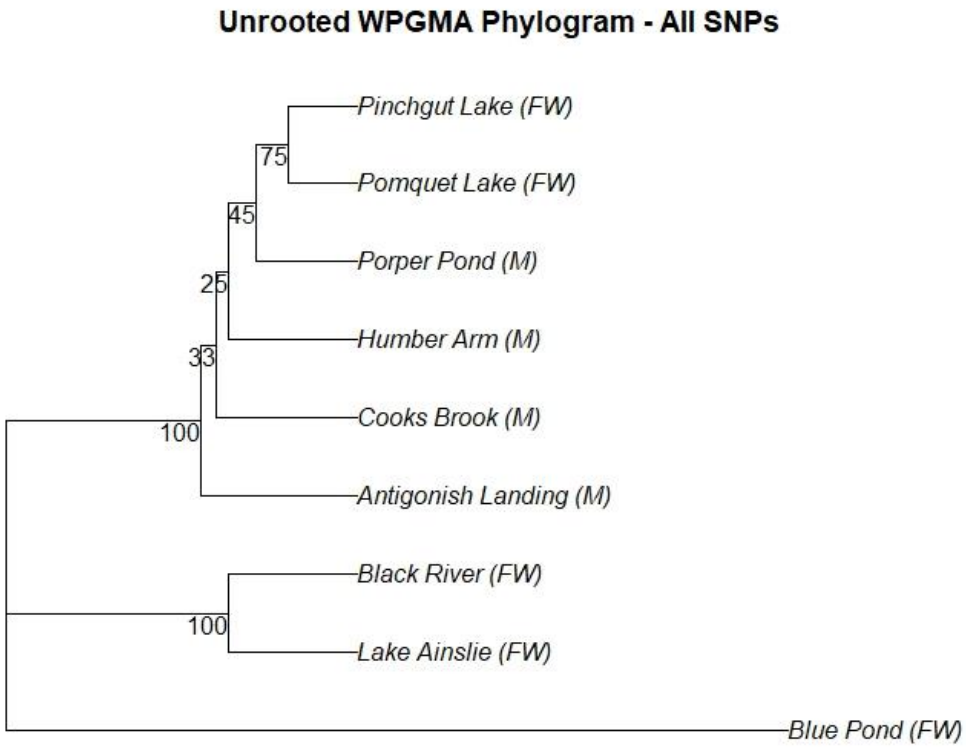


Figure 5. Unrooted phylogenetic tree of marine (M) and freshwater (FW) populations using genome-wide SNPs.

Discussion

In this study, we investigated marine and freshwater populations of stickleback along Canada's east coast to gain insights into their evolutionary history and the potential genetic basis of their adaptation to freshwater. Contrary to the expectation of low genome-wide differentiation between marine and freshwater populations along the Atlantic (Fang et al. 2020a), we found that differentiation among M-FW populations varied widely on the east coast of North America. This suggests a more complex evolutionary history that likely involved multiple colonization events or different refugia,

or variable levels of ongoing gene flow. Evidence of parallel evolution was found as repeated differences in allele frequency at SNP loci across multiple M-FW populations. Notably, we found consistent differentiation of one SNP near a dopamine receptor previously found to be differentially expressed in freshwater (Di Poi et al. 2016). Likely lacking the alleles for reduced armour found in western North America, our results suggest selection on behaviour or endocrine function might have played a more important role in eastern populations persisting in fresh water.

Heterogeneous marine-freshwater differentiation in Eastern Canada

Current evidence on the evolutionary history of threespine stickleback indicates this species colonized the Atlantic during the late Pleistocene, as recently as 36.9 thousand years ago (Fang et al. 2018; Fang et al. 2020). During range expansion into the Atlantic, repeated bottlenecks appear to have led to the stochastic loss of some alleles associated with freshwater adaptation, such as the *Eda* allele for the low-plated morph (Fang et al. 2020). Loci with alleles conferring higher fitness in freshwater are expected to become rapidly differentiated after freshwater colonization (Marques et al. 2018; Kingman et al. 2021). Therefore, the low M-FW differentiation observed in the Atlantic is thought to be caused by a lack of these alleles in the standing variation present in Atlantic populations relative to Pacific populations (Fang et al. 2020).

Here, we show that although genome-wide M-FW differentiation is indeed low in some population comparisons in eastern North America, other comparisons show a level of differentiation more similar to that observed in the Pacific (Figure 3). We found the highest level of differentiation between different freshwater populations. For example, in Newfoundland, stickleback from two lakes isolated from the ocean and only 5 km apart

from each other, Pinchgut Lake and Blue Pond, differ considerably in their degree of differentiation relative to marine populations (Figure 4). Blue Pond is a relatively small lake with no input or output streams and is only fed by rainwater, thus immigration from neighbouring sites is unlikely. Barriers to gene-flow likely played a role in differentiating Blue Pond, as genetic drift drives populations apart in the absence of gene flow. An alternative explanation is that there were multiple waves of colonization in Eastern Canada, as occurred in the Japanese Archipelago (Kakioka et al. 2020). The clustering of freshwater populations is consistent with three colonization events, one by the lineage in Blue Pond, and one by the lineages in Lake Ainslie and Black River, followed by a more recent expansion to Pomquet and Pinchgut lakes. However, our results are also consistent with freshwater populations that evolved in isolation in different glacial refugia during the last ice age (Mäkinen and Merilä 2008).

Evidence of selection on dopamine receptors

Adaptation to similar ecological conditions through re-use of standing genetic variation in a source population is thought to be the dominant mechanism leading to parallel evolution (Schlötterer 2023). The low-plated allele of the *Eda* gene, which is found at low frequencies in marine populations, is known to confer higher fitness in freshwater (Marchinko and Schluter 2007; Barrett et al. 2008; Rouzic et al. 2011). Loss of this allele throughout much of the Atlantic gene pool helps explain why parallel evolution of the low-plated phenotype is not observed in lakes of eastern Canada. However, *Eda* is not the

only gene known to be selected in freshwater, as osmoregulatory genes have also been implicated in freshwater adaptation at a global scale, including Na/K ATPase pump and $Na^{+/}K^{+/}2Cl^{-}$ cotransporter (Deagle et al. 2013; Divino et al. 2016; McCairns and Bernatchez 2010; Garcia-Elfring et al. 2023).

We did not find evidence of repeated differentiation near the *Eda* locus, which is consistent with historical observations of stickleback phenotypes in lakes from this region of North America and stochastic loss of allele during range expansion (Hagen and Moddie 1982; Fang et al. 2020). Nonetheless, the west Atlantic stickleback populations sampled here show evidence of parallel evolution, with 1.1% of SNPs showing some evidence of repeated differentiation in M-FW comparisons in Nova Scotia, Newfoundland, and between the two provinces. Notably, we found repeated M-FW differentiation of SNPs near dopamine receptors, *Drd4a* and *Drd2l*, and enrichment of gene functions related dopamine receptor activity and nervous system development. These results suggest that stickleback might have adapted to freshwater through parallel changes in behaviour, like aggression.

Individual variation in endocrinology tends to be underutilized in evolutionary biology despite its potential to bridge physiology with ecology and genetics (Williams 2008). In zebrafish (*Danio rerio*), the expression profile of dopamine receptors *Drd2c* and *Drd3* have been found to correlate with individual differences in aggression, particularly between dominant and subordinate males (Filby et al. 2010). In stickleback, exposure to a simulated predator lowers the expression of *Drd4a* (Abbey-Lee et al. 2018). At the population level, marine and freshwater populations show extensive differences in the behavioural parameters of sociability, aggressiveness, and levels of

activity (Di-Poi et al. 2014). A study comparing physiological regulatory networks in freshwater and marine populations of stickleback indicates that evolution of receptors, like *Drd2,* and not ligands (e.g., monoamine levels or metabolites) are associated with population divergence.

The dopaminergic system is also known to inhibit prolactin which, in fish, plays an important function in osmoregulation (Liu et al. 2006; Yamamoto and Vernier 2011; Mancera and McCormick 2019). Prolactin increases retention of ions like Na^+ and Cl^- and decreases water uptake, giving it the name the 'freshwater-adapting hormone' (Liu et al. 2006). Prolactin mRNA levels correlate with local ion concentrations in natural populations of the black-chinned tilapia (*Sarotherodon melanotheron*) adapted to different salinities (Tine et al. 2007). In the spotted scat (*Scatophagus argus*), a euryhaline fish, the renal dopaminergic system responds via *Drd1* to hypo-osmotic shock by inhibiting Na^+/K^+-ATPase (Su et al. 2016). In stickleback, prolactin is linked to freshwater migration and tolerance (Lam and Hoar 1967; Ishikawa et al. 2016; Pavlova et al. 2020; Taugbøl et al. 2022). Thus, the parallel differentiation of regions near dopamine receptors in M-FW comparisons may reflect adaptation mediated by phenotypic changes to behaviour, osmoregulation, or both.

Conclusion

Using genome-wide SNPs, our study showed a complex evolutionary history of threespine stickleback in eastern Canada. We found evidence of parallel evolution near dopamine receptors, suggesting that behavioural traits might have been an important target of selection upon colonization of freshwater lakes. Alternatively, individuals with alleles associated with specific behaviour profiles may have been predisposed for

undergoing range expansion into freshwater. Genetic changes near genes for dopamine regulation may also be connected to osmoregulatory adaptation. However, phenotypic data and measures of fitness are quired to establish a causal link between genotype, phenotype, and adaptive evolution.

Bridging text

Chapter 3 presented an investigation of freshwater adaptation and the evolutionary history of post-glacial populations of stickleback from Eastern Canada. In Chapter 4, we focus on the lower limits of contemporary evolution. We use estuary stickleback as a system for glimpsing the first steps of adaptive evolution.

CHAPTER 4

Using seasonal genomic changes to understand historical adaptation to new environments: parallel selection on stickleback in highly variable estuaries

*Alan Garcia-Elfring, *¶Antoine Paccard, *ŦTimothy J. Thurman, †Ben A. Wasserman, †Eric P. Palkovacs, *Andrew P. Hendry, and *Rowan D. H. Barrett

*Redpath Museum and Department of Biology, McGill University, Montreal, QC, Canada

¶McGill University Genome Center, McGill University, Montreal, QC, Canada

†Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA, USA

ŦCurrent address: Division of Biological Sciences, University of Montana, Missoula, MT, USA.

Abstract

Parallel evolution is considered strong evidence for natural selection. However, few studies have investigated the process of parallel selection as it plays out in real time. The common approach is to study historical signatures of selection in populations already well adapted to different environments. Here, to document selection under natural conditions, we study six populations of threespine stickleback (*Gasterosteus aculeatus*) inhabiting bar-built estuaries that undergo seasonal cycles of environmental changes. Estuaries are periodically isolated from the ocean due to sandbar formation during dry summer months, with concurrent environmental shifts that resemble the long-term changes associated with postglacial colonization of freshwater habitats by marine populations. We used pooled whole-genome sequencing (Pool-WGS) to track seasonal allele frequency changes in six of these populations and search for signatures of natural selection. We found consistent changes in allele frequency across estuaries, suggesting a potential role for parallel selection. Functional enrichment among candidate genes included transmembrane ion transport and calcium binding, which are important for osmoregulation and ion balance. The genomic changes that occur in threespine stickleback from bar-built estuaries could provide a glimpse into the early stages of adaptation that have occurred in many historical marine to freshwater transitions.

Introduction

Knowledge of the genomic targets of natural selection is central to understanding the mechanisms responsible for adaptive evolution at the molecular level. Generating this knowledge often involves comparing patterns of genomic differentiation (e.g., F_{ST}) between populations adapted to distinct ecological conditions (e.g. Hoekstra et al. 2006; llardo et al. 2018). In such studies, natural selection is considered a strong candidate for the mechanism driving phenotypic diversification when multiple closely related but independently evolved populations use the same genetic pathways to reach a shared adaptive solution to an environmental challenge. Here, we refer to this phenomenon as parallel evolution (Elmer and Meyer 2011). In vertebrates, studies of parallel evolution have provided insights into the genetic mechanisms underlying adaptation to freshwater (Colosimo et al. 2005; Barrett et al. 2008; Schluter et al. 2010; Hohenlohe et al. 2010; Kitano et al. 2010; Lescak et al. 2015; Rudman et al. 2019; Fang et al. 2020), novel pathogens (Alves et al. 2019), low oxygen availability (McCracken et al. 2009; Foll et al. 2014; Wang et al. 2014; Graham and McCracken 2019; Lim et al. 2019), crypsis (Comeault et al. 2016; Jones et al. 2018; Barrett et al. 2019), nutrient-limited environments (Riddle et al. 2018), and dissolved ion (H^+) profiles (Haenel et al. 2019). However, nearly all studies of parallel evolution are retrospective in the sense that they investigate reasonably well adapted populations long after selection for successful habitat transition occurred.

Retrospective approaches thus have difficulty detecting the specific genetic changes that were under natural selection during the initial habitat shift amongst the noise from local effects and stochastic processes that accumulates afterward (Elmer and Meyer

2011). A valuable addition to the inferential toolbox, then, is to study natural selection that takes place *during* parallel habitat shifts. Most studies adopting this selection-based approach have used artificial perturbations of genotypes or environments, which have revealed genomic targets of strong selection (e.g. Soria-Carrasco et al. 2014; Nosil et al. 2018; Barrett et al. 2019). However, the manipulations involved in such tests leaves uncertain the extent to which similar selection would attend natural habitat shifts. One solution is to take advantage of serendipitous events, like studying populations before and after extreme weather events (Grant et al. 2017), like hurricanes (Donihue et al. 2018), heat waves (Coleman et al. 2020; Gurgel et al. 2020) and cold snaps (Campbell-Staton et al. 2017; Card et al. 2018). However, the location and timing of extreme weather events are unpredictable by nature, limiting the study of how natural populations respond to such events. The optimal situation, then, might be to study selection as it occurs in natural populations experiencing large, yet predictable, shifts between alternative environments, like those that occur during seasonal changes (e.g. Behrman et al. 2018; Tourneur et al. 2020). Here, we search for the signature of natural selection in a set of populations that experience parallel seasonal changes in local conditions that likely resemble the early phase of a classic habitat transition. Specifically, we study allele frequency changes in stickleback populations in environments that alternate between marine-like (brackish) and freshwater conditions.

Study system

The threespine stickleback (*Gasterosteus aculeatus*, 'stickleback' hereafter) is a classic model system for studying parallel evolution (Boughman et al. 2005; Colosimo et al.

2005; Jones et al. 2012a, 2012b; Deagle et al. 2013; Hendry et al. 2009, 2013; Lescak et al. 2015; Paccard et al. 2018; Haenel et al. 2019; Smith et al. 2020). Over the past approximately 12,000 years, marine stickleback have repeatedly colonized and become adapted to freshwater environments, often through parallel phenotypic changes (Reimchen 1983; Colosimo et al. 2005) linked to predator defence (Reimchen 2000; Marchinko 2009; Miller et al. 2019; Wasserman et al. 2020) and ion regulation (Gibbons et al. 2016, 2017; Hasan et al. 2017). The genomic basis of such adaptation is partly known. For instance, researchers have identified genes of large effect underlying differences in the number of bony armour plates (e.g. *Eda* gene, e.g. Colosimo et al. 2005), pelvic spines (e.g. *Pitx1* gene, Chan et al. 2010), and the ability to osmoregulate (e.g. *Kcnh4* gene, Jones et al. 2012a; *Atp1a1a* gene, McCairns and Bernatchez 2010; Jones et al. 2012b; Hasan et al. 2017). These genes have been found primarily through a retrospective approach of studying signatures of selection millennia after the initial colonization.

Stickleback populations in bar-built estuaries along the coast of California represent a natural system for studying parallel selection over seasonal timescales. These populations experience repeated bouts of strong and abrupt temporal changes driven by wet winters and dry summers. With heavy winter rains, increased water flow breaches the wall of sediment (i.e., ‘sandbar’) that, during the summer, typically isolates estuaries from the ocean. When rains subside, sandbars are re-built by wave action and sand deposition in the spring or summer, isolating estuaries from the ocean and creating coastal lagoons (Behrens et al. 2009; Behrens et al. 2013; Rich and Keller 2013). The changes in precipitation that lead to the build-up of sandbars and the subsequent

breaching can result in drastic environmental shifts in, for example, predator abundance (Becker and Reining 2008), salinity (Williams 2014), and habitat structure (Heady et al. 2015). For example, in between breaching events, a shift takes place from lotic (i.e. moving) brackish water to lentic (i.e. pond-like) freshwater (Heady et al. 2015; Des Roches et al. 2020). Salinity also becomes stratified on the water column (Williams 2014), with freshwater forming the top layer. During the time that the estuary is closed, the surface freshwater layer progressively increases in thickness (see Figure 2.17 of Williams 2014). Following a breaching event there is a mixing of freshwater and saltwater, resulting in drastic increases and decreases in salinity in the top and bottom of the water column, respectively (see Figures 2.19, 2.24 - 2.27 of Williams 2014). These seasonal habitat shifts may be analogous to the environmental changes experienced by stickleback populations during postglacial marine-to-freshwater colonization events and are replicated both spatially (in different estuaries) and temporally (with seasonal changes in precipitation).

To study natural selection in action, we sampled stickleback from six bar-built estuaries at two time points between breaching events, when the estuaries were isolated from the ocean. Using a whole-genome SNP dataset, we characterized the extent of allele frequency change between the sampling times, which should reflect, at least in part, natural selection taking place during that interval. Then, to evaluate evidence of parallel natural selection, we ask several questions. (1) What genomic regions show relatively large changes in allele frequency across time in multiple estuaries? (2) For these regions, do the changes in allele frequency occur in parallel across estuaries? (3) Do genes putatively under parallel selection show enrichment of genetic functions consistent with

the changes in the environment? By obtaining this information over a seasonal timescale in multiple estuaries, we hope to gain insight into the genetic changes driven by selection that may have occurred when postglacial stickleback populations first colonized freshwater environments from the ocean.

Methods and Materials

Field sampling and DNA extraction

In 2016, we sampled stickleback from six bar-built estuaries (Figure 1), three from small coastal watersheds (< 7 km^2; Old Dairy, Lombardi, and Younger) and three that are relatively large watersheds (> 22 km^2; Laguna, Scott, and Waddell; see Table S1 for full estuary names and size metrics). We sampled in the spring at the end of the breaching season upon completion of the sandbar (after winter rain), and again in the fall before the breaching season (before the winter rain, Figure D1). Thus, we are testing for selection during the part of the year that estuaries are isolated from the ocean, which provides a single-season analogue of the marine to freshwater transition that is a classic theme in stickleback research (e.g. Colosimo et al. 2005; Bassham et al. 2018; reviewed by Hohenlohe and Magalhaes 2019). At each time point in each estuary, we collected 40 adult stickleback (> 30 mm in length) by means of minnow traps and beach seines. Although sampling time is less than one generation (~6 months), some stickleback may have given birth shortly after our first sample, with progeny growing large enough to be sampled as adults in our second sample. Selection during our sampling period therefore reflects both differential mortality and reproduction. Fish were euthanized with tricaine

methanosulphonate (MS-222) and tissue samples (pectoral fin) were stored in 95% ethanol prior to DNA extraction. Collections were made in accordance with California Scientific Collector's Permit SC-12752. Animal handling protocols were approved by the University of California, Santa Cruz IACUC under protocols Palke-1306 and Palke-1310. We extracted DNA following a standard phenol-chloroform procedure. Briefly, tissue samples were placed in digestion buffer containing proteinase K and incubated at 55 °C. We then isolated DNA using an isoamyl-phenol-chloroform solution, followed with ethanol precipitation.

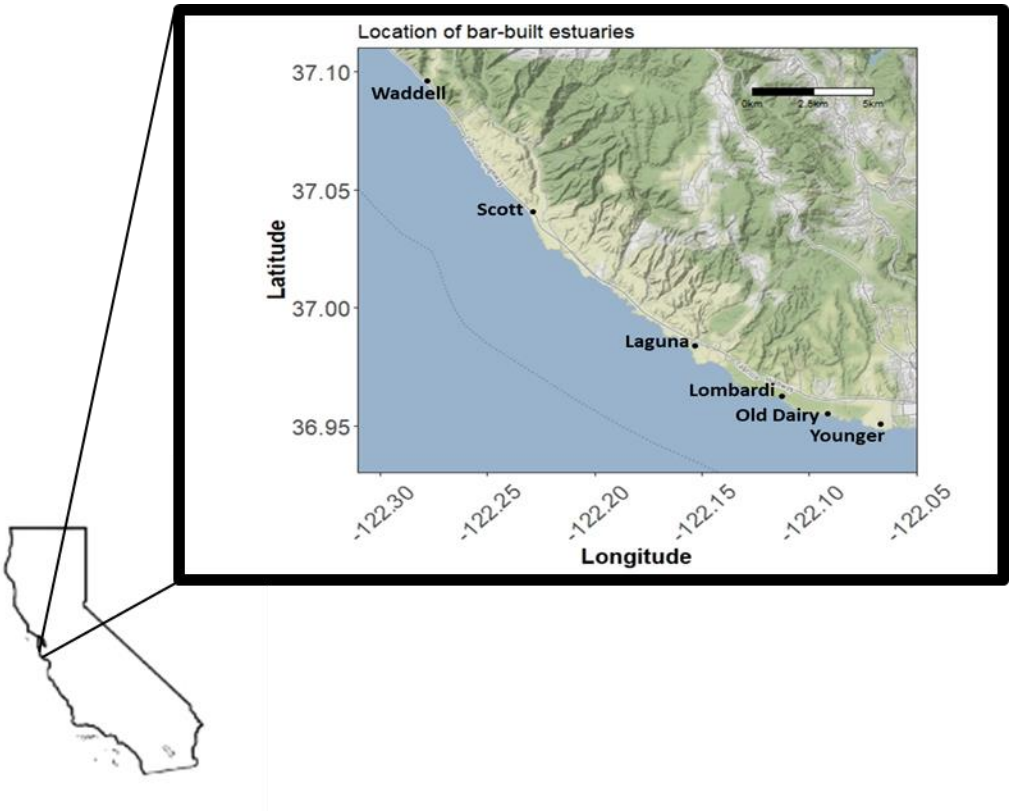


Figure 1. Locations of six bar-built estuaries sampled along the coast of California USA.

Sequencing

We quantified all samples using a Picogreen® ds DNA assay (Thermo Fisher Scientific, Waltham, USA) on an Infinite® 200 Nanoquant (Tecan Group Ltd. Männedorf, Switzerland). Samples were normalized to a dsDNA concentration of 15ng/µl, re-quantified, and pooled according to sampling location and time of sampling. Thus, we created 12 pools of 40 individuals each (i.e. six estuaries sampled at two time points). Whole-genome libraries of each pool were prepared at the McGill University and Genome Quebec Innovation Center, Montreal, Canada, and sequenced across five lanes of Illumina HiSeq2500 with paired-end, 125bp reads.

Bioinformatics

We filtered raw reads based on quality (--quality-threshold 20) and length (--min-length 50) with the *trim-fastq.pl* script of *Popoolation* (Kofler et al. 2011a). The resulting reads were mapped to the stickleback reference genome (BROADS S1) using *BWA mem* v. 0.7.13 (Li and Durbin 2009). We then used *SAMtools* (Li et al. 2009) to convert SAM files to BAM format and remove reads with mapping quality below 20 (samtools view -q 20). We then generated a mpileup file (samtools mpilep -B) and filtered for a minimum depth of coverage of 5X. We converted the mpileup file to the synchronized (sync) format using *Popoolation2* (Kofler et al. 2011b) for downstream analysis.

Analysis of repeated genomic differentiation

In this study we use F_{ST} to measure changes in allele frequency within a lineage across time points (rather than differentiation between two lineages, Burri 2017). First, within each estuary, we calculated F_{ST} (Hartl and Clark 1997) at the SNP level to identify variants showing relatively large changes in allele frequency (i.e. outliers) between the

two time points. We then quantified the extent of overlap in these outliers among estuaries. To obtain genome-wide F_{ST} estimates, we applied the *fst-sliding.pl* script of *Popoolation2* for each of the six estuaries post-breaching (i.e. brackish conditions) versus pre-breaching (freshwater) (--min-count 2, --min-coverage 5, --max-coverage 100, --min-covered-fraction 0, --window-size 1, --step-size 1, --pool-size 40:40:40:40:40:40:40:40:40:40:40:40, --suppress-noninformative). We only analyzed genomic regions assembled at the chromosome level (i.e. scaffolds excluded). We included data from chromosome 19 (allosome) as we did not find evidence of any artefact on this chromosome or large differences in the coverage (mean = 23.92, SD: 6.48, range 5-74) relative to the genome-wide average (see Results). We identified F_{ST} outliers as SNPS that fell in the top 5% of the F_{ST} distribution. These loci were excluded from calculations of genome-wide F_{ST} and allele frequency change distributions to obtain estimates for putatively neutral SNPs (e.g. Batista et al. 2016). To discover candidates potentially under selection, we focused on SNPs that showed large allele frequency changes in multiple estuaries. Because drift and sampling variance will affect loci at random across the genome within any particular estuary, it is unlikely that consistent genetic changes across three or more different estuaries will be due to stochastic processes. We used a custom bash script to quantify F_{ST} outlier overlap across estuaries and identify SNPs that qualify as outliers in at least three out of the six estuaries. However, because evidence of repeated changes in allele frequency in the same SNP (as shown by F_{ST}) does not necessarily mean that these changes were parallel (i.e. in estuary *X* an allele shows a large *increase* in frequency, while in estuary *Y* the same allele

experiences a large *decrease* in frequency), we also tested for parallelism in allele frequency change.

Parallel changes in allele frequencies in response to seasonality

We identified SNPs showing consistent directional changes in allele frequency across our estuaries using the program *PoolFreqDiff* (Wiberg et al. 2017). *PoolFreqDiff* uses a generalized linear model with a quasibinomial error distribution (qGLM). Wiberg et al. (2017) showed that the qGLM has a substantially lower false positive rate than the Cochran-Mantel-Haenszel test, a method commonly used in pool-seq studies to identify consistent changes in allele frequency across replicates. We used the same flags (e.g. minimum read count and coverage settings) in the *PoolFreqDiff* program as in our F_{ST} analysis. The qGLM test implemented in *PoolFreqDiff* has the potential to identify parallel selection reflected as subtle, but consistent, changes in allele frequency in multiple estuaries. Such small changes in allele frequency are unlikely to be identified as F_{ST} outliers in individual estuaries. We used the 'no rescaling' option of *PoolFreqDiff* (re-scaling allele counts relative to the effective sample size gave similar results). We corrected for population structure using the empirical null-hypothesis approach (Caye et al. 2016; François et al. 2016). Visual inspection of the histogram of corrected P-values confirmed a uniform distribution, indicating that confounders were controlled (Figure D2). Next, we corrected for multiple hypothesis testing using the false discovery rate (FDR) procedure implemented in the R package *qvalue* V2.14 (Storey et al. 2018). We analyzed three sets of outliers to study potential targets of selection. First, to look for strong and parallel changes in allele frequency, we categorized 'F_{ST}-qGLM outliers' as SNPs that are an F_{ST} outlier in at least three estuaries and also significant (FDR = 1%)

under the qGLM model. Second, we identified outliers from each of the two distinct approaches (F_{ST} and qGLM) but with more conservative thresholds than those used in the overlapping F_{ST}-qGLM outlier set. For 'F_{ST} candidates', we identified F_{ST} outliers (SNPs in the top 5% of the F_{ST} distribution in a single estuary) that were shared across at least four of six estuaries (as opposed to the three estuaries minimum requirement in the F_{ST}-qGLM outlier set). Note that the frequency changes across these estuaries may not be parallel and thus this outlier set accounts for potential causes of selection that may differ in direction among estuaries. We also tested whether estuary size may influence the likelihood of shared targets of selection (F_{ST} outliers; see supplemental information). Finally, 'qGLM outliers' are SNPs identified as highly significant using the quasibinomial GLM test for parallel changes in allele frequency, here using an FDR = 0.01% as opposed to the less conservative FDR = 1% used for the F_{ST}-qGLM outlier set. We obtained estimates of allele frequency change across time points for F_{ST}-qGLM outliers and putatively neutral loci with respect to F_{ST} (not in top 5% F_{ST} distribution) using the *snp-frequency-diff.pl* script of *Popoolation2*.

Identification of candidate genes and analysis of molecular function

To identify genes putatively under parallel selection, we used a custom bash script to map outliers (i.e. F_{ST}-qGLM outliers , F_{ST} candidates, and qGLM outliers) to protein-coding genes in the reference genome. We limited our search to a set of 14,252 protein-coding gene annotations with attributes "ID=gene" and "biotype=protein_coding". To gain insights into the traits under selection, we analyzed candidate genes for enrichment of molecular functions. To do this, we obtained gene names and gene ontology (GO) information from the stickleback reference genome on *Ensembl* using the R package

biomaRt (Smedley et al. 2009). We compared the three lists of candidate genes with the reference set of 14,252 genes ('gene universe') and tested for functional enrichment using the package *TopGO* 2.34.0 (Rahnenfuhrer 2018) and the Fisher's exact test. To reduce false positives, we pruned the GO hierarchy by requiring that each GO term have at least 10 annotated genes in our reference list ("nodeSize = 10").

Results

Data processing and F_{ST} estimates in response to seasonal sandbar formation

Our sequencing efforts led to an average of 23,914,973,875 bases sequenced per pool (SD: 1,760,685,042). After filtering data, we obtained 101,911,501 bases for variable sites, providing F_{ST} estimates for 4,024,542 SNPs distributed across 21 stickleback chromosomes. The average minimum coverage per SNP was 25.32 (SD = 6.96, range: 5 – 84, Figure D3) among pools. Overall, allele frequency changed relatively little within estuaries, showing a mean 'neutral' F_{ST} of 0.0253 across time points (Waddell = 0.0224; Lombardi = 0.0230; Old Dairy = 0.0216; Younger = 0.0243; Scott = 0.0236; Laguna = 0.0369).

Consistent changes in allele frequency and the signature of parallel selection

To identify candidates under temporally varying parallel selection, we looked for an overlap among estuaries of SNPs that fall in the top 5% of the F_{ST} distribution. As expected, most SNPs found in the top 5% of the distribution in an estuary only reach this threshold in a single estuary (Figure 2). Yet, we find 22,111 SNPs in the top 5% in three or more estuaries. The majority of these SNPs, 19,390 SNPs (87.7%), are confined to exactly three estuaries, with 2,721 'F_{ST} candidates' found in four or more estuaries. At a

FDR of 1%, we identified 37,687 SNPs using *PoolFreqDiff*, 705 of which overlapped with the SNPs that were F_{ST} outliers in at least three estuaries ('F_{ST}-qGLM outliers'; Figure 3 and Figure 4). We also identified 2,411 SNPs with a FDR of 0.01% ('qGLM outliers').

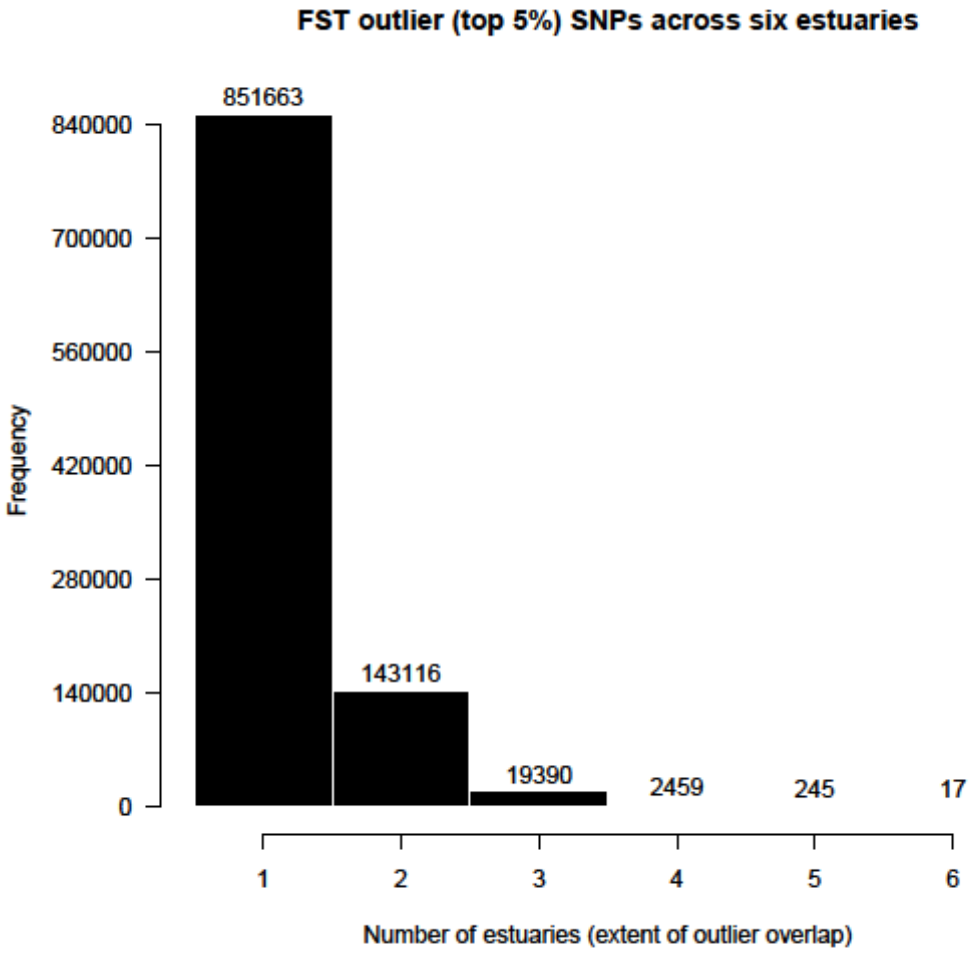


Figure 2. Extent of overlap of F_{ST} outlier loci across six estuaries.

Candidate genes and analysis of molecular function

We mapped outlier SNPs to genes and found 710 genic F_{ST} candidates in 579 different genes (Table S2), 704 qGLM outliers in 569 genes (Table S3), and 190 F_{ST}-qGLM outliers in 169 different genes (Table S4). All three sets of outliers have candidate genes associated with ion transfer, including *Wnk4* (Table S2 and Table S3) and *Nalcn* (Table S3 and Table S4). We find consistent changes in allele frequency in ATPase genes that code for proteins that transport, for example, sodium and potassium (e.g., *Atp1b1a*, Table S2) and phospholipids (*Atp8b5a*, Table S2 and Table S4). Potassium transport channels are also found among our candidate genes (e.g., *Kcnma1a*, Table S3 and Table S4; *Kcnn1a*, Table S2) as well as genes involved in calcium binding or transport, like the calcitonin receptor (*Calcr*, Table S2 and Table S3), calmodulin (*Calm1b*, Table S2) and the calcium channel *Cacna1d* (Table S4). Yet others code for various mitogen-activated protein kinases (e.g., *Map3k12*, Table S3; *Mapkbp1*, Table S4). We also found an F_{ST} outlier mapping to *Ccny* (Table S2).

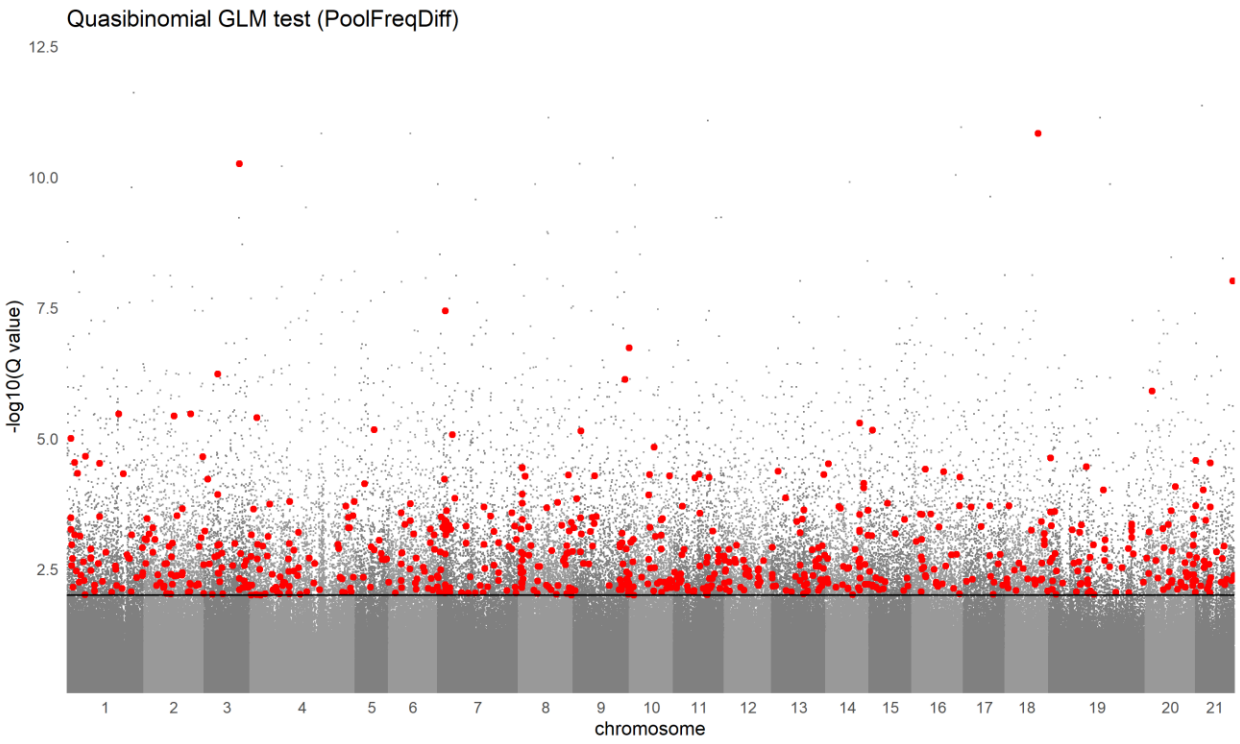


Figure 3. Result of qGLM test (*PoolFreqDiff*) for parallel changes in allele frequency. Each dot represents a single SNP (MT chromosome and unplaced scaffolds excluded). The 705 loci identified by F_{ST} analysis as candidates in at least three estuaries and by the qGLM test as significant (FDR = 1%) are shown in red ('F_{ST}-qGLM outliers'). Black line demarks the 1% false discovery rate.

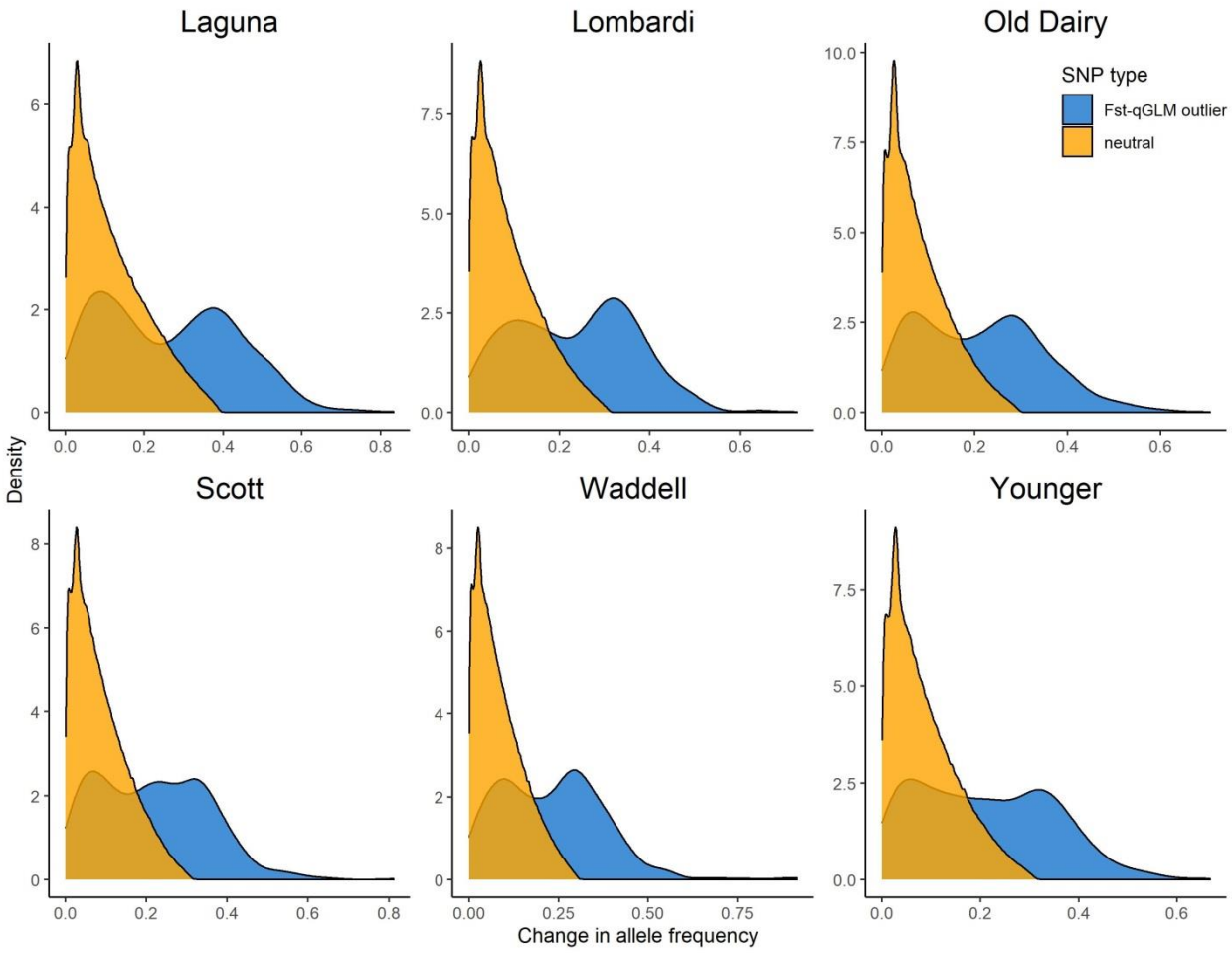


Figure 4. Distribution of allele frequency change in genome-wide SNPs and F_{ST}-qGLM outliers.

GO analysis

For an overall assessment of the gene functions that are most represented among our three sets of candidate genes, we tested for enrichment of molecular function. We find that the candidate genes from all three outlier sets have in their top ten most significant GO terms molecular functions related to ion channel activity. For example, F_{ST} candidates show enriched GO functions related to calcium ion binding (Table D5).

Among the most enriched molecular functions among qGLM candidate genes are metal ion transmembrane transporter activity and calcium ion transmembrane transporter activity (Table D6). Similarly, F_{ST}-qGLM outlier genes are enriched for ion transmembrane transporter activity and ligand-gated ion channel activity (Table S7). We also found extracellular matrix structural constituent (Table S5) among the significant molecular functions (see Discussion).

Discussion

To document natural selection, we studied stickleback populations from estuaries experiencing seasonal fluctuations in environmental conditions. We found evidence of parallel selection on genes linked to ion transport and salinity adaptation. Consistent with a change in the ionic environment, we found that the most statistically significant functions were related to ionic homeostasis. Our findings suggest that intermittent connectivity with the ocean results in episodic shifts in selection regime, a change that may resemble the initial phase of freshwater colonization that occurred during the marine-freshwater transitions of postglacial stickleback populations. More generally, our study adds to a growing literature that collectively shows that natural selection can drive genetic change over very short timespans.

Parallel selection over a seasonal timescale

We found evidence of natural selection in the form of consistent changes in allele frequency in stickleback populations from different bar-built estuaries. Our results

showed changes in allele frequency on genes with functions related to ion balance. This suggests that seasonal fluctuations of environmental conditions shift the selection regime within bar-built estuaries. The time scale involved in the overall subtle but consistent changes in allele frequency detected here conforms with studies that show selection for freshwater adaptation can be detected within a single year rather to decades (Bell 2001; Kristjánsson et al. 2002; Bell et al. 2004; Lescak et al. 2015; Marques et al 2018). These results are also consistent with the idea that adaptation to a particular environment, like freshwater, likely happens in the first few generations after colonization (e.g. reviewed by Reimchen et al. 2013). Importantly, our results set a new bar for how quickly selection can result in genetic adaptation during freshwater colonization by stickleback.

Seasonal allele frequency changes in genes involved in ion balance

Over the last 10,000 years, post-glacial stickleback populations have adapted to the different osmotic conditions found in freshwater (Spence et al. 2012). In this study, we found that seasonal isolation from the ocean led to repeated changes in allele frequency in many genes linked to ion balance. For example, genes such as *Nalcl* and *Wnk4* show signs of potential selection as shared outliers across multiple estuaries. *Nalcn* is a salt-sensing gene that was recently found to be rapidly evolving in saline-alkaline lake-dwelling fish (Tong and Li 2020). Similarly, *Wnk4* codes for an intracellular chloride sensor (Chen et al. 2019) implicated in salinity-tolerance in stickleback (Wang et al. 2014). Additionally, we found evidence of parallel selection on the gene *Ccny*. A recent epigenomic study linked *Ccny* to salinity adaptation, showing that *Ccny* is differentially methylated in stickleback populations along a gradient of decreasing salinity (Heckwolf et al. 2020). We also found evidence of selection on genes from the *Mapk* family. These

genes are differentially expressed in fish in response to many environmental stressors (Mateus et al. 2017), including osmotic stress (Tse et al. 2011; Tian et al. 2019; reviewed by Kültz 2012).

We also found evidence of selection on genes for calcium balance, for example, the calcium sensing gene calmodulin (*calm1b*, Chin and Means 2000). Another gene we found to be putatively under seasonal parallel selection is the calcitonin receptor (*Calcr*), part of a family of genes known to regulate calcium homeostasis (Naot et al. 2019). Indeed, a gene from the same family, the calcitonin gene-related peptide type 1 receptor (*Calcrl*), has been implicated in salinity tolerance or osmoregulatory adaptation in postglacial stickleback populations (Kusakabe et al. 2017). These findings suggest a potential role for selection on these loci. However, to avoid false conclusions based on selective assessment of particular genes we analyzed the overall genetic functions of candidate genes.

Temporal changes in bar-built estuaries likely select for ionic homeostasis

We focused on the top ten most significant GO terms in each outlier set and found functions related to ion transmembrane transport among candidate genes. Our findings are consistent with previous studies on freshwater adaptation that have documented divergence in genes linked to ion channels (e.g. sodium/potassium channels) during parallel adaptation to freshwater (DeFaveri et al. 2011; Jones et al. 2012a; Jones et al. 2012b; Gibbons et al. 2016; Gibbons et al. 2017; Hasan et al. 2017; Rudman et al. 2019; Heckwolf et al. 2020). Freshwater adaptation has occurred independently in a wide range of taxa through selection on genes involved in osmoregulation. For instance, annelids (Horn et al. 2019), arthropods (Lee et al. 2011), and fish (Velotta et al. 2016), including

sculpin (Dennenmoser et al. 2016) and stickleback (Jones et al. 2012b; Kusakabe et al. 2016; Hasan et al. 2017), have all shown genetic changes in ion channel genes following colonization of freshwater habitats. Our findings are also in line with a recent study (Tong and Li 2020) on adaptation of fish to a saline-alkaline lake, which showed that rapidly-evolving genes, those with an elevated rate of non-synonymous substitutions, are overwhelmingly involved in ion transport. Our findings of gene functions related to transmembrane ion transport could also be in part driven by changes in temperature over the time period sampled (during the summer months). Increases in temperature may disrupt osmoregulation. A study on estuary fish from California showed that experimentally exposing fish to higher temperature results in differential expression of some of the same genetic functions we found overrepresented among our candidate genes, including ion channel activity and extracellular matrix structural constituent (Jeffries et al. 2016). Thus, changes in temperature could amplify the osmoregulatory challenges experienced by populations responding to changes in salinity.

Freshwater adaptation in stickleback also could involve genetic changes to maintain calcium homeostasis (Gibbons et al. 2016). Calcium binding proteins play an important physiological role in maintaining calcium balance in fish (Evans et al. 2005). Calcium must be continuously absorbed from freshwater, which is hypoosmotic relative to fish plasma (Liem et al. 2001; reviewed by Evans et al. 2005). Fittingly, not only did we find consistent changes in allele frequency in relevant genes, but we also found an overall enrichment of gene functions related to calcium ion binding. This provides additional evidence that genes for ion regulation are targets of selection during freshwater transitions that last from months to millennia.

Conclusion

We found evidence of natural selection for osmoregulatory adaptation, likely brought into operation by seasonal changes in the ionic environment within estuaries. Repeated changes in allele frequency across estuaries suggests parallel selection is occurring, highlighting the power of this system for studying adaptive evolution over very short timescales. Our results are consistent with the idea that cyclical isolation and exposure to the ocean results in seasonally oscillating selection, although time-series data over multiple instances of sandbar formation is needed for confirmation. The threespine stickleback found in bar-built estuaries along coastal California thus provide the rare opportunity to study parallel selection in real-time, *in natura*.

THESIS DISCUSSION

The research presented in this thesis sought to add to our knowledge of the genetics of phenotypic differentiation using two complimentary model systems under artificial and natural selection. Chapters 1 and 2 used the history of artificial selection on ball pythons to investigate the genetic basis of reptile colouration. In Chapter 1, we established that a nonsense mutation in *tfec* is associated with the piebald phenotype. Functional validation on a reptile model showed that *tfec* function is required for structural colouration. We did not find evidence of protein-coding variants other than the nonsense mutation in *tfec* associated with the piebald pattern. It is likely that all piebald ball pythons carry this mutation. Piebald ball pythons in Africa are rare but not unheard of. There are reports of observations of piebald ball pythons in Ghana, one dating back to the 1960s (www.reptileadvisor.com/piebald-ball-python/). It is likely the variant discovered in captivity is segregating in nature where it likely experiences negative natural selection given the low frequency of the recessive phenotype and the hinderance to camouflage caused by white spotting. This contrasts sharply with the positive artificial selection imposed in captivity. It remains an open question whether other variants in *tfec* associated with piebaldism are segregating in the native distribution of ball pythons. The primary limitation of Chapter 1 is doing functional validation in a lizard rather than a snake. That is, not having *tfec* mutant ball pythons or TEM images of scales, either from piebald or wild-type individuals, leaves open questions about the effects of *tfec* variation in ball pythons and snakes in general. Developing a CRSIPR/Cas9 protocol for use in ball pythons will be an important step in establishing ball pythons as a model organism in

pigmentation research. From a commercial perspective, such gene editing technology in ball pythons is expected to have a disruptive effect on the reptile industry.

Chapter 2 focused on the pigment-producing melanophores and xanthophores by studying the clown and VPI axanthic phenotype, respectively. We found that a missense mutation in *mc1r* is linked to the clown morph. This finding showed that *mc1r* not only regulates the ratio of eumelanin to pheomelanin, but also plays a role in pattern formation in snakes, perhaps by controlling the density or distribution of melanophores. A limitation of this chapter is that we did not analyze chromatophore content of clown ball python scales or apply functional validation via gene editing, as we did in Chapter 1. An open question that also remains is the melanin content of scales. Variation in *mc1r* is associated with red pheomelanin phenotypes. The clown morph displays a reddish-brown colouring. Historically, it was thought that only birds and mammals produce pheomelanin. Work on a tortoise species showed the presence of pheomelanin on the shell (Roulin et al. 2013), although such analyses in snakes are lacking. Analysis of scales via alkaline hydrogen peroxide is needed to determine if snakes have pheomelanin (Ito et al. 2011). It thus remains an open question as to whether the reddish-brown colour of clown ball pythons reflects a greater proportion of pheomelanin content relative to eumelanin or just a decrease in eumelanin relative to other pigments, like pteridines.

Chapters 3 and 4 use natural selection operating in the real word and at different time scales to study how evolutionary forces shape genetic variation. In Chapter 3, I studied stickleback populations from Eastern Canada, in Nova Scotia and Newfoundland. Some lake populations are highly differentiated relative to marine and other freshwater populations, suggesting strong selection or long periods of isolation and drift. Other lake

populations show little differentiation, suggesting a more recent colonization or gene flow. The results presented in Chapter 3 indicate that these freshwater populations have a complex evolutionary history, although further study of demographic models is needed to resolve the demographic history of lake stickleback from Eastern Canada. We also found the signature of parallel evolution at loci near dopamine receptors, which suggest behavioural evolution in freshwater.

Meanwhile, Chapter 4 showed that such signatures of selection begin to take form within a single year. During seasonal isolation from the ocean, independent estuary stickleback experience genetic differentiation in parallel. We found that these changes occur at genes previously implicated in freshwater adaptation in post-glacial populations, such as osmoregulatory genes. This study only used two time points within a year. Future directions include using multi-year time series data to test for repeatedly oscillating allele frequencies year-to-year.

THESIS CONCLUSION

This thesis investigated the genetic basis of phenotypic differentiation using systems in which genetic variation has been shaped by either artificial or natural selection. Artificial selection imposed on ball pythons in captivity was used to study patterning, melanin pigmentation, pteridine pigmentation, and structural colouration in reptiles. We discovered the genetic basis of white spotting in ball pythons, adding *tfec* to the list of genes affecting patterning in vertebrates. We also established that *tfec* is required for structural colouration in lizards, although further work is required to test for this function across squamates.

We found evidence consistent with the hypothesis that in snakes it is the melanophore that is key to pattern formation. This hypothesis is supported by evidence of variation in *mc1r* being linked to changes in melanin pigmentation and reduced patterning in ball pythons (with the clown colour morph). This research has also provided evidence that axanthism in snakes can result if the first step in the conversion of GTP to pteridines is prevented, helping fill a gap in the pigmentation literature on axanthism. The lack of pteridine pigments does not influence pattern, as we expect the *mc1r* variant does by regulating melanophore differentiation. Many other variants for Mendelian phenotypes, including axanthic loci (e.g., 'MJ axanthic' colour morph), remain to be discovered in ball pythons. The work in Chapter 2 establishes the ball python as an emerging model for the study of axanthism and xanthophore biology in general.

This thesis also shed light on how colonization of new environments can impose natural selection that shapes patterns of genetic variation across multiple timescales. Lake populations of stickleback in Eastern Canada show a complex evolutionary history where

patterns of genetic variation have likely been affected by multiple colonization events, gene flow, and periods of isolation and drift. Repeated marine-freshwater differentiation near genes for dopamine receptors suggest behavioural traits under selection, although behavioural assays are required to test this idea. However, the work presented in Chapter 4 using estuary stickleback showed that such genetic and phenotypic differentiation can be observed, and thus tested, within a year.

In summary, ball pythons represent an exciting emerging model organism in pigmentation research, a model for reptiles as the zebrafish is for teleost fish. The stickleback system provided novel loci undergoing parallel evolution and allowed us to glimpse the first steps of evolution, showing that changes in allele frequency across seasons are a subtle reflection of the deeper divergence and parallel evolution of post-glacial populations.

REFERENCES

Adachi, K., Kato, K., Wakamatsu, K., Ito, S., Ishimaru, K., Murata, O. and Kumai, H., 2006. Low temperature induced discoloration of juvenile red sea bream, Pagrus major. Aquaculture Science, 54(1), pp.31-35.

Abbey-Lee, R.N., Uhrig, E.J., Zidar, J., Favati, A., Almberg, J., Dahlbom, J., Winberg, S. and Løvlie, H., 2018. The influence of rearing on behavior, brain monoamines, and gene expression in three-spined sticklebacks. Brain, behavior and evolution, 91(4), pp.201-213.

Adachi, K., Wakamatsu, K., Ito, S., Matsubara, H., Nomura, K., Tanaka, H. and Kato, K., 2010. A close relationship between androgen levels and eumelanogenesis in the teleost red seabream (Pagrus major): Quantitative analysis of its seasonal variation and effects of oral treatment with methyl-testosterone. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology, 156(2), pp.184-189.

Adhikari, K., Mendoza-Revilla, J., Sohail, A., Fuentes-Guajardo, M., Lampert, J., Chacón-Duque, J.C., Hurtado, M., Villegas, V., Granja, V., Acuña-Alonzo, V. and Jaramillo, C., 2019. A GWAS in Latin Americans highlights the convergent evolution of lighter skin pigmentation in Eurasia. Nature communications, 10(1), p.358.

Agostini, F., Agostinis, R., Medina, D.L., Bisaglia, M., Greggio, E. and Plotegher, N., 2022. The Regulation of MiTF/TFE Transcription Factors Across Model Organisms: from Brain Physiology to Implication for Neurodegeneration. Molecular Neurobiology, 59(8), pp.5000-5023.

Ahi, E.P. and Sefc, K.M., 2017. A gene expression study of dorso-ventrally restricted pigment pattern in adult fins of Neolamprologus meeli, an African cichlid species. PeerJ, 5, p.e2843.

Alves JM, Carneiro M, Cheng JY, de Matos AL, Rahman MM, Loog L, Campos PF, Wales N, Eriksson A, Manica A, Strive T. Parallel adaptation of rabbit populations to myxoma virus. Science. 2019 Mar 22;363(6433):1319-26.

Ancans, J., Tobin, D.J., Hoogduijn, M.J., Smit, N.P., Wakamatsu, K. and Thody, A.J., 2001. Melanosomal pH controls rate of melanogenesis, eumelanin/phaeomelanin ratio and melanosome maturation in melanocytes and melanoma cells. Experimental cell research, 268(1), pp.26-35.

Ando, H., Ryu, A., Hashimoto, A., Oka, M. and Ichihashi, M., 1998. Linoleic acid and α-linolenic acid lightens ultraviolet-induced hyperpigmentation of the skin. Archives of dermatological research, 290, pp.375-381.

Andrade, P. and Carneiro, M., 2021. Pterin-based pigmentation in animals. Biology Letters, 17(8), p.20210221.

Arnheiter, H. and Debbache, J., 2021. Development of Melanin-Bearing Pigment Cells in Birds and Mammals. Pigments, Pigment Cells and Pigment Patterns, pp.185-208.

Arora, N., Siddiqui, E.M. and Mehan, S., 2021. Involvement of adenylate cyclase/cAMP/CREB and SOX9/MITF in melanogenesis to prevent vitiligo. Molecular and Cellular Biochemistry, pp.1-9.

Avilés, J.M., Cruz-Miralles, Á., Ducrest, A.L., Simon, C., Roulin, A., Wakamatsu, K. and Parejo, D., 2019. Redness variation in the Eurasian scops-owl otus scops is due to pheomelanin but is not associated with variation in the melanocortin-1 receptor gene (MC1R). Ardeola, 67(1), pp.3-13.

Bailleul, R., Curantz, C., Desmarquet-Trin Dinh, C., Hidalgo, M., Touboul, J. and Manceau, M., 2019. Symmetry breaking in the embryonic skin triggers directional and sequential plumage patterning. PLoS biology, 17(10), p.e3000448.

Baker, B.I., 1993. The role of melanin-concentrating hormone in color change. Annals of the New York Academy of Sciences, 680(1), pp.279-289.

Bamshad, M.J., Nickerson, D.A. and Chong, J.X., 2019. Mendelian gene discovery: fast and furious with no end in sight. The American Journal of Human Genetics, 105(3), pp.448-455.

Baranowska Körberg, I., Sundström, E., Meadows, J.R., Rosengren Pielberg, G., Gustafson, U., Hedhammar, Å., Karlsson, E.K., Seddon, J., Söderberg, A., Vilà, C. and Zhang, X., 2014. A simple repeat polymorphism in the MITF-M promoter is a key regulator of white spotting in dogs. PLoS One, 9(8), p.e104363.

Barker, D.G., and Barker, T.M., 2006. Ball pythons: the history, natural history, care and breeding (VPI Library).

Barrett, R.D., Laurent, S., Mallarino, R., Pfeifer, S.P., Xu, C.C., Foll, M., Wakamatsu, K., Duke-Cohan, J.S., Jensen, J.D. and Hoekstra, H.E., 2019. Linking a mutation to survival in wild mice. Science, 363(6426), pp.499-504.

Barrett, R.D., Rogers, S.M. and Schluter, D., 2008. Natural selection on a major armor gene in threespine stickleback. Science, 322(5899), pp.255-257.

Barrett, R.D.H., 2010. Adaptive evolution of lateral plates in three-spined stickleback Gasterosteus aculeatus: a case study in functional analysis of natural variation. Journal of Fish Biology, 77(2), pp.311-328.

Barrett, R.D., Laurent, S., Mallarino, R., Pfeifer, S.P., Xu, C.C., Foll, M., Wakamatsu, K., Duke-Cohan, J.S., Jensen, J.D. and Hoekstra, H.E., 2019. Linking a mutation to survival in wild mice. Science, 363(6426), pp.499-504.

Barrett, R.D., Paccard, A., Healy, T.M., Bergek, S., Schulte, P.M., Schluter, D. and Rogers, S.M., 2011. Rapid evolution of cold tolerance in stickleback. Proceedings of the Royal Society B: Biological Sciences, 278(1703), pp.233-238.

Barrett, R.D., Rogers, S.M. and Schluter, D., 2008. Natural selection on a major armor gene in threespine stickleback. Science, 322(5899), pp.255-257.

Basch, M.L., Bronner-Fraser, M. and García-Castro, M.I., 2006. Specification of the neural crest occurs during gastrulation and requires Pax7. Nature, 441(7090), pp.218-222.

Bassham S, Catchen J, Lescak E, von Hippel FA, Cresko WA. Repeated selection of alternatively adapted haplotypes creates sweeping genomic remodeling in stickleback. Genetics. 2018 Jul 1;209(3):921-39.

Bastiaens, M.T., Ter Huurne, J.A., Kielich, C., Gruis, N.A., Westendorp, R.G., Vermeer, B.J., Bavinck, J.N.B. and Leiden Skin Cancer Study Team, 2001. Melanocortin-1 receptor

gene variants determine the risk of nonmelanoma skin cancer independently of fair skin and red hair. The American Journal of Human Genetics, 68(4), pp.884-894.

Batista, P.D., Janes, J.K., Boone, C.K., Murray, B.W. and Sperling, F.A., 2016. Adaptive and neutral markers both show continent-wide population structure of mountain pine beetle (Dendroctonus ponderosae). Ecology and evolution, 6(17), pp.6292-6300.

Baumgartner, J.V. and Bell, M.A., 1984. Lateral plate morph variation in California populations of the threespine stickleback, Gasterosteus aculeatus. Evolution, pp.665-674.

Baxter, L.L. and Pavan, W.J., 2013. The etiology and molecular genetics of human pigmentation disorders. Wiley Interdisciplinary Reviews: Developmental Biology, 2(3), pp.379-392.

Baxter, L.L., Hou, L., Loftus, S.K. and Pavan, W.J., 2004. Spotlight on spotted mice: a review of white spotting mouse mutants and associated human pigmentation disorders. Pigment Cell Research, 17(3), pp.215-224.

Baxter, L.L., Watkins-Chow, D.E., Pavan, W.J. and Loftus, S.K., 2019. A curated gene list for expanding the horizons of pigmentation biology. Pigment cell & melanoma research, 32(3), pp.348-358.

Bechtel, H.B., 1978. Color and pattern in snakes (Reptilia, Serpentes). Journal of Herpetology, pp.521-532.

Bechtel, H.B., 1991. Inherited color defects: Comparison between humans and snakes. International journal of dermatology, 30(4), pp.243-246.

Becker, G.S. & Reining, I.J. 2008. Steelhead/Rainbow Trout (Oncorhynchus mykiss). Resources South of the Golden Gate, California. Center for Ecosystem Restoration and Management, Oakland, CA.

Behrens, D.K., Bombardelli, F.A., Largier, J.L. and Twohy, E., 2009. Characterization of time and spatial scales of a migrating rivermouth. Geophysical Research Letters, 36(9).

Behrens, D.K., Bombardelli, F.A., Largier, J.L. and Twohy, E., 2013. Episodic closure of the tidal inlet at the mouth of the Russian River—A small bar-built estuary in California. Geomorphology, 189, pp.66-80.

Behrman, E.L., Howick, V.M., Kapun, M., Staubach, F., Bergland, A.O., Petrov, D.A., Lazzaro, B.P. and Schmidt, P.S., 2018. Rapid seasonal evolution in innate immunity of wild Drosophila melanogaster. Proceedings of the Royal Society B: Biological Sciences, 285(1870), p.20172599.

Bell, G., 2010. Fluctuating selection: the perpetual renewal of adaptation in variable environments. Philosophical Transactions of the Royal Society B: Biological Sciences, 365(1537), pp.87-97.

Bell, M.A., Aguirre, W.E. and Buck, N.J., 2004. Twelve years of contemporary armor evolution in a threespine stickleback population. Evolution, 58(4), pp.814-824.

Bell, M.A., 2001. Lateral plate evolution in the threespine stickleback: getting nowhere fast. Microevolution rate, pattern, process, pp.445-461.

Bell, M.A. and Aguirre, W.E., 2013. Contemporary evolution, allelic recycling, and adaptive radiation of the threespine stickleback. Evolutionary Ecology Research, 15(4), pp.377-411.

Bell, M.A. and Richkind, K.E., 1981. Clinal variation of lateral plates in threespine stickleback fish. The American Naturalist, 117(2), pp.113-132.

Bellono, N.W., Escobar, I.E., Lefkovith, A.J., Marks, M.S. and Oancea, E., 2014. An intracellular anion channel critical for pigmentation. Elife, 3, p.e04543.

Benirschke, K., 2004. Francis Galton: Pioneer of Heredity and Biometry.

Bennett, D.C. and Lamoreux, M.L., 2003. The color loci of mice–a genetic century. Pigment cell research, 16(4), pp.333-344.

Bergland, A.O., Behrman, E.L., O'Brien, K.R., Schmidt, P.S. and Petrov, D.A., 2014. Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in Drosophila. PLoS genetics, 10(11), p.e1004775.

Bertin, L., 1925. Recherches bionomiques, biométriques et systématiques sur les épinoches gastérostéidés). Am. Inst. Oceanograph. N. S. Z. Frac. 1: 1-204.

Betancur-R, R., Ortí, G. and Pyron, R.A., 2015. Fossil-based comparative analyses reveal ancient marine ancestry erased by extinction in ray-finned fishes. Ecology Letters, 18(5), pp.441-450.

Bharti, K., Gasper, M., Ou, J., Brucato, M., Clore-Gronenborn, K., Pickel, J. and Arnheiter, H., 2012. A regulatory loop involving PAX6, MITF, and WNT signaling controls retinal pigment epithelium development. PLoS genetics, 8(7), p.e1002757.

Bissegger, M., Laurentino, T.G., Roesti, M. and Berner, D., 2020. Widespread intersex differentiation across the stickleback genome–The signature of sexually antagonistic selection?. Molecular Ecology, 29(2), pp.262-271.

Bizzo, N. and El-Hani, C.N., 2009. Darwin and Mendel: evolution and genetics. Journal of Biological Education, 43(3), pp.108-114.

Borteiro, C., Diesel Abegg, A., Hirouki Oda, F., Cardozo, D.E., Kolenc, F., Etchandy, I., Bisaiz, I., Prigioni, C. and Baldo, J.D., 2021. Aberrant colouration in wild snakes: case study in Neotropical taxa and a review of terminology.

Boughman, J.W., Rundle, H.D. and Schluter, D., 2005. Parallel evolution of sexual isolation in sticklebacks. Evolution, 59(2), pp.361-373.

Bridges, C.B., 1916. Non-disjunction as proof of the chromosome theory of heredity (concluded). Genetics, 1(2), p.107.

Brooks, B.P., Zand, D.J., Hufnagel, R.B., Sharma, R., Sergeev, Y.V., Gamm, D.M., Bharti, K. and George, A., 2018. COMMAD: a novel syndrome caused by biallelic mutation of the MITF gene. Journal of American Association for Pediatric Ophthalmology and Strabismus {JAAPOS}, 22(4), p.e28.

Brown, A.R., Comai, K., Mannino, D., McCullough, H., Donekal, Y., Meyers, H.C., Graves, C.W., Seidel, H.S. and BIO306W Consortium, 2022. A community-science approach identifies genetic variants associated with three color morphs in ball pythons (Python regius). Plos one, 17(10), p.e0276376.

Bruders, R., Van Hollebeke, H., Osborne, E.J., Kronenberg, Z., Maclary, E., Yandell, M. and Shapiro, M.D., 2020. A copy number variant is associated with a spectrum of pigmentation patterns in the rock pigeon (Columba livia). Plos Genetics, 16(5), p.e1008274.

Bulmer, M., 1998. Galton's law of ancestral heredity. Heredity, 81(5), pp.579-585.

Burri R. Interpreting differentiation landscapes in the light of long-term linked selection. Evolution Letters. 2017 Aug;1(3):118-31.

Campbell-Staton SC, Cheviron ZA, Rochette N, Catchen J, Losos JB, Edwards SV. Winter storms drive rapid phenotypic, regulatory, and genomic shifts in the green anole lizard. Science. 2017 Aug 4;357(6350):495-8.

Card DC, Perry BW, Adams RH, Schield DR, Young AS, Andrew AL, Jezkova T, Pasquesi GI, Hales NR, Walsh MR, Rochford MR. Novel ecological and climatic conditions drive rapid adaptation in invasive Florida Burmese pythons. Molecular ecology. 2018 Dec;27(23):4744-57.

Caro, T., 2017. Wallace on coloration: contemporary perspective and unresolved insights. Trends in ecology & evolution, 32(1), pp.23-30.

Caro, T., Hill, G., Lindström, L. and Speed, M., 2008. The Colours of Animals: From Wallace to the Present Day II. Conspicuous Coloration. NATURAL SELECTION AND BEYOND, p.144.

Caro, T., Stoddard, M.C. and Stuart-Fox, D., 2017. Animal coloration: production, perception, function and application. Philosophical Transactions of the Royal Society B: Biological Sciences, 372(1724), p.20170047.

Castle, W.E. and Little, C.C., 1909. The Peculiar Inheritance pf Pink Eyes Among Colored Mice. Science, 30(766), pp.313-314.

Castle, W.E. and Little, C.C., 1910. On a modified Mendelian ratio among yellow mice. Science, 32(833), pp.868-870.

Castle, W.E., 1903. Mendel's law of heredity. Science, 18(456), pp.396-406.

Castle, W.E., 1919. Piebald rats and the theory of genes. Proceedings of the National Academy of Sciences, 5(4), pp.126-130.

Castoe, T.A., De Koning, A.J., Hall, K.T., Card, D.C., Schield, D.R., Fujita, M.K., Ruggiero, R.P., Degner, J.F., Daza, J.M., Gu, W. and Reyes-Velasco, J., 2013. The Burmese python genome reveals the molecular basis for extreme adaptation in snakes. Proceedings of the National Academy of Sciences, 110(51), pp.20645-20650.

Cattaneo, A., 2015. The Zamenis longissimus (Laurenti) axanthic phenotype found on the Castelporziano Presidential Estate: considerations on its morphology, genetic nature and probable extinction (Serpentes: Colubridae). Rendiconti Lincei, 26, pp.385-389.

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal Jr, G., Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J. and Myers, R.M., 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. science, 327(5963), pp.302-305.

Chen, J.C., Lo, Y.F., Lin, Y.W., Lin, S.H., Huang, C.L. and Cheng, C.J., 2019. WNK4 kinase is a physiological intracellular chloride sensor. Proceedings of the National Academy of Sciences, 116(10), pp.4502-4507.

Chin D, Means AR. Calmodulin: a prototypical calcium sensor. Trends in cell biology. 2000 Aug 1;10(8):322-8.

Chong, J.X., Buckingham, K.J., Jhangiani, S.N., Boehm, C., Sobreira, N., Smith, J.D., Harrell, T.M., McMillin, M.J., Wiszniewski, W., Gambin, T. and Akdemir, Z.H.C., 2015. The genetic basis of Mendelian phenotypes: discoveries, challenges, and opportunities. The American Journal of Human Genetics, 97(2), pp.199-215.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X. and Ruden, D.M., 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly, 6(2), pp.80-92.

Coleman, M.A., Minne, A.J., Vranken, S. and Wernberg, T., 2020. Genetic tropicalisation following a marine heatwave. Scientific reports, 10(1), p.12726.

Colosimo, P.F., Hosemann, K.E., Balabhadra, S., Villarreal, G., Dickson, M., Grimwood, J., Schmutz, J., Myers, R.M., Schluter, D. and Kingsley, D.M., 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. science, 307(5717), pp.1928-1933.

Comeault, A.A., Carvalho, C.F., Dennis, S., Soria-Carrasco, V. and Nosil, P., 2016. Color phenotypes are under similar genetic control in two distantly related species of Timema stick insect. Evolution, 70(6), pp.1283-1296.

Concordet, J.P. and Haeussler, M., 2018. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. Nucleic acids research, 46(W1), pp.W242-W245.

Cooper, C.D., 2017. Insights from zebrafish on human pigment cell disease and treatment. Developmental dynamics, 246(11), pp.889-896.

Cott, H.B., 1940. Adaptive coloration in animals.

Cuenot, L.C.M.J., 1905. Les races pures et leurs combinaisons chez les Souris (Notes et Revue). Archives de Zoologie Experimentale, 3(7).

Cuthill, I.C., Allen, W.L., Arbuckle, K., Caspers, B., Chaplin, G., Hauber, M.E., Hill, G.E., Jablonski, N.G., Jiggins, C.D., Kelber, A. and Mappes, J., 2017. The biology of color. Science, 357(6350), p.eaan0221.

Côte, J., Boniface, A., Blanchet, S., Hendry, A.P., Gasparini, J. and Jacquin, L., 2018. Melanin-based coloration and host–parasite interactions under global change. Proceedings of the Royal Society B: Biological Sciences, 285(1879), p.20180285.

d'Ischia, M., Wakamatsu, K., Napolitano, A., Briganti, S., Garcia-Borron, J.C., Kovacs, D., Meredith, P., Pezzella, A., Picardo, M., Sarna, T. and Simon, J.D., 2013. Melanins and melanogenesis: methods, standards, protocols. Pigment cell & melanoma research, 26(5), pp.616-633.

Dabney, A., Storey, J.D. and Warnes, G.R., 2010. qvalue: Q-value estimation for false discovery rate control. R package version, 1(0).

Dannemann, M. and Kelso, J., 2017. The contribution of Neanderthals to phenotypic variation in modern humans. The American journal of human genetics, 101(4), pp.578-589.

Dao, U.M., Lederer, I., Tabor, R.L., Shahid, B., Graves, C.W. and Seidel, H.S., 2022. Leucism and stripe formation in ball pythons (Python regius) are associated with variants affecting endothelin signaling. bioRxiv.

Darwin, C., 1887. The life and letters of Charles Darwin: including an autobiographical chapter (Vol. 1). New York: D. Appleton 1887..

Davidson, G., Shen, J., Huang, Y.L., Su, Y., Karaulanov, E., Bartscherer, K., Hassler, C., Stannek, P., Boutros, M. and Niehrs, C., 2009. Cell cycle control of wnt receptor activation. Developmental cell, 17(6), pp.788-799.

Davison, A., Jackson, H.J., Murphy, E.W. and Reader, T., 2019. Discrete or indiscrete? Redefining the colour polymorphism of the land snail Cepaea nemoralis. Heredity, 123(2), pp.162-175.

Dawes, J.H. and Kelsh, R.N., 2021. Cell fate decisions in the neural crest, from pigment cell to neural development. International Journal of Molecular Sciences, 22(24), p.13531.

DeFaveri, J., Shikano, T., Shimada, Y., Goto, A. and Merilä, J., 2011. Global analysis of genes involved in freshwater adaptation in threespine sticklebacks (Gasterosteus aculeatus). Evolution, 65(6), pp.1800-1807.

Deagle, B.E., Jones, F.C., Absher, D.M., Kingsley, D.M. and Reimchen, T.E., 2013. Phylogeography and adaptation genetics of stickleback from the H’aida G’waii

archipelago revealed using genome-wide single nucleotide polymorphism genotyping. Molecular ecology, 22(7), pp.1917-1932.

del Marmol V, Ito S, Bouchard B et al (1996) Cysteine deprivation promotes eumelanogenesis in human melanoma cells. J Invest Dermatol 107:698–702

Dennenmoser, S., Vamosi, S.M., Nolte, A.W. and Rogers, S.M., 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(1), pp.25-42.

Des Roches, S., Bell, M.A. and Palkovacs, E.P., 2020. Climate-driven habitat change causes evolution in Threespine Stickleback. Global Change Biology, 26(2), pp.597-606.

Des Roches, S., Sollmann, R., Calhoun, K., Rothstein, A.P. and Rosenblum, E.B., 2017. Survival by genotype: patterns at Mc1r are not black and white at the W hite S ands ecotone. Molecular Ecology, 26(1), pp.320-329.

Di Poi, C., Bélanger, D., Amyot, M., Rogers, S. and Aubin-Horth, N., 2016. Receptors rather than signals change in expression in four physiological regulatory networks during evolutionary divergence in threespine stickleback. Molecular ecology, 25(14), pp.3416-3427.

Di-Poi, C., Lacasse, J., Rogers, S.M. and Aubin-Horth, N., 2014. Extensive behavioural divergence following colonisation of the freshwater environment in threespine sticklebacks. PloS one, 9(6), p.e98980.

Di-Poï, N. and Milinkovitch, M.C., 2016. The anatomical placode in reptile scale morphogenesis indicates shared ancestry among skin appendages in amniotes. Science advances, 2(6), p.e1600708.

Ding, Q., Hu, Y., Xu, S., Wang, C.C., Li, H., Zhang, R., Yan, S., Wang, J. and Jin, L., 2014. Neanderthal origin of the haplotypes carrying the functional variant Val92Met in the MC1R in modern humans. Molecular biology and evolution, 31(8), pp.1994-2003.

Divino, J.N., Monette, M.Y., McCormick, S.D., Yancey, P.H., Flannery, K.G., Bell, M.A., Rollins, J.L., von Hippel, F.A. and Schultz, E.T., 2016. Osmoregulatory physiology and rapid evolution of salinity tolerance in threespine stickleback recently introduced to fresh water. Evolutionary ecology research, 17(2), pp.179-201.

Domyan, E.T., Hardy, J., Wright, T., Frazer, C., Daniels, J., Kirkpatrick, J., Kirkpatrick, J., Wakamatsu, K. and Hill, J.T., 2019. SOX10 regulates multiple genes to direct eumelanin versus pheomelanin production in domestic rock pigeon. Pigment Cell & Melanoma Research, 32(5), pp.634-642.

Donihue, C.M., Herrel, A., Fabre, A.C., Kamath, A., Geneva, A.J., Schoener, T.W., Kolbe, J.J. and Losos, J.B., 2018. Hurricane-induced selection on the morphology of an island lizard. Nature, 560(7716), pp.88-91.

Ducrest, A.L., Keller, L. and Roulin, A., 2008. Pleiotropy in the melanocortin system, coloration and behavioural syndromes. Trends in ecology & evolution, 23(9), pp.502-510.

Dudchenko, O., Batra, S.S., Omer, A.D., Nyquist, S.K., Hoeger, M., Durand, N.C., Shamim, M.S., Machol, I., Lander, E.S., Aiden, A.P. and Aiden, E.L., 2017. De novo assembly of

the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. Science, 356(6333), pp.92-95.

Dudchenko, O., Shamim, M.S., Batra, S.S., Durand, N.C., Musial, N.T., Mostofa, R., Pham, M., Hilaire, B.S., Yao, W., Stamenova, E. and Hoeger, M., 2018. The Juicebox Assembly Tools module facilitates de novo assembly of mammalian genomes with chromosome-length scaffolds for under $1000. BioRxiv, 254797. Preprint.

Dunn, L.C. and Charles, D.R., 1937. Studies on spotting patterns I. Analysis of quantitative variations in the pied spotting of the house mouse. Genetics, 22(1), p.14.

Dunn, L.C., 1920. Types of white spotting in mice. The american naturalist, 54(635), pp.465-495.

Dunn, L.C., 1937. Studies on spotting patterns II. Genetic analysis of variegated spotting in the house mouse. Genetics, 22(1), p.43.

D'Cruze, N., Paterson, S., Green, J., Megson, D., Warwick, C., Coulthard, E., Norrey, J., Auliya, M. and Carder, G., 2020. Dropping the ball? The welfare of ball pythons traded in the EU and North America. Animals, 10(3), p.413.

Efron, B., 2004. Large-scale simultaneous hypothesis testing: the choice of a null hypothesis. Journal of the American Statistical Association, 99(465), pp.96-104.

Elmer KR, Meyer A. Adaptation in the age of ecological genomics: insights from parallelism and convergence. Trends in ecology & evolution. 2011 Jun 1;26(6):298-306.

Endler, J.A. and Mappes, J., 2017. The current and future state of animal coloration research. Philosophical Transactions of the Royal Society B: Biological Sciences, 372(1724), p.20160352.

Eom, D.S., 2020. Airinemes: thin cellular protrusions mediate long-distance signalling guided by macrophages. Open biology, 10(8), p.200039.

Eom, D.S., Bain, E.J., Patterson, L.B., Grout, M.E. and Parichy, D.M., 2015. Long-distance communication by specialized cellular projections during pigment pattern development and evolution. Elife, 4, p.e12401.

Estes, S. and Arnold, S.J., 2007. Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence on all timescales. The american naturalist, 169(2), pp.227-244.

Evans, D.H., Piermarini, P.M. and Choe, K.P., 2005. The multifunctional fish gill: dominant site of gas exchange, osmoregulation, acid-base regulation, and excretion of nitrogenous waste. Physiological reviews, 85(1), pp.97-177.

Fang, B., Kemppainen, P., Momigliano, P., Feng, X. and Merilä, J., 2020. On the causes of geographically heterogeneous parallel evolution in sticklebacks. Nature Ecology & Evolution, 4(8), pp.1105-1115.

Fang, B., Merilä, J., Matschiner, M. and Momigliano, P., 2020. Estimating uncertainty in divergence times among three-spined stickleback clades using the multispecies coalescent. Molecular phylogenetics and evolution, 142, p.106646.

Fang, B., Merilä, J., Ribeiro, F., Alexandre, C.M. and Momigliano, P., 2018. Worldwide phylogeny of three-spined sticklebacks. Molecular phylogenetics and evolution, 127, pp.613-625.

Fang, W., Huang, J., Li, S. and Lu, J., 2022. Identification of pigment genes (melanin, carotenoid and pteridine) associated with skin color variant in red tilapia using transcriptome analysis. Aquaculture, 547, p.737429.

Feder, A.F., Petrov, D.A. and Bergland, A.O., 2012. LDx: estimation of linkage disequilibrium from high-throughput pooled resequencing data. PloS one, 7(11), p.e48588.

Feng, Y., McQuillan, M.A. and Tishkoff, S.A., 2021. Evolutionary genetics of skin pigmentation in African populations. Human Molecular Genetics, 30(R1), pp.R88-R97.

Filby, A.L., Paull, G.C., Hickmore, T.F. and Tyler, C.R., 2010. Unravelling the neurophysiological basis of aggression in a fish model. BMC genomics, 11(1), pp.1-17.

Fofonjka, A. and Milinkovitch, M.C., 2021. Reaction-diffusion in a growing 3D domain of skin scales generates a discrete cellular automaton. Nature communications, 12(1), pp.1-13.

Foll, M., Gaggiotti, O.E., Daub, J.T., Vatsiou, A. and Excoffier, L., 2014. Widespread signals of convergent adaptation to high altitude in Asia and America. The American Journal of Human Genetics, 95(4), pp.394-407.

François, O., Martins, H., Caye, K. and Schoville, S.D., 2016. Controlling false discoveries in genome scans for selection. Molecular ecology, 25(2), pp.454-469.

Frohnhöfer, H.G., Krauss, J., Maischein, H.M. and Nüsslein-Volhard, C., 2013. Iridophores and their interactions with other chromatophores are required for stripe formation in zebrafish. Development, 140(14), pp.2997-3007.

Frost, S.K., Epp, L.G. and Robinson, S.J., 1986. The pigmentary system of developing axolotls: IV. An analysis of the axanthic phenotype.

Fuller, B.B., Spaulding, D.T. and Smith, D.R., 2001. Regulation of the catalytic activity of preexisting tyrosinase in black and Caucasian human melanocyte cell cultures. Experimental cell research, 262(2), pp.197-208.

Gallone, A., Sagliano, A., Guida, G., Ito, S., Wakamatsu, K., Capozzi, V.I.T.O., Perna, G., Zanna, P. and Cicero, R., 2007. The melanogenic system of the liver pigmented macrophages of Rana esculenta L.-Tyrosinase activity. Histology and histopathology.

Galloway, J., Cresko, W.A. and Ralph, P., 2020. A few stickleback suffice for the transport of alleles to new lakes. G3: Genes, Genomes, Genetics, 10(2), pp.505-514.

Garcia-Elfring, A., Sabin, C. E., Iouchmanov, A. L., Roffey, H. L., Samudra, S. P., Alcala, A. J., Osman, R. S., Lauderdale, J. D., Hendry, A. P., Menke, D. B. and Barrett, R. D. H. 2023. Piebaldism and chromatophore development in reptiles are linked to the tfec gene. Current Biology, 33, 755-763.e3.

Gaudel, C., Soysouvanh, F., Leclerc, J., Bille, K., Husser, C., Montcriol, F., Bertolotto, C. and Ballotti, R., 2020. Regulation of melanogenesis by the amino acid transporter SLC7A5. Journal of Investigative Dermatology, 140(11), pp.2253-2259.

Ge, S.X., Jung, D. and Yao, R., 2020. ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics, 36(8), pp.2628-2629.

Geneva, A.J., Park, S., Bock, D.G., de Mello, P.L., Sarigol, F., Tollis, M., Donihue, C.M., Reynolds, R.G., Feiner, N., Rasys, A.M. and Lauderdale, J.D., 2022. Chromosome-scale genome assembly of the brown anole (Anolis sagrei), an emerging model species. Communications Biology, 5(1), p.1126.

George, A., Zand, D.J., Hufnagel, R.B., Sharma, R., Sergeev, Y.V., Legare, J.M., Rice, G.M., Schwoerer, J.A.S., Rius, M., Tetri, L. and Gamm, D.M., 2016. Biallelic mutations in MITF cause coloboma, osteopetrosis, microphthalmia, macrocephaly, albinism, and deafness. The American Journal of Human Genetics, 99(6), pp.1388-1394.

Gibbons, T.C., Rudman, S.M. and Schulte, P.M., 2017. Low temperature and low salinity drive putatively adaptive growth differences in populations of threespine stickleback. Scientific reports, 7(1), p.16766.

Gibbons, T.C., Rudman, S.M. and Schulte, P.M., 2016. Responses to simulated winter conditions differ between threespine stickleback ecotypes. Molecular ecology, 25(3), pp.764-775.

Goda, M. and Kuriyama, T., 2021. Physiological and morphological color changes in teleosts and in reptiles. Pigments, Pigment Cells and Pigment Patterns, pp.387-423.

Goding, C.R. and Arnheiter, H., 2019. MITF—the first 25 years. Genes & development, 33(15-16), pp.983-1007.

Graham, A.M. and McCracken, K.G., 2019. Convergent evolution on the hypoxia-inducible factor (HIF) pathway genes EGLN1 and EPAS1 in high-altitude ducks. Heredity, 122(6), pp.819-832.

Grant PR, Grant BR, Huey RB, Johnson MT, Knoll AH, Schmitt J. Evolution caused by extreme events. Philosophical Transactions of the Royal Society B: Biological Sciences. 2017 Jun 19;372(1723):20160146.

Greenhill, E.R., Rocco, A., Vibert, L., Nikaido, M. and Kelsh, R.N., 2011. An iterative genetic and dynamical modelling approach identifies novel features of the gene regulatory network underlying melanocyte development. PLoS Genetics, 7(9), p.e1002265.

Griffiths SP, West RJ. Preliminary assessment of shallow water fish in three small intermittently open estuaries in southeastern Australia. Fisheries Management and Ecology. 1999 Aug;6(4):311-21.

Gross, J.B., Borowsky, R. and Tabin, C.J., 2009. A novel role for Mc1r in the parallel evolution of depigmentation in independent populations of the cavefish Astyanax mexicanus. PLoS genetics, 5(1), p.e1000326.

Guernsey, M.W., Ritscher, L., Miller, M.A., Smith, D.A., Schöneberg, T. and Shapiro, M.D., 2013. A Val85Met mutation in melanocortin-1 receptor is associated with reductions in eumelanic pigmentation and cell surface expression in domestic rock pigeons (Columba livia). PLoS One, 8(8), p.e74475.

Guo, B., DeFaveri, J., Sotelo, G., Nair, A. and Merilä, J., 2015. Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. BMC biology, 13, pp.1-18.

Guo, L., Bloom, J., Sykes, S., Huang, E., Kashif, Z., Pham, E., Ho, K., Alcaraz, A., Xiao, X.G., Duarte-Vogel, S. and Kruglyak, L., 2021. Genetics of white color and iridophoroma in "Lemon Frost" leopard geckos. PLoS genetics, 17(6), p.e1009580.

Gur, D., Bain, E.J., Johnson, K.R., Aman, A.J., Pasolli, H.A., Flynn, J.D., Allen, M.C., Deheyn, D.D., Lee, J.C., Lippincott-Schwartz, J. and Parichy, D.M., 2020. In situ differentiation of iridophore crystallotypes underlies zebrafish stripe patterning. Nature Communications, 11(1), p.6391..

Gur, D., Palmer, B.A., Weiner, S. and Addadi, L., 2017. Light manipulation by guanine crystals in organisms: biogenic scatterers, mirrors, multilayer reflectors and photonic crystals. Advanced functional materials, 27(6), p.1603514.

Gurgel CF, Camacho O, Minne AJ, Wernberg T, Coleman MA. Marine heatwave drives cryptic loss of genetic diversity in underwater forests. Current Biology. 2020 Feb 27.

Gutiérrez-Gil, B., Wiener, P. and Williams, J.L., 2007. Genetic effects on coat colour in cattle: dilution of eumelanin and phaeomelanin pigments in an F2-Backcross Charolais× Holstein population. BMC genetics, 8(1), pp.1-12

Haenel Q, Roesti M, Moser D, MacColl AD, Berner D. Predictable genome-wide sorting of standing genetic variation during parallel adaptation to basic versus acidic environments in stickleback fish. Evolution letters. 2019 Feb;3(1):28-42.

Haffter, P., Odenthal, J., Mullins, M.C., Lin, S., Farrell, M.J., Vogelsang, E., Haas, F., Brand, M., van Eeden, F.J., Furutani-Seiki, M. and Granato, M., 1996. Mutations affecting pigmentation and shape of the adult zebrafish. Development genes and evolution, 206, pp.260-276.

Hagen, D.W. and Moodie, G.E.E., 1982. Polymorphism for plate morphs in Gasterosteus aculeatus on the east coast of Canada and an hypothesis for their global distribution. Canadian Journal of Zoology, 60(5), pp.1032-1042.

Haines, G.E., 2022. Intraspecific diversity of threespine stickleback (Gasterosteus aculeatus) populations in eastern Canada. Environmental Biology of Fishes, pp.1-18.

Hall, B.K., 2009. The neural crest and neural crest cells in vertebrate development and evolution (Vol. 11). Springer Science & Business Media.

Hartl, D.L., Clark, A.G. and Clark, A.G., 1997. Principles of population genetics (Vol. 116). Sunderland, MA: Sinauer associates.

Hasan MM, DeFaveri J, Kuure S, Dash SN, Lehtonen S, Merilä J, McCairns RS. Sticklebacks adapted to divergent osmotic environments show differences in plasticity for kidney morphology and candidate gene expression. Journal of Experimental Biology. 2017 Jun 15;220(12):2175-86.

Hasegawa, M., Kuriyama, T., Brandley, M., and Murakami, A. (2020). Blue, black, and stripes: evolution and development of color production and pattern formation in lizards and snakes. Frontiers in Ecology and Evolution 8, 232.

Hashimoto, H., Goda, M. and Kelsh, R.N., 2021. Pigment cell development in teleosts. Pigments, pigment cells and pigment patterns, pp.209-246.

Haupaix, N., Curantz, C., Bailleul, R., Beck, S., Robic, A. and Manceau, M., 2018. The periodic coloration in birds forms through a prepattern of somite origin. Science, 361(6408).

Hauser, M., Signer-Hasler, H., Küttel, L., Capitan, A., Guldbrandtsen, B., Hinrichs, D., Flury, C., Seefried, F.R. and Drögemüller, C., 2022. Identification of two new recessive MC1R alleles in red-coloured Evolèner cattle and other breeds. Animal genetics, 53(3), pp.427-435.

Hauswirth, R., Haase, B., Blatter, M., Brooks, S.A., Burger, D., Drögemüller, C., Gerber, V., Henke, D., Janda, J., and Jude, R., 2019. Correction: Mutations in MITF and PAX3 Cause "SplashedWhite" and Other White Spotting Phenotypes in Horses. PLoS genetics 15, e1008321.

Heckwolf MJ, Meyer BS, Häsler R, Höppner MP, Eizaguirre C, Reusch TB., 2019. Two different epigenetic pathways detected in wild three-spined sticklebacks are involved in salinity adaptation. bioRxiv. 1:649574.

Hedley, J., Johnson, R. and Yeates, J., 2018. Reptiles (Reptilia). Companion Animal Care and Welfare: The UFAW Companion Animal Handbook, pp.371-394.

Hejna, M., Moon, W.M., Cheng, J., Kawakami, A., Fisher, D.E., and Song, J.S. (2019). Local genomic features predict the distinct and overlapping binding patterns of the bHLH-Zip family oncoproteins MITF and MYC-MAX. Pigment cell & melanoma research 32, 500-509.

Hellström, A.R., Watt, B., Fard, S.S., Tenza, D., Mannström, P., Narfström, K., Ekesten, B., Ito, S., Wakamatsu, K., Larsson, J. and Ulfendahl, M., 2011. Inactivation of Pmel alters melanosome shape but has only a subtle effect on visible pigmentation. PLoS genetics, 7(9), p.e1002285.

Hendry AP, Bolnick DI, Berner D, Peichel CL. Along the speciation continuum in sticklebacks. Journal of fish biology. 2009 Nov;75(8):2000-36.

Hendry AP, Peichel CL, Matthews B, Boughman JW, Nosil P. Stickleback research: the now and the next. Evolutionary Ecology Research. 2013;15(2):111-41.

Hendry, A.P. and Kinnison, M.T., 1999. Perspective: the pace of modern life: measuring rates of contemporary microevolution. Evolution, 53(6), pp.1637-1653.

Hendry, A.P., Peichel, C.L., Matthews, B., Boughman, J.W. and Nosil, P., 2013. Stickleback research: the now and the next. Evolutionary Ecology Research, 15(2), pp.111-141.

Henson, S.M., Cushing, J.M. and Hayward, J., 2015. Introduction to special issue on eco-evolutionary dynamics.

Heuts, M.J., 1947. Experimental studies on adaptive evolution in Gasterosteus aculeatus L. Evolution, pp.89-102.

Heuts, M.J., 1947. The Phenotypical Variability of Gasterosteus Aculeatus (L.) Populations in Belgium.... NV Standaard-boekhandel.

Hirano, M. and Ueno, S., 1999. Mutant GTP cyclohydrolase I in autosomal dominant dystonia and recessive hyperphenylalaninemia. Neurology, 52(1), pp.182-182.

Hirobe, T., 2011. How are proliferation and differentiation of melanocytes regulated?. Pigment cell & melanoma research, 24(3), pp.462-478.

Hirobe, T., Ito, S. and Wakamatsu, K., 2011. The mouse pink-eyed dilution allele of the P-gene greatly inhibits eumelanin but not pheomelanin synthesis. Pigment Cell & Melanoma Research, 24(1), pp.241-246.

Hirobe, T., Wakamatsu, K. and Ito, S., 2007. The eumelanin and pheomelanin contents in dorsal hairs of female recessive yellow mice are greater than in male. Journal of dermatological science, 45(1), pp.55-62.

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP., 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. Science. 7;313(5783):101-4.

Hoekstra, H.E., 2006. Genetics, development and evolution of adaptive pigmentation in vertebrates. Heredity, 97(3), pp.222-234.

Hohenlohe PA, Magalhaes IS., 2019. The Population Genomics of Parallel Adaptation: Lessons from Threespine Stickleback. In: Oleksiak M., Rajora O. (eds) Population Genomics: Marine Organisms. Population Genomics. Springer, Cham

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A. and Cresko, W.A., 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS genet, 6(2), p.e1000862.

Horn KM, Williams BW, Erséus C, Halanych KM, Santos SR, Creuzé des Châtelliers M, Anderson FE., 2019. Na+/K+-ATP ase gene duplications in clitellate annelids are associated with freshwater colonization. Journal of evolutionary biology.

Hu, S., Bai, S., Dai, Y., Yang, N., Li, J., Zhang, X., Wang, F., Zhao, B., Bao, G., and Chen, Y. (2021). Deubiquitination of MITF-M Regulates Melanocytes Proliferation and Apoptosis. Frontiers in Molecular Biosciences 8, 566.

Huang, D., Lewis, V.M., Foster, T.N., Toomey, M.B., Corbo, J.C., and Parichy, D.M. (2021). Development and genetics of red coloration in the zebrafish relative Danio albolineatus. Elife 10. 10.7554/eLife.70253.

Hudon, J., 2005. Considerations in the conservation of feathers and hair, particularly their pigments. Fur trade legacy. The preservation of organic materials, pp.127-147.

Morgan, T.H., 1910. Sex limited inheritance in Drosophila. Science, 32(812), pp.120-122.

Ignatius, M.S., Moose, H.E., El-Hodiri, H.M. and Henion, P.D., 2008. colgate/hdac1 Repression of foxd3 expression is required to permit mitfa-dependent melanogenesis. Developmental biology, 313(2), pp.568-583.

Ilardo MA, Moltke I, Korneliussen TS, Cheng J, Stern AJ, Racimo F, de Barros Damgaard P, Sikora M, Seguin-Orlando A, Rasmussen S, van den Munckhof IC. Physiological and genetic adaptations to diving in sea nomads. Cell. 2018 Apr 19;173(3):569-80.

Irion, U., and Nüsslein-Volhard, C., 2022. Developmental genetics with model organisms. Proceedings of the National Academy of Sciences 119, e2122148119.

Irizarry, K.J. and Bryden, R.L., 2016. In silico analysis of gene expression network components underlying pigmentation phenotypes in the Python identified evolutionarily conserved clusters of transcription factor binding sites. Advances in bioinformatics.

Ishikawa, A., Sugiyama, M., Hondo, E., Kinoshita, K., and Yamagishi, Y. 2015. Development of a novel pink-eyed dilution mouse model showing progressive darkening of the eyes and coat hair with aging. Exp Anim 64, 207-220. 10.1538/expanim.14-0101.

Ishikawa, A., Kusakabe, M., Kume, M. and Kitano, J., 2016. Comparison of freshwater tolerance during spawning migration between two sympatric Japanese marine threespine stickleback species. Evolutionary Ecology Research, 17(4), pp.525-534.

Ito, S. and Wakamatsu, K., 2011. Diversity of human hair pigmentation as studied by chemical analysis of eumelanin and pheomelanin. Journal of the European Academy of Dermatology and Venereology, 25(12), pp.1369-1380.

Ito, S. and Wakamatsu, K., 2011. Human hair melanins: what we have learned and have not learned from mouse coat color pigmentation. Pigment cell & melanoma research, 24(1), pp.63-74.

Ito, S., Nakanishi, Y., Valenzuela, R.K., Brilliant, M.H., Kolbe, L. and Wakamatsu, K., 2011. Usefulness of alkaline hydrogen peroxide oxidation to analyze eumelanin and pheomelanin in various tissue samples: application to chemical analysis of human hair melanins. Pigment cell & melanoma research, 24(4), pp.605-613.

Ito, S., Suzuki, N., Takebayashi, S., Commo, S. and Wakamatsu, K., 2013. Neutral p H and copper ions promote eumelanogenesis after the dopachrome stage. Pigment cell & melanoma research, 26(6), pp.817-825.

Itou, T., 2018. Morphological changes in hair melanosomes by aging. Pigment Cell & Melanoma Research, 31(5), pp.630-635.

Iwanishi, S., Zaitsu, S., Shibata, H. and Nitasaka, E., 2018. An albino mutant of the Japanese rat snake (Elaphe climacophora) carries a nonsense mutation in the tyrosinase gene. Genes & Genetic Systems, 93(4), pp.163-167.

Jablonski, D., Alena, A., Vlček, P. and Jandzik, D., 2014. Axanthism in amphibians: A review and the first record in the widespread toad of the Bufotes viridis complex (Anura: Bufonidae). Belgian Journal of Zoology, 144(2).

Jackson, I.J. 1997. Homologous pigmentation mutations in human, mouse and other model organisms. Hum Mol Genet 6, 1613-1624. 10.1093/hmg/6.10.1613.

Jahanbakhsh, E., and Milinkovitch, M.C. (2022). Modeling convergent scale-by-scale skin color patterning in multiple species of lizards. Current Biology 32, 5069-5082. e5013.

Jeffries KM, Connon RE, Davis BE, Komoroske LM, Britton MT, Sommer T, Todgham AE, Fangue NA. Effects of high temperatures on threatened estuarine fishes during periods of extreme drought. Journal of Experimental Biology. 2016 Jun 1;219(11):1705-16.

Jin, Y., Tong, H., Shao, G., Li, J., Lv, Y., Wo, Y., Brown, R.P. and Fu, C., 2020. Dorsal pigmentation and its association with functional variation in MC1R in a lizard from different elevations on the Qinghai–Tibetan plateau. Genome Biology and Evolution, 12(12), pp.2303-2313.

Johnson, O.L., Tobler, R., Schmidt, J.M. and Huber, C.D., 2023. Fluctuating selection and the determinants of genetic variation. Trends in Genetics.

Jones FC, Chan YF, Schmutz J, Grimwood J, Brady SD, Southwick AM, Absher DM, Myers RM, Reimchen TE, Deagle BE, Schluter D.2012 A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. Current biology. 2012b Jan 10;22(1):83-90.

Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, Johnson J, Swofford R, Pirun M, Zody MC, White S, Birney E. 2012a The genomic basis of adaptive evolution in threespine sticklebacks. Nature. Apr;484(7392):55.

Jones MR, Mills LS, Alves PC, Callahan CM, Alves JM, Lafferty DJ, Jiggins FM, Jensen JD, Melo-Ferreira J, Good JM., 2018. Adaptive introgression underlies polymorphic seasonal camouflage in snowshoe hares. Science. Jun 22;360(6395):1355-8.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S. and Birney, E., 2012. The genomic basis of adaptive evolution in threespine sticklebacks. Nature, 484(7392), pp.55-61.

Kakioka, R., Mori, S., Kokita, T., Hosoki, T.K., Nagano, A.J., Ishikawa, A., Kume, M., Toyoda, A. and Kitano, J., 2020. Multiple waves of freshwater colonization of the three-spined stickleback in the Japanese Archipelago. BMC evolutionary biology, 20(1), pp.1-14.

Kawauchi, H., Kawazoe, I., Tsubokawa, M., Kishida, M. and Baker, B.I., 1983. Characterization of melanin-concentrating hormone in chum salmon pituitaries. Nature, 305(5932), pp.321-323.

Kelsh, R.N., Brand, M., Jiang, Y.J., Heisenberg, C.P., Lin, S., Haffter, P., Odenthal, J., Mullins, M.C., Van Eeden, F.J., Furutani-Seiki, M. and Granato, M., 1996. Zebrafish pigmentation mutations and the processes of neural crest development. Development, 123(1), pp.369-389.

Kelsh, R.N., Camargo Sosa, K., Farjami, S., Makeev, V., Dawes, J.H. and Rocco, A., 2021. Cyclical fate restriction: a new view of neural crest cell fate specification. Development, 148(22), p.dev176057.

Kimura, T., 2021. Pigments in Teleosts and their Biosynthesis. In Pigments, Pigment Cells and Pigment Patterns (pp. 127-148). Springer, Singapore.

Kitano J, Lema SC, Luckenbach JA, Mori S, Kawagishi Y, Kusakabe M, Swanson P, Peichel CL., 2010. Adaptive divergence in the thyroid hormone signaling pathway in the stickleback radiation. Current Biology. Dec 7;20(23):2124-30.

Kofler, R., Orozco-terWengel, P., De Maio, N., Pandey, R.V., Nolte, V., Futschik, A., Kosiol, C. and Schlötterer, C., 2011. PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. PloS one, 6(1), p.e15925.

Kofler, R., Pandey, R.V. and Schlötterer, C., 2011. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics, 27(24), pp.3435-3436.

Kolaczkowski, B., Kern, A.D., Holloway, A.K. and Begun, D.J., 2011. Genomic differentiation between temperate and tropical Australian populations of Drosophila melanogaster. Genetics, 187(1), pp.245-260.

Kolenda, K., Najbar, B., Najbar, A., Kaczmarek, P., Kaczmarski, M. and Skawiński, T., 2017. Rare colour aberrations and anomalies of amphibians and reptiles recorded in Poland. Herpetology Notes, 10, pp.103-109.

Kondo, S., Watanabe, M. and Miyazawa, S., 2021. Studies of Turing pattern formation in zebrafish skin. Philosophical Transactions of the Royal Society A, 379(2213), p.20200274.

Kristjánsson BK, Skúlason S, Noakes DL., 2002. Rapid divergence in a recently isolated population of threespine stickleback (Gasterosteus aculeatus L.). Evolutionary Ecology Research.;4(5):659-72.

Kuiper, R.P., Schepens, M., Thijssen, J., Schoenmakers, E.F., and van Kessel, A.G., 2004. Regulation of the MiTF/TFE bHLH-LZ transcription factors through restricted spatial expression and alternative splicing of functional domains. Nucleic acids research 32, 2315-2322.

Kunieda, T., Nakagiri, M., Takami, M., Ide, H. and Ogawa, H., 1999. Cloning of bovine LYST gene and identification of a missense mutation associated with Chediak-Higashi syndrome of cattle. Mammalian Genome, 10(12), pp.1146-1149.

Kuriyama, T. and Hasegawa, M., 2017. Embryonic developmental process governing the conspicuousness of body stripes and blue tail coloration in the lizard Plestiodon latiscutatus. Evolution & development, 19(1), pp.29-39.

Kuriyama, T., Misawa, H., Miyaji, K., Sugimoto, M. and Hasegawa, M., 2013. Pigment cell mechanisms underlying dorsal color-pattern polymorphism in the Japanese four-lined snake. Journal of morphology, 274(12), pp.1353-1364.

Kuriyama, T., Murakami, A., Brandley, M. and Hasegawa, M., 2020. Blue, Black, and Stripes: Evolution and Development of Color Production and Pattern Formation in Lizards and Snakes.. Front. Ecol. Evol. 8:232.doi: 10.3389/fevo.2020.00232

Kusakabe M, Ishikawa A, Ravinet M, Yoshida K, Makino T, Toyoda A, Fujiyama A, Kitano J., 2017 Genetic basis for variation in salinity tolerance between stickleback ecotypes. Molecular ecology. Jan;26(1):304-19.

Kvon, E.Z., Kamneva, O.K., Melo, U.S., Barozzi, I., Osterwalder, M., Mannion, B.J., Tissières, V., Pickle, C.S., Plajzer-Frick, I., Lee, E.A. and Kato, M., 2016. Progressive loss of function in a limb enhancer during snake evolution. Cell, 167(3), pp.633-642.

Kõressaar, T. and Remm, M., 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics, 23(10), pp.1289-1291.

Kõressaar, T., Lepamets, M., Kaplinski, L., Raime, K., Andreson, R. and Remm, M., 2018. Primer3_masker: integrating masking of template sequence with primer design software. Bioinformatics, 34(11), pp.1937-1938.

Kühn, C. and Weikard, R., 2007. An investigation into the genetic background of coat colour dilution in a Charolais× German Holstein F2 resource population. Animal genetics, 38(2), pp.109-113.

Kültz D. The combinatorial nature of osmosensing in fishes. Physiology. 2012 Aug;27(4):259-75.

Lam, T.J. and Hoar, W.S., 1967. Seasonal effects of prolactin on freshwater osmoregulation of the marine form (trachurus) of the stickleback Gasterosteus aculeatus. Canadian Journal of Zoology, 45(4), pp.509-516.

Lambert, H., Carder, G. and D'Cruze, N., 2019. Given the Cold Shoulder: A review of the scientific literature for evidence of reptile sentience. Animals, 9(10), p.821.

Langmead, B. and Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nature methods, 9(4), p.357.

Lao, O., De Gruijter, J.M., van Duijn, K., Navarro, A. and Kayser, M., 2007. Signatures of positive selection in genes associated with human skin pigmentation as revealed from analyses of single nucleotide polymorphisms. Annals of human genetics, 71(3), pp.354-369.

Le Rouzic, A., Østbye, K., Klepaker, T.O., Hansen, T.F., Bernatchez, L., Schluter, D. and Vøllestad, L.A., 2011. Strong and consistent natural selection associated with armour reduction in sticklebacks. Molecular ecology, 20(12), pp.2483-2493.

Le, L., Escobar, I.E., Ho, T., Lefkovith, A.J., Latteri, E., Haltaufderhyde, K.D., Dennis, M.K., Plowright, L., Sviderskaya, E.V., Bennett, D.C. and Oancea, E., 2020. SLC45A2 protein stability and regulation of melanosome pH determine melanocyte pigmentation. Molecular biology of the cell, 31(24), pp.2687-2702.

Leal, F. and Cohn, M.J., 2016. Loss and re-emergence of legs in snakes by modular evolution of sonic hedgehog and HOXD enhancers. Current Biology, 26(21), pp.2966-2973.

Lee CE, Kiergaard M, Gelembiuk GW, Eads BD, Posavi M. 2011 Pumping ions: rapid parallel evolution of ionic regulation following habitat invasions. Evolution: International Journal of Organic Evolution. Aug;65(8):2229-44.

Lescak, E.A., Bassham, S.L., Catchen, J., Gelmond, O., Sherbick, M.L., von Hippel, F.A. and Cresko, W.A., 2015. Evolution of stickleback in 50 years on earthquake-uplifted islands. Proceedings of the National Academy of Sciences, 112(52), pp.E7204-E7212.

Levy, C., Khaled, M., and Fisher, D.E. 2006. MITF: master regulator of melanocyte development and melanoma oncogene. Trends in molecular medicine 12, 406-414.

Lewis, A.C., Rankin, K.J., Pask, A.J., and Stuart-Fox, D. 2017. Stress-induced changes in color expression mediated by iridophores in a polymorphic lizard. Ecology and Evolution 7, 8262-8272.

Lewis, V.M., Saunders, L.M., Larson, T.A., Bain, E.J., Sturiale, S.L., Gur, D., Chowdhury, S., Flynn, J.D., Allen, M.C., Deheyn, D.D. and Lee, J.C., 2019. Fate plasticity and reprogramming in genetically distinct populations of Danio leucophores. Proceedings of the National Academy of Sciences, 116(24), pp.11806-11811.

Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. bioinformatics. 2009 Jul 15;25(14):1754-60.

Li, H. and Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. bioinformatics, 25(14), pp.1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R., 2009. The sequence alignment/map format and SAMtools. Bioinformatics, 25(16), pp.2078-2079.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078-2079.

Liem KF, Bemis WE, Walker WF, Grande L. Functional anatomy of the vertebrates: an evolutionary perspective.

Lim MC, Witt CC, Graham CH, Dávalos LM. Parallel molecular evolution in pathways, genes, and sites in high-elevation hummingbirds revealed by comparative transcriptomics. Genome biology and evolution. 2019 May 22.

Lin, J.Y. and Fisher, D.E., 2007. Melanocyte biology and skin pigmentation. Nature, 445(7130), pp.843-850.

Linnen, C.R., Kingsley, E.P., Jensen, J.D. and Hoekstra, H.E., 2009. On the origin and spread of an adaptive allele in deer mice. Science, 325(5944), pp.1095-1098.

Lister, J.A., Lane, B.M., Nguyen, A., and Lunney, K. 2011. Embryonic expression of zebrafish MiT family genes tfe3b, tfeb, and tfec. Developmental Dynamics 240, 2529-2538.

Lister, J.A., Robertson, C.P., Lepage, T., Johnson, S.L., and Raible, D.W. 1999. Nacre encodes a zebrafish microphthalmia-related protein that regulates neural-crest-derived pigment cell fate. Development 126, 3757-3767.

Liu, N.A., Liu, Q., Wawrowsky, K., Yang, Z., Lin, S. and Melmed, S., 2006. Prolactin receptor signaling mediates the osmotic response of embryonic zebrafish lactotrophs. Molecular Endocrinology, 20(4), pp.871-880.

Liu, S., Hansen, M.M. and Jacobsen, M.W., 2016. Region-wide and ecotype-specific differences in demographic histories of threespine stickleback populations, estimated from whole genome sequences. Molecular Ecology, 25(20), pp.5187-5202.

Liu, S., Hansen, M.M. and Jacobsen, M.W., 2016. Region-wide and ecotype-specific differences in demographic histories of threespine stickleback populations, estimated from whole genome sequences. Molecular Ecology, 25(20), pp.5187-5202.

Logan, D.W., Burn, S.F., and Jackson, I.J. (2006). Regulation of pigmentation in zebrafish melanophores. Pigment cell research 19, 206-213.

Lopes, S.S., Yang, X., Müller, J., Carney, T.J., McAdow, A.R., Rauch, G.J., Jacoby, A.S., Hurst, L.D., Delfino-Machín, M., Haffter, P. and Geisler, R., 2008. Leukocyte tyrosine kinase functions in pigment cell development. PLoS genetics, 4(3), p.e1000026.

Luo, M., Lu, G., Yin, H., Wang, L., Atuganile, M. and Dong, Z., 2021. Fish pigmentation and coloration: Molecular mechanisms and aquaculture perspectives. Reviews in Aquaculture.

Lyerla, T.A. and Dalton, H.C., 1971. Genetic and developmental characteristics of a new color variant, axanthic, in the Mexican axolotl, Ambystoma mexicanum Shaw. Developmental Biology, 24(1), pp.1-18.

Lyon, M.F., 1962. Sex chromatin and gene action in the mammalian X-chromosome. American journal of human genetics, 14(2), p.135.

Mack, K.L., Square, T.A., Zhao, B., Miller, C.T. and Fraser, H.B., 2023. Evolution of spatial and temporal cis-regulatory divergence in sticklebacks. Molecular Biology and Evolution, 40(3), p.msad034.

Maloney, S.K., Fuller, A. and Mitchell, D., 2009. Climate change: is the dark Soay sheep endangered?. Biology Letters, 5(6), pp.826-829.

Manceau, M., Domingues, V.S., Linnen, C.R., Rosenblum, E.B. and Hoekstra, H.E., 2010. Convergence in pigmentation at multiple levels: mutations, genes and function.

Philosophical Transactions of the Royal Society B: Biological Sciences, 365(1552), pp.2439-2450.

Mancera, J.M. and McCormick, S.D., 2019. Role of prolactin, growth hormone, insulin-like growth factor I and cortisol in teleost osmoregulation. In Fish osmoregulation (pp. 497-515).

Mandal, B.K., Chen, H., Si, Z., Hou, X., Yang, H., Xu, X., Wang, J. and Wang, C., 2020. Shrunk and scattered black spots turn out due to MC1R knockout in a white-black Oujiang color common carp (Cyprinus carpio var. color). Aquaculture, 518, p.734822.

Marchinko KB 2009. Predation's role in repeated phenotypic and genetic divergence of armor in threespine stickleback. Evolution: International Journal of Organic Evolution. Jan;63(1):127-38.

Marchinko, K.B. and Schluter, D., 2007. Parallel evolution by correlated response: lateral plate reduction in threespine stickleback. Evolution: International Journal of Organic Evolution, 61(5), pp.1084-1090.

Marques, D.A., Jones, F.C., Di Palma, F., Kingsley, D.M. and Reimchen, T.E., 2018. Experimental evidence for rapid genomic adaptation to a new niche in an adaptive radiation. Nature ecology & evolution, 2(7), pp.1128-1138.

Martina, J.A., and Puertollano, R. (2017). TFEB and TFE3: the art of multi-tasking under stress conditions. Transcription 8, 48-54.

Masada, M., Matsumoto, J. and Akino, M., 1990. Biosynthetic pathways of pteridines and their association with phenotypic expression in vitro in normal and neoplastic pigment cells from goldfish. Pigment Cell Research, 3(2), pp.61-70.

Masselink, W. and Tanaka, E.M., 2021. Toward whole tissue imaging of axolotl regeneration. Developmental Dynamics, 250(6), pp.800-806.

Mateus AP, Power DM, Canário AV. Stress and disease in fish. In Fish Diseases. 2017 Jan 1 (pp. 187-220). Academic Press.

Matschiner, M., Hanel, R. and Salzburger, W., 2011. On the origin and trigger of the notothenioid adaptive radiation. PLoS one, 6(4), p.e18911.

Matsuda, N., Kasagi, S., Nakamaru, T., Masuda, R., Takahashi, A. and Tagawa, M., 2018. Left-right pigmentation pattern of Japanese flounder corresponds to expression levels of melanocortin receptors (MC1R and MC5R), but not to agouti signaling protein 1 (ASIP1) expression. General and Comparative Endocrinology, 262, pp.90-98.

Matsumoto, H., Kojya, M., Takamuku, H., Kimura, S., Kashimura, A., Imai, S., Yamauchi, K. and Ito, S., 2020. MC1R c. 310G>-and c. 871G> A determine the coat color of Kumamoto sub-breed of Japanese Brown cattle. Animal Science Journal, 91(1), p.e13367.

McCairns, R.S. and 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: International Journal of Organic Evolution, 64(4), pp.1029-1047.

McCracken KG, Barger CP, Bulgarella M, Johnson KP, Sonsthagen SA, Trucco J, Valqui TH, Wilson RE, Winker K, Sorenson MD. Parallel evolution in the major haemoglobin genes of eight species of Andean waterfowl. Molecular ecology. 2009 Oct;18(19):3992-4005.

McGraw, K.J., 2006. Mechanics of uncommon colors: pterins, porphyrins, and psittacofulvins. Bird coloration, 1, pp.354-398.

McKinnon, J.S. and Rundle, H.D., 2002. Speciation in nature: the threespine stickleback model systems. Trends in ecology & evolution, 17(10), pp.480-488.

McLean, C.A., Lutz, A., Rankin, K.J., Elliott, A., Moussalli, A. and Stuart-Fox, D., 2019. Red carotenoids and associated gene expression explain colour variation in frillneck lizards. Proceedings of the Royal Society B, 286(1907), p.20191172.

McMenamin, S.K., Bain, E.J., McCann, A.E., Patterson, L.B., Eom, D.S., Waller, Z.P., Hamill, J.C., Kuhlman, J.A., Eisen, J.S. and Parichy, D.M., 2014. Thyroid hormone–dependent adult pigment cell lineage and pattern in zebrafish. Science, 345(6202), pp.1358-1361.

McNamara, M.E., Rossi, V., Slater, T.S., Rogers, C.S., Ducrest, A.L., Dubey, S. and Roulin, A., 2021. Decoding the evolution of melanin in vertebrates. Trends in Ecology & Evolution.

McNamara, M.E., Rossi, V., Slater, T.S., Rogers, C.S., Ducrest, A.L., Dubey, S., and Roulin, A. (2021). Decoding the Evolution of Melanin in Vertebrates. Trends Ecol Evol 36, 430-443. 10.1016/j.tree.2020.12.012.

McPhail, J.D., 1969. Predation and the evolution of a stickleback (Gasterosteus). Journal of the Fisheries Board of Canada, 26(12), pp.3183-3208.

Megía-Palma, R., Jorge, A. and Reguera, S., 2018. Raman spectroscopy reveals the presence of both eumelanin and pheomelanin in the skin of lacertids. Journal of Herpetology, 52(1), pp.67-73.

Mehravar, M., Shirazi, A., Nazari, M., and Banan, M. (2019). Mosaicism in CRISPR/Cas9-mediated genome editing. Developmental biology 445, 156-162.

Miller SE, Roesti M, Schluter D. A Single Interacting Species Leads to Widespread Parallel Evolution of the Stickleback Genome. Current Biology. 2019 Jan 24.

Miller, M.W., Duhl, D.M., Vrieling, H., Cordes, S.P., Ollmann, M.M., Winkes, B.M. and Barsh, G.S., 1993. Cloning of the mouse agouti gene predicts a secreted protein ubiquitously expressed in mice carrying the lethal yellow mutation. Genes & development, 7(3), pp.454-467.

Mills, M.G. and Patterson, L.B., 2009, February. Not just black and white: pigment pattern development and evolution in vertebrates. In Seminars in cell & developmental biology (Vol. 20, No. 1, pp. 72-81). Academic Press.

Minchin, J.E. and Hughes, S.M., 2008. Sequential actions of Pax3 and Pax7 drive xanthophore development in zebrafish neural crest. Developmental biology, 317(2), pp.508-522.

Moodie, G.E.E. and Reimchen, T.E., 1976. Phenetic variation and habitat differences in Gasterosteus populations of the Queen Charlotte Islands. Systematic Zoology, 25(1), pp.49-61.

Moodie, G.E.E., 1972. Morphology, life history, and ecology of an unusual stickleback (Gasterosteus aculeatus) in the Queen Charlotte Islands, Canada. Canadian Journal of Zoology, 50(6), pp.721-732.

Morgan, T.H., 1911. The origin of five mutations in eye color in Drosophila and their modes of inheritance. Science, 33(849), pp.534-537.

Morris, M.R., Richard, R., Leder, E.H., Barrett, R.D., Aubin-Horth, N. and Rogers, S.M., 2014. Gene expression plasticity evolves in response to colonization of freshwater lakes in threespine stickleback. Molecular ecology, 23(13), pp.3226-3240.

Murakami, A., Hasegawa, M. and Kuriyama, T., 2016. Pigment cell mechanism of postembryonic stripe pattern formation in the Japanese four-lined snake. Journal of Morphology, 277(2), pp.196-203.

Mäkinen, H.S. and Merilä, J., 2008. Mitochondrial DNA phylogeography of the three-spined stickleback (Gasterosteus aculeatus) in Europe—evidence for multiple glacial refugia. Molecular Phylogenetics and Evolution, 46(1), pp.167-182.

Mäkinen, H.S., Cano, J.M. and Merilä, J., 2008. Identifying footprints of directional and balancing selection in marine and freshwater three-spined stickleback (Gasterosteus aculeatus) populations. Molecular Ecology, 17(15), pp.3565-3582.

Münzing, J., 1963. The evolution of variation and distributional patterns in European populations of the three-spined stickleback, Gasterosteus aculeatus. Evolution, pp.320-332.

Nagatsu, T. and Ichinose, H., 1996. GTP cyclohydrolase I gene, tetrahydrobiopterin, and tysorine hydroxylase gene: Their relations to dystonia and parkinsonism. Neurochemical research, 21, pp.245-250.

Nakamasu, A., Takahashi, G., Kanbe, A. and Kondo, S., 2009. Interactions between zebrafish pigment cells responsible for the generation of Turing patterns. Proceedings of the National Academy of Sciences, 106(21), pp.8429-8434.

Nam, I.S., Oh, M.G., Nam, M.S. and Kim, W.S., 2021. Specific mutations in the genes of MC1R and TYR have an important influence on the determination of pheomelanin pigmentation in Korean native chickens. Journal of Advanced Veterinary and Animal Research, 8(2), p.266.

Naot D, Musson DS, Cornish J. The activity of peptides of the calcitonin family in bone. Physiological reviews. 2019 Jan 1;99(1):781-805.

Neuffer, S.J., and Cooper, C.D. (2022). Zebrafish Syndromic Albinism Models as Tools for Understanding and Treating Pigment Cell Disease in Humans. Cancers 14, 1752.

Nicolaï, M.P., D'Alba, L., Goldenberg, J., Gansemans, Y., Van Nieuwerburgh, F., Clusella-Trullas, S. and Shawkey, M.D., 2021. Untangling the structural and molecular mechanisms underlying colour and rapid colour change in a lizard, Agama atra. Molecular Ecology, 30(10), pp.2262-2284.

Nicolaï, M.P., D'Alba, L., Goldenberg, J., Gansemans, Y., Van Nieuwerburgh, F., Clusella-Trullas, S., and Shawkey, M.D. (2021). Untangling the structural and molecular mechanisms underlying colour and rapid colour change in a lizard, Agama atra. Molecular Ecology 30, 2262-2284.

Nikaido, M., Subkhankulova, T., Uroshlev, L.A., Kasianov, A.J., Sosa, K.C., Bavister, G., Yang, X., Rodrigues, F.S., Carney, T.J., Schwetlick, H. and Dawes, J.H., 2021. Zebrafish pigment cells develop directly from persistent highly multipotent progenitors. bioRxiv, pp.2021-06.

Nilsen TO, Ebbesson LO, Madsen SS, McCormick SD, Andersson E, Björnsson BT, Prunet P, Stefansson SO. Differential expression of gill Na+, K+-ATPaseα-and β-subunits, Na+, K+, 2Cl-cotransporter and CFTR anion channel in juvenile anadromous and landlocked Atlantic salmon Salmo salar. Journal of Experimental Biology. 2007 Aug 15;210(16):2885-96.

Nord, H., Dennhag, N., Muck, J. and von Hofsten, J., 2016. Pax7 is required for establishment of the xanthophore lineage in zebrafish embryos. Molecular biology of the cell, 27(11), pp.1853-1862.

Norton, H.L., Kittles, R.A., Parra, E., McKeigue, P., Mao, X., Cheng, K., Canfield, V.A., Bradley, D.G., McEvoy, B. and Shriver, M.D., 2007. Genetic evidence for the convergent evolution of light skin in Europeans and East Asians. Molecular biology and evolution, 24(3), pp.710-722.

Nosil P, Villoutreix R, de Carvalho CF, Farkas TE, Soria-Carrasco V, Feder JL, Crespi BJ, Gompert Z. Natural selection and the predictability of evolution in Timema stick insects. Science. 2018 Feb 16;359(6377):765-70.

Nunes, V.L., Miraldo, A., Beaumont, M.A., Butlin, R.K. and Paulo, O.S., 2011. Association of Mc1r variants with ecologically relevant phenotypes in the European ocellated lizard, Lacerta lepida. Journal of Evolutionary Biology, 24(10), pp.2289-2298.

O'Brown NM, Summers BR, Jones FC, Brady SD, Kingsley DM. A recurrent regulatory change underlying altered expression and Wnt response of the stickleback armor plates gene EDA. Elife. 2015 Jan 28;4:e05290.

Odenthal, J., Rossnagel, K., Haffter, P., Kelsh, R.N., Vogelsang, E., Brand, M., Van Eeden, F.J., Furutani-Seiki, M., Granato, M., Hammerschmidt, M. and Heisenberg, C.P., 1996. Mutations affecting xanthophore pigmentation in the zebrafish, Danio rerio. Development, 123(1), pp.391-398.

Olsson, M., Stuart-Fox, D. and Ballen, C., 2013, June. Genetics and evolution of colour patterns in reptiles. In Seminars in cell & developmental biology (Vol. 24, No. 6-7, pp. 529-541). Academic Press.

Orti, G., Bell, M.A., Reimchen, T.E. and Meyer, A., 1994. Global survey of mitochondrial DNA sequences in the threespine stickleback: evidence for recent migrations. Evolution, 48(3), pp.608-622.

Owen, J., Yates, C. and Kelsh, R.N., 2021. Pigment patterning in teleosts. Pigments, pigment cells and pigment patterns, pp.247-292.

Paccard A, Wasserman BA, Hanson D, Astorg L, Durston D, Kurland S, Apgar TM, El-Sabaawi RW, Palkovacs EP, Hendry AP, Barrett RD. Adaptation in temporally variable environments: stickleback armor in periodically breaching bar-built estuaries. Journal of evolutionary biology. 2018 May;31(5):735-52.

Parichy, D.M., Rawls, J.F., Pratt, S.J., Whitfield, T.T. and Johnson, S.L., 1999. Zebrafish sparse corresponds to an orthologue of c-kit and is required for the morphogenesis of a

subpopulation of melanocytes, but is not essential for hematopoiesis or primordial germ cell development. Development, 126(15), pp.3425-3436.

Passeron, T., Valencia, J.C., Bertolotto, C., Hoashi, T., Le Pape, E., Takahashi, K., Ballotti, R. and Hearing, V.J., 2007. SOX9 is a key player in ultraviolet B-induced melanocyte differentiation and pigmentation. Proceedings of the National Academy of Sciences, 104(35), pp.13984-13989.

Patterson, L.B. and Parichy, D.M., 2013. Interactions with iridophores and the tissue environment required for patterning melanophores and xanthophores during zebrafish adult pigment stripe formation. PLoS genetics, 9(5), p.e1003561.

Pavan, W.J. and Sturm, R.A., 2019. The genetics of human skin and hair pigmentation. Annual review of genomics and human genetics, 20, pp.41-72.

Pavlova, N.S., Neretina, T.V. and Smirnova, O.V., 2020. Dynamics of prolactin axis genes in the brain of male and female three-spined stickleback Gasterosteus aculeatus (Gasterostaidae) during short-term freshwater adaptation. Journal of Ichthyology, 60, pp.299-304.

Penczak, T., 1965. Morphological variation of the stickleback (Gasterosteus aculeatus L.) in Poland. Zoologica Poloniae, 15, pp.3-49.

Peterson BZ, DeMaria CD, Yue DT. Calmodulin is the Ca2+ sensor for Ca2+-dependent inactivation of L-type calcium channels. Neuron. 1999 Mar 1;22(3):549-58.

Petratou, K., Spencer, S.A., Kelsh, R.N. and Lister, J.A., 2021. The MITF paralog tfec is required in neural crest development for fate specification of the iridophore lineage from a multipotent pigment cell progenitor. PLoS One, 16(1), p.e0244794.

Phelps, G.B., Hagen, H.R., Amsterdam, A., and Lees, J.A. 2022. MITF deficiency accelerates GNAQ-driven uveal melanoma. Proceedings of the National Academy of Sciences 119, e2107006119.

Piegorsch, W.W., 1990. Fisher's contributions to genetics and heredity, with special emphasis on the Gregor Mendel controversy. Biometrics, pp.915-924.

Pizzatto, L. and Dubey, S., 2012. Colour-polymorphic snake species are older. Biological Journal of the Linnean Society, 107(1), pp.210-218.Irizarry, K.J. and Bryden, R.L., 2016. In silico analysis of gene expression network components underlying pigmentation phenotypes in the Python identified evolutionarily conserved clusters of transcription factor binding sites. Advances in bioinformatics, 2016.

Poulton, E.B., 1890. The colours of animals: their meaning and use, especially considered in the case of insects (Vol. 67). D. Appleton.

Quigley, I.K. and Parichy, D.M., 2002. Pigment pattern formation in zebrafish: a model for developmental genetics and the evolution of form. Microscopy research and technique, 58(6), pp.442-455.

Raeymaekers, J.A., Konijnendijk, N., Larmuseau, M.H., Hellemans, B., De Meester, L. and Volckaert, F.A., 2014. A gene with major phenotypic effects as a target for selection vs. homogenizing gene flow. Molecular Ecology, 23(1), pp.162-181.

Rafiński, J., Bańgbura, J. and Przybylski, M., 1989. Genetic differentiation of freshwater and marine sticklebacks,(Gasterosteus aculeatus) of Eastern Europe. Journal of Zoological Systematics and Evolutionary Research, 27(1), pp.33-43.

Rahnenfuhrer AAJ. topGO: Enrichment analysis for Gene Ontology 2018. R package version 2.34.

Rasys, A.M., Park, S., Ball, R.E., Alcala, A.J., Lauderdale, J.D., and Menke, D.B. (2019). CRISPR-Cas9 gene editing in lizards through microinjection of unfertilized oocytes. Cell reports 28, 2288-2292. e2283.

Recio, P., Rodríguez-Ruiz, G., López, P. and Martín, J., 2022. Size-related changes and chemical basis of melanin-based body coloration in the amphisbaenian Trogonophis wiegmanni. Amphibia-Reptilia, 1(aop), pp.1-9.

References

Reid, J.B. and Ross, J.J., 2011. Mendel's genes: toward a full molecular characterization. Genetics, 189(1), pp.3-10.

Reid, K., Bell, M.A. and Veeramah, K.R., 2021. Threespine stickleback: A model system for evolutionary genomics. Annual review of genomics and human genetics, 22, pp.357-383.

Reimchen TE, Bergstrom C, Nosil P. Natural selection and the adaptive radiation of Haida Gwaii stickleback. Evol. Ecol. Res. 2013 Mar 1;15:241-69.

Reimchen TE. Predator handling failures of lateral plate morphs in Gasterosteus aculeatus: functional implications for the ancestral plate condition. Behaviour. 2000 Jul 1;137(7):1081-96.

Reimchen TE. Structural relationships between spines and lateral plates in threespine stickleback (Gasterosteus aculeatus). Evolution. 1983 Sep;37(5):931-46.

Reimchen, T.E., 1980. Spine deficiency and polymorphism in a population of Gasterosteus aculeatus: an adaptation to predators?. Canadian Journal of Zoology, 58(7), pp.1232-1244.

Reimchen, T.E., 1983. Structural relationships between spines and lateral plates in threespine stickleback (Gasterosteus aculeatus). Evolution, pp.931-946.

Reynolds, E.S. (1963). The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. The Journal of cell biology 17, 208.

Rich A, Keller EA. A hydrologic and geomorphic model of estuary breaching and closure. Geomorphology. 2013 Jun 1;191:64-74.

Riddle MR, Aspiras AC, Gaudenz K, Peuß R, Sung JY, Martineau B, Peavey M, Box AC, Tabin JA, McGaugh S, Borowsky R. Insulin resistance in cavefish as an adaptation to a nutrient-limited environment. Nature. 2018 Mar;555(7698):647-51.

Roberts Kingman, G.A., Vyas, D.N., Jones, F.C., Brady, S.D., Chen, H.I., Reid, K., Milhaven, M., Bertino, T.S., Aguirre, W.E., Heins, D.C. and Von Hippel, F.A., 2021. Predicting future from past: The genomic basis of recurrent and rapid stickleback evolution. Science Advances, 7(25), p.eabg5285.

Rodríguez, A., Mundy, N.I., Ibáñez, R. and Pröhl, H., 2020. Being red, blue and green: the genetic basis of coloration differences in the strawberry poison frog (Oophaga pumilio). BMC genomics, 21(1), pp.1-16.

Romanov, G.A. and Sukhoverov, V.S., 2017. Arginine CGA codons as a source of nonsense mutations: a possible role in multivariant gene expression, control of mRNA quality, and aging. Molecular Genetics and Genomics, 292(5), pp.1013-1026.

Rose, M.P. and Williams, D.L., 2014. Neurological dysfunction in a ball python (Python regius) colour morph and implications for welfare. Journal of Exotic Pet Medicine, 23(3), pp.234-239.

Rosenblum, E.B., Hoekstra, H.E. and Nachman, M.W., 2004. Adaptive reptile color variation and the evolution of the MCIR gene. Evolution, 58(8), pp.1794-1808.

Rosenblum, E.B., Römpler, H., Schöneberg, T. and Hoekstra, H.E., 2010. Molecular and functional basis of phenotypic convergence in white lizards at White Sands. Proceedings of the National Academy of Sciences, 107(5), pp.2113-2117.

Roulin, A., Mafli, A. and Wakamatsu, K., 2013. Reptiles produce pheomelanin: evidence in the eastern Hermann's tortoise (Eurotestudo boettgeri). Journal of Herpetology, 47(2), pp.258-261.

Rowan, S., Chen, C.-M.A., Young, T.L., Fisher, D.E., and Cepko, C.L. (2004). Transdifferentiation of the retina into pigmented cells in ocular retardation mice defines a new function of the homeodomain gene Chx10.

Rowe, C. and Halpin, C., 2013. Why are warning displays multimodal?. Behavioral Ecology and Sociobiology, 67, pp.1425-1439.

Rudman SM, Goos JM, Burant JB, Brix KV, Gibbons TC, Brauner CJ, Jeyasingh PD. Ionome and elemental transport kinetics shaped by parallel evolution in threespine stickleback. Ecology letters. 2019 Apr;22(4):645-53.

Rudman, S.M., Goos, J.M., Burant, J.B., Brix, K.V., Gibbons, T.C., Brauner, C.J. and Jeyasingh, P.D., 2019. Ionome and elemental transport kinetics shaped by parallel evolution in threespine stickleback. Ecology letters, 22(4), pp.645-653.

Russell, E.S., 1985. A history of mouse genetics. Annual review of genetics, 19(1), pp.1-29.

Saenko, S.V., Lamichhaney, S., Barrio, A.M., Rafati, N., Andersson, L. and Milinkovitch, M.C., 2015. Amelanism in the corn snake is associated with the insertion of an LTR-retrotransposon in the OCA2 gene. Scientific reports, 5(1), pp.1-9.

Schlötterer, C., 2023. How predictable is adaptation from standing genetic variation? Experimental evolution in Drosophila highlights the central role of redundancy and linkage disequilibrium. Philosophical Transactions of the Royal Society B, 378(1877), p.20220046.

Schluckebier, R., Sachs, M. and Vences, M., 2022. Axanthic green toads, Bufotes viridis (Anura: Bufonidae), from Cologne, Germany. Herpetology Notes, 15, pp.345-348.

Schluter D, Marchinko KB, Barrett RD, Rogers SM. Natural selection and the genetics of adaptation in threespine stickleback. Philosophical Transactions of the Royal Society B: Biological Sciences. 2010 Aug 27;365(1552):2479-86.

Schluter, D. and Conte, G.L., 2009. Genetics and ecological speciation. Proceedings of the National Academy of Sciences, 106(Supplement 1), pp.9955-9962.

Schluter, D., Marchinko, K.B., Barrett, R.D.H. and Rogers, S.M., 2010. Natural selection and the genetics of adaptation in threespine stickleback. Philosophical Transactions of the Royal Society B: Biological Sciences, 365(1552), pp.2479-2486.

Schott, R.K., Müller, J., Yang, C.G., Bhattacharyya, N., Chan, N., Xu, M., Morrow, J.M., Ghenu, A.H., Loew, E.R., Tropepe, V. and Chang, B.S., 2016. Evolutionary transformation of rod photoreceptors in the all-cone retina of a diurnal garter snake. Proceedings of the National Academy of Sciences, 113(2), pp.356-361.

Scott, R.J., Haines, G.E. and Trask, C.A., 2023. Armour reduction and pelvic girdle loss in a population of threespine stickleback (Gasterosteus aculeatus) from western Newfoundland, Canada. Environmental Biology of Fishes, pp.1-11.

Sedlazeck, F.J., Rescheneder, P., and Von Haeseler, A. (2013). NextGenMap: fast and accurate read mapping in highly polymorphic genomes. Bioinformatics 29, 2790-2791.

Seruggia, D., Josa, S., Fernández, A., and Montoliu, L. (2021). The structure and function of the mouse tyrosinase locus. Pigment Cell & Melanoma Research 34, 212-221.

Settembre, C., Di Malta, C., Polito, V.A., Arencibia, M.G., Vetrini, F., Erdin, S., Erdin, S.U., Huynh, T., Medina, D., and Colella, P. (2011). TFEB links autophagy to lysosomal biogenesis. science 332, 1429-1433.

Shanfelter, A.F., Archambeault, S.L. and White, M.A., 2019. Divergent fine-scale recombination landscapes between a freshwater and marine population of threespine stickleback fish. Genome biology and evolution, 11(6), p.1552.

Shapiro MD, Bell MA, Kingsley DM. Parallel genetic origins of pelvic reduction in vertebrates. Proceedings of the National Academy of Sciences. 2006 Sep 12;103(37):13753-8.

Shapiro MD, Marks ME, Peichel CL, Blackman BK, Nereng KS, Jónsson B, Schluter D, Kingsley DM. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature. 2004 Apr;428(6984):717-23.

Silvers, W.K., 2012. The coat colors of mice: a model for mammalian gene action and interaction. Springer Science & Business Media.

Simon JD, Peles D, Wakamatsu K et al (2009) Current challenges in understanding melanogenesis: bridging chemistry, biological control, morphology, and function. Pigment Cell Melanoma Res 22:563–579

Simões-Costa, M. and Bronner, M.E., 2015. Establishing neural crest identity: a gene regulatory recipe. Development, 142(2), pp.242-257.

Singh, A.P. and Nüsslein-Volhard, C., 2015. Zebrafish stripes as a model for vertebrate colour pattern formation. Current Biology, 25(2), pp.R81-R92.

Singh, A.P., Schach, U. and Nüsslein-Volhard, C., 2014. Proliferation, dispersal and patterned aggregation of iridophores in the skin prefigure striped colouration of zebrafish. Nature cell biology, 16(6), pp.604-611.

Slade, L., and Pulinilkunnil, T. (2017). The MiTF/TFE family of transcription factors: master regulators of organelle signaling, metabolism, and stress adaptation. Molecular Cancer Research 15, 1637-1643.

Smedley D, Haider S, Ballester B, Holland R, London D, Thorisson G, Kasprzyk A. BioMart–biological queries made easy. BMC genomics. 2009 Dec;10(1):22.

Smith, M. and Bannerman, J. 2016. The Entomological Society of Manitoba.

Smit, A.K., Collazo-Roman, M., Vadaparampil, S.T., Valavanis, S., Del Rio, J., Soto, B., Flores, I., Dutil, J. and Kanetsky, P.A., 2020. MC1R variants and associations with pigmentation characteristics and genetic ancestry in a Hispanic, predominately Puerto Rican, population. Scientific Reports, 10(1), p.7303.

Smith C, Zięba G, Spence R, Klepaker T, Przybylski M. Three-spined stickleback armour predicted by body size, minimum winter temperature and pH. Journal of Zoology. 2020 Jan 8.

Smith, D.R., Spaulding, D.T., Glenn, H.M. and Fuller, B.B., 2004. The relationship between Na+/H+ exchanger expression and tyrosinase activity in human melanocytes. Experimental cell research, 298(2), pp.521-534.

Soria-Carrasco V, Gompert Z, Comeault AA, Farkas TE, Parchman TL, Johnston JS, Buerkle CA, Feder JL, Bast J, Schwander T, Egan SP. Stick insect genomes reveal natural selection's role in parallel speciation. Science. 2014 May 16;344(6185):738-42.

Spence R, Wootton RJ, Przybylski M, Zięba G, Macdonald K, Smith C. Calcium and salinity as selective factors in plate morph evolution of the three-spined stickleback (G asterosteus aculeatus). Journal of evolutionary biology. 2012 Oct;25(10):1965-74.

Steingrímsson, E., Copeland, N.G. and Jenkins, N.A., 2004. Melanocytes and the microphthalmia transcription factor network. Annu. Rev. Genet., 38, pp.365-411.

Steingrímsson, E., Copeland, N.G. and Jenkins, N.A., 2006. Mouse coat color mutations: from fancy mice to functional genomics. Developmental Dynamics: An Official Publication of the American Association of Anatomists, 235(9), pp.2401-2411.

Steingrímsson, E., Moore, K.J., Lamoreux, M.L., Ferré-D'Amaré, A.R., Burley, S.K., Zimring, D.C.S., Skow, L.C., Hodgkinson, C.A., Arnheiter, H., and Copeland, N.G. (1994). Molecular basis of mouse microphthalmia (mi) mutations helps explain their developmental and phenotypic consequences. Nature genetics 8, 256-263.

Steingrímsson, E., Tessarollo, L., Pathak, B., Hou, L., Arnheiter, H., Copeland, N.G., and Jenkins, N.A. (2002). Mitf and Tfe3, two members of the Mitf-Tfe family of bHLH-Zip transcription factors, have important but functionally redundant roles in osteoclast development. Proceedings of the National Academy of Sciences 99, 4477-4482.

Stoltzfus, A. and Cable, K., 2014. Mendelian-mutationism: the forgotten evolutionary synthesis. Journal of the History of Biology, 47(4), pp.501-546.

Storey JD, Bass AJ, Dabney A, Robinson D (2018) qvalue: Q-value estimation for false discovery rate control. R package version 2.14.0. http://github.com/jdstorey/qvalue

Storey, J.D., 2003. The positive false discovery rate: a Bayesian interpretation and the q-value. The Annals of Statistics, 31(6), pp.2013-2035.

Strickland, K., Räsänen, K., Kristjánsson, B.K., Phillips, J.S., Einarsson, A., Snorradóttir, R.G., Bartrons, M. and Jónsson, Z.O., 2022. Genome-phenotype-environment associations identify signatures of selection in a panmictic population of threespine stickleback. Molecular Ecology.

Sturm, R.A. 2006. A golden age of human pigmentation genetics. Trends Genet 22, 464-468. 10.1016/j.tig.2006.06.010.

Sturm, R.A., 2009. Molecular genetics of human pigmentation diversity. Human molecular genetics, 18(R1), pp.R9-R17.

Sturtevant, A.H., 1913. A third group of linked genes in Drosophila ampelophila. Science, 37(965), pp.990-992.

Subkhankulova, T., Camargo Sosa, K., Uroshlev, L.A., Nikaido, M., Shriever, N., Kasianov, A.S., Yang, X., Rodrigues, F.S., Carney, T.J., Bavister, G. and Schwetlick, H., 2023. Zebrafish pigment cells develop directly from persistent highly multipotent progenitors. Nature Communications, 14(1), p.1258.

Sumner, F.B., 1937. Color and Pigmentation. The Scientific Monthly, 44(4), pp.350-352.

Sussmilch, F.C., Ross, J.J. and Reid, J.B., 2022. Mendel: From genes to genome. Plant Physiology, 190(4), pp.2103-2114.

Szydłowski, P., Madej, J.P., and Mazurkiewicz-Kania, M. (2016). Ultrastructure and distribution of chromatophores in the skin of the leopard gecko (E ublepharis macularius). Acta Zoologica 97, 370-375.

Tamate, H.B., Hirobe, T., Wakamatsu, K., Ito, S., Shibahara, S. and Ishikawa, K., 1989. Levels of tyrosinase and its mRNA in coat-color mutants of C57BL/10J congenic mice: effects of genic substitution at the agouti, brown, albino, dilute, and pink-eyed dilution loci. Journal of Experimental Zoology, 250(3), pp.304-311.

Tanaka, S.A.T.O.S.H.I., Yamamoto, H.I.R.O.A.K.I., Takeuchi, S.A.K.A.E. and Takeuchi, T.A.K.U.J.I., 1990. Melanization in albino mice transformed by introducing cloned mouse tyrosinase gene. Development, 108(2), pp.223-227.

Tang, C.Y., Zhang, X., Xu, X., Sun, S., Peng, C., Song, M.H., Yan, C., Sun, H., Liu, M., Xie, L. and Luo, S.J., 2023. Genetic mapping and molecular mechanism behind color variation in the Asian vine snake. Genome Biology, 24(1), pp.1-21.

Tassabehji, M., Newton, V.E., and Read, A.P. (1994). Waardenburg syndrome type 2 caused by mutations in the human microphthalmia (MITF) gene. Nature genetics 8, 251-255.

Taugbøl, A., Solbakken, M.H., Jakobsen, K.S. and Vøllestad, L.A., 2022. Salinity-induced transcriptome profiles in marine and freshwater threespine stickleback after an abrupt 6-hour exposure. Ecology and Evolution, 12(10), p.e9395.

Terekhanova, N.V., Logacheva, M.D., Penin, A.A., Neretina, T.V., Barmintseva, A.E., Bazykin, G.A., Kondrashov, A.S. and Mugue, N.S., 2014. Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback Gasterosteus aculeatus. PLoS Genetics, 10(10), p.e1004696.

Teyssier, J., Saenko, S.V., Van Der Marel, D. and Milinkovitch, M.C., 2015. Photonic crystals cause active colour change in chameleons. Nature communications, 6(1), p.6368.

Theron, E., Hawkins, K., Bermingham, E., Ricklefs, R.E. and Mundy, N.I., 2001. The molecular basis of an avian plumage polymorphism in the wild: a melanocortin-1-receptor point mutation is perfectly associated with the melanic plumage morph of the bananaquit, Coereba flaveola. Current Biology, 11(8), pp.550-557.

Thomas, A.J. and Erickson, C.A., 2008. The making of a melanocyte: the specification of melanoblasts from the neural crest. Pigment cell & melanoma research, 21(6), pp.598-610.

Thompson, J.N., 1998. Rapid evolution as an ecological process. Trends in ecology & evolution, 13(8), pp.329-332.

Tian Y, Wen H, Qi X, Zhang X, Li Y. Identification of mapk gene family in Lateolabrax maculatus and their expression profiles in response to hypoxia and salinity challenges. Gene. 2019 Feb 5;684:20-9.

Tine, M., De Lorgeril, J., Panfili, J., Diop, K., Bonhomme, F. and Durand, J.D., 2007. Growth hormone and Prolactin-1 gene transcription in natural populations of the black-chinned tilapia Sarotherodon melanotheron acclimatised to different salinities. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology, 147(3), pp.541-549.

Toews, D.P., Hofmeister, N.R. and Taylor, S.A., 2017. The evolution and genetics of carotenoid processing in animals. Trends in Genetics, 33(3), pp.171-182.

Tong C, Li M. Genomic signature of accelerated evolution in a saline-alkaline lake-dwelling Schizothoracine fish. International Journal of Biological Macromolecules. 2020 Apr 15;149:341-7.

Tse WKF, Lai KP, Takei Y. Medaka osmotic stress transcription factor 1b (Ostf1b/TSC22D3-2) triggers hyperosmotic responses of different ion transporters in medaka gill and human embryonic kidney cells via the JNK signalling pathway. The international journal of biochemistry & cell biology. 2011 Dec 1;43(12):1764-75.

Tsetskhladze, Z.R., Canfield, V.A., Ang, K.C., Wentzel, S.M., Reid, K.P., Berg, A.S., Johnson, S.L., Kawakami, K., and Cheng, K.C. (2012). Functional assessment of human coding mutations affecting skin pigmentation using zebrafish.

Tzika, AC. On the role of MITF and TFEC in reptilian coloration. 2023. *Current Biology*. In press.

Ullate-Agote A, Tzika AC. Characterization of the leucistic Texas rat snake Pantherophis obsoletus. Frontiers in Ecology and Evolution. 2021 Feb 9;9:583136.

Ullate-Agote, A., Burgelin, I., Debry, A., Langrez, C., Montange, F., Peraldi, R., Daraspe, J., Kaessmann, H., Milinkovitch, M.C. and Tzika, A.C., 2020. Genome mapping of a LYST mutation in corn snakes indicates that vertebrate chromatophore vesicles are lysosome-related organelles. Proceedings of the National Academy of Sciences, 117(42), pp.26307-26317.

Ullate-Agote, A., Milinkovitch, M.C. and Tzika, A.C., 2015. The genome sequence of the corn snake (Pantherophis guttatus), a valuable resource for EvoDevo studies in squamates. International Journal of Developmental Biology, 58(10-11-12), pp.881-888.

Ullate-Agote, A., and Tzika, A.C. (2021). Characterization of the leucistic Texas rat snake Pantherophis obsoletus. Frontiers in Ecology and Evolution 9, 583136.

Velotta JP, Wegrzyn JL, Ginzburg S, Kang L, Czesny S, O'Neill RJ, McCormick SD, Michalak P, Schultz ET. Transcriptomic imprints of adaptation to fresh water: parallel evolution of osmoregulatory gene expression in the Alewife. Molecular ecology. 2017 Feb;26(3):831-48.

Vickrey, A.I., Bruders, R., Kronenberg, Z., Mackey, E., Bohlender, R.J., Maclary, E.T., Maynez, R., Osborne, E.J., Johnson, K.P., Huff, C.D. and Yandell, M., 2018. Introgression of regulatory alleles and a missense coding mutation drive plumage pattern diversity in the rock pigeon. Elife, 7, p.e34803.

Vonk, F.J., Admiraal, J.F., Jackson, K., Reshef, R., de Bakker, M.A., Vanderschoot, K., van den Berge, I., van Atten, M., Burgerhout, E., Beck, A. and Mirtschin, P.J., 2008. Evolutionary origin and development of snake fangs. Nature, 454(7204), pp.630-633.

Vonk, F.J., Casewell, N.R., Henkel, C.V., Heimberg, A.M., Jansen, H.J., McCleary, R.J., Kerkkamp, H.M., Vos, R.A., Guerreiro, I., Calvete, J.J. and Wüster, W., 2013. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. Proceedings of the National Academy of Sciences, 110(51), pp.20651-20656.

Wakamatsu, K. and Ito, S., 2021. Melanins in vertebrates. Pigments, Pigment Cells and Pigment Patterns, pp.45-89.

Wallace, A.R., 1877. The colors of animals and plants. The American Naturalist, 11(11), pp.641-662.

Wallace, A.R., 1882. Dr. Fritz Müller on some difficult cases of mimicry. Nature, 26(656), pp.86-87.

Wang G, Yang E, Smith KJ, Zeng Y, Ji G, Connon R, Fangue NA, Cai JJ. Gene expression responses of threespine stickleback to salinity: implications for salt-sensitive hypertension. Frontiers in genetics. 2014 Sep 11;5:312.

Wang GD, Fan RX, Zhai W, Liu F, Wang L, Zhong L, Wu H, Yang HC, Wu SF, Zhu CL, Li Y. Genetic convergence in the adaptation of dogs and humans to the high-altitude environment of the Tibetan plateau. Genome biology and evolution. 2014 Apr 4;6(8):2122-8.

Wang, T. and Rindom, E., 2021. The physiological response to digestion in snakes: A feast for the integrative physiologist. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology, 254, p.110891.

Wasserman BA, Paccard A, Apgar TM, Des Roches S, Barrett RD, Hendry AP, Palkovacs EP. Ecosystem size shapes antipredator trait evolution in estuarine threespine stickleback. Oikos. 2020 Sep.

Wen, B., Chen, Y., Li, H., Wang, J., Shen, J., Ma, A., Qu, J., Bismuth, K., Debbache, J., Arnheiter, H. and Hou, L., 2010. Allele-specific genetic interactions between Mitf and Kit affect melanocyte development. Pigment cell & melanoma research, 23(3), pp.441-447.

Wiberg, R.A.W., Gaggiotti, O.E., Morrissey, M.B. and Ritchie, M.G., 2017. Identifying consistent allele frequency differences in studies of stratified populations. Methods in ecology and evolution, 8(12), pp.1899-1909.

Williams M. Hydrodynamics and salt dispersion in intermittently closed bar-built estuaries (Doctoral dissertation, UC Berkeley).

Williams, T.D., 2008. Individual variation in endocrine systems: moving beyond the 'tyranny of the Golden Mean'. Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1497), pp.1687-1698.

Woodcock, M.R., Vaughn-Wolfe, J., Elias, A., Kump, D.K., Kendall, K.D., Timoshevskaya, N., Timoshevskiy, V., Perry, D.W., Smith, J.J., Spiewak, J.E. and Parichy, D.M., 2017. Identification of mutant genes and introgressed tiger salamander DNA in the laboratory axolotl, Ambystoma mexicanum. Scientific reports, 7(1), p.6.

Wright, S., 1917. Color inheritance in mammals: V. The Guinea-pig—Great Diversity in Coat-pattern, Due to Interaction of Many Factors in Development—Some Factors Hereditary, Others of the Nature of Accidents in Development. Journal of Heredity, 8(10), pp.476-480.

Wright, S., 1920. The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs. Proceedings of the National Academy of Sciences, 6(6), pp.320-332.

Wu, S., Huang, J., Li, Y., Zhao, L. and Liu, Z., 2022. Analysis of yellow mutant rainbow trout transcriptomes at different developmental stages reveals dynamic regulation of skin pigmentation genes. Scientific Reports, 12(1), p.256.

Wysoker, A., Tibbetts, K., and Fennell, T. (2013). Picard tools version 1.90. Available online at picard. sourceforge. net.

Yamamoto, H., Takeuchi, S., Kudo, T., Sato, C. and Takeuchi, T., 1989. Melanin production in cultured albino melanocytes transfected with mouse tyrosinase cDNA. The Japanese Journal of Genetics, 64(2), pp.121-135.

Yamamoto, K. and Vernier, P., 2011. The evolution of dopamine systems in chordates. Frontiers in neuroanatomy, 5, p.21.

Yamasaki, Y.Y., Mori, S., Kokita, T. and Kitano, J., 2019. Armour plate diversity in Japanese freshwater threespine stickleback (Gasterosteus aculeatus). Evolutionary Ecology Research, 20(1), pp.51-67.

Yoshioka, S. and Akiyama, T., 2021. Mechanisms of Feather Structural Coloration and Pattern Formation in Birds. In Pigments, Pigment Cells and Pigment Patterns (pp. 343-366). Springer, Singapore.

Zhou, D., Ota, K., Nardin, C., Feldman, M., Widman, A., Wind, O., Simon, A., Reilly, M., Levin, L.R., Buck, J. and Wakamatsu, K., 2018. Mammalian pigmentation is regulated by a distinct cAMP-dependent mechanism that controls melanosome pH. Science signaling, 11(555), p.eaau7987.

Ziegler, I., 2003. The pteridine pathway in zebrafish: regulation and specification during the determination of neural crest cell-fate. Pigment cell research, 16(3), pp.172-182.

Zorina-Lichtenwalter, K., Lichtenwalter, R.N., Zaykin, D.V., Parisien, M., Gravel, S., Bortsov, A. and Diatchenko, L., 2019. A study in scarlet: MC1R as the main predictor of red hair and exemplar of the flip-flop effect. Human molecular genetics, 28(12), pp.2093-2106.

APPENDIX

Appendix A: Chapter 1

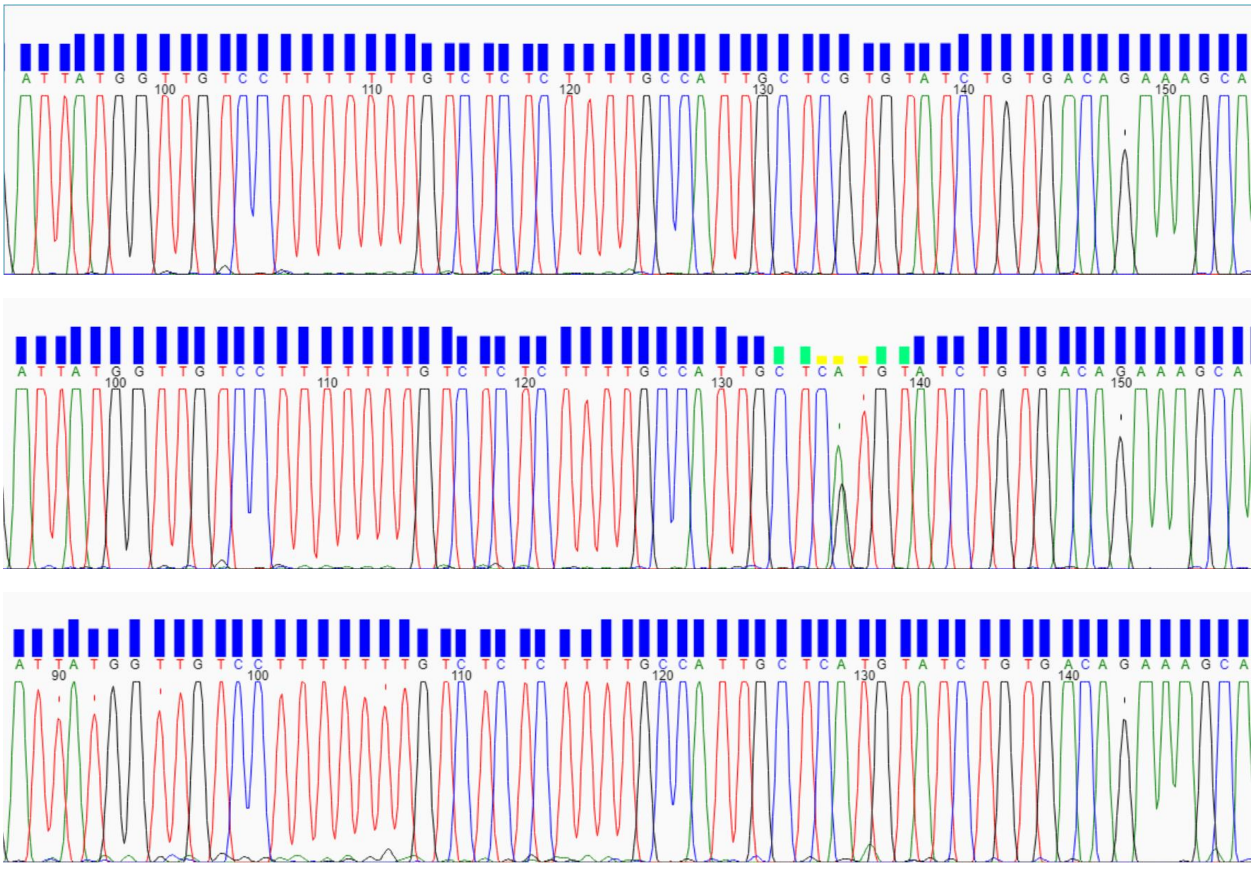


Figure A1. Chromatographs from Sanger sequencing of a known non-piebald (top), heterozygote (middle), and piebald.

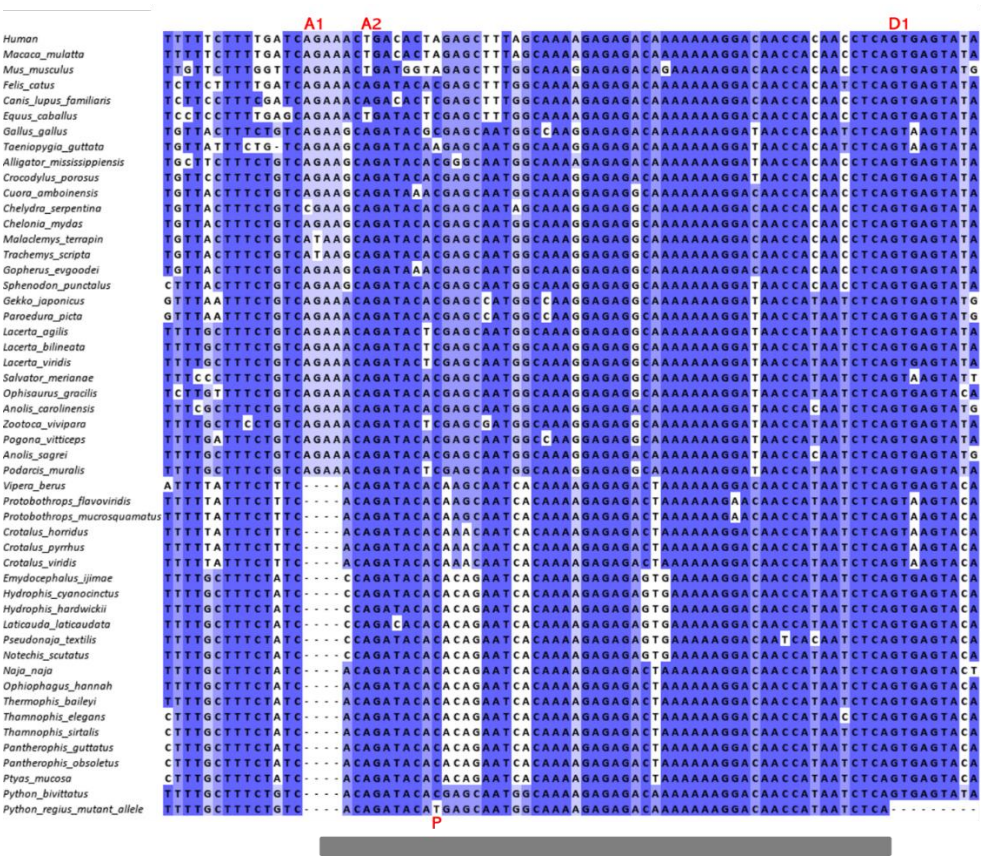


Figure A2. Multispecies (vertebrates) sequence alignment. *tfec* coding exon 5 (delineated by gray rectangle) shows a nonsense mutation (P annotation). Splice acceptor sites shown by A1 and A2 annotations. Splice donor sites indicated with D1 annotation. A 4 bp deletion results in loss of A1 splice acceptor site in snakes. The alignment is coloured by percentage identity.

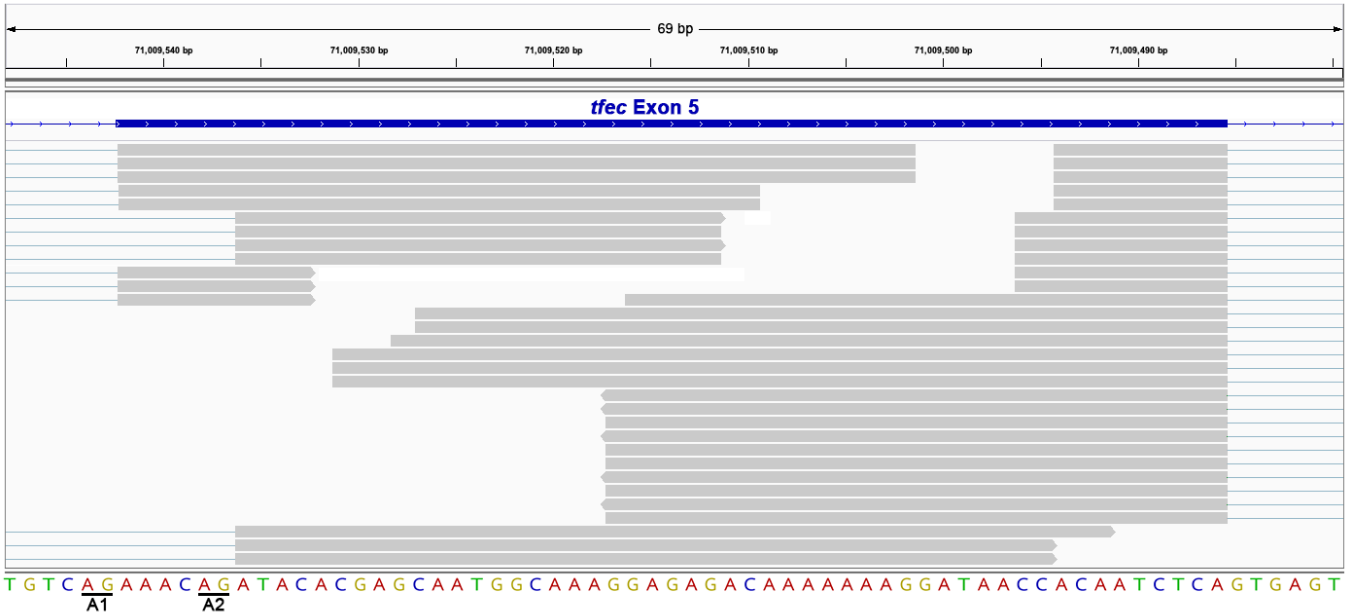


Figure A3. Alternative splicing of *tfec* exon 5 in *Anolis sagrei*. RNA-seq reads from embryonic (limbs and eyes) and adult (brain, liver, and skin) tissues of *A. sagrei* demonstrate the presence of two alternative splice acceptor sites at the 5' end of *tfec* exon 5. Blue = *tfec* exon 5, gray = individual RNA-seq reads, A1 = Acceptor site 1, A2 = Acceptor site 2.

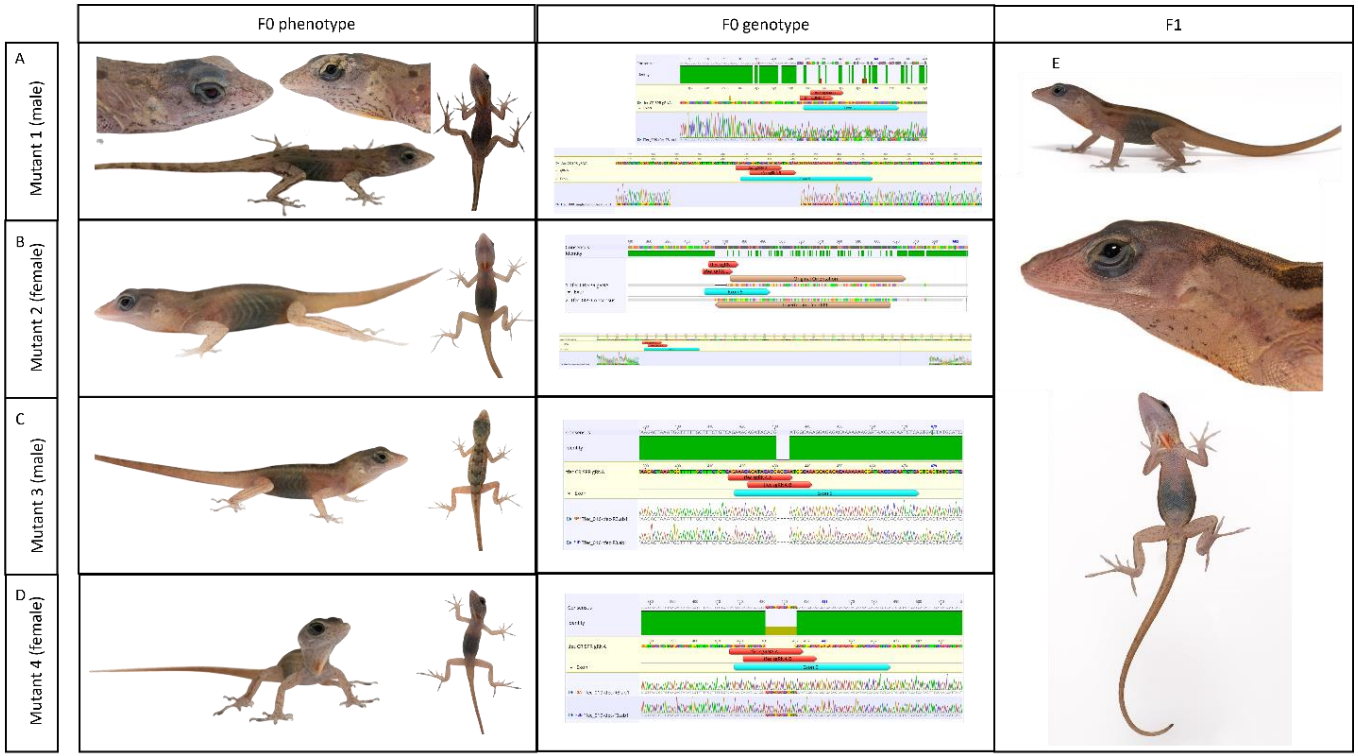


Figure A4. Phenotypes and chromatograms of *tfec* exon 5 from mutant *A. sagrei* (A) Mutant 1: Small patches of skin above the left eye are wild type in appearance. Overlapping signal indicates presence of multiple alleles (top panel). 56bp deletion shown in bottom panel. 1bp deletion not shown. (B) Mutant 2: Two mutant alleles detected, a local ~190bp inversion (middle panel) and a 295bp deletion (bottom panel). (C) Mutant 3: 4bp deletion. (D) Mutant 4: 13bp insertion. (E) F1 female showing that the pigmentation phenotype is transmitted to the F1 generation.

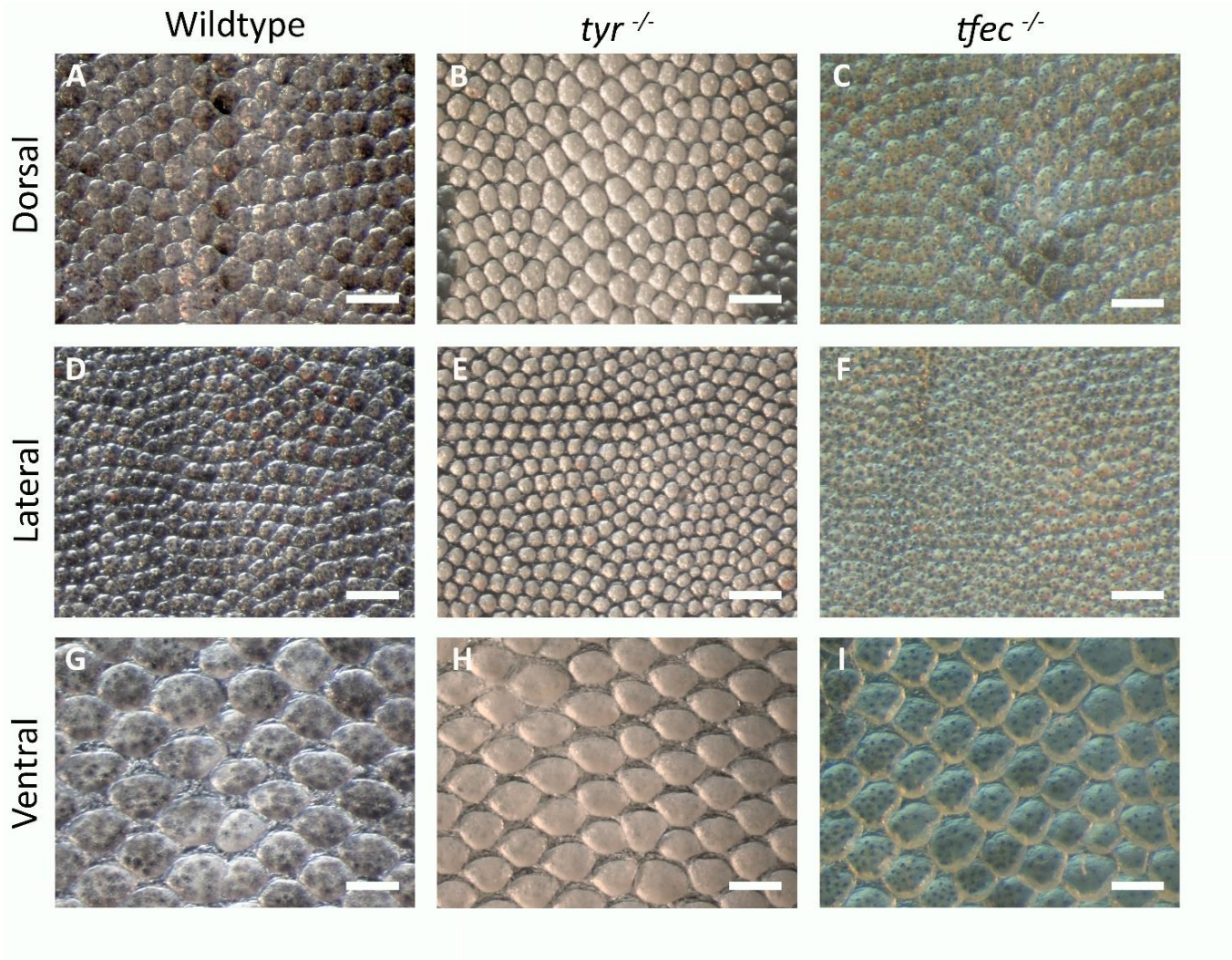


Figure A5. Close up views of *A. sagrei* skin from wild type, *tyr*$^{-/-}$, and *tfec* $^{-/-}$ hatchlings. Panels display (A, D, G) wild type, (B, E, H) tyrosinase and (C, F, I) *tfec* mutants. (A-C) scales along the dorsal midline. (D-F) scales along the side of the lizard. (G-

I) scales along the belly of the lizard. Scale bar for all panels, 250 µm. Reading frame mutations in *tyr* and *tfec* do not affect scale patterning or size but do impact coloration.

Appendix B: Chapter 2

Table B1. Ball python samples collected with clown phenotype

Sample_ID	Target_morph	Other_1	Other_2	Other_3	Other_4
HR0050	Clown	Pastel	.	.	.
MC0063	Clown	Caramel	.	.	Possible_het
MC0070	Clown	Pastel	.	.	.
MC0085	Clown	Blade	Butter	.	.
MC0125	Clown	Super_Pastel	Mojave	.	.
MC0132	Clown	Orange_Dream	.	.	.
MC0143	Clown	Lesser	Spotnose	.	.
MC0152	Clown	Blade_(Super?)	Pastel	Spider	.
MC0237	Clown	Pastel	Leopard	.	.
MC0253	Clown	Lesser	.	.	.
MC0254	Clown	Enchi	.	.	.
MC0257	Clown	.	.	.	.
MC0357	Clown	Pastel	Poss_leopard	.	.
MC0358	Clown	Banana	.	.	.
MC0397	Clown	piebald?	.	.	.
HR021-1	Clown	.	.	.	.
MC0408	Clown	Lace	piebald	.	.
MC0414	Clown	Enchi	Pastel	Yellowbelly	.
MC0417	Clown	Pastel	Het Lavendar	50% het pied	.

MC0418	Clown	Enchi	Pastel	Yellowbelly	Asphalt
MC0420	Clown	Asphalt	Cinnamon	Pastel	.
TDOT060	Clown	piebald?	.	.	.
21HR014-1	Clown	Black pastel	.	.	.
21HR014-3	Clown	Black pastel	Enchi	.	.
21HR020-1	Clown	Blade	.	.	.
21HR020-2	Clown	Blade	.	.	.
21HR020-3	Clown	.	.	.	.
HR0049	Clown	Super pastel	.	.	.
18HR021-1	Clown	Enchi	Lesser	.	.
20HR034-5	Clown	Blade	.	.	.
21HR020-4	Clown	Banana	blade	.	.
DM00021	Clown	Blade	Ultramel	.	.
Miller-01	Clown	Blade	Pastel	.	.
DPR0003	Clown	orange dream		.	.
21HR021-1	Clown	lesser	possible_red_stripe	.	.
21HR021-3	Clown	.	.	.	.
21HR0020-4	Clown	Banana	Blade		
21HR0050-2	Clown				

Table B2. Ball python samples collected with VPI axanthic phenotype.

Sample_ID	Target_morph	Other_1	Other_2	Other_3	Other_4
SBS007	VPI_Axanthic	.	.	.	.
SBS008	VPI_Axanthic	.	.	.	.
MC0075	VPI_Axanthic	.	.	.	.
MC0321	VPI_Axanthic	.	.	.	.
MC0345	VPI_Axanthic	.	.	.	.

Table B3. Reference samples (genetic stripe) compared to clown.

Sample_ID	Target_morph	Other_1	Other_2	Other_3	Other_4
DM0013	Genetic_Stripe	Pastel	Het_Enhancer	.	.
MC0043	Genetic_Stripe	Pastel	Yellowbelly	.	.
MC0108	Genetic_Stripe	Het_Enhancer	.	.	.
MC0330	Genetic_Stripe	Pastel	Het_Enhancer	.	.
MC0195	Genetic_Stripe	Yellowbelly	.	.	.
MC0308	Genetic_Stripe	Albino	.	.	.
MC0422	Genetic_Stripe	Asphalt	Yellowbelly	.	.
MC0423	Genetic_Stripe	Piebald	.	.	.
19HR0053-1	Genetic_Stripe	Het_Enhancer	Pastel	.	.
HR0036	Genetic_Stripe	Banana	HRA	.	.
20HR053-3	Genetic_Stripe	Het_Enhancer	Pastel	.	.
DPR0006	Genetic_Stripe	Lavender_Albino	Pastel	.	.
21HR053-1	Genetic_Stripe	Enhancer	Pastel	.	.
HR0038	Genetic stripe	het enhancer			
MC0173	Genetic_Stripe	Enhancer	.	.	.
20HR0038-2	Genetic_stripe	Enhancer	.	.	.

Table B4. Refence samples (Champagne) compared to clown.

Sample_ID	Target_morph	Other_1	Other_2	Other_3
HR0008	Champagne	.	.	.
HR0013	Champagne	HRA	het_pied	.
HR0035	Champagne	Albino	Enchi	.
20HR013-3	Champagne	Mojave	Possible_HRA	Possible_Het_Pied
20HR013-6	Champagne	Banana	Possible_Het_Pied	.
20HR016-5	Champagne	Albino	Enchi	.
20HR016-6	Champagne	Albino	Enchi	Possible.super.enchi
21HR0008-2	Champagne	Lesser	.	.
21HR006-2	Champagne	Super_pastel	.	.
21HR006-3	Champagne	Super_cinnamon	.	possible_black_pastel
21HR046-1	Champagne	Cinnamon	Pastel	.
21HR046-2	Champagne	Champagne	Pastel	Super cinnamon
DPR0007	Champagne	Pastel	Ghost	.
MC0001	Champagne	Asphalt	Enchi	.
MC0190	Champagne	HRA	Het_piebald	.
TDOT019	Champagne	HRA	Pastel	.
TDOT046	Champagne	Pastel	.	.
19HR006-1	Champagne	Super pastel	Possible black pastel and/or cinnamon and/or HRA	
21HR008-1	Champagne	Possible red stripe	Possible het clown	

Table B5. Refence samples (ivory) compared to VPI axanthic.

Sample_ID	Target_morph	Other_1	Other_2	Other_3	Other_4

MC0031	Ivory	Het_Piebald	Pastel	.	.
19HR0002-4	Ivory	Spider	.	.	.
TOOT002	Ivory	Pastel	.	.	.
TOOT058	Ivory	.	.	.	.
MC0104	Ivory	Pastel	Het_Piebald	.	.
MC0166	Ivory	.	.	.	.
MC0174	Ivory	Citrus_Pastel	.	.	.
MC0233	Ivory	Pastel	Spider	.	.
MC0240	Ivory	Enchi	.	.	.
MC0250	Ivory	Casper_Ghost	.	.	.
MC0376	Ivory	.	.	.	.
MC0377	Ivory	.	.	.	.
MC0386	Ivory	.	.	.	.
DPR0001	Ivory	Super_enchi	.	.	.
HR063	Ivory	Het_Clown	.	.	.

Table B6. Refence samples (black pastel) compared to VPI axanthic.

Sample_ID	Target_morph	Other_1	Other_2	Other_3	Other_4
HR0031	Black_Pastel	Het Lavender			
20HR057-4	Black_Pastel	Banana	Possible het pied		
21HR014-6	Black_Pastel	Enchi	Het clown		
MC0014	Black_Pastel	Banana	Pastel		
MC0048	Black_Pastel	Banana			
MC0211	Black_Pastel				
MC0302	Black_Pastel	Spider	Pastel	Banana	
MC0360	Black_Pastel				
MC0436	Black_Pastel	Yellowbelly	Het Lavender	Poss Het	

				Pied	
MC0437	Black_Pastel	Het Lavender	Poss Het Pied		
TDOT018	Black_Pastel	GHI			
DM0002	Black_Pastel	.	Het_Piebald	.	.
DM0009	BLack_Pastel	Het_Piebald	.	.	.
HR0006	Black_Pastel	Het_Red_Axanthic	Pastel	.	.
HR0006-2	Black_Pastel	Cinnamon	.	.	.
HR0006-6	Black_Pastel	HRA	.	.	.
HR0014	Black_Pastel	Het_Clown	.	.	.
HR0018	Black_Pastel	Pinstripe	Het_Albino	.	.
HR0019	Black_Pastel	.	.	.	.
HR0024	Black_Pastel	Piebald	.	.	.
MC0188	Black_Pastel	Het_Lavender	Het_Piebald	.	.
HR0055	Black_Pastel	Lavender	.	.	.
MC0137	Black_Pastel	Lavender			
MC0429	Black_Pastel	Lavender	Black_Head		
TDOT030	Black_Pastel	Asphalt	GHI	.	.
MC0067	Black_Pastel	Super_Pastel			

Appendix C: Chapter 3

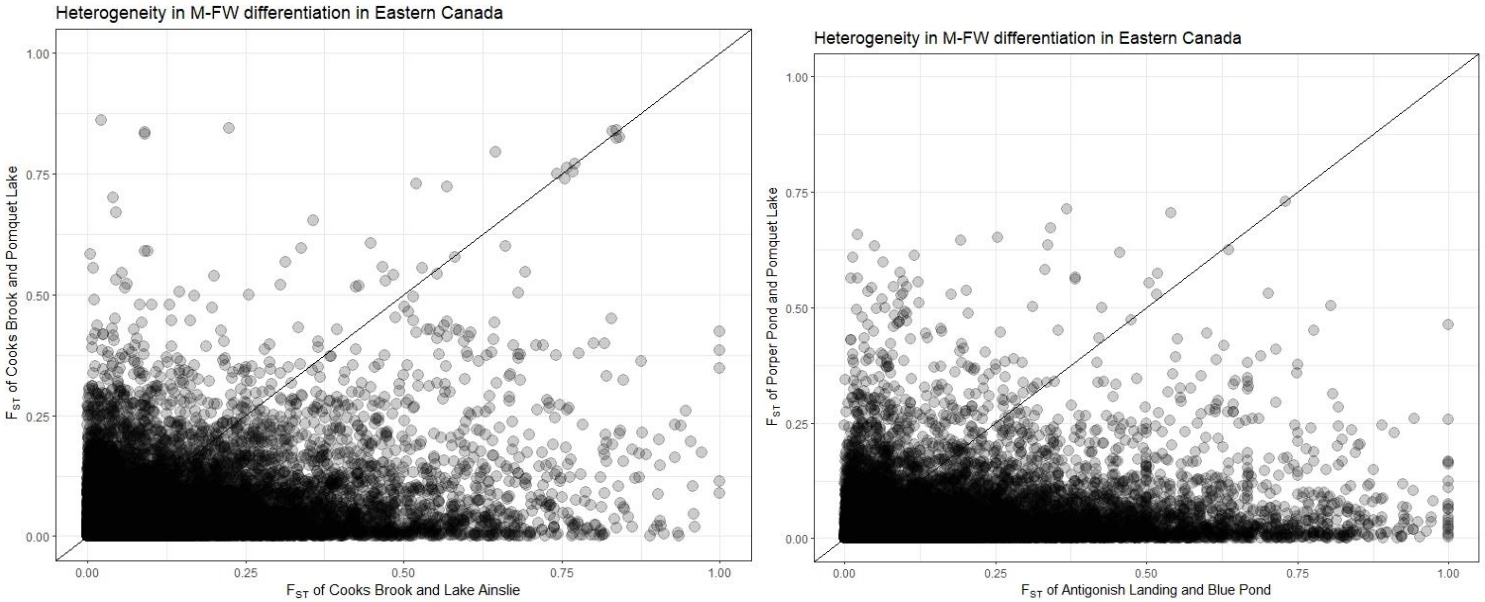


Figure C1. A sample of the heterogeneity of M-FW differentiation in Eastern Canada (Newfoundland and Nova Scotia).

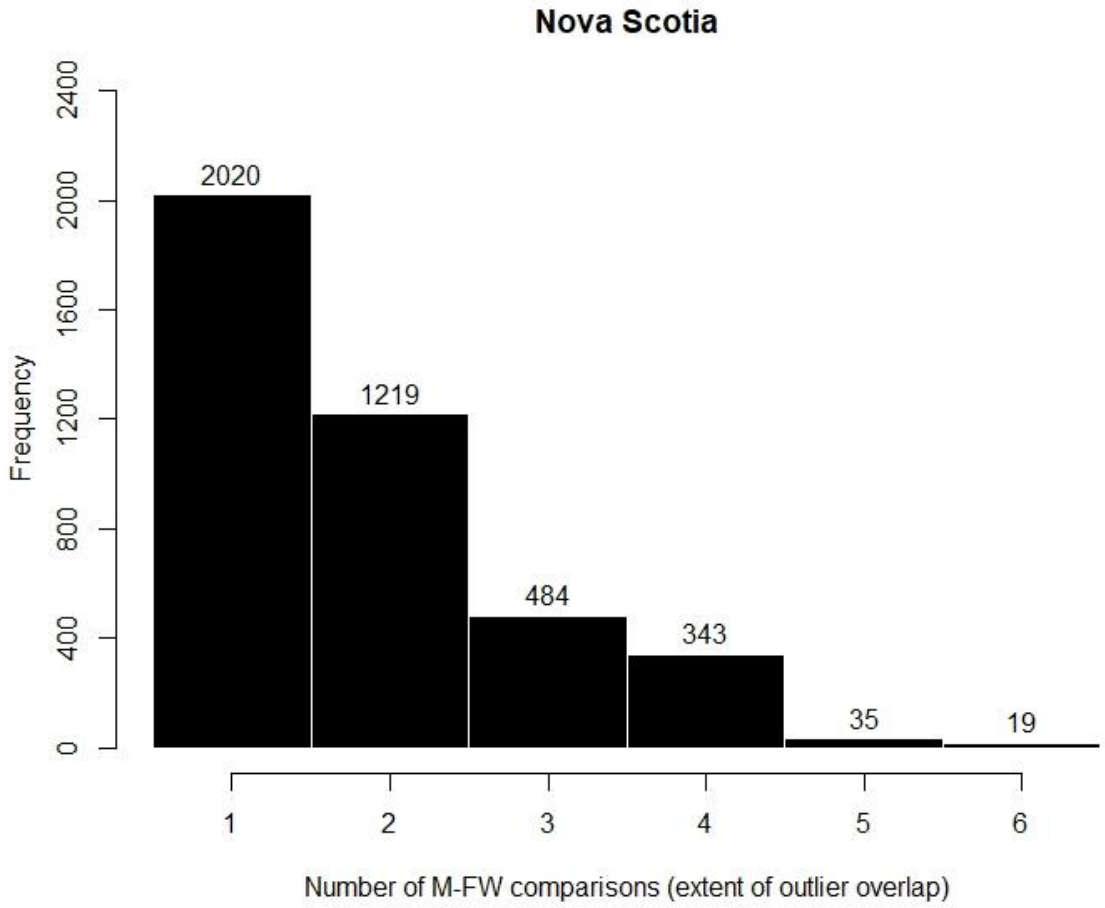


Figure C2a. Overlap of F_{ST} outliers in marine-freshwater (M-FW) comparisons in Nova Scotia.

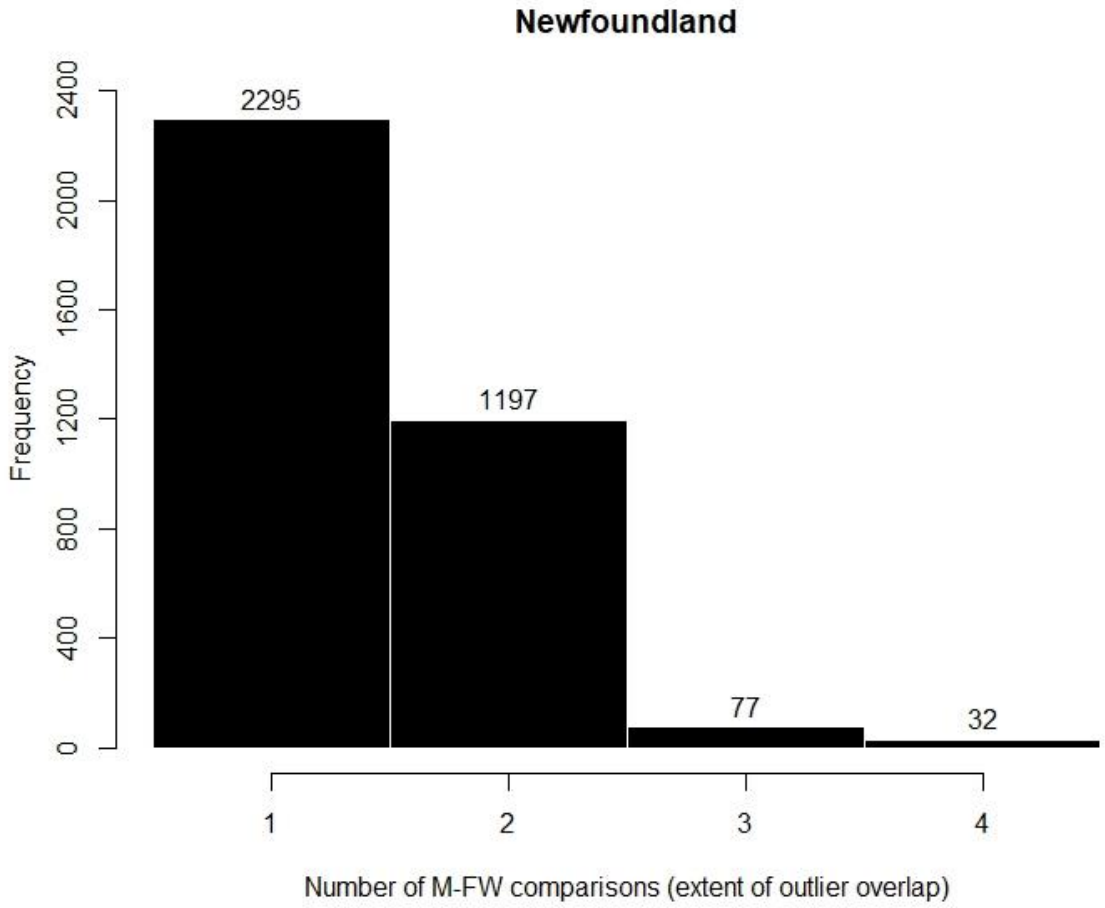


Figure C2b. Overlap of FST outliers from M-FW comparisons in Newfoundland.

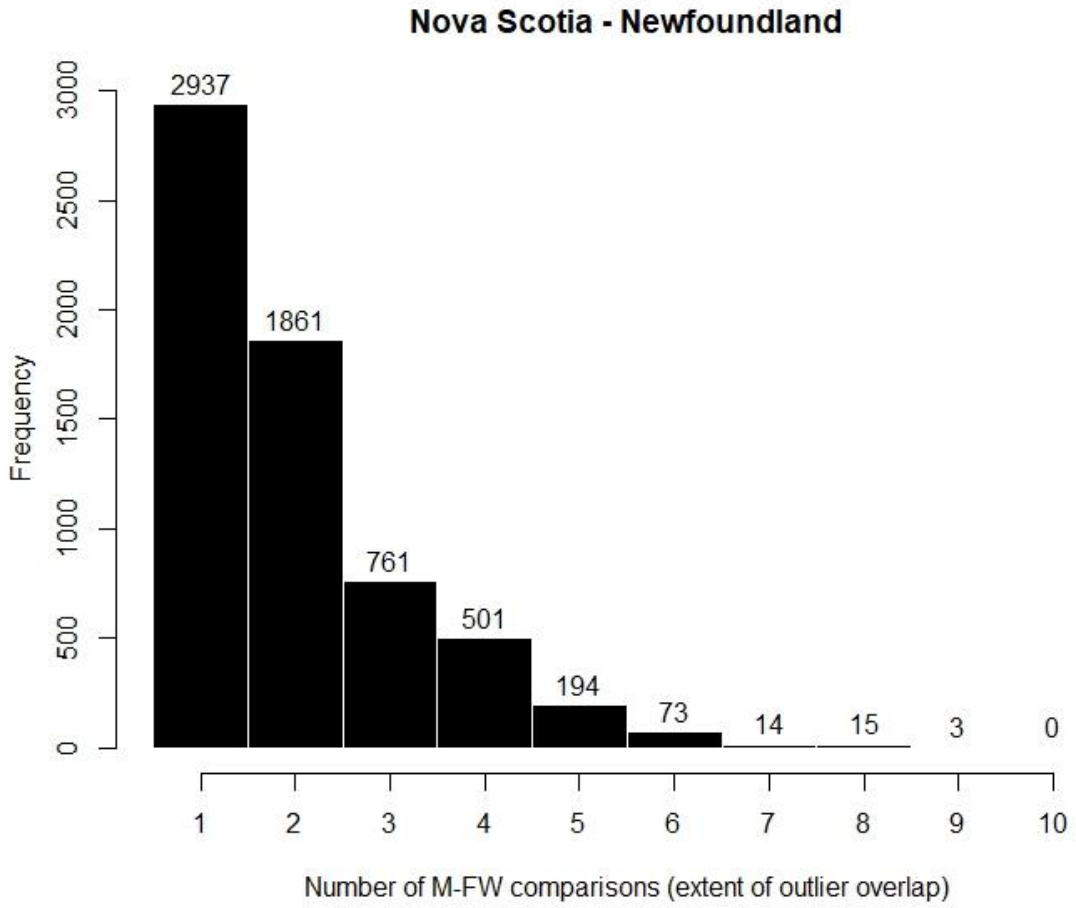


Figure C2c. Overlap of FST outliers from M-FW comparisons between populations in Nova Scotia and Newfoundland.

Table C1. Median pair-wise FST estimates.

	Antigonish Landing	Porper Pond	Pomquet Lake	Lake Ainslie	Black River	Humber Arm	Cooks Brook	Pinchgut Lake	Blue Pond
Antigonish Landing	0	0.025582	0.024025	0.053188	0.048141	0.024419	0.031298	0.023474	0.075191
Porper Pond	0.025582	0	0.015813	0.047885	0.041689	0.023149	0.025055	0.019944	0.07266
Pomquet Lake	0.024025	0.015813	0	0.048578	0.040121	0.021179	0.022015	0.011818	0.07245
Lake Ainslie	0.053188	0.047885	0.048578	0	0.021105	0.051514	0.053934	0.049448	0.082419
Black River	0.048141	0.041689	0.040121	0.021105	0	0.048189	0.046671	0.043252	0.078213
Humber Arm	0.024419	0.023149	0.021179	0.051514	0.048189	0	0.024473	0.021789	0.073272
Cooks Brook	0.031298	0.025055	0.022015	0.053934	0.046671	0.024473	0	0.025338	0.070571
Pinchgut Lake	0.023474	0.019944	0.011818	0.049448	0.043252	0.021789	0.025338	0	0.071363
Blue Pond	0.075191	0.07266	0.07245	0.082419	0.078213	0.073272	0.070571	0.071363	0

Table C2. F_{ST} outliers and genes with outliers or within 5 kb of outliers.

Chromosome	Position	Gene
groupI	1331845	ywhae1
groupI	6719755	gria4a
groupI	8648106	sik2b
groupI	12340803	ca4b
groupI	20277749	kcnh7
groupI	20277807	kcnh7
groupI	24627952	ARHGAP6
groupII	853492	lto1
groupII	2443237	slc7a6
groupII	2443443	slc7a6
groupII	3259808	sh3gl3a
groupII	3259808	cx36.7
groupII	5579787	phkb
groupII	8309975	cmtm3
groupII	8309975	galns
groupII	13471989	tspan15
groupII	18677147	skor1a
groupII	19748833	mpped2a
groupII	22161363	iqgap1
groupIII	6006502	dusp12
groupIII	6006502	prg4a
groupIII	10756260	st6galnac5a
groupIII	10756267	st6galnac5a
groupIII	12726940	cacnb2b
groupIV	2468647	rapgef2
groupIV	4426498	fat2
groupIV	6432893	mid1ip1l
groupIV	6432893	tspan7

groupIV	6432927	mid1ip1l
groupIV	6432927	tspan7
groupIV	7540301	si:ch211-221f10.2
groupIV	20570132	tbc1d22a
groupIV	22794835	LARGE1
groupIV	29711546	iqsec3a
groupIX	12921104	CACNG4
groupIX	14146937	si:dkey-16m19.1
groupIX	18039513	zgc:153247
groupIX	20190341	soul4
groupV	4335166	wnt3
groupV	9616054	taok2b
groupV	9616054	cabp5a
groupV	9616055	taok2b
groupV	9616055	cabp5a
groupV	9616056	taok2b
groupV	9616056	cabp5a
groupV	9616057	taok2b
groupV	9616057	cabp5a
groupV	9616059	taok2b
groupV	9616059	cabp5a
groupV	9616061	taok2b
groupV	9616061	cabp5a
groupV	10017653	mybpc2b
groupV	11332724	pms2
groupV	11332724	eif2ak1
groupVI	1462374	STK32C
groupVI	2440049	nanp
groupVI	2440049	polr1b
groupVI	3315969	cdh23
groupVI	3327118	cdh23

groupVI	7831714	lrrc1
groupVI	9493911	lgi1a
groupVII	14195649	cltca
groupVII	17121794	ar
groupVII	23887141	stk10
groupVII	23887142	stk10
groupVII	23887143	stk10
groupVII	24477131	ncam1a
groupVIII	2093733	tprg1
groupVIII	3336195	adcy1a
groupVIII	3648884	GMDS
groupVIII	6320700	ATP11B
groupVIII	6320701	ATP11B
groupVIII	12982440	dmap1
groupVIII	12982440	hnrnpm
groupVIII	12982505	dmap1
groupVIII	12982505	hnrnpm
groupVIII	12982510	dmap1
groupVIII	12982510	hnrnpm
groupVIII	12982637	dmap1
groupVIII	12982637	hnrnpm
groupVIII	12982683	dmap1
groupVIII	12982683	hnrnpm
groupVIII	15926670	ssbp4
groupVIII	15926670	fkbp8
groupX	1235804	gsk3aa
groupX	2109575	AGO3
groupX	3427268	ext1b
groupX	8541803	gabbr2
groupX	8545943	gabbr2
groupX	13170833	ST3GAL1

groupX	13170834	ST3GAL1
groupX	15439245	zgc:100906
groupX	15439245	si:dkey-27c15.3
groupXI	8531205	cdc42ep4a
groupXI	14593614	si:ch211-278a6.1
groupXI	15205337	pctp
groupXI	15205342	pctp
groupXII	2799938	cpne9
groupXII	7058307	wdr6
groupXII	7889552	ybx1
groupXII	7889552	arhgef16
groupXII	10562652	SCN8A
groupXII	10562653	SCN8A
groupXII	13484687	hsd17b10
groupXII	16890050	foxp3b
groupXIII	843366	olfml2a
groupXIII	5852755	ipo11
groupXIII	7315597	ntrk2a
groupXIII	9927130	si:ch211-130m23.3
groupXIII	12341448	zgc:77112
groupXIII	12341448	gins4
groupXIII	12341590	zgc:77112
groupXIII	12341590	gins4
groupXIII	12341662	zgc:77112
groupXIII	12341662	gins4
groupXIII	15556908	arvcfb
groupXIII	15556943	arvcfb
groupXIII	16916924	aacs
groupXIV	2512633	lamc3
groupXIV	2712891	usp20

groupXIV	6693066	lrrc75bb
groupXIV	9276197	snx18b
groupXIX	290484	ntrk3a
groupXIX	4346553	syt12
groupXIX	10096107	drd4a
groupXIX	10096107	deaf1
groupXIX	10096160	drd4a
groupXIX	10096160	deaf1
groupXIX	10096162	drd4a
groupXIX	10096162	deaf1
groupXIX	10302716	met
groupXIX	10317294	met
groupXIX	10798959	nts
groupXIX	12527947	OTUD7A
groupXV	2560	hikeshi
groupXV	2660373	dnajc17
groupXV	2660373	INF2
groupXV	5969349	ivd
groupXV	5969349	itpka
groupXV	8157598	slitrk3b
groupXV	8438378	hs6st1b
groupXV	12580687	lrfn5a
groupXVI	6583300	IMPG2
groupXVI	7672101	sytl5
groupXVI	7892340	col4a1
groupXVI	14668657	lrp1bb
groupXVI	17339020	dgkg
groupXVI	17339022	dgkg
groupXVII	829298	dnase1l4.1
groupXVII	1698078	cept1b
groupXVII	1698078	dram2b

groupXVII	1698078	zgc:162255
groupXVII	3377127	tulp1b
groupXVII	5074480	ppp1r16b
groupXVII	5074521	ppp1r16b
groupXVII	5074704	ppp1r16b
groupXVII	9157856	slc25a26
groupXVII	12472238	nfasca
groupXVIII	728783	cd2ap
groupXVIII	728872	cd2ap
groupXVIII	729091	cd2ap
groupXVIII	2308262	nrxn3b
groupXVIII	2308263	nrxn3b
groupXVIII	7122964	LAMA2
groupXVIII	10864543	mtmr9
groupXVIII	13214109	clvs2
groupXVIII	13263902	gtf3c2
groupXVIII	13263902	slc30a2
groupXVIII	13472763	si:ch1073-155h21.1
groupXVIII	13472763	zgc:112001
groupXVIII	13490437	lrfn2b
groupXX	7443937	cdkal1
groupXX	14009648	drd2l
groupXX	16943035	cnr2
groupXX	16943035	akirin1
groupXX	18784958	cyb5r4
groupXXI	3025247	POMP
groupXXI	6728474	KCNB2
groupXXI	7924735	zbtb47b
groupXXI	10057609	si:ch1073-406l10.2

Table C3. Enrichment of biological process of genes within 5 kb of F_{ST} outliers. Best matching species.

Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment	Pathway	Genes
0.043506064	2	9	51.97541	Adenylate cyclase-inhibiting dopamine receptor signaling pathway	drd2l drd4a
0.043506064	2	8	51.97541	Neg. reg. of cytosolic calcium ion concentration	drd2l drd4a
0.043506064	2	8	51.97541	Neg. reg. of synaptic transmission, glutamatergic	drd2l drd4a
0.060704103	2	10	34.65027	Oocyte maturation	ybx1 ar
0.060704103	2	10	34.65027	Neg. reg. of adenylate cyclase activity	drd2l drd4a
0.037306674	3	24	25.9877	Dopamine secretion	drd2l drd4a syt12
0.037306674	3	24	25.9877	Reg. of dopamine secretion	drd2l drd4a syt12
0.037306674	3	25	25.9877	Catecholamine secretion	drd2l drd4a syt12
0.037306674	3	25	25.9877	Reg. of amine transport	drd2l drd4a syt12
0.038039399	3	28	23.98865	Amine transport	drd2l drd4a syt12
0.038039399	3	26	22.27518	Dopamine transport	drd2l drd4a syt12
0.038039399	3	27	22.27518	Catecholamine transport	drd2l drd4a syt12
0.041969835	3	22	20.79016	Reg. of voltage-gated calcium channel activity	drd2l drd4a cacnb2b
0.043506064	3	27	19.49078	Reg. of calcium ion transmembrane transporter activity	drd2l drd4a cacnb2b
0.04737799	3	38	17.32514	Monoamine transport	drd2l drd4a syt12
0.057983305	6	260	5.241218	Reg. of ion transport	drd2l drd4a kcnb2 cacnb2b kcnh
0.049696951	21	1814	2.131804	Cell differentiation	ybx1 rapgef2 cx36.7 gsk3aa drd4 taok2b
0.052170446	21	1828	2.113231	Cellular developmental proc.	ybx1 rapgef2 cx36.7 gsk3aa drd4 taok2b

Table C4. Enrichment of GO terms associated with cellular component in genes within 5 kb of FST outliers. Best matching species.

Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment	Pathway	Genes
0.081106685	2	10	34.65027	Intrinsic component of synaptic membrane	drd2l lrfn2b
0.081106685	2	10	34.65027	Integral component of synaptic membrane	drd2l lrfn2b
0.0983624	3	58	10.75353	Basement membrane	col4a1 lamc3 lama2
0.081106685	13	1244	2.624001	Integral component of plasma membrane	drd2l tspan7 gria4a drd4a kcnb2 tspan15 ntrk2a gabbr lrfn2b
0.081106685	13	1265	2.549737	Intrinsic component of plasma membrane	drd2l tspan7 gria4a drd4a kcnb2 tspan15 ntrk2a gabbr lrfn2b

Appendix D: Chapter 4

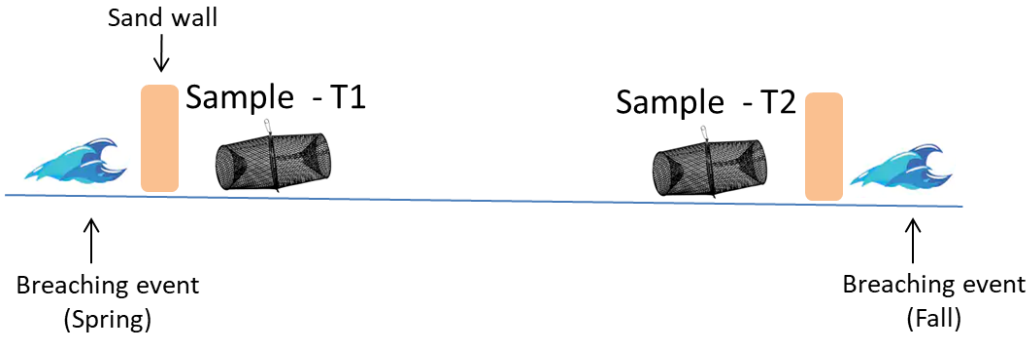


Figure D1. Diagram of sampling strategy of stickleback in each estuary between two breaching events. The first sample was taken shortly after the sandbar was rebuilt from a breaching event in the spring (T1). The second sampling time point was in the fall before the sandbar was broken in a new breaching event (T2). During this time while the sandbar is intact (T1 – T2), lentic estuaries transition from brackish to freshwater.

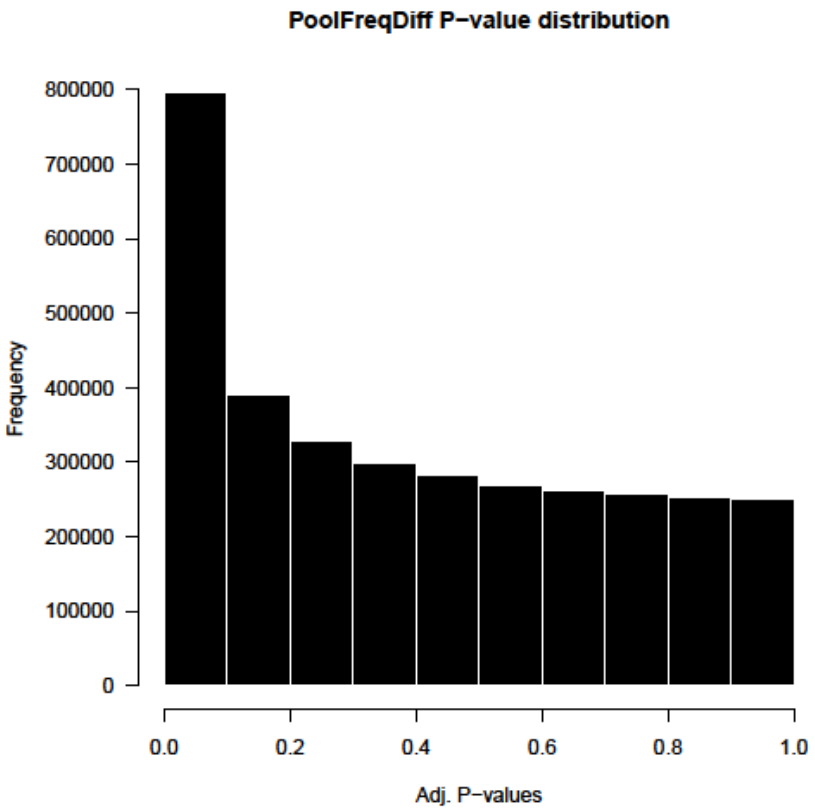


Figure D2. Histogram of P-values from the qGLM test (*PoolFreqDiff*) after adjusting for population structure using lambda (λ = 0.3) and prior to controlling for multiple hypothesis testing with *qvalue* program.

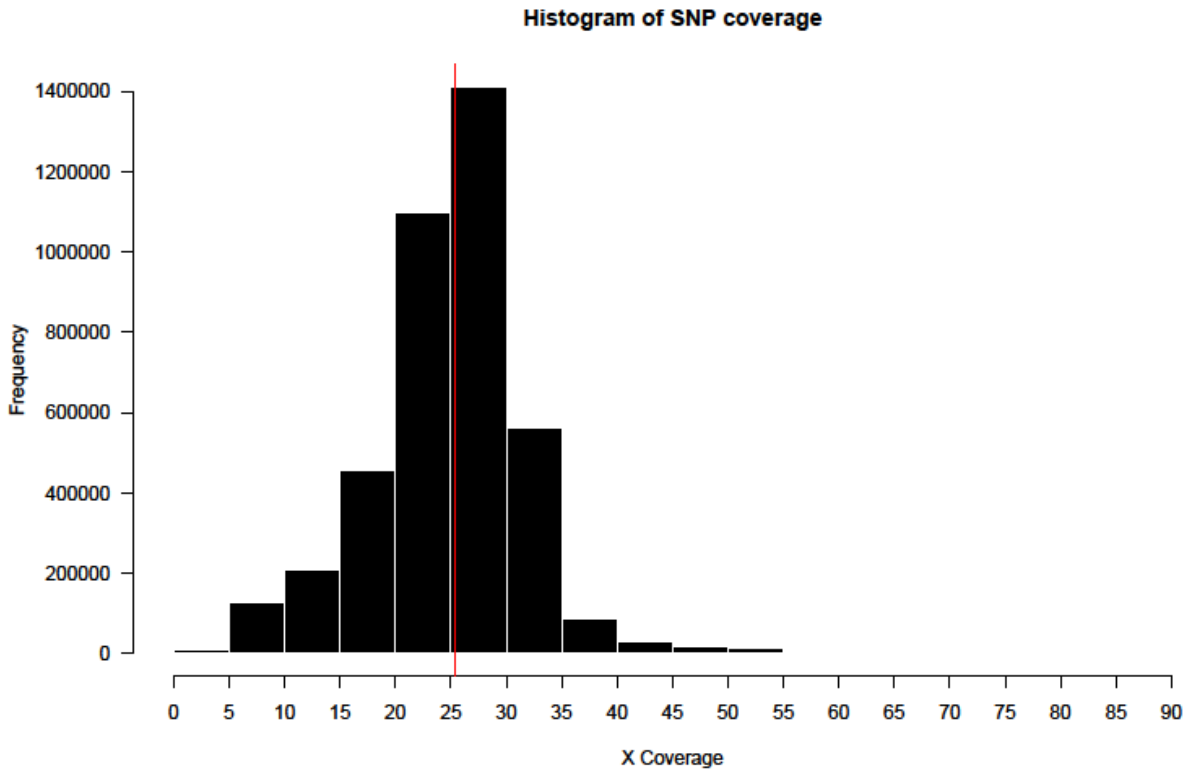


Figure D3. Histogram of minimum SNP coverage across estuaries. The average minimum coverage of 25.32 (SD = 6.96, range: 5 – 84) is denoted by the vertical red line.

Table D1. Site coordinates and size characteristics of each estuary. The size of each estuary was evaluated using Channel Area (CA), Permanently Flooded Area (CA + lakes, ponds, and side channels), and Total Wetland Area (maximum area that becomes seasonally flooded).

Site	Latitude	Longitude	Channel Area (km^2)	Permanently Flooded Area (km^2)	Total Wetland Area (km^2)
Scott Creek	37.04	-122.23	0.00635	0.01283	0.12509
Laguna Creek	36.984	-122.154	0.02032	0.02185	0.05354
Old Dairy Creek	36.955	-122.091	0.00004	4.00E-05	0.01331
Younger Lagoon	36.951	-122.067	0.01081	0.01081	0.02092
Lombardi Creek	36.962	-122.113	0.00117	0.00117	0.028
Waddell Creek	37.096	-122.278	0.02004	0.0204	0.16908

Table D2. Genic F_{ST} outliers.

Chromosome	Position	Overlapping gene
groupI	85261	col4a4
groupI	762830	nf1a
groupI	924144	ubash3bb
groupI	1210147	sgsm2
groupI	1222323	sgsm2
groupI	1356415	myo1cb
groupI	1356419	myo1cb
groupI	1612316	si:ch211-276c2.2
groupI	2925222	si:ch73-265h17.5
groupI	4497884	slc33a1
groupI	5872569	hip1
groupI	6494330	DYNC2H1
groupI	6649525	gria4a
groupI	6931938	FCHSD2
groupI	7877774	limk1a
groupI	9207104	npas1
groupI	10040313	ppm1lb
groupI	10075425	kpna4
groupI	10099523	smc4
groupI	10176672	schip1
groupI	12048484	si:dkey-243i1.1
groupI	13702383	frem2b
groupI	14725652	gab2
groupI	16276019	NOVA1
groupI	16712599	HNRNPUL1
groupI	17311060	capn5a
groupI	17977964	zmp:0000001073
groupI	18039618	zgc:101731

groupI	18383974	aldh3a2a
groupI	18418633	ulk2
groupI	18462682	usp32
groupI	18534255	brip1
groupI	19617593	ksr1b
groupI	20253216	kcnh7
groupI	21680480	obsl1a
groupI	22638012	erbb4a
groupI	25231590	rap2ab
groupI	26530632	cd99
groupI	26558938	dhrsx
groupI	27013742	hsf2bp
groupI	27459733	xdh
groupI	27783813	DES
groupII	168168	arnt2
groupII	484327	ano9b
groupII	1932057	dpy19l3
groupII	1996614	si:ch211-186j3.6
groupII	2023253	si:ch211-186j3.6
groupII	2036781	si:ch211-186j3.6
groupII	2687554	CDH8
groupII	2716104	CDH8
groupII	3450974	ano1
groupII	3561457	pcsk6
groupII	3632139	pop4
groupII	5064721	kif13ba
groupII	6008546	scml2
groupII	6310189	itgb6
groupII	9489069	plekha7b
groupII	10829463	zfhx3
groupII	11800158	furina

groupII	13939981	kif26ba
groupII	17868436	nod2
groupII	17887414	nkd1
groupII	18376096	FBN1
groupII	18579021	smad6a
groupII	18962456	KIF23
groupII	19168167	CHRNA7
groupII	19168186	CHRNA7
groupII	19357502	MYO9A
groupII	19511916	neo1b
groupII	19529077	neo1b
groupII	19723509	mpped2a
groupII	20020936	prmt3
groupII	20154359	NELL1
groupII	20163214	NELL1
groupII	20164745	NELL1
groupII	20223139	NELL1
groupII	20859948	kcnq1
groupII	21725408	MEGF11
groupII	21898229	otog
groupII	22080909	serinc4
groupII	22086970	serinc4
groupII	22294374	rbpms2b
groupII	22538893	fhod1
groupII	22632723	wdr59
groupIII	995837	ZNF521
groupIII	3255287	rem2
groupIII	3973089	zswim5
groupIII	4120217	ctnnd2b
groupIII	4593029	rnf220a
groupIII	6104099	chd7

groupIII	6179834	ca8
groupIII	7994872	itga6b
groupIII	8551306	kcnn1a
groupIII	8551317	kcnn1a
groupIII	11476793	acvr2bb
groupIII	12656591	si:ch211-106h4.9
groupIII	13585006	mib1
groupIII	13591612	mib1
groupIII	15214945	lig1
groupIII	15354432	kirrel1b
groupIII	16065943	wrnip1
groupIII	16098153	prkacbb
groupIII	16139254	si:ch1073-186i23.1
groupIII	16436611	snai2
groupIII	16611642	AGAP3
groupIII	16737274	nol7
groupIV	177633	rest
groupIV	196289	matr3l1.1
groupIV	1433258	cxcl14
groupIV	2135893	si:dkey-237h12.3
groupIV	2628302	FSTL5
groupIV	2670202	FSTL5
groupIV	2675175	FSTL5
groupIV	2952251	LIMCH1
groupIV	3199408	fgf24
groupIV	3337184	nudt6
groupIV	3337197	nudt6
groupIV	3381695	SPATA5
groupIV	5213913	fermt3b
groupIV	5490046	efemp2a

groupIV	6284022	drp2
groupIV	8329956	SLC16A2
groupIV	8424582	mid2
groupIV	8424584	mid2
groupIV	8436677	mid2
groupIV	10492440	spock1
groupIV	10961873	ndst1a
groupIV	11372731	clk4a
groupIV	12108815	irg1l
groupIV	12158965	irg1l
groupIV	13406322	dctn4
groupIV	13406323	dctn4
groupIV	13571693	ctnna1
groupIV	13578848	ctnna1
groupIV	14223518	ablim3
groupIV	14225308	ablim3
groupIV	14227500	ablim3
groupIV	15005199	TENM2
groupIV	15141467	EFNB1
groupIV	15867851	cnot7
groupIV	16007799	gpc3
groupIV	16013845	gpc3
groupIV	16104413	brd8
groupIV	16358824	si:ch73-49o8.1
groupIV	17630508	irf2
groupIV	17942415	ano6
groupIV	18234006	grm8a
groupIV	20995782	SLC2A13
groupIV	21642308	lrig3
groupIV	24122473	scube1
groupIV	24182241	scube1

groupIV	25003349	OSBPL8
groupIV	25003475	OSBPL8
groupIV	25068359	zdhhc17
groupIV	25713675	nxpe3
groupIV	25713799	nxpe3
groupIV	25713900	nxpe3
groupIV	25713920	nxpe3
groupIV	25713960	nxpe3
groupIV	25714242	nxpe3
groupIV	25828152	slc38a2
groupIV	25906741	tm7sf3
groupIV	26006387	setd7
groupIV	30100102	hgfa
groupIV	30228236	snd1
groupIV	30776503	anks1b
groupIV	31092938	si:ch73-352p4.8
groupIV	31109232	TULP3
groupIV	31662074	ppfia2
groupIV	31679576	ppfia2
groupIV	31741054	lin7a
groupIV	31934666	syt1a
groupIV	32127569	nup37
groupIX	610299	rbfox2
groupIX	1667470	klhl5
groupIX	1816241	slit2
groupIX	2286243	SLC24A3
groupIX	2377058	sh3pxd2aa
groupIX	4859265	cyp3a65
groupIX	4859265	cyp3a65
groupIX	4861626	cyp3a65
groupIX	4861626	cyp3a65

groupIX	4861730	cyp3a65
groupIX	4861730	cyp3a65
groupIX	4867044	cyp3a65
groupIX	4867044	cyp3a65
groupIX	4934834	sdk1a
groupIX	5208341	si:dkeyp-9d4.3
groupIX	5959655	tll1
groupIX	6407015	lrba
groupIX	7260230	inpp4b
groupIX	7263639	inpp4b
groupIX	7283138	inpp4b
groupIX	13978810	smc3
groupIX	14967746	pkd1a
groupIX	15314114	afmid
groupIX	15728645	xylt1
groupIX	18118315	micall2b
groupIX	18255428	mala
groupIX	18757824	ppp3r1b
groupIX	18940021	tbce
groupIX	18947442	nid1b
groupIX	19952391	psen2
groupV	257699	ADK
groupV	262824	ADK
groupV	1246806	lhpp
groupV	1564568	dock1
groupV	1990980	opn4b
groupV	3071211	sdk2b
groupV	3642949	WNK4
groupV	4393370	MYO1D
groupV	4771292	acsf2
groupV	4771292	chad

groupV	6028735	SPAG9
groupV	6039736	SPAG9
groupV	8333154	cpn1
groupV	8362617	dnmbp
groupV	8371507	dnmbp
groupV	8526046	ATRNL1
groupV	8788338	cep55l
groupV	8964836	pcdh15b
groupV	9676174	aldoaa
groupV	10260074	ntn2
groupV	10355684	im:7160594
groupV	10355685	im:7160594
groupV	10472878	ift140
groupV	11236388	trrap
groupV	11757092	lipf
groupV	12105748	arg1
groupVI	65315	pdzd8
groupVI	65320	pdzd8
groupVI	535902	piezo1
groupVI	628846	degs1
groupVI	1469002	STK32C
groupVI	2224653	btaf1
groupVI	2832952	sorbs1
groupVI	2977696	MYPN
groupVI	3054537	ATE1
groupVI	3315380	cdh23
groupVI	3532982	mlh1
groupVI	5214539	jag1b
groupVI	5561692	macrod2
groupVI	6171741	MCU
groupVI	7263134	adgrb3

groupVI	10695465	pcdh15a
groupVI	10720947	pcdh15a
groupVI	10720979	pcdh15a
groupVI	11196458	add3b
groupVI	11196522	add3b
groupVI	11972866	dachc
groupVI	15176816	prkceb
groupVI	15222130	epas1b
groupVI	16489290	ddx43
groupVI	16698814	entpd6
groupVI	16786362	ppp2r5d
groupVII	1533	cnpy3
groupVII	734163	hdac8
groupVII	1166395	exoc6b
groupVII	1527191	supt16h
groupVII	2069023	arhgef11
groupVII	2262912	zgc:55262
groupVII	2281161	pacs1a
groupVII	2927070	efs
groupVII	3225940	tbc1d19
groupVII	3475396	pcdh7b
groupVII	3506539	pcdh7b
groupVII	4111347	dvl2
groupVII	7590548	slc8a4b
groupVII	10022174	NRXN2
groupVII	12416367	wscd1b
groupVII	12839808	arhgap32b
groupVII	14033304	med13a
groupVII	15251866	akap10
groupVII	15655377	gabra3
groupVII	17132205	ar

groupVII	17751863	rb1
groupVII	18060796	dpf2
groupVII	18428275	grk1b
groupVII	19030918	asl
groupVII	19037675	asl
groupVII	19041464	asl
groupVII	20697034	map3k7cl
groupVII	21450620	slc47a1
groupVII	21450624	slc47a1
groupVII	21464005	slc47a1
groupVII	23608574	gabrg2
groupVII	24285841	zbtb16a
groupVII	26201343	gemin5
groupVII	26830639	doc2b
groupVII	27786221	chm
groupVIII	705499	eps15l1a
groupVIII	1366702	slc1a8a
groupVIII	1420692	podn
groupVIII	1439673	slc1a7b
groupVIII	1571263	si:ch211-247n2.1
groupVIII	2019099	lpp
groupVIII	2070270	lpp
groupVIII	2140385	aldh9a1a.1
groupVIII	3901055	prrx1b
groupVIII	5145150	rgs8
groupVIII	7185822	mpl
groupVIII	7185827	mpl
groupVIII	9584441	arid3a
groupVIII	13574967	acbd6
groupVIII	13585724	acbd6
groupVIII	14756664	shdb

groupVIII	14764400	shdb
groupVIII	14966625	trabd2b
groupVIII	15943909	ell
groupVIII	16221404	mbd3b
groupVIII	16677624	cbarpb
groupVIII	16677625	cbarpb
groupVIII	17043307	NCAN
groupVIII	17173890	hmha1b
groupVIII	17173893	hmha1b
groupVIII	17448751	dph5
groupVIII	17704983	atp1b1a
groupVIII	17715065	nme7
groupVIII	17953010	cyfip1
groupVIII	17959965	cyfip1
groupVIII	17966434	cyfip1
groupVIII	18218521	AP1S2
groupVIII	18228657	SHROOM2
groupVIII	18311767	nek1
groupVIII	18536876	scfd2
groupVIII	19036986	uap1
groupVIII	19206380	echdc2
groupVIII	19206380	zyg11
groupX	1031088	pou2f2a
groupX	1434269	rnf19b
groupX	2022424	xkr8.2
groupX	2236081	atp9b
groupX	2372157	pleca
groupX	3397057	med30
groupX	3946811	TGFBR2
groupX	6434188	AKAP9
groupX	6467748	krit1

groupX	6959237	efna1a
groupX	8346636	elmo1
groupX	8912520	calcr
groupX	9268603	col1a2
groupX	9313919	ppp1r9a
groupX	9559058	igf2bp3
groupX	10643412	col16a1
groupX	10651654	col16a1
groupX	11313196	gstr
groupX	11411476	arid1ab
groupX	12140845	MANEAL
groupX	12376549	top2b
groupX	13158201	ST3GAL1
groupX	13158217	ST3GAL1
groupX	13346262	KHDRBS3
groupX	15057778	vps52
groupXI	2784937	zgc:63569
groupXI	3037606	nsfa
groupXI	3181135	cpt1cb
groupXI	4127932	si:ch1073-322p19.1
groupXI	5242096	zgc:113411
groupXI	6319058	vat1
groupXI	8141369	dnm2a
groupXI	9138672	cant1a
groupXI	9863208	coro7
groupXI	9892926	coro7
groupXI	12085729	gtpbp1
groupXI	12097682	sgsm3
groupXI	12967327	sec14l1
groupXI	14638719	LLGL1

groupXI	14692631	drg2
groupXI	14862190	cacng3b
groupXI	14868809	cacng3b
groupXI	15669080	CACNG2
groupXI	15820278	med15
groupXI	15853154	btr30
groupXI	16239896	mfsd11
groupXII	173234	prpf6
groupXII	219262	nkain4
groupXII	778543	magi1a
groupXII	1064128	nr2c2
groupXII	1390432	adcy6a
groupXII	1582014	pfkfb2a
groupXII	1607297	scube3
groupXII	1661807	ANKS1A
groupXII	1665691	ANKS1A
groupXII	2054151	mitfb
groupXII	3965736	ntsr1
groupXII	4021290	slco4a1
groupXII	4137490	ralgapb
groupXII	4167543	prex1
groupXII	4167960	prex1
groupXII	4822717	plxna1a
groupXII	5159212	agrn
groupXII	6460320	dnajc16
groupXII	6796067	sema3h
groupXII	9692047	noc2l
groupXII	9836616	rerea
groupXII	13491489	hsd17b10
groupXII	13764043	cs
groupXII	14138697	plxna2

groupXII	14143778	plxna2
groupXII	14989561	acap3a
groupXII	15764340	ythdf1
groupXII	16832979	ephb2a
groupXII	16833056	ephb2a
groupXII	17501954	skia
groupXII	17502365	skia
groupXII	17531731	skia
groupXII	18301421	ccdc22
groupXIII	1000849	zgc:64051
groupXIII	1048433	gtf2h2
groupXIII	1951818	ndufs4
groupXIII	2619063	rad17
groupXIII	3217071	NSMF
groupXIII	4836805	DMXL1
groupXIII	4842212	DMXL1
groupXIII	5081419	loxhd1b
groupXIII	6211357	zswim6
groupXIII	6532396	smarcad1a
groupXIII	6532401	smarcad1a
groupXIII	6629991	grid2
groupXIII	7469484	ndc80
groupXIII	9245655	p2rx2
groupXIII	9245657	p2rx2
groupXIII	11053039	fam172a
groupXIII	11293846	mctp1a
groupXIII	12344504	zgc:77112
groupXIII	13798927	angptl2b
groupXIII	14334245	ncor2
groupXIII	14747990	MSI1
groupXIII	14892645	htr7c

groupXIII	15336090	chfr
groupXIII	16056038	hk2
groupXIII	16098803	mxd1
groupXIII	16372407	si:dkey-91m11.5
groupXIII	16374537	si:dkey-91m11.5
groupXIII	16554401	citb
groupXIII	17713078	col27a1b
groupXIII	18861948	cds1
groupXIII	19081696	adamts3
groupXIII	19113887	adamts3
groupXIII	19258058	dmrt1
groupXIII	19368639	atp8b5a
groupXIII	19389293	cnnm4b
groupXIV	191297	apc
groupXIV	670391	adamts6
groupXIV	902522	pde8b
groupXIV	1233560	pik3r1
groupXIV	1676000	si:dkey-226m8.10
groupXIV	2022655	col5a1
groupXIV	3450849	plpp1a
groupXIV	6336247	WDR70
groupXIV	7479984	rc3h2
groupXIV	7560515	strbp
groupXIV	7691338	cercam
groupXIV	7691412	cercam
groupXIV	7724003	dbh
groupXIV	7724067	dbh
groupXIV	7724068	dbh
groupXIV	7928407	cdc37l1
groupXIV	8230710	si:ch73-287m6.1

groupXIV	8317774	si:dkey-112m2.1
groupXIV	8627095	sfswap
groupXIV	8802409	mmp17a
groupXIV	9438357	GRIN3A
groupXIV	10128595	cntfr
groupXIV	11047129	man1b1a
groupXIV	11073386	whrna
groupXIV	11394032	tsc1b
groupXIV	11394034	tsc1b
groupXIV	13226808	WDR7
groupXIV	13256683	ccng2
groupXIV	13558515	dcc
groupXIV	13562065	dcc
groupXIV	13786838	fbxw2
groupXIV	14341827	slc25a25b
groupXIV	14530148	fpgs
groupXIV	14530148	cdk9
groupXIV	14656013	crata
groupXIX	155399	AGBL1
groupXIX	535966	zgc:158366
groupXIX	696909	prmt7
groupXIX	696910	prmt7
groupXIX	767590	smpd3
groupXIX	855929	gins3
groupXIX	1177562	btbd10a
groupXIX	1497498	ppp6r3
groupXIX	1604388	ces3
groupXIX	1605843	ces3
groupXIX	2101576	siah1
groupXIX	2111122	siah1
groupXIX	2123707	lonp2

groupXIX	2486385	myh7bb
groupXIX	2693940	necab2
groupXIX	2768433	mbtps1
groupXIX	2910599	ranbp10
groupXIX	2919345	ranbp10
groupXIX	3058100	klhdc4
groupXIX	3165547	fa2h
groupXIX	3696916	tead1a
groupXIX	3729185	parvaa
groupXIX	3729194	parvaa
groupXIX	3729222	parvaa
groupXIX	4601439	mon2
groupXIX	4603471	mon2
groupXIX	6187781	dhtkd1
groupXIX	6187938	dhtkd1
groupXIX	6936668	LRRC10B
groupXIX	6956001	syt7a
groupXIX	6956003	syt7a
groupXIX	6956013	syt7a
groupXIX	7341537	ano3
groupXIX	7375690	clpxb
groupXIX	7375691	clpxb
groupXIX	7375693	clpxb
groupXIX	7375707	clpxb
groupXIX	7432798	MYO5A
groupXIX	7582137	tjp1b
groupXIX	8714499	lamb1a
groupXIX	8763053	frs2b
groupXIX	8763087	frs2b
groupXIX	8860525	gpia
groupXIX	9325487	tnni2a.3

groupXIX	9338343	lsp1
groupXIX	9356436	tnnt3a
groupXIX	10225695	asz1
groupXIX	10237889	asz1
groupXIX	10844748	alx1
groupXIX	11450107	RELN
groupXIX	11467173	RELN
groupXIX	11656169	PDE3A
groupXIX	12158600	tln2b
groupXIX	12170304	tln2b
groupXIX	12396182	roraa
groupXIX	13231403	osbpl5
groupXIX	13355991	mrpl23
groupXIX	13420421	cpt1ab
groupXIX	13760083	ldha
groupXIX	13779758	hrasb
groupXIX	13783376	hrasb
groupXIX	13783774	hrasb
groupXIX	13916379	LDLRAD3
groupXIX	13916381	LDLRAD3
groupXIX	13916396	LDLRAD3
groupXIX	13916400	LDLRAD3
groupXIX	14035773	pamr1
groupXIX	14035877	pamr1
groupXIX	14037793	pamr1
groupXIX	14037805	pamr1
groupXIX	14679371	banp
groupXIX	14893013	rassf8a
groupXIX	14934058	itpr2
groupXIX	15004158	nr1h4
groupXIX	15235216	magi2b

groupXIX	15260020	magi2b
groupXIX	15348919	prr5a
groupXIX	15348921	prr5a
groupXIX	16083573	rbm28
groupXIX	16675756	lmf2a
groupXIX	16917923	IGF1R
groupXIX	16917928	IGF1R
groupXIX	18144486	hipk2
groupXIX	18154488	hipk2
groupXIX	18555879	tmtc2b
groupXIX	18566263	tmtc2b
groupXIX	18566268	tmtc2b
groupXIX	18669924	bicd1a
groupXIX	18676323	bicd1a
groupXIX	18808896	ada2a
groupXIX	18979705	zgc:173742
groupXIX	18981760	zgc:173742
groupXIX	19002277	mlycd
groupXIX	19063381	cdh13
groupXIX	19529735	e2f4
groupXIX	19532706	e2f4
groupXIX	19799949	cbfb
groupXIX	19799953	cbfb
groupXIX	19868929	GNAO1
groupXV	420963	slc4a11
groupXV	623475	fgfr3
groupXV	878013	mark3a
groupXV	1006490	BRF1
groupXV	1634088	galnt16
groupXV	2227371	RGS6
groupXV	2472773	FMN1

groupXV	2715782	akt1
groupXV	3071522	ZDHHC14
groupXV	3660135	SCAF8
groupXV	4347319	ldlrap1b
groupXV	4347322	ldlrap1b
groupXV	5993917	ltk
groupXV	6042640	si:dkey-13p1.4
groupXV	6245246	kidins220a
groupXV	6929983	znf292a
groupXV	8210359	crlf1a
groupXV	10439279	snap23.1
groupXV	10856058	slc25a21
groupXV	11927596	pacs2
groupXV	11964154	actn1
groupXV	14239298	babam2
groupXV	15455501	setd3
groupXV	16025397	kcnk5a
groupXVI	3648993	sestd1
groupXVI	4150750	RAMP1
groupXVI	6455386	slc40a1
groupXVI	6455392	slc40a1
groupXVI	6455500	slc40a1
groupXVI	6711719	cmss1
groupXVI	7012607	rcbtb1
groupXVI	7734394	appb
groupXVI	8201834	GSK3B
groupXVI	8543530	nqo1
groupXVI	9578110	ADAM23
groupXVI	11786779	trpm2
groupXVI	12645970	erbb4b
groupXVI	14401010	stxbp5l

groupXVI	15927404	lrch1
groupXVI	15927437	lrch1
groupXVI	15994142	sh3bp4
groupXVI	16098029	caska
groupXVI	16558988	abcb6b
groupXVI	17584426	plcl1
groupXVII	1379422	ATP6AP1L
groupXVII	1677561	atp2b2
groupXVII	2321752	ephb2b
groupXVII	3617667	si:dkey-238f9.1
groupXVII	6303657	znf362a
groupXVII	6999690	hdac7b
groupXVII	10053721	cdh4
groupXVII	10420063	zgc:92107
groupXVII	10447506	ndrg3a
groupXVII	10987464	phactr3b
groupXVII	11679213	cfap57
groupXVII	11917198	fbln2
groupXVII	12211780	CPNE9
groupXVII	12227031	mtmr14
groupXVII	12620833	cyb561d1
groupXVII	12620834	cyb561d1
groupXVII	13731462	pdzrn3b
groupXVIII	724190	cd2ap
groupXVIII	747441	tnfrsf21
groupXVIII	1298178	cdc42bpab
groupXVIII	2016697	calm1b
groupXVIII	2317408	nrxn3b
groupXVIII	2334678	nrxn3b
groupXVIII	2955853	cgrrf1
groupXVIII	3355320	nt5dc1

groupXVIII	4709916	cdk19
groupXVIII	5385454	tbc1d32
groupXVIII	9208673	mrap2b
groupXVIII	9398199	matn3a
groupXVIII	10169758	fdft1
groupXVIII	10708381	snx17
groupXVIII	10965226	eya4
groupXVIII	10965334	eya4
groupXVIII	11368094	nbas
groupXVIII	13449171	si:ch1073-155h21.1
groupXVIII	13650635	runx2b
groupXVIII	13768203	fkbp1b
groupXVIII	14165771	PLCB1
groupXVIII	15030531	pigh
groupXVIII	15431671	rnf217
groupXVIII	16173901	mthfd1a
groupXX	83790	dennd4b
groupXX	174683	pik3r4
groupXX	399080	IQGAP3
groupXX	563010	myo1eb
groupXX	2009335	ano10a
groupXX	2978274	RIMS2
groupXX	6277129	cobl
groupXX	6934056	nrsn1
groupXX	7470424	cdkal1
groupXX	7484682	cdkal1
groupXX	7606880	COLEC10
groupXX	7606890	COLEC10
groupXX	7686849	CARMIL1
groupXX	7686851	CARMIL1

groupXX	7933731	col14a1a
groupXX	8793862	si:ch73-380l3.2
groupXX	8814434	si:ch73-380l3.2
groupXX	10265744	steap4
groupXX	10518699	tmem245
groupXX	10909957	pafah1b3
groupXX	11117793	cicb
groupXX	11608973	lpcat3
groupXX	11723404	zgc:92912
groupXX	12088534	foxj2
groupXX	12819596	crabp2a
groupXX	14703661	VWDE
groupXX	16834661	MAN1C1
groupXX	17207593	fhl3a
groupXX	17772190	smpdl3b
groupXX	18664829	ascc3
groupXX	18664871	ascc3
groupXX	18939655	cdk14
groupXXI	1068422	PARD3
groupXXI	1323162	ccny
groupXXI	1861396	adarb2
groupXXI	2183125	abca4a
groupXXI	2308001	col11a1a
groupXXI	4929720	zfhx4
groupXXI	5928041	pola1
groupXXI	5928082	pola1
groupXXI	6849448	eya1
groupXXI	9058787	lmbr1
groupXXI	9259242	esyt2b
groupXXI	9485833	COLEC12
groupXXI	9961350	asap1b

groupXXI	10093284	puf60b
groupXXI	10292686	gli3
groupXXI	10294135	gli3

Table D3. qGLM outliers (FDR = 0.01%) mapped to genes.

Chromosome	Position	Overlapping gene
groupI	194284	ANO7
groupI	194297	ANO7
groupI	1356417	myo1cb
groupI	2079573	numbl
groupI	2079695	numbl
groupI	2225424	CEP164
groupI	2225425	CEP164
groupI	2294609	MPZL2
groupI	2921889	si:ch73-265h17.5
groupI	4359566	dscamb
groupI	4486100	RPS6KB1
groupI	6503833	DYNC2H1
groupI	6652237	gria4a
groupI	8623856	sik2b
groupI	9151616	slc1a5
groupI	11028941	ROBO3
groupI	11997404	pitpnm3
groupI	12205295	synrg

groupI	14268117	cog6
groupI	16878384	aplp2
groupI	17023775	MAP4K1
groupI	18834669	bcas3
groupI	19552437	akap1b
groupI	19761643	PIPOX
groupI	19761644	PIPOX
groupI	19777379	PIPOX
groupI	20058982	ical1
groupI	20075909	bmpr2a
groupI	22857862	kpnb3
groupI	24615905	ARHGAP6
groupI	24647536	ARHGAP6
groupI	26035523	pudp
groupI	26123756	nlgn4
groupI	27198465	gucy1b2
groupI	27795801	map3k2
groupI	27892870	ahr1b
groupII	64906	cemip
groupII	72713	cemip
groupII	1386127	znf536
groupII	1895302	znf507
groupII	1976574	si:ch211-186j3.6
groupII	2069386	si:ch211-186j3.6
groupII	2482071	adgrg1
groupII	3096000	adamts18
groupII	3386926	zgc:153993
groupII	3386930	zgc:153993
groupII	3810111	ccdc102a
groupII	4307052	CSK
groupII	5507405	itfg1

groupII	5507408	itfg1
groupII	5646182	wtip
groupII	6072158	ace2
groupII	7405938	pcdh17
groupII	7845640	GPC5
groupII	9205254	sord
groupII	9352682	dagla
groupII	9360871	dagla
groupII	9464226	plekha7b
groupII	9484845	plekha7b
groupII	9485992	plekha7b
groupII	11339059	aldh1a2
groupII	11413867	cgnl1
groupII	11486721	znf280d
groupII	11486724	znf280d
groupII	12887822	elp4
groupII	13091572	CSMD1
groupII	14027788	fzd3b
groupII	14759789	mdga1
groupII	16360482	ugt5f1
groupII	18426961	SECISBP2L
groupII	19317537	gramd2aa
groupII	19529077	neo1b
groupII	21905813	otog
groupII	22473755	PLEKHG4
groupII	22516645	fhod1
groupII	22640762	wdr59
groupIII	109352	col6a1
groupIII	532867	PHACTR1
groupIII	1339603	dpp6b
groupIII	1349077	arhgap21b

groupIII	1357304	arhgap21b
groupIII	2025815	eif2b3
groupIII	2028072	eif2b3
groupIII	2325889	gpc5c
groupIII	3187533	carmil3
groupIII	3245336	rem2
groupIII	4944608	crispld1b
groupIII	7401013	onecut3b
groupIII	8295261	upf1
groupIII	8875390	rpap2
groupIII	8876148	rpap2
groupIII	9580896	MIER1
groupIII	10334379	abi1b
groupIII	11368126	PTPRM
groupIII	11395637	PTPRM
groupIII	11742815	vav3b
groupIII	12115738	PTPRF
groupIII	12115741	PTPRF
groupIII	12458632	prkci
groupIII	12527450	samd7
groupIII	12735410	cacnb2b
groupIII	12915935	lpin2
groupIII	13274093	b4galt2
groupIII	13585312	mib1
groupIII	13881945	pth1ra
groupIII	15164917	kcnj9
groupIII	15202767	lig1
groupIII	16710805	cdk5
groupIII	16767818	retreg1
groupIII	16767819	retreg1
groupIII	16768113	retreg1

groupIV	215897	tgfbi
groupIV	1019927	gpc2
groupIV	2243320	pdgfc
groupIV	2601190	FSTL5
groupIV	2620952	FSTL5
groupIV	2725245	ldb2a
groupIV	3344338	SPATA5
groupIV	3518208	ANKRD50
groupIV	3897816	PCGF3
groupIV	4522355	glra1
groupIV	4950162	mgat4b
groupIV	4950163	mgat4b
groupIV	5185177	fermt3b
groupIV	5567545	slc25a43
groupIV	5872589	mcf2a
groupIV	5875137	mcf2a
groupIV	6162198	arhgef6
groupIV	7962155	BEND4
groupIV	8529654	si:ch211-220e11.3
groupIV	11356900	RASGEF1C
groupIV	11363916	RASGEF1C
groupIV	12568372	nlgn3a
groupIV	14910821	TENM2
groupIV	14931022	TENM2
groupIV	14931034	TENM2
groupIV	15087182	stag2b
groupIV	15552599	si:ch211-26b3.4
groupIV	16755575	si:ch211-159i8.4
groupIV	18294419	grm8a
groupIV	22466149	ptprz1b

groupIV	22904945	mkrn1
groupIV	22927334	tmem178b
groupIV	23122945	tmtc1
groupIV	23122946	tmtc1
groupIV	23837305	hspa14
groupIV	24380863	mkln1
groupIV	25592793	ANO4
groupIV	28487084	wnt5b
groupIV	28514992	wnt5b
groupIV	28607445	ERC1
groupIV	28698922	wnk1b
groupIV	29365747	chchd3a
groupIV	29383782	chchd3a
groupIV	29713181	iqsec3a
groupIV	30041546	cacna2d1a
groupIV	30161642	snd1
groupIV	31019585	atp6v1e1b
groupIV	31163979	ptn
groupIV	32039274	btbd11a
groupIV	32124668	parpbp
groupIV	32379845	acot16
groupIX	1233489	pcdh7a
groupIX	1519539	pax5
groupIX	2371846	sh3pxd2aa
groupIX	4059070	kdm2aa
groupIX	5665301	fat1a
groupIX	5680724	fat1a
groupIX	5957375	tll1
groupIX	6033487	prom1b
groupIX	6789945	pou4f2
groupIX	7312774	rnf150a

groupIX	8002116	lef1
groupIX	8395482	mxd4
groupIX	9189845	ctnna2
groupIX	10822865	wbp1lb
groupIX	11188437	dok7
groupIX	12341680	NAA15
groupIX	12861345	prkca
groupIX	12938477	helz
groupIX	15454806	gcdha
groupIX	15456561	gcdha
groupIX	15578032	C16orf45
groupIX	15892809	ccdc85al
groupIX	17068902	gga3
groupIX	18849813	strn
groupIX	19216545	si:ch211-133l11.10
groupIX	19233714	USP34
groupIX	19520170	myom2a
groupIX	19895755	gpr108
groupIX	20141599	zgc:109889
groupV	249160	ADK
groupV	1028846	CDHR1
groupV	1657571	ptprea
groupV	1873318	mta3
groupV	2549872	prph2a
groupV	3530315	TANC2
groupV	3652480	WNK4
groupV	3652485	WNK4
groupV	3652488	WNK4
groupV	3985769	tmem106a
groupV	3987676	tmem106a

groupV	4610493	hoxb5b
groupV	4610517	hoxb5b
groupV	4610526	hoxb5b
groupV	4610538	hoxb5b
groupV	4610541	hoxb5b
groupV	4610543	hoxb5b
groupV	4610559	hoxb5b
groupV	4610561	hoxb5b
groupV	4610562	hoxb5b
groupV	4610573	hoxb5b
groupV	4661278	fbrs
groupV	6480734	cacna1g
groupV	6783702	SLC39A11
groupV	6829574	usp22
groupV	8172811	DNAJC7
groupV	8257791	VAT1
groupV	8931495	pcdh15b
groupV	10375451	MPRIP
groupV	11239174	trrap
groupVI	360048	cdh30
groupVI	362892	cdh30
groupVI	383491	cdh30
groupVI	674490	polr1c
groupVI	674493	polr1c
groupVI	2817311	sorbs1
groupVI	2939368	bag3
groupVI	2970516	MYPN
groupVI	3188922	wdr11
groupVI	3776332	fbxw4
groupVI	4182520	b3gat2
groupVI	4343168	PTK7

groupVI	4725102	sipa1l2
groupVI	7082913	fam135a
groupVI	7088860	fam135a
groupVI	7088868	fam135a
groupVI	7534195	khdrbs2
groupVI	7853672	gclc
groupVI	8201351	WDFY4
groupVI	10059688	kcnma1a
groupVI	10125565	kcnma1a
groupVI	11042169	neurl1ab
groupVI	11042202	neurl1ab
groupVI	11042204	neurl1ab
groupVI	11573676	adam12
groupVI	12204800	bicc1a
groupVI	12209702	bicc1a
groupVI	12212527	bicc1a
groupVI	13821175	CCDC88A
groupVI	13944537	kcnh1a
groupVI	13976992	cyp3c3
groupVI	14016829	USP54
groupVI	14121472	kif20ba
groupVI	14490864	SPTBN1
groupVI	15543216	nckap1
groupVI	15684936	napba
groupVI	15685231	napba
groupVI	15725087	mcm8
groupVI	16729856	ninl
groupVI	16887743	ubr2
groupVII	170884	tec
groupVII	1152920	exoc6b
groupVII	1633750	si:dkey-46i9.6

groupVII	1633751	si:dkey-46i9.6
groupVII	1853284	ltb4r2a
groupVII	2060777	arhgef11
groupVII	2070514	arhgef11
groupVII	2264587	zgc:55262
groupVII	2491155	intu
groupVII	3115871	chd3
groupVII	7279206	pcxb
groupVII	8915527	fgf11a
groupVII	9372934	ndufs2
groupVII	10491101	cblb
groupVII	10827138	picalma
groupVII	10870723	dlg2
groupVII	10939050	dlg2
groupVII	10985051	dlg2
groupVII	10987012	dlg2
groupVII	10995435	dlg2
groupVII	10995450	dlg2
groupVII	11022315	dlg2
groupVII	11496027	ntm
groupVII	13128851	dscama
groupVII	15498647	slc8a4a
groupVII	16924111	fgf11b
groupVII	17068916	TBC1D8B
groupVII	18010043	slc25a35
groupVII	19037705	asl
groupVII	19046228	asl
groupVII	20101469	nbeaa
groupVII	20645749	tiam1a
groupVII	20655065	tiam1a
groupVII	22692904	lrch2

groupVII	24337667	zbtb16a
groupVII	24480979	ncam1a
groupVII	25566233	wwc1
groupVII	26822186	doc2b
groupVII	26867183	rcc1l
groupVII	27872941	tmprss15
groupVII	27872950	tmprss15
groupVII	27872952	tmprss15
groupVIII	1029214	nfia
groupVIII	1067947	nfia
groupVIII	1255699	sgta
groupVIII	1261461	SLC39A3
groupVIII	1444901	slc1a7b
groupVIII	1914639	si:ch73-194h10.2
groupVIII	2065096	lpp
groupVIII	2119853	magoh
groupVIII	2498771	tgfbr3
groupVIII	3513731	irf4b
groupVIII	3717572	GMDS
groupVIII	3729233	GMDS
groupVIII	4509668	gpx7
groupVIII	4570277	atpaf1
groupVIII	6561046	meltf
groupVIII	6684784	slc25a24
groupVIII	7946080	raver2
groupVIII	10306874	ptprsa
groupVIII	10310570	ptprsa
groupVIII	10550808	nr5a2
groupVIII	11785209	ankrd13c
groupVIII	11795426	cthl

groupVIII	11925305	negr1
groupVIII	12630691	actl6a
groupVIII	12866299	si:dkey-30h22.11
groupVIII	12866318	si:dkey-30h22.11
groupVIII	13723230	cacna1ea
groupVIII	14636941	tle2b
groupVIII	15002573	si:dkey-97o5.1
groupVIII	15101479	olfml2ba
groupVIII	15101480	olfml2ba
groupVIII	15101511	olfml2ba
groupVIII	15136378	GHRHR
groupVIII	16815001	si:ch211-212d10.2
groupVIII	16933459	tmem161a
groupVIII	17159656	tjp3
groupVIII	17449991	dph5
groupVIII	17820046	gabrb3
groupVIII	17962542	cyfip1
groupVIII	18044970	tbc1d23
groupVIII	18497497	sgcb
groupVIII	18542341	scfd2
groupVIII	18970188	fsd1
groupVIII	19045555	ddr2a
groupVIII	19062127	serbp1a
groupVIII	19218952	ralgps2
groupVIII	19247716	ralgps2
groupX	1258350	bmp8a
groupX	1431303	rnf19b
groupX	1632737	DPY19L4

groupX	1635764	DPY19L4
groupX	1872357	hivep1
groupX	1872395	hivep1
groupX	2096192	slc9a3.2
groupX	2651577	tfap2e
groupX	3893681	STT3B
groupX	4885550	mylipa
groupX	5119835	bloc1s4
groupX	7976349	mterf3
groupX	8921684	calcr
groupX	9043725	cdk6
groupX	9136826	TMEM196
groupX	9222085	sp4
groupX	9263082	col1a2
groupX	9705450	rftn1a
groupX	11348351	fhod3a
groupX	11365416	fhod3a
groupX	13560411	TRAPPC9
groupX	14515807	ccdc106b
groupX	14675401	rnf139
groupX	15588482	bop1
groupXI	44440	pdxdc1
groupXI	44441	pdxdc1
groupXI	299682	mettl22
groupXI	437592	prrg2
groupXI	764525	nags
groupXI	1382735	carm1
groupXI	1387661	carm1
groupXI	2175473	PSME3
groupXI	2313798	caskin1
groupXI	3571147	ntn1b

groupXI	3974501	GRIN2B
groupXI	5013362	TST
groupXI	6507868	st8sia6
groupXI	8663375	ttyh2l
groupXI	8717546	aatka
groupXI	8717548	aatka
groupXI	8726277	aatka
groupXI	10003166	ADCY9
groupXI	10098607	prkar1b
groupXI	10817841	arhgap17b
groupXI	10935631	noxo1a
groupXI	10966681	srebf1
groupXI	11449609	09-Sep
groupXI	11519770	tnrc6c1
groupXI	12093819	gtpbp1
groupXI	12093831	gtpbp1
groupXI	12656280	mapk8ip3
groupXI	14254415	si:dkey-43p13.5
groupXI	15562330	fam83fa
groupXI	15612231	si:ch73-233k15.2
groupXI	15612252	si:ch73-233k15.2
groupXI	15612253	si:ch73-233k15.2
groupXI	15612265	si:ch73-233k15.2
groupXI	15821746	med15
groupXI	16487641	tmem104
groupXII	1594422	scube3
groupXII	1594424	scube3
groupXII	2621029	tm9sf4

groupXII	2621171	tm9sf4
groupXII	3562883	usp19
groupXII	3562888	usp19
groupXII	4060908	ampd1
groupXII	4060911	ampd1
groupXII	4306905	VSTM2L
groupXII	5661366	mib2
groupXII	5661843	mib2
groupXII	5839118	epha2b
groupXII	5846364	epha2b
groupXII	5846559	epha2b
groupXII	7018377	celsr3
groupXII	7100749	mgll
groupXII	7471705	CACNA1D
groupXII	7511555	nsun5
groupXII	8433707	KAZN
groupXII	8835052	phactr3a
groupXII	9996128	GSTM4
groupXII	11523684	rarga
groupXII	13531664	slc16a1b
groupXII	14077868	plxna2
groupXII	14154789	plxna2
groupXII	14268544	magi3a
groupXII	15489921	pip4k2ca
groupXII	15823000	opn7d
groupXII	15935495	magixb
groupXII	18272380	rtel1
groupXIII	302905	nelfb
groupXIII	1762349	hic2
groupXIII	2959586	VPS37B
groupXIII	3787612	zmat4a

groupXIII	4379022	msh3
groupXIII	4549297	lhfpl2b
groupXIII	7190703	frmd3
groupXIII	8461578	ebf2
groupXIII	10266430	cox7c
groupXIII	11442150	adgrv1
groupXIII	12071325	bmp1a
groupXIII	12116272	gfpt1
groupXIII	12139588	aak1a
groupXIII	12392757	tcf7l1a
groupXIII	12863489	rph3ab
groupXIII	13129213	trpm6
groupXIII	13182723	grk5l
groupXIII	13185867	grk5l
groupXIII	13185885	grk5l
groupXIII	13185886	grk5l
groupXIII	14354401	ncor2
groupXIII	14850519	kcnv2b
groupXIII	16347129	specc1la
groupXIII	16372968	si:dkey-91m11.5
groupXIII	17236935	jak2b
groupXIII	17405512	poli
groupXIII	17568669	notch1b
groupXIII	17736887	col27a1b
groupXIII	19303002	fbp2
groupXIII	19642070	coq5
groupXIV	549351	ZSWIM6
groupXIV	808642	col4a3bpb
groupXIV	916126	wdr41
groupXIV	1031933	ARSB
groupXIV	1405959	fcho2

groupXIV	2538572	aif1l
groupXIV	2732354	kyat1
groupXIV	2780479	NTNG2
groupXIV	4411253	TMEM38B
groupXIV	4960232	ASTN2
groupXIV	7011755	KCNN2
groupXIV	7375043	morn5
groupXIV	7925226	cdc37l1
groupXIV	8130524	bcr
groupXIV	8307713	si:dkey-112m2.1
groupXIV	8341222	si:dkey-112m2.1
groupXIV	8612841	sfswap
groupXIV	9394041	plppr1
groupXIV	9394042	plppr1
groupXIV	9403163	plppr1
groupXIV	10153615	rab27b
groupXIV	13361141	spina
groupXIV	13494921	dcc
groupXIV	13541022	dcc
groupXIV	13802010	CRB2
groupXIV	14608983	ntmt1
groupXIV	15077079	auh
groupXIX	279445	ntrk3a
groupXIX	309244	ntrk3a
groupXIX	315383	ntrk3a
groupXIX	815471	ndrg4
groupXIX	1002818	cry2
groupXIX	1002823	cry2
groupXIX	1002824	cry2
groupXIX	1002836	cry2
groupXIX	1149661	galnt18a

groupXIX	1240432	PPFIBP2
groupXIX	1284472	far1
groupXIX	1485001	ric8a
groupXIX	1580409	zgc:91860
groupXIX	1607502	ces3
groupXIX	2123911	lonp2
groupXIX	2216564	st3gal2
groupXIX	2259685	hist1h4l
groupXIX	3055852	klhdc4
groupXIX	3343656	kiaa1549la
groupXIX	5448302	nell2a
groupXIX	6275171	usp3
groupXIX	6306183	commd4
groupXIX	6646006	smad3b
groupXIX	7581386	tjp1b
groupXIX	7799647	INSC
groupXIX	8132582	hsd17b12a
groupXIX	8586228	ptpn9a
groupXIX	10488868	si:dkey-106n21.1
groupXIX	10919586	lrrk2
groupXIX	11256691	tspan18b
groupXIX	11744277	PRKAR2B
groupXIX	12704045	si:ch211-13k12.2
groupXIX	12704047	si:ch211-13k12.2
groupXIX	13112677	dusp8a
groupXIX	13808858	rassf7a
groupXIX	15891442	shank3a
groupXIX	17610748	uri1

groupXIX	17610751	uri1
groupXIX	17610753	uri1
groupXIX	17610757	uri1
groupXIX	18253089	mical3a
groupXIX	18568168	tmtc2b
groupXIX	18720711	nup205
groupXIX	18907475	kcng4a
groupXV	141476	arid4a
groupXV	187245	BAZ1A
groupXV	352022	atrn
groupXV	361817	atrn
groupXV	391375	eif2ak3
groupXV	923132	klc1a
groupXV	986791	BRF1
groupXV	994919	BRF1
groupXV	1329085	mrps26
groupXV	1397019	slc39a8
groupXV	2369927	RYR3
groupXV	2915763	rbbp4
groupXV	4126558	pabpc4
groupXV	5506022	hspa4l
groupXV	5882707	eml1
groupXV	6034997	rpap1
groupXV	7048473	si:ch211-15d5.11
groupXV	7048495	si:ch211-15d5.11
groupXV	8912484	fermt2
groupXV	8952097	STXBP6
groupXV	10021510	fndc3ba
groupXV	10561857	elmsan1b

groupXV	10874243	slc25a21
groupXV	11634105	sos2
groupXV	11928511	pacs2
groupXV	12056687	slc8a3
groupXV	12570065	lrfn5a
groupXV	15509173	gabrb1
groupXV	16002575	iars2
groupXVI	147492	slc37a1
groupXVI	147494	slc37a1
groupXVI	195347	pde9a
groupXVI	1629241	clasp1a
groupXVI	2474886	mycbp2
groupXVI	3030302	nalcn
groupXVI	4490055	asb1
groupXVI	4549159	hdac4
groupXVI	4565118	hdac4
groupXVI	5755922	gulp1a
groupXVI	6394706	glsb
groupXVI	6583007	IMPG2
groupXVI	6621698	abi3bpa
groupXVI	6622524	abi3bpa
groupXVI	6645406	tfg
groupXVI	7726845	appb
groupXVI	9603361	prkra
groupXVI	10223788	SLC25A12
groupXVI	10554458	stk39
groupXVI	11685983	nme8
groupXVI	11686070	nme8
groupXVI	11696630	SMARCAL1
groupXVI	11739083	pfkla
groupXVI	11741264	pfkla

groupXVI	11799716	trpm2
groupXVI	12337419	vwc2l
groupXVI	13042576	SH3RF3
groupXVI	15761756	ABI2
groupXVI	16555594	abcb6b
groupXVI	16624511	kalrna
groupXVI	16910614	parp4
groupXVI	17292751	itgb5
groupXVI	17347979	dgkg
groupXVI	17462557	pikfyve
groupXVI	17685457	rab5b
groupXVI	18091809	mao
groupXVII	486408	ctsa
groupXVII	558224	mkrn2
groupXVII	833691	dnase1l4.1
groupXVII	1664612	atp2b2
groupXVII	1978579	arhgef10la
groupXVII	2999932	pacsin1b
groupXVII	3506319	slc5a8l
groupXVII	4011290	ASIC1
groupXVII	4041973	ASIC1
groupXVII	4183577	eya2
groupXVII	4366611	cdh22
groupXVII	4759550	st7l
groupXVII	5555572	MAP3K12
groupXVII	6549208	cpne5b
groupXVII	7158503	krt4
groupXVII	10275514	dido1
groupXVII	10413238	zgc:92107
groupXVII	11027907	txnrd3
groupXVII	11255412	etnk2

groupXVII	11308336	klhdc8a
groupXVII	11955141	wnt7aa
groupXVII	13238989	EPB41L1
groupXVII	14069629	mitfa
groupXVII	14215418	itpr1b
groupXVII	14249407	arl8bb
groupXVIII	293689	lbr
groupXVIII	496239	cnih3
groupXVIII	631355	ptchd4
groupXVIII	639121	ptchd4
groupXVIII	2332942	nrxn3b
groupXVIII	2394844	nrxn3b
groupXVIII	2544694	cyp46a1.4
groupXVIII	3033960	mcm9
groupXVIII	3446131	mdn1
groupXVIII	4499462	si:dkey-119m7.4
groupXVIII	4499526	si:dkey-119m7.4
groupXVIII	7671230	slc24a4b
groupXVIII	8566642	afg1lb
groupXVIII	8579462	afg1lb
groupXVIII	9543664	rhag
groupXVIII	9763683	slc35f6
groupXVIII	10240054	daam1b
groupXVIII	10241528	daam1b
groupXVIII	10843673	pinx1
groupXVIII	11889168	ust
groupXVIII	12136724	stxbp5a
groupXVIII	12179367	rab32a
groupXVIII	12721986	myt1la
groupXVIII	13771632	wdcp
groupXX	85188	dennd4b

groupXX	172527	pik3r4
groupXX	633589	tmem145
groupXX	1984914	snrka
groupXX	2463673	spire1b
groupXX	2557791	ubr5
groupXX	2614088	oxr1b
groupXX	2705937	si:dkey-122a22.2
groupXX	3669526	gdf6a
groupXX	3669532	gdf6a
groupXX	3669614	gdf6a
groupXX	4980646	derl1
groupXX	5025141	atad2
groupXX	5025181	atad2
groupXX	5036765	psma2
groupXX	6155595	RBMS3
groupXX	6965505	nrsn1
groupXX	7033062	nrsn1
groupXX	7048536	nrsn1
groupXX	8808003	si:ch73-380l3.2
groupXX	10712460	rad54b
groupXX	10824480	fam171a1
groupXX	12961899	kcnn3
groupXX	13183746	flad1
groupXX	14703868	VWDE
groupXX	15626825	smg5
groupXX	15626826	smg5
groupXX	15626830	smg5
groupXX	15626833	smg5
groupXX	16270708	onecutl
groupXX	17274290	si:ch211-1i11.3

groupXX	18131353	tgfb2
groupXXI	1091376	PARD3
groupXXI	3762843	fen1
groupXXI	3766074	fen1
groupXXI	4579413	dap
groupXXI	5686739	mllt10
groupXXI	5966696	pcyt1ba
groupXXI	6708043	KCNB2
groupXXI	7549622	pan3
groupXXI	7879936	vipr1b
groupXXI	8687826	acad11
groupXXI	8829563	dlgap1b
groupXXI	8928532	insig1
groupXXI	9261362	esyt2b
groupXXI	9261363	esyt2b
groupXXI	9261365	esyt2b
groupXXI	9261366	esyt2b
groupXXI	10219841	DNAJC13
groupXXI	11076874	gpr158a

Table D4. F_{ST}-qGLM outliers that map to protein-coding genes.

Chromosome	Position	Overlapping gene
groupI	1356415	myo1cb
groupI	1356417	myo1cb

groupI	1356419	myo1cb
groupI	2225425	CEP164
groupI	2250676	sik3
groupI	11997404	pitpnm3
groupI	15136838	nectin1b
groupI	18856683	bcas3
groupI	20075909	bmpr2a
groupI	22824368	cacnb4b
groupI	23628696	GABRG3
groupI	27700431	cybb
groupII	267489	aldh1a3
groupII	2023253	si:ch211-186j3.6
groupII	5046745	kif13ba
groupII	6086839	pir
groupII	7214728	diaph3
groupII	9373814	ext2
groupII	9489069	plekha7b
groupII	10284123	galnt18b
groupII	10284126	galnt18b
groupII	11413867	cgnl1
groupII	13419593	supv3l1
groupII	19363379	MYO9A
groupII	19373274	MYO9A
groupII	19529077	neo1b
groupII	21682346	si:ch211-112g6.4
groupIII	4120217	ctnnd2b
groupIII	5202005	aglb
groupIII	5593347	sdr16c5a
groupIII	6977810	wdr47a

groupIII	12400167	insra
groupIII	16578504	AGAP3
groupIV	2063943	si:dkey-237h12.3
groupIV	2945152	LIMCH1
groupIV	3525588	ANKRD50
groupIV	7963436	BEND4
groupIV	18340744	grm8a
groupIV	29377073	chchd3a
groupIV	30675359	FUT9
groupIV	30675360	FUT9
groupIV	32062097	btbd11a
groupIV	32330935	ERI1
groupIX	170530	adgrl3.1
groupIX	2282295	SLC24A3
groupIX	5374892	trim2a
groupIX	6047838	anxa5b
groupIX	6047839	anxa5b
groupIX	6789945	pou4f2
groupIX	7554047	sh3gl2a
groupIX	18849813	strn
groupIX	20141599	zgc:109889
groupV	3530315	TANC2
groupV	5771266	rangap1b
groupV	6301445	fdxr
groupVI	65315	pdzd8
groupVI	4144174	smap1
groupVI	4146999	smap1
groupVI	4147000	smap1
groupVI	5854986	rrbp1a
groupVI	10098698	kcnma1a

groupVI	10512756	polh
groupVI	10695465	pcdh15a
groupVI	14491371	SPTBN1
groupVI	16698814	entpd6
groupVI	16724969	ninl
groupVII	734163	hdac8
groupVII	2070514	arhgef11
groupVII	2281161	pacs1a
groupVII	2297755	pacs1a
groupVII	10336150	senp3b
groupVII	14493165	p2rx1
groupVII	16078035	MSI2
groupVII	18430464	grk1b
groupVII	21446485	slc47a1
groupVII	21450557	slc47a1
groupVII	21464005	slc47a1
groupVII	26806302	doc2b
groupVII	27094361	pou2f3
groupVII	27182881	slc22a5
groupVIII	1025715	nfia
groupVIII	1235981	nwd1
groupVIII	1252454	thop1
groupVIII	1261437	SLC39A3
groupVIII	1261461	SLC39A3
groupVIII	1446795	slc1a7b
groupVIII	1914639	si:ch73-194h10.2
groupVIII	6344609	mccc1
groupVIII	7117808	znhit6
groupVIII	10330503	tbxa2r
groupVIII	13898442	faf1

groupVIII	14408125	lpl
groupVIII	17323822	grin3bb
groupVIII	17357959	med16
groupVIII	17357961	med16
groupVIII	17820046	gabrb3
groupVIII	18066094	sft2d2a
groupVIII	18605846	lnx1
groupX	1265572	MACF1
groupX	6434251	AKAP9
groupX	7428667	KCNQ4
groupX	7796551	spire1a
groupX	9712971	rftn1a
groupX	13158201	ST3GAL1
groupX	13612904	TRAPPC9
groupX	14529968	cnot3b
groupX	15057778	vps52
groupXI	4768115	smcr8a
groupXI	9089704	cyth1a
groupXI	9093791	cyth1a
groupXI	11460524	09-Sep
groupXI	13038226	akt1s1
groupXI	14802584	tnrc6a
groupXI	14903507	prkcbb
groupXII	334413	dnah12
groupXII	1406570	g6pd
groupXII	1439605	asic1a
groupXII	7474444	CACNA1D
groupXII	13564965	lima1a
groupXII	15674173	osbpl2a
groupXIII	277188	ehmt1a
groupXIII	6275830	elovl7a

groupXIII	8091956	pcsk5b
groupXIII	8583186	tcn2
groupXIII	13395404	DTX1
groupXIII	13888895	zgc:154046
groupXIII	18809089	wdfy3
groupXIII	19100744	adamts3
groupXIII	19368639	atp8b5a
groupXIV	916126	wdr41
groupXIV	7724003	dbh
groupXIV	7925224	cdc37l1
groupXIV	7925227	cdc37l1
groupXIV	12051392	rxraa
groupXIV	14709730	YTHDC2
groupXIV	14854047	slc38a9
groupXIX	702930	exoc3l1
groupXIX	758493	smpd3
groupXIX	1259055	kdelc2
groupXIX	1481801	ric8a
groupXIX	2822284	nrn1la
groupXIX	6006011	zgc:136858
groupXIX	6773313	tle3b
groupXIX	7058825	pgghg
groupXIX	7432798	MYO5A
groupXIX	7781050	INSC
groupXIX	7803011	INSC
groupXIX	10201370	cftr
groupXIX	10225694	asz1
groupXIX	10225695	asz1
groupXIX	11438458	RELN
groupXIX	13808858	rassf7a
groupXIX	14035769	pamr1

groupXIX	14035773	pamr1
groupXIX	14073941	slc1a2b
groupXIX	16675756	lmf2a
groupXIX	17562354	srgap1a
groupXV	56921	mapkbp1
groupXV	985430	BRF1
groupXV	5388528	OTUB2
groupXV	10244764	grin3ba
groupXV	13281395	pdss2
groupXVI	3059660	nalcn
groupXVI	3069359	ITGBL1
groupXVI	3209162	p2ry8
groupXVI	6455500	slc40a1
groupXVI	10222778	SLC25A12
groupXVI	12337419	vwc2l
groupXVII	5344466	ARHGAP9
groupXVII	11237013	sox13
groupXVIII	747441	tnfrsf21
groupXVIII	10843666	pinx1
groupXVIII	11827808	METTL24
groupXVIII	12137126	stxbp5a
groupXVIII	13768203	fkbp1b
groupXVIII	14165771	PLCB1
groupXX	518577	si:dkey-23o4.6
groupXX	6575071	adamts16
groupXX	6934056	nrsn1
groupXX	7484682	cdkal1
groupXX	8793862	si:ch73-380l3.2
groupXX	8814434	si:ch73-380l3.2
groupXX	10710861	rad54b
groupXX	10865868	abcb4

groupXX	16253625	ctss2.1
groupXX	18939681	cdk14
groupXX	18957089	cdk14
groupXX	18957777	cdk14
groupXXI	1068422	PARD3
groupXXI	9447756	SLC39A12

Table D5. Enrichment of molecular functions among F_{ST} outlier genes (P-value < 0.05).

GO.ID	Term	Annotated	Significant	Expected	classic
GO:0005509	calcium ion binding	407	39	18.87	1.20E-05
GO:0005201	extracellular matrix structural constitu...	23	6	1.07	0.00049
GO:0005515	protein binding	3356	187	155.62	0.00101
GO:0043167	ion binding	3318	181	153.85	0.00374
GO:0043169	cation binding	1937	112	89.82	0.00528
GO:0046872	metal ion binding	1924	111	89.22	0.00591
GO:0015085	calcium ion transmembrane transporter ac...	74	9	3.43	0.00695
GO:0005488	binding	7189	355	333.35	0.0101
GO:0004000	adenosine deaminase activity	11	3	0.51	0.01237
GO:0003779	actin binding	190	16	8.81	0.01523
GO:0015368	calcium:cation antiporter activity	12	3	0.56	0.01593
GO:0008092	cytoskeletal protein binding	307	22	14.24	0.02813
GO:0005096	GTPase activator activity	109	10	5.05	0.02958
GO:0003774	motor activity	94	9	4.36	0.02988
GO:0016814	hydrolase activity, acting on carbon-nit...	26	4	1.21	0.03035
GO:0030695	GTPase regulator activity	110	10	5.1	0.03125

GO:0016462	pyrophosphatase activity	513	33	23.79	0.03463
GO:0016887	ATPase activity	112	10	5.19	0.03479
GO:0016817	hydrolase activity, acting on acid anhyd...	515	33	23.88	0.03629
GO:0016818	hydrolase activity, acting on acid anhyd...	515	33	23.88	0.03629
GO:0008081	phosphoric diester hydrolase activity	68	7	3.15	0.03738
GO:0032559	adenyl ribonucleotide binding	1095	63	50.77	0.03985
GO:0030554	adenyl nucleotide binding	1097	63	50.87	0.04112
GO:0005524	ATP binding	1079	62	50.03	0.0422
GO:0060589	nucleoside-triphosphatase regulator acti...	116	10	5.38	0.04268
GO:0016409	palmitoyltransferase activity	29	4	1.34	0.04328
GO:0005544	calcium-dependent phospholipid binding	18	3	0.83	0.04813
GO:0016810	hydrolase activity, acting on carbon-nit...	72	7	3.34	0.04863

Table D6. Enrichment of molecular functions of genes harbouring qGLM outliers (P-value < 0.05).

GO.ID	Term	Annotated	Significant	Expected	classic
GO:0046873	metal ion transmembrane transporter acti...	247	29	11.13	2.00E-06
GO:0015075	ion transmembrane transporter activity	506	46	22.79	3.70E-06
GO:0015085	calcium ion transmembrane transporter ac...	74	14	3.33	4.40E-06
GO:0005216	ion channel activity	281	29	12.66	2.50E-05
GO:0022838	substrate-specific channel activity	283	29	12.75	2.90E-05
GO:0022890	inorganic cation transmembrane transport...	331	32	14.91	3.60E-05
GO:0008324	cation transmembrane transporter activit...	350	33	15.77	4.50E-

					05
GO:0015318	inorganic molecular entity transmembrane...	461	40	20.77	4.80E-05
GO:0005261	cation channel activity	198	22	8.92	8.10E-05
GO:0015267	channel activity	305	29	13.74	0.00011
GO:0022803	passive transmembrane transporter activi...	305	29	13.74	0.00011
GO:0022839	ion gated channel activity	236	24	10.63	0.00016
GO:0005262	calcium channel activity	56	10	2.52	0.00017
GO:0005515	protein binding	3356	187	151.18	0.00018
GO:0022836	gated channel activity	241	24	10.86	0.00022
GO:0005516	calmodulin binding	51	9	2.3	0.0004
GO:0019899	enzyme binding	237	22	10.68	0.00102
GO:0022857	transmembrane transporter activity	748	52	33.7	0.00103
GO:0005215	transporter activity	771	53	34.73	0.00119
GO:0051020	GTPase binding	189	17	8.51	0.00508
GO:0017048	Rho GTPase binding	26	5	1.17	0.00544
GO:0005488	binding	7189	347	323.85	0.00573
GO:0015079	potassium ion transmembrane transporter ...	102	11	4.59	0.00606
GO:0005085	guanyl-nucleotide exchange factor activi...	133	13	5.99	0.00687
GO:0004709	MAP kinase kinase kinase activity	10	3	0.45	0.0086
GO:0005347	ATP transmembrane transporter activity	29	5	1.31	0.00879
GO:0022843	voltage-gated cation channel activity	93	10	4.19	0.00886
GO:0000295	adenine nucleotide transmembrane transpo...	30	5	1.35	0.01017
GO:0005346	purine ribonucleotide transmembrane tran...	30	5	1.35	0.01017
GO:0015215	nucleotide transmembrane transporter act...	30	5	1.35	0.01017
GO:0015216	purine nucleotide transmembrane transpor...	30	5	1.35	0.01017
GO:0015605	organophosphate ester transmembrane tran...	30	5	1.35	0.01017
GO:0015491	cation:cation antiporter activity	20	4	0.9	0.01108
GO:0005227	calcium activated cation channel activit...	11	3	0.5	0.01143

GO:0015278	calcium-release channel activity	11	3	0.5	0.01143
GO:0099604	ligand-gated calcium channel activity	11	3	0.5	0.01143
GO:0015298	solute:cation antiporter activity	21	4	0.95	0.01321
GO:0015276	ligand-gated ion channel activity	114	11	5.14	0.01356
GO:0022834	ligand-gated channel activity	114	11	5.14	0.01356
GO:0015368	calcium:cation antiporter activity	12	3	0.54	0.01474
GO:0005245	voltage-gated calcium channel activity	22	4	0.99	0.01558
GO:0043167	ion binding	3318	171	149.47	0.01594
GO:0005509	calcium ion binding	407	28	18.33	0.0167
GO:0015077	monovalent inorganic cation transmembran...	200	16	9.01	0.01851
GO:0005088	Ras guanyl-nucleotide exchange factor ac...	13	3	0.59	0.01853
GO:0004672	protein kinase activity	484	32	21.8	0.01856
GO:0004674	protein serine/threonine kinase activity	251	19	11.31	0.0186
GO:0016773	phosphotransferase activity, alcohol gro...	564	36	25.41	0.02122
GO:0016301	kinase activity	604	38	27.21	0.02233
GO:0005267	potassium channel activity	92	9	4.14	0.02244
GO:0015238	drug transmembrane transporter activity	37	5	1.67	0.02408
GO:0098772	molecular function regulator	525	33	23.65	0.03227
GO:0005244	voltage-gated ion channel activity	130	11	5.86	0.03248
GO:1901505	carbohydrate derivative transmembrane tr...	40	5	1.8	0.03259
GO:0017016	Ras GTPase binding	68	7	3.06	0.0327
GO:0031267	small GTPase binding	68	7	3.06	0.0327
GO:0004890	GABA-A receptor activity	16	3	0.72	0.03283
GO:0005217	intracellular ligand-gated ion channel a...	16	3	0.72	0.03283
GO:0032559	adenyl ribonucleotide binding	1095	62	49.33	0.03285
GO:0030554	adenyl nucleotide binding	1097	62	49.42	0.03393
GO:0015932	nucleobase-containing compound transmemb...	41	5	1.85	0.0358
GO:0043168	anion binding	1632	88	73.52	0.036
GO:0022832	voltage-gated channel activity	134	11	6.04	0.03923

GO:0008144	drug binding	1147	64	51.67	0.03944
GO:0099516	ion antiporter activity	29	4	1.31	0.03956
GO:0008514	organic anion transmembrane transporter ...	57	6	2.57	0.04224
GO:0008195	phosphatidate phosphatase activity	18	3	0.81	0.04479
GO:0016917	GABA receptor activity	18	3	0.81	0.04479
GO:0004725	protein tyrosine phosphatase activity	73	7	3.29	0.0455
GO:0005524	ATP binding	1079	60	48.61	0.04806
GO:0008509	anion transmembrane transporter activity	139	11	6.26	0.04898

Table D7. Molecular functions enriched among F_{ST}-qGLM genes (P-value < 0.05)

GO.ID	Term	Annotated	Significant	Expected	classic
GO:0015075	ion transmembrane transporter activity	506	21	7	5.60E-06
GO:0005215	transporter activity	771	25	10.66	4.80E-05
GO:0015318	inorganic molecular entity transmembrane...	461	18	6.38	6.20E-05
GO:0022857	transmembrane transporter activity	748	24	10.34	8.10E-05
GO:0015276	ligand-gated ion channel activity	114	8	1.58	0.00017
GO:0022834	ligand-gated channel activity	114	8	1.58	0.00017
GO:0022839	ion gated channel activity	236	11	3.26	0.00042
GO:0022836	gated channel activity	241	11	3.33	0.0005
GO:0005216	ion channel activity	281	12	3.89	0.00051
GO:0022838	substrate-specific channel activity	283	12	3.91	0.00054
GO:0005230	extracellular ligand-gated ion channel a...	78	6	1.08	0.00071
GO:0015267	channel activity	305	12	4.22	0.00104

GO:0022803	passive transmembrane transporter activi...	305	12	4.22	0.00104
GO:0008324	cation transmembrane transporter activit...	350	13	4.84	0.00109
GO:0003779	actin binding	190	9	2.63	0.00126
GO:0004890	GABA-A receptor activity	16	3	0.22	0.00127
GO:0016917	GABA receptor activity	18	3	0.25	0.00181
GO:0003774	motor activity	94	6	1.3	0.00188
GO:0008509	anion transmembrane transporter activity	139	7	1.92	0.00311
GO:0005254	chloride channel activity	49	4	0.68	0.00458
GO:0022890	inorganic cation transmembrane transport...	331	11	4.58	0.00612
GO:0005253	anion channel activity	55	4	0.76	0.00692
GO:0046873	metal ion transmembrane transporter acti...	247	9	3.42	0.00724
GO:0015108	chloride transmembrane transporter activ...	57	4	0.79	0.00785
GO:0035381	ATP-gated ion channel activity	10	2	0.14	0.00794
GO:0008092	cytoskeletal protein binding	307	10	4.25	0.01009
GO:0015368	calcium:cation antiporter activity	12	2	0.17	0.01144
GO:0005543	phospholipid binding	102	5	1.41	0.01344
GO:0015297	antiporter activity	38	3	0.53	0.01534
GO:0001614	purinergic nucleotide receptor activity	14	2	0.19	0.0155
GO:0016502	nucleotide receptor activity	14	2	0.19	0.0155
GO:0005096	GTPase activator activity	109	5	1.51	0.01747
GO:0030695	GTPase regulator activity	110	5	1.52	0.0181
GO:0008289	lipid binding	202	7	2.79	0.02195
GO:0060589	nucleoside-triphosphatase regulator acti...	116	5	1.6	0.02223
GO:0005544	calcium-dependent phospholipid binding	18	2	0.25	0.02513
GO:0015491	cation:cation antiporter activity	20	2	0.28	0.03066
GO:0035586	purinergic receptor activity	20	2	0.28	0.03066
GO:0015103	inorganic anion transmembrane	87	4	1.2	0.03227

	transporte...				
GO:0005516	calmodulin binding	51	3	0.71	0.03327
GO:0015298	solute:cation antiporter activity	21	2	0.29	0.03358
GO:0008047	enzyme activator activity	135	5	1.87	0.03915
GO:0022804	active transmembrane transporter activit...	182	6	2.52	0.04044
GO:0004970	ionotropic glutamate receptor activity	24	2	0.33	0.04297
GO:0008066	glutamate receptor activity	24	2	0.33	0.04297
GO:0008514	organic anion transmembrane transporter ...	57	3	0.79	0.04406
GO:0017048	Rho GTPase binding	26	2	0.36	0.0497