© Copyright 2019

Sophie Louise Archambeault

Interrogating the genotype-phenotype-fitness map of an adaptive haplotype in threespine stickleback

Sophie Louise Archambeault

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Catherine L. Peichel, Chair

Chris T. Amemiya

Harmit S. Malik

Program Authorized to Offer Degree:

Molecular and Cellular Biology

University of Washington

Abstract

Interrogating the genotype-phenotype-fitness map of an adaptive haplotype in threespine stickleback

Sophie Louise Archambeault

Chair of the Supervisory Committee:
Affiliate Professor Catherine L. Peichel
Department of Biology

Evolutionary biology is challenged with understanding how genotypes shape phenotypes, how and when selection acts on phenotypes, and how those selective forces translate to changes in the underlying frequencies of those genotypes: the genotype-phenotype-fitness map. Theoretical and technological advances have facilitated progress connecting genotypes to phenotypes and phenotypes to fitness, yet we have few cases of a complete genotype-phenotype-fitness map, limiting our ability to accurately predict evolutionary outcomes. This dissertation describes my work on the genotype-phenotype-fitness map of an adaptive and pleiotropic haplotype in threespine stickleback (*Gasterosteus aculeatus*), a powerful model of adaptive evolution. In the introduction (Chapter 1) I provide background relevant to the genetics of adaptation and my work, including progress on answering outstanding questions in the field. In Chapter 2 I describe my work on the genetic architecture of adaptation within a pleiotropic haplotype containing the

developmental signaling gene, *Ectodysplasin* (*Eda*). Specifically, I present the results of fine-mapping multiple traits in two stickleback populations aimed at disentangling the roles of pleiotropy and linkage. I find that the 16 kilobase *Eda* haplotype is significantly associated with three phenotypes in both populations of stickleback, and that all three phenotypes show the same pattern of association with the genetic markers within a small 1.4 kb region of the haplotype, suggestive of a pleiotropic mutation. In Chapter 3, I present the results of an empirical test of a leading hypothesis about selection on *Eda*; namely I test the effect of an abiotic agent of selection (dietary phosphorus) on a component of fitness (juvenile growth rate) in experimental crosses between fish that differ in their genotype at the *Eda* haplotype. The results of this experiment suggest that phosphorus limitation is not the agent of selection acting on the *Eda* haplotype, and highlight the importance of testing hypotheses of selection by experimentally connecting genotypes to fitness through selection on phenotypes. I end in Chapter 4 with my interpretation of these results, the implications of these findings, and my suggestions for future work in this system on the genetics of adaptation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This dissertation is the culmination of many years of work that was only possible due to the communities of colleagues and friends that have mentored, enabled, supported, and collaborated with me over the years. First, I want to thank the Molecular and Cellular Biology program at the University of Washington for the opportunity to do this work and for the support to pursue research, teaching, and moving to Switzerland. I am particularly grateful to Maia Low at the University of Washington and Therese Oesch at the University of Bern who both went above and beyond to make my move from Seattle to Bern as smooth as possible.

This research experience has taken me around the world to several institutions, and I have been warmly welcomed into these science communities. First, I am thankful to the UW MCB and Fred Hutchinson Cancer Research Center communities who have helped guide this work and my development as a confident and independent scientist, particularly Harmit Malik, Chris Amemiya, and Celeste Berg. I would particularly like to thank the members of Katie Peichel and Harmit Malik's labs for their interest in my research and for many helpful discussions and feedback. I would like to thank Anna Greenwood, Mike White, Jen Cech, Hilary Poore, Dianna Rennison and Shaun McCann for introducing me to the challenging and practical sides of stickleback research in the lab and in the field.

I am thankful to the University of Bern community who welcomed the North American members of the Peichel lab into the Institute of Ecology and Evolution, especially Verena Saladin and Myles Menz. I would also like to thank two additional science communities that have supported me in the final year of my dissertation work. Thank you to David Kingsley and the members of his lab who have invited me to weekly lab meetings, provided a stickleback home-away-from-home, and given me feedback on my research. And thank you to Brian Simison,

Rebecca Johnson, and Gio Rapacciuolo at the California Academy of Sciences who have provided me with a workspace and science community, and who have re-ignited my interest in science outreach and communication.

There are several individuals who have inspired me in and out of the lab. I would particularly like to thank my advisor and mentor, Katie Peichel, for her unwavering support and her dedication to thoughtful and interesting biology. She has set an excellent example of a woman in science and a mentor for me to emulate in my career. I am also thankful to her for building a community of scientists that enabled this work, since it was too large to complete by myself.

In particular, I am indebted to those who helped me on the projects described in the following pages. Madi Gamble, Casey Clark, Dave Beauchamp, Taylor Frierson, and Will Dezan helped me catch wild stickleback. I am grateful to Shaugnessy McCann, Hilary Poore, Jake Whitehurse, Verena Saladin, and Melanie Hiltbrunner for excellent fish husbandry, Soraya Villalba for genotyping hundreds of fish, Aurélie Merminod for carefully dissecting and phenotyping fish, and especially Hilary Poore and Melanie Hiltbrunner for help with fish husbandry, genotyping, and phenotyping. A huge thank you to Luis Bärtschi for his continued dedication to this work including photographing and phenotyping over a thousand fish. Thank you to all the Peichel lab members who carefully fed pellet food to the fish for a year- a tedious but important team effort. A huge thank you to Dan Durston for flying to Switzerland and teaching me to collect fish excretion, to Alex Wan for formulating the special phosphorus diets, and to Daniel Steiner for performing the phosphorus measurements. I would also like to thank Jun Kitano, David Kingsley, Craig Miller, Rowan Barrett, Rana El-Sabaawi, Blake Matthews, Harmit Malik, and Chris Amemiya for helpful feedback on this dissertation and the manuscripts contained within.

Finally, I would like to acknowledge the efforts of many friends and family members towards keeping me motivated and relatively sane during this process. Thank you to Jen Cech, Lauren Saunders, Abby Keller, Brittany Whitley, Michelle Hayes, Rebecca Resnick, Anna Feller, Rishi De-Kayne, and Amy Herbert for their friendship, as well as for setting inspiring examples of how to be a successful PhD student, and for pushing me to believe in myself. I especially want to thank Lauren, Abby, Brittany, Michelle, and Rebecca who made life outside of grad school entertaining and adventurous. Thank you to Scott Roy, who has never once doubted my ability to finish this work and who has unquestioningly accommodated my moving to Seattle and later to Switzerland. I know we will enjoy the benefits of his well-honed flight purchasing skills for years to come, but I also want him to know that I appreciate him, his support, and his love. I would especially like to thank my parents, Paul and Elizabeth, and my sister, Emilie, who have made this all possible. Thank you for your love and understanding throughout this journey.

DEDICATION

I dedicate this work to my sister, Emilie.

Thank you for going through life's adventures with me.

What's next?

Chapter 1. INTRODUCTION

A central goal of evolutionary biology is to characterize the processes that generate and maintain phenotypic diversity among individuals, populations, and species. These processes include both inherently evolutionary processes, such as drift, mutation, recombination, migration and natural selection, as well as non-evolutionary features of individual organisms such as plasticity and developmental noise. Although it has long been recognized that "natural selection is not evolution" (Fisher, 1930), selection is often invoked to explain the suitability of organisms to their environments, and insofar as a large fraction of traits are at least partially heritable (Hansen *et al.*, 2011; Hendry, 2013; Cava *et al.*, 2019), natural selection has the opportunity to shape many traits. I have been interested in the genetic and developmental bases of phenotypic novelty and diversity for many years, and therefore was particularly attracted to studying the genetic basis of adaptation during my PhD training.

As a general background to my thesis on the evolution of bony armor in stickleback fish, this Introduction provides a synopsis of the relevant aspects of Evolution and Adaptation. As Morgan wrote in 1903 "I am fully aware of the danger in attempting to cover so wide a field as that of 'Evolution and Adaptation,' and I cannot hope to escape the criticism that is certain to be directed against a specialist who ventures nowadays beyond the immediate field of his own researches" (Morgan, 1903), nonetheless I have attempted to tease out some of the large questions most relevant to the work presented in the remainder of the thesis.

Since the integration of Evolution and Genetics (i.e., the Modern Synthesis), it has been recognized that adaptation entails changes in allele frequencies in a population that occur due to non-random survival and/or reproduction of individuals carrying different alleles (Endler, 1986;

Provine, 2001), thus integrating phenotype, genotype and fitness. For multicellular organisms, development is also closely tied to adaptation, since selection acts on phenotypes produced via developmental processes that integrate genetic and environmental signals (Alberch, 1991). Thus, a comprehensive understanding of the process of adaptation requires connecting the genetic and developmental bases of phenotypes to the fitness effects of possessing the underlying alleles in specific environments: the genotype-phenotype-fitness map. In turn, characterizing the genotype-phenotype-fitness map minimally requires characterizing the genetic architecture of complex phenotypic variation and the features of adaptive alleles, as well as the selective pressures that shape phenotypic evolution (Barrett and Hoekstra, 2011). I devote the first part of this Introduction to the rich question of the nature of adaptive alleles ("1.1 THE NATURE OF ADAPTIVE MUTATIONS"), followed by an argument for the importance of experimentally connecting the agents of selection to the underlying alleles via the phenotypes ("1.2 THE SELECTIVE FORCES THAT DRIVE ADAPTATION"). I have included examples from the relevant literature in my model organism, the threespine stickleback (*Gasterosteus aculeatus*) throughout these sections.

1.1 THE NATURE OF ADAPTIVE MUTATIONS

The central question of the genetic basis of adaptation is which mutations contribute to adaptation. This large question in turn can be broken down into a large number of smaller questions as to the features of adaptive mutations. For instance, relative to other mutations (those that do not contribute to adaptation), are adaptive mutations more or less (i) large in effect size; (ii) likely to be regulatory or coding (iii) likely to arise from standing genetic variation or from *de novo* mutations; and (iv) pleiotropic. In this section I treat each of these questions in turn.

I. Effect size and number of adaptive mutations.

a. Background and the Fisher model

> "Slow though the process of selection may be, if feeble man can do much by his powers of artificial selection, I can see no limit to the amount of change, to the beauty and infinite complexity of the coadaptations between all organic beings, one with another and with their physical conditions of life, which may be effected in the long course of time by nature's power of selection." (Darwin, 1859)

Ever since Darwin and others expounded the principle of natural selection acting on phenotypic variation, debates have raged about the origins and nature of this variation. One key axis in this debate has focused on the number and sizes of the differences that underpin observed phenotypic diversity.

The debate over whether evolutionary change and adaptation occurs through few mutations of large effect or many mutations of small effect dates back over 100 years. Darwin believed that evolution via natural selection would be a slow and gradual process, an idea embraced by the so-called micromutationalists, who interpreted the continuous distribution of quantitative traits as evidence in support of gradual evolution (Pearson *et al.*, 1901). However, documentation of discontinuous traits such as melanism in peppered moths (Ford, 1937; Kettlewell, 1955), measurable shifts in phenotypic traits of house sparrows in response to short-term environmental changes (Bumpus, 1899), and Mendelian inheritance of certain traits was at odds with the notion of slow, incremental evolution (Gillham, 2001). The tension between the seemingly contrasting implications of Mendelian inheritance and continuous biometrical traits was resolved mathematically by R.A. Fisher (1918) who showed that it is possible to explain continuous phenotypic variation through the Mendelian inheritance of multiple genetic loci, thereby laying the groundwork for quantitative genetics (Visscher and Goddard, 2019).

Fisher viewed adaptation to an environment as a complex process, often requiring changes in suites of traits. He framed adaptation in mathematical terms, viewing adaptation as a process of movement within a multidimensional space with the fitness optimum at the origin and in which dimensions represent traits, and mutations are represented by vectors through that fitness space (Fisher, 1930). Fisher further stipulated that mutations are random in direction and size. Because of this randomness, adaptation does not involve a single leap to the optimum, but instead proceeds through a series of mutations along a potentially quite indirect path.

Treatment of adaptation in these terms leads to novel conclusions about the expected distributions of effect sizes specifically of adaptive mutations. In particular, the distribution of effect sizes of adaptive mutations is expected to be different from the distribution of newly arising mutations. This is because, whereas newly occurring mutations of small effect (relative to the distance to the optimum) are roughly equally likely to be beneficial as they are to be detrimental, large effect mutations are extremely unlikely to be beneficial (Fisher, 1930). These conclusions have been refined through the inclusion of the probability of loss due to drift as well as the change in the probability of mutations being beneficial as the optimum is approached (Kimura, 1983; Orr, 1998). As such, while larger mutations may be adaptive at early steps in the walk, only small mutations will be adaptive as the population approaches the optimum. This updated geometric model predicts an exponential distribution of phenotypic effects, with few mutations of large effect and many mutations of relatively small effect.

b. Empirical data on the number and effect sizes of adaptive mutations

Data to test the predictions of Fisher's model have come from mapping of suspected or confirmed adaptive phenotypes or components of fitness (Ågren *et al.*, 2013; Selby and Willis, 2018). In either case, there are two ways to calculate the distribution of effect sizes: quantitative trait loci

(QTL) mapping and genome wide association studies (GWAS). An example of the former comes from QTL mapping of lateral plate loss in freshwater stickleback, which is largely controlled by a single locus surrounding the developmental signaling molecule, *Ectodysplasin* (*Eda*) (Colosimo *et al.*, 2005). While *Eda* controls 76% of the variation in plate number, there are three additional small-effect QTL on separate chromosomes that control 2.9, 3.5, and 4.9% of the variation in plate number, consistent with few mutations of large effect, and many more mutations of small effects (Colosimo *et al.*, 2004). More generally, meta-analyses of the available QTL data also provide support for Fisher's geometric model (Rockman, 2012; Hendry, 2013; Peichel and Marques, 2017), which, after accounting for publication bias of large-effect mutations and the limited ability to detect mutations with small effects (Xu, 2003; Rockman, 2012; Hendry, 2013; Wellenreuther and Hansson, 2016), roughly approximates an exponential distribution of effect sizes across all putatively adaptive traits (Orr, 1998; Otto and Jones, 2000).

We can also quantify the number of QTL that underlie a single adaptive trait using the data gathered by Peichel and Marques (2017) (Figure 1.1). Out of 818 mapped traits, only a single QTL per trait was detected for 112 traits, and the most QTL detected for a single trait was 10, suggesting that most traits are polygenic with a limited number of underlying QTL. There are two caveats to these data. First, QTL are usually large genomic regions that may include multiple causative mutations and/or genes, so these data likely underrepresent the number of mutations underlying an adaptive trait. Second, the power to detect QTL with small effects is limited by the number of F2 mapping individuals so QTL of small effect would be expected to go undetected in these studies.

The geometric model further predicts that adaptation of populations further from the optimum will result in fixation of mutations with larger effect sizes (Orr, 1998; Dittmar *et al.*, 2016). This

prediction has held in both laboratory experiments (Sousa *et al.*, 2011; McGee *et al.*, 2016) and in stickleback adaptation to freshwater lakes with and without predatory sculpins (Rogers *et al.*, 2012). To date, data from both experimental studies and studies of natural populations have qualitatively supported many of the predictions of the geometric model.

Additional power in QTL and GWAS studies is needed to identify all genomic loci contributing to adaptation, which is likely to raise the number of loci that underlie adaptive quantitative traits. Future progress on this central question will also come from categorizing the adaptive mutations from different adaptive scenarios, including local adaptation in the face of gene flow versus population-wide directional selection, adaptation from standing variation versus from *de novo* mutations, and adaptation of different types of traits (morphological, physiological, behavioral and life-history traits) (Stern and Orgogozo, 2008; Dittmar *et al.*, 2016). Finally, meta-analyses of these future studies of adaptive evolution will be required to contextualize and answer these questions.

II. MOLECULAR NATURE OF ADAPTIVE MUTATIONS AND THE *CIS*-REGULATORY HYPOTHESIS

Another central question in the study of the genetics of adaptation concerns the molecular impacts of adaptive mutations. Changes in the genome (i.e., mutations) can influence the molecular biology of the organism in a variety of ways. Along one axis, mutations may cause changes in the sequence of the encoded product while other mutations may influence the expression of genes. Along another axis, mutations may affect genes that reside in different positions within regulatory networks, ranging from 'switch' genes or master regulators that influence the expression of tens or hundreds of additional genes to 'target' genes that do not directly alter expression of other genes but instead code for functional products such as enzymes. Finally, mutations may alter genes that are expressed in a range of cell types, from genes involved in proper function of only a single cell

type to genes required for proper function of all cells. The impacts of mutations with different molecular impacts may have very different implications on phenotypes, fitness, and therefore adaptation.

There has been much debate over what type of mutations are expected to drive phenotypic evolution (King and Wilson, 1975; Carroll, 2005; Davidson and Erwin, 2006; Hoekstra and Coyne, 2007; Wray, 2007; Stern and Orgogozo, 2008; Martin and Orgogozo, 2013). For example, morphological evolution is predicted to occur via mutations in developmental genes imbedded within gene regulatory networks, i.e. switch genes (Davidson and Erwin, 2006), and this has largely been borne out in stickleback, albeit from only a small number of characterized cases (Peichel and Marques, 2017).

Adaptation by changes in switch genes presents a paradox. Such switch genes often have severe deleterious consequences when mutated in laboratory animals due to the pleiotropic effects of switch genes like *Eda* (Srivastava *et al.*, 1997; Harris *et al.*, 2008; Wucherpfennig *et al.*, 2019), raising the question of how their function can be perturbed without severe deleterious pleiotropic consequences. In order to reconcile evolution of switch genes with their pleiotropic effects, the *cis*-regulatory hypothesis theorizes that mutations that change when and where a switch protein is expressed would be favored over coding mutations that change whether the protein functions under all conditions. In contrast, physiological trait evolution is hypothesized to occur via either coding or *cis*-regulatory mutations in proteins at the edge of networks that have fewer pleiotropic effects, i.e. target genes (Carroll, 2005; Davidson and Erwin, 2006; Stern and Orgogozo, 2008). To answer this debate, we need to identify the underlying mutations in many cases of phenotypic evolution- the quantitative trait nucleotides (QTN)- which requires significant research effort. Despite this

hurdle, there is continued effort to compile the relevant data for this question (gephebase.org) and to periodically assess the patterns within the compiled data.

Within this context, the question of adaptation can be framed in terms of how adaptive mutations differ from the full set of mutations that have phenotypic effects, i.e. which types of changes are over- and underrepresented among adaptive molecular changes. Data in many systems, including stickleback, are growing (Peichel and Marques, 2017) and meta-analyses generally support the importance of *cis*-regulatory mutations in morphological evolution at the interspecific level (Stern and Orgogozo, 2008, 2009; Martin and Orgogozo, 2013). Data from known adaptive traits also support the *cis*-regulatory hypothesis, since most identified mutations affect the regulation of developmental genes controlling adaptive, morphological traits (Hendry, 2013). These compiled data also highlight the flexibility of adaptive physiological traits, for which mutations that affect expression level, coding sequences and gene copy number have been identified (Stern and Orgogozo, 2008; Howes *et al.*, 2017; Marques *et al.*, 2017; Peichel and Marques, 2017; Cleves *et al.*, 2018; Ishikawa *et al.*, 2019). However, these data include only eight stickleback traits, so more effort to identify the QTN of adaptive traits, especially behavioral, physiological and life-history traits, is required before strong conclusions can be drawn.

III. THE ORIGINS OF ADAPTIVE MUTATION: *DE NOVO* MUTATION VS. STANDING VARIATION

Another feature that can affect what kind of mutations contribute to adaptive walks is the source of those mutations (Sella and Barton, 2019). In particular, mutations that contribute to an adaptive walk can either occur *de novo* during that walk or can be selected from variation that existed in the population before an environmental shift rendered the mutations adaptive. Notably, Fisher's geometric model of adaptation was based solely on mutations arising *de novo* during adaptation,

however several features of the adaptive process may be considerably different if adaptation instead occurs substantially through standing genetic variation.

Theory predicts that adaptation may frequently occur through use of standing genetic variation. This is for at least two reasons. First, since selection can act on phenotypic variation already present in the population, adaptation need not wait for new mutations to occur. Waiting for new mutations may be particularly problematic in vertebrates due to the relatively small effective population sizes (Charlesworth, 2009). Secondly, mutations at higher frequency are more likely to fix (i.e., less likely to be lost by genetic drift) than a *de novo* mutation with the same fitness effect (Hermisson and Pennings, 2005; Patwa and Wahl, 2008).

So how often does adaptation occur through standing genetic variation? The current data suggest that both standing variation and *de novo* mutations play a role in adaptation. One attempt to quantify the fraction of adaptive loci that comes from standing genetic variation in stickleback estimated that 35% of putatively adaptive freshwater alleles in one population are shared globally across freshwater populations, perhaps representing a lower bound on the contribution of standing genetic variation (Jones *et al.*, 2012b). Similarly, a recent comparison of genomic divergence between marine and freshwater stickleback found that > 50% of putatively adaptive loci are shared between recently evolved, island freshwater populations and much older and geographically distant, mainland populations (Bassham *et al.*, 2018), emphasizing the importance of standing genetic variation in adaptation of stickleback to freshwater.

Importantly, adaptation from standing genetic variation may lead to a different distribution of effect sizes of beneficial mutations than adaptation from *de novo* mutations, depending on the type of selection acting on the relevant trait(s) before the environmental shift. The type of selection in standing populations, i.e. disruptive, stabilizing, or balancing, is expected to influence which

mutations are immediately available to selection after an environmental shift (Lande and Arnold, 1983; Josephs *et al.*, 2017). Evolutionary geneticists have often assumed that most populations evolve under stabilizing selection (Barton and Keightley, 2002), which predicts purging of mutations with large phenotypic effects leading to a prediction of adaptation from highly polygenic standing phenotypic variation dominated by mutations of small effect (Barrett and Schluter, 2008; Hendry, 2013). GWAS data from humans and the current data in plants are consistent with this model, finding evidence for pervasive negative selection against phenotype-altering mutations (Boyle *et al.*, 2017; Josephs *et al.*, 2017; Visscher *et al.*, 2017; Zeng *et al.*, 2018; Watanabe *et al.*, 2019). On the other hand, direct measurements of selection on phenotypes in wild populations finds that disruptive selection is as common as stabilizing selection (Kingsolver *et al.*, 2001; Kingsolver and Diamond, 2011; Kingsolver *et al.*, 2012), leading to different predictions of standing variation. Under disruptive selection, standing variation might not be biased away from mutations of large effect. For example, disruptive selection maintains a large effect polymorphism within a freshwater population of stickleback (Marchinko *et al.*, 2014). The same large effect polymorphism is a classic example of adaptation to freshwater from standing genetic variation in marine stickleback populations. This polymorphism is the large-effect *Eda* locus that controls 76% of the variation in lateral plate number in threespine stickleback (Colosimo *et al.*, 2005), and on which this thesis focuses. While disruptive selection maintains this polymorphism in certain freshwater populations, it is unclear what selective forces, if any, maintain this polymorphism in the marine environment. Clearly, a better understanding of the relative contributions of *de novo* mutations versus standing genetic variation, as well as the selective forces that shaped the standing variation prior to the environmental shift, is needed (Dittmar *et al.*, 2016).

IV. PLEIOTROPY AND ADAPTIVE MUTATIONS

Another key issue in shaping which mutations contribute to adaptation is pleiotropy. Pleiotropy is the phenomenon where a single gene or mutation affects multiple traits or components of fitness (Paaby and Rockman, 2013). An interesting paradox has arisen as data have accumulated on the genetic basis of adaptive mutations. On the one hand, theory predicts that adaptive evolution should be constrained by pleiotropy through both "a cost of complexity" where more complex organisms are predicted to adapt more slowly, and through the phenomenon of antagonistic pleiotropy in which the beneficial effects of a mutation may be partially or completely canceled out by additional deleterious effects on other phenotypes (Orr, 2000; Welch and Waxman, 2003; Carroll, 2005). Yet empirical studies have found that pleiotropy is a relatively common feature of adaptive genomic loci (Hawthorne and Via, 2001; Albertson *et al.*, 2003; Bratteler *et al.*, 2006; Hall *et al.*, 2006; Albert *et al.*, 2008; Rogers *et al.*, 2012; Liu *et al.*, 2014; Friedman *et al.*, 2015; McGee *et al.*, 2016; Frachon *et al.*, 2017).

How can we reconcile these two contrasting findings about pleiotropy? First, the theory is based on the pleiotropy of individual *mutations*, while the empirical QTL and GWAS studies identified pleiotropic genomic regions. Thus, one possibility is that these pleiotropic adaptive regions contain multiple mutations, each of which affects a single trait. Addressing this possibility requires identifying the QTNs underlying multiple adaptive phenotypes that map to a pleiotropic genomic locus (Chapter 2). This has been achieved in a few cases, revealing that the pleiotropic nature of some loci is indeed caused by the clustering of multiple apparently less pleiotropic mutations (Rebeiz *et al.*, 2009; Frankel *et al.*, 2011; Hermann *et al.*, 2013; Erickson *et al.*, 2018).

The clustering of adaptive mutations is predicted by theory in certain cases of adaptation (Charlesworth and Charlesworth, 1979; Yeaman and Whitlock, 2011). In particular, when a

population is adapting to a novel environment or niche in the face of gene flow, reduction in recombination that increases the co-inheritance of alleles adapted to the same environment should be favored. In the extreme, natural selection could favor genomic rearrangements that facilitate linkage disequilibrium or suppress recombination between adaptive alleles, such as inversions (Kirkpatrick and Barton, 2006; Kirkpatrick, 2010; Via *et al.*, 2012; Fishman *et al.*, 2013; Yeaman, 2013). This reduced recombination would appear as a pleiotropic locus in a mapping study even in the absence of pleiotropy of the constituent alleles.

However, a surprising pattern has emerged for pleiotropic loci that may be hard to explain by the clustering of mutations. There is synergistic scaling between the degree of pleiotropy (how many traits a locus affects) and the per-trait effect size, such that pleiotropic loci tend to have larger per trait effects (Wagner *et al.*, 2008; Wang *et al.*, 2010; Frachon *et al.*, 2017). The cause of this pattern is unclear, though it seems unlikely to be due to linked mutations, suggesting that some adaptive loci of large effect are driven by pleiotropic mutations. Perhaps most surprisingly, a single *cis*-regulatory mutation was recently found to have pleiotropic effects on both genital bristle and sex comb tooth numbers in *Drosophila* (Nagy *et al.*, 2018).

A correlation between pleiotropy and effect size raises the question of the individual and combined fitness effects of phenotypic variation driven by pleiotropic mutations. The fitness effects of two or more phenotypic effects of a pleiotropic locus could tend to be synergistic (both costly or both beneficial), antagonistic (one of each), neutral (random), or any of these depending on the environment, population or traits affected. Theory predicts that pleiotropic mutations are likely to have antagonistic fitness effects, thereby reducing the probability of fixation during adaptation (Otto, 2004). Evidence of antagonistic pleiotropy slowing or preventing adaptation certainly exists (Scarcelli *et al.*, 2007), but so does evidence for synergistic pleiotropy facilitating

adaptation (McGee *et al.*, 2016). A meta-analysis of selection in natural populations found little evidence of constraint on phenotypic evolution in response to directional selection due to correlated traits (e.g., from pleiotropy) (Kingsolver and Diamond, 2011). To understand and predict the adaptive potential of or constraint imposed by pleiotropic loci, we need to interrogate the genotype-phenotype-fitness map of known, pleiotropic loci, which is the focus of this work.

1.2 THE SELECTIVE FORCES THAT DRIVE ADAPTATION

Much progress has been made on the genetics of putatively adaptive phenotypes (the genotype-phenotype map). Studies around the turn of the century confirmed that genomic mapping of traits in QTL crosses was possible (Bradshaw *et al.*, 1998; Peichel *et al.*, 2001; Albertson *et al.*, 2003; Colosimo *et al.*, 2004; Protas *et al.*, 2008), and advances in genomic sequencing have improved our ability to detect the loci that underlie phenotypic evolution and produced an abundance of data (Martin and Orgogozo, 2013). As discussed in 1.1 THE NATURE OF ADAPTIVE MUTATIONS, this work has provided the data to test longstanding hypotheses about the forces that shape phenotypic variation.

Despite this abundance of data, our understanding of the fitness effects of alleles and the agents of selection that act on them is lacking. To address this, we need to connect the phenotypes for which we understand the genetic basis to the fitness consequences of possessing the alleles. This has been attempted in several systems where experimental manipulations or long-term study was possible (Gratten *et al.*, 2008; Gratten *et al.*, 2012; Anderson *et al.*, 2013; Lovell *et al.*, 2013; Gompert *et al.*, 2014; Soria-Carrasco *et al.*, 2014; Milesi *et al.*, 2016; Barrett *et al.*, 2019; Rennison *et al.*, 2019). As predicted from the genotype-phenotype and phenotype-fitness maps, some of these studies document selection acting on the alleles that control adaptive phenotypes in the predicted direction, such as coat color in deer mice, spine length in stickleback, cryptic color

pattern in stick insects, and insecticide resistance in mosquitos (Gompert *et al.*, 2014; Milesi *et al.*, 2016; Barrett *et al.*, 2019; Rennison *et al.*, 2019).

While these heroic efforts are confirming some of our expectations of adaptive evolution, other studies are highlighting the complexity and unpredictability of the genotype-phenotype-fitness map (Barrett and Hoekstra, 2011). For example, genome-wide association studies identified *ALX1* as a candidate gene associated with beak morphology and body size, traits that experience strong selection during repeated droughts in Darwin's finches (Boag and Grant, 1981; Price *et al.*, 1984; Gibbs and Grant, 1987; Grant and Grant, 2002, 2006; Lamichhaney *et al.*, 2015). However, despite the presence of genetic variation at *ALX1* within medium ground finches on Daphne Major, this locus did not experience changes in allele frequencies consistent with the phenotypic changes during the 2004 - 2005 drought (Lamichhaney *et al.*, 2016b).

A second example of the complexity of putative cases of phenotypic adaptation comes from *Timema* stick insects. Selection on color and pattern morphs is driven by predation (Nosil and Crespi, 2006). Recent work identified a genomic locus associated with color and pattern morphs. The expectation is that under divergent predation regimes, the genomic locus would show predictable changes in allele frequencies. As predicted, a shift in allele frequency at the locus was observed in three separate experimental studies, confirming that selection acts on the a genomic locus underlying traits under selection. However, despite these links between phenotype, genotype and selection, phenotypic evolution over long periods of time was difficult to predict (Nosil *et al.*, 2018). The authors contribute this unpredictability of evolution to "multiple, complex and counteracting sources of selection", highlighting the need to identify the agents of selection most associated with fitness of individuals with alternate alleles in different environments.

Another complex genotype-fitness relationship exists for a single coding mutation in the *TYRP1* gene (Gratten *et al.*, 2007). The mutation is nearly perfectly associated with dark coat color. Dark coat color and the mutation are associated with larger birth weight, likely due to linkage with additional allele(s) (Gratten *et al.*, 2008), and selection favors lambs with larger birth weights (Wilson *et al.*, 2005). Therefore, an increase in the frequency of the *TYRP1* mutation is expected over time since there is no evidence for direct selection on coat color that might counteract the linked effects of larger birth weights (Gratten *et al.*, 2008). However, the frequency of the *TYRP1* mutation has been declining, due to reduced fitness of homozygous dark sheep compared with phenotypically indistinguishable heterozygous sheep and homozygous light sheep (Gratten *et al.*, 2008). The selective agent(s) causing this fitness differential are unknown.

Our final example comes from the threespine stickleback. As introduced above, one of the most striking morphological adaptations to freshwater is the repeated loss of lateral bony plates. Loss of bony plates is associated with a 16 kilobase freshwater haplotype on chromosome IV, which contains the *Eda* gene, and controls most of the variation in plate number (Colosimo *et al.*, 2005). The freshwater haplotype is nearly fixed in most freshwater populations, consistent with repeated and strong selection for the freshwater version of *Eda* in freshwater environments (Hohenlohe *et al.*, 2010; Jones *et al.*, 2012b). A 2008 study put wild-caught marine fish heterozygous at the *Eda* haplotype into four replicate, semi-natural, freshwater ponds and measured change in allele frequency at *Eda* in the F1 generation for a year. The expectation was that selection would favor the freshwater *Eda* allele at some or all life stages. Instead, selection favored the homozygous marine genotype (and disfavored the heterozygous genotype) in all four ponds during juvenile stages (Barrett *et al.*, 2008). Selection then shifted in favor of the freshwater allele at later life stages to overcome this early disadvantage. Additional analyses detected

selection favoring the freshwater allele, even among fish with the same number of lateral plates, suggesting that selection was acting on additional phenotypes linked to (or affected by) *Eda* (Rennison *et al.*, 2015).

These examples highlight the complexity of the genotype-phenotype-fitness map. Despite our ability to map the genetic basis of adaptive traits and measure selection in the wild, some cases of adaptation remain untested and unpredictable (Barrett and Hoekstra, 2011). One explanation for this may be the pleiotropic nature of the genotype-phenotype and/or genotype-fitness map, as highlighted by the Soay sheep and stickleback examples. Therefore, dissecting the genetic architecture of traits within a pleiotropic locus is important, and the focus of Chapter 2. An additional hinderance to predicting adaptation could be our lack of understanding of the agents of selection acting on phenotypes and their underlying genotypes in the wild. Though there are likely to be many different agents of selection acting over time and space in the wild (Siepielski *et al.*, 2009; Kingsolver *et al.*, 2012), identifying the most important agent(s) of selection and measuring their selective effects on alleles and phenotypes will allow for better predictive models of adaptation. This is the motivating rationale for the work presented in Chapter 3.

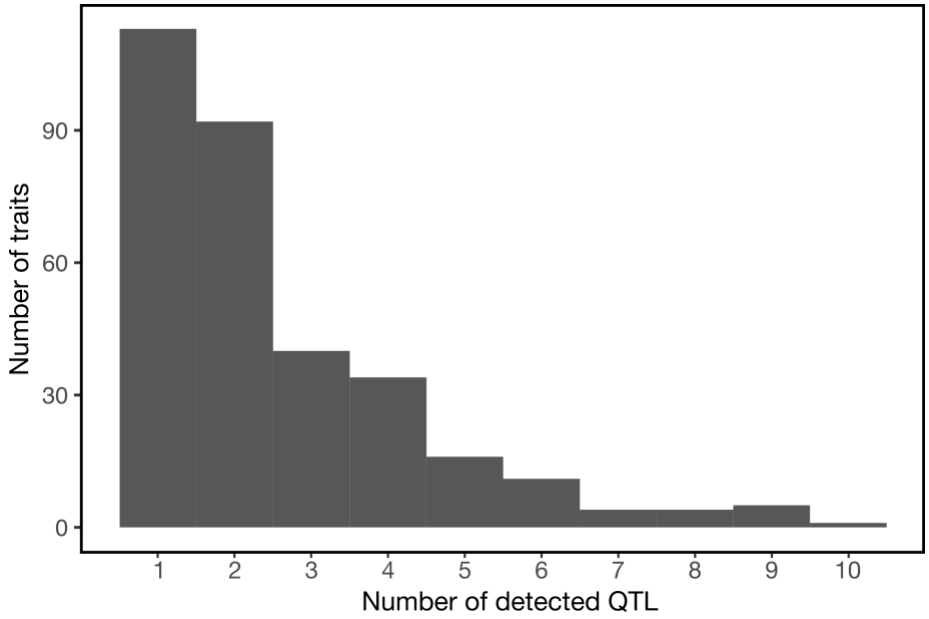


Figure 1.1. The distribution of number of QTL detected per trait. The data presented here were gathered by Peichel and Marques (2017) and reflect 28 mapping studies in threespine stickleback. A single QTL was detected for many traits, but most traits are polygenic.

Chapter 2. ADAPTATION VIA PLEIOTROPY AND LINKAGE: ASSOCIATION MAPPING REVEALS A COMPLEX GENETIC ARCHITECTURE WITHIN THE STICKLEBACK *EDA* LOCUS

2.1 SUMMARY

Genomic mapping of the loci associated with phenotypic evolution has revealed genomic "hotspots", or regions of the genome that control multiple phenotypic traits. This clustering of traits has important implications for the speed and maintenance of adaptation and could be due to pleiotropic effects of a single mutation or tight genetic linkage of multiple causative mutations affecting different traits. The threespine stickleback (*Gasterosteus aculeatus*) is a powerful model for the study of adaptive evolution because the marine ecotype has repeatedly adapted to freshwater environments across the northern hemisphere in the last 12,000 years. Freshwater ecotypes have repeatedly fixed a minimal 16 kilobase haplotype on chromosome IV that contains *Ectodysplasin* (*Eda*), a gene known to affect multiple traits, including defensive armor plates, lateral line sensory hair cells, and schooling behavior. Many additional traits have previously been mapped to a larger region of chromosome IV that encompasses the *Eda* freshwater haplotype. We used crosses of marine fish heterozygous for the freshwater haplotype to identify which of these traits specifically map to this adaptive haplotype. Further, we performed fine-scale association mapping in a fully interbreeding, polymorphic population of freshwater stickleback to disentangle the effects of pleiotropy and linkage on the phenotypes affected by this haplotype. Although we find evidence that linked mutations have small effects on a few phenotypes, a small 1.4 kb region within the first intron of *Eda* has large effects on three phenotypic traits: lateral plate count, and both the number and patterning of the posterior lateral line neuromasts. Thus, the *Eda* haplotype

is a hotspot of adaptation in stickleback due to both a small, pleiotropic region affecting multiple traits as well as multiple linked mutations affecting additional traits.

2.2 INTRODUCTION

Adaptation to divergent environments is often associated with changes in many traits (Darwin, 1859; Fisher, 1930; Orr, 2000). For example, when a marine fish colonizes freshwater, it not only encounters a new abiotic environment, but also a new biotic environment with different predators, prey, and parasites. Thus, adaptation to freshwater is expected to involve a suite of morphological, behavioral and physiological changes. Co-inheritance of these suites of traits under selection in a particular environment is predicted to facilitate adaptation (Charlesworth and Charlesworth, 1979; Kirkpatrick and Barton, 2006; Hoffmann and Rieseberg, 2008; Schwander *et al.*, 2014). Indeed, genetic mapping of phenotypic changes in systems of adaptive evolution has identified clustering of traits, such that multiple traits are affected by a single genomic region or gene (Hawthorne and Via, 2001; Albertson *et al.*, 2003; McKay *et al.*, 2003; Bratteler *et al.*, 2006; Hall *et al.*, 2006; Scarcelli *et al.*, 2007; Protas *et al.*, 2008; Lowry and Willis, 2010; Joron *et al.*, 2011; Parnell *et al.*, 2012; Yoshizawa *et al.*, 2012; Friedman *et al.*, 2015; Peichel and Marques, 2017). In most cases, it is unknown whether these phenotypic hotspots are due to the effects of a single pleiotropic mutation; or multiple, linked causative mutations; or a combination of both (but see Carbone *et al.* (2006); Hermann *et al.* (2013); Kamberov *et al.* (2013); Linnen *et al.* (2013); Lee *et al.* (2017); Dong *et al.* (2018); Erickson *et al.* (2018); Nagy *et al.* (2018); Butelli *et al.* (2019)). Knowing if a single mutation is affecting multiple traits or if multiple, linked mutations within a pleiotropic gene or genomic region underlie adaptive traits is crucial to understanding and predicting the adaptability of populations.

Pleiotropy could facilitate rapid, adaptive evolution, or adaptation in the face of gene flow, if all or most of the phenotypic changes were beneficial in the new environment. However, the classical view is that pleiotropy is more likely to constrain adaptation because the probability that

a mutation with beneficial effects on one trait has detrimental effects on other traits and overall fitness (i.e. antagonistic pleiotropy) is predicted to increase with the degree of pleiotropy imposing a "cost of complexity" (Fisher, 1930; Orr, 2000; Otto, 2004). Indeed, the pleiotropic effects of many developmental genes and disease mutations has led to the expectation that morphological evolution is more likely to occur through mutations that reduce pleiotropy, such as tissue-specific regulatory mutations (Carroll, 2008; Stern and Orgogozo, 2008), but see (Hoekstra and Coyne, 2007). Empirical work has linked morphological evolution to mutations in modular or tissue-specific enhancers (Rebeiz *et al.*, 2009; Chan *et al.*, 2010; Frankel *et al.*, 2011; Wallbank *et al.*, 2016), but recent work has also identified pleiotropic mutations in regulatory DNA (Nagy *et al.*, 2018; Ramaekers *et al.*, 2019). Meta-analyses of genome-wide association studies report widespread pleiotropy of genomic loci (Boyle *et al.*, 2017; Chesmore *et al.*, 2018; Watanabe *et al.*, 2019). However, data from gene knockout studies in yeast, nematodes and mice as well as quantitative trait locus (QTL) mapping in mice find that most genes or QTL exhibit little or no pleiotropy, but loci that are pleiotropic show a positive correlation between the per trait effect sizes of mutations with the number of traits affected (Wagner *et al.*, 2008; Wang *et al.*, 2010). Together these findings suggest that the "cost of complexity" imposed by pleiotropy can be overcome. In fact, with the synergistic scaling of per trait effect size and degree of pleiotropy, intermediate levels of pleiotropy are expected to be the most adaptive (Wang *et al.*, 2010; Wagner and Zhang, 2011). This expectation of rapid adaptation occurring through mutations with intermediate levels of pleiotropy has been found in wild populations of *Arabidopsis thaliana* (Frachon *et al.*, 2017). However, this pleiotropy is measured at the level of QTL or genomic regions, and the extent to which the pleiotropic effects of a single mutation contribute to adaptation remain unknown.

Linkage of multiple adaptive mutations has also been proposed as a mechanism to facilitate rapid adaptation, particularly in the face of gene flow (Charlesworth and Charlesworth, 1979; Kirkpatrick and Barton, 2006; Hoffmann and Rieseberg, 2008; Yeaman and Whitlock, 2011; Ortiz-Barrientos *et al.*, 2016). Consistent with the theory that there is selection for tight linkage between alleles that contribute to adaptation, genomic regions of low recombination, such as inversions, often harbor loci important for many different traits (Lowry and Willis, 2010; Joron *et al.*, 2011; Fishman *et al.*, 2013; Hermann *et al.*, 2013; Wang *et al.*, 2013; Kunte *et al.*, 2014; Küpper *et al.*, 2016; Lamichhaney *et al.*, 2016a; Tuttle *et al.*, 2016; Lee *et al.*, 2017; Cocker *et al.*, 2018). With a few exceptions (Hermann *et al.*, 2013), the number and nature of the causative mutation(s) in these regions have not yet been identified.

Linkage and pleiotropy have different implications for the speed, acquisition, and maintenance of phenotypic effects during adaptive evolution. For example, linked mutations are likely acquired one at a time, which could require longer to gain multiple phenotypic effects than a single, pleiotropic mutation. Furthermore, in contrast to the phenotypic effects of pleiotropic mutations, the effects of linked mutations can be separated by mutation or recombination given enough time, lowering the long-term maintenance of genetic correlations. Therefore, disentangling the roles of pleiotropy and linkage has important implications for our understanding of adaptation (Barrett and Hoekstra, 2011).

RAPID AND REPEATED FRESHWATER COLONIZATION BY THREESPINE STICKLEBACK

The threespine stickleback (*Gasterosteus aculeatus*) has become a model system for studying the genetic basis of adaptive evolution (Kingsley and Peichel, 2007; Peichel and Marques, 2017). Across its holarctic range, marine stickleback have repeatedly invaded and adapted to freshwater environments, which has resulted in the parallel evolution of many traits. For example, most

freshwater populations have reduced defensive bony armor, including bony lateral plates and spines, and changes in feeding morphology, including reduction of gill raker length and number (Hagen and Gilbertson, 1972; Hendry *et al.*, 2013). This parallel phenotypic evolution is mirrored at the genetic level by parallel fixation of shared freshwater haplotypes (Hohenlohe *et al.*, 2010; Jones *et al.*, 2012b; Terekhanova *et al.*, 2014; Bassham *et al.*, 2018). These freshwater haplotypes exist as standing genetic variation in marine populations that persists through ongoing migration of alleles to and from freshwater (Colosimo *et al.*, 2005; Hohenlohe *et al.*, 2010; Jones *et al.*, 2012b; Nelson and Cresko, 2018). Importantly, these freshwater haplotypes often overlap with known quantitative trait loci (QTL) for traits that differ between marine and freshwater sticklebacks, suggesting that they are key for the rapid, parallel evolution of many traits in freshwater (Hohenlohe *et al.*, 2010; Jones *et al.*, 2012b).

THE ROLE OF THE *EDA* GENE AS A MAJOR PLEIOTROPIC GENE IN FRESHWATER ADAPTATION

One of the strongest molecular signals of divergent selection between marine and freshwater stickleback genomes is in a 16 kb region on chromosome IV (Jones *et al.*, 2012b). This 16 kb region has a number of fixed sequence differences between the marine and freshwater haplotypes and contains three protein-coding genes, *Ectodysplasin* (*Eda*), *Tumor necrosis factor super-family member 13b* (*Tnfsf13b*) and *Glycoprotein A repetitions predominant* (*Garp*) (Figure 2.1). The first of these genes, *Eda*, controls at least three traits (number of bony lateral plates, schooling behavior, and lateral line patterning) and resides mostly within the haplotype (exon 1 and the 5' UTR are outside the minimally shared freshwater haplotype) (Colosimo *et al.*, 2004; Colosimo *et al.*, 2005; Wark *et al.*, 2012; Greenwood *et al.*, 2013; Mills *et al.*, 2014; Greenwood *et al.*, 2016). *Eda* is required for the development of epithelial appendages in vertebrates, including hair, teeth, scales, and lateral plates (Srivastava *et al.*, 1997; Harris *et al.*, 2008; Aman *et al.*, 2018; Wucherpfennig

et al., 2019). Protein sequence comparison and tissue-specific expression level analysis suggest that the phenotypic differences driven by *Eda* are due to reduced expression of the freshwater allele (Colosimo *et al.*, 2005; O'Brown *et al.*, 2015). Consistent with this hypothesis, overexpression of *Eda* in freshwater sticklebacks partially recovers the ancestral marine phenotypes for number of lateral plates, schooling behavior and lateral line patterning, confirming the pleiotropic role of *Eda* in phenotypic evolution of freshwater stickleback (Colosimo *et al.*, 2005; Mills *et al.*, 2014; Greenwood *et al.*, 2016).

Despite clear evidence for the role of *Eda* in these phenotypes, the causative mutation(s) are unknown. Compared to marine fish, there are four coding changes in *Eda* shared by North American low-plated stickleback. However, a low-plated Japanese population (NAKA stream population) lacking the North American amino acid changes fails to complement for lateral plate formation, suggesting that plate loss is due to regulatory, not coding, changes (Colosimo *et al.*, 2005). Consistent with this model, allele-specific expression analysis revealed a *cis*-regulatory downregulation of *Eda* in F1 hybrid flank tissue (O'Brown *et al.*, 2015). This study also identified a regulatory element 3' of the *Eda* gene containing a single base-pair mutation that is homozygous and shared by all low-plated, freshwater stickleback populations, including the NAKA Japanese population (O'Brown *et al.*, 2015). Hereafter, we will refer to this mutation as the NAKA SNP (single nucleotide polymorphism). Rare, wild-caught marine fish heterozygous at the NAKA SNP are completely-plated (O'Brown *et al.*, 2015), yet it remains unknown whether this single SNP is necessary or sufficient in the homozygous state to cause plate loss, or if other mutations are involved. It is also unknown whether lateral plate, lateral line patterning, and schooling behavior changes are caused by a single pleiotropic mutation or by linked mutations.

In addition to the three known phenotypes affected by *Eda*, chromosome IV harbors QTL for more traits than expected by chance (Albert *et al.*, 2008; Miller *et al.*, 2014; Peichel and Marques, 2017). Many of these QTL have confidence intervals that overlap with the *Eda* haplotype. This 'phenotypic hotspot' is therefore an interesting and important test case for disentangling the relative roles of pleiotropy and linkage during adaptive evolution. Here, we addressed two complementary questions: 1) how many and which traits map to the 16 kb freshwater haplotype on a marine genomic background; and 2) are these phenotypes affected by the same pleiotropic mutation, or by separate tightly linked mutations? To address the first question, we identified wild-caught marine fish carrying the freshwater *Eda* haplotype, made crosses between these heterozygous fish, and phenotyped and genotyped the offspring to determine which of the traits previously associated with QTL overlapping the *Eda* haplotype are specifically associated with this 16 kb haplotype. To address the second question, we first tested whether the previously identified regulatory NAKA SNP is sufficient to generate the freshwater lateral plate and lateral line phenotypes in a marine genomic background. We also performed association mapping of these lateral plate and lateral line traits, as well as other phenotypes associated with QTL overlapping *Eda*, in a polymorphic, interbreeding freshwater population of stickleback in which recombination has occurred between the marine and freshwater *Eda* haplotypes.

2.3 Materials and Methods

Ethics statement

Fish were collected under the Washington Department of Fish and Wildlife scientific collection permits 14-311b, 15-033, and 16-066. Animal care and handling protocols were approved by the Fred Hutchinson Cancer Research Center Institutional Animal Care and Use Committee (protocol

1575) or the Veterinary Service of the Department of Agriculture and Nature of the Canton of Bern (VTHa# BE4/16).

PUGET SOUND FISH COLLECTIONS, CROSSES, AND CARE

In the summer of 2015, marine fish were caught during a multi-day trawling trip in the Whidbey Basin and Bellingham Bay areas of Puget Sound, WA. In June 2016, marine fish were collected in a beach seine in full saltwater in Clam Bay, near Manchester, WA, with the help of the Washington Department of Fish and Wildlife. Fish were transported and housed in the stickleback facility at the Fred Hutchinson Cancer Research Center. Animals were kept in standard 29 gallon aquarium tanks, each aerated with an air stone and containing 0.35% saltwater (3.7gL^{-1} Instant Ocean sea salt, Instant Ocean Spectrum Brands, Inc., Blacksburg, VA, USA; 0.003gL^{-1} NaHCO3; 0.0003gL^{-1} $Ca(OH)_2$). Water was filtered through an external filter (AquaClear Power Filter, Rolf C. Hagan Inc., Baie d'Urfé, Quebec, Canada). The rooms were programmed to mimic summer light conditions (16h light: 8h dark) and maintained at approximately 17.5˚C. Fish were kept at a density of up to 24 adults per tank and were fed live brine shrimp nauplii in the mornings and frozen *Mysis* shrimp in the afternoons.

In both 2015 and 2016, fish were marked with a combination of spine clipping and elastomer tagging, fin clipped, and genotyped following a HotSHOT DNA extraction (see "DNA EXTRACTIONS" section below). In 2015, fish were first genotyped at the NAKA SNP, an intergenic SNP 3' of the *Eda* gene previously suggested to be involved in *Eda* regulation and plate reduction (O'Brown *et al.*, 2015), and heterozygotes were then genotyped at Stn382, an indel polymorphism within the first intron of *Eda* that differentiates most marine from freshwater haplotypes (Colosimo *et al.*, 2005)(Figure 2.1, Figure 2.2, Table 2.1). Four crosses were made between fish that were heterozygous carriers of the freshwater allele at the NAKA SNP but that were homozygous for the

marine allele at Stn382. One cross was made between a homozygous marine fish and the single heterozygous carrier of freshwater alleles at both the NAKA SNP and Stn382. Given the higher frequency of NAKA heterozygotes to Stn382 heterozygotes in the 2015 sample, we reversed our genotyping strategy in 2016. Wild-caught fish from 2016 were first genotyped at the marker Stn382 (Figure 2.1, Table 2.1) to find heterozygous carriers of the freshwater *Eda* haplotype. Eleven crosses were made between fish that were heterozygous carriers of the freshwater allele at Stn382. Later genotyping (see "GENOTYPING ASSAYS" section below) revealed that five of these were crosses between carriers of the full freshwater *Eda* haplotype, five were crosses between carriers of a short haplotype with a carrier of a full haplotype, and one cross was between carriers of different short haplotypes (Figure 2.3). Fish were considered heterozygous carriers of full haplotypes if they were heterozygous for all markers tested between SNP1 through SNP13, except the NAKA SNP. In contrast, short haplotypes varied in size, extending from SNP1 through somewhere between the NAKA SNP to SNP 11 (Figure 2.2). In addition, an F2 cross was made between heterozygous F1 siblings from the single *Eda* full haplotype by homozygous F0 marine cross.

Offspring of these crosses were reared in the stickleback facility at the Fred Hutchinson Cancer Research Center under conditions similar to those for wild-caught adults described above. The differences were that entire crosses were housed in a single tank and young fish were fed brine shrimp nauplii twice per day. The offspring of the 2016 crosses were shipped to the stickleback facility at the University of Bern, Switzerland, between the ages of 2 and 5 months (corresponding to 18 - 35 mm standard length). Animals were housed in 100-liter tanks on a recirculating system. Conductivity (5.3 millisiemens/cm) and pH (7.5) were automatically monitored and maintained using saturated solutions of Instant Ocean sea salt (Instant Ocean Spectrum Brands, Inc.,

Blacksburg, VA, USA) and sodium bicarbonate (75gL^{-1}). Lighting was programmed with 11 hours full sunlight, 1 hour sunrise, 1 hour sunset, and a moon light for nighttime. Water temperature was maintained near 15˚C. Crosses were housed in a single tank until they reached approximately 30 mm, and then they were split among multiple tanks to maintain approximately 45 - 55 fish per tank. Young fish were fed brine shrimp nauplii twice per day, and adult fish were fed brine shrimp in the mornings and frozen *Mysis* shrimp three times per week in the afternoon. Fish were grown to about 2 years old before genotyping and phenotyping.

LAKE WASHINGTON FISH COLLECTIONS AND CARE

Lake Washington is a large freshwater lake near Seattle, WA that contains both completely- and low-plated sticklebacks, and genotype at Stn382 in the *Eda* haplotype explains 75.2% of the variation in plate phenotype in this population (Kitano *et al.*, 2008). Adult fish were collected from different locations around Lake Washington between April 2015 and March 2016. Unbaited minnow traps were used to catch fish (n=52) at Mercer Slough south of Bellevue, WA. A Merwin trap was set in Kenmore, WA, on the northern edge of Lake Washington (n=136). Nighttime purse seining (n=129) and trawling (n=560) were conducted in the northern half of Lake Washington. Nearshore trapping (minnow and Merwin traps) was conducted in May, June and July when adult sticklebacks are in nearshore habitat. Seining and trawling were used in October and March, respectively, when adult sticklebacks are found offshore. We sampled fish from a subset of locations around the lake because collections from 2005 (Kitano *et al.* (2008): locations 2, 4, 7, 9) demonstrated that while fish from different locations have varying degrees of marine ancestry, there is no correlation between genotypes at neutral markers and plate phenotype. Based on this previous work, we expected low and completely plated fish to be interbreeding, and the frequency of the freshwater *Eda* allele to be ~40%.

Animals were housed in the stickleback facility at the Fred Hutchinson Cancer Research Center as described above for Puget Sound wild-caught fish. Fish were kept in the lab for between 0 and 57 days before phenotyping.

PHENOTYPING

To disentangle the relative roles of pleiotropy and linkage within the highly divergent marine-freshwater haplotype on chromosome IV, we focused on phenotypes that had previously mapped to chromosome IV in QTL crosses between marine and freshwater stickleback (Peichel and Marques, 2017). To fully analyze the geometric morphometric landmarks and lateral line neuromast counts (see below), we also included some specific traits in these categories that were not previously mapped to chromosome IV (Table 2.2). For logistical reasons, we did not measure schooling behavior, which has previously been associated with variation in *Eda* (Greenwood *et al.*, 2016).

PHENOTYPING: LATERAL LINE NEUROMASTS

The Puget Sound NAKA SNP crosses (n=86 individuals) and a subset of Lake Washington fish (244 of 768) were stained to visualize lateral line neuromasts using the fluorescent vital dye 2-(4-(dimethylamino)styrl)-N-ethylpyridinium iodide (DASPEI; Invitrogen/Molecular Probes, Carlsbad, CA) as described in Wark et al. (2012). Neuromasts were visualized and counted as described below using a Leica dissecting scope with fluorescent light and a FITC filter set (Leica Microsystems Inc., Banncokburn, IL, USA). After staining, fish were euthanized with a lethal dose of MS-222, fins were clipped and saved in 95% ethanol for DNA extraction, and fish were placed in a T-Sac Tea Filter Bag (Magic Teafit, Columbus, OH, USA) with a waterproof, unique ID tag.

Fish were then stored in Mason jars in 10% buffered formalin for at least one week before staining with Alizarin red to visualize bony structures, as described by Peichel *et al.* (2001).

DASPEI staining is time-intensive and variable because each fish is stained and screened live, and the quality of staining can abruptly change if the fish is anesthetized for too long. We therefore adapted a method of neuromast staining using alkaline phosphatase to allow for the bulk preservation, staining and storage of fish, plus reliable and convenient phenotyping (adapted from Villablanca *et al.* (2006)). This method was used to phenotype the Puget Sound *Eda* haplotype crosses (n=498) and the remaining Lake Washington fish (524 of 768). Fish were first placed in aerated epinephrine for 10 minutes to contract their melanophores (0.07 gL^{-1} epinephrine in fish water). Fish were transferred to a lethal dose of MS-222 and left until all operculum movement had stopped for 4 minutes. Standard length was recorded, and fins were clipped and placed in 95% ethanol for DNA extraction. Fish bodies were then placed with a waterproof ID tag in custom staining chambers (Figure 2.4), and submerged in freshly made, cold, 4% paraformaldehyde. Fish were kept at 4˚C for 24 - 36 hours, then rinsed 3 times for 10 minutes in PBS and returned to 4˚C for 12 -72 hours. On the day of neuromast staining, fresh coloration buffer was made (0.1M NaCl; 0.1M Tris-HCl, pH 9.5; 0.05M $MgCl_2$; 0.1% Tween-20 in reverse osmosis water). Fish were rinsed in coloration buffer for 20 minutes, then submerged in half-strength NBT/BCIP solution (0.225% NBT, 0.175% BCIP in coloration buffer; NBT/BCIP from Promega Corporation, Madison, WI, USA). Staining was carried out in the dark for 3 - 4 hours on a rotational shaker. The preopercular and infraorbital lateral lines were checked periodically on a dissecting microscope to assess staining, as these lines are easy to find and often stain well. High quality staining produces a circle around the outer edge of the neuromast, and sometimes a line or dot in the middle of the neuromast (Figure 2.4). Once this pattern is dark, staining can be terminated by a 10 minute rinse in PBS,

followed by two methanol rinses to intensify staining (30 minutes in 25:75 methanol:PBTw (0.1% Tween in PBS), 20 minutes in 43:57 methanol:PBTw). Fish were then rinsed twice for 10 minutes in PBS and could be stored at 4˚C for up to 3 days in PBS. Fish could be screened at this point, but we found it easier to screen them after Alizarin red staining. Fish were transferred to tea bags and stored in 10% buffered formalin at room temperature before Alizarin red staining as described previously (Peichel *et al.*, 2001). Fish stained with alkaline phosphatase/Alizarin red were submerged in water and gently brushed with a paintbrush to remove background alkaline phosphatase stain on the bones surrounding the neuromasts. Neuromasts were viewed and counted on a Leica dissecting scope using visible light as described below (Leica Microsystems Inc., Bannockburn, IL, USA).

Neuromasts were counted in the 12 stickleback lateral lines, as described previously (Wark and Peichel, 2010; Wark *et al.*, 2012) (Figure 2.4, Table 2.2). In addition to counting the neuromasts in the anterior and posterior main trunk lateral lines (Ma and Mp, respectively), the neuromast pattern was drawn for each plate and/or body segment. These drawings were later used to count neuromasts in the Ma and Mp lines and to quantify the dorsal-ventral (D-V) neuromast patterning. Since staining quality varied along the anterior-posterior axis and some segments lacked neuromasts due to poor staining quality, the anterior and posterior trunk line (Ma and Mp) neuromast counts were averaged per body segment with stained neuromasts. Additionally, each body segment of the Mp line was phenotyped as having a midline neuromast pattern (with neuromasts lying directly on the midline) or a dorsal-ventral neuromast pattern with at least one neuromast off the midline. This trait is reported as the fraction of segments displaying the dorsal-ventral pattern (D-V patterned segments).

PHENOTYPING: METRICS AND MERISTICS

Meristic and metric traits were scored under a Leica dissecting scope. Meristic traits included plate count on the left side of the fish, anal fin ray number and dorsal fin ray number (Table 2.2). Metric traits were measured using S_Cal PRO IP67 digital calipers (Sylvac SA, Crissier, Switzerland), and included standard length, pelvic girdle length, ectocorocoid length, spine lengths (1st, 2nd and 3rd dorsal spines; anal spine; and left pelvic spine), middle gill raker length, and supraoccipital notch length (Table 2.2). To expose the gill rakers for measurement, the branchiostegal rays were cut and the operculum was peeled up to expose the branchial bones and gill rakers. The length of the middle anterior facing gill raker on the first gill arch was measured with the digital calipers.

PHENOTYPING: BRANCHIAL TRAITS

The dissections and mounting of branchial skeletons were done following the protocol of Ellis and Miller (2016). Three categories of branchial traits were phenotyped following Miller *et al.* (2014): pharyngeal tooth counts, gill raker counts and branchial bone lengths (Table 2.2). The counting of pharyngeal teeth on the right two dorsal and single ventral toothplates was done under a binocular microscope using a black background to increase contrast. Teeth were counted between two and eight times by the same researcher to improve precision. Anterior-facing gill rakers were counted on the first left branchial arch following Miller *et al.* (2014). The fourth ceratobranchial bone on the left side was measured using digital calipers on the mounted branchial skeleton.

PHENOTYPING: MORPHOMETRICS

Fish were pinned to a dissecting mat for consistency between fish. Insect pins were inserted at landmarks 2, 8, 23, and 29 corresponding to the posterior insertion of the last anal fin ray, the anterior edge of the ecotocoracoid, the supraoccipital notch, and the posterior insertion of the last

dorsal fin ray, respectively (Figure 2.5). Photos were taken of each fish with a 1 cm reference scale in the picture. Photos were converted into .tps files using the *tps.util* program. Twenty-nine landmarks were placed on each fish using *tps.dig2* (Figure 2.5, Table 2.2). If landmarks could not be placed without doubt, placeholder landmarks were used and these fish were removed from the data before the morphometric analysis. Landmark coordinates were scaled and rotated using the *geomorph* package in R (https://cran.r-project.org/package=geomorph), and the scaled and rotated X and Y coordinate values were used as trait values in further analyses. Two linear measurements were extracted from the scaled landmark positions: body depth and maxilla length. Body depth was calculated as the distance between landmarks 6 and 25, and maxilla length as the distance between landmarks 18 and 16 (Figure 2.5).

DNA EXTRACTIONS

DNA from fin tissue of wild-caught Puget Sound fish was extracted using a modified HotSHOT DNA extraction method (Meeker *et al.*, 2007). Fin tissue was placed directly into 30uL of NaOH, digested at 95°C for 30 minutes, cooled on ice and then pH neutralized with 8uL of 1mM TrisHCl. Samples were spun down for 10 minutes and 0.5uL of the supernatant was used in 5 or 10uL PCR reactions.

Fin tissue from Puget Sound crosses and Lake Washington fish were kept in 95% ethanol for DNA extractions. DNA was either extracted following a standard phenol-chloroform protocol or using the Wizard® SV 96 Genomic DNA Purification System (Promega, Madison, WI, USA). Following resuspension in TE buffer, DNA was quantified on a NanoDrop and diluted to 10 ng/uL in water for genotyping assays.

GENOTYPING ASSAYS

Wild-caught Puget Sound and Lake Washington fish were initially genotyped at a subset of known SNPs and indels across the 16 kb region (Stn382, SNP5, NAKA SNP, SNPs10/11, and SNPs12-16), and the Puget Sound fish were additionally genotyped at two flanking markers outside the 16 kb region (Cnv767 and SNP19, ~ 10 kb 5' and 3' of the 16 kb region, respectively) (Colosimo *et al.*, 2005; Lowe *et al.*, 2017) (Table 2.1). Fish were also genotyped for a marker in the 3' UTR of the isocitrate dehydrogenase gene (*Idh*) that distinguishes males from females (Peichel *et al.*, 2004). Puget Sound F1 offspring from parents carrying various-sized freshwater *Eda* haplotypes were genotyped at 2 markers: Stn382 and SNPs12/13 (Table 2.1). This allowed us to distinguish between offspring carrying a long haplotype from those carrying a short haplotype. Puget Sound F1 offspring of parents heterozygous only at the NAKA SNP were genotyped only at the NAKA SNP since there was no known variation at any other marker in the haplotype in the F0 parents.

Following initial genotyping at this subset of known polymorphisms, additional genotyping was performed on recombinant Lake Washington fish to fine-map their recombination breakpoints to increase mapping resolution. Both known and novel polymorphisms were used in fine-mapping (SNP1, LP3621, Cnv770, Stn381, SNP7, LP13173, SNP8) (Table 2.1). Novel putative shared polymorphisms were selected using 3 datasets. First, published genomic and BAC sequences from one saltwater (Salmon River marine fish, GenBank: AC144489.2) and two freshwater individuals (Paxton benthic and Bear Paw fish, GenBank: AY897589.1 and AANH00000000.1, respectively) were aligned in Geneious version R11.1.5 (http://www.geneious.com), and putative SNPs and indels were identified. Next, the sequencing data from Jones *et al.* (2012b) was used to confirm the usefulness of SNPs or indels across multiple freshwater and marine populations because it contains low-coverage whole genome sequences from 10 paired freshwater and saltwater

populations (http://sticklebrowser.stanford.edu/). A putative SNP was confirmed if the same freshwater allele was found in all freshwater populations with data and not in any marine populations. We also looked at insertions or deletions (indels) on the freshwater haplotype, as these may play an important role in phenotypic evolution (Lowe *et al.*, 2017). The raw reads from the Jones dataset were used to confirm putative indels. For each putative indel, marine and freshwater bait queries were created from the alignment of published genomes and BAC clones described above. Next, raw sequencing reads from the Jones dataset were separated into marine and freshwater databases and searched for matches to the queries. A putative indel was confirmed if the freshwater query accumulated multiple matches to the freshwater database and the marine query accumulated multiple matches to the marine database. This much reduced set of putative haplotype indels and SNPs were then used to design additional genotyping assays for screening recombinant fish (LP3621, Cnv770 and LP13173) (Table 2.1).

CURATION OF GENOTYPE DATA FOR ASSOCIATION MAPPING

Puget Sound offspring were successfully genotyped at the chosen informative markers for all fish. For the Lake Washington fish, missing genotypes were dealt with in the following manner. Missing genotypes were imputed if the genotypes at both genotyped flanking markers were the same, and no genotyped sample within the dataset had an incongruous genotype at that marker (with the assumption of no double recombination events). In addition, two or three missing genotypes in a row were filled in following the same guidelines as for a single missing marker. Figure 2.6 shows the genotypes at each marker (with missing data in white) prior to imputation. The following numbers for each marker are the numbers of successfully genotyped samples / imputed genotypes SNP1: 884/0, Stn382: 883/2, LP3621: 860/18, Cnv770: 852/0, SNP5: 874/0, Stn381: 863/21, NAKA SNP: 884/0, SNP7: 386/498, LP13173: 838/46, SNP8: 706/178,

SNPs10/11: 877/7, SNP13: 885/0. Our method of imputing missing data could lead to incorrect genotype information; therefore, a second trimmed genotype file was created in which all missing genotypes were deleted. This mostly involved deleting individual fish from the analysis, however all genotypes for SNP7 and SNP8 were deleted since they had missing data for over half of the fish. We do not report the association mapping results using the filled genotype file here because it produced qualitatively the same results but is less conservative.

ANALYSES

Puget Sound crosses and Lake Washington wild-caught fish were analyzed in largely the same way and differences will be noted below. Outliers were assessed in two ways and then removed. First, prior to analyses, data were explored by eye for categories of fish that should be removed. For example, for lateral line traits, fish with poor staining quality (assessed subjectively during phenotyping on a scale of 1 - 7) were removed from analyses (n= 161 in Lake Washington; n=86 in Puget Sound). For non-lateral line traits: the following samples were removed from Lake Washington analyses: June and July males (n=6 and n=3, respectively), October samples larger than 65 mm (n=9; these fish are likely a year older than all other fish, since they are bigger than reproductive summer fish, and much bigger than October fish), May fish less than 50 mm (n=1). Second, trait values >4 standard deviations from the mean trait value were also removed prior to analysis of the Lake Washington and Puget Sound datasets (between 0 - 2 samples per trait were removed totaling 30 trait values across all analyses). The number of fish analyzed for each phenotype and dataset is provided in Table 2.3, Table 2.4, and Table 2.5.

To assess whether genotype is significantly associated with phenotype in the Puget Sound crosses and the wild-caught Lake Washington fish, linear models representing the null and alternate hypotheses were compared. The null model included presumed covariates. The alternate

hypothesis included covariates plus the genotype at the focal marker (i.e. Stn382). The models were compared using both a Chi-square test, and by calculated the log-odds likelihood ratio (LOD) of the two models following the formula:

$$LOD = \frac{n}{2} * log_{10}\left(\frac{RSS0}{RSS1}\right) \quad (2.1)$$

where n is the number of samples and $RSS0$ and $RSS1$ are the residual sums of squares of the null and alternate models, respectively (Broman and Sen, 2009). We then calculated the percent of phenotypic variation explained (PVE) by each marker using the formula (Broman and Sen, 2009):

$$PVE = 1 - 10^{\left(\frac{-2*LOD}{n}\right)} \quad (2.2)$$

Using this method, we asked how much additional phenotypic variation is explained by genotype at the focal marker, after accounting for possible covariates. Variables that possibly contributed to variation in the raw trait values include biological variables such as sex, standard length, and family; collection variables including month of collection, lab tank housed in, and days kept in tank before staining; and method of staining (DASPEI vs. alkaline phosphatase staining for lateral line neuromasts). Therefore, covariates in the Puget Sound cross analyses included family, standard length, sex, and interactions between standard length and family, and sex and family. Covariates in the Lake Washington fine-mapping varied by trait. Geometric morphometric data was already corrected for size, so covariates were sex and month of collection, plus their interaction. Meristic and metric traits were corrected for size, sex and month of collection, plus interactions between sex and month and between size and month. These interactions allow for changes between sex or size and trait value that could change across life stages, such as pre-reproductive and reproductive fish collected in October and June, respectively. Since lateral line traits were correlated with staining quality, which was susceptible to holding tank and staining method, covariates were sex, size, month of collection, tank, staining method, days held in tank,

tank, and staining quality. Interactions between sex and month, and size and month were also included.

Significance of LOD values were assessed via permutation testing. For the Puget Sound crosses, we stratified the data for the permutation testing to account for possible differences between the size of the haplotype in each cross. Genotypes were only permuted within families with roughly the same haplotype sizes. The Lake Washington data was stratified by collection month, to adjust for non-random differences in allele frequencies between collections. 5000 permutations were performed. P-values for each LOD score were calculated using these permutations, and the p-value threshold was adjusted for multiple comparisons using a Bonferroni correction for number of traits tested (0.05/ # of traits analyzed).

2.4 RESULTS

FRESHWATER HAPLOTYPE SIZE IS VARIABLE IN MARINE FISH

Genotyping of marine fish revealed that some fish in the Puget Sound marine population are heterozygous carriers of freshwater *Eda* haplotypes (Figure 2.2). Here, freshwater alleles are designated "L", due to their association with the low-plated phenotype, while alleles typically shared by marine fish are designated as "C" for completely-plated (Barrett *et al.*, 2008). Marine fish from 2016 were initially genotyped at Stn382, a marker that differentiates the C and L *Eda* haplotypes (Colosimo *et al.*, 2005; O'Brown *et al.*, 2015). The frequency of the L allele at Stn382 was 4.4% in 2016 (n = 86 of 1938 chromosomes). Further genotyping across the haplotype region revealed that most fish carried the full 16 kb freshwater haplotype, but some fish carried a shorter version (47 "full L" haplotypes, 39 "short L" haplotypes; Figure 2.2). Similar genotyping of the 2015 Puget Sound fish found frequencies of the L allele at Stn382 to be 0.16% and at the NAKA SNP to be 9.4% (n = 1 and 59 of 628 chromosomes for Stn382 and NAKA, respectively). These

shorter versions of the haplotype varied in size from nearly the full 16 kb region to somewhere between 10 - 12 kb, down to possibly just a single base pair at the NAKA SNP.

FOUR TRAITS MAP TO THE *EDA* HAPLOTYPE IN THE PUGET SOUND MARINE CROSSES

In the offspring of crosses between heterozygous carriers of the *Eda* full and short freshwater haplotypes, genotype at the *Eda* haplotype was significantly associated with four of the 17 measured phenotypes: left plate count (LOD = 344, PVE = 93.6), dorsal-ventral (D-V) neuromast patterning (LOD = 292, PVE = 93.3), Mp neuromasts per segment (LOD = 77.1, PVE = 51.0), and gill raker length (LOD = 4.65, PVE = 3.68) (Figure 2.3, Table 2.3). The C allele is dominant for these traits; heterozygous CL and homozygous CC fish have similar phenotypes and differ from LL fish. No other measured traits that have been previously mapped to QTL on chromosome IV were significantly associated with the *Eda* haplotype in these crosses, including branchial skeletal traits, dorsal and pelvic spine lengths, pelvic girdle length, or ectocoracoid length.

Crosses between heterozygous carriers of different-sized freshwater haplotypes allowed us to assess which part of the haplotype contains the causative variants by comparing the average phenotypic residuals of offspring with different combinations of haplotypes. We found that gill raker length is not significantly associated with Stn382, but is associated with SNPs12/13 (LOD = 4.65, Figure 2.3, Figure 2.7, and Table 2.3). In contrast, the variation in plates and the two lateral line traits is more strongly associated with Stn382 than SNPs12/13 (e.g. LOD = 344 at Stn382 versus 114 at SNPs12/13 for plates, Table 2.3).

MULTIPLE MUTATIONS WITHIN THE HAPLOTYPE AFFECT PLATE AND LATERAL LINE TRAITS

As described above, genotype at Stn382 is highly correlated with variation in plate number, neuromast patterning and neuromast number in the Puget Sound marine crosses. However,

additional variation in these three traits is explained by genotype at SNPs12/13 (Figure 2.3). For example, fish that are LL at both Stn382 and SNPs12/13 have significantly fewer D-V patterned segments, Mp neuromasts per segment, and plates than fish that are LL at Stn382 and CL at SNPs12/13. Interestingly, there is no effect of genotype at SNPs12/13 (CC versus CL) on these three traits when the genotype at Stn382 is CL (Figure 2.3). These data suggest that there are at least two mutations within the haplotype that affect these three traits. The smaller effect mutation acts either epistatically with the larger effect mutation (an effect of SNPs12/13 is only observed when the genotype at Stn382 is LL), or the freshwater allele of the mutation near SNPs12/13 is recessive to the marine allele.

THE NAKA SNP IS NOT SUFFICIENT FOR PLATE OR LATERAL LINE PHENOTYPES

We next tested whether fish homozygous for the L allele at the NAKA SNP (contained within the short L haplotypes, Figure 2.2) had significantly different trait values for plate number, as hypothesized by O'Brown *et al.* (2015), or neuromast number or neuromast patterning, given the strong correlation of these traits during development (Mills *et al.*, 2014). In crosses between Puget Sound marine fish heterozygous only at the NAKA SNP, we found no significant effect of genotype on plate count, neuromast number or neuromast pattern (Figure 2.8, Table 2.4). These data are consistent with the finding that the two wild-caught Puget Sound marine fish homozygous for the L allele only at the NAKA SNP were completely-plated, while the three wild-caught Puget Sound fish homozygous for the short L haplotype were low-plated (Figure 2.2).

REDUCED LINKAGE WITHIN THE HAPLOTYPE IN LAKE WASHINGTON STICKLEBACK

In order to determine whether the genotypic correlations of lateral plates and lateral line traits are due to pleiotropic effects of a causative mutation(s) or due to tight genetic linkage of multiple

causative mutations, we performed fine-mapping in a polymorphic, interbreeding, freshwater population from Lake Washington in Seattle (Kitano *et al.*, 2008). Genotyping of 885 fish within the 16 kb haplotype identified 198 fish (22%) with an historical recombination event within the 16 kb haplotype (Figure 2.6).

ASSOCIATION MAPPING OF TRAITS IN LAKE WASHINGTON STICKLEBACK

In the Lake Washington mapping population, we again found that left plate count (LOD = 218, PVE = 76.4), dorsal-ventral (D-V) neuromast patterning (LOD = 113, PVE = 67.9), and Mp neuromasts per segment mapped to the *Eda* haplotype (LOD = 30.7, PVE = 26.5)(Figure 2.9, Table 2.5). In this population, we did not recover a significant correlation with gill raker length after correction for multiple comparisons (LOD = 1.17, PVE = 0.78). Some additional traits show an association but do not survive multiple correction, including anal spine length (LOD = 2.68, PVE = 1.77), left pelvic spine length (LOD = 1.94, PVE = 1.28) and pelvic girdle length (LOD = 2.75, PVE = 1.80). None of the other branchial, lateral line, metric or meristic traits previously mapped to chromosome IV and measured in the Lake Washington fish show an association with the *Eda* haplotype (Table 2.2 and Table 2.5).

We also performed fine-mapping for geometric morphometric traits in Lake Washington. This analysis revealed that landmarks Y10, X16 and Y18 are associated with the *Eda* haplotype (Figure 2.10, Table 2.5). Landmark 10 is the dorsal posterior extent of the operculum, and the X position has previously been mapped to near *Eda* (Albert *et al.*, 2008). Landmarks 18 and 16 mark the anterior and posterior extent of the maxilla, respectively. The Y position of landmark 16 has previously been mapped to chromosome IV (Albert *et al.*, 2008).

LARGE EFFECT TRAITS MAP TO THE SAME INTRAGENIC REGION OF *EDA*

We find that the phenotypes that map strongly to *Eda* all have the same association pattern with the markers across the 16 kb haplotype (Figure 2.9). The strongest associations are with indels LP3621 and Cnv770, which are located in the first intron of *Eda* (Figure 2.1). These data suggest that one or more closely linked mutations in a 1.4 kb region mediate the effects of the *Eda* haplotype on left plate count, D-V neuromast patterning, and Mp neuromasts per segment. Notably, the lowest LOD score for all three traits is at the NAKA SNP, consistent with the fact that the NAKA SNP is not sufficient to affect plate count or lateral line phenotypes in the Puget Sound crosses (Figure 2.8).

MULTIPLE MUTATIONS WITHIN THE HAPLOTYPE AFFECT NUMBER OF LATERAL PLATES

Mapping of lateral plate count revealed a second LOD score peak at SNP7, which has a higher LOD score than both Stn381 and LP13173 (Figure 2.9). Together with the finding that Puget Sound F1 offspring with the full L haplotype have fewer plates than F1 offspring with the short L haplotype (Figure 2.3), we hypothesized that there may be more than one mutation within the haplotype that affects plate number. To test this idea, we plotted adjusted left plate counts for the Lake Washington fish by genotype at Cnv770 (highest LOD score in the association mapping) and SNPs12/13 (to be consistent with the analysis of the Puget Sound crosses). We found that genotype at SNPs12/13 explains additional variation in adjusted plate number within fish that are heterozygous at Cnv770 (Figure 2.11). Although the phenotypic values for neuromast pattern and neuromast number trended in the same direction as the pattern described for lateral plates, significant effects of SNPs12/13 were only observed on neuromast pattern.

2.5 DISCUSSION

The goal of this study was to determine whether a single, highly adaptive haplotype within a phenotypic hotspot is responsible for multiple phenotypic changes, or whether the overlap of these phenotypic and genomic divergence signals is due to the low resolution of QTL mapping. Further, we asked whether the pleiotropic effects of the *Eda* gene are due to pleiotropic effects of a single mutation, or due to linked causative mutations within the adaptive haplotype. First, we found that the haplotype is responsible for two traits previously known to be affected by *Eda* (lateral plate count and neuromast patterning), plus the number of neuromasts per body segment. We did not find evidence that the haplotype explains significant variation in additional phenotypes that have previously mapped to chromosome IV. Second, we found evidence for a small 1.4 kb region with large effects on both neuromast and bony armor traits, as well as additional linked, small effect mutations that affect bony armor, feeding and shape traits.

PLEIOTROPIC EFFECTS OF A 1.4 KB REGION ON PLATE AND LATERAL LINE TRAITS

Both the Puget Sound and Lake Washington datasets suggest that variation in lateral bony plate number, dorsal-ventral patterning and number of neuromasts per segment in the main posterior lateral line map to a small genomic region with large effects. These effects are therefore due to the shared freshwater haplotype, rather than population specific alleles or gene by environment interactions. The strong similarity of the LOD curves for these traits suggests that the pleiotropic effects of a single, or few tightly linked mutations at or between LP3621 and Cnv770 are responsible. This is consistent with the Puget Sound data that also show that these traits map to the "short L" haplotype, which includes LP3621 and Cnv770.

The pleiotropic effects of this region may be mediated either via: 1) clustered mutations that alter distinct but closely linked regulatory regions that independently control armor plate formation

and neuromast number and patterning; 2) direct interactions between the developing tissues/cell types responsible for the different traits; or 3) via independent responses by the relevant tissues to the same external signal. Testing the first hypothesis will require further resolution of the sequence changes responsible for armor plate and neuromast changes in sticklebacks. We favor the latter two possibilities because plate development and neuromast patterning are tightly correlated in space and time during stickleback development. Armor plates form around neuromasts, and neuromasts display the dorsal-ventral arrangement only on body segments with plates (Wark and Peichel, 2010; Wark *et al.*, 2012; Mills *et al.*, 2014). In support of the second hypothesis, work in zebrafish has shown that neuromasts and dermal bone, such as the opercular bone and scales, interact (Wada *et al.*, 2010; Wada *et al.*, 2014). However, the interaction between lateral scale formation (thought to be homologous to lateral plate formation in stickleback) and patterning or number of neuromasts has not been investigated. In support of the third hypothesis, Wnt signaling is required for neuromast proliferation and scale formation in zebrafish (Wada *et al.*, 2013; Lush and Piotrowski, 2014; Aman *et al.*, 2018), as well as for normal plate patterning in sticklebacks (O'Brown *et al.*, 2015). *Eda* is required for refinement of Wnt expression prior to scale formation (Aman *et al.*, 2018), but its role in neuromast patterning or proliferation has not been investigated. Therefore, lateral plate and neuromast traits could be responding independently to the modification of the same external (Wnt) signal caused by a reduction in *Eda* expression in freshwater sticklebacks. Studying neuromast patterning in *Eda* mutants or studying the interactions between scale/plate formation and neuromast patterning in species with natural variation in these traits may be informative for disentangling these two hypotheses.

Few adaptive pleiotropic mutations are known (Sabeti *et al.*, 2007; Kamberov *et al.*, 2013; Nagy *et al.*, 2018; Butelli *et al.*, 2019; Ramaekers *et al.*, 2019), so identifying the particular

mutation(s) behind adaptive plate loss and neuromast changes will be an informative endeavor. Previous sequence conservation analysis and allele-specific expression work suggested that the causative mutation for plate loss is a regulatory mutation that drives higher expression of the marine *Eda* allele in the flank of the fish, but not in other body regions, such as the midline fins (Colosimo *et al.*, 2005; O'Brown *et al.*, 2015). This suggests that the causative mutation is in a tissue-specific enhancer of *Eda*. A single SNP that is shared between North American and Japanese freshwater, low-plated stickleback populations, the NAKA SNP, is within a flank-specific enhancer and displayed reduced responsiveness to Wnt signaling compared to the marine enhancer (O'Brown *et al.*, 2015). We found the NAKA SNP in the Puget Sound marine population at a frequency of 9.4%. However, our crosses between heterozygous carriers of the NAKA SNP demonstrated that the NAKA SNP is not sufficient to cause phenotypic changes in plate number or neuromast patterning (Figure 2.8). Reinforcing this finding, the NAKA SNP had the lowest LOD score within the haplotype for plate count and neuromast patterning in our association mapping study within Lake Washington (Figure 2.9). Recent population sampling in Japan confirms that the freshwater allele of the NAKA SNP is found in all sampled freshwater populations with the low-plated phenotype, but also occurs in the closely-related marine species *G. nipponicus*, which is completely-plated (Yamasaki *et al.*, 2019). In addition, the freshwater allele of the NAKA SNP was found in multiple percomorph fish taxa that have scales, including tilapia and platyfish (Figure 2.12), suggesting that the marine allele of the SNP is derived. Thus, the functional consequences of the NAKA SNP remain a mystery.

In addition to ruling out the NAKA SNP as the causative mutation for plate loss and neuromast number and patterning, we were able to narrow the putative causative region to intron 1 of *Eda*, which is consistent with the causative mutation being in a tissue-specific enhancer of *Eda*. The

LOD curves from the association mapping study suggest that the causative mutation(s) are conservatively between Stn382 and SNP5, corresponding to bases 12,802,847 - 12,808,303 in the gasAcu1 genome assembly. Additionally, alignment of all low-plated Lake Washington fish has identified a smaller putative causative region that includes approximately 19 polymorphisms that differentiate the marine and freshwater *Eda* haplotypes and spans a 1,401 bp region (Figure 2.13). This region includes two indels from this study: LP3621 and Cnv770, which encode a 16 bp and a 107 bp deletion in the freshwater and marine haplotypes, respectively (Lowe *et al.*, 2017). Genetic manipulations are now underway to test whether this intronic region contains one or more tissue-specific enhancers active near developing lateral plates or neuromasts, and if so, to test which SNPs or indels may alter the activity of these enhancers and drive the phenotypic differences in plate number, neuromast number and neuromast patterning.

LINKED MUTATIONS IN THE *EDA* HAPLOTYPE HAVE SMALLER EFFECTS ON MULTIPLE TRAITS

In addition to identifying a small 1.4 kb region within intron 1 that causes phenotypic changes in three traits, we found evidence of linked mutations with effects on the same phenotypes as well as additional phenotypes. Data from the Puget Sound crosses suggest that there is a mutation in the downstream portion (~3000 bp) of the haplotype that has a small effect on gill raker length (Figure 2.7). We did not recover this correlation in the Lake Washington mapping population, either due to absence of that mutation, epistatic interactions, or phenotypic plasticity in the wild-caught fish (Day *et al.*, 1994). However, we did find associations in the Lake Washington population between three geometric morphometric landmarks and the *Eda* haplotype. Two of these landmark positions (Y10 and Y18) had very similar LOD curves (Figure 2.10), suggestive of another mutation with pleiotropic effects. Geometric morphometric analysis of the Puget Sound crosses would tell us if these associations hold for the freshwater haplotype in different genetic backgrounds and

populations. Because the PVE of these linked mutations are much lower than the PVE from previous QTL mapping studies (Albert *et al.*, 2008), we do not think we have mapped the previously identified QTL for these traits.

We also found evidence in both the Puget Sound crosses and the Lake Washington mapping population that additional linked mutation(s) within the haplotype have effects on plate count and neuromast patterning. Together with linked mutations affecting morphometric landmarks, this could explain why the 16 kb "full" version of the minimal haplotype is commonly favored and fixed in most low-plated populations. The size of the 16 kb minimal shared freshwater haplotype could be due to either selection in freshwater for the full haplotype, or due to lack of recombination events that make the haplotype smaller. We found that short haplotypes were nearly as abundant as full freshwater haplotypes in the Puget Sound marine population, suggesting that recombination has not limited the availability of shorter haplotypes in anadromous fish (and consequently in freshwater populations). Furthermore, in Lake Washington, where there has been a recent shift in selection towards favoring completely-plated individuals and therefore the marine genotype at *Eda* (Kitano *et al.*, 2008), we find that 22% of individuals have a recombination event within the 16 kb haplotype. These data suggest that recombination is not limiting in this genomic region. Therefore, selection may be preserving the full 16 kb haplotype in typical freshwater environments due to the multiple mutations with phenotypic effects.

SELECTION ON THE *EDA* HAPLOTYPE IN FRESHWATER

Because changes in neuromast number, patterning and plate count all map to the same small genomic interval, we cannot conclude which trait(s) provide a fitness advantage in freshwater. Neuromast number in the posterior lateral line is variable across freshwater populations (Wark and Peichel, 2010; Jiang *et al.*, 2016), suggesting that selection on this trait is not driven by shared

conditions across freshwater environments, and is likely not the main target of selection within the *Eda* haplotype. The fitness effects of the dorsal-ventral patterning of neuromasts are unknown, although it may play a role in schooling behavior (Greenwood *et al.*, 2016). In contrast, much attention has been given to the possible role of plate loss in adaptation to freshwater. Many bony elements are reduced in freshwater, suggesting that an overall reduction in bone may be advantageous (Giles, 1983; Bell *et al.*, 1993; Bell and Foster, 1994; Myhre and Klepaker, 2009). In addition, plates play a known functional role in predation survival, which confirms their visibility to selection in certain environments (Reimchen, 1992). However, while there are known ecological correlates with plate reduction in freshwater environments, such as ion concentration, distance from ocean, and co-occurrence with predators, the selection pressure acting to reduce number of lateral plates is unknown (Bell *et al.*, 1993; Bourgeois *et al.*, 1994; Gelmond *et al.*, 2009). It is also possible that there are pleiotropic effects of this mutation on other, as yet unmeasured phenotypes, which are the direct targets of selection (Rennison *et al.*, 2015).

More broadly our results suggest that the phenotypic hotspot on chromosome IV (Peichel and Marques, 2017) is not explained solely by the *Eda* haplotype. A few traits in previous QTL mapping studies had their highest association within the *Eda* haplotype (Stn382), including ceratobranchial length and dorsal pharyngeal tooth number (Erickson *et al.*, 2014; Miller *et al.*, 2014), and induced coding region mutations in the stickleback *Eda* gene eliminate armor plates and significantly reduce pharyngeal teeth (Wucherpfennig *et al.*, 2019). However, these traits did not map to the *Eda* haplotype in our study. While we cannot rule out a mutation within the freshwater haplotype that is specific to the populations used in the previous QTL crosses (Japan Pacific Ocean and Paxton benthic freshwater fish), our results suggest that additional linked mutations outside, but near, the *Eda* haplotype contribute to other phenotypic traits that map to the

Eda region of chromosome IV. For example, there is a dorsal spine QTL approximately 1 Mb downstream from *Eda* corresponding to the gene *Msx2a* (Howes *et al.*, 2017).

2.6 CONCLUSIONS

Pleiotropic mutations have been proposed as a limitation to adaptation due to their potential deleterious effects on additional traits, yet pleiotropic loci appear to be common. We have measured the phenotypic effects of an adaptive haplotype within a phenotypic hotspot in two separate populations and found that it contains a small genomic region controlling three traits – lateral plate count, neuromast number and neuromast pattern – and additional mutations with small effects on lateral plate count and body shape. We propose that selection favors the entire haplotype in freshwater due to the linkage of multiple mutations and that the multiple phenotypes controlled by this small genomic region facilitates rapid adaptation.

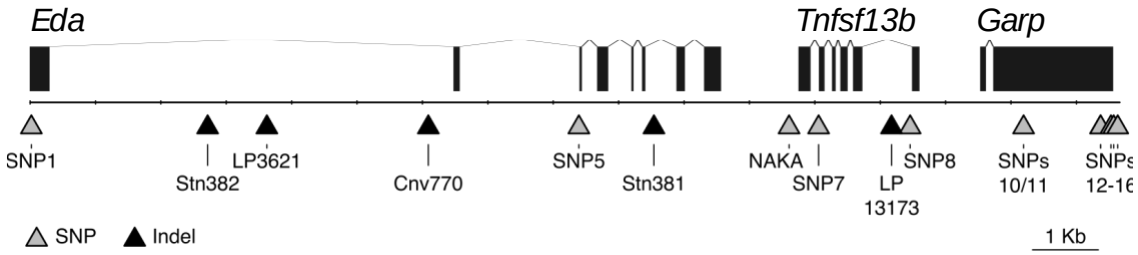


Figure 2.1. Gene structure and genetic markers within the 16 kb *Eda* haplotype. The minimal region shared by most low-plated, freshwater populations of threespine stickleback is 16 kilobases long. It contains three protein coding genes: *Ectodysplasin* (*Eda*), *Tumor necrosis factor super-family member 13b* (*Tnfsf13b*) and *Glycoprotein A rich protein* (*Garp*). Genotyping assays were designed for 17 markers (triangles) that distinguish the marine or "C" allele (associated with the completely-plated phenotype) from the freshwater or "L" allele (associated with the low-plated phenotype). These markers are a mix of SNPs and indels indicated by grey and black triangles, respectively, and are listed in Table 2.1.

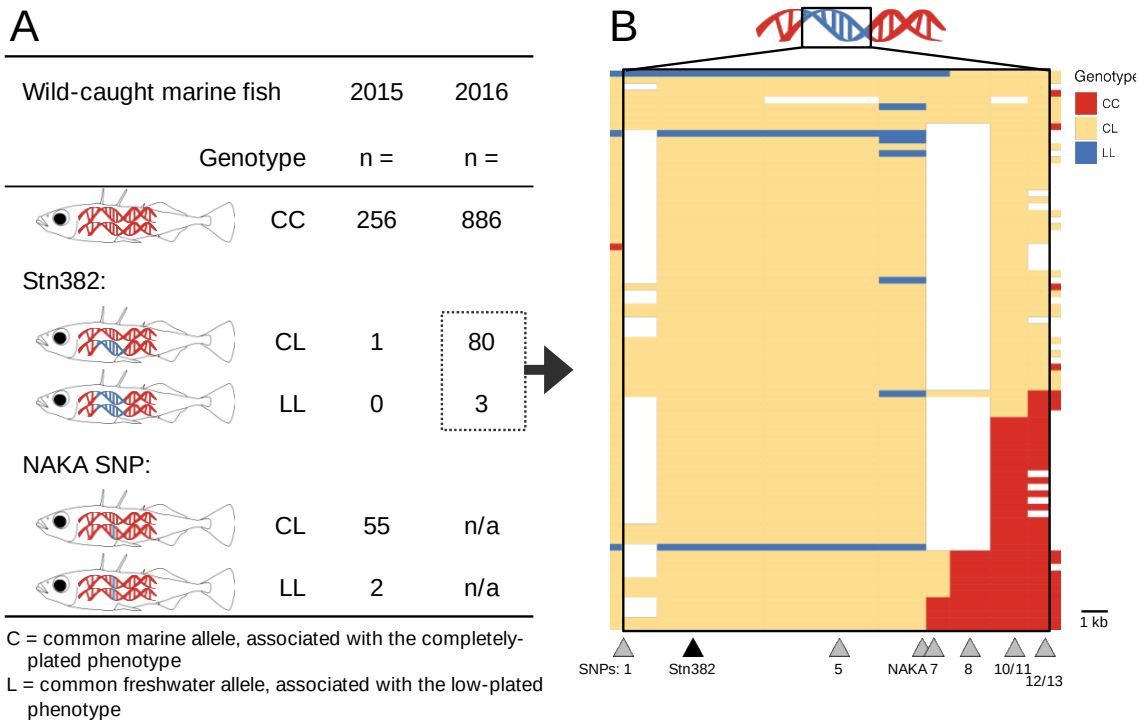


Figure 2.2. Puget Sound marine stickleback collections and genotypes. (A) Marine stickleback were collected and genotyped from Puget Sound in two consecutive summers, 2015 and 2016. In 2015, wild-caught fish were genotyped first at the NAKA SNP, and subsequently at Stn382. The frequency of the L allele at the NAKA SNP was 9.4%, while the frequency of the full freshwater *Eda* haplotype was 0.16%. In 2016, all fish were genotyped at Stn382, and the frequency of the L allele was 4.4%. (B) The carriers of the L allele from 2016 were genotyped at a subset of additional markers, and their genotypes are represented visually. Each row represents a single fish (n=83), and the subset of markers genotyped within the haplotype (designated by the box) is shown below the plot. Triangles mark the physical location of each marker within the haplotype. Coloring representing the genotypes extends halfway to the next marker location. Additional markers on either side of the haplotype (Cnv767 and SNP19, ~10 kb 5' and 3' of the region, respectively) were genotyped and are represented visually outside the box. Missing data are left blank.

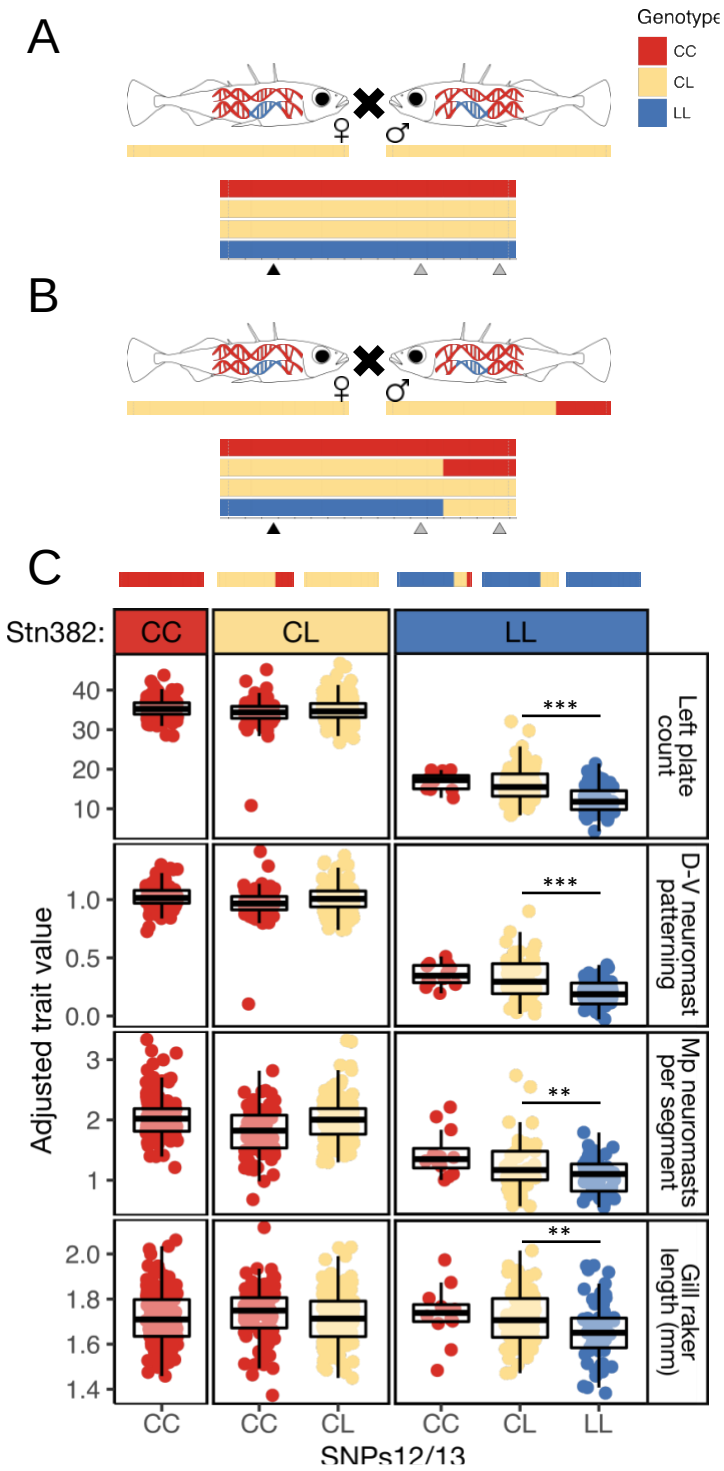


Figure 2.3. Puget Sound crosses schematic and trait values by genotype. (A) Schematic of the five crosses between heterozygous marine carriers of the full freshwater haplotype (CL genotype). (B) Schematic of the five crosses between one carrier of the full haplotype and one carrier of a short haplotype. In both (A) and (B), the possible genotypes of offspring are visualized below the parents, along with triangles identifying the three markers at which offspring were genotyped –

Stn382, NAKA SNP and SNPs12/13. Not pictured is the single cross between two carriers of different short haplotypes, which produced some offspring with a tricolored haplotype, depicted in (C). (C) Trait values for four phenotypes are plotted by offspring genotypes at Stn382 and SNPs12/13. Representative haplotypes are drawn above the plots. Left plate count is strongly associated with genotype at Stn382 (LOD = 344, PVE = 93.6), and the C allele is dominant (CC and CL fish have similar trait values). However, SNPs12/13 are also associated with plate number in fish that are homozygous LL at Stn382 (fish LL at SNPs12/13 have fewer plates than fish CC and CL at SNPs12/13, $t_{151} = -6.54$, $p = 4.4e-10$). The same pattern is observed in the dorsal-ventral (D-V) patterning of neuromasts and the number of neuromasts per segment in the Mp line. In contrast, gill raker length is not associated with genotype at Stn382, but it is associated with genotype at SNPs12/13 (LOD = 4.6, PVE = 3.7). Adjusted trait values were calculated by adding the residual trait value for each individual to the predicted trait value when all the covariates are equal (sex, standard length and family).

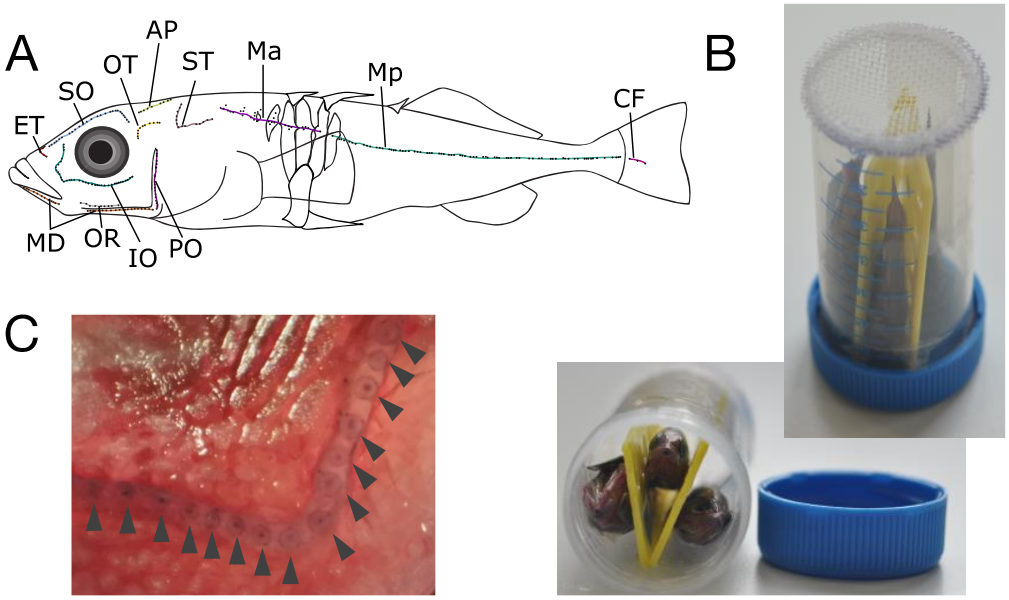
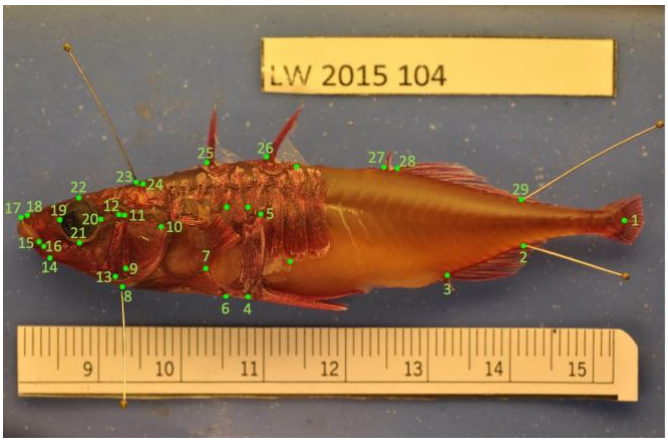


Figure 2.4. Lateral line screening using alkaline phosphatase staining. Twelve neuromast lateral lines were phenotyped using either DASPEI or alkaline phosphatase staining to identify and count neuromasts. (A) Lateral lines were identified following Wark and Peichel (2010). The twelve lines are: AP-anterior pit; CF-caudal fin; ET-ethmoid; IO-infraorbital; Ma-anterior main trunk; MD-mandibular; Mp-posterior main trunk; OR-oral; OT-otic; PO-preopercular; SO-supraorbital; ST-supratemporal. (B) Fish were stained with alkaline phosphatase in custom staining chambers made from 50 mL conical tubes, nylon mesh, and pipette tip refill wafers. Each chamber fit 3 fish with their ID tags, and washes were easily exchanged through the mesh without opening the tubes. (C) A representative photo of alkaline phosphatase staining of neuromasts (arrowheads) and Alizarin red staining of bone at the intersection of the PO and MD lines.



Landmark	Description
1	Posterior extent of caudal peduncle
2	Posterior insertion of anal fin
3	Anterior insertion of anal fin
4	Insertion point of pelvic spine into the pelvic girdle
5	Dorsal extent of the ascending branch of the pelvis
6	Posterior extent of ectocoracoid
7	Dorsal extent of ectocoracoid
8	Anterior extent of ectocoracoid
9	Ventral extent of operculum
10	Posteriodorsal extent of operculum
11	Anteriodorsal extent of operculum
12	Dorsal extent of preopercular
13	Posterioventral extent of preopercular
14	Anterioventral extent of preopercular
15	Posterior extent of premaxilla
16	Posterior extent of maxilla
17	Anterior-most extent of the premaxilla
18	Anterior extent of maxilla
19	Anterior extent of orbit
20	Posterior extent of orbit
21	Ventral extent of orbit
22	Dorsal margin of the orbit in line with the eye's midpoint
23	Supraoccipital notch lateral to the dorsal midline
24	Posterior extent of supraoccipital
25	Anterior insertion of first dorsal spine
26	Anterior insertion of second dorsal spine
27	Anterior insertion of third dorsal spine
28	Anterior insertion of the dorsal fin at the first soft ray
29	Posterior insertion of dorsal fin at the first soft ray

Figure 2.5. Landmark positions and descriptions used for geometric morphometrics. Landmarks were chosen based on previous QTL mapping studies between marine and freshwater fish (Table 2.2).

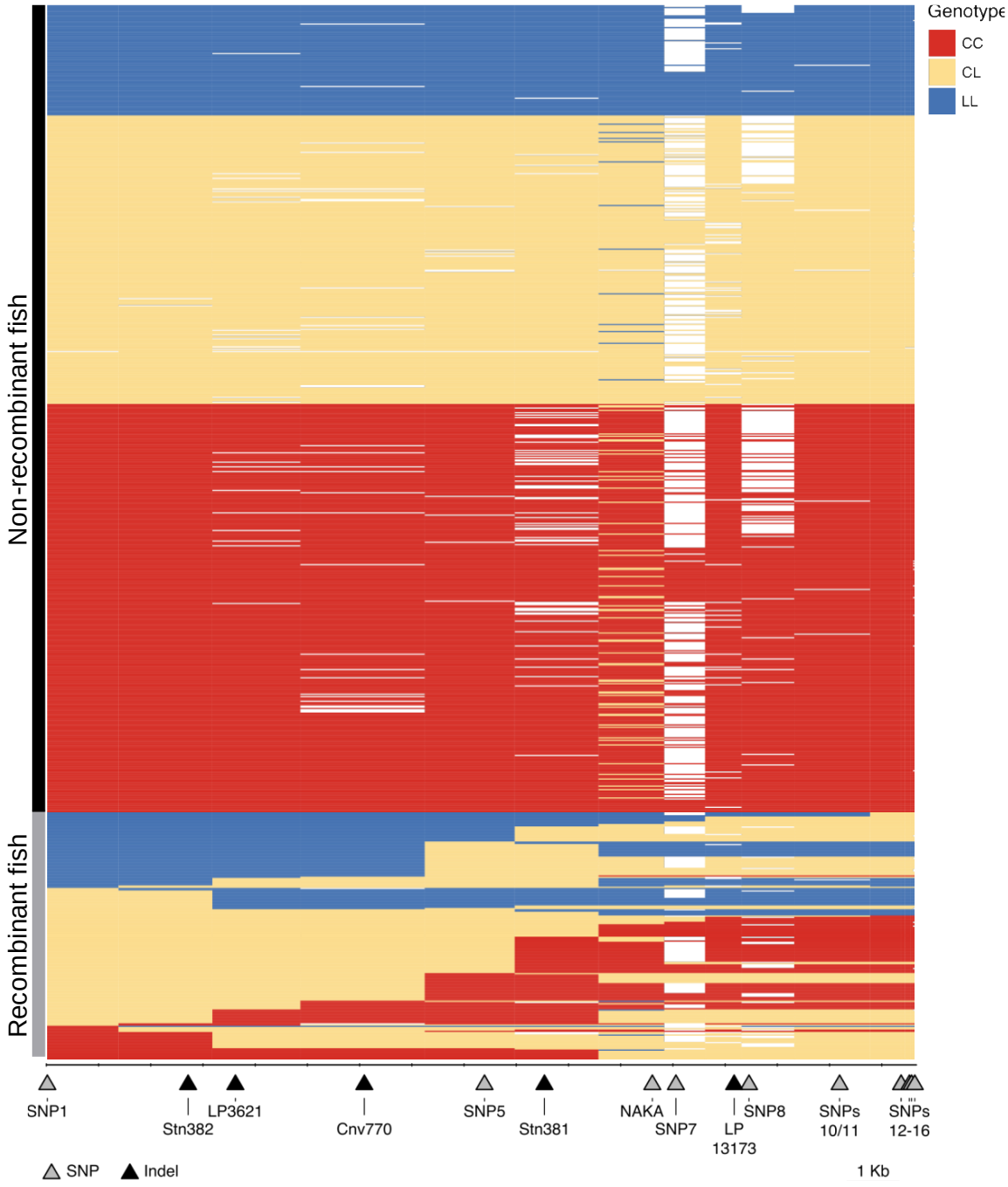


Figure 2.6. Visual genotypes of wild-caught Lake Washington fish. Wild-caught adult stickleback from Lake Washington (n = 885) were genotyped at a subset of 17 SNPs or indels across the haplotype. Most fish showed no evidence of recombination within the 16 kb haplotype: n = 95 (11%) were LL at all screened markers; n = 245 (28%) were CL at all screened markers except NAKA; and n = 347 (39%) were CC at all screened markers except NAKA (top). However, we did find evidence of at least one historical recombination event within the haplotype in 198, or 22%, of the fish. In addition to visualizing recombination breakpoints, this figure highlights the

frequency of the freshwater allele of the NAKA SNP on an otherwise marine haplotype. For example, we see complete correspondence between NAKA and the other genotypes within the LL non-recombinant fish (blue, top). However, within the generally CL non-recombinant fish, there are 12 blue LL genotypes at NAKA (n = 12 of 245 C haplotypes, 4.9%). Within the generally CC non-recombinant fish there are 56 yellow CL genotypes at NAKA (n = 56 of 694 C haplotypes, 8.1%). The frequencies of the L allele at the NAKA SNP on generally C haplotype is 9.4% in Puget Sound (Figure 2.2) and 7.2% in Lake Washington. The markers are depicted as triangles at their physical location within the haplotype and labelled at the bottom of the figure. Tick marks are spaced every 1000 bases for scale. Genotypes are represented visually as CC (homozygous for the completely-plated allele), CL (heterozygous) or LL (homozygous for the low-plated, typically freshwater, allele).

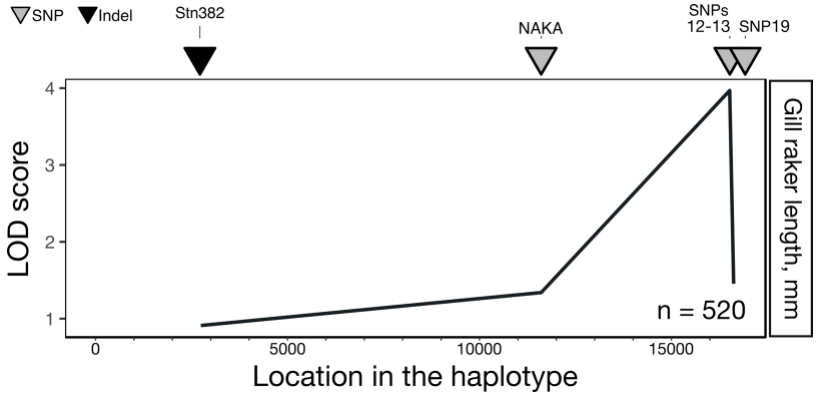


Figure 2.7. LOD plot of gill raker length from Puget Sound crosses. Gill raker length was significantly associated with SNPs12/13 in the Puget Sound crosses when only Stn382 and SNPs12/13 were considered. In Lake Washington, the strongest signal of association between genotype and gill raker length was at the NAKA SNP (Figure 2.9). To assess the association between the NAKA SNP and gill raker length in Puget Sound, we genotyped the F1 offspring at the NAKA SNP, which differed from the other genotypes within the haplotype in two of the crosses, and reran the analysis. This confirmed the first analysis- the association of gill raker length and genotype is driven by the genotype at SNPs12/13 in Puget Sound, not the NAKA SNP or something outside of the haplotype (SNP19).

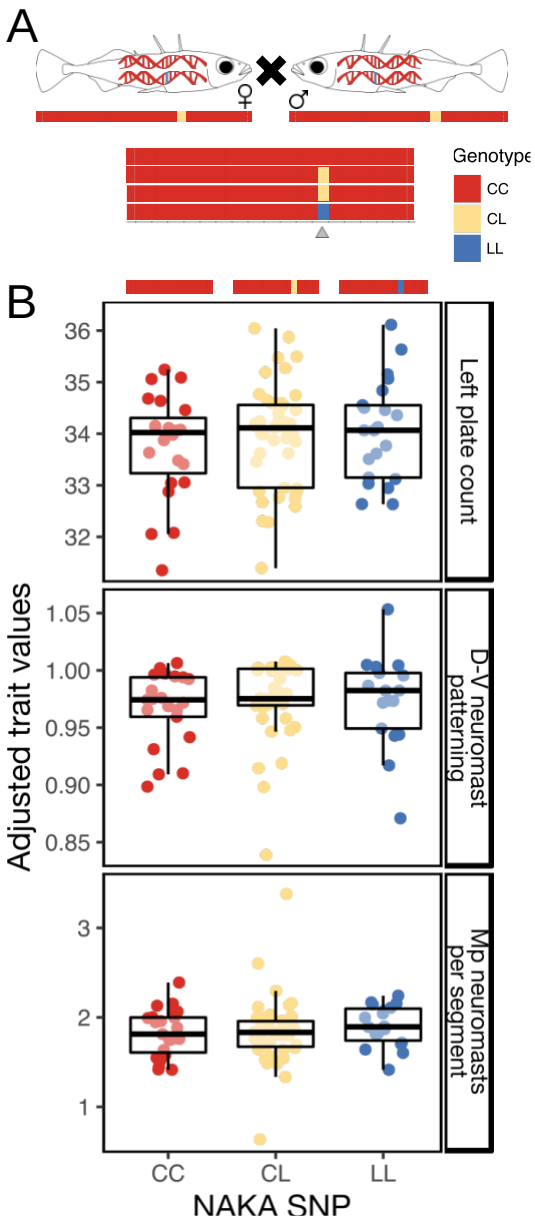


Figure 2.8. Puget Sound NAKA SNP crosses schematic and trait values by genotype. (A) Schematic of crosses between heterozygous marine carriers of the NAKA SNP (CL genotype). Possible offspring genotypes are visualized below the parents, along with a triangle marking the NAKA SNP. (B) Trait values for three phenotypes are plotted by offspring genotypes at the NAKA SNP. Representative genotypes are drawn above the plots. There is no association between genotype at the NAKA SNP and any of the three phenotypes – left plate count, dorsal-ventral (D-V) patterning of neuromasts, or neuromasts per body segment along the posterior main trunk line (Mp). Adjusted trait values were calculated by adding the residual trait value for each individual to the predicted trait value when all the covariates are equal (sex, standard length and family).

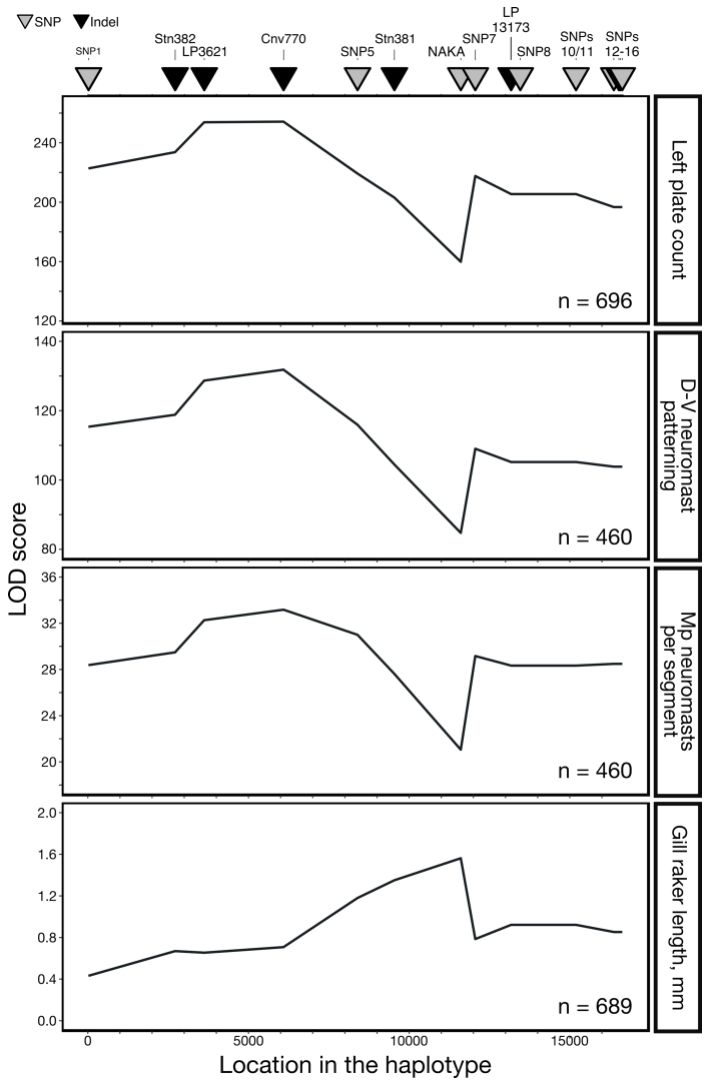


Figure 2.9. LOD plots from association mapping in Lake Washington stickleback. The strength of association between each marker and each phenotype was calculated as a log odds likelihood (LOD) score compared with the model of no association between marker and phenotype. These LOD curves are plotted for four traits. Three traits – left plate count, dorsal-ventral patterned segments, and Mp neuromasts per segment – display very similar LOD curves, suggesting these traits are controlled by the same mutation(s). The LOD curve for gill raker length is shown, since this trait significantly mapped to the 3' region (SNPs12-13) of the haplotype in the Puget Sound crosses. In Lake Washington, this trait does not significantly map to any marker within the haplotype, and the highest LOD score for gill raker length is with the NAKA SNP.

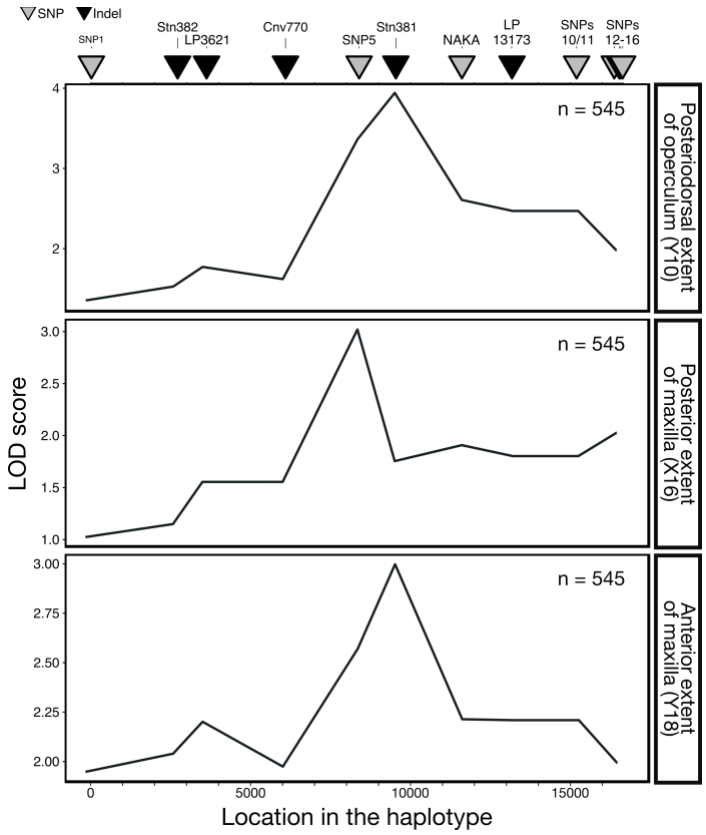


Figure 2.10. LOD plots from association mapping of geometric morphometric landmarks in Lake Washington stickleback. The strength of association between each marker and each phenotype was calculated as a log odds likelihood (LOD) score compared with the model of no association between marker and phenotype. These LOD curves are plotted for three geometric morphometric landmarks that are significant following correction for multiple comparisons (Figure 2.5). Posteriodorsal extent of the operculum (Y10) and anterior extent of maxilla (Y18) map most strongly to Stn381. The posterior extent of the maxilla (X16) maps most strongly to SNP5. These patterns of association are consistent with either separate, linked causative mutations, or a single pleiotropic mutation between SNP5 and Stn381.

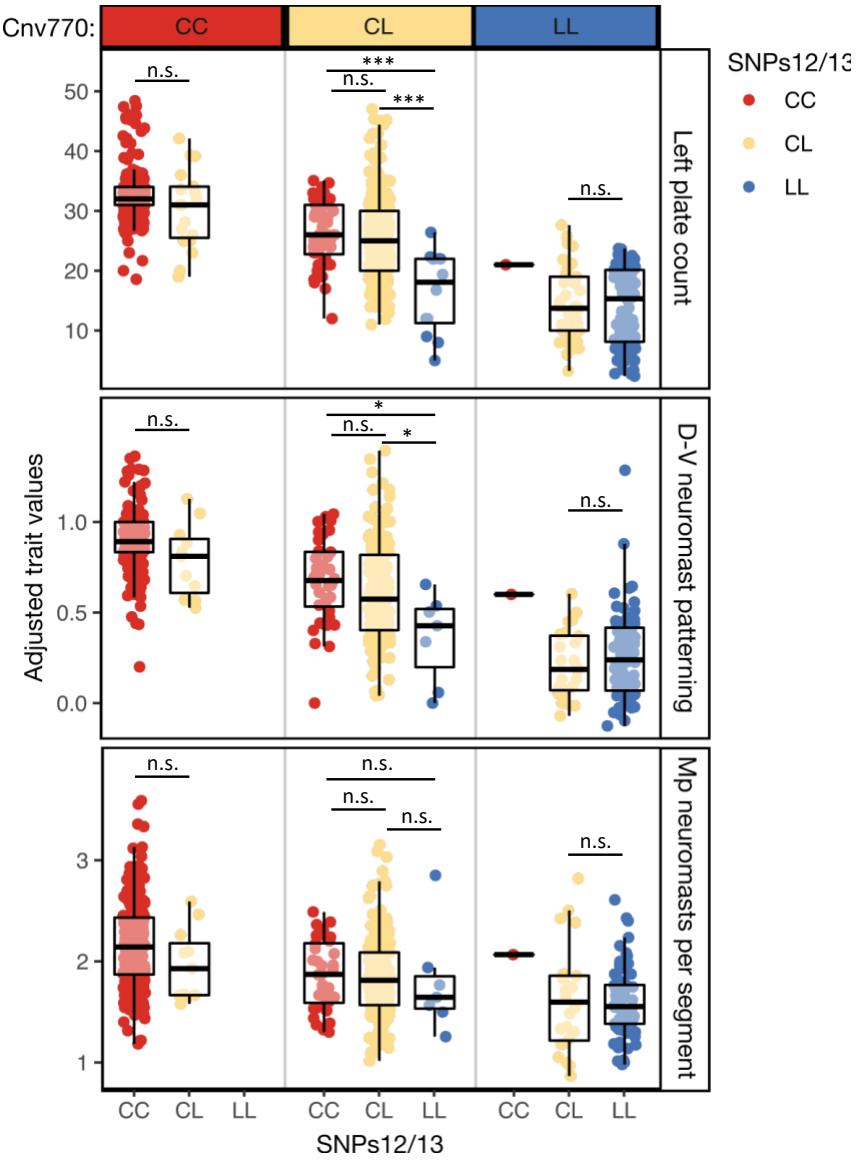


Figure 2.11. Trait values by genotype for plate count in Lake Washington stickleback. Adjusted left plate count, D-V neuromast patterning, and Mp neuromasts per segment are plotted for Lake Washington fish by both genotype at Cnv770 (the marker with the highest single LOD score in the association mapping) and SNPs12/13, which controlled additional variation in plate count in the Puget Sound crosses (Figure 2.3). Fish that are heterozygous at both Cnv770 and SNPs12/13 have significantly more plates than fish heterozygous at Cnv770 and homozygous LL at SNPs12/13 (t_{12} = 4.4, p = 0.00085) and have more segments displaying D-V neuromast patterning (t_{12} = 2.7, p = 0.03). Adjusted trait values were calculated by adding the residual trait value for each individual to the predicted trait value when all the covariates are equal (sex, standard length and collection), and therefore vary outside the normal range of left plate counts.

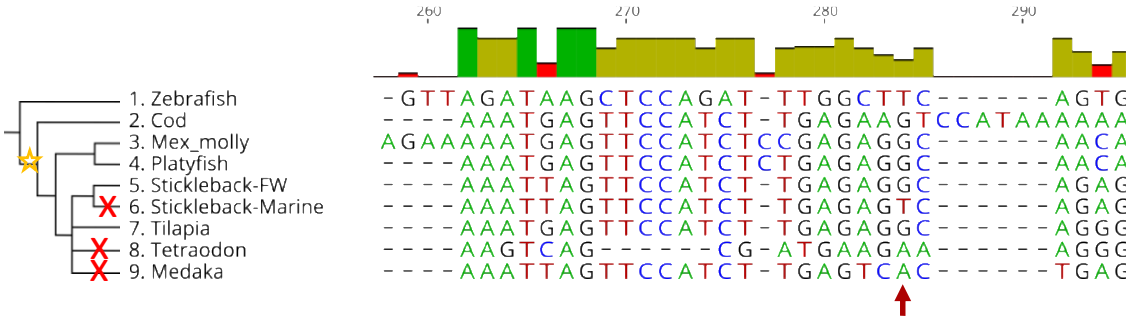


Figure 2.12. Sequence alignment of the NAKA SNP from stickleback and seven additional teleost species. Alignment of a portion of the intergenic region downstream of *Eda*. Genomic sequences downstream of *Eda* were retrieved from NCBI using the stickleback Eda protein sequence as a query. Sequences were aligned using VISTA and corrected at the locus by eye in Geneious R11.1.5 (https://www.geneious.com). Zebrafish and cod are not members of the percomorph clade, yet cod share a "G" at the NAKA SNP (red arrow) with other percomorphs (including freshwater, low-plated stickleback), suggesting that the NAKA SNP arose in the common ancestor of cod and percomorphs (yellow star). Three percomorphs do not have a "G" at the NAKA SNP position, suggesting it was lost in marine stickleback, tetraodon and medaka (red X's).

A

Left plate count	SNP1	Stn382	LP3621	Cnv770	SNP5	Stn381	NAKA	SNP7	LP13173	SNP8	SNPs10/11	SNP12	SNP13	SNP14	SNP15	SNP16
* 7	CL	CL	LL	LL	LL	LL	LL	-	LL	LL	LL	LL	LL	LL	LL	LL
* 7	CL	CL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL
* 7	CL	CL	LL	LL	LL	LL	LL	-	LL	LL	LL	LL	LL	LL	LL	LL
7	LL	LL	LL	LL	LL	LL	LL	-	LL	LL	LL	CL	CL	CL	CL	CL
7	CL	CL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL
→ 7	LL	LL	LL	CL	CL	CL	CL	CL	-	CL	CL	CL	CL	CL	CL	CL
7	LL	LL	LL	LL	CL	CL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL
7	CL	LL	LL	-	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL
7	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	CL	CL	CL	CL	CL
7	CC	LL	LL	LL	LL	LL	LL	-	-	LL	LL	LL	LL	LL	LL	LL
7	LL	LL	LL	LL	LL	LL	LL	-	LL	LL	LL	CL	CL	CL	CL	CL
7	LL	LL	LL	LL	CL	CL	CL	-	CL	CL	CL	CL	CL	CL	CL	CL
8	LL	LL	LL	LL	LL	LL	LL	CL	CL	CL	CL	CL	CL	CL	CL	CL
7	LL	LL	LL	LL	LL	LL	LL	LL	CL	CL	CL	CL	CL	CL	CL	CL
7	LL	LL	LL	LL	LL	LL	LL	LL	CL	CL	CL	CL	CL	CL	CL	CL
8	LL	LL	LL	LL	LL	CL	CL	-	CL	-	CL	CL	CL	CL	CL	CL
7	CL	CL	LL	LL	LL	LL	LL	-	LL	-	LL	LL	LL	LL	LL	LL
8	CL	LL	LL	LL	LL	LL	LL	-	LL	LL	LL	LL	LL	LL	LL	LL
8	LL	LL	LL	LL	LL	LL	LL	-	LL	LL	LL	CL	CL	CL	CL	CL
8	LL	LL	LL	LL	LL	LL	LL	CL	CL	CL	CL	CL	CL	CL	CL	CL
8	LL	LL	LL	LL	CL	CL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL
8	LL	LL	LL	LL	LL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	-
8	LL	LL	LL	LL	LL	LL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL
* 8	CL	CL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL
9	LL	LL	LL	LL	CL	CL	CL	-	CL	-	CL	CL	CL	CL	CL	CL
9	LL	LL	LL	LL	CL	CL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL
* 8	CL	CL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL
9	LL	LL	LL	LL	LL	LL	LL	LL	CL	CL	CL	CL	CL	CL	CL	CL
9	LL	LL	LL	LL	CL	CL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL
10	LL	LL	LL	LL	CL	CL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL
9	CL	CL	LL	LL	LL	LL	LL	-	LL	LL	LL	LL	LL	LL	LL	LL
10	LL	LL	LL	LL	CL	CL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL
10	LL	LL	LL	LL	LL	CL	CL	-	CL	CL	CL	CL	CL	CL	CL	CL
11	LL	LL	LL	LL	LL	CL	CL	-	CL	CL	CL	CL	CL	CL	CL	CL
10	LL	LL	LL	LL	LL	CL	CL	-	CL	CL	CL	CL	CL	CL	CL	CL
# 7	LL	LL	CL	CL	CL	CL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL
# 8	LL	LL	CL	CL	CL	CL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL
# 8	LL	LL	CL	CL	CL	CL	LL	LL	-	LL	LL	LL	LL	LL	LL	LL
9	LL	LL	CL	CL	CL	CL	LL	LL	LL	LL	-	LL	LL	LL	LL	LL
# 11	LL	LL	CL	CL	CL	CL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL
10	LL	LL	CL	CL	CL	CL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL

B

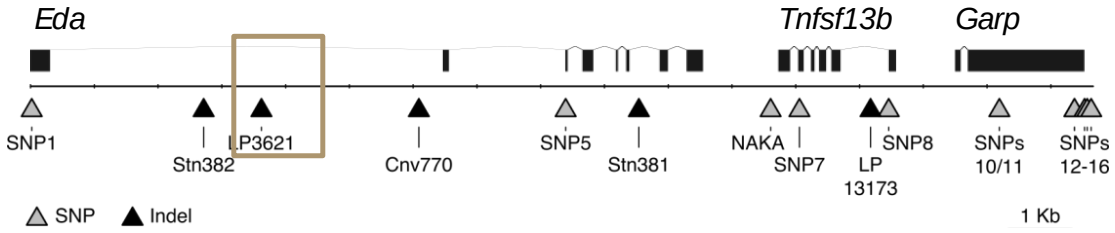


Figure 2.13. Fine-mapping of breakpoints in low-plated Lake Washington fish. Genotypes of low-plated fish at the 16 screened markers are shown for fish with evidence of at least one historical recombination. Together, these genotypes suggest that the causative allele is located between Stn382 and LP3621. Cloning and sequencing of the five fish with an asterisk next to their plate count identified the 5' recombination breakpoint at base pair 12,803,450 of the gasAcu1 genome. Analysis of the sequences of the four fish with a pound sign next to their plate count found no mutation for which these low-plated fish all carried two copies of the freshwater allele. We

therefore conclude that the six fish on the bottom of the table may carry a unique set of marine alleles and refocused on the region between LP3621 and Cnv770, since all other low-plated fish are homozygous for the freshwater allele at LP3621. A single fish (arrow) recombines between LP3621 and Cnv770. This recombination breakpoint was mapped to base pair 12,804,851 in the gasAcu1 genome by cloning the recombinant allele. Therefore, the causative mutation lies within a region of 1,401 base pairs, marked with a box in (B).

Table 2.1. Table of markers and genotyping assays. The 17 markers within the 16 kb haplotype, two markers flanking the haplotype, and the two markers used to determine sex in these studies are listed. Chromosome IV location refers to the Broad/gasAcu1 genome assembly. Abbreviations used in the table include: Cnv = copy number variant; SNP = single nucleotide polymorphism; Stn = (Stanford) short tandem repeat; LP = length polymorphism; FW = freshwater; PS = Puget Sound marine population; LW = Lake Washington freshwater population; RFLP = restriction fragment length polymorphism assay; Seq = Sanger sequencing; and MAMA= mismatched amplification mutation assay.

MARKER NAME	DESCRIPTION	CHR. IV LOCATION	HAPLOTYPE LOCATION	POPULATION(S) SCREENED (CHAPTER 2)	GENERATION SCREENED (CHAPTER 3)
Cnv767[1]	Intergenic FW insertion (68 bp)	12,790,042-12,790,110	-	PS / --	F0
SNP1[2]	Nonsynonymous SNP in *Eda*	12,800,238	19	-- / LW	F0
SNP4[2]	Nonsynonymous SNP in *Eda*	12,800,508	289	-- / --	F0 / F1
Stn382[2]	Intronic FW deletion (66 bp)	12,802,847	2,715	PS / LW	F0 / F1 / F2
LP3621	Intronic FW deletion (16 bp)	12,803,862	3,621	-- / LW	--
Cnv770[1]	Intronic FW insertion (107 bp)	12,806,052-12,806,138	6,090	-- / LW	--
SNP5[2]	Nonsynonymous SNP in *Eda*	12,808,303	8,392	-- / LW	F0
Stn381[2]	Intronic FW insertion (19 bp)	12,809,414-12,809,601	9,540	-- / LW	--
SNP11110	Intergenic SNP	12,810,985	11,110	-- / --	F0 / F1
NAKA.SNP[3]	Intergenic SNP	12,811,481	11,606	-- / LW	F0 / F1
SNP7[2]	Nonsynonymous SNP in *Tnfsf13b*	12,811,933	12,058	-- / LW	F0
LP13173	Intronic FW insertion (20 bp)	12,813,043-12,813,066	13,173	-- / LW	--
SNP8[2]	Synonymous SNP in *Tnfsf13b*	12,813,328	13,458	-- / LW	F0
SNPs10/11[2]	Nonsynonymous SNP in *Garp*	12,815,024	15,193	-- / LW	F0
SNP12[2]	Nonsynonymous SNP in *Garp*	12,816,201	16,370	PS / LW	F0 / F1
SNP13[2]	Synonymous SNP in *Garp*	12,816,202	16,371	PS / LW	F0 / F1
SNP14[2]	Nonsynonymous SNP in *Garp*	12,816,360	16,529	-- / LW	--
SNP15[2]	SNP in *Garp* 3' UTR	12,816,402	16,571	-- / LW	--
SNP16[2]	SNP in *Garp* 3' UTR	12,816,464	16,633	-- / LW	--
SNP19[2]	Nonsynonymous SNP in *Gjb1*	12,826,854	-	PS / --	F0
LRR	Y-chr. deletion	-	-	-- / --	F2
IDH[4]	Y-chr. deletion in 3'UTR of *Idh*	-	-	PS / LW	F0 / F1 / F2

[1]Lowe *et al.*, 2018; [2]Colosimo *et al.*, 2005; [3]O'Brown *et al.*, 2015; [4]Peichel *et al.*, 2004

Table 2.1 (continued). Table of genotyping markers and assays.

MARKER NAME	FORWARD PRIMER (5' - 3')	REVERSE PRIMER (5' - 3')	SCREENING METHOD	RFLP ENZYME
Cnv767[1]	TGTTGTGCAGTCGGGTGCCAGT	GGCTTCCATTGTGCCTGGTGCCTT	PCR	
SNP1[2]	AGATGCCGAGGTAGAGAGCA	GTAGCAGACGAGCGTGACAG	RFLP	BtsCI
SNP4[2]	GACATTTGCTGGCATCGACC	CAGCCTCCGTGCGAGAAAT	RFLP	SfaNI
Stn382[2]	CCCTTAGAGAATTTCCTAGCAG	CTTGTCCCGGATCATACGC	PCR	
LP3621	CCTGGTGGACGGATAGAGCA	ATGTTGCCTGTCAGCAGCCA	PCR	
Cnv770[1]	AGGCATCGCGCTCACGTTGA	TCGCCCGYTGAGTTATGCCCC	PCR	
SNP5[2]	GTTCAGGAGAACGTTTCAAGCT	GCCGCTTTTTCCCTGTGAAG	RFLP	Taqa1
Stn381[2]	CCGCTACACACGGACTTACA*	ATTCGAGGGTTCAGCTCTGG	PCR	
SNP11110	AAAGGACACTGCGCTGACAT	TCTGCAGGGATGCTGTAGTG	RFLP	EcoRV
NAKA[3]	TTGGACACTGCTGGCACGGG	TCACACTCTGRTTAACCCCCGGA	RFLP	MnlI
SNP7[2]	AGGATGTAGTCATCTTGGACATTGT AGGATGTAGTCATCTTGGACATTGT	CTTCAGAGAACTCAGACTGTTTGTG TCAGAGAACTCAGACTGTTTGCC	MAMA PCR - FW allele - Mar. allele	
LP13173	TACGCTAGTTAGCTGCTGGC	AAATGAGGGAATGGGGCCTG	PCR	
SNP8[2]	CAAACAGCTTCAGAGACTGTGC CAAACAGCTTCAGAGACTGTGC	CCTGAAGCAGCTGCACAAGA CCTGAAGCAGCTGCACAAAG	MAMA PCR - FW allele - Mar. allele	
SNPs10/11[2]	ACGACTTGGGTGATGATGCA	TCAAGCCTGTACAACTGGTCA	RFLP	AluI
SNP12[2]	TCTGCTACCAAGCAGTTTGA	AGTTTGCTATTCAGTTTCTCAC	Seq in LW & RFLP in PS	Mfe-HF
SNP13[2]	TCTGCTACCAAGCAGTTTGA	AGTTTGCTATTCAGTTTCTCAC	Seq in LW & RFLP in PS	Mfe-HF
SNP14[2]	TCTGCTACCAAGCAGTTTGA	AGTTTGCTATTCAGTTTCTCAC	Seq	
SNP15[2]	TCTGCTACCAAGCAGTTTGA	AGTTTGCTATTCAGTTTCTCAC	Seq	
SNP16[2]	TCTGCTACCAAGCAGTTTGA	AGTTTGCTATTCAGTTTCTCAC	Seq	
SNP19[2]	CAGCTCATCCTGGTCTCCAC	TCCGGTGATCTGGAACTTCT	RFLP	Eco0109I
LRR	TGGGGAGTGTATGAGTTTAGAGAA	CTCAAAGGCTGAAGGCAGTG	PCR	
IDH[4]	GGGACGAGCAAGATTTATTG	TTATCGTTAGCCAGGAGATGG	PCR	

[1]Lowe *et al.*, 2018; [2]Colosimo *et al.*, 2005; [3]O'Brown *et al.*, 2015; [4]Peichel *et al.*, 2004

*Redesigned primer; previously published primer sat on a polymorphic site in the Lake Washington population

Table 2.2. Descriptions of traits mapped in this study.

#	TRAIT TYPE	TRAIT ABBREVIATION	TRAIT DESCRIPTION	STUDY MAPPING TRAIT TO CHR. IV
1	Branchial	DTP1	Right dorsal toothplate 1 tooth count	Miller *et al.*, 2014
2	Branchial	DTP2	Right dorsal toothplate 2 tooth count	Miller *et al.*, 2014
3	Branchial	VTP	Right ventral toothplate tooth count	Miller *et al.*, 2014; Cleves *et al.*, 2014; Ellis *et al.*, 2015
4	Branchial	Gill raker count	1st row gill raker count	Glazer *et al.*, 2014
5	Branchial	CB4L	Ceratobranchial 4 length	Miller *et al.*, 2014; Erickson *et al.*, 2014
6	Lateral line	LL.AP	Anterior pit neuromasts (AP)	
7	Lateral line	LL.ET	Ethmoid neuromasts (ET)	
8	Lateral line	LL.IO	Infraorbital neuromasts (IO)	Wark *et al.*, 2012
9	Lateral line	LL.MD	Mandibular neuromasts (MD)	
10	Lateral line	LL.OR	Oral neuromasts (OR)	
11	Lateral line	LL.OT	Otic neuromasts (OT)	
12	Lateral line	LL.PO	Preopercular neuromasts (PO)	Wark *et al.*, 2012
13	Lateral line	LL.SO	Supraorbital neuromasts (SO)	
14	Lateral line	LL.ST	Supratemporal neuromasts (ST)	Wark *et al.*, 2012
15	Lateral line	LL.CF	Caudal fin neuromasts (CF)	
16	Lateral line	LL.Ma N per segment	Ma (anterior main trunk line) neuromasts per segment	
17	Lateral line	LL.Mp N per segment	Mp (posterior main trunk line) neuromasts per segment	
18	Lateral line	LL.Mp % DV	Dorsal-Ventral (D-V) neuromast patterning	Wark *et al.*, 2012
19	Meristic	Plates	Left plate count	Colosimo *et al.*, 2004; Cresko *et al.*, 2004; Wark *et al.*, 2012; Glazer *et al.*, 2015; Lui *et al.*, 2014
20	Meristic	Anal rays	Anal fin ray count	Miller *et al.*, 2014
21	Meristic	Dorsal rays	Dorsal fin ray count	Miller *et al.*, 2014
22	Metric	X1DS	Dorsal spine 1 length	Miller *et al.*, 2014; Rogers *et al.*, 2012
23	Metric	X2DS	Dorsal spine 2 length	Miller *et al.*, 2014
24	Metric	X3DS	Dorsal spine 3 length	Miller *et al.*, 2014
25	Metric	LPS	Left pelvic spine length	Rogers *et al.*, 2012; Shapiro *et al.*, 2004
26	Metric	Pel girdle	Pelvic girdle length	Rogers *et al.*, 2012
27	Metric	Ectocorocoid	Ectocoracoid length	Rogers *et al.*, 2012
28	Metric	Raker length	Middle gill raker length	Berner *et al.*, 2014; Glazer *et al.*, 2015
29	Metric	Body depth	Body depth	
30	Metric	AS	Anal spine length	Miller *et al.*, 2014
31	Metric	Maxilla	Maxilla length	
32	Metric	Supraoccipital	Supraoccipital notch length	Miller *et al.*, 2014
33	Morphometric	Centroid size	Centroid size	

#	TRAIT TYPE	TRAIT ABBREVIATION	TRAIT DESCRIPTION	STUDY MAPPING TRAIT TO CHR. IV
34	Morphometric	X1	Posterior extent of caudal peduncle (X1)	
35	Morphometric	X2	Posterior insertion of anal fin (X2)	
36	Morphometric	X3	Anterior insertion of anal fin (X3)	Rogers *et al.*, 2012
37	Morphometric	X4	Insertion point of pelvic spine into the pelvic girdle (X4)	Albert *et al.*, 2008
38	Morphometric	X5	Dorsal extent of the ascending branch of the pelvis (X5)	Albert *et al.*, 2008
39	Morphometric	X6	Posterior extent of ectocoracoid (X6)	Rogers *et al.*, 2012
40	Morphometric	X7	Dorsal extent of ectocoracoid (X7)	
41	Morphometric	X8	Anterior extent of ectocoracoid (X8)	Rogers *et al.*, 2012
42	Morphometric	X9	Ventral extent of operculum (X9)	
43	Morphometric	X10	Posteriodorsal extent of operculum (X10)	Albert *et al.*, 2008
44	Morphometric	X11	Anteriodorsal extent of operculum (X11)	
45	Morphometric	X12	Dorsal extent of preopercular (X12)	
46	Morphometric	X13	Posterioventral extent of preopercular (X13)	
47	Morphometric	X14	Anterioventral extent of preopercular (X14)	
48	Morphometric	X15	Posterior extent of premaxilla (X15)	
49	Morphometric	X16	Posterior extent of maxilla (X16)	
50	Morphometric	X17	Anterior-most extent of the premaxilla (X17)	
51	Morphometric	X18	Anterior extent of maxilla (X18)	
52	Morphometric	X19	Anterior extent of orbit (X19)	
53	Morphometric	X20	Posterior extent of orbit (X20)	
54	Morphometric	X21	Ventral extent of orbit (X21)	
55	Morphometric	X22	Dorsal margin of the orbit in line with the eye's midpoint (X22)	
56	Morphometric	X23	Supraoccipital notch lateral to the dorsal midline (X23)	
57	Morphometric	X24	Posterior extent of supraoccipital (X24)	Albert *et al.*, 2008
58	Morphometric	X25	Anterior insertion of first dorsal spine (X25)	Albert *et al.*, 2008; Rogers *et al.*, 2012
59	Morphometric	X26	Anterior insertion of second dorsal spine (X26)	
60	Morphometric	X27	Anterior insertion of third dorsal spine (X27)	Rogers *et al.*, 2012
61	Morphometric	X28	Anterior insertion of the dorsal fin at the first soft ray (X28)	
62	Morphometric	X29	Posterior insertion of dorsal fin at the first soft ray (X29)	
63	Morphometric	Y1	Posterior extent of caudal peduncle (Y1)	
64	Morphometric	Y2	Posterior insertion of anal fin (Y2)	
65	Morphometric	Y3	Anterior insertion of anal fin (Y3)	
66	Morphometric	Y4	Insertion point of pelvic spine into the pelvic girdle (Y4)	Rogers *et al.*, 2012
67	Morphometric	Y5	Dorsal extent of the ascending branch of the pelvis (Y5)	Albert *et al.*, 2008; Rogers *et al.*, 2012
68	Morphometric	Y6	Posterior extent of ectocoracoid (Y6)	Albert *et al.*, 2008
69	Morphometric	Y7	Dorsal extent of ectocoracoid (Y7)	Albert *et al.*, 2008; Rogers *et al.*, 2012
70	Morphometric	Y8	Anterior extent of ectocoracoid (Y8)	
71	Morphometric	Y9	Ventral extent of operculum (Y9)	
72	Morphometric	Y10	Posteriodorsal extent of operculum (Y10)	
73	Morphometric	Y11	Anteriodorsal extent of operculum (Y11)	Albert *et al.*, 2008

#	TRAIT TYPE	TRAIT ABBREVIATION	TRAIT DESCRIPTION	STUDY MAPPING TRAIT TO CHR. IV
74	Morphometric	Y12	Dorsal extent of preopercular (Y12)	Albert *et al.*, 2008
75	Morphometric	Y13	Posterioventral extent of preopercular (Y13)	
76	Morphometric	Y14	Anterioventral extent of preopercular (Y14)	Albert *et al.*, 2008
77	Morphometric	Y15	Posterior extent of premaxilla (Y15)	Albert *et al.*, 2008; Rogers *et al.*, 2012
78	Morphometric	Y16	Posterior extent of maxilla (Y16)	Albert *et al.*, 2008
79	Morphometric	Y17	Anterior-most extent of the premaxilla (Y17)	
80	Morphometric	Y18	Anterior extent of maxilla (Y18)	
81	Morphometric	Y19	Anterior extent of orbit (Y19)	
82	Morphometric	Y20	Posterior extent of orbit (Y20)	
83	Morphometric	Y21	Ventral extent of orbit (Y21)	
84	Morphometric	Y22	Dorsal margin of the orbit in line with the eye's midpoint (Y22)	
85	Morphometric	Y23	Supraoccipital notch lateral to the dorsal midline (Y23)	
86	Morphometric	Y24	Posterior extent of supraoccipital (Y24)	Albert *et al.*, 2008
87	Morphometric	Y25	Anterior insertion of first dorsal spine (Y25)	Rogers *et al.*, 2012
88	Morphometric	Y26	Anterior insertion of second dorsal spine (Y26)	
89	Morphometric	Y27	Anterior insertion of third dorsal spine (Y27)	
90	Morphometric	Y28	Anterior insertion of the dorsal fin at the first soft ray (Y28)	
91	Morphometric	Y29	Posterior insertion of dorsal fin at the first soft ray (Y29)	

Table 2.3. Summary of results from linkage mapping in Puget Sound stickleback crosses. Puget Sound fish were F1 and F2 offspring of heterozygous marine fish. The marker with the highest LOD score for each trait is listed. P-values and p-value significance thresholds are based on 5000 permutations of the data. Covariates used to map traits in the Puget Sound crosses were Sex + Standard length + Family + Family x Standard length interaction + Family x Sex interaction. LOD = log odds likelihood score; PVE = percent variation explained.

#	TRAIT TYPE	TRAIT ABBREVIATION	N	MARKER WITH HIGHEST LOD	LOD	PVE	P-VALUE	P-VALUE SIG. THRESHOLD
1	Branchial	DTP1	98	Stn382	1.09	5.01	0.095	0.003
2	Branchial	DTP2	98	Stn382	0.38	1.78	0.444	0.003
3	Branchial	VTP	98	Stn382	0.24	1.13	0.591	0.003
4	Branchial	Gill raker count	572	Stn382	0.79	0.64	0.188	0.003
5	Branchial	CB4L	98	Stn382	0.25	1.14	0.602	0.003
15	Lateral line	LL.CF	496	SNPs12/13	0.70	0.65	0.233	0.003
16	Lateral line	LL.Ma N per segment	498	Stn382	2.37	2.17	0.005	0.003
17	Lateral line	LL.Mp N per segment	497	Stn382	77.09	51.05	0.000	0.003
18	Lateral line	LL.Mp % DV	498	Stn382	292.33	93.30	0.000	0.003
19	Meristic	Plates	578	Stn382	344.26	93.56	0.000	0.003
22	Metric	X1DS	570	SNPs12/13	0.99	0.80	0.098	0.003
23	Metric	X2DS	574	SNPs12/13	0.76	0.61	0.175	0.003
24	Metric	X3DS	569	SNPs12/13	1.43	1.15	0.047	0.003
25	Metric	LPS	576	Stn382	1.28	1.02	0.064	0.003
26	Metric	Pel girdle	577	SNPs12/13	0.35	0.28	0.468	0.003
27	Metric	Ectocorocoid	577	SNPs12/13	0.05	0.04	0.891	0.003
28	Metric	Raker length	571	SNPs12/13	4.65	3.68	0.000	0.003

Table 2.4. Summary of results from linkage mapping in NAKA SNP stickleback crosses. NAKA SNP crosses were crosses between Puget Sound marine fish heterozygous at the NAKA SNP only. The marker with the highest LOD score for each trait is listed. P-values are based on 5000 permutations of the data. Covariates used to map traits in the NAKA SNP crosses depended on the trait type. The single meristic trait was mapped using Sex + Family + Standard length + Family x Standard length + Family x Sex as covariates. Lateral line traits were mapped using Sex + Family + Standard length + Staining quality + Family x Standard length + Family x Sex as covariates. LOD = log odds likelihood score; PVE = percent variation explained.

#	TRAIT TYPE	TRAIT ABBREVIATION	N	MARKER WITH HIGHEST LOD	LOD	PVE	P-VALUE	P-VALUE SIG. THRESHOLD
6	Lateral line	LL.AP	86	NAKA	1.19	6.19	0.101	0.004
7	Lateral line	LL.ET	86	NAKA	0.57	3.00	0.331	0.004
8	Lateral line	LL.IO	86	NAKA	1.50	7.73	0.056	0.004
9	Lateral line	LL.MD	86	NAKA	1.24	6.41	0.092	0.004
10	Lateral line	LL.OR	86	NAKA	0.57	3.02	0.343	0.004
11	Lateral line	LL.OT	86	NAKA	0.33	1.73	0.545	0.004
12	Lateral line	LL.PO	86	NAKA	2.97	14.72	0.005	0.004
13	Lateral line	LL.SO	86	NAKA	0.50	2.64	0.393	0.004
14	Lateral line	LL.ST	86	NAKA	0.04	0.24	0.919	0.004
15	Lateral line	LL.CF	86	NAKA	0.07	0.38	0.869	0.004
17	Lateral line	LL.Mp N per segment	86	NAKA	0.24	1.28	0.636	0.004
18	Lateral line	LL.Mp % DV	80	NAKA	0.17	0.97	0.732	0.004
19	Meristic	Plates	89	NAKA	0.26	1.34	0.597	0.004

Table 2.5. Summary of results from association mapping in Lake Washington stickleback. Lake Washington fish were wild-caught, freshwater adults. The marker with the highest LOD score for each trait is listed. P-values and p-value thresholds are based on 5000 permutations of the data. Covariates used to map traits in the Lake Washington adults depended on the trait type. Branchial, meristic, and metric traits were mapped using Sex + Standard length + Collection + Collection x Standard length + Collection x Sex as covariates. Lateral line traits were mapped using Sex + Standard length + Collection + Staining method + Days in tank + Staining quality + Holding tank + Collection x Standard length + Collection x Sex as covariates. And morphometric traits were mapped using Sex + Collection + Sex x Collection as covariates, since these data were corrected for size prior to analysis. LOD = log odds likelihood score; PVE = percent variation explained.

#	TRAIT TYPE	TRAIT ABBREVIATION	N	MARKER WITH HIGHEST LOD	LOD	PVE	P-VALUE	P-VALUE SIG. THRESHOLD
1	Branchial	DTP1	76	SNP12	0.17	1.04	0.7080	0.0005
2	Branchial	DTP2	76	SNP1	0.26	1.58	0.5742	0.0005
3	Branchial	VTP	76	SNP5	0.50	3.01	0.3508	0.0005
4	Branchial	Gill raker count	690	Cnv770	0.43	0.29	0.3660	0.0005
5	Branchial	CB4L	76	SNP5	0.16	0.98	0.7356	0.0005
6	Lateral line	LL.AP	429	Stn381	0.27	0.29	0.5788	0.0005
7	Lateral line	LL.ET	443	SNP12	1.51	1.56	0.0474	0.0005
8	Lateral line	LL.IO	450	SNP12	0.71	0.72	0.2276	0.0005
9	Lateral line	LL.MD	409	NAKA	0.75	0.84	0.1974	0.0005
10	Lateral line	LL.OR	445	SNP5	0.56	0.57	0.3016	0.0005
11	Lateral line	LL.OT	433	SNPs10/11	0.70	0.75	0.2238	0.0005
12	Lateral line	LL.PO	442	SNP1	1.44	1.49	0.0476	0.0005
13	Lateral line	LL.SO	449	SNP12	0.46	0.47	0.4064	0.0005
14	Lateral line	LL.ST	436	SNP5	0.10	0.10	0.8112	0.0005
15	Lateral line	LL.CF	365	SNP1	0.91	1.14	0.1330	0.0005
16	Lateral line	LL.Ma N per segment	460	Stn381	3.21	3.16	0.0018	0.0005
17	Lateral line	LL.Mp N per segment	460	Cnv770	30.70	26.46	0.0000	0.0005
18	Lateral line	LL.Mp % DV	460	Cnv770	113.42	67.87	0.0000	0.0005
19	Meristic	Plates	696	Cnv770	218.25	76.40	0.0000	0.0005
20	Meristic	Anal rays	697	SNP1	0.75	0.49	0.1898	0.0005
21	Meristic	Dorsal rays	696	Stn382	0.42	0.28	0.4042	0.0005
22	Metric	X1DS	693	Stn382	1.03	0.68	0.1058	0.0005
23	Metric	X2DS	691	NAKA	0.61	0.41	0.2628	0.0005
24	Metric	X3DS	687	NAKA	0.61	0.41	0.2652	0.0005
25	Metric	LPS	695	Stn382	1.94	1.28	0.0150	0.0005
26	Metric	Pel girdle	698	SNP1	2.75	1.80	0.0018	0.0005
27	Metric	Ectocorocoid	698	SNP1	0.28	0.19	0.5226	0.0005
28	Metric	Raker length	689	NAKA	1.17	0.78	0.0766	0.0005

#	Trait type	Trait abbreviation	N	Marker with highest LOD	LOD	PVE	P-value	P-value sig. threshold
29	Metric	Body depth	545	Cnv770	0.58	0.49	0.2694	0.0005
30	Metric	AS	691	Stn381	2.68	1.77	0.0030	0.0005
31	Metric	Maxilla	545	NAKA	0.95	0.80	0.1158	0.0005
32	Metric	Supraoccipital	691	Stn382	1.35	0.89	0.0490	0.0005
33	Morphometric	Centroid size	545	NAKA	1.67	1.40	0.0236	0.0005
34	Morphometric	X1	545	SNP1	0.15	0.13	0.7178	0.0005
35	Morphometric	X2	545	SNP1	1.26	1.06	0.0648	0.0005
36	Morphometric	X3	545	SNP12	0.56	0.47	0.2968	0.0005
37	Morphometric	X4	545	Stn382	0.80	0.67	0.1618	0.0005
38	Morphometric	X5	545	Stn382	0.54	0.46	0.2924	0.0005
39	Morphometric	X6	545	Stn382	2.86	2.39	0.0022	0.0005
40	Morphometric	X7	545	NAKA	1.21	1.02	0.0622	0.0005
41	Morphometric	X8	545	Stn382	0.81	0.68	0.1588	0.0005
42	Morphometric	X9	545	Stn382	1.66	1.40	0.0252	0.0005
43	Morphometric	X10	545	SNP12	1.67	1.40	0.0288	0.0005
44	Morphometric	X11	545	SNP5	2.25	1.88	0.0066	0.0005
45	Morphometric	X12	544	SNP5	2.18	1.82	0.0092	0.0005
46	Morphometric	X13	545	Stn381	1.55	1.30	0.0274	0.0005
47	Morphometric	X14	544	SNP5	1.99	1.67	0.0072	0.0005
48	Morphometric	X15	544	SNP12	2.40	2.02	0.0038	0.0005
49	Morphometric	X16	545	SNP5	3.02	2.52	0.0004	0.0005
50	Morphometric	X17	545	SNP1	0.46	0.39	0.3588	0.0005
51	Morphometric	X18	545	SNP1	0.75	0.63	0.1990	0.0005
52	Morphometric	X19	545	SNP12	0.72	0.61	0.2026	0.0005
53	Morphometric	X20	545	SNP12	0.82	0.69	0.1690	0.0005
54	Morphometric	X21	545	SNP12	1.89	1.58	0.0126	0.0005
55	Morphometric	X22	545	LP3621	0.53	0.45	0.3046	0.0005
56	Morphometric	X23	545	SNP5	2.71	2.26	0.0014	0.0005
57	Morphometric	X24	545	SNP12	2.81	2.35	0.0024	0.0005
58	Morphometric	X25	545	Stn381	2.31	1.94	0.0060	0.0005
59	Morphometric	X26	543	NAKA	1.00	0.84	0.1086	0.0005
60	Morphometric	X27	545	SNP5	0.53	0.45	0.3070	0.0005
61	Morphometric	X28	545	SNP5	1.02	0.86	0.0916	0.0005
62	Morphometric	X29	544	NAKA	1.29	1.09	0.0480	0.0005
63	Morphometric	Y1	545	Stn381	3.29	2.74	0.0006	0.0005
64	Morphometric	Y2	545	Stn381	1.48	1.24	0.0364	0.0005
65	Morphometric	Y3	545	SNPs10/11	0.24	0.20	0.5888	0.0005
66	Morphometric	Y4	545	SNP12	2.18	1.83	0.0086	0.0005
67	Morphometric	Y5	545	NAKA	1.11	0.93	0.0870	0.0005

#	Trait type	Trait abbreviation	N	Marker with highest LOD	LOD	PVE	p-value	p-value sig. threshold
68	Morphometric	Y6	545	LP13173	1.42	1.19	0.0360	0.0005
69	Morphometric	Y7	545	Stn381	1.44	1.21	0.0356	0.0005
70	Morphometric	Y8	544	SNP5	1.47	1.23	0.0370	0.0005
71	Morphometric	Y9	545	Stn381	2.22	1.86	0.0050	0.0005
72	Morphometric	Y10	545	Stn381	3.94	3.27	0.0000	0.0005
73	Morphometric	Y11	545	Stn381	0.90	0.76	0.1244	0.0005
74	Morphometric	Y12	545	SNP5	1.07	0.90	0.0760	0.0005
75	Morphometric	Y13	545	SNP5	0.43	0.36	0.3806	0.0005
76	Morphometric	Y14	545	Stn381	1.32	1.11	0.0444	0.0005
77	Morphometric	Y15	544	Stn381	2.63	2.20	0.0010	0.0005
78	Morphometric	Y16	544	Stn381	2.37	1.99	0.0032	0.0005
79	Morphometric	Y17	545	Stn381	2.78	2.32	0.0020	0.0005
80	Morphometric	Y18	545	Stn381	3.00	2.50	0.0004	0.0005
81	Morphometric	Y19	545	Stn382	0.92	0.77	0.1238	0.0005
82	Morphometric	Y20	545	NAKA	1.07	0.90	0.0800	0.0005
83	Morphometric	Y21	544	SNPs10/11	0.90	0.76	0.1296	0.0005
84	Morphometric	Y22	545	SNP5	2.94	2.46	0.0006	0.0005
85	Morphometric	Y23	545	Stn381	1.10	0.92	0.0938	0.0005
86	Morphometric	Y24	545	Stn381	1.54	1.29	0.0326	0.0005
87	Morphometric	Y25	545	Stn381	2.28	1.91	0.0040	0.0005
88	Morphometric	Y26	545	Stn381	2.10	1.76	0.0060	0.0005
89	Morphometric	Y27	545	Stn381	1.93	1.62	0.0110	0.0005
90	Morphometric	Y28	545	SNP1	2.26	1.89	0.0058	0.0005
91	Morphometric	Y29	545	Stn381	1.22	1.03	0.0632	0.0005

Chapter 3. PHOSPHORUS LIMITATION DOES NOT DRIVE LOSS OF BONY LATERAL PLATES IN FRESHWATER STICKLEBACK

3.1 SUMMARY

Connecting the selective forces that drive evolution of phenotypes to their underlying genotypes is key to understanding adaptation to changing environments. However, there are few cases in which the selective force acting on an adaptive allele has been tested experimentally. Threespine stickleback (*Gasterosteus aculeatus*) are a powerful model for such studies because genotypes with phenotypic effects and putative adaptive value have been identified. For example, a regulatory mutation in the *Ectodysplasin* (*Eda*) gene causes a reduction in the number of bony armor plates. Loss of plates occurs both rapidly and repeatedly when marine sticklebacks invade freshwater, but the source of selection on plate loss in freshwater is unknown. Here, we tested whether dietary limitation of phosphorus, which is a key elemental component of bone, can account for the adaptive value of plate loss in freshwater due to a growth advantage of fish with low-plated phenotypes. We crossed marine fish heterozygous for the 16 kilobase freshwater *Eda* haplotype and compared the growth of offspring of different genotypes under phosphorus limitation in both saltwater and freshwater. We found no difference in growth between fish carrying the marine and freshwater *Eda* alleles in any environmental treatment. Moreover, we found no evidence that mechanisms which could mediate the impacts of phosphorus limitation, such as differential phosphorus deposition in bone or excretion rates, are linked to *Eda* genotype. This study highlights the importance of rigorous experimental testing to identify the selective forces acting on genotypes and phenotypes in the wild.

3.2 INTRODUCTION

Understanding and predicting the process of evolution by natural selection requires connecting phenotypes, genotypes and fitness (Barrett and Hoekstra, 2011). Experimental evolution suggests that evolution can be predictable (Lässig *et al.*, 2017). However, predicting phenotypic changes in nature has been difficult due, in part, to the unpredictability of the ecological factors driving selection (Grant and Grant, 2002; Ozgul *et al.*, 2009; Nosil *et al.*, 2018). This difficulty highlights the importance of understanding the ecological mechanisms driving adaptive change. While conceptually straightforward, direct tests of selective agents on the response of genotypes and phenotypes requires both experimental manipulations of selective agents and knowledge of the genetic basis of phenotypes (Wade and Kalisz, 1990; MacColl, 2011).

Threespine stickleback (*Gasterosteus aculeatus*) are an informative model of adaptive evolution due to a century of behavioral and ecological studies, ease of producing large crosses in the laboratory, genomic and genetic tools, and their rapid and repeated adaptation to freshwater, which has resulted in extensive phenotypic diversity (Wootton, 1976; Bell and Foster, 1994; Cresko *et al.*, 2007; Kingsley and Peichel, 2007). Marine stickleback have repeatedly colonized and adapted to freshwater environments across the Northern hemisphere in the last 10,000 - 20,000 years, which has led to repeated and rapid shifts in allele frequencies at the genes under selection (Hohenlohe *et al.*, 2010; Jones *et al.*, 2012a; Jones *et al.*, 2012b). Many of these adaptive alleles are present in the marine population at low frequency and are repeatedly selected in freshwater (Colosimo *et al.*, 2005; Bassham *et al.*, 2018). One of the most consistent allele frequency shifts in response to freshwater environments is at the gene *Ectodysplasin* (*Eda*), which is found in a 16 kb haplotype that is highly divergent between marine and freshwater populations (Hohenlohe *et al.*, 2010; Jones *et al.*, 2012b). *Eda* affects four phenotypes, including number of bony armor

plates, patterning and number of neuromasts in the posterior lateral line (a mechanosensory system), and schooling ability (Chapter 2) (Colosimo *et al.*, 2005; Mills *et al.*, 2014; Greenwood *et al.*, 2016). Lateral plate number is the most intensively studied of these phenotypes, and it is one of many bony armor traits that are reduced or lost in freshwater (Bell and Foster, 1994). Shifts in lateral plate number and *Eda* freshwater allele frequency occur over a few generations in natural and experimental populations, indicative of strong and predictable selection on both the phenotype and the genotype (Bell *et al.*, 2004; Barrett *et al.*, 2008; Kitano *et al.*, 2008; Gelmond *et al.*, 2009; Rennison *et al.*, 2015). This repeatable and strong selection on *Eda* in freshwater environments provides an opportunity to identify the selective agent(s) acting on the *Eda* locus.

The selective agent or agents driving armor loss in freshwater environments remain unclear. Previous work has shown that armor plating is protective against predatory attacks by birds and fish (Reimchen, 1983, 1992), but several other hypotheses about armor evolution have been proposed, including avoidance of insect predators (Reimchen, 1980; Marchinko, 2009), faster escape velocity (Taylor and McPhail, 1986; Bergstrom, 2002), salinity tolerance (Heuts, 1947; Marchinko *et al.*, 2007), increase in buoyancy by reducing body density (Myhre and Klepaker, 2009), and a growth trade-off with building bone in environments with low ion concentrations (Giles, 1983; Bell *et al.*, 1993; Bourgeois *et al.*, 1994; Marchinko *et al.*, 2007; Barrett *et al.*, 2008, 2009a). Although a few direct tests of a single selective agent on armor plate loss have been performed (Marchinko, 2009; Rennison *et al.*, 2019), none have isolated the effects of the *Eda* allele in a relevant (marine) genomic background.

Previous experiments in freshwater ponds have manipulated *Eda* genotype in a marine genomic background using crosses from populations which are polymorphic for lateral plate phenotypes and *Eda* genotypes. In both laboratory and pond experiments, low-plated fish have a

growth advantage over completely-plated fish in freshwater, and this is partially mediated by *Eda* genotype (Marchinko *et al.*, 2007; Barrett *et al.*, 2008, 2009a; Rennison *et al.*, 2015). Rapid juvenile growth can be beneficial for predator avoidance (Foster *et al.*, 1988; Marchinko, 2009) and overwinter survival (Carlson *et al.*, 2010), while adult size is positively correlated with increased reproductive output (reviewed by Baker *et al.* (2015)), suggesting that growth rate mediated by genotype at *Eda* could be a target of selection in freshwater sticklebacks.

What could cause the growth differential between *Eda* genotypes in freshwater (Marchinko *et al.*, 2007; Barrett *et al.*, 2009a)? Bone is primarily composed of hydroxyapatite, a calcium- and phosphorus-rich crystal, and accounts for most of the calcium and phosphorus in fishes (Hendrixson *et al.*, 2007). Both calcium and phosphorus are less abundant in freshwater habitats than in marine environments. Calcium and phosphorus availability in freshwater stickleback environments is variable, and Bourgeois *et al.* (1994) found that they were among the top predictors of number of lateral plates in freshwater stickleback populations from the Cook Inlet of Alaska. Both calcium and phosphorus are therefore reasonable selective agents to test for their effects on growth rate, but only phosphorus has been found to limit the growth of consumers in freshwater environments (Elser *et al.*, 2000). Growth is particularly phosphorus demanding due to its requirement of P-rich RNA for transcription and translation and may be of particular importance during early juvenile growth (Sterner and Elser, 2002; Boros *et al.*, 2015). Indeed RNA content tracks with specific growth rate in many invertebrate and vertebrate consumers and dietary phosphorus limitation causes growth and mineralization defects in trout (Acharya *et al.*, 2004; Hessen *et al.*, 2013; Boros *et al.*, 2015; Witten *et al.*, 2016).

We hypothesized that dietary phosphorus is a limiting element for completely-plated marine stickleback in freshwater environments. We predict that this limitation imposes a cost on these fish

that could present in a variety of ways, including as a reduction in (i) survival, (ii) somatic growth, (iii) bone mineral density, (iv) P excretion rate, or an increase in (v) absorption rates via longer intestine lengths. Further, we hypothesized that reduction in lateral plate count would alleviate phosphorus demands on marine fish in freshwater, and therefore marine fish homozygous for the freshwater allele at *Eda* would not display these costs. To test our hypotheses, we caught marine stickleback that were heterozygous at *Eda*. We made F2 crosses between heterozygous F1 parents and split families between four conditions of water and diet: saltwater and near freshwater plus high phosphorus and low phosphorus pellet diet. We compared the survival and growth rates of fish that carried homozygous marine, heterozygous, and homozygous freshwater alleles at *Eda*, both before and during bony plate growth.

3.3 Materials and Methods

Ethics statement

Marine stickleback were collected under the Washington Department of Fish and Wildlife scientific collection permits 15-033, 15-213, and 16-066. Animal care and handling procedures were approved by the Fred Hutchinson Cancer Research Center Institutional Animal Care and Use Committee (protocol 1575) or the Veterinary Service of the Department of Agriculture and Nature of the Canton of Bern (VTHa# BE4/16 and BE82/17).

Puget Sound fish collection and crosses

Wild marine fish were collected in Puget Sound, WA, in the summers of 2015 and 2016 as previously described (Chapter 2). In 2015, fish were collected in at least 30 meters of water during a multi-day trawling trip in the Whidbey Basin and Bellingham Bay areas. In 2016, fish were collected with the help of the Washington Department of Fish and Wildlife in a beach seine in

Clam Bay, near Manchester, WA. Wild-caught fish were transferred to the Fred Hutchinson Cancer Research Center and initially genotyped at a single marker within the *Ectodysplasin* (*Eda*) haplotype (Stn382, see "DNA EXTRACTIONS AND GENOTYPING*"* below). Fish carrying one or two freshwater alleles at this marker were uniquely tagged using spine clips and elastomer tagging. Subsequent genotyping of these carriers revealed that some fish carried a partial freshwater haplotype, while other fish were heterozygous at all markers within the haplotype. 30 crosses between these heterozygous fish carrying 37 unique freshwater haplotypes were made and subsequently shipped to the stickleback facility at the University of Bern, Switzerland, in the fall of 2016. These F1 fish were grown to reproductive maturity and genotyped to identify heterozygous carriers of the full 16 kb haplotype (i.e. fish that were heterozygous at all markers within the haplotype, Figure 3.1). Of the 37 unique freshwater haplotypes transmitted to the F1 generation, 24 of these were full 16 kb haplotypes and transmitted to the F2 experimental generation.

FISH HOUSING AND CARE

Wild-caught and F1 fish were housed in Seattle and Bern in 3.5 parts per thousand (ppt) saltwater, as previously described (Chapter 2). F2 crosses were made as described below (see "F2 EXPERIMENTAL CROSS SETUP") and housed in either full (35 ppt) saltwater or reduced (3.5 ppt) saltwater conditions (referred to as saltwater and freshwater, respectively, from here on). Fish were kept in 100-liter tanks on saltwater or freshwater recirculating systems. Conductivity was automatically monitored and maintained at 41.5 and 5.3 millisiemens/cm for the saltwater and freshwater systems, respectively, using a saturated solution of Instant Ocean sea salt (Instant Ocean, Aquarium Systems, Sarrebourg, France). Water temperature (15˚C) and pH (7.8 - 8.0) were monitored and maintained automatically for both the saltwater and freshwater systems. Lighting

was programmed with 11 hours full sunlight, 1 hour sunrise, 1 hour sunset, and a moon light for nighttime for both recirculating systems.

Adult wild-caught and F1 fish were fed a mixture of freshly hatched artemia nauplii every morning and frozen mysis shrimp in the afternoons, three times per week. Newly hatched larvae and juvenile F1 fish were fed artemia nauplii twice per day. Heterozygous F1s were fed mysis shrimp every afternoon to induce reproduction prior to the start of the experiment.

DNA EXTRACTIONS AND GENOTYPING ASSAYS

DNA was extracted from fin tissue using a modified HotSHOT DNA extraction method (modified from Meeker *et al.* (2007) as previously described in Chapter 2). Wild-caught fish were genotyped at markers listed in Table 2.1. SNP4, SNP11110 and the NAKA SNP are segregating polymorphisms in the wild population. Although not useful for determining whether a fish carries a marine or a freshwater *Eda* haplotype, these SNPs were informative for distinguishing among the haplotypes in the F1s. F1 fish were genotyped at Stn382, IDH or LRR (sex markers; see Peichel *et al.* (2004) and Table 2.1, respectively), and any informative markers that would distinguish the transmitted haplotype (i.e. SNPs12/13 distinguish short haplotypes from full haplotypes). Heterozygous carriers of full haplotypes were marked within each cross as potential parents for the F2 experimental crosses.

F2 experimental fish were genotyped at Stn382 and IDH (or LRR) to identify their genotype at the *Eda* haplotype and their sex, respectively.

F2 EXPERIMENTAL CROSS SETUP

Experimental F2 crosses were made between F1 *Eda* heterozygotes, so that the resulting offspring would have a 1:2:1 ratio of homozygous marine:heterozygous:homozygous freshwater genotypes

at *Eda*. The marine allele is referred to as C, and CC homozygotes have a completely-plated phenotype of 31 - 36 plates per side of the fish. L alleles are usually found in freshwater populations and LL fish are low-plated with between 0 - 10 plates per side (Figure 3.1). Crosses of at least 102 eggs were fertilized, separated into groups of ~10 embryos for better nutrient exchange and kept in petri dishes containing clean freshwater for 2 days. Water was exchanged daily, and after 2 - 3 days, unfertilized or dead embryos were discarded, and the remaining developing embryos were split between four clean tanks: freshwater + LowP diet, freshwater + HighP diet, saltwater + LowP diet, and saltwater + HighP diet. Embryos were kept in plastic cups with mesh replacing the bottom of the cup to allow for air and water exchange but confine the embryos and hatchlings to the cup. Embryos generally hatched 9 - 10 days post fertilization (dpf), so between 10 - 12 dpf, hatchlings were counted within each tank, and equal numbers of larvae were released into each of the four tanks for that family. Released hatchlings per tank ranged from 20 - 25 per tank for Experiment 1, and 20 - 38 per tank for Experiments 2 and 3 (see "SAMPLING AND PHENOTYPING OF F2 EXPERIMENTAL FISH" below).

EXPERIMENTAL DIETS

Fish obtain the majority of their phosphorus through their diet (Vielma and Lall, 1998; Avila *et al.*, 2000; Lall, 2002; Sullivan *et al.*, 2007), therefore we designed the experimental diets to differ only in phosphorus content, while being *iso*-caloric, *iso*-nitrogenous, and *iso*-lipidic (Table 3.1). We targeted environmentally relevant levels of phosphorus, albeit near the low end of measured natural foods (Hall *et al.*, 2004; Evans-White *et al.*, 2005). We targeted 0.5% and 1% P in the diets, however, testing of the pellet diets found them to be 0.48% and 0.61%. These restrictive levels of phosphorus are closer together than expected, yet they nonetheless provide the test of the relevant

question: does dietary phosphorus restriction impose a growth cost on stickleback genetically programmed to build more bony lateral plates?

Experimental F2 fish were fed live artemia twice per day on 11 - 13 dpf (approximately 1 day after hatching). At 14 dpf, fish were started on their experimental diets. Larvae were given experimental diet in the morning and received live artemia in the afternoon at 14 - 16 dpf. At 17 dpf, fish were fed pellet food twice per day, and no longer received any artemia. The amount of pellet food provided per day was estimated to be 4% of body weight per day based on number of larvae in the tank and was not adjusted as fish died. Therefore, fish received food effectively ad libitum.

Pellet diets were ground and sieved to provide appropriate sizes for different age classes of stickleback. Larval fish were fed ground pellets that were smaller than 200 µm until 87 dpf. At 88 dpf, fish were moved to pellets sized between 320 and 500 µm.

SAMPLING AND PHENOTYPING OF F2 EXPERIMENTAL FISH

In our crosses, new plates start forming at approximately 13 mm standard length (SL), and continue to form until 25 mm SL. After this stage, plates continue to grow in size, but new plates do not form (Figure 3.2). Fish from Experiment 1 were sampled between 37 and 39 dpf, so that fish were under 13.1 mm. Fish from Experiment 2 were sampled between 128 and 132 dpf, so that the majority of fish in the freshwater tanks were between 21 and 30 mm SL. Fish from Experiment 3 were sampled between 155 and 175 dpf when freshwater fish were between 18 and 55 mm SL.

Fish were euthanized with a lethal dose of MS-222. Experiment 1 fish were measured with manual calipers for SL to the nearest 0.05 mm and weighed to the nearest 0.01 mg. The caudal fin (and sometimes a portion of the posterior trunk) was sampled and placed in 95% ethanol for later DNA extraction. Experiment 2 fish were measured to the nearest 0.05 mm with manual calipers

for SL, weighed to the nearest mg, and photographed. A portion of the caudal fin was stored in ethanol for DNA extraction. Experiment 3 fish were placed into an excretion collection experiment (see "EXCRETION COLLECTION AND MEASUREMENT" below), then euthanized, measured to the nearest 0.05 mm, weighed to the nearest mg, and then frozen at -20°C. Later, Experiment 3 fish were thawed in order to remove the pelvic girdle for phosphorus analysis. Additionally, the intestine was dissected, measured to the nearest 0.05 mm, and weighed to the nearest 0.01 mg. The liver and spleens were also weighed to the nearest 0.01 mg for most Experiment 3 fish, and the plates on the left side of the body were counted.

EXCRETION COLLECTION AND MEASUREMENT

Fish were fed 2 hours before the experiment began. They were left in the tank with the food for 1 hour, then caught and held in buckets until the start of the experiment. The experiment was carried out in a climate-controlled room at 15°C using custom made chambers. These chambers were Ziploc® Twist/n Loc® 16 ounce containers (Johnson and Son, Inc., Racine,WI) with a hole in the lid through which a hose for water exchange could pass.

System water (freshwater and saltwater) was filtered to 0.2 μm using a gravity filter (Platypus GravityWorks). The hoses and chambers for excretion collection were rinsed with DI water from the tap and then filled with 300 mL of filtered system water. Fish introduction to chambers was staggered by 1 minute (2 minutes if there were not as many fish to process) to allow for sample handling between fish. Up to 30 chambers were processed in a set (26 fish, 2 freshwater controls and 2 saltwater controls). Samples from each chamber (experimental and controls) were taken at 0 (just after fish introduction), 60, and 120 minutes.

Phosphorus does not homogenize as well as ammonium in the chamber, so unmixed samples can be very inaccurate. We spent approximately 40 seconds slowly withdrawing and then

reinserting two full syringes (60 mL each) of water to mix the chambers before taking the 30 mL sample on the third pull. The 30 mL sample was then handled in the following way. First, a filter cartridge was attached to the syringe containing a GF-75 glass fiber filter (0.3 μm pore, AM.D. Manufacturing Inc., Mississauga, Ontario, Canada). Next, the filter was rinsed by pushing 5 mL of the sample through the filter cartridge. The collection tube was then rinsed twice with 5 mL of the sample, and then the final 15 mL of sample was filtered into the collection tube and sealed. Once all the samples were completed, the samples were frozen and stored at -20°C until analysis.

BONE AND EXCRETION PHOSPHORUS MEASUREMENTS

All water and bone samples for phosphorus analysis were digested and oxidized using a peroxydisulfate solution (10 grams of $K_2S_2O_8$ and 1.5 grams of NaOH in 1 liter of water), and analyzed on a Skalar San++ Continuous Flow P/N analyzer (as in Leal et al. (2017)).

SURVIVAL ANALYSES

Survival was calculated for each tank from the number of hatched larvae released into the tank and the number of live fish collected at the end of each experiment. Differences in survival between conditions were analyzed using a linear mixed effects model with family and tank as random effects, and water, diet and water x diet interaction as fixed effects.

We also tested whether the CC, CL and LL genotypes had differential survival during the experiment. The null hypothesis was that there is no difference in survival between genotypes. Assuming that each tank initially received a 1:2:1 ratio of genotypes, our expected ratios of genotypes at the end of the experiment under the null hypothesis was 1:2:1. Due to low numbers of surviving fish in most experiments (n often < 20 except in FW tanks in Experiment 1), the surviving fish were summed within genotype, but across families within each condition to produce

a three genotype (CC, CL, LL) x four condition table (FW-HighP, FW-LowP, SW-HighP, SW-LowP). Survival was then analyzed in two ways: 1) a chi-square analysis across the 3x4 table to see if survival of genotypes varied by condition; and 2) a chi-square analysis comparing the observed genotype ratios within each condition to the 1:2:1 expected ratio.

GROWTH, INTESTINE LENGTH, AND BONE PHOSPHORUS ANALYSES

Growth was measured in two ways for fish from all three experiments: SL (mm) and body condition. Intestine length and bone phosphorus content (%P) were measured in Experiment 3 only. Each of these measures were then analyzed for differences between conditions and genotypes using a linear mixed effects model with family and tank as random effects, and water, diet, sex, genotype, and the interactions between water x genotype, water x diet, and diet x genotype as fixed effects. Analyses of fish from Experiment 3 also included age as a fixed effect, since the fish ranged in age between 155 - 175 dpf. Body condition and intestine length were analyzed by running the model for the natural log of weight (or intestine length) and including the natural log of standard length as an additional fixed effect.

EXCRETION DATA ANALYSES

Excretion rates were calculated for the 0 - 60 minute sampling interval as well as the 60 - 120 minute sampling interval using the equation:

$$Excretion\ rate = \frac{P_f * V_i - P_i * V_i}{t} \quad (3.1)$$

where P_i and P_f are the μg/L phosphorus in the water at the start and end of the sampling interval, respectively, V_i is the volume of water (in liters) at the start of the sampling interval (before the end sample was removed), and t is the length of the sampling interval in minutes. The phosphorus excretion rate is therefore reported as μg/minute.

During data exploration, we found that the excretion measurements varied from batch to batch on the Skalar San++ Continuous Flow P/N analyzer. We also found that excretion rates in the second time interval (60 - 120 minutes) were lower than in the first time interval (0 – 60 minutes). Therefore, the data were trimmed to only include samples from the first time interval, and in which the 60 minute sample was analyzed in the same batch as the 0 minute sample for each individual fish. Therefore, effects of time interval and analysis batch were removed.

Exactly as for the analysis of bone phosphorus content, excretion rates for the 0 – 60 minute time interval were analyzed for differences between conditions and genotypes using a linear mixed effects model. Family and collection day were set as random effects; water, diet, sex, weight, and genotype were fixed effects; and the interactions between water x genotype, water x diet, and diet x genotype were also fixed effects.

3.4 RESULTS

PLATE NUMBER IS CONTROLLED BY GENOTYPE AT *EDA*

The *Eda* locus controls between 74 - 100% of the variation in number of plates, depending on the genetic modifiers present in the population or cross (Avise, 1976; Colosimo *et al.*, 2005). In the F2 fish in Experiment 3, fish homozygous for the marine allele (CC) are completely-plated, fish homozygous for the freshwater allele (LL) are low-plated, and heterozygous fish (CL) are phenotypically indistinguishable from CC fish (Figure 3.1). Genotype at the *Eda* haplotype controls 93% of the variation in the number of lateral plates in this study.

HATCHING RATES EQUAL BETWEEN SALTWATER AND FRESHWATER CONDITIONS

We observed no differences in hatching rates (% eggs that hatched) between saltwater and freshwater conditions (Figure 3.3). Hatching rates were analyzed separately for Experiment 1,

since this experiment was run earlier than Experiments 2 and 3, which were setup and hatched concurrently. Hatching rates in Experiment 1 were 95% in both freshwater and saltwater ($F_{1,56}$ = 0.1428, p = 0.70). Hatching rates in Experiments 2 and 3 were 93% and 94% in freshwater and saltwater, respectively ($F_{1,130}$ = 0.1762, p = 0.6754).

SURVIVAL HIGHER IN FRESHWATER THAN IN SALTWATER

Water salinity, but not diet, affected survival rates during all three experiments, and survival decreased over time (Figure 3.4). In Experiment 1, which lasted 37 days, survival was 84% and 55% in freshwater and saltwater, respectively ($F_{1,54}$ = 64.9, p = 8.0e-11), while diet had no effect ($F_{1,54}$ = 0.3861, p = 0.5370). Survival in Experiment 2 was 45% and 30% in freshwater and saltwater, respectively ($F_{1,78}$ = 10.9, p = 0.001), and again diet had no effect ($F_{1,78}$ = 0.5499, p = 0.460596). Survival in Experiment 3, which lasted between 155 and 175 days, was 60% and 12% in freshwater and saltwater, respectively ($F_{1,20}$ = 67.1, p = 8.1e-08), with no significant effect of diet ($F_{1,20}$ = 2.8115, p = 0.1092).

NO DIFFERENTIAL SURVIVAL OF *EDA* GENOTYPES

We also analyzed the ratios of genotypes that were alive at the end of each experiment, and found no difference in the ratios of genotypes across the conditions (Experiment 1: χ_{26} = 3.6, p = 0.73; Experiment 2: χ_{26} = 5.5, p = 0.48; Experiment 3: χ_{26} = 3.4, p = 0.76). Comparing of surviving genotypes in each condition, we also found no deviation from the expected 1:2:1 ratio of genotypes (Figure 3.5, Figure 3.6, and Figure 3.7).

NO DIFFERENTIAL GROWTH OF *EDA* GENOTYPES PRE-PLATE DEVELOPMENT (EXPERIMENT 1)

Experiment 1 represented a timepoint before the lateral bony plates began to develop. The results of the mixed effects linear model revealed significant effects of water on SL and body condition (the natural log of weight with natural log of SL in the model) at this timepoint (Table 3.2, Figure 3.8). Fish were larger and had higher body condition if they were grown in saltwater. In addition, there was a significant effect of sex on SL; males were larger. Overall, there was no effect of *Eda* genotype alone on any measure of growth at this timepoint before the plates have started to develop.

NO DIFFERENTIAL GROWTH OF *EDA* GENOTYPES DURING PLATE DEVELOPMENT (EXPERIMENTS 2 & 3)

Fish were an average of 27.85 mm long (range 13.05 - 44.00 mm) at the end of Experiment 2. Plate number is generally set by 25 mm SL (Figure 3.2); importantly LL fish have fewer plates than CC and CL fish. Therefore, we considered this a reasonable time to assess the differential effects of genotype on growth. However, the mixed effects linear models for SL and body condition show no effect of genotype on size (Table 3.3, Figure 3.9). Water salinity had a significant effect on SL (but not on body condition); however, the fish raised in freshwater were larger than fish grown in saltwater - a reversal from Experiment 1.

At the end of Experiment 3, fish were an average of 34.59 mm long (range 15.35 - 50.40 mm). The results of the linear mixed effect model suggest that water and sex have significant effects on size (Table 3.4, Figure 3.10). Similar to Experiment 2, Experiment 3 fish grown in freshwater were larger than their siblings grown in saltwater. Additionally, males were larger than females. At this stage, genotype has an effect on size as measured by SL ($F_{2,259} = 3.02$, $p = 0.05$) or body condition ($F_{2,259} = 4.14$, $p = 0.02$). However, contrary to our hypothesis, CC fish were larger than their

siblings (CL and LL fish were on average 1.98 and 1.77 mm smaller than CC fish in SL, respectively) and had a higher body condition. This effect is even greater in saltwater, where LL fish have an even lower body condition (water x genotype interaction) but these effects are slight, at best, and hard to see, likely due to low numbers of surviving fish (Figure 3.10).

NO DIFFERENCE IN PHOSPHORUS CONTENT OF BONE BY *EDA* GENOTYPE (EXPERIMENT 3)

There was no effect of diet or water on the number of lateral plates at the end of experiment 3 ($F_{1,261} = 0.33$, $p = 0.56$ and $F_{1,263} = 1.04$, $p = 0.31$, respectively), suggesting that there was not a plastic response of plate development to the environmental treatments. However, to further investigate whether phosphorus deposition in bony structures changed in response to the treatments and/or due to *Eda* genotype, we also analyzed the phosphorus content of the bones of the pelvic girdle. The results of the mixed-effects linear model show that water and age have significant effects on percent phosphorus (%P) of the bone and sex has a nearly significant effect (Table 3.5, Figure 3.11). Fish grown in saltwater have lower %P bone content, age is positively correlated with %P content, and males have lower %P bone content than females. Importantly, phosphorus content of the pelvic girdle bones does not differ between fish with different genotypes.

NO DIFFERENCE IN INTESTINE LENGTHS BY *EDA* GENOTYPE (EXPERIMENT 3)

Both interspecific and intraspecific intestine length is associated with nutrient composition of diet in cichlids (Wagner *et al.*, 2009), so we examined whether intestine length varied in response to either lower phosphorus content in the diet or for higher phosphorus demands in the bone driven by genotype. We found that intestine length was strongly positively correlated with standard length, and there was no effect of *Eda* genotype on intestine length (Table 3.5).

NO DIFFERENCE IN EXCRETION RATES BY *EDA* GENOTYPE (EXPERIMENT 3)

Fish may compensate for dietary limitations by adjusting their excretion rates, and consequently the retention rates of limiting elements. We therefore compared excretion rates of fish across conditions and genotypes. The mixed-effects linear model shows that water and diet have significant effects on phosphorus excretion rate (Table 3.6, Figure 3.12). Fish reared in saltwater excrete less phosphorus, while fish fed the HighP diet excrete more phosphorus. Importantly, phosphorus excretion rates do not differ between fish with different genotypes at *Eda*.

3.5 DISCUSSION

Ancient haplotypes containing the gene *Ectodysplasin* (*Eda*) control bony plate number in threespine stickleback. Fish carrying the common marine allele, C, are completely-plated with 30 - 36 plates per side. Fish homozygous for the freshwater *Eda* allele, L, are low-plated with between 0 - 10 plates per side. Previous work has found growth differences between CC and LL fish in freshwater but have not identified what causes this growth differential (Barrett *et al.*, 2009a). In both wild populations and in lab-reared fish, CC fish have higher %P than LL adults (El-Sabaawi *et al.*, 2016; Durston and El-Sabaawi, 2017; Leal *et al.*, 2017) and consume diets with different %P content to match those demands (Durston and El-Sabaawi, 2019). We therefore asked whether the growth differential of marine fish with different alleles at *Eda* was driven by their differential demands for dietary phosphorus. However, we did not detect growth differences between fish carrying different alleles at *Eda* under reduced phosphorus availability.

Ecological stoichiometry predicts that when an element is limiting, consumers will be forced to cope with that limitation. Mechanisms for coping with a limitation include (i) survival, (ii) growth, (iii) allocation to tissues (e.g. bone density), (iv) excretion and (v) element assimilation (e.g. intestine length). We found no differences in (i) survival or (ii) growth rate driven by diet or

genotype at either developmental stage. We did observe a significantly lower survival in saltwater, and an interaction between water and life stage on growth rates, with faster larval growth in saltwater and faster juvenile growth in freshwater. In addition, we did not detect differences in (iii) bone mineral content, (iv) P excretion rates or (v) intestine lengths between sibling fish with different genotypes at *Eda*. Together, these data suggest that phosphorus availability does not impose a cost on marine stickleback in freshwater and is unlikely to be driving evolutionary changes in genotype at *Eda*, the major effect locus underlying plate phenotype.

In polymorphic freshwater populations of stickleback, CC fish at *Eda* forage in littoral habitats possibly allowing them to consume more P-rich foods than fish LL, suggesting that P limitation may be driving behavioral shifts in feeding to compensate for low P availability in freshwater plus high demand for P due to bony armor (Durston and El-Sabaawi, 2019). Therefore, consumption rate could be an alternate coping mechanism for a dietary limitation. It is possible that our lab-reared fish with different genotypes employed differential feeding strategies to compensate for phosphorus demands. This is possible because we did not remove uneaten fish after feeding. Instead, we fed each tank a measured supply of food based on approximately 4% of body weight per fish per day, which appeared to be in excess. In addition, algae was available to both saltwater and freshwater fish, but the %P composition of the algae ranged from 0.37 - 0.47% in freshwater tanks and 0.14 - 0.21% in saltwater tanks. Algae in freshwater tanks was both more abundant and contained a higher %P than algae in saltwater tanks, yet the %P of the algae was lower than the experimental diets. Therefore, we do not think that consumption rate was a coping mechanism utilized in our experiment to compensate for reduced phosphorus availability.

We conclude that limited phosphorus availability is not the selective agent favoring plate loss in freshwater habitats. This could be the case for several reasons. First, stickleback may not be

phosphorus limited in the wild or in our experiment. Most fish thought to be living under P-limited growth conditions are herbivores (Hood *et al.*, 2005; Frost *et al.*, 2006), though P limitation is predicted to occur occasionally in carnivorous and insectivorous fishes as well (Schindler and Eby, 1997; McIntyre and Flecker, 2010). Wild stickleback forage in habitats that match their body composition, particularly the N:P ratio, suggesting that armor plates may be associated with a dietary limitation. Interestingly in our experiment, the fish fed a HighP diet (0.61% P) had a higher phosphorus excretion rate than fish fed the LowP diet (0.48% P) (Figure 3.12). Both of these diets are relatively low in phosphorus compared to wild prey items (0.4 - 2.2% P), and an increase in phosphorus excretion with an increase in %P diet in this range is consistent with either no phosphorus limitation at these dietary levels, or an inability to adjust the absorption rate of dietary phosphorus. If absorption rates are static, then we would expect to see growth differentials between fish with different *Eda* genotypes unless growth limitation is driven by something other than phosphorus. This hypothesis is consistent with the observation that P excretion rates are not inversely related to body P content in the wild, as is expected under limiting dietary P conditions (Durston and El-Sabaawi, 2019; Rudman *et al.*, 2019).

Second, it is possible that the previously identified growth advantage of *Eda* LL sticklebacks in freshwater may have been driven by a gene or genes linked to *Eda* and not by *Eda* itself. In our study, we found no evidence for increased growth of *Eda* LL fish in either freshwater or saltwater. We precisely defined the boundaries of the freshwater *Eda* haplotypes used in the crosses, whereas the previous studies did not determine whether additional freshwater alleles were in linkage disequilibrium with the *Eda* freshwater haplotype (Marchinko *et al.*, 2007; Barrett *et al.*, 2008, 2009a). The authors of these studies raise the possibility that growth trade-offs could be facilitated by a neighboring, downstream gene, *Gjb1*, that was unlinked from *Eda* in about half of the F0 fish

used in our crosses (Chapter 2). Consistent with this, selection on *Eda* in the freshwater pond experiment (Barrett *et al.*, 2008) was attributed to both direct selection for plate loss and independent selection on phenotypes mediated by pleiotropic effects of *Eda* and/or linked gene(s) (Rennison *et al.*, 2015).

Direct selection on plate loss in freshwater (Rennison *et al.*, 2015) suggests that there may still be a trade-off between bone development and growth in freshwater. The majority of calcium, like phosphorus, is found in the bone, and calcium homeostasis is critical for a range of cellular processes. Our data are not inconsistent with calcium being the limiting element in freshwater stickleback. The water in Switzerland contains high levels of calcium (> 20 ppm), and in addition, our freshwater laboratory condition was 0.35ppt seawater, approximately 1.4 ppm calcium. The salinity and calcium content of the water conditions is an important difference between our study and previous studies that used 0ppt for the freshwater conditions in Canada, where freshwater calcium levels are lower (~4 ppm) (Marchinko *et al.*, 2007; Barrett *et al.*, 2009a; Weyhenmeyer *et al.*, 2019). This has two potentially important consequences. First, our freshwater condition was unlikely to mimic the calcium levels of most freshwater environments, thereby not imposing a calcium limitation. And secondly, calcium levels (and perhaps other trace elements) were higher in our saltwater condition than most marine (or freshwater) environments, potentially outside natural ranges. The former may explain why we did not replicate a growth trade-off with bone in freshwater, and the latter might explain why we observed poor survival and growth of fish in our saltwater condition.

It is also important to consider the limitations of the study and whether these limitation in the experiment could explain our results. For example, it is possible that we did not measure fish at the correct timepoint to detect a difference between genotypes. For example, fish in saltwater grew

more slowly, and were therefore smaller with fewer plates at the end of Experiment 2 (mean SL = 22 mm, range = 13 - 32 mm, Figure 3.2 for relationship between size and number of plates). Therefore, fish in saltwater may not have experienced strong differential P demands by this timepoint. However, we were primarily interested in the selective agent(s) acting in freshwater, and the fish in freshwater at the end of Experiment 2 had different numbers of lateral plates and most were at the size when plate development is complete (mean SL = 32 mm, range = 13 – 44 mm; Figure 3.2). Since number of lateral plates is a strong predictor of total body P in the field (El-Sabaawi *et al.*, 2016; Durston and El-Sabaawi, 2017), we likely measured the freshwater fish while they were experiencing differential demands for P, driven by their genotype at *Eda*. In addition, our sampled standard lengths were in between the sizes sampled by Barrett *et al.* (2009a)(mean SL = 27 mm for juvenile growth, 44 mm for adult timepoint), and they reported growth rate differences by genotype in freshwater at both timepoints. Therefore, we likely selected the right stage to detect a growth differential due to phosphorus limitation.

Another possibility is that our targeted dietary P limitation was either too strict or not strict enough. The difference between the high (0.61% P) and low (0.48% P) diets was not as large as originally planned; nonetheless, no effect of genotype was seen in either diet treatment. Diets available to stickleback in freshwater environments likely range from ~0.4 - 2.2% P (Hall *et al.*, 2004; Evans-White *et al.*, 2005). The growth deficits observed in previous laboratory experiments with stickleback in freshwater resulted from diets of artemia, *Daphnia* and/or bloodworms (Marchinko *et al.*, 2007; Barrett *et al.*, 2009a), all of which have estimated %P content of 1.1 - 1.3% (Andersen and Hessen, 1991; Hall *et al.*, 2004; El-Sabaawi *et al.*, 2016). Additionally, studies in trout and haddock observed growth trade offs with dietary P contents of 0.5% and 0.42%, respectively, (and at least partially rescue it at 1.0% P) (Roy *et al.*, 2002; Witten *et al.*, 2016).

Therefore, our targeted 0.5% LowP diet should have been P-limited and biologically relevant. It is possible that we imposed too strong of a phosphorus limitation. However, the fish grew and developed normally, at least in freshwater (for saltwater, see below). Based on work in trout, we would expect a severe P limitation to cause skeletal deformities (Fontagné *et al.*, 2009; Deschamps *et al.*, 2014; Deschamps *et al.*, 2016; Witten *et al.*, 2016), but a chi-square analysis of "abnormal" vs "normal" fish at the end of Experiment 2 detected no differences in the frequencies of abnormal fish (assessed qualitatively) between the four conditions (SW-HighP, SW-LowP, FW-HighP, FW-LowP; $\chi_2 = 5.31$, df=3, n=1209, p=0.15), suggesting that the P limitation was not too severe.

There is also the possibility that we did not accurately assess the compensatory mechanisms, excretion and assimilation of P. For example, most of the fish used for the excretion analyses were smaller than have been used for other studies of wild-caught fish (El-Sabaawi *et al.*, 2016). Consequently, we observed low values of excreted ug P. In addition, these values dropped in the second hour of excretion (roughly 4 hours after feeding; data not shown). It is possible that we did not measure differences in excretion because of timing of the assays and size of the fish. Additionally, we measured P assimilation in the pelvic girdle, which was robust and accessible on all fish, whereas the plates on most fish were too small to sample reliably. However, it is possible that a difference in assimilation only occurred in the plates themselves.

As discussed, these caveats are unlikely to completely explain our results. We therefore conclude that 0.48% dietary phosphorus is not limiting to marine stickleback, and therefore phosphorus availability is not the selective agent driving plate loss in freshwater stickleback populations. An intriguing result from the work by Barrett *et al.* (2009a) suggests that dietary and water conditions both impose a growth cost to CC fish during juvenile growth. A recent study of stickleback ionomes found that rare elements, particularly barium, differ consistently between

marine and freshwater fish and their native habitats (Rudman *et al.*, 2019), suggesting a broader search for the selective agent(s) are justified.

3.6 CONCLUSIONS

Connecting genotypes to their selective agents is a difficult, yet important challenge for evolutionary biology. Here we empirically tested the role of an abiotic agent (dietary phosphorus) on survival and juvenile growth of marine stickleback carrying different alleles at *Eda*, the locus that controls the bony plate phenotype. We did not find a growth advantage of freshwater *Eda* alleles or low-plated fish in our experiment, suggesting that phosphorus demand does not limit growth of marine stickleback in freshwater. This work highlights the difficulty and importance of directly testing hypotheses of selection.

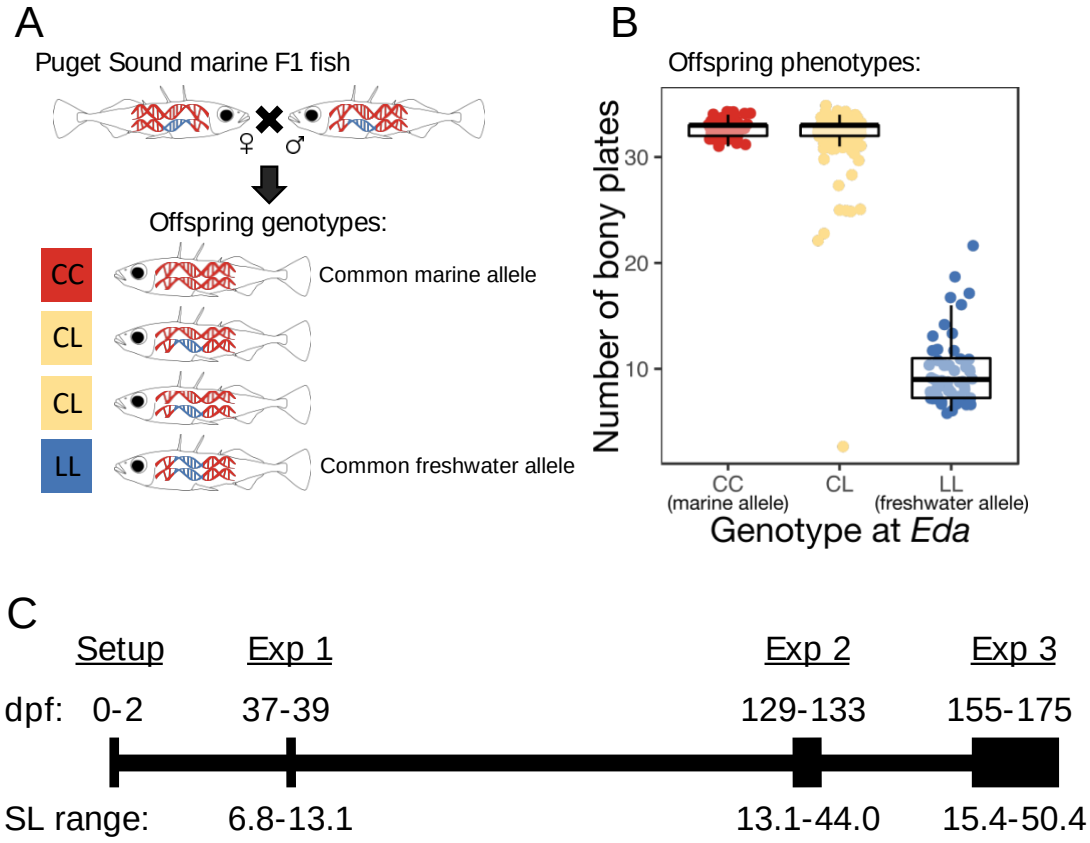


Figure 3.1. Crosses and experimental design for test of dietary phosphorus. (A) Crosses were established between laboratory reared F1 fish heterozygous for the L allele, which produced the expected 1:2:1 ratios of genotypes at *Eda*. (B) The numbers of lateral bony plates were counted on F2 offspring older than 140 days and > 25 mm standard length. Note the single CL fish with 3 plates on the left side; this fish had 31 plates on the right side. (C) For each experiment (Exp) eggs were transferred into experimental tanks at 2 days post fertilization (dpf). Fish were collected, measured and genotyped for Experiment 1 between 37 - 39 dpf (prior to bony plate formation) and for Experiment 2 between 129 - 133 dpf (during and after bony plate formation). For Experiment 3, fish were collected for excretion and bone samples, measured and genotyped between 155 - 175 dpf.

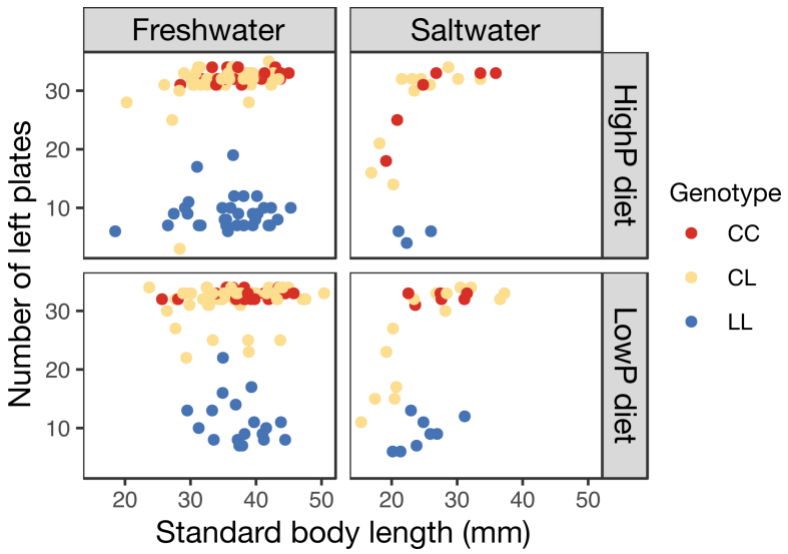


Figure 3.2. Lateral plate number across conditions and ontogeny. Numbers of plates were counted in Experiment 3 fish (155 - 175 dpf) and plotted by standard length. Number of plates is stable in freshwater, since these fish are past the size at which plate development is complete (25 mm). However, in saltwater, smaller fish have fewer plates, particularly fish with genotypes CL and CC, until 25 mm when plate development is complete.

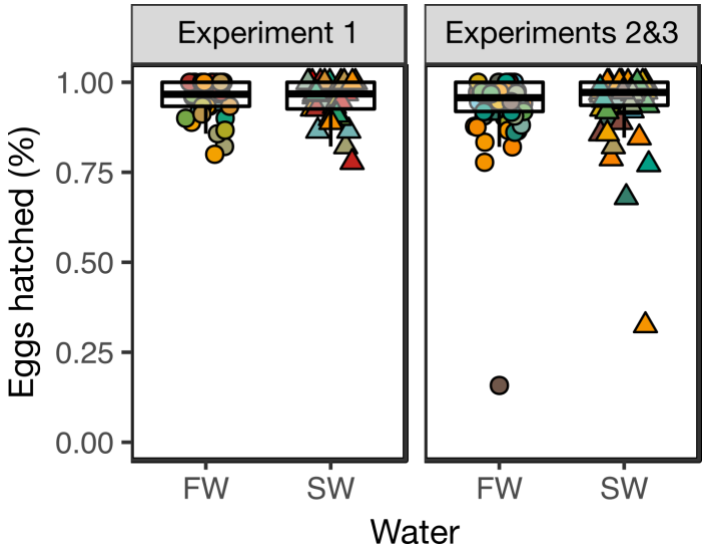


Figure 3.3. Hatching rates across water conditions and experiments. Percentage of eggs that hatched is plotted for each tank and water condition. Families were divided equally between two freshwater (FW) and two saltwater (SW) tanks, and each family is represented by a single color. Experiments 2 and 3 are combined here, because they were started concurrently. There are two tanks with low hatching rates in Experiments 2 and 3, however, these tanks were from different families and different conditions. Overall, hatching rates did not vary between freshwater and saltwater tanks.

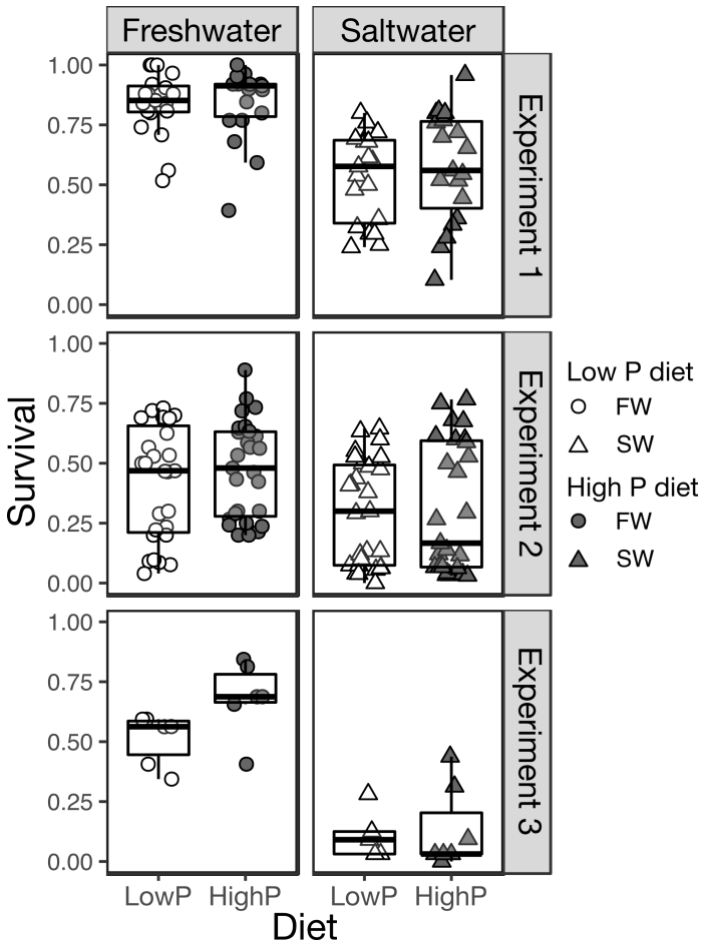


Figure 3.4. Survival across conditions and experiments. Survival varied between freshwater and saltwater tanks, but not between high and low phosphorus diets. Survival is plotted for each family and condition. Survival was significantly lower in saltwater (SW) than in freshwater (FW) in all experiments. Survival was highest in Experimental 1 at 37 days post fertilization (dpf)), lower at 129 - 133 dpf in Experiment 2, and lowest in Experiment 3, especially in SW conditions. There was no effect of diet or water x diet interaction on survival in any of the experiments.

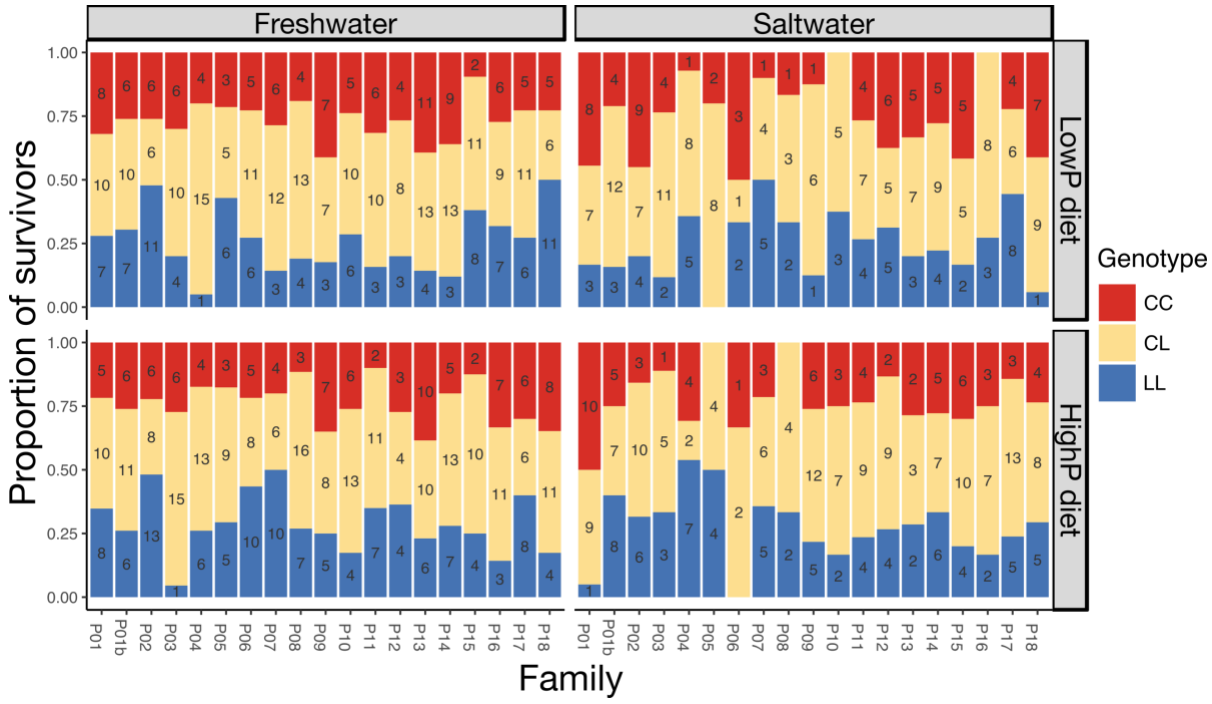


Figure 3.5. Survival of fish by genotype, condition and tank for Experiment 1. Survival is plotted by water, diet, family and genotype for Experiment 1. Numbers of fish of each genotype are shown in the colored bars. Genotype ratios within each condition did not deviate from the expected 1:2:1 ratios (FW-LowP: X-squared = 1.2, df = 2, p-value = 0.54; FW-HighP: X-squared = 3.2, df = 2, p-value = 0.20; SW-LowP: X-squared = 0.8, df = 2, p-value = 0.67; SW-HighP: X-squared = 0.9, df = 2, p-value = 0.65).

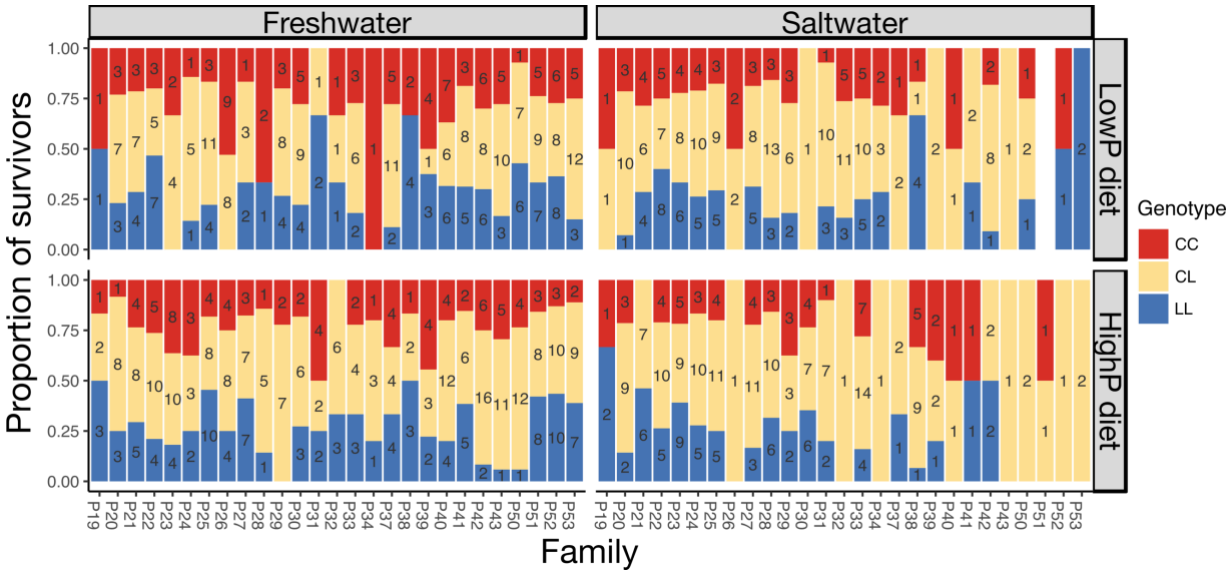


Figure 3.6. Survival of fish by genotype, condition and tank for Experiment 2. Survival is plotted by water, diet, family and genotype for Experiment 2. Numbers of fish of each genotype are shown on the colored bars. Genotype ratios within each condition did not deviate from the expected 1:2:1 ratios (FW-LowP: X-squared = 1.7, df = 2, p-value = 0.42; FW-HighP: X-squared = 2.0, df = 2, p-value = 0.37; SW-LowP: X-squared = 1.5, df = 2, p-value = 0.46; SW-HighP: X-squared = 2.4, df = 2, p-value = 0.30).

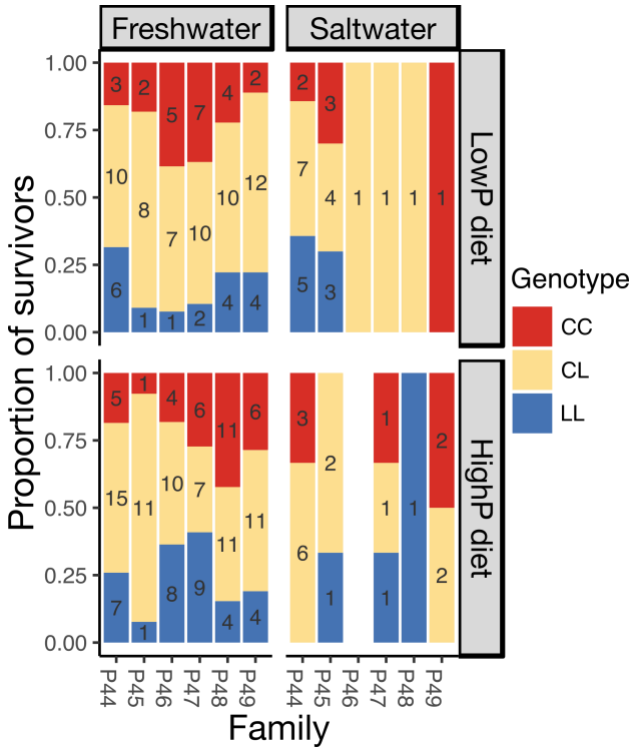


Figure 3.7. Survival of fish by genotype, condition and tank for Experiment 3. Survival is plotted by water, diet, family and genotype for Experiment 3. Numbers of fish of each genotype are shown on the colored bars. Genotype ratios within each condition did not deviate from the expected 1:2:1 ratios (FW-LowP: X-squared = 3.1, df = 2, p-value = 0.21; FW-HighP: X-squared = 0.0, df = 2, p-value = 1.0; SW-LowP: X-squared = 0.3, df = 2, p-value = 0.87; SW-HighP: X-squared = 1.1, df = 2, p-value = 0.58).

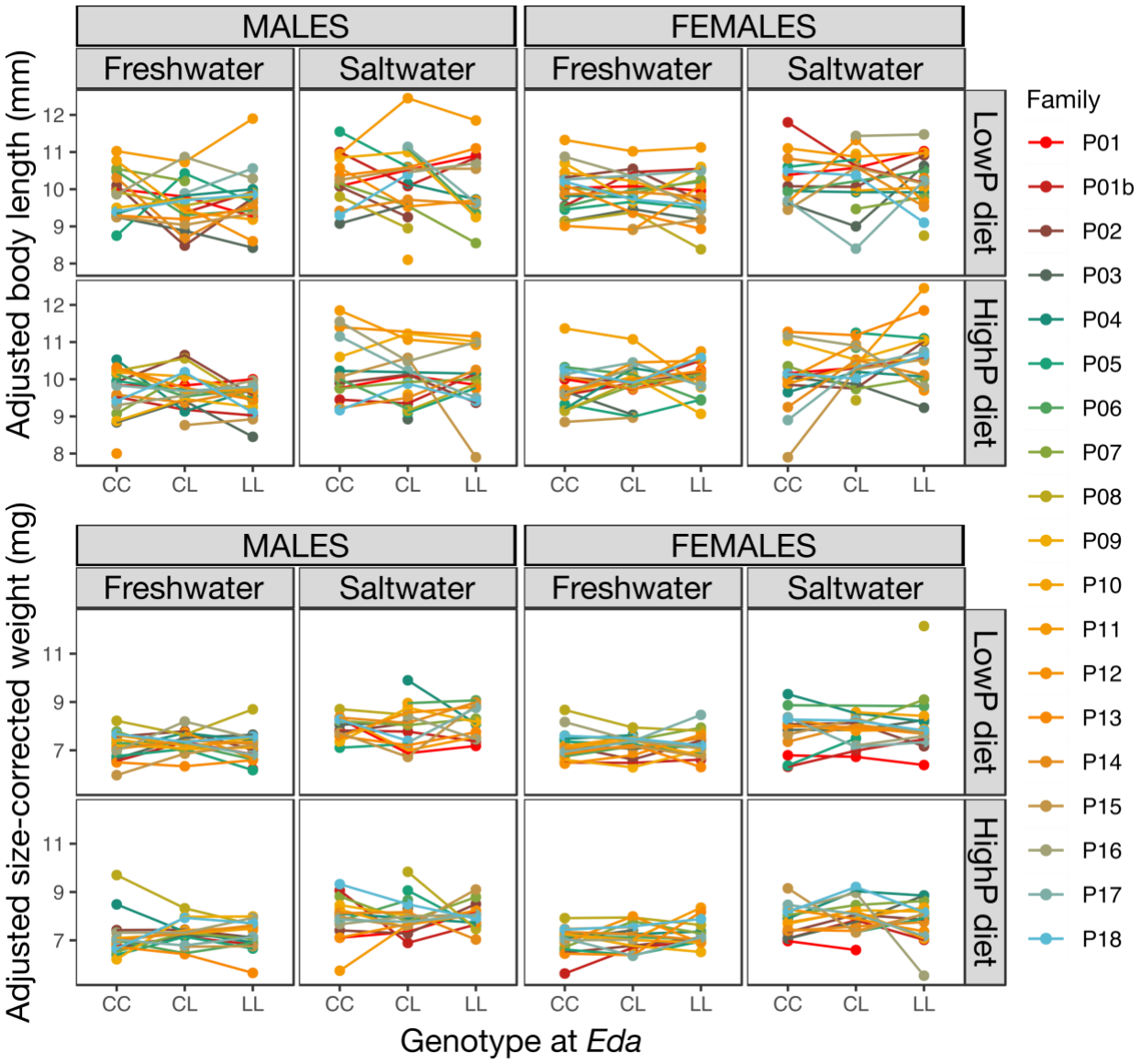


Figure 3.8. Adjusted mean length and size-corrected weight by family, sex, water, diet and genotype in Experiment 1. These results visualize the ANOVA results in Table 3.2. Fish in saltwater tend to be slightly larger and have higher size-corrected weight than fish in freshwater. Additionally, males are slightly smaller than females, but this is hard to see. The most important result is that size or size-corrected weight does not consistently vary between genotypes (x axis). Adjusted trait values were calculated by adding the residual trait value for each individual to the predicted trait value when the covariates are equal (standard length in the case of size-corrected weight).

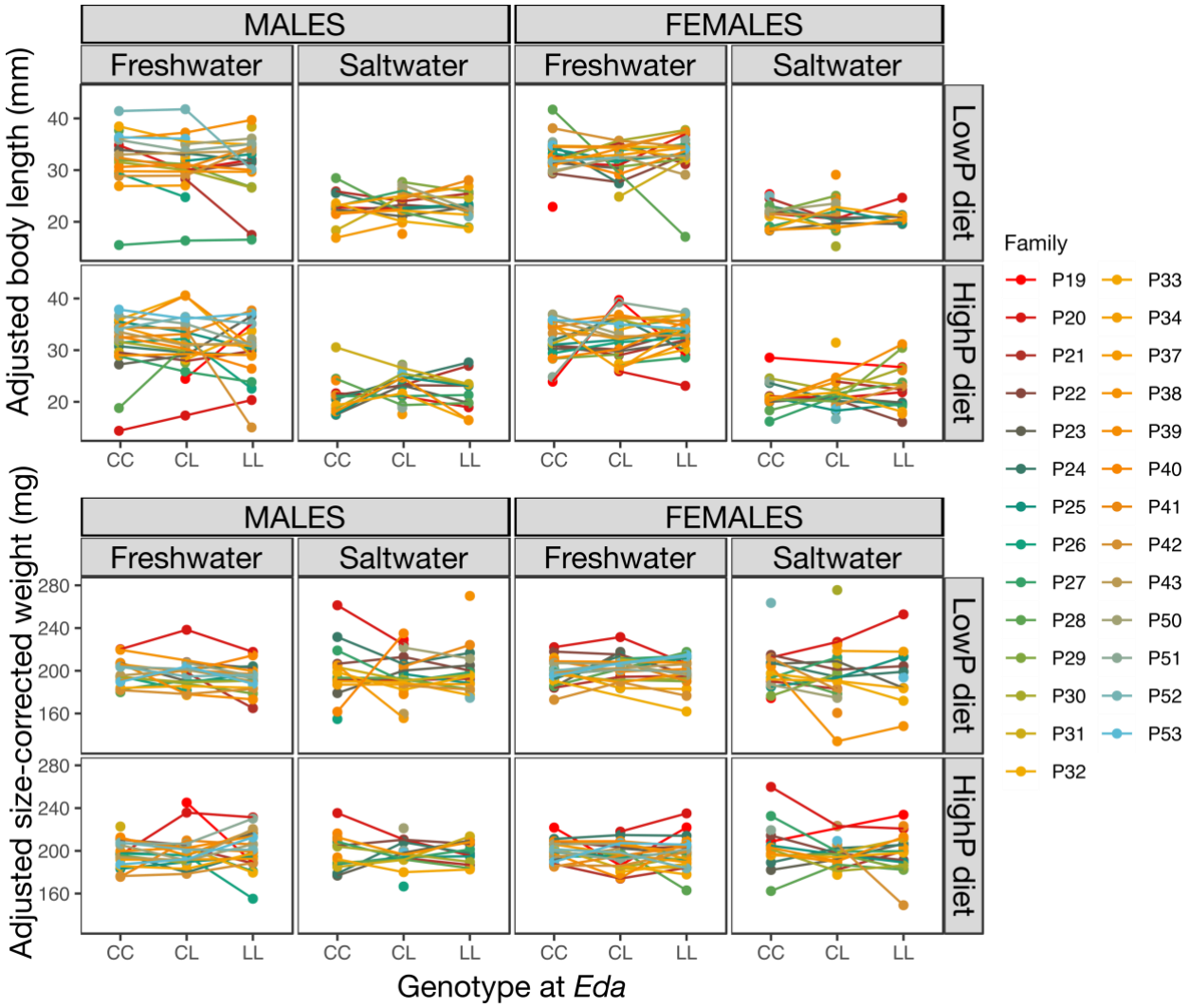


Figure 3.9. Adjusted mean length and size-corrected weight by family, sex, water, diet and genotype in Experiment 2. These results are a visual representation of the ANOVA results in Table 3.3. Fish in saltwater are smaller than freshwater fish at this timepoint but have similar size-corrected weights (body condition). Again, size and size-corrected weights do not vary between genotypes. Adjusted trait values were calculated by adding the residual trait value for each individual to the predicted trait value when the covariates are equal (standard length in the case of size-corrected weight).

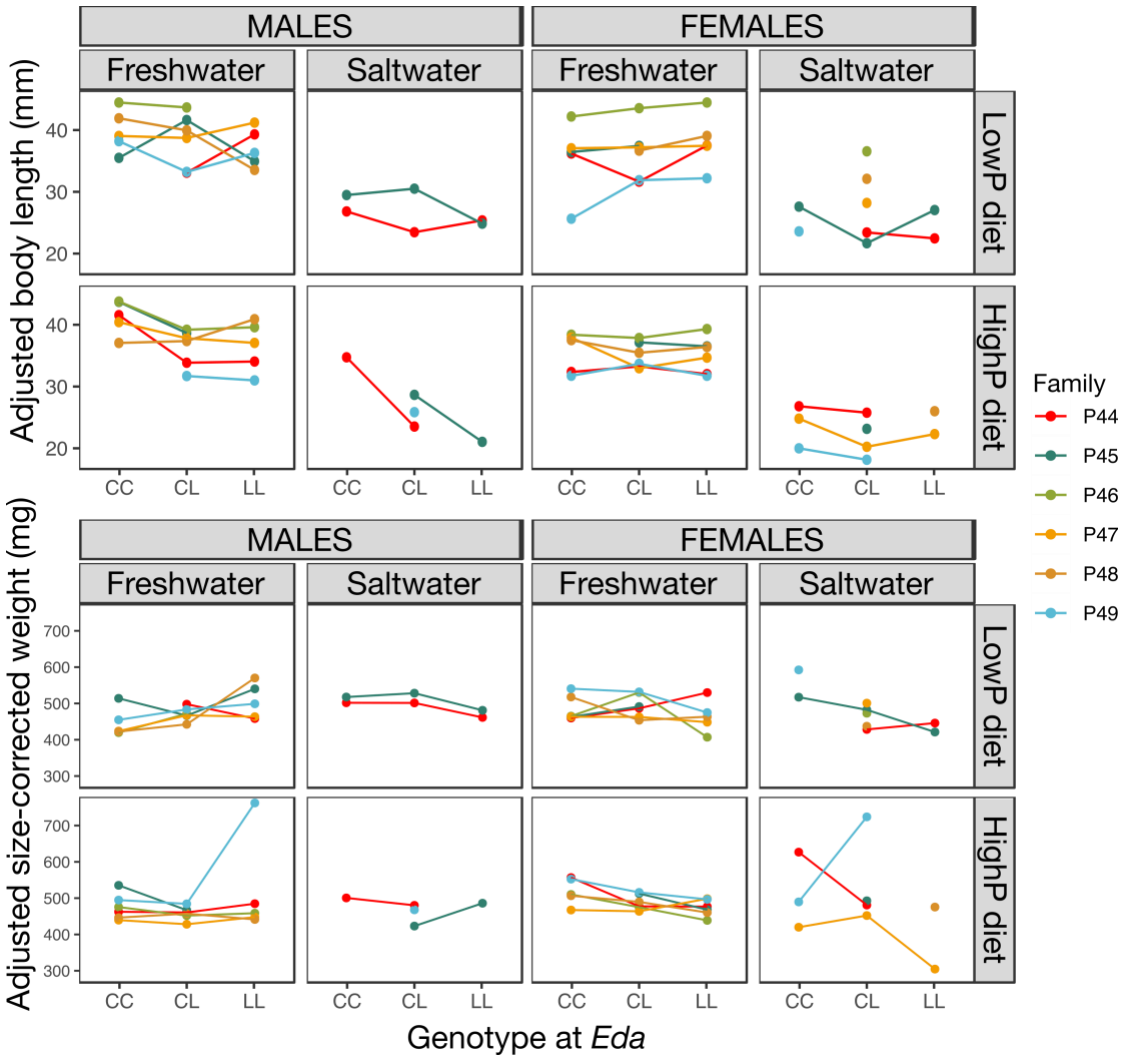


Figure 3.10. Adjusted mean length and size-corrected weight by family, sex, water, diet and genotype in Experiment 3. These results are a visual representation of the ANOVA results in Table 3.4. Fish in saltwater are smaller than freshwater fish at this timepoint but have similar size-corrected weights (body condition). Males are slightly longer than females but have similar size-corrected weights. The effect of genotype on standard length is nearly significant and is significant on size-corrected weight ($p = 0.017$), with CC fish being the largest and the highest body condition. This effect is amplified in saltwater (LL fish have even lower body conditions in saltwater), but these effects are slight, at best, and hard to see. Adjusted trait values were calculated by adding the residual trait value for each individual to the predicted trait value when the covariates are held constant (standard length in the case of size-corrected weight).

Figure 3.11. Adjusted mean phosphorus content of bone is plotted by family, sex, water, diet and genotype in Experiment 3. These results are a visual representation of the ANOVA results in Table 3.5. Phosphorus content of bone is strongly affected by water (lower % phosphorus in saltwater) but does not differ between fish with different genotypes. Adjusted trait values were calculated by adding the residual trait value for each individual to the predicted trait value when the covariate, standard length, is held constant.

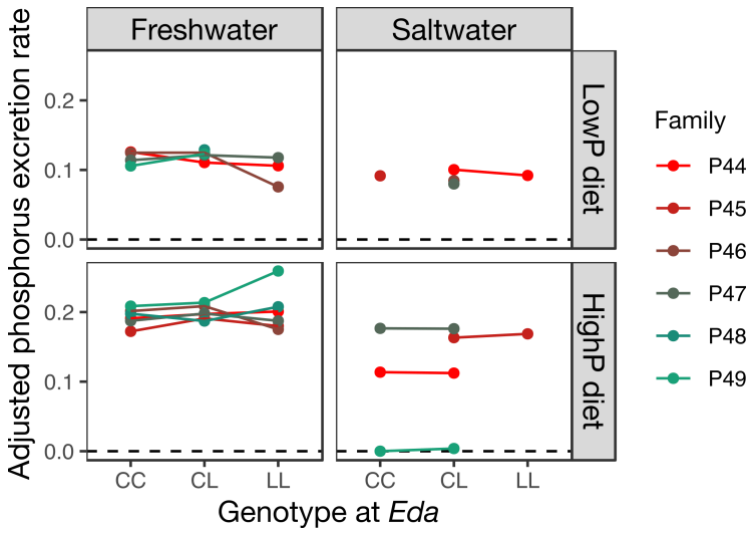


Figure 3.12. Adjusted mean phosphorus excretion rate by family, water, diet and genotype in Experiment 3. These results are a visual representation of the ANOVA results in Table 3.6. Fish in freshwater had a slightly higher excretion rate than fish in saltwater (μg P excreted / minute). In addition, fish fed the HighP diet had higher phosphorus excretion rates than fish in the LowP diet. Importantly, there was no difference in excretion rates between fish with different genotypes at *Eda*. Adjusted trait values were calculated by adding the residual trait value for each individual to the predicted trait value when the covariate, standard length, is held constant.

Table 3.1. Diet composition for LowP and HighP diets. Composition of diets by each ingredient is shown as a percentage. Diets were *iso*-nitrogenous (44% protein) and *iso*-lipidic (14.5% lipid). The LowP diet had a final % phosphorus of 0.48% while the HighP diet was 0.61%.

INGREDIENT	LOWP DIET	HIGHP DIET
Potato starch	24.00	24.03
Soybean meal	14.94	14.94
Pea protein	14	14
Hydrolyzed fish meal	12	12
Fish oil	11.94	11.44
Wheat gluten	9	9
Brewer yeast	6.7	6.7
LT94 Fish meal	5.38	5.38
Vitamin & mineral Premix	2	2
Ammonium phosphate	-	0.51

Table 3.2. ANOVA table for standard length and body condition in Experiment 1. Results from the type III ANOVA using Satterthwaite's method of the mixed effects linear models on standard length and body condition are shown for the 1342 fish from Experiment 1. Standard length was used as a measure of growth, while body condition was assessed as natural log of weight as the response variable with natural log of standard length as a predictor variable. For a visual representation of these data, see Figure 3.8.

RESPONSE	PREDICTOR	N.D.F.	D.D.F.	F-VALUE	P-VALUE	
Standard length	Water	1	56	22.49	1.5E-05	***
	Diet	1	55	0.10	0.75	
	Sex	1	1288	20.06	8.2E-06	***
	Genotype	2	1291	0.24	0.78	
	Water x Genotype	2	1291	2.06	0.13	
	Water x Diet	1	52	0.01	0.92	
	Diet x Genotype	2	1290	2.55	0.08	
Body condition	ln(Standard length)	1	1160	10414.50	<2e-16	***
	Water	1	52	194.40	<2e-16	***
	Diet	1	45	1.84	0.18	
	Sex	1	1263	0.68	0.41	
	Genotype	2	1262	0.77	0.46	
	Water x Genotype	2	1263	0.78	0.46	
	Water x Diet	1	39	0.00	0.95	
	Diet x Genotype	2	1261	0.98	0.38	

Table 3.3. ANOVA table for standard length and body condition in Experiment 2. Results from the type III ANOVA using Satterthwaite's method of the mixed effects linear models on standard length and body condition are shown for the 1209 fish from Experiment 2. Standard length was used as a measure of growth, while body condition was assessed as natural log of weight as the response variable with natural log of standard length as a predictor variable. For a visual representation of these data, see Figure 3.9.

RESPONSE	PREDICTOR	N.D.F.	D.D.F.	F-VALUE	P-VALUE	
Standard length	Water	1	66	577.70	<2e-16	***
	Diet	1	52	0.23	0.63	
	Sex	1	1152	0.63	0.43	
	Genotype	2	1150	0.57	0.57	
	Water x Genotype	2	1149	0.39	0.67	
	Water x Diet	1	47	0.03	0.85	
	Diet x Genotype	2	1151	1.10	0.33	
Body condition	ln(Standard length)	1	1142	50637.90	<2e-16	***
	Water	1	62	1.00	0.32	
	Diet	1	44	0.79	0.38	
	Sex	1	1109	0.59	0.44	
	Genotype	2	1106	2.20	0.11	
	Water x Genotype	2	1107	0.75	0.47	
	Water x Diet	1	43	0.22	0.64	
	Diet x Genotype	2	1104	1.13	0.32	

Table 3.4. ANOVA table for standard length and body condition in Experiment 3. Results from the type III ANOVA using Satterthwaite's method of the mixed effects linear models on standard length and body condition are shown for the 277 fish from Experiment 3. Standard length was used as a measure of growth, while body condition was assessed as natural log of weight as the response variable with natural log of standard length as a predictor variable. For a visual representation of these data, see Figure 3.10.

RESPONSE	PREDICTOR	N.D.F.	D.D.F.	F-VALUE	P-VALUE	
Standard length	Water	1	17	135.51	2.19E-09	***
	Diet	1	15	1.83	0.20	
	Sex	1	259	14.00	0.00	***
	Genotype	2	259	3.02	0.05	.
	Age	1	4	6.31	0.06	
	Water x Genotype	2	259	1.61	0.20	
	Water x Diet	1	13	0.19	0.67	
	Diet x Genotype	2	258	1.88	0.15	
Body condition	ln(Standard length)	1	258	5291.11	<2e-16	***
	Water	1	18	0.62	0.44	
	Diet	1	10	0.45	0.52	
	Sex	1	251	3.63	0.06	
	Genotype	2	259	4.14	0.02	*
	Age	1	4	0.62	0.47	
	Water x Genotype	2	260	4.16	0.02	*
	Water x Diet	1	9	0.76	0.41	
	Diet x Genotype	2	250	1.25	0.29	

Table 3.5. ANOVA table for intestine length and bone phosphorus content in Experiment 3. Results from the type III ANOVA using Satterthwaite's method of the mixed effects linear models for intestine length and bone % phosphorus content are shown for the 277 fish from Experiment 3. Both intestine length and standard length were log transformed prior to analysis. For a visual representation of these data, see Figure 3.11.

RESPONSE	PREDICTOR	N.D.F.	D.D.F.	F-VALUE	P-VALUE	
Intestine length	ln(Standard length)	1	260	916.15	<2e-16	***
	Water	1	70	0.17	0.68	
	Diet	1	28	0.88	0.36	
	Sex	1	255	1.65	0.20	
	Genotype	2	257	1.31	0.27	
	Age	1	4	3.02	0.16	
	Water x Genotype	2	256	2.64	0.07	
	Water x Diet	1	24	0.90	0.35	
	Diet x Genotype	2	258	2.84	0.06	
Bone %P	Water	1	17	82.60	5.878E-08	***
	Diet	1	15	0.00	0.96	
	Sex	1	257	3.33	0.07	
	Genotype	2	262	1.36	0.26	
	Age	1	4	10.00	0.03	*
	Water x Genotype	2	263	1.62	0.20	
	Water x Diet	1	14	0.12	0.73	
	Diet x Genotype	2	258	0.03	0.97	

Table 3.6. ANOVA table for phosphorus excretion rates in Experiment 3. Results from the type III ANOVA using Satterthwaite's method of the mixed effects linear models for phosphorus excretion rate are shown for the 201 fish from Experiment 3 that were used in this analysis. For a visual representation of these data, see Figure 3.12.

RESPONSE	PREDICTOR	N.D.F.	D.D.F.	F-VALUE	P-VALUE	
P excretion rate	Water	1	189	8.58	0.00	**
	Diet	1	74	5.71	0.02	*
	Weight	1	188	0.06	0.81	
	Sex	1	179	1.68	0.20	
	Genotype	2	187	0.63	0.54	
	Water x Genotype	2	186	0.97	0.38	
	Water x Diet	1	185	3.57	0.06	
	Diet x Genotype	2	183	0.41	0.66	

Chapter 4. DISCUSSION

Stickleback's unique biology provides the rare opportunity to connect genotypes to phenotypes to fitness in natural, vertebrate populations. Plate loss in freshwater stickleback and the underlying gene, *Eda*, have been at the center of this work in stickleback nearly since its inception (Hagen, 1973; Avise, 1976; Colosimo *et al.*, 2004; Colosimo *et al.*, 2005). Because of these early advances and the large phenotypic effect that *Eda* has on plate loss, *Eda* has remained at the center of much stickleback literature (Barrett *et al.*, 2008; Barrett *et al.*, 2009b; Marchinko, 2009; Zeller *et al.*, 2012a, 2012b; Marchinko *et al.*, 2014; O'Brown *et al.*, 2015; Pujolar *et al.*, 2017; Robertson *et al.*, 2017; Paccard *et al.*, 2018; Yamasaki *et al.*, 2019). I hope that by the end of this Discussion I will have convinced you that my work on *Eda* has both addressed long standing and important questions about the genetics of adaptation but has also highlighted additional questions that are worth exploring. Thus, as our understanding of adaptation continues to advance, I argue for the usefulness of examples centered around studying known, adaptive loci like *Eda*.

When I started on this PhD work, I hypothesized that: 1) the recently identified NAKA SNP controlled plate number; 2) the *Eda* haplotype controlled adaptive variation in many traits, including possibly spine length, gill rakers and body shape; 3) plate loss and neuromast patterning are caused by a single pleiotropic mutation; and 4) dietary phosphorus limitation was the selective agent that drove plate loss in freshwater environments. I was excited to test these hypotheses, mostly because I naively thought they were true. While most of my hypotheses are not supported by my data, my results are more informative and generalizable for the genetics of adaptation than I had hoped.

Just after I joined the lab and started on these projects, a paper came out linking a single SNP downstream of *Eda* (the so-called NAKA SNP because of its recurrence in the NAKA population) to reduced enhancer activity on the flank of fish where plates form, suggesting it was the causative mutation (O'Brown *et al.*, 2015). I was therefore thrilled that I caught so many marine fish that were heterozygous at the NAKA SNP (Figure 2.2), allowing me to test the functional implication of the SNP in a variety of backgrounds. I was quite surprised when these studies failed to detect any phenotypic differences driven by this single SNP (Figure 2.8). In addition to overturning one of the more important previous results in the search for specific phenotype-altering mutations, these results underscore that functional data for single mutations, though difficult to obtain, are critically important to make conclusions about the genetic basis of traits.

Instead of documenting the effects of a causative mutation, I was instead left with the question of what this common genetic variant (the NAKA SNP) does in freshwater. However, recent findings turn this question around: the so-called marine version of the NAKA SNP appears to be derived when compared to sequences from other percomorphs and Gasterosteids (Figure 2.12 and Yamasaki *et al.* (2019)). I am now left wondering if the derived marine allele at the NAKA SNP has phenotypic effects we have not yet measured (e.g. plate height or spine length), or if it is under selection in the marine environment. I found that the NAKA SNP was in Hardy Weinberg equilibrium in the Puget Sound population in 2015 (χ^2=0.37, d.f.=3, p=0.94; Figure 2.2), suggesting that selection is not acting on the NAKA SNP in the marine environment, despite the evidence that it modulates an enhancer (O'Brown *et al.*, 2015). I am left with the hypothesis that the NAKA SNP is an example of near neutral variation that exists widely across the globe. Future work on this SNP could further explore phenotypic effects of this SNP as well as the evolutionary history and spread of this allele.

The next results I obtained were the mapping results in the Lake Washington wild-caught stickleback population. I had hypothesized that variation within the adaptive haplotype containing *Eda* would have phenotypic effects on many traits. I based this hypothesis on previous work showing that tissues with marine-freshwater phenotypic differences (e.g. spines, fins, jaw) express *Eda* or are affected by *Eda* knockouts (Harris *et al.*, 2008; O'Brown *et al.*, 2015; Wucherpfennig *et al.*, 2019) and QTL studies that mapped the genetic bases of these traits to genomic regions containing *Eda*. However, this mapping dataset was noisy because these fish were collected in different months and processed using different methods, and I thought that the noise may conceal the true genetic signal. Therefore, I re-analyzed the data every way I could to try to mitigate the effects of this sampling noise, but I could not detect an effect of *Eda* on additional phenotypes in the data. I held out hope that the Puget Sound F1 lab crosses would detect associations with *Eda* that had been obscured by environmental noise in the Lake Washington dataset. Importantly, the lab crosses produced qualitatively the same results: the freshwater *Eda* haplotype has reproducible and large effects on plate count, neuromast patterning and neuromast number, and small, population-specific effects on gill raker length and three body shape landmarks. This was not what I was expecting, but I think this makes the *Eda* haplotype a good example of adaptive evolution instead of the magic locus I had hypothesized it to be, and highlights the poor mapping resolution obtained in QTL mapping studies.

One way in which these results reinforce *Eda* as an interesting example of an adaptive locus is that it is surrounded by QTL for additional traits (the traits whose QTL overlap with *Eda* but that did not map to the haplotype). This clustering of locally adapted traits in populations experiencing gene flow is predicted to select for reduced recombination (Yeaman and Whitlock, 2011). In several other systems, and in other regions of the stickleback genome, selection for

reduced recombination may have resulted in the fixation of inversions (Mather, 1950; Lowry and Willis, 2010; Jones *et al.*, 2012b; Lamichhaney *et al.*, 2016a; Tuttle *et al.*, 2016; Lee *et al.*, 2017). Inversions, however, pose a problem for adaptive trait mapping: high linkage disequilibrium makes disentangling the genetic basis of multiple traits extremely difficult, and cannot inform us about an adaptive walk towards a new fitness optimum through fixation of multiple alleles. No evidence of an inversion has been detected between marine and freshwater versions of chromosome IV, therefore further study of this stickleback chromosome during early stages of adaptation to freshwater would shed light on the process of adaptation.

There are a couple of studies that are currently underway that will shed light on this process. First, there is a long-term study tracking allele frequency shifts genome-wide as marine stickleback adapt to new freshwater habitats. I am interested in the results because it will shed light on which genomic and biological factors best predict changes in allele frequencies. *Eda* is a good candidate to expect to fix quickly because it contains a large-effect mutation, selection on the locus can be very strong and it has repeatedly fixed in freshwater in the past (Barrett *et al.*, 2008; Kitano *et al.*, 2008; Jones *et al.*, 2012b). On the other hand, the freshwater allele of *Eda* may be in linkage disequilibrium with the marine allele at other nearby loci, such as the spine QTL at *Msx2a* (Howes *et al.*, 2017), and this could retard the fixation of both freshwater alleles. Therefore, measuring how well factors such as recombination rate, abundance of standing variation, gene or QTL density, etc. can predict shifts in allele frequencies will be informative of the process of adaptation.

The second ongoing study will investigate the abundance and distribution of standing variation in the marine population. Previous work found that many regions of the genome are shared across freshwater populations worldwide, suggesting that these freshwater alleles exist as standing variation in the marine populations (Jones *et al.*, 2012b). The next question is: do these

adaptive freshwater alleles exist in the same fish in the marine environment, or do they reassemble in each new freshwater population? Our Puget Sound data suggest that the plate QTL (the *Eda* freshwater haplotype) segregates separately from the other trait QTL in marine fish, suggesting that these adaptive alleles must recombine onto the same genomic background in each new population of freshwater stickleback. However, there could be some standing variation in the marine population that is often or always found in the same individual fish (possibly hybrid individuals lost at sea). Together these projects will tell us about the extent of standing variation in marine populations and the process of reassembling these adaptive alleles during adaptation. These studies may also inform us about the strength, speed and type of selection that acts on these standing allelic variants in both the marine environment and freshwater environments. These studies would not be as informative in systems where most of the phenotypic variation exists in inversions or if *Eda* controlled variation in many traits, as I had naively hypothesized.

While we found that the *Eda* haplotype consistently controls variation in only three traits, we also found that these effects are the result of multiple mutations within the haplotype. For example, the effect that *Eda* has on plate number is due to at least two separate mutations within the haplotype (Figure 2.3). This is similar to findings in other systems where large effect QTL are also composed of several to many smaller effect mutations that act together to produce the composite phenotype (Gompel *et al.*, 2005; Rebeiz *et al.*, 2009; Linnen *et al.*, 2013). The evolutionary histories of these composite QTL are intriguing. In his reanalysis of the geometric model of adaptation (Fisher, 1930), Orr (1998) predicted that large effect mutations are more likely to occur early in the adaptive walk, while smaller effect mutations would likely occur later. This leads to a clear prediction of the two separate mutations we mapped: the smaller effect mutation near SNPs12/13 would have arisen later than the larger effect mutation near LP3621. Once we have

identified the causative mutations for plate loss, we should be able to address this hypothesis using refugial freshwater populations of stickleback that may have freshwater haplotypes that diverged at different times.

Further, we are close to identifying the actual mutations that cause plate loss. The mapping study allowed us to narrow down the large effect causative region to 1400 bases containing about 18 fixed differences between multiple marine and freshwater populations of fish, but we were limited from further resolution by recombination events captured in the wild fish. The exciting next step is to use CRISPR to disrupt these 18 candidate sites and measure their individual and combined effects on plate count, neuromast number and patterning and schooling behavior. We have begun this effort by using CRISPR to induce deletions in a 16 bp sequence that is shared by completely-plated marine fish and deleted in low-plated freshwater fish. Very preliminary evidence from founder individuals suggests that disrupting this indel has effects on plate count.

My third hypothesis for my PhD work was that plate loss and neuromast patterning would be caused by the same, pleiotropic mutation(s). I believed this might be true because studies have shown that neuromasts and dermal bone are tightly correlated in both zebrafish and stickleback (Wada *et al.*, 2010; Mills *et al.*, 2014; Wada *et al.*, 2014). The corresponding prediction for our study was that plates and neuromast patterning would show the same pattern of association between phenotypes and markers, and this has held. However, until we identify the causative mutation(s) for both traits, we cannot exclude that the different traits are controlled by multiple tightly linked sites. Again, obtaining CRISPR-modified fish for the various candidate sites will allow final resolution of the twin questions of pleiotropy (are multiple traits controlled by the same site) and linkage (are multiple sites controlling the same trait).

If in fact it does prove true that a single pleiotropic site controls multiple traits, then I think this opens up some interesting developmental and evolutionary questions. For example, *Eda* has not been found to play a role in neuromast development, so how is it that this mutation is driving this phenotypic difference in stickleback? Perhaps *Eda* does not play a direct role, but instead signals to the plate forming cells, which in turn signal to the neuromasts. On the other hand, *Eda* may play direct and independent roles in both neuromast patterning and plate formation, but its role in neuromast patterning has previously gone undetected. If the former is true, then blocking plate formation via tissue-specific knockdown of Eda's receptors (*Edar*, *Xedar* or *Troy*) would have cascading effects on neuromast patterning. If instead *Eda* has a direct role in both tissues independently, disruption of plates would have no effect on neuromast patterning (and vice versa). Either mechanism has implications for the evolvability of these two traits. If both the plate and the neuromast precursors are responding to the same signal (*Eda*), then there may be more constraint on their coevolution. But if the plate precursor cells, for example, signal to the neuromast precursor cells, then I would hypothesize there would be molecular mechanisms able to modify this second signal.

All of this work still leaves us with the question of what agents of selection act on which combination of phenotypes- plate reduction, neuromast repatterning, and/or neuromast number- in freshwater environments. I hypothesized that dietary phosphorus limitation was the selective agent that drove plate loss in freshwater environments, and consequently the repatterning of neuromasts. Before this work, neuromast number had not mapped to chromosome IV, likely due to other genomic loci that may have masked the effects of this locus (Wark *et al.*, 2012). Freshwater stickleback populations show variation in neuromast numbers, but not in a consistent pattern to suggest that selection favors the reduction in neuromast number caused by the *Eda* locus (Wark

and Peichel, 2010). Therefore, I do not think that the reduction in neuromast number caused by *Eda* is the phenotype that selection is acting on and driving fixation of the *Eda* haplotype.

Neuromast patterning and plate reduction continue to be inseparable, meaning we cannot rule out selection acting on either trait. Instead, we can experimentally manipulate possible agents of selection acting on one trait or the other and measure the fitness response. This is the approach others have taken to identify agents of selection acting on the *Eda* allele, though none have tested possible selective agents acting on neuromast patterning (Marchinko *et al.*, 2007; Barrett *et al.*, 2008, 2009a; Marchinko, 2009; Zeller *et al.*, 2012a). These studies have found that fish with marine *Eda* alleles show a growth deficit in freshwater compared with fish with freshwater *Eda* alleles (both in the lab or in semi-natural ponds) (Marchinko *et al.*, 2007; Barrett *et al.*, 2008, 2009a). However, freshwater and saltwater differ in a number of ways. We tried to further disentangle the effects of freshwater into diet and salinity but were unable to reproduce the growth advantage of freshwater *Eda* alleles (Chapter 3). This leads me to conclude that another selective agent acts on the *Eda* locus. There are a few interesting possibilities.

First, as discussed in Chapter 3, I think the next logical agent of selection to test is calcium. In addition to testing the role of calcium on growth and survival, I think it would be worthwhile to test the selective pressure of parasites on the *Eda* haplotype. Within the 16 kb haplotype, there are three protein coding genes, one of which (*Tnfsf13b*, Figure 2.1) is associated with *Ascaris* infection in human populations (Williams-Blangero *et al.*, 2002; Williams-Blangero *et al.*, 2008). We attempted this analysis by measuring infection rates of two parasites in the polymorphic Lake Washington stickleback population. However, infection rates of fish by *Eustrongylides* spp. (6%) and *Schistocephalus solidus* (11%) were low and varied by season. Our uneven sampling coverage across seasons and genotypes combined with low infection rates may explain why we did not

detect a significant association between the presence/absence of these parasites and genotype at *Eda*. However, the lack of association between the presence of these two parasites and genotype at *Eda* could be driven by additional loci obscuring the effects of *Tnfsf13b* in Lake Washington or by no role of natural variation in *Tnfsf13b* in the immune response to these parasites. A different way to address this question would be to measure gene expression levels of *Tnfsf13b* in relevant tissues (e.g. B cells) in the offspring of our Puget Sound crosses both with and without immune challenges. If natural variation in *Tnfsf13b* plays a role in differential immune responses between marine and freshwater fish, then we would expect expression level differences of *Tnfsf13b* associated with infection status and/or genotype.

Documenting the role of selection on natural genotypic variation is a hurdle to understanding the process of adaptation in wild populations. In many systems selection may fluctuate in time or space, or be too complex to easily disentangle (Siepielski *et al.*, 2009; Kingsolver and Diamond, 2011; Nosil *et al.*, 2018). Selection on *Eda* in freshwater also fluctuates over developmental stages (Barrett *et al.*, 2008), and yet is strongly directional over multiple generations (Kitano *et al.*, 2008). I believe these two observations reinforce the importance of studying the *Eda* haplotype as both a representative and tractable locus of adaptive evolution.

Despite much attention and progress on understanding the genetics of adaptation and the *Eda* haplotype in threespine stickleback, I am continually surprised by the findings and implications of this locus. For example, what maintains the polymorphism at the NAKA SNP in marine populations, how do mutations in multiple enhancers disrupt the molecular function of *Eda* to produce multiple adaptive phenotypes, and what are the agents of selection that act on these phenotypes in fresh- and saltwater? The tractability of stickleback as a model organism in the lab combined with advances in genetic engineering and the growing interest in predicting the action

of selection on alleles in response to predicted environmental changes will keep stickleback in the foreground of research on the genetics of adaptation.

REFERENCES

Acharya, K., Kyle, M., & Elser, J.J. (2004). Biological stoichiometry of *Daphnia* growth: an ecophysiological test of the growth rate hypothesis. *Limnol. Oceanogr.* 49:656-665.

Ågren, J., Oakley, C.G., McKay, J.K., Lovell, J.T., & Schemske, D.W. (2013). Genetic mapping of adaptation reveals fitness tradeoffs in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 110:21077-21082.

Alberch, P. (1991). From genes to phenotypes: dynamical systems and evolvability. *Genetica* 84:5-11.

Albert, A.Y.K., Sawaya, S., Vines, T.H., Knecht, A.K., Miller, C.T., Summers, B.R., *et al.* (2008). The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. *Evolution* 62:76-85.

Albertson, R.C., Streelman, J.T., & Kocher, T.D. (2003). Directional selection has shaped the oral jaws of Lake Malawi cichlid fishes. *Proc. Natl. Acad. Sci. USA* 100:5252-5257.

Aman, A.J., Fulbright, A.N., & Parichy, D.M. (2018). Wnt/β-catenin regulates an ancient signaling network during zebrafish scale development. *eLife* 7:e37001.

Andersen, T., & Hessen, D.O. (1991). Carbon, nitrogen, and phosphorus content of freshwater zooplankton. *Limnol. Oceanogr.* 36:807-814.

Anderson, J.T., Lee, C.-R., Rushworth, C.A., Colautti, R.I., & Mitchell-Olds, T. (2013). Genetic trade-offs and conditional neutrality contribute to local adaptation. *Mol. Ecol.* 22:699-708.

Avila, E.M., Tu, H., Basantes, S., & Ferraris, R.P. (2000). Dietary phosphorus regulates intestinal transport and plasma concentrations of phosphate in rainbow trout. *J. Comp. Physiol. B* 170:201-209.

Avise, J.C. (1976). Genetics of plate morphology in an unusual population of threespine sticklebacks (*Gasterosteus aculeatus*). *Genet. Res.* 27:33-46.

Baker, J.A., Wund, M.A., Heins, D.C., King, R.W., Reyes, M.L., & Foster, S.A. (2015). Life-history plasticity in female threespine stickleback. *Heredity* 115:322-334.

Barrett, R.D.H., & Hoekstra, H.E. (2011). Molecular spandrels: tests of adaptation at the genetic level. *Nat. Rev. Genet.* 12:767-780.

Barrett, R.D.H., Laurent, S., Mallarino, R., Pfeifer, S.P., Xu, C.C.Y., Foll, M., *et al.* (2019). Linking a mutation to survival in wild mice. *Science* 363:499-504.

Barrett, R.D.H., Rogers, S.M., & Schluter, D. (2008). Natural selection on a major armor gene in threespine stickleback. *Science* 322:255-257.

Barrett, R.D.H., Rogers, S.M., & Schluter, D. (2009a). Environment specific pleiotropy facilitates divergence at the *Ectodysplasin* locus in threespine stickleback. *Evolution* 63:2831-2837.

Barrett, R.D.H., & Schluter, D. (2008). Adaptation from standing genetic variation. *Trends Ecol. Evol.* 23:38-44.

Barrett, R.D.H., Vines, T.H., Bystriansky, J.S., & Schulte, P.M. (2009b). Should I stay or should I go? The Ectodysplasin locus is associated with behavioural differences in threespine stickleback. *Biol. Lett.* 5:788-791.

Barton, N.H., & Keightley, P.D. (2002). Understanding quantitative genetic variation. *Nat. Rev. Genet.* 3:11-21.

Bassham, S., Catchen, J., Lescak, E., von Hippel, F.A., & Cresko, W.A. (2018). Repeated selection of alternatively adapted haplotypes creates sweeping genomic remodeling in stickleback. *Genetics* 209:921-939.

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

Bell, M.A., & Foster, S.A. (1994). *The evolutionary biology of the threespine stickleback*. Oxford University Press, New York.

Bell, M.A., Orti, G., Walker, J.A., & Koenings, J.P. (1993). Evolution of pelvic reduction in threespine stickleback fish: a test of competing hypotheses. *Evolution* 47:906-914.

Bergstrom, C.A. (2002). Fast-start swimming performance and reduction in lateral plate number in threespine stickleback. *Can. J. Zool.* 80:207-213.

Boag, P.T., & Grant, P.R. (1981). Intense natural selection in a population of Darwin's finches (Geospizinae) in the Galápagos. *Science* 214:82-85.

Boros, G., Sály, P., & Vanni, M.J. (2015). Ontogenetic variation in the body stoichiometry of two fish species. *Oecologia* 179:329-341.

Bourgeois, J.F., Blouw, D.M., Koenings, J.P., & Bell, M.A. (1994). Multivariate analysis of geographic covariance between phenotypes and environments in the threespine stickleback, *Gasterosteus aculeatus*, from the Cook Inlet area, Alaska. *Can. J. Zool.* 72:1497-1509.

Boyle, E.A., Li, Y.I., & Pritchard, J.K. (2017). An expanded view of complex traits: from polygenic to omnigenic. *Cell* 169:1177-1186.

Bradshaw, H.D., Jr., Otto, K.G., Frewen, B.E., McKay, J.K., & Schemske, D.W. (1998). Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (*Mimulus*). *Genetics* 149:367-382.

Bratteler, M., Lexer, C., & Widmer, A. (2006). Genetic architecture of traits associated with serpentine adaptation of *Silene vulgaris*. *J. Evol. Biol.* 19:1149-1156.

Broman, K.W., & Sen, S. (2009). *A Guide to QTL Mapping with R/qtl*. Springer-Verlag, New York.

Bumpus, H.C. (1899). The elimination of the unfit as illustrated by the introduced sparrow, *Passer domesticus:* (a fourth contribution to the study of variation). *Biol. Lect., Woods Hole Mar. Biol. Sta.* 6:209-226.

Butelli, E., Licciardello, C., Ramadugu, C., Durand-Hulak, M., Celant, A., Reforgiato Recupero, G., *et al.* (2019). *Noemi* controls production of flavonoid pigments and fruit acidity and illustrates the domestication routes of modern citrus varieties. *Curr. Biol.* 29:158-164.

Carbone, M.A., Jordan, K.W., Lyman, R.F., Harbison, S.T., Leips, J., Morgan, T.J., *et al.* (2006). Phenotypic variation and natural selection at *Catsup*, a pleiotropic quantitative trait gene in *Drosophila*. *Curr. Biol.* 16:912-919.

Carlson, S.M., Kottas, A., & Mangel, M. (2010). Bayesian analysis of size-dependent overwinter mortality from size-frequency distributions. *Ecology* 91:1016-1024.

Carroll, S.B. (2005). Evolution at two levels: on genes and form. *PLoS Biol.* 3:1159-1166.

Carroll, S.B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134:25-36.

Cava, J.A., Perlut, N.G., & Travis, S.E. (2019). Heritability and evolvability of morphological traits of Savannah sparrows (*Passerculus sandwichensis*) breeding in agricultural grasslands. *PLoS ONE* 14:e0210472.

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Jr., Shapiro, M.D., Brady, S.D., *et al.* (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* 327:302-305.

Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nat. Rev. Genet.* 10:195-205.

Charlesworth, D., & Charlesworth, B. (1979). Selection on recombination in clines. *Genetics* 91:581-589.

Chesmore, K., Bartlett, J., & Williams, S.M. (2018). The ubiquity of pleiotropy in human disease. *Hum. Genet.* 137:39-44.

Cleves, P.A., Hart, J.C., Agoglia, R.M., Jimenez, M.T., Erickson, P.A., Gai, L., *et al.* (2018). An intronic enhancer of *Bmp6* underlies evolved tooth gain in sticklebacks. *PLoS Genet.* 14:e1007449.

Cocker, J.M., Wright, J., Li, J., Swarbreck, D., Dyer, S., Caccamo, M., *et al.* (2018). *Primula vulgaris* (primrose) genome assembly, annotation and gene expression, with comparative genomics on the heterostyly supergene. *Sci. Rep.* 8:17942.

Colosimo, P.F., Hosemann, K.E., Balabhadra, S., Villarreal, G., Jr., Dickson, M., Grimwood, J., *et al.* (2005). Widespread parallel evolution in sticklebacks by repeated fixation of *Ectodysplasin* alleles. *Science* 307:1928-1933.

Colosimo, P.F., Peichel, C.L., Nereng, K., Blackman, B.K., Shapiro, M.D., Schluter, D., *et al.* (2004). The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biol.* 2:635-641.

Cresko, W.A., McGuigan, K.L., Phillips, P.C., & Postlethwait, J.H. (2007). Studies of threespine stickleback developmental evolution: progress and promise. *Genetica* 129:105-126.

Darwin, C. (1859). *On the origin of species by means of natural selection*. John Murray, London.

Davidson, E.H., & Erwin, D.H. (2006). Gene regulatory networks and the evolution of animal body plans. *Science* 311:796-800.

Day, T., Pritchard, J., & Schluter, D. (1994). Ecology and genetics of phenotypic plasticity: a comparison of two sticklebacks. *Evolution* 48:1723-1734.

Deschamps, M.-H., Poirier Stewart, N., Demanche, A., & Vandenberg, G.W. (2016). Preliminary study for description of bone tissue responsiveness to prolonged dietary phosphorus deficiency in rainbow trout, Oncorhynchus mykiss (Walbaum). *Aquac. Res.* 47:900-911.

Deschamps, M.H., Poirier Stewart, N., Demanche, A., & Vandenberg, G.W. (2014). Preliminary study for phenotypic description of vertebral abnormalities in triploid trout subjected to prolonged deficiency in phosphorus. *J. Appl. Ichthyol.* 30:833-839.

Dittmar, E.L., Oakley, C.G., Conner, J.K., Gould, B.A., & Schemske, D.W. (2016). Factors influencing the effect size distribution of adaptive substitutions. *Proc. R. Soc. Lond. B* 283:20153065.

Dong, Y., Liu, J., Li, P.-W., Li, C.-Q., Lü, T.-F., Yang, X., *et al.* (2018). Evolution of Darwin's *peloric* Gloxinia (*Sinningia speciosa*) is caused by a null mutation in a pleiotropic TCP gene. *Mol. Biol. Evol.* 35:1901-1915.

Durston, D.J., & El-Sabaawi, R.W. (2017). Bony traits and genetics drive intraspecific variation in vertebrate elemental composition. *Funct. Ecol.* 31:2128-2137.

Durston, D.J., & El-Sabaawi, R.W. (2019). The utility of stoichiometric and metabolic theory for understanding the foraging habitat and excretion of threespine stickleback (*Gasterosteus aculeatus*). *Evol. Ecol. Res.* 20:193-211.

El-Sabaawi, R.W., Warbanski, M.L., Rudman, S.M., Hovel, R., & Matthews, B. (2016). Investment in boney defensive traits alters organismal stoichiometry and excretion in fish. *Oecologia* 181:1209-1220.

Ellis, N.A., & Miller, C.T. (2016). Dissection and flat-mounting of the threespine stickleback branchial skeleton. *J. Vis. Exp.*, doi:10.3791/54056.

Elser, J.J., Sterner, R.W., Gorokhova, E., Fagan, W.F., Markow, T.A., Cotner, J.B., *et al.* (2000). Biological stoichiometry from genes to ecosystems. *Ecol. Lett.* 3:540-550.

Endler, J.A. (1986). *Natural selection in the wild*. Princeton University Press, Princeton, NJ.

Erickson, P.A., Baek, J., Hart, J.C., Cleves, P.A., & Miller, C.T. (2018). Genetic dissection of a supergene implicates *Tfap2a* in craniofacial evolution of threespine sticklebacks. *Genetics* 209:591-605.

Erickson, P.A., Glazer, A.M., Cleves, P.A., Smith, A.S., & Miller, C.T. (2014). Two developmentally temporal quantitative trait loci underlie convergent evolution of increased branchial bone length in sticklebacks. *Proc. R. Soc. Lond. B* 281:20140822.

Evans-White, M.A., Stelzer, R.S., & Lamberti, G.A. (2005). Taxonomic and regional patterns in benthic macroinvertebrate elemental composition in streams. *Freshw. Biol.* 50:1786-1799.

Fisher, R.A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* 52:399-433.

Fisher, R.A. (1930). *The genetical theory of natural selection*. Clarendon Press, Oxford.

Fishman, L., Stathos, A., Beardsley, P.M., Williams, C.F., & Hill, J.P. (2013). Chromosomal rearrangements and the genetics of reproductive barriers in *Mimulus* (monkeyflowers). *Evolution* 67:2547-2560.

Fontagné, S., Silva, N., Bazin, D., Ramos, A., Aguirre, P., Surget, A., *et al.* (2009). Effects of dietary phosphorus and calcium level on growth and skeletal development in rainbow trout (Oncorhynchus mykiss) fry. *Aquaculture* 297:141-150.

Ford, E.B. (1937). Problems of heredity in the *Lepidoptera*. *Biol. Rev.* 12:461-501.

Foster, S.A., Garcia, V.B., & Town, M.Y. (1988). Cannibalism as the cause of an ontogenetic shift in habitat use by fry of the threespine stickleback. *Oecologia* 74:577-585.

Frachon, L., Libourel, C., Villoutreix, R., Carrère, S., Glorieux, C., Huard-Chauveau, C., *et al.* (2017). Intermediate degrees of synergistic pleiotropy drive adaptive evolution in ecological time. *Nat. Ecol. Evol.* 1:1551-1561.

Frankel, N., Erezyilmaz, D.F., McGregor, A.P., Wang, S., Payre, F., & Stern, D.L. (2011). Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. *Nature* 474:598-603.

Friedman, J., Twyford, A.D., Willis, J.H., & Blackman, B.K. (2015). The extent and genetic basis of phenotypic divergence in life history traits in *Mimulus guttatus*. *Mol. Ecol.* 24:111-122.

Frost, P.C., Benstead, J.P., Cross, W.F., Hillebrand, H., Larson, J.H., Xenopoulos, M.A., *et al.* (2006). Threshold elemental ratios of carbon and phosphorus in aquatic consumers. *Ecol. Lett.* 9:774-779.

Gelmond, O., Von Hippel, F.A., & Christy, M.S. (2009). Rapid ecological speciation in three-spined stickleback *Gasterosteus aculeatus* from Middleton Island, Alaska: the roles of selection and geographic isolation. *J. Fish Biol.* 75:2037-2051.

Gibbs, H.L., & Grant, P.R. (1987). Oscillating selection on Darwin's finches. *Nature* 327:511-513.

Giles, N. (1983). The possible role of environmental calcium levels during the evolution of phenotypic diversity in Outer Hebridean populations of the three-spined stickleback, *Gasterosteus aculeatus*. *J. Zool.* 199:535-544.

Gillham, N.W. (2001). Evolution by jumps: Francis Galton and William Bateson and the mechanism of evolutionary change. *Genetics* 159:1383-1392.

Gompel, N., Prud'homme, B., Wittkopp, P.J., Kassner, V.A., & Carroll, S.B. (2005). Chance caught on the wing: *cis*-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* 433:481-487.

Gompert, Z., Comeault, A.A., Farkas, T.E., Feder, J.L., Parchman, T.L., Buerkle, C.A., *et al.* (2014). Experimental evidence for ecological selection on genome variation in the wild. *Ecol. Lett.* 17:369-379.

Grant, P.R., & Grant, B.R. (2002). Unpredictable evolution in a 30-year study of Darwin's finches. *Science* 296:707-711.

Grant, P.R., & Grant, B.R. (2006). Evolution of character displacement in Darwin's finches. *Science* 313:224-226.

Gratten, J., Beraldi, D., Lowder, B.V., McRae, A.F., Visscher, P.M., Pemberton, J.M., *et al.* (2007). Compelling evidence that a single nucleotide substitution in *TYRP1* is responsible for coat-colour polymorphism in a free-living population of Soay sheep. *Proc. R. Soc. Lond. B* 274:619-626.

Gratten, J., Pilkington, J.G., Brown, E.A., Clutton-Brock, T.H., Pemberton, J.M., & Slate, J. (2012). Selection and microevolution of coat pattern are cryptic in a wild population of sheep. *Mol. Ecol.* 21:2977-2990.

Gratten, J., Wilson, A.J., McRae, A.F., Beraldi, D., Visscher, P.M., Pemberton, J.M., *et al.* (2008). A localized negative genetic correlation constrains microevolution of coat color in wild sheep. *Science* 319:318-320.

Greenwood, A.K., Mills, M.G., Wark, A.R., Archambeault, S.L., & Peichel, C.L. (2016). Evolution of schooling behavior in threespine sticklebacks is shaped by the *Eda* gene. *Genetics* 203:677-681.

Greenwood, A.K., Wark, A.R., Yoshida, K., & Peichel, C.L. (2013). Genetic and neural modularity underlie the evolution of schooling behavior in threespine sticklebacks. *Curr. Biol.* 23:1884-1888.

Hagen, D.W. (1973). Inheritance of numbers of lateral plates and gill rakers in Gasterosteus aculeatus. *Heredity* 30:303-312.

Hagen, D.W., & Gilbertson, L.G. (1972). Geographic variation and environmental selection in *Gasterosteus aculeatus* L. in the Pacific Northwest, America. *Evolution* 26:32-51.

Hall, M.C., Basten, C.J., & Willis, J.H. (2006). Pleiotropic quantitative trait loci contribute to population divergence in traits associated with life-history variation in *Mimulus guttatus*. *Genetics* 172:1829-1844.

Hall, S.R., Leibold, M.A., Lytle, D.A., & Smith, V.H. (2004). Stoichiometry and planktonic grazer composition over gradients of light, nutrients, and predation risk. *Ecology* 85:2291-2301.

Hansen, T.F., Pélabon, C., & Houle, D. (2011). Heritability is not evolvability. *Evol. Biol.* 38:258-277.

Harris, M.P., Rohner, N., Schwarz, H., Perathoner, S., Konstantinidis, P., & Nusslein-Volhard, C. (2008). Zebrafish *eda* and *edar* mutants reveal conserved and ancestral roles of ectodysplasin signaling in vertebrates. *PLoS Genet.* 4:e1000206.

Hawthorne, D.J., & Via, S. (2001). Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature* 412:904-907.

Hendrixson, H.A., Sterner, R.W., & Kay, A.D. (2007). Elemental stoichiometry of freshwater fishes in relation to phylogeny, allometry and ecology. *J. Fish Biol.* 70:121-140.

Hendry, A.P. (2013). Key questions in the genetics and genomics of eco-evolutionary dynamics. *Heredity* 111:456-466.

Hendry, A.P., Peichel, C.L., Matthews, B., Boughman, J.W., & Nosil, P. (2013). Stickleback research: the now and the next. *Evol. Ecol. Res.* 15:111-141.

Hermann, K., Klahre, U., Moser, M., Sheehan, H., Mandel, T., & Kuhlemeier, C. (2013). Tight genetic linkage of prezygotic barrier loci creates a multifunctional speciation island in *Petunia*. *Curr. Biol.* 23:873-877.

Hermisson, J., & Pennings, P.S. (2005). Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* 169:2335-2352.

Hessen, D.O., Elser, J.J., Sterner, R.W., & Urabe, J. (2013). Ecological stoichiometry: An elementary approach using basic principles. *Limnol. Oceanogr.* 58:2219-2236.

Heuts, M.J. (1947). Experimental studies on adaptive evolution in *Gasterosteus aculeatus* L. *Evolution* 1:89-102.

Hoekstra, H.E., & Coyne, J.A. (2007). The locus of evolution: evo devo and the genetics of adaptation. *Evolution* 61:995-1016.

Hoffmann, A.A., & Rieseberg, L.H. (2008). Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? *Ann. Rev. Ecol. Evol. Syst.* 39:21-42.

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

Hood, J.M., Vanni, M.J., & Flecker, A.S. (2005). Nutrient Recycling by Two Phosphorus-Rich Grazing Catfish: The Potential for Phosphorus-Limitation of Fish Growth. *Oecologia* 146:247-257.

Howes, T.R., Summers, B.R., & Kingsley, D.M. (2017). Dorsal spine evolution in threespine sticklebacks via a splicing change in *MSX2A*. *BMC Biol.* 15:115.

Ishikawa, A., Kabeya, N., Ikeya, K., Kakioka, R., Cech, J.N., Osada, N., *et al.* (2019). A key metabolic gene for recurrent freshwater colonization and radiation in fishes. *Science* 364:886-889.

Jiang, Y., Peichel, C.L., Ling, F., & Bolnick, D.I. (2016). Sensory trait variation contributes to biased dispersal of threespine stickleback in flowing water. *J. Evol. Biol.* 30:681-695.

Jones, F.C., Chan, Y.F., Schmutz, J., Grimwood, J., Brady, S.D., Southwick, A.M., *et al.* (2012a). A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Curr. Biol.* 22:83-90.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., *et al.* (2012b). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484:55-61.

Joron, M., Frezal, L., Jones, R.T., Chamberlain, N.L., Lee, S.F., Haag, C.R., *et al.* (2011). Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature* 477:203-206.

Josephs, E.B., Stinchcombe, J.R., & Wright, S.I. (2017). What can genome-wide association studies tell us about the evolutionary forces maintaining genetic variation for quantitative traits? *New Phytol.* 214:21-33.

Kamberov, Yana G., Wang, S., Tan, J., Gerbault, P., Wark, A., Tan, L., *et al.* (2013). Modeling recent human evolution in mice by expression of a selected EDAR variant. *Cell* 152:691-702.

Kettlewell, H.B.D. (1955). Selection experiments on industrial melanism in the *Lepidoptera*. *Heredity* 9:323-342.

Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge University Press, Cambridge.

King, M., & Wilson, A. (1975). Evolution at two levels in humans and chimpanzees. *Science* 188:107-116.

Kingsley, D.M., & Peichel, C.L. (2007). The molecular genetics of evolutionary change in sticklebacks. In: *Biology of the Three-Spined Stickleback*, eds. Ostlund-Nilsson, S., Mayer, I., & Huntingford, F.A. CRC Press, Boca Raton, FL, USA, pp. 41-81.

Kingsolver, J.G., & Diamond, S.E. (2011). Phenotypic selection in natural populations: what limits directional selection? *Am. Nat.* 177:346-357.

Kingsolver, J.G., Diamond, S.E., Siepielski, A.M., & Carlson, S.M. (2012). Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. *Evol. Ecol.* 26:1101-1118.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., *et al.* (2001). The strength of phenotypic selection in natural populations. *Am. Nat.* 157:245-261.

Kirkpatrick, M. (2010). How and why chromosome inversions evolve. *PLoS Biol.* 8:e1000501.

Kirkpatrick, M., & Barton, N. (2006). Chromosome inversions, local adaptation and speciation. *Genetics* 173:419-434.

Kitano, J., Bolnick, D.I., Beauchamp, D.A., Mazur, M.M., Mori, S., Nakano, T., *et al.* (2008). Reverse evolution of armor plates in the threespine stickleback. *Curr. Biol.* 18:769-774.

Kunte, K., Zhang, W., Tenger-Trolander, A., Palmer, D.H., Martin, A., Reed, R.D., *et al.* (2014). *doublesex* is a mimicry supergene. *Nature* 507:229-232.

Küpper, C., Stocks, M., Risse, J.E., dos Remedios, N., Farrell, L.L., McRae, S.B., *et al.* (2016). A supergene determines highly divergent male reproductive morphs in the ruff. *Nat. Genet.* 48:79-83.

Lall, S.P. (2002). The minerals. In: *Fish Nutrition*, 3rd edition, eds. Halver, J.E. & Hardy, R.W. Academic Press Inc., San Diego, pp. 259-308.

Lamichhaney, S., Berglund, J., Almen, M.S., Maqbool, K., Grabherr, M., Martinez-Barrio, A., *et al.* (2015). Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* 518:371-375.

Lamichhaney, S., Fan, G., Widemo, F., Gunnarsson, U., Thalmann, D.S., Hoeppner, M.P., *et al.* (2016a). Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). *Nat. Genet.* 48:84-88.

Lamichhaney, S., Han, F., Berglund, J., Wang, C., Almén, M.S., Webster, M.T., *et al.* (2016b). A beak size locus in Darwin's finches facilitated character displacement during a drought. *Science* 352:470-474.

Lande, R., & Arnold, S.J. (1983). The measurement of selection on correlated characters. *Evolution* 37:1210-1226.

Lässig, M., Mustonen, V., & Walczak, A.M. (2017). Predicting evolution. *Nat. Ecol. Evol.* 1:0077.

Leal, M.C., Best, R.J., Durston, D., El-Sabaawi, R.W., & Matthews, B. (2017). Stoichiometric traits of stickleback: effects of genetic background, rearing environment, and ontogeny. *Ecol. Evol.* 7:2617-2625.

Lee, C.-R., Wang, B., Mojica, J.P., Mandáková, T., Prasad, K.V.S.K., Goicoechea, J.L., *et al.* (2017). Young inversion with multiple linked QTLs under selection in a hybrid zone. *Nat. Ecol. Evol.* 1:0119.

Linnen, C.R., Poh, Y.P., Peterson, B.K., Barrett, R.D., Larson, J.G., Jensen, J.D., *et al.* (2013). Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science* 339:1312-1316.

Liu, J., Shikano, T., Leinonen, T., Cano, J.M., Li, M.-H., & Merilä, J. (2014). Identification of major and minor QTL for ecologically important morphological traits in three-spined sticklebacks (*Gasterosteus aculeatus*). *G3* 4:595-604.

Lovell, J.T., Juenger, T.E., Michaels, S.D., Lasky, J.R., Platt, A., Richards, J.H., *et al.* (2013). Pleiotropy of *FRIGIDA* enhances the potential for multivariate adaptation. *Proc. R. Soc. Lond. B* 280:20131043.

Lowe, C.B., Sanchez-Luege, N., Howes, T.R., Brady, S.D., Richardson, R.R., Jones, F.C., *et al.* (2017). Detecting differential copy number variation between groups of samples. *Genome Res.* 28:256-265.

Lowry, D.B., & Willis, J.H. (2010). A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol.* 8:e1000500.

Lush, M.E., & Piotrowski, T. (2014). ErbB expressing Schwann cells control lateral line progenitor cells via non-cell-autonomous regulation of Wnt/β-catenin. *eLife* 3:e01832.

MacColl, A.D.C. (2011). The ecological causes of evolution. *Trends Ecol. Evol.* 26:514-522.

Marchinko, K.B. (2009). Predation's role in repeated phenotypic and genetic divergence of armor in threespine stickleback. *Evolution* 63:127-138.

Marchinko, K.B., Matthews, B., Arnegard, M.E., Rogers, S.M., & Schluter, D. (2014). Maintenance of a genetic polymorphism with disruptive natural selection in stickleback. *Curr. Biol.* 24:1289-1292.

Marchinko, K.B., Schluter, D., & Schwenk, K. (2007). Parallel evolution by correlated response: lateral plate reduction in threespine stickleback. *Evolution* 61:1084-1090.

Marques, D.A., Taylor, J.S., Jones, F.C., Di Palma, F., Kingsley, D.M., & Reimchen, T.E. (2017). Convergent evolution of SWS2 opsin facilitates adaptive radiation of threespine stickleback into different light environments. *PLoS Biol.* 15:e2001627.

Martin, A., & Orgogozo, V. (2013). The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* 67:1235-1250.

Mather, K. (1950). The genetical architecture of heterostyly in *Primula sinensis*. *Evolution* 4:340-352.

McGee, L.W., Sackman, A.M., Morrison, A.J., Pierce, J., Anisman, J., & Rokyta, D.R. (2016). Synergistic pleiotropy overrides the costs of complexity in viral adaptation. *Genetics* 202:285-295.

McIntyre, P.B., & Flecker, A.S. (2010). Ecological stoichiometry as an integrative framework in stream fish ecology. *Am. Fish. Soc. Symp.* 73:539-558.

McKay, J.K., Richards, J.H., & Mitchell-Olds, T. (2003). Genetics of drought adaptation in *Arabidopsis thaliana*: I. Pleiotropy contributes to genetic correlations among ecological traits. *Mol. Ecol.* 12:1137-1151.

Meeker, N.D., Hutchinson, S.A., Ho, L., & Trede, N.S. (2007). Method for isolation of PCR-ready genomic DNA from zebrafish tissues. *BioTechniques* 43:610-614.

Milesi, P., Lenormand, T., Lagneau, C., Weill, M., & Labbe, P. (2016). Relating fitness to long-term environmental variations in natura. *Mol. Ecol.* 25:5483-5499.

Miller, C.T., Glazer, A.M., Summers, B.R., Blackman, B.K., Norman, A.R., Shapiro, M.D., *et al.* (2014). Modular skeletal evolution in sticklebacks is controlled by additive and clustered quantitative trait Loci. *Genetics* 197:405-420.

Mills, M.G., Greenwood, A.K., & Peichel, C.L. (2014). Pleiotropic effects of a single gene on skeletal development and sensory system patterning in sticklebacks. *EvoDevo* 5:5.

Morgan, T.H. (1903). *Evolution and adaptation*. Norwood Press, Norwood, MA.

Myhre, F., & Klepaker, T. (2009). Body armour and lateral-plate reduction in freshwater three-spined stickleback *Gasterosteus aculeatus*: adaptations to a different buoyancy regime? *J. Fish Biol.* 75:2062-2074.

Nagy, O., Nuez, I., Savisaar, R., Peluffo, A.E., Yassin, A., Lang, M., *et al.* (2018). Correlated evolution of two copulatory organs via a single *cis*-regulatory nucleotide change. *Curr. Biol.* 28:3450-3457.

Nelson, T.C., & Cresko, W.A. (2018). Ancient genomic variation underlies repeated ecological adaptation in young stickleback populations. *Evol. Lett.* 2:9-21.

Nosil, P., & Crespi, B.J. (2006). Experimental evidence that predation promotes divergence in adaptive radiation. *Proc. Natl. Acad. Sci. USA* 103:9090-9095.

Nosil, P., Villoutreix, R., de Carvalho, C.F., Farkas, T.E., Soria-Carrasco, V., Feder, J.L., *et al.* (2018). Natural selection and the predictability of evolution in *Timema* stick insects. *Science* 359:765-770.

O'Brown, N.M., Summers, B.R., Jones, F.C., Brady, S.D., & Kingsley, D.M. (2015). A recurrent regulatory change underlying altered expression and Wnt response of the stickleback armor plates gene *EDA*. *eLife* 4:e05290.

Orr, H.A. (1998). The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution* 52:935-949.

Orr, H.A. (2000). Adaptation and the cost of complexity. *Evolution* 54:13-20.

Ortiz-Barrientos, D., Engelstädter, J., & Rieseberg, L.H. (2016). Recombination rate evolution and the origin of species. *Trends Ecol. Evol.* 31:226-236.

Otto, S.P. (2004). Two steps forward, one step back: the pleiotropic effects of favoured alleles. *Proc. R. Soc. Lond. B* 271:705-714.

Otto, S.P., & Jones, C.D. (2000). Detecting the undetected: estimating the total number of loci underlying a quantitative trait. *Genetics* 156:2093-2107.

Ozgul, A., Tuljapurkar, S., Benton, T.G., Pemberton, J.M., Clutton-Brock, T.H., & Coulson, T. (2009). The dynamics of phenotypic change and the shrinking sheep of St. Kilda. *Science* 325:464-467.

Paaby, A.B., & Rockman, M.V. (2013). The many faces of pleiotropy. *Trends Genet.* 29:66-73.

Paccard, A., Wasserman, B.A., Hanson, D., Astorg, L., Durston, D., Kurland, S., *et al.* (2018). Adaptation in temporally variable environments: stickleback armor in periodically breaching bar-built estuaries. *J. Evol. Biol.* 31:735-752.

Parnell, N.F., Hulsey, C.D., & Streelman, J.T. (2012). The genetic basis of a complex functional system. *Evolution* 66:3352-3366.

Patwa, Z., & Wahl, L.M. (2008). The fixation probability of beneficial mutations. *J. Royal Soc. Interface* 5:1279-1289.

Pearson, K., Lee, A., Warren, E., Fry, A., & Fawcett, C.D. (1901). VIII. Mathematical contributions to the theory evolution. —IX. On principle of homotyposis and its relation, the variability of the individual, and to that of the race. Part I. —Homotyposis in the vegetable Kingdom. *Phil. Trans. R. Soc. A* 197:285-379.

Peichel, C.L., & Marques, D.A. (2017). The genetic and molecular architecture of phenotypic diversity in sticklebacks. *Phil. Trans. R. Soc. Lond. B* 372:20150486.

Peichel, C.L., Nereng, K.S., Ohgi, K.A., Cole, B.L.E., Colosimo, P.F., Buerkle, C.A., *et al.* (2001). The genetic architecture of divergence between threespine stickleback species. *Nature* 414:901-905.

Peichel, C.L., Ross, J.A., Matson, C.K., Dickson, M., Grimwood, J., Schmutz, J., *et al.* (2004). The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. *Curr. Biol.* 14:1416-1424.

Price, T.D., Grant, P.R., Gibbs, H.L., & Boag, P.T. (1984). Recurrent patterns of natural selection in a population of Darwin's finches. *Nature* 309:787-789.

Protas, M., Tabansky, I., Conrad, M., Gross, J.B., Vidal, O., Tabin, C.J., *et al.* (2008). Multi-trait evolution in a cave fish, *Astyanax mexicanus*. *Evol. Dev.* 10:196-209.

Provine, W.B. (2001). *The origins of theoretical population genetics*. University of Chicago Press, Chicago, IL.

Pujolar, J.M., Ferchaud, A.L., Bekkevold, D., & Hansen, M.M. (2017). Non-parallel divergence across freshwater and marine three-spined stickleback *Gasterosteus aculeatus* populations. *J. Fish Biol.* 91:175-194.

Ramaekers, A., Claeys, A., Kapun, M., Mouchel-Vielh, E., Potier, D., Weinberger, S., *et al.* (2019). Altering the temporal regulation of one transcription factor drives evolutionary trade-offs between head sensory organs. *Dev. Cell* 50:1-13.

Rebeiz, M., Pool, J.E., Kassner, V.A., Aquadro, C.F., & Carroll, S.B. (2009). Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. *Science* 326:1663-1667.

Reimchen, T.E. (1980). Spine deficiency and polymorphism in a population of *Gasterosteus aculeatus*: an adaptation to predators? *Can. J. Zool.* 58:1232-1244.

Reimchen, T.E. (1983). Structural relationships between spines and lateral plates in threespine stickleback (*Gasterosteus aculeatus*). *Evolution* 37:931-946.

Reimchen, T.E. (1992). Injuries on stickleback from attacks by a toothed predator (*Oncorhynchus*) and implications for the evolution of lateral plates. *Evolution* 46:1224-1230.

Rennison, D.J., Heilbron, K., Barrett, R.D.H., & Schluter, D. (2015). Discriminating selection on lateral plate phenotype and its underlying gene, *Ectodysplasin*, in threespine stickleback. *Am. Nat.* 185:150-156.

Rennison, D.J., Rudman, S.M., & Schluter, D. (2019). Genetics of adaptation: experimental test of a biotic mechanism driving divergence in traits and genes. *Evol. Lett.* 3:513-520.

Robertson, S., Bradley, J.E., & MacColl, A.D.C. (2017). Eda haplotypes in three-spined stickleback are associated with variation in immune gene expression. *Sci. Rep.* 7:42677.

Rockman, M.V. (2012). The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* 66:1-17.

Rogers, S.M., Tamkee, P., Summers, B., Balabahadra, S., Marks, M., Kingsley, D.M., *et al.* (2012). Genetic signature of adaptive peak shift in threespine stickleback. *Evolution* 66:2439-2450.

Roy, P.K., Witten, P.E., Hall, B.K., & Lall, S.P. (2002). Effects of dietary phosphorus on bone growth and mineralisation of vertebrae in haddock (*Melanogrammus aeglefinus* L.). *Fish Physiol. Biochem.* 27:35-48.

Rudman, S.M., Goos, J.M., Burant, J.B., Brix, K.V., Gibbons, T.C., Brauner, C.J., *et al.* (2019). Ionome and elemental transport kinetics shaped by parallel evolution in threespine stickleback. *Ecol. Lett.* 22:645-653.

Sabeti, P.C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., *et al.* (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature* 449:913-918.

Scarcelli, N., Cheverud, J.M., Schaal, B.A., & Kover, P.X. (2007). Antagonistic pleiotropic effects reduce the potential adaptive value of the FRIGIDA locus. *Proc. Natl. Acad. Sci. USA* 104:16986-16991.

Schindler, D.E., & Eby, L.A. (1997). Stoichiometry of fishes and their prey: implications for nutrient recycling. *Ecology* 78:1816-1831.

Schwander, T., Libbrecht, R., & Keller, L. (2014). Supergenes and complex phenotypes. *Curr. Biol.* 24:R288-R294.

Selby, J.P., & Willis, J.H. (2018). Major QTL controls adaptation to serpentine soils in *Mimulus guttatus*. *Mol. Ecol.* 27:5073-5087.

Sella, G., & Barton, N.H. (2019). Thinking about the evolution of complex traits in the era of genome-wide association studies. *Annu. Rev. Genom. Hum. Genet.* 20:461-493.

Siepielski, A.M., DiBattista, J.D., & Carlson, S.M. (2009). It's about time: the temporal dynamics of phenotypic selection in the wild. *Ecol. Lett.* 12:1261-1276.

Soria-Carrasco, V., Gompert, Z., Comeault, A.A., Farkas, T.E., Parchman, T.L., Johnston, J.S., *et al.* (2014). Stick insect genomes reveal natural selection's role in parallel speciation. *Science* 344:738-742.

Sousa, A., Magalhães, S., & Gordo, I. (2011). Cost of antibiotic resistance and the geometry of adaptation. *Mol. Biol. Evol.* 29:1417-1428.

Srivastava, A.K., Pispa, J., Hartung, A.J., Du, Y., Ezer, S., Jenks, T., *et al.* (1997). The Tabby phenotype is caused by mutation in a mouse homologue of the *EDA* gene that reveals novel mouse and human exons and encodes a protein (ectodysplasin-A) with collagenous domains. *Proc. Natl. Acad. Sci. USA* 94:13069-13074.

Stern, D.L., & Orgogozo, V. (2008). The loci of evolution: how predictable is genetic evolution? *Evolution* 62:2155-2177.

Stern, D.L., & Orgogozo, V. (2009). Is genetic evolution predictable? *Science* 323:746-751.

Sterner, R.W., & Elser, J.J. (2002). *Ecological stoichiometry: the biology of elements from molecules to the biosphere*. Princeton University Press, Princeton, NJ.

Sullivan, M., Reid, S.W.J., Ternent, H., Manchester, N.J., Roberts, R.J., Stone, D.A.J., *et al.* (2007). The aetiology of spinal deformity in Atlantic salmon, *Salmo salar* L.: influence of different commercial diets on the incidence and severity of the preclinical condition in salmon parr under two contrasting husbandry regimes. *J. Fish Dis.* 30:759-767.

Taylor, E.B., & McPhail, J.D. (1986). Prolonged and burst swimming in anadromous and freshwater threespine stickleback, *Gasterosteus aculeatus*. *Can. J. Zool.* 64:416-420.

Terekhanova, N.V., Logacheva, M.D., Penin, A.A., Neretina, T.V., Barmintseva, A.E., Bazykin, G.A., *et al.* (2014). Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback *Gasterosteus aculeatus*. *PLoS Genet.* 10:e1004696.

Tuttle, Elaina M., Bergland, Alan O., Korody, Marisa L., Brewer, Michael S., Newhouse, Daniel J., Minx, P., *et al.* (2016). Divergence and functional degradation of a sex chromosome-like supergene. *Curr. Biol.* 26:344-350.

Via, S., Conte, G., Mason-Foley, C., & Mills, K. (2012). Localizing F_{ST} outliers on a QTL map reveals evidence for large genomic regions of reduced gene exchange during speciation-with-gene-flow. *Mol. Ecol.* 21:5546-5560.

Vielma, J., & Lall, S.P. (1998). Control of phosphorus homeostasis of Atlantic salmon (*Salmo salar*) in fresh water. *Fish Physiol. Biochem.* 19:83-93.

Villablanca, E.J., Renucci, A., Sapède, D., Lec, V., Soubiran, F., Sandoval, P.C., *et al.* (2006). Control of cell migration in the zebrafish lateral line: implication of the gene "tumour-associated calcium signal transducer," tacstd. *Dev. Dyn.* 235:1578-1588.

Visscher, P.M., & Goddard, M.E. (2019). From R.A. Fisher's 1918 paper to GWAS a century later. *Genetics* 211:1125-1130.

Visscher, P.M., Wray, N.R., Zhang, Q., Sklar, P., McCarthy, M.I., Brown, M.A., *et al.* (2017). 10 Years of GWAS discovery: biology, function, and translation. *Am. J. Hum. Genet.* 101:5-22.

Wada, H., Dambly-Chaudière, C., Kawakami, K., & Ghysen, A. (2013). Innervation is required for sense organ development in the lateral line system of adult zebrafish. *Proc. Natl. Acad. Sci. USA* 110:5659-5664.

Wada, H., Ghysen, A., Satou, C., Higashijima, S.-i., Kawakami, K., Hamaguchi, S., *et al.* (2010). Dermal morphogenesis controls lateral line patterning during postembryonic development of teleost fish. *Dev. Biol.* 340:583-594.

Wada, H., Iwasaki, M., & Kawakami, K. (2014). Development of the lateral line canal system through a bone remodeling process in zebrafish. *Dev. Biol.* 392:1-14.

Wade, M.J., & Kalisz, S. (1990). The causes of natural selection. *Evolution* 44:1947-1955.

Wagner, C.E., McIntyre, P.B., Buels, K.S., Gilbert, D.M., & Michel, E. (2009). Diet predicts intestine length in Lake Tanganyika's cichlid fishes. *Funct. Ecol.* 23:1122-1131.

Wagner, G.P., Kenney-Hunt, J.P., Pavlicev, M., Peck, J.R., Waxman, D., & Cheverud, J.M. (2008). Pleiotropic scaling of gene effects and the 'cost of complexity'. *Nature* 452:470-472.

Wagner, G.P., & Zhang, J. (2011). The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. *Nat. Rev. Genet.* 12:204-213.

Wallbank, R.W.R., Baxter, S.W., Pardo-Diaz, C., Hanly, J.J., Martin, S.H., Mallet, J., *et al.* (2016). Evolutionary novelty in a butterfly wing pattern through enhancer shuffling. *PLoS Biol.* 14:e1002353.

Wang, J., Wurm, Y., Nipitwattanaphon, M., Riba-Grognuz, O., Huang, Y.-C., Shoemaker, D., *et al.* (2013). A Y-like social chromosome causes alternative colony organization in fire ants. *Nature* 493:664-668.

Wang, Z., Liao By Fau - Zhang, J., & Zhang, J. (2010). Genomic patterns of pleiotropy and the evolution of complexity. *Proc. Natl. Acad. Sci. USA* 107:18034-18039.

Wark, A.R., Mills, M.G., Dang, L.H., Chan, Y.F., Jones, F.C., Brady, S.D., *et al.* (2012). Genetic architecture of variation in the lateral line sensory system of threespine sticklebacks. *G3* 2:1047-1056.

Wark, A.R., & Peichel, C.L. (2010). Lateral line diversity among ecologically divergent threespine stickleback populations. *J. Exp. Biol.* 213:108-117.

Watanabe, K., Stringer, S., Frei, O., Umićević Mirkov, M., de Leeuw, C., Polderman, T.J.C., *et al.* (2019). A global overview of pleiotropy and genetic architecture in complex traits. *Nat. Genet.* 51:1339-1348.

Welch, J.J., & Waxman, D. (2003). Modularity and the cost of complexity. *Evolution* 57:1723-1734.

Wellenreuther, M., & Hansson, B. (2016). Detecting polygenic evolution: problems, pitfalls, and promises. *Trends Genet.* 32:155-164.

Weyhenmeyer, G.A., Hartmann, J., Hessen, D.O., Kopáček, J., Hejzlar, J., Jacquet, S., *et al.* (2019). Widespread diminishing anthropogenic effects on calcium in freshwaters. *Sci. Rep.* 9:10450.

Williams-Blangero, S., VandeBerg, J.L., Subedi, J., Aivaliotis, M.J., Rai, D.R., Upadhayay, R.P., *et al.* (2002). Genes on chromosomes 1 and 13 have significant effects on *Ascaris* infection. *Proc. Natl. Acad. Sci. USA* 99:5533-5538.

Williams-Blangero, S., VandeBerg, J.L., Subedi, J., Jha, B., Corrêa-Oliveira, R., & Blangero, J. (2008). Localization of multiple quantitative trait loci influencing susceptibility to infection with *Ascaris lumbricoides*. *J. Infect. Dis.* 197:66-71.

Wilson, A.J., Pilkington, J.G., Pemberton, J.M., Coltman, D.W., Andrew, D.J.O., Byrne, K.A., *et al.* (2005). Selection on mothers and offspring: whose phenotype is it and does it matter? *Evolution* 59:451-463.

Witten, P.E., Owen, M.A.G., Fontanillas, R., Soenens, M., McGurk, C., & Obach, A. (2016). A primary phosphorus-deficient skeletal phenotype in juvenile Atlantic salmon *Salmo salar*: the uncoupling of bone formation and mineralization. *J. Fish Biol.* 88:690-708.

Wootton, R.J. (1976). *The biology of the sticklebacks*. Academic Press, London.

Wray, G.A. (2007). The evolutionary significance of cis-regulatory mutations. *Nat. Rev. Genet.* 8:206-216.

Wucherpfennig, J.I., Miller, C.T., & Kingsley, D.M. (2019). Efficient CRISPR-Cas 9 editing of major evolutionary loci in sticklebacks. *Evol. Ecol. Res.* 20:107-132.

Xu, S. (2003). Theoretical basis of the Beavis effect. *Genetics* 165:2259-2268.

Yamasaki, Y.Y., Mori, S., Kokita, T., & Kitano, J. (2019). Armour plate diversity in Japanese freshwater threespine stickleback (*Gasterosteus aculeatus*). *Evol. Ecol. Res.* 20:51-67.

Yeaman, S. (2013). Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proc. Natl. Acad. Sci. USA* 110:E1743-E1751.

Yeaman, S., & Whitlock, M.C. (2011). The genetic architecture of adaptation under migration-selection balance. *Evolution* 65:1897-1911.

Yoshizawa, M., Yamamoto, Y., O'Quin, K.E., & Jeffery, W.R. (2012). Evolution of an adaptive behavior and its sensory receptors promotes eye regression in blind cavefish. *BMC Biol.* 10:108.

Zeller, M., Lucek, K., Haesler, M.P., Seehausen, O., & Sivasundar, A. (2012a). Little evidence for a selective advantage of armour-reduced threespined stickleback individuals in an invertebrate predation experiment. *Evol. Ecol.* 26:1293-1309.

Zeller, M., Lucek, K., Haesler, M.P., Seehausen, O., & Sivasundar, A. (2012b). Signals of predation-induced directional and disruptive selection in the threespine stickleback. *Evol. Ecol. Res.* 14:193-205.

Zeng, J., de Vlaming, R., Wu, Y., Robinson, M.R., Lloyd-Jones, L.R., Yengo, L., *et al.* (2018). Signatures of negative selection in the genetic architecture of human complex traits. *Nat. Genet.* 50:746-753.

VITA

Sophie L. Archambeault received her B.A. in Zoology from Connecticut College, where she discovered her interest in molecular biology and evolution while pursuing her desire to work with animals and the natural world. Her thesis work, under the direction of Phillip Barnes, investigated whether life history traits of *Drosophila melanogaster* had evolved in response to long term laboratory culture. Sophie got her first experience of combining molecular biology and conservation biology during her internship at the Center for Reproduction of Endangered Species at the San Diego Zoo, which expanded her research interests to include the evolution of organismal diversity. After graduation from Connecticut College, Sophie worked in medical genetics at UC San Francisco in the laboratory of Mignon Loh, until she took a year off to indulge her love of travel. After traveling around the world with her sister, Sophie returned to San Francisco and earned a Master's degree in Conservation Biology, Ecology and Systematics at San Francisco State University, in the laboratory of Karen Crow. It was through this work that Sophie discovered the amazing diversity of fishes and the beautiful molecular evolutionary story of Hox genes. Her newfound interest in evolutionary developmental biology nudged her towards a PhD program and the exciting field of stickleback research. Sophie will continue exploring the genetic bases and molecular mechanisms that drive developmental differences in natural populations of stickleback in her next research position and continue to explore the world when she finds the time.