UNIVERSITY OF CALIFORNIA
SANTA CRUZ

ADAPTATION GENOMICS OF SURFPERCH POPULATIONS IN THE CONTEXT OF RAPID ENVIRONMENTAL CHANGE

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ECOLOGY AND EVOLUTIONARY BIOLOGY

by

Jason A. Toy

December 2022

The Dissertation of Jason A. Toy is approved:

_____
Professor Kristy J. Kroeker, chair

_____
Professor Giacomo Bernardi

_____
Professor Peter T. Raimondi

_____
Professor Malin L. Pinsky

_____
Peter Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures & Tables

CHAPTER 1

CHAPTER 2

Abstract

ADAPTATION GENOMICS OF SURFPERCH POPULATIONS IN THE CONTEXT OF RAPID ENVIRONMENTAL CHANGE

by Jason A. Toy

While the immediate impacts of global climate change are of serious concern, the outcomes of these environmental changes for populations will ultimately play out over multiple generations. Despite this, our understanding of the evolutionary impacts of climate-related environmental change is still in its early stages, particularly in the marine realm. Furthermore, evolutionary processes can act over short, ecological timescales, such that they may play a key role in both the short-term resistance and long-term resilience of natural populations. Therefore, the objective of this dissertation is to better incorporate evolutionary processes into the developing understanding of the ecological effects of global change. In particular, I focus my investigations on the relationships between genetic diversity, local adaptation, and environmental change to better understand the evolutionary factors that contribute to a population's adaptive capacity and resilience. A population's genetic diversity may translate to response (phenotypic) diversity, the level of which will determine the likelihood of evolutionary rescue via the portfolio effect. Adaptation of

subpopulations to their local conditions (local adaptation) has the potential to enhance broad-scale genetic diversity within a species, potentially increasing resilience in the face of environmental change, but can simultaneously reduce local diversity, increasing the risk of extinction for subpopulations if gene flow is low. In this dissertation, I use a pair of marine fish species with unique life-histories (*Embiotoca jacksoni and Brachyistius frenatus*; family *Embiotocidae*) to 1) test the molecular impact of environmental change on an important temperate fish group, 2) investigate the scale of genetic diversity and admixture along the Pacific coast of North America, 3) provide evidence of local adaptation among subpopulations, and 4) associate genomic differences between subpopulations with regional environmental differences to better understand the physiological pressures imposed by climate variables and form hypotheses for the genetic mechanisms that may underlie ongoing adaptation to climate change.

To my friends and family, who have always supported me in finding my own path.

To Celine, who stuck with me on this journey through the highs and lows.

- and -

To the coastal ecosystems of the Northeast Pacific, which first inspired me to take this path and continue to fill me with wonder to this day.

Acknowledgements

Advisors and Committee

I cannot express enough gratitude to the members of my Dissertation Committee, whose guidance was indispensable in the completion of this dissertation. I first thank my advisors, Kristy Kroeker, who chaired my committee, and Giacomo Bernardi, who co-advised me. Kristy, thank you for welcoming me into the field of subtidal research and for putting your faith in me as a young graduate student. I cannot thank you enough for your endless support in all aspects of completing this degree and beginning my career. Thank you for encouraging me to go the extra mile, and for always having my best interest in mind. I am also deeply grateful for your open-mindedness in taking on a co-advisor and your encouragement to explore my scientific interests, which allowed me to investigate the evolutionary themes in this dissertation. Thank you, Giacomo, for welcoming me into your lab as a co-advised student. Thank you also for your willingness to discuss evolutionary processes at any time or place, and for being up for a dive, sail, or a motorcycle ride. I am grateful as well for your infectious enthusiasm for evolution and fish, which reminded me why I started this journey and helped restore my motivation when I needed it the most. I also extend my thanks to Pete Raimondi. Pete, thank you for your genuine interest in the details of my research. Your broad-scale perspective was always insightful and helped improve the quality of my work. Thank you also for your willingness to "talk stats" with me, even over the phone at a moment's notice. Finally, I thank Malin

Pinsky for his support for and engagement in my research. Malin, your dedication to mentorship and genuine interest in my research helped me push through a tight timeline and get this dissertation over the finish line. Thank you for graciously hosting me in your lab for two straight weeks. It was truly a pleasure to work with you and your lab members.

Non-Committee Mentors, Advisors, and Collaborators

There are several additional people I would like to thank who, despite not officially being part of my committee, were incredibly generous with their time and assistance. First, Cheryl Logan was instrumental in the completion of the gene expression analyses in Chapter 1. Cheryl, thank you for giving your time and lab space to teach someone with no experience with next-generation sequencing how to analyze RNAseq data. You and your lab (especially Jake Cline) made that chapter possible. Thank you to Yui Takeshita and Joe Warren, who patiently answered my questions about carbonate chemistry analysis and oceanographic sensing equipment and helped me to analyze my water samples. I also am deeply grateful to Nina Overgaard Therkildsen and Nicolas Lou. Your guidance on the analysis of whole-genome sequencing data was indispensable and I thank you for your incredible generosity. Finally, I would like to recognize the support of the many undergraduate volunteers that worked with me over the course of my degree, as well as my peers within the Department of Ecology and Evolutionary Biology (EEB) at UCSC. I particularly thank the members of the Kroeker and Bernardi Labs, who have offered countless hours of feedback, advice, and moral support.

Furthermore, I would like to thank those who mentored me even before beginning my graduate studies. Eric Sanford, Jay Stachowicz, and Matt Whalen were all instrumental, not only in the development of my interest in marine science, but also in helping me establish myself in the field. Thank you all for your continued support over the many years since we first worked together. I would not be where I am today without it.

Funding Sources

Funding for this work was supplied by many sources including the American Fisheries Society (through the Stephen Berkeley Marine Conservation Fellowship), the Earl and Ethel Myers Oceanographic Trust, Friends of Long Marine Lab Student Research and Education Award, and the EEB Department at UCSC. My stipend and tuition were supported by grants from the National Science Foundation, the Packard Foundation, and the California Conservation Genomics Project. During the final analysis and writing for this dissertation I was also supported by the Dissertation Year Fellowship from the EEB Department.

Friends and Family

Finally, I would like to express my deepest gratitude to my family and friends who have supported me through this journey. To my parents, thank you for a lifetime of support in all its innumerable forms. Your love and efforts allowed me the opportunity to follow this path. To my siblings, Nick, Lauren, and Alex, thank you for your emotional and inspirational support. You all kept me smiling during difficult

moments. I am fortunate to also have many close friends on whom I leaned over the past 6+ years. Thank you all for keeping my life balanced and healthy through this journey, and for your understanding when I needed to be absent. To J.B. Ballard, David Dito, and Gustavo Navarro, thank you, especially, for the heartfelt advice and conversation that kept me grounded in trying times.

To my partner, Celine de Jong, thank you for the innumerable ways in which you have supported me through this endeavor. Your love and emotional support kept me moving forward from day to day. Thank you for keeping me grounded and engaged in the things I love. Thank you also for your direct support of my research. As a marine scientist yourself, your intellectual support and technical skill were invaluable to the completion of this dissertation. For everything from acting as my dependable dive buddy and helping me think through field logistics to your input on manuscript word choice, thank you for all your assistance in bringing this dissertation to fruition.

Chapter publications

Chapter 1 is reproduced from the following published article:

Toy, J. A., Kroeker, K. J., Logan, C. A., Takeshita, Y., Longo, G. C., & Bernardi, G. (2022). Upwelling-level acidification and pH/pCO_2 variability moderate effects of ocean acidification on brain gene expression in the temperate surfperch, *Embiotoca jacksoni*. *Molecular Ecology*, 31, 4707– 4725. https://doi.org/10.1111/mec.16611

Chapter 2 is reproduced from the following submitted manuscript:

Toy, J.A., Bernardi, G. A high-quality reference genome of the kelp surfperch, *Brachyistius frenatus* (Embiotocidae), a wide-ranging Eastern Pacific reef fish with no pelagic larval stage.

Introduction

Over the past several decades, a dramatically improved understanding of the potential ecological impacts of climate change has emerged (Bellard et al., 2012; Root et al., 2003). Populations are largely considered to have three possible pathways to persistence in the face of large-scale climate change, none of which are mutually exclusive. Populations may tolerate local environmental changes through plastic (non-heritable) alterations in the behavior and physiology of individuals (acclimatization), shift their ranges through migration to match changing distributions of suitable habitat, or adapt genetically through the combined action of selection and heritable trait variation. If none of these response pathways are effective, populations may go extinct. This general framework is often referred to as "tolerate, move, adapt, or die".

Present populations often have some ability to tolerate climate-induced changes in their physical and biological environments through plastic changes in their physiology or behavior (Sandblom et al., 2014; Seebacher et al., 2015; Snell-Rood et al., 2018). In the ocean, organisms may make physiological and behavioral changes in response to the metabolic and biochemical costs of increased temperature, acidification, and deoxygenation, with varied results. For example, some species may exhibit compensatory feeding in response to increased metabolic costs, while others seem to show no evidence of compensation (Kindinger et al., 2022). Behavioral plasticity is common (Beever et al., 2017; Nagelkerken & Munday, 2016), and many species will adjust their position in the water column or among habitat types to avoid

stressful conditions (e.g. low dissolved-oxygen; Breitburg et al., 1997). Additionally, individuals may be able to confer adaptive phenotypic changes on their offspring in response to changing environmental conditions through transgenerational plasticity (Eirin-Lopez & Putnam, 2018; Hofmann, 2017; Hoshijima & Hofmann, 2019). The organism-level effects of global climate change are clearly widespread and pervasive, with effects cascading through many direct and indirect pathways that have already led to significant and concerning alterations to natural ecosystems (e.g., McPherson et al., 2021).

At the same time, it is now well accepted that evolutionary processes, facilitated by genomic diversity, can act over short ecological timescales (Schoener, 2011; Thompson, 1998). These evolutionary processes (i.e., migration, mutation, selection, and drift) may therefore play a critical role in both the short-term resistance and long-term resilience of natural populations to environmental change. Despite this understanding, investigations into the evolutionary impacts of global change on natural populations and the evolutionary processes that mediate population responses remain underrepresented in the literature (Sunday et al., 2014). To better understand and more capably predict the outcomes of rapid environmental change, evolutionary processes must be better incorporated into the existing ecological framework of global change impacts.

Evolutionary responses to environmental change

The relative effectiveness of the "tolerate", "move", and "adapt" strategies for a given population will often depend on the life-history characteristics of the species

(Carlson et al., 2014). For example, rapid adaptive evolution will be less likely among large, long-lived organisms with long generation times ("k-selected" species), and migration – if possible – is likely to be a more effective response. On the other hand, adaptive evolutionary change in "r-selected" species with short generation times, high fecundity, and large population sizes is much more likely to keep pace with rapid rates of environmental change (Berg et al., 2010; Buoro & Carlson, 2014; Carlson et al., 2014; Travis et al., 2013). Phenotypic plasticity is likely to be beneficial in populations where plasticity has already evolved and where environmental variability is predictable (Bitter et al., 2021; Kroeker et al., 2020).

If environmental change is too widespread, rapid, or unpredictable, plastic and migratory responses may be ineffectual or insufficient for population persistence. In this case, an evolutionary response may be a population's last line of defense against extirpation (Holt, 1990; Smith, 1989). The first aim of this dissertation is therefore to better understand the circumstances under which a population will either successfully adapt or die and to what extent populations may fit such criteria. In other words, I seek to answer the question: How likely are likely are populations to successfully undergo evolutionary rescue? – defined here as the persistence of a population through a disturbance due to the spread of new or pre-adapted genotypes. Factors that may determine the likelihood of evolutionary rescue may be grouped into three general categories: demographic factors, extrinsic factors, and genetic factors (Carlson et al., 2014). Although this dissertation will focus primarily on the dynamics between genetic factors, I briefly summarize all three categories below to describe the

context under which genetic factors may act to determine the likelihood of evolutionary rescue.

Perhaps the most important demographic factor affecting the adaptive capacity of a population is its effective size (Bell, 2013; Gomulkiewicz & Holt, 1995). A large absolute population size (N) promotes demographic resilience, but the effective size (N_e) will determine the relative strengths of genetic drift and selection. As effective population size increases, selection becomes more effective, increasing the substitution rate of advantageous mutations and decreasing the rate for deleterious mutations (Lanfear et al., 2014). The effective size of a population therefore impacts its rate of adaptive evolution, which – all else being equal – then determines the maximum rate of environmental change it can withstand (Lynch & Lande, 1993).

The most obvious extrinsic factor is the rate of environmental change. A rapid rate of change causes the level of population maladaptation to be large, giving the population less time to adapt before falling below the theoretical size threshold where stochastic extirpation becomes likely (Barrett & Schluter, 2008; Gomulkiewicz & Holt, 1995; Lande & Shannon, 1996; Lynch & Lande, 1993). This rate can also describe changes in biotic interactions, which can have significant impacts on the likelihood of evolutionary rescue. Intraspecific competition, for example, may limit population growth rate and therefore size and absolute mutation rate, diminishing the likelihood of evolutionary rescue, while interspecific competition may actually promote evolutionary rescue by strengthening selection on advantageous alleles (Osmond & de Mazancourt, 2013). Similarly, predation is predicted to increase the

strength of purifying selection on maladaptive alleles (Jones, 2008), further increasing the rate of adaptive evolution.

Finally, genetic factors include those that impact the genetic variation available for selection to act upon, as well as genetic interactions among this variation. Standing variation is critical to the evolutionary rescue of a population, as it is immediately available to selective processes (Barrett & Schluter, 2008; Bell, 2013; Gomulkiewicz & Holt, 1995), but new (de novo) variation produced by mutation can also play an important role in adaptive responses. For example, if genetic variation is limiting, a greater mutation rate may increase the likelihood of evolutionary rescue (Orr & Unckless, 2008). Additionally, the architecture of genetic variation present within a population may affect the rate of evolution. Linkage between alleles that act antagonistically (i.e., interlocus conflict) will slow the rate of adaptive evolution, as can antagonistic pleiotropy (Lanfear et al., 2014; Otto, 2004). Positive correlation between either deleterious or advantageous alleles, however, can speed the rate of adaptive evolution (Chevin, 2012).

Genetic variation provides the raw material necessary for adaptation and evolutionary rescue in the face of anthropogenic environmental change, but this variation exists within the context of demographic and extrinsic factors. How much genetic variation and time is required for evolutionary rescue to occur, the prominent sources of this variation, and how standing variation will be indirectly impacted by other population- and community-level climate change responses are all still unclear. This dissertation therefore seeks to provide empirical insight into the process of

climate adaptation using a combination of controlled laboratory experiments and a species-wide analysis of genomic diversity and local adaptation.

Study system: Climate change and drivers of genetic diversity in the Northeast Pacific

The Pacific Coast of North America contains some of the most valuable and productive marine ecosystems in the world. Commercial fisheries fuel economies, coastal state parks and MPAs provide immense cultural value, and local ecosystems provide additional critical services unrelated to fisheries for coastal communities. As discussed above, the resilience of these important systems to environmental change may in part depend on the genomic diversity present within their constituent species. To better characterize and quantify this within-species genetic variation, the ecological processes involved in its preservation must be more closely considered. Within the coastal ecosystems of the NE Pacific, the extensive spatial heterogeneity in the physical and chemical environment represents an obvious candidate for spatially variable selection, a key factor behind the maintenance of within-species variation, (Hofmann et al., 2014). Driving this spatial heterogeneity are latitudinal gradients in seawater temperature, light availability, and wave exposure, as well as a mosaic of upwelling occurrence and intensity, which creates corresponding mosaics of pH, dissolved oxygen, and nutrient availability (Chan, Barth, Blanchette, Byrne, Chavez, Cheriton, Feely, Friederich, Gaylord, Gouhier, & others, 2017). Importantly, it is also some of these same environmental drivers (temperature, pH, and dissolved oxygen) that are expected to continue to change rapidly as a result of anthropogenic CO_2 emissions, making the NE Pacific an ideal system in which current populations

may be used to investigate potential adaptive responses to future environmental change.

The mosaic of environmental conditions across the NE Pacific allows for the possibility of local adaptation within subpopulations of a species by creating spatial differences in selective pressures (Hofmann et al., 2014). Local adaptation can be defined as the maintenance of different alleles, or different frequencies of alleles, in different geographic regions. Although this differential selection can reduce genetic diversity within a local subpopulation, it can promote a greater diversity at the species level (Slatkin, 1987), as different alleles are favored and therefore preserved in different regions. While differential selection is a key component of local adaptation, it is not the only requirement. If individuals move and mate freely between subpopulations, gene flow will occur and differential selection may have a limited effect on the genetic structuring of a species (panmixia). Local adaptation therefore also requires gene flow to be limited (e.g., through limited dispersal) to prevent outbreeding depression and genetic homogenization.

Given the extent of environmental heterogeneity along the NE Pacific and the potential of local adaptation to promote within species diversity, I seek to address the following question: How do environmental heterogeneity and local adaptation interact to mediate a species' response to global environmental change? To answer this question, one must first determine the strength and prevalence of local adaptation in coastal ecosystems. The prevalence of local adaptation in marine populations was once poorly understood, as the openness and continuity of the marine realm was

historically believed to strongly promote gene flow across broad geographic scales. More recent work, however, has revealed that local adaptation in marine populations is much more prevalent than previously assumed (Conover et al., 2006; Sanford & Kelly, 2011), even in species with large dispersal distances. This indicates that the local adaptation may be even more pronounced in marine species with limited dispersal – for example, those that lack a pelagic larval stage (e.g., Bernardi, 2000).

The surfperches (family: Embiotocidae), are an abundant group of viviparous (live bearing) nearshore fishes in the NE Pacific that all lack a pelagic larval phase. These unique life-history traits lead to a greatly reduced capacity for dispersal (and therefore gene flow) compared to ecologically similar species, greatly increasing the likelihood that local adaptation has occurred at some point in their evolutionary histories. Though there is a paucity of data on adult dispersal in these species, high site fidelity has been demonstrated in some species (Hixon, 1981; Bernardi, pers. comm.). Additionally, Bernardi (2000) demonstrated the presence of coast-wide genetic structure in black surfperch (*Embiotoca jacksoni*) using mitochondrial markers.

For the most part, these embiotocids are shallow-water species, living along sandy substrate, rocky reef, kelp forest, and even estuarine habitats. Among the productive rocky reef and kelp forest communities present from Alaska to Baja California, many species are abundant, including *E. jacksoni*, *E. lateralis*, *E. caryi*, *B. frenatus*, *Phanaerodon vacca*, *P. furcatus*, *P. atripes*, *Rhacochilus toxotes*, and *Cymatogaster aggregata*. All species are mesopredators, primarily consuming a

variety of crustaceans, especially amphipods (Laur & Ebeling, 1983; Love, 2011). Their abundance, combined with their tendency to feed on herbivorous invertebrates can also exert significant top-down control in these systems (Davenport & Anderson, 2007). Because of their ecological importance, and because their life histories increase the likelihood of local adaptation, the research conducted in this dissertation focused on two species in this family: *E. jacksoni* and *B. frenatus* (kelp surfperch).

Overview of research

Because local adaptation requires differential selection across geographic space, **Chapter 1** of this dissertation focused on testing whether and how individuals of a single subpopulation respond physiologically (molecularly) to different environmental regimes. The goal was to test whether observed differences in coastal environmental regimes translate into differences in selective pressures, which would in turn have some effect on the fitness of individuals. As a proof of concept, and to simplify the experimental system to a manageable level, I took two of the factors listed above that exhibit strong spatial heterogeneity – the mean and temporal variability in seawater acidity – and tested their effects on brain gene expression in *E. jacksoni*. I combined the results of this study with a dataset from a previously conducted pilot study to disentangle the relative impacts of moderate, extreme, and variable pH/pCO_2 on neurological function.

My results from Chapter 1 demonstrated the differential response of a single population to different pH regimes, which can be interpreted as evidence for spatial variation in selective pressures. I therefore proceeded to a population genomic

investigation of local adaptation in subpopulations with similarly divergent environments. To make use of current technical and statistical advances in genomic research, I used a whole genome resequencing approach to conduct a range-wide characterization of genomic diversity and population structure in the widespread embiotocid species, *B. frenatus*. To facilitate these analyses, I first needed to assemble a reference genome sequence for the species, which I describe in **Chapter 2**. A combination of long-read nanopore sequencing and short-read shotgun sequencing resulted in a high quality and highly complete reference assembly. The assembly revealed a relatively small genome size for this species (596 Mb) compared to similar species, which provides further motivation for the use of this species for resequencing-based studies. The comparison of this genome to the reference sequence of *E. jacksoni* revealed high synteny between the two species with no indications of major inversions or rearrangements. This similarity of the genomes of these two species could facilitate future studies investigating the effects of ecological differences between the two species on genome structure, as well as species-wide patterns of genomic diversity.

In **Chapter 3**, I resequenced 158 *B. frenatus* individuals from 13 locations along the west coast of North America and characterized the genetic structure and diversity present in the species across most of its current range. Using multiple allele frequency-based outlier methods, I conducted a genome-wide scan for loci under selection across the sampled populations. Furthermore, I coupled this genomic data with regional environmental data to test for correlations between allele frequency

differences and climate-related variables. A genomic study such as this is valuable in several ways. First, my analysis of population genetic structure across the species' range has provided insight into the extent of gene flow between geographic regions. Second, this study identified many candidates for genes under selection that have provided hypotheses for the mechanisms through which *B. frenatus* may adaptively respond to environmental change. Third, analysis of the association of loci under selection with environmental variables provided valuable insight into the relative importance of several climate variables in shaping the adaptive variation in this species. Finally, this genomic approach inherently isolates genetic sources of trait variation from non-genetic sources (i.e., plasticity), allowing for more confident conclusions about genetic adaptation when compared to phenotype-based studies.

Chapter 1: Upwelling-level acidification and pH/pCO_2 variability moderate effects of ocean acidification on brain gene expression in the temperate surfperch, *Embiotoca jacksoni*

This chapter was originally published in a peer reviewed journal and is reproduced here for inclusion in this dissertation. The citation for the original publication is:

Toy, J. A., Kroeker, K. J., Logan, C. A., Takeshita, Y., Longo, G. C., & Bernardi, G. (2022). Upwelling-level acidification and pH/pCO_2 variability moderate effects of ocean acidification on brain gene expression in the temperate surfperch, *Embiotoca jacksoni*. *Molecular Ecology*, 31, 4707– 4725. https://doi.org/10.1111/mec.16611

ABSTRACT

Acidification-induced changes in neurological function have been documented in several tropical marine fishes. Here, we investigate whether similar patterns of neurological impacts are observed in a temperate Pacific fish that naturally experiences regular and often large shifts in environmental pH/pCO_2. In two laboratory experiments, we tested the effect of acidification, as well as pH/pCO_2 variability, on gene expression in the brain tissue of a common temperate kelp forest/estuarine fish, *Embiotoca jacksoni*. Experiment 1 employed static pH treatments (target pH = 7.85/7.30), while Experiment 2 incorporated two variable treatments that oscillated around corresponding static treatments with the same mean (target pH = 7.85/7.70) in an eight-day cycle (amplitude ± 0.15). We found that patterns of global gene expression differed across pH level treatments. Additionally, we identified differential expression of specific genes and enrichment of specific gene sets (GSEA) in comparisons of static pH treatments and in comparisons of static and variable pH treatments of the same mean pH. Importantly, we found that pH/pCO_2 variability decreased the number of differentially expressed genes detected between high and low pH treatments, and that inter-individual variability in gene expression was greater in variable treatments than static treatments. These results provide important confirmation of neurological impacts of acidification in a temperate fish species and, critically, that natural environmental variability may mediate the impacts of ocean acidification.

INTRODUCTION

Ocean acidification (OA; here defined as both increased ocean pCO_2 and decreased pH) has been identified as a major threat to marine species (Kroeker et al., 2013). Several studies have documented changes in neurological functioning, including altered cognition, sensory function and behavior, in marine fish (e.g., Domenici et al., 2012; S. L. Hamilton et al., 2017; T. J. Hamilton et al., 2013; Munday et al., 2010; Pistevos et al., 2015), raising concerns about neurological impacts leading to changes in the strength of species interactions (e.g., predation). In contrast, more recent work has questioned the generality and replicability of such impacts across studies and species (Clark et al., 2020a). Additionally, much of the evidence of neurological impacts comes from studies of a few tropical reef species under static pH/pCO_2 regimes (Nagelkerken & Munday, 2016). Physiological evidence indicates these neurological impacts may be the result of a hypercapnia-driven reversal of electrochemical gradients in GABAergic neurons. This has been hypothesized to result from internal acid-base balance processes that lead to an accumulation of intracellular [HCO_3^-] and/or a decrease in extracellular [Cl^-] (Heuer & Grosell, 2014; Nilsson et al., 2012). This shift in ion concentrations is thought to cause neuron depolarization upon $GABA_A$ receptor activation rather than the hyperpolarization expected under non-acidified conditions, reversing the functional nature of these neurons from inhibitory to excitatory and presumably causing the observed shifts in cognition and behavior (Heuer & Grosell, 2014; Nilsson et al., 2012; Schunter et al., 2019). Given this body of evidence, altered neurological

function may be a major pathway through which changing seawater carbonate chemistry will impact fitness in marine fish. Continued work elucidating the molecular mechanisms underlying these changes is therefore critical for moving the field forward.

Many coastal ecosystems experience significant environmental variability over a range of temporal scales, including fluctuations in seawater pH/pCO_2 (Chan, Barth, Blanchette, Byrne, Chavez, Cheriton, Feely, Friederich, Gaylord, Gouhier, Hacker, et al., 2017; Hofmann et al., 2011; Kang et al., 2022; Kroeker et al., 2020). In upwelling regions, where deeper, more acidic water is brought to the ocean surface, pH can vary by half a unit over a period of weeks (Hirsh et al., 2020; Hofmann et al., 2011). In seagrass beds, pH can vary by a whole unit over a period of hours to weeks due to diurnal fluctuations in photosynthesis and respiration and tidal movement (Duarte et al., 2013; Hofmann et al., 2011). These fluctuations often reach or exceed predictions for the mean future ocean pH under OA (Gruber et al., 2012; Hauri et al., 2013; Takeshita et al., 2015). In previous studies, exposure to low pH/high pCO_2 seawater has affected indicators of fish neurological function anywhere from 2-12 days after exposure has ceased (T. J. Hamilton et al., 2013; Munday et al., 2010), but it is unclear what duration of exposure elicits these effects. How temporal environmental variability moderates fish responses to low pH/high pCO_2 remains critically understudied, leaving many unanswered questions about how OA may realistically affect populations in nature (but see Jarrold et al., 2017; Jarrold & Munday, 2019; Schunter et al., 2021). Investigations into the effects of pH/pCO_2

variability are important for accurate prediction of the severity of impacts acidification will have on natural populations and ecosystems (Kroeker et al., 2020). For example, if variability dampens the effects documented in studies using only static pH/pCO_2 treatments (as seen in Jarrold et al., 2017 and Jarrold & Munday, 2019, where diel pCO_2 fluctuations ameliorated impairments in behavior and growth seen under static decreases in pCO_2), we may be overestimating the effects of OA, and overlooking an important role that variability may play as a provider of temporal refuge. Conversely, if variability exacerbates negative impacts of acidification, acting as an additional stressor, we may be underestimating the potential impact of acidification on natural populations.

We expect physiological responses to OA to be reflected in the gene expression of the affected organism (Griffiths et al., 2019; S. L. Hamilton et al., 2017). In particular, we expect changes in brain gene expression to be associated with shifts in neurological and cognitive function (Schunter et al., 2016). Given the proposed mechanism of OA-induced cognitive impairment described above, we expect experimental acidification to impact expression in genes related to the maintenance of homeostasis and neuronal signaling, such as ion transporter and signal receptor genes, and those involved in the GABAergic signaling pathway. Changes in expression in these gene categories have been noted in spiny damselfish (Schunter et al., 2018) and three-spined stickleback (Lai et al., 2016), but this has not yet been investigated in a temperate reef species with an evolutionary history of exposure to fluctuating pCO_2.

Here, we present two experimental studies of the effects of acidification on brain gene expression in a common temperate reef fish, the black surfperch (*Embiotoca jacksoni*). Surfperches make up a large proportion of fish biomass on California rocky reefs (Laur & Ebeling, 1983) and support an immensely popular recreational fishery. *E. jacksoni* is found in both upwelling reef systems and estuarine seagrass ecosystems, and therefore has an evolutionary history of exposure to variable pH conditions that are often more extreme than those experienced by tropical reef fish (Hofmann et al., 2011). A few studies have investigated the effects of acidification on temperate reef fishes (e.g. S. L. Hamilton et al., 2017; Kwan et al., 2017; Cline et al., 2020), but surfperches are unique because they exhibit viviparity and no pelagic larval phase, with young born as developed juveniles. Additionally, *E. jacksoni* has limited adult dispersal (Bernardi, 2000, 2005; Hixon, 1981). These two life-history traits increase the likelihood of adaptation to local environmental conditions in *E. jacksoni*, which may lead to divergent effects of acidification in this species compared to other temperate fish (e.g., greater physiological adaptation to OA in populations that have historically experienced local acidification).

Experiment 1 was designed to determine the presence and extent of any impacts of acidification on *E. jacksoni* brain gene expression and used a static and more extreme acidified treatment (pH ~7.30). In Experiment 2, we used a less extreme static treatment and incorporated two variable treatments with different mean pH levels to mimic upwelling-scale pH variability. This experiment was designed to test the potential role of temporal pH/pCO_2 variability in mediating any neurological

effects of acidification. Together, the results of these experiments provide a more comprehensive understanding of the impacts of acidification on marine organisms, particularly in dynamic, temperate ecosystems.

METHODS

Collections and Acclimation

We collected young-of-the-year *E. jacksoni* from Elkhorn Slough (Monterey County, CA) using a beach seine. Collected fish were placed in coolers and driven back to UCSC-CSC, where they were kept in outdoor flow-through containers until the start of each experiment. For exact dates of collections, acclimation periods, and experimental manipulations see Table S1.1.

Experimental Design

We conducted two separate experiments with similar methods in November 2015 and September 2017 at University of California, Santa Cruz's Coastal Science Campus (UCSC-CSC). Both experiments treated *E. jacksoni* juveniles in outdoor flow-through seawater systems. In Experiment 1 (2015), we set target treatments at pH 7.85, representing a common current upwelling condition along the coast of Central California, and pH 7.30, representing a current extreme estuarine event or future extreme upwelling event (Chan, Barth, Blanchette, Byrne, Chavez, Cheriton, Feely, Friederich, Gaylord, Gouhier, Hacker, et al., 2017; Hofmann et al., 2011; Lowe et al., 2019; Takeshita et al., 2015). Both treatments in this experiment held the

target pH constant (static) over the course of the experiment. Five randomly assigned juvenile *E. jacksoni* were distributed across two replicate tanks at pH 7.30 and four were distributed across two replicate tanks at pH 7.85 (opaque 200 L plastic drums). Seawater pH treatments were replicated only at the level of holding tanks. Replicates were then brought down to experimental pH levels over seven days. Tissue sampling was conducted after 23 days of treatment (Table S1.1).

In Experiment 2 (2017), we incorporated upwelling-scale pH variability into two of the treatments, and target pH levels were set at more conservative levels. Two static pH treatments were set at target pH levels of 7.85 and 7.70, approximating conservative present and future reef conditions during the upwelling season (Chan, Barth, Blanchette, Byrne, Chavez, Cheriton, Feely, Friederich, Gaylord, Gouhier, Hacker, et al., 2017; Takeshita et al., 2015). For each static treatment, there was a corresponding variable treatment that oscillated around the same mean pH as the static treatment with an amplitude of ± 0.15 pH and a period of eight days (Figure 1.1), approximating a typical upwelling pattern (Hofmann et al., 2011). An additional treatment, hereafter referred to as "ambient", had a static target pH of 8.00. However, because our pH control system was not capable of increasing pH above that of the incoming seawater, periodic natural decreases in the pH of the input seawater below 8.00 caused this treatment to exhibit an intermediate level of variability between that of the static and variable treatments (Figure S1.1). We used 10 header buckets (two per treatment) to create the five different pH treatments, with two replicate tanks per header. We randomly assigned six juvenile *E. jacksoni* to replicate tanks (translucent

61 L plastic containers; 6 individuals x 4 replicate tanks x 5 pH treatments) and allowed them to acclimate at ambient pH for 2-3 days. We then allowed the pH of each treatment to slowly approach its starting pH (target pH 8.00, 7.85, or 7.70) over a period of two days. After an additional 4-5 days, the variable treatments began their programmed oscillations (Figure S1.1). Due to logistical restrictions, the treatments were separated into two groups (pH 7.85 and ambient treatments, pH 7.70 treatments) that were staggered in their timing by one day (Figure S1.1). Using a custom-built LabView program, set points for the variable treatments were changed throughout the experiment at intervals of 0.003125 pH per hour to create 8-day cycles. During this experiment, fish were removed from their treatment tanks on two occasions to conduct behavioral assays, after which they were returned to their treatment tanks. Because we were met with logistical challenges that precluded the proper execution of these trials, these data were not analyzed. Eight days (1 full cycle of the variable treatments) were allowed to elapse between the last trial and tissue sampling, which was conducted after 22 days (Figure S1.1; Table S1.1).

pH Control System & Sampling of Seawater Chemistry

Seawater pH was manipulated using a custom-built feedback control system. Two large sumps received a continuous flow of ambient seawater. One of these sumps ("low pH") was continuously bubbled with CO_2 gas, while the other ("ambient") was left untreated. Lines from both sumps fed seawater into header buckets at varying rates to create pH treatments. The pH of each bucket was

continuously measured by Honeywell Durafet II sensors connected to Honeywell Universal Dual Analyzers (UDAs; see Kapsenberg et al., 2017). Seawater pH in each header was controlled through a feedback system, where a solenoid valve determined the flow of low pH seawater to the header to either increase or decrease pH. The mixed treatment water from each header then flowed out into two replicate holding tanks. We oxygenated and mixed seawater in each header using air pumps/stones and/or water pumps (Experiment 2 only).

Prior to beginning each experiment, we calibrated the Durafet sensors from the header buckets using equimolar Tris buffer (DelValls & Dickson, 1998) obtained from the Dickson Lab (Scripps Institution of Oceanography). In Experiment 1, discrete water samples were taken from the replicate tanks at seven time points and used for characterization of carbonate chemistry via spectrophotometric pH analysis and open cell total alkalinity titration (Dickson et al., 2007). In Experiment 2, samples were taken from headers at five time points and used for post-hoc calibration of Durafet pH measurements. Using a handheld sensor (YSI), we also measured temperature, pH, dissolved oxygen, and salinity in each replicate tank daily and, in Experiment 2, in each header as well to allow for calibration of YSI pH measurements to calibrated Durafet measurements. See Tables 1.1 & 1.2 for measured and calculated seawater parameters.

Experimental Considerations

A heat wave struck Santa Cruz during Experiment 2. This added stress may have contributed to the juvenile mortality observed during this experiment (48 out of 120 fish died, unrelated to treatment; ANOVA, $F = 2.088$, $p = 0.133$). Additionally, some of the Durafet sensors (four of ten) used to control the pH in the headers experienced heavy fouling by microalgae toward the end of Experiment 2. This likely led to artificially high pH measurements for about two hours around midday due to photosynthesis, and thus a corresponding over-correction by the pH control system. Because of this issue, the pH of certain tanks was likely lower during midday than their respective set points. To better understand the scale of this over-correction, we conducted a test 11 days after tissue sampling, in which all Durafets were placed in the same header with no active pH control (Figure S1.2). Though the effect of fouling on recorded pH appeared to strengthen over the 11 days since tissue sampling, this post-hoc test revealed variability in the impact across headers, and a relatively even distribution of fouling across treatments (Figure S1.2). The greatest spike during this test occurred in one of the two ambient headers with a magnitude of ~0.5 pH units, but this treatment was not included in most of our analyses, and thus we believe it does not affect our conclusions. Examination of experiment Durafet readings from the ambient header (which, due to its high set point and limited pH control, displayed the true extent of the pH spikes) revealed that significant spikes (deviation of ~ 0.05 pH units or greater) in this most affected treatment only began occurring approximately three days before tissue dissection. Because our test indicated that the other headers

were affected to a much smaller degree (Figure S1.2), the other treatments likely did not experience midday spikes of greater than 0.05 pH units for any significant duration prior to the end of the experiment. To prevent the inclusion of spurious pH data points in the characterization of the experimental treatments, we used the continuous Durafet pH and temperature data for the dates September 1 – 17, after which we used pH and temperature data from daily YSI readings taken from each header. This shift in sampling frequency likely explains much of the apparent increased variability of the pH treatments after September 17 (Figure S1.1), as the YSI data represents only a daily snapshot of the pH of each header. The pH sensor within the YSI is also functionally different from those within Durafets (Martz et al., 2010). Finally, outside of the daily spikes, the Durafet pH data collected September 17-24 showed no obvious departure from the precision seen earlier in the experiment. We therefore contend that, apart from the midday overcorrections experienced at the end of the experiment, the true precision and variability of the pH treatments was unchanged after September 17, and any apparent changes reflect only a change in the pH-sensing instrument used.

Fish Care and Handling

This experiment was run under the approval of UCSC IACUC project proposals BERNG1312 and KROEK1503_A2. We performed system checks at least daily and fed fish frozen shrimp every day (Experiment 1) or a mix of frozen brine shrimp, S*pirulina* brine shrimp, and mysis shrimp every other day (Experiment 2).

Tanks were cleaned and excess food removed approximately 7 times per week (Experiment 1) and at least once per week (Experiment 2). To minimize stress, a shelter was placed in each replicate. To reduce heat and sun exposure, shade cloth was kept over the top of the replicate tanks whenever water monitoring, cleaning, or feeding was not occurring.

Tissue Sampling

At the end of each experiment, we dissected tissue from four individuals from each treatment. Individuals were dissected one at a time, with all dissections taking less than 10 min from the time of fish removal from its tank. Brain and lateral muscle tissue were dissected and sequenced in Experiment 1 for use in the transcriptome assembly, but only brain tissue was sequenced in Experiment 2. In Experiment 2, the whole brain was dissected at the approximate time when the variable pH treatments were crossing (in the ascending direction) their target mean pH levels (Figure S1.1). Only brain gene expression analysis will be further discussed here. Tissue was stored in screw-cap tubes and flash frozen in liquid nitrogen. We stored all tissue samples at -80°C until RNA extraction.

RNA Extraction and Library Preparation

Dissected whole brains were arbitrarily subsampled and homogenized using a Qiagen TissueLyser. A discussion of the potential effects of subsampling are included below in the 'Inter-Individual Variability in Gene Expression' section of the

Discussion. We extracted RNA using the Qiagen RNeasy® Mini extraction kit. RNA quality and quantity were assessed using a NanoDrop spectrophotometer and Qubit fluorometer. RNA was stored in DEPC-treated water at -80°C. cDNA libraries were prepared from 1 µg of total RNA using the New England Biolabs NEBNext® Ultra™ II RNA Library Prep kit. Prepared libraries were sequenced on an Illumina HiSeq 4000 (150bp SE) at the QB3 Vincent J. Coates Genomics Sequencing Laboratory at the University of California, Berkeley.

Read Processing & Transcriptome Assembly

We removed adapters and trimmed/removed low quality reads using the Trimmomatic software (v0.36; Bolger et al., 2014; parameters = LEADING:2 TRAILING:2 SLIDINGWINDOW:4:2 MINLEN:25) and quality checked the trimmed sequences using FastQC (v0.11.7; Andrews, 2010). We used the trimmed reads from all sequenced samples from both experiments to assemble a brain/muscle tissue combined transcriptome for *E. jacksoni* using the genome-guided TopHat/Cuffmerge/Cufflinks pipeline (default parameters; TopHat v2.1.1, Cufflinks v2.2.1; Trapnell et al., 2012). This pipeline creates separate assemblies for each sample, which are then merged. A draft, scaffold-level *E. jacksoni* genome assembly was used as the reference (see Supplementary Materials). We annotated the assembly by running a blastx query (e-value cutoff = 1e-3; NCBI, Altschul et al., 1990) against the SwissProt database (uniprot_sprot.dat.gz downloaded April 25, 2020; The UniProt Consortium, 2021).

Multivariate Analysis of Gene Expression

In Experiment 1, we sequenced transcripts from both muscle and brain tissue, but only brain gene expression will be discussed here. To characterize global gene expression of individuals, trimmed reads were aligned and quantified into gene-level expression data using bowtie (v1.2.3; Langmead et al., 2009) and RSEM (v1.3.3; Li & Dewey, 2011) within the Trinity software package (v2.9.1; Haas et al., 2013). Raw read counts were then filtered to remove genes with low expression using the default parameters of the filterByExpr function (min.count = 10, min.total.count = 15, large.n = 10, min.prop = 0.7) in the R package, edgeR (v3.34.0; R Core Team, 2021; Robinson et al., 2010). We normalized the read counts using the TPM method, as implemented by the calcNormFactors function in the edgeR package, then $\log_2$-transformed the data using the cpm function (prior.count = 2). The transformed data were dimensionally reduced through multidimensional scaling (metric MDS in Experiment 1, nMDS in Experiment 2) using Manhattan distances, as implemented through the wcmdscale and metaMDS functions in the vegan package for R (v2.5.7; Oksanen et al., 2020).

To test whether global gene expression profiles differed among treatments, we ran a permutational multivariate analysis of variance (PERMANOVA; Anderson, 2001, 2017) on the transformed expression data using the adonis function of the vegan package (method="manhattan", perm=1,000,000). In Experiment 1, the sole model factor was pH level (7.85, 7.30). In Experiment 2, the model was run with two

factors: pH level (7.85, 7.70) and pH variability (static, variable). For Experiment 2, pairwise comparisons between treatments were conducted using the pairwise.adonis function of the pairwiseAdonis package (sim.method = "manhattan", perm = 1,000,000) (Martinez Arbizu, 2020).

Differential Gene Expression Analysis

Using the gene-level counts matrix created by RSEM, we identified differentially expressed genes (DEGs) between all pairwise treatment comparisons using the edgeR package, as implemented through Trinity. To buffer against false positives and noise due to the experimental conditions described above, we used a conservative FDR cutoff value of 0.001 (-P parameter) and a fold-change cutoff of 1.5 (-C parameter) to create the final list of DEGs for each treatment comparison. We then repeated MDS procedures and PERMANOVA tests as described above, using only these DEGs.

Functional Enrichment Analysis

To identify gene sets (groups of functionally related genes) that were significantly enriched in a given treatment comparison (e.g., pH 7.85 vs. pH 7.30) we used the threshold-free analytical method, Gene Set Enrichment Analysis (GSEA; Subramanian & Tamayo et al., 2005) as implemented through the FGSEA package for R (Korotkevich et al., 2021). Given a ranked list of genes derived from differential expression analysis, this method yields a list of gene sets from user-

supplied gene set databases - in this case GO (The Gene Ontology Consortium 2020), KEGG (Kanehisa & Goto, 2000), and the MSigDB Hallmark collection (Liberzon et al., 2015) - that are enriched among upregulated and downregulated genes. Enrichment analysis was completed for Experiment 1 and 2 separately, and the resulting enriched gene sets from analogous treatment comparisons were then contrasted across experiments to identify commonly enriched gene sets.

RESULTS

Seawater pH Manipulation

Mean seawater pH levels were maintained near their target set points in each experiment. Seawater parameters and information on how they were calculated are given in Tables 1.1 and 1.2. Note that for consistency, we continue to use target pH levels to refer to each treatment.

Sequencing & Transcriptome Assembly

The TopHat/Cufflinks pipeline yielded a transcriptome made up of 71,933 assembled transcripts grouped into 39,258 putative genes. Aligning the trimmed reads back to the assembled transcriptome resulted in an 82.98% alignment rate. A blastx search of the transcriptome against the SwissProt database revealed that 8,836 transcripts represented nearly full-length transcripts (>80% alignment coverage), and that 16,574 proteins were represented in the transcriptome at some level of alignment coverage. We used BUSCO (v4.0.6; Simão et al., 2015) to quantify the completeness

of our transcriptome and found that of the 3,640 BUSCO orthologs in the Actinopterygii dataset, 76.9% were found complete (41.2% single-copy, 35.7% duplicated), 6.5% were found fragmented, and 16.6% were missing. For more statistics on the assembly, see Table S1.2.

Total sequenced reads per sample are listed in Tables S1.3 and S1.4. Of the 39,258 putative genes in the assembled transcriptome, 22,961 (Experiment 1) and 33,597 (Experiment 2) remained in each dataset after filtering for genes with low expression.

Multivariate Analyses of Global Gene Expression

Single-factor PERMANOVA analysis identified a strong and significant effect of pH level on global gene expression in both Experiment 1 (single-factor; r^2 = 0.811, F = 25.766, p = 0.029) and Experiment 2 (two-factor, ambient excluded; r^2 = 0.159, F = 2.53, p = 0.021), with pH explaining 81% and 16% of the observed variation, respectively (Tables S1.5 & S1.6). In Experiment 2, we did not detect an effect of pH variability on global gene expression (r^2 = 0.037, F = 0.59 p = 0.890) or an interaction of pH level and variability (r^2 = 0.052, F = 0.84, p = 0.524). Pairwise comparisons of all treatments (including ambient) revealed two comparisons were nearly significantly different (Table S1.7): ambient vs. 7.70 static (r^2 = 0.230, F = 1.79, p = 0.086) and 7.85 static vs. 7.70 static (r^2 = 0.292, F = 2.47, p = 0.057). We visualized the differences in global gene expression patterns between treatments using MDS (Figures S1.3 & S1.4).

Differential Gene Expression Analysis

We found 10,656 DEGs between the treatments in Experiment 1 (Figure 1.2). In Experiment 2, we found a total of 200 DEGs across all treatment comparisons (Table 1.3; Figure S1.5). The 7.85 static vs. 7.70 static comparison produced the majority of DEGs (159). The 7.85 static vs 7.70 variable and 7.85 variable vs. 7.70 static comparisons produced 11 DEGs each and six genes each were differentially expressed in the static vs. variable comparisons of both the 7.85 and 7.70 pH levels (one gene was consistently differentially expressed across the two comparisons).

Since we did not expect transcriptome-wide shifts in gene expression across pH variability treatments in Experiment 2 (Figure S1.4), and were instead interested in how the expression of acidification response genes was affected by the introduction of environmental variability, we repeated our multivariate analyses of gene expression for only the DEG subset of Experiment 2. For consistency, analogous analyses were also performed for the Experiment 1 DEG subset (Table S1.8; Figure S1.6).

DEG expression differed among pH levels in Experiment 2 ($r^2 = 0.388$, $F = 10.77$, $p = 0.004$), with pH explaining 39% of the observed variation (Figure 1.3). We did not detect an effect of pH variability ($r^2 = 0.041$, $F = 1.15$ $p = 0.291$). The interaction of pH level and variability, however, was marginally significant ($r^2 = 0.139$, $F = 3.85$, $p = 0.052$) (Table S1.9). Pairwise comparisons of all treatments revealed a significant difference between the pH 7.85 and pH 7.70 static treatments (p

= 0.029) and nearly-significant differences between the ambient and 7.85 static treatments (p = 0.057) and between the ambient and 7.70 static treatments (p = 0.086; Table S1.10). The comparison of 7.85 static vs. 7.85 variable was also nearly significant (p = 0.057), but the comparison of 7.70 static vs. 7.70 variable was less so (p = 0.171).

Analysis of Within-Treatment Variances

To test the effect of pH variability on within-treatment variability in gene expression, we calculated the variance of normalized gene expression for each DEG within each treatment in Experiment 2 (excluding ambient) and averaged the variance across all DEGs (Figure 1.4). For each pH level (7.85, 7.70), we then calculated F-ratios by dividing the mean variance of the variable treatment by that of the static treatment. We log-transformed these variances and compared the distributions to a t-distribution centered at 0 using one-tailed t-tests (Table S1.11). We found that at both pH levels, the average variance in DEG expression was greater in the variable pH treatment than in the static pH treatment (pH 7.85: p = 0.0001; pH 7.70: p = 0.03487).

Functional Enrichment Analysis

In Experiment 1, gene set enrichment analysis using FGSEA revealed 240 enriched gene sets among the upregulated genes and 343 enriched gene sets among the downregulated genes in the pH 7.30 treatment compared to the 7.85 treatment (FDR < 0.05; Table S1.12). In Experiment 2, 61 gene sets were enriched among upregulated

genes and 71 among downregulated genes in the 7.70 static treatment compared to the 7.85 static treatment (Table S1.13). At the 7.85 pH level, we found 44 enriched gene sets among upregulated genes and 202 among downregulated genes in the variable treatment compared to the static treatment (Table S1.14). At the 7.70 pH level, 115 and 22 gene sets were enriched among the upregulated and downregulated genes, respectively, in the variable treatment compared to the static treatment (Table S1.15). To aid interpretation, the enriched gene sets for the 7.85/7.30 comparison in Experiment 1 and the 7.85 static/7.70 static comparison in Experiment 2 were further collapsed into clusters of gene sets (using a gene set similarity coefficient) using the AutoAnnotate and clusterMaker2 applications for the Cytoscape software platform. These clusters were manually summarized based on their constituent gene sets (Tables 1.4 and 1.5).

To assess consistency of response to acidification across experiments, we determined the overlap in enriched gene sets between the pH 7.85/pH 7.30 comparison of Experiment 1 and the pH 7.85 static/pH 7.70 static comparison of Experiment 2 (Figure 1.5, Table 1.6). To assess consistency in expression response across the two static vs. variable treatment comparisons of Experiment 2 (7.85 static/7.85 variable, 7.70 static/7.70 variable), we determined the overlap in enriched gene sets between these comparisons (Figure 1.6, Table S1.16).

DISCUSSION

Impacts of Static Acidification

Numerous studies have demonstrated impaired behavior and sensory function in fish and other marine organisms when exposed to low pH/high pCO_2, (Domenici et al., 2012; S. L. Hamilton et al., 2017; T. J. Hamilton et al., 2013; Munday et al., 2010; Pistevos et al., 2015)(but see Clark et al., 2020). Though the mechanisms behind these changes are still poorly understood, significant effects of low pH/increased pCO_2 on brain gene expression have been documented in a few marine fish species that demonstrate associated impairments in behavior (Lai et al., 2016; Schunter et al., 2016, 2018, 2021). In this study we tested whether brain gene expression is similarly impacted in a temperate reef fish that experiences prolonged periods of natural acidification. Across both experiments presented here, we found that global gene expression was significantly affected by acidification (high vs. low pH). Comparing results across experiments, the number of detected DEGs increased with more extreme acidification, as did the number of enriched gene sets. A similar increase in DEGs with increased intensity of acidification has also been reported in the olfactory bulb of coho salmon (Williams et al., 2019). This marked increase in effect size indicates that further acidification past the already-low pH of 7.70 can have a substantial additional impact on the physiology of marine fish. This pattern may have important implications for the management of marine ecosystems and the services they provide as our global society struggles to control CO_2 emissions.

Although a greater number of gene sets were enriched in Experiment 1 than in the comparison of the static treatments of Experiment 2, similar enrichment themes emerged. In both experiments, static acidification led to the upregulation of gene sets related to turnover in the proteome and transcriptome that may reflect ongoing physiological adaptation to altered environmental conditions (Tables 1.4 & 1.5). Additionally, static acidification in both experiments led to the downregulation of gene sets related to the MAPK cascade, G protein-coupled receptor signaling pathways, plasma membrane components, secretory vesicles and granules, neuroactive ligand-receptor interaction, and calcium ion binding, indicating a general reduction in cell signaling, including neuroactive signaling, in response to high pCO_2. In general, this is the opposite of the response seen in similar gene sets in spiny damselfish (*Acanthochromis polyacanthus*) (Schunter et al., 2018) and the olfactory bulb of coho salmon (Williams et al., 2019). Schunter et al. (2019) proposed that high *p*CO2-induced changes in electrochemical gradients across GABAergic neuron membranes may initiate a "vicious cycle" of feedbacks and ultimately an increase in excitatory activity in the brain that may explain behavioral changes seen in other species. If this is indeed the case, the downregulation of gene sets related to neuroactive signaling seen here may represent a species-specific adaptive response aimed at combating maladaptive runaway excitation in acidic waters (see discussion of $GABA_A$ receptor related genes below). Finally, downregulation of gene sets related to growth and morphogenesis, cell-cell adhesion, and the cytoskeleton indicate potential disruption of cell growth and development due to increased cellular stress.

Similar themes of upregulated transcription and cellular stress response have also been documented in the muscle tissue of Pacific rockfish (S. L. Hamilton et al., 2017).

We also identified divergent sets of genes enriched between the moderate (Experiment 2, target pH 7.70) and extreme (Experiment 1, target pH 7.30) acidification treatments, indicative of a potential threshold effect as static pH decreases. In comparison to the static acidification in Experiment 2, static acidification in Experiment 1 resulted in the up- and downregulation of additional gene sets related to metabolic processes (Table 1.4). These changes may again indicate further shifts to the synthesis of stress response proteins, or to isoforms that are better suited to an altered cellular environment. Because, at least in humans, there can be interaction/crosstalk between cellular stress response pathways and the innate immune system signaling pathways (Muralidharan & Mandrekar, 2013), the upregulation of an additional 6 gene sets related to the innate immune response may further indicate increased cellular stress. The acidification in Experiment 1 also resulted in the downregulation of broad categories of gene sets related to basic neurological functions, behavior, and cognition, which supports the hypothesis that acidification can lead to behavioral impairment in this species though, as mentioned above, the specific mechanisms through which OA induced alterations in neurobiology might impact fish behavior are still not well understood (Tresguerres & Hamilton, 2017).

In Experiment 2, additional gene sets related to the regulation of gene expression (including epigenetic regulation) were upregulated, again indicating a systemic shift in gene expression and response to cellular stress. We also found a unique downregulation of a large number of gene sets related to immune response, which may in part reflect the external conditions of this experiment, particularly the unusually warm ambient temperatures. The combination of physical and chemical stressors may have led to the suppression of the immune system in fish in the acidified treatment. Suppression or dysregulation of immune function is a well-established response to stress (Dhabhar, 2014), and heat stress-induced immunosuppression, specifically, has been noted across various animal systems (Nardone et al., 2010).

The biological themes of the enriched gene sets in both experiments are consistent with enriched categories identified in previous studies in tropical reef fish (e.g. Schunter et al., 2016, 2018) and salmon (Williams et al., 2019). Interestingly, however, the pattern of enrichment in *E. jacksoni* under acidified conditions is generally opposite to the enrichment pattern found by Schunter et al. (2018) when comparing acute or developmentally (together: cis-generationally) exposed spiny damselfish to control (untreated) individuals, but closely resembles the pattern of gene set enrichment that Schunter et al. found when comparing transgenerationally exposed *A. polyacanthus* to those that were developmentally exposed to acidified conditions (not the control treatment). The contrast of our results may reflect the transgenerational and evolutionary exposure history of *E. jacksoni* populations to

naturally acidic environments. While *A. polyacanthus* on coral reefs may experience diurnal pCO_2 fluctuations on the scale of ± 50-150 µatm (Schunter et al., 2021 and references therein), *E. jacksoni* in upwelling regions are likely to regularly experience prolonged increases in pCO_2 (days to weeks) from as low as ~300 µatm to >1000 µatm (Chavez et al., 2018; Kroeker et al., unpublished data). Those living in estuaries can experience even greater shifts in carbonate chemistry over even shorter timescales (Duarte et al., 2013; Hofmann et al., 2011). It is therefore likely that the juvenile *E. jacksoni* in our experiments were transgenerationally exposed to acidified conditions *in-situ*. Additionally, while both species lack a pelagic larval stage, *A. polyacanthus* is a substrate spawner and *E. jacksoni* is a live-bearing species. This means that the *E. jacksoni* used for this experiment may also have developmentally experienced their mothers' natural environmental exposures prior to their birth. Because of this potential in-situ transgenerational exposure, our experimental design may be more comparable to the transgenerational treatment used by Schunter et al (2018).

As mentioned above, it has been suggested that the cause of previously documented acidification-induced behavioral changes in fish may not only be due to a reversal of electrochemical gradients that flip the nature of GABAergic neurons from inhibitory to excitatory (Nilsson et al., 2012), but also due to a positive feedback cycle that may develop as a response to this increase in excitatory activity in the brain (Schunter et al., 2018, 2019). This proposed response consists of an increase in GABA release and in the abundance of $GABA_A$ receptors, which under non-acidified

conditions would serve to reduce overactivity in the brain, but under acidified conditions likely act to exacerbate the overactivity. Some previous studies in fish have seen changes in expression consistent with this response, such as increased expression of $GABA_A$ receptor subunits and transporter genes (e.g., Schunter et al., 2016; 2018; Lai et al., 2016). However, the fish in Experiment 1 showed the opposite response in GABA-related genes. In Experiment 1, $GABA_A$ receptor subunit isoforms α (1–6), β (1–3), γ (1–3), ρ (2), and π were all downregulated in the pH 7.30 treatment, along with many other GABA signaling genes, including glutamate decarboxylases *gad1* and *gad2* (Lai et al., 2016) and *gabarapl2*. Interestingly, a similar general downregulation of GABAergic signaling pathways was recently noted in *A. polyacanthus* at CO_2 seeps, but not in other reef fish species (Kang et al., 2022). A study on Pacific coho salmon (Williams et al., 2019) also found no changes in $GABA_A$ receptor subunit expression in the olfactory bulb under increased pCO_2, but did find an increase in the expression of a $GABA_B$ receptor subunit (*gabbr2*), which was instead downregulated in *E. jacksoni* in our Experiment 1. Williams et al. (2019) also found significant changes in the expression of other genes associated with GABA signaling, including downregulation of the *slc6a13* gene involved in GABA uptake, which we also saw downregulated in *E. jacksoni* in Experiment 1 (see Table S1.17 for all differentially expressed GABA-related genes). These divergent responses between *E. jacksoni,* Pacific salmon, and tropical fish species could represent species-specific adaptation to differing environmental conditions. In the case of *E. jacksoni*, which frequently experiences periods of high pCO_2, the

downregulation of GABA-related genes under high pCO_2 may be an adaptation that prevents or interrupts the excitatory positive feedback cycle proposed by Schunter et al. (2019). Previous studies have also noted opposite responses in gene expression across species of the same taxa (Kang et al., 2022; Strader et al., 2020), and even across populations of the same species (Goncalves et al., 2016), but the extent of the role that transgenerational effects play in creating divergent responses is still unclear (but see Goncalves et al., 2016; Schunter et al., 2018). Importantly, however, our seemingly species-specific results may indicate that *E. jacksoni* is preadapted to acidified conditions, whether through long-term local adaptation or transgenerational plasticity. Because of its limited adult dispersal and lack of a pelagic larval phase, *E. jacksoni* may be more likely to be genetically adapted to local conditions than other species (Warner, 1997), and its live-bearing reproduction may also facilitate adaptation through maternal effects. Kang et al. (2022) recently proposed a similar hypothesis to explain why *A. polyacanthus* (which also lacks a pelagic larval stage) differed from other co-occurring damselfish species in its molecular response to elevated pCO_2.

Interestingly, the response of GABA-related genes to acidification varied between Experiments 1 and 2 (which used different levels of acidification). In response to the more moderate static acidification in Experiment 2, *E. jacksoni* showed an upregulation of two subunits of the $GABA_A$ receptor (*gabra6* and *gabrb3*), which were instead downregulated in Experiment 1. Interestingly, the *gabra6* subunit is also upregulated in spiny damselfish transgenerationally exposed to

high *p*CO2 when compared to those that were only developmentally exposed, an effect opposite to that seen in the expression of other $GABA_A$ subunits in the same experiment (Schunter et al., 2018). No other GABA-related genes were significantly affected by this treatment. In Experiment 1, the greater magnitude change in pH resulted in an opposite and much broader response of GABA-related genes. These conflicting responses in the transcription of $GABA_A$ receptor subunits and other $GABA_A$-related genes indicate that in addition to varying across species, the response of the GABA signaling pathway to acidification/high pCO_2 may also depend on the magnitude of the environmental change. Further study is needed to determine how the divergent transcriptomic response of *E. jacksoni* seen in our experiments translates to behavior and overall fitness, how the magnitude of any emergent effects compare to those observed in other species, and the role transgenerational exposure plays in *E. jacksoni* response to acidification.

In both Experiment 1 and Experiment 2, an additional group of gene sets related to muscle tissue were identified as enriched in the acidified treatment. In Experiment 1, this included the upregulation of gene sets related to muscle development, contraction, and adaptation and muscle cell components, as well as the downregulation of the vascular smooth muscle contraction GO gene set. In Experiment 2, the A band GO gene set was upregulated, while the smooth muscle contraction gene set was downregulated. While smooth muscle is present in blood vessels in the brain, it is possible that the identification of some of these pathways (such as those related to striated muscle tissue) as enriched is in part due to the

misannotation of genes in this non-model species to orthologous reference genes. Alternatively, because our transcriptome was assembled using both brain and muscle tissue, it is possible that some brain transcripts were misaligned to muscle-exclusive reference transcripts during differential expression analysis.

Impacts of pH Variability

Overall, we found that variability in pH moderated the differential gene expression seen under static acidification. pH variability decreased the number of DEGs detected by the edgeR analysis between the pH 7.85 and pH 7.70 treatments in Experiment 2 (from 159 genes when treatments were static to 9 genes when both treatments were variable). This aligns with two previous studies that found that effects of pH on fish gene expression and behavior were diminished by the incorporation of diel pH fluctuations (Jarrold et al., 2017; Schunter et al., 2021).

Functional enrichment analysis revealed many up- and downregulated gene sets between static and variable treatments at each pH level, though there were more enriched gene sets in the more moderate pH 7.85 comparison (329 gene sets) than the pH 7.70 comparison (139 gene sets). This difference may represent an acidification threshold nearer to the 7.85 treatment, where the majority of transcriptional adaptation to acidification is activated. Such a threshold effect in gene expression patterns has also been observed in the gill tissue of spider crabs exposed to two levels of acidification (Harms et al., 2014), as well as in the muscle tissue of blue rockfish (*Sebastes mystinus*; S. L. Hamilton et al., 2017), and thresholds in OA response have

been noted across taxa (Bednaršek et al., 2021; Castillo et al., 2014; Wittmann & Pörtner, 2013). Interestingly, although 33 gene sets were commonly enriched across the pH 7.85 and pH 7.70 static-variable comparisons, the majority of them (30) were enriched in opposite directions depending on the pH level, with variability at pH 7.85 eliciting a directional response mirroring that of static acidification, and variability at pH 7.70 eliciting the opposite response (Figure 1.6, Table S1.16). For example, at the 7.85 pH level, variability led to a downregulation of gene sets related to morphogenesis, development, cell differentiation, exocytosis, cell-cell adhesion, molecular transducer activity, and leukocyte mediated immunity, while variability at pH 7.70 led to upregulation in these gene sets compared to the static treatment. These contrasting responses indicate that pH variability can have opposing effects on brain physiology depending on the underlying mean pH level. This interactive effect of acidification and variability may again reflect a threshold in the neural response of fish to acidification. It may be that at more moderate pH levels, variability exacerbates the negative effects of acidification by temporarily dropping the pH further below the average, but under more extreme acidification, perhaps past a biological tipping point, any negative effects of further acidification introduced by temporary oscillations may be outweighed by the temporary relief provided by the upswing of the oscillations above the mean pH.

It is important to note that our interpretation of these results could be limited by the scope of our experimental design. In Experiment 2, we sampled tissue from individuals in each treatment when the variable treatments were increasing in pH and

intersecting their corresponding static treatments. While this design keeps the pH at the time of sampling consistent between the static and variable treatments, it assesses expression at only a single time point, and therefore does not account for likely divergent expression patterns at different positions in the pH cycles of the variable treatments. Additional experiments are necessary to determine if and how gene expression differs in *E. jacksoni* depending on the trajectory and value of the pH at the time of sampling.

Inter-Individual Variability in Gene Expression

A particularly striking finding from our experiments is the observation that gene expression variability across individuals was greater in the variable pH treatments of Experiment 2 than in the static treatments (Figure 1.4). This pattern indicates that the environmental variability introduced by the pH oscillations may be revealing significant "cryptic variation" (Rutherford, 2000, 2003; Rutherford & Lindquist, 1998) in the transcriptomic response of *E. jacksoni* to acidification. In the context of climate change, such phenotypic variation, if beneficial and heritable, could represent potential adaptive variation on which selection may act, allowing populations to adapt to ongoing changes in environmental conditions (Rutherford & Lindquist, 1998; Rutherford, 2000, 2003; Queitsch et al., 2002; reviewed in Ghalambor et al., 2007).

Patterns of expression across individuals within static treatments, and across functionally related genes within individuals, were notably consistent. This

consistency provides evidence of a conserved stress response as described by Kültz (2005), and may again reflect a biochemical "switch" type response, activated at a certain environmental threshold. This idea is further supported in Experiment 2 by the similarity of expression profiles of some individuals in the pH 7.85 variable treatment to the expression profiles exhibited by those in the pH 7.85 static treatment, while others in the variable treatment exhibited expression profiles similar to those in the pH 7.70 static treatment (Figure S1.5, Figure 1.3).

Because we did not use the whole brain, and instead arbitrarily subsampled brain tissue from each individual, some of the inter-individual variability in expression profiles may be the result of variability in the exact section(s) of the brain that was sampled for each individual. Conversely, it is possible that this sampling method could introduce treatment-level bias in the brain region sampled that could lead to misleading signals of differential expression between treatments. However, expression profiles within static treatments were remarkably consistent across individuals, especially in Experiment 1 (Figure 1.2), indicating a low probability of sampling bias, and all individuals were subsampled in an arbitrary manner by a single researcher for each experiment. We therefore maintain that alternative sampling methods would have been unlikely to change the major patterns and conclusions presented here.

Conclusions

Overall, our results indicate that both acidification and pH/pCO_2 variability can have significant impacts on the brain gene expression of a nearshore temperate fish species. Given recent debate regarding the generality of neurological impacts of OA on marine fish (Clark et al., 2020a, 2020b; Munday et al., 2020), our study provides evidence of neurological impacts, even in a species with a high likelihood for local adaptation to naturally low pH/pCO_2. We found a significant effect of acidification on global gene expression in *E. jacksoni* brain tissue, and that the transcriptomic response was similar to a previous experiment that compared transgenerationally exposed tropical damselfish to individuals that were developmentally exposed (Schunter et al., 2018). These results suggest that the *E. jacksoni* in our experiments were exposed to ecologically relevant pH/pCO_2 variability *in situ*, which may have influenced their response to acidification in the lab. Additionally, our results demonstrate that the incorporation of upwelling-scale pH variability into acidification treatments has a substantial impact on the number of DE genes detected between moderate and low levels of acidification, indicating that temporal pH variability can moderate the impacts of acidification. Interestingly, we also found that the direction of the effect of variability on gene expression in certain genes depended on the degree of acidification. These opposing patterns of gene expression indicate that the impact of pH variability on fish brain physiology may be context-dependent, perhaps serving as an additional stressor at more moderate levels of acidification, but as an ameliorating factor when the mean pH is more extreme.

Finally, we observed significant variation in gene expression across individuals, and found that upwelling-scale pH variability revealed additional cryptic phenotypic variation. This finding indicates that studies employing only static treatments may underestimate standing genetic variation in traits related to the response of fish to acidification. This cryptic variation may provide additional genetic variation on which selection may act and therefore increase the likelihood of successful adaptation of fish populations to acidification. In summary, our results emphasize the importance of considering environmental variability in global change experiments and demonstrate that a species with an evolutionary history of exposure to acidified and variable conditions exhibits a distinctive transcriptomic response in gene sets similar to those affected in species that have shown behavioral impairment.

ACKNOWLEDGEMENTS

Work was funded by a fellowship from the David and Lucile Packard Foundation to KJK, as well as a Packard Endowment Award administered through UCSC to GB, KJK, CAL and YT, and a UCSC Committee on Research (COR) grant to GB. YT was also supported by the Packard Foundation. We would like to thank the many lab members, staff, and volunteers who helped carry out these experiments, including Sarah Lummis, Tye Kindinger, Emily Donham, Evan O'Brien, Hector Alvarado, Jim Freed, Melissa Gutterman, Savannah Mangold, Cassandra Powell, Nicole Lenoski, Jake Cline, Jacoby Baker, Eric Garcia, Nicholas Toy, Joseph Warren,

Nate Moore, and Randolph Skrovan. We also thank Malin Pinsky and Pete Raimondi for insightful comments, discussion, and assistance with statistical analyses.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article are available through online data repositories. Raw sequence data are deposited in the SRA (BioProject PRJNA757398). The annotated *E. jacksoni* transcriptome assembly (FASTA file) and environmental data from both experiments are available on Dryad (https://doi.org/10.7291/D19H5G). An alternatively trimmed version of the transcriptome assembly is also deposited in the TSA (BioProject PRJNA757398, accession GJIV00000000). The scaffold-level reference genome assembly for *E. jacksoni* is also available on GenBank (BioProject PRJNA810428, accession JALAZG000000000). Scripts for read processing and data analyses are available at https://github.com/jtoy7 .

AUTHOR CONTRIBUTIONS

JAT, KJK, GB, CAL, and YT conceptualized and designed the experiments. JAT and YT designed and tested the pH-manipulation system. GB carried out and led the sequencing project for Experiment 1. JAT carried out Experiment 2 and led the sequencing project with laboratory assistance from CAL. GCL led the reference genome sequencing project. JAT conducted all data processing and analyses with

input from KJK, GB, CAL, and YT. JAT wrote the paper. All authors reviewed and edited the drafts.

FIGURES

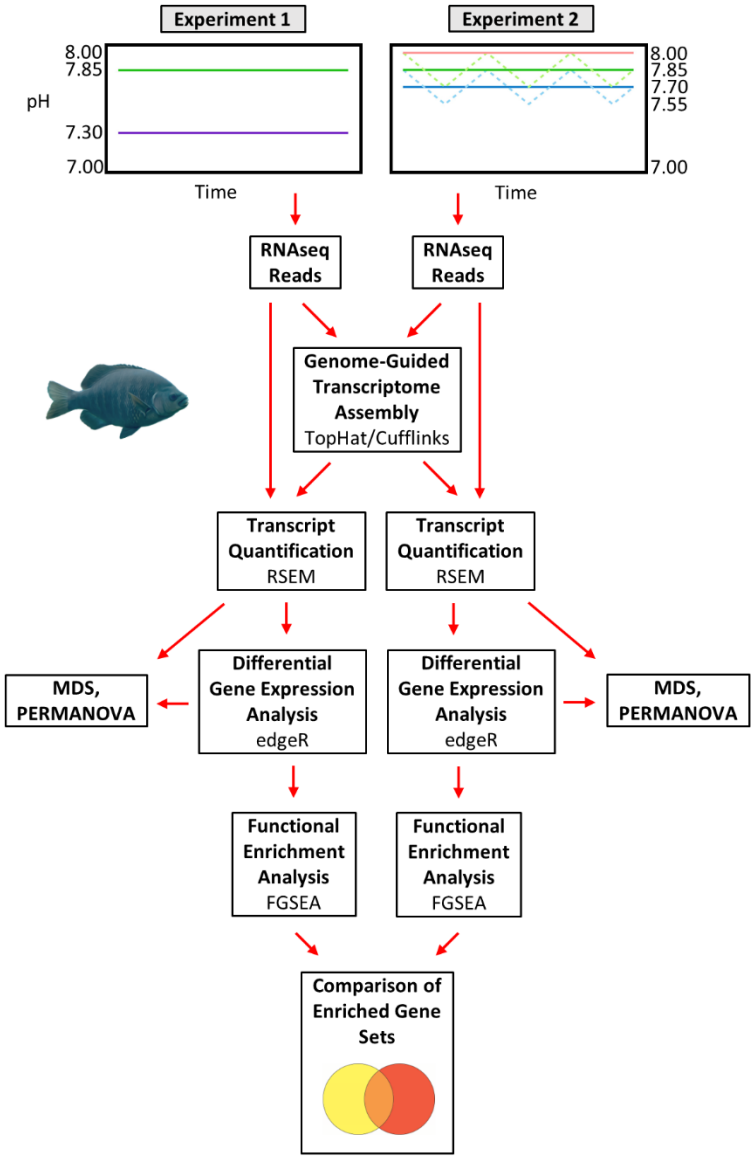


Figure 1.1 - Experiment design and data analysis pipeline for Experiments 1 and 2.

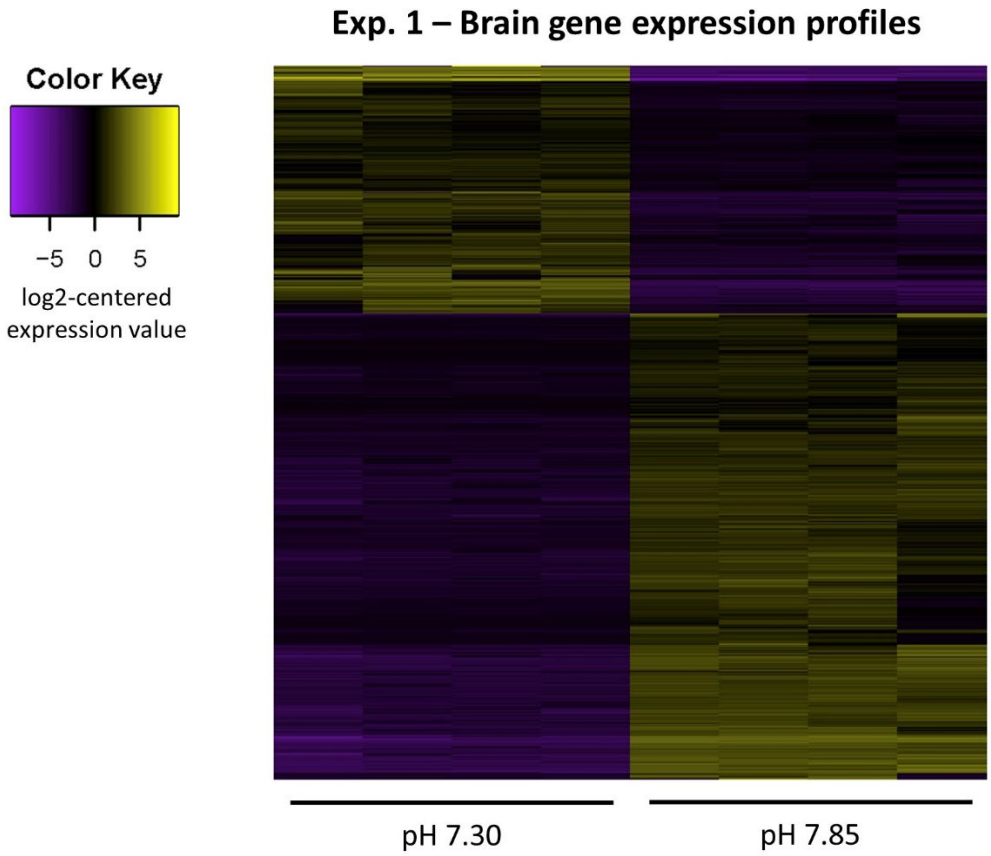


Figure 1.2 - Heatmap of gene expression profiles for each individual in Experiment 1. Each column represents an individual fish, and each row represents a differentially expressed gene. Each column represents an individual fish. Yellow colors represent upregulation in a given treatment and purple colors represent downregulation. Brighter hues represent larger differences in relative gene expression across the treatments.

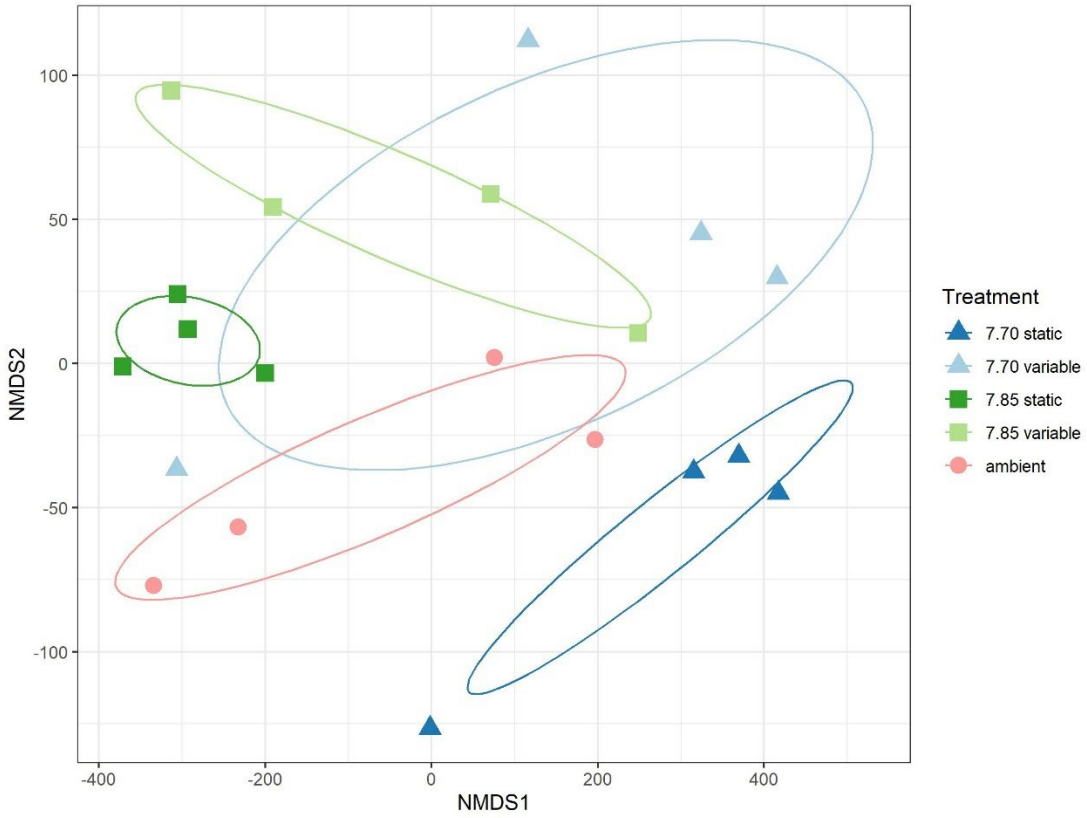


Figure 1.3 - nMDS plot of DEG expression in Experiment 2. Points represent single individuals. Ellipses are 95% confidence ellipses.

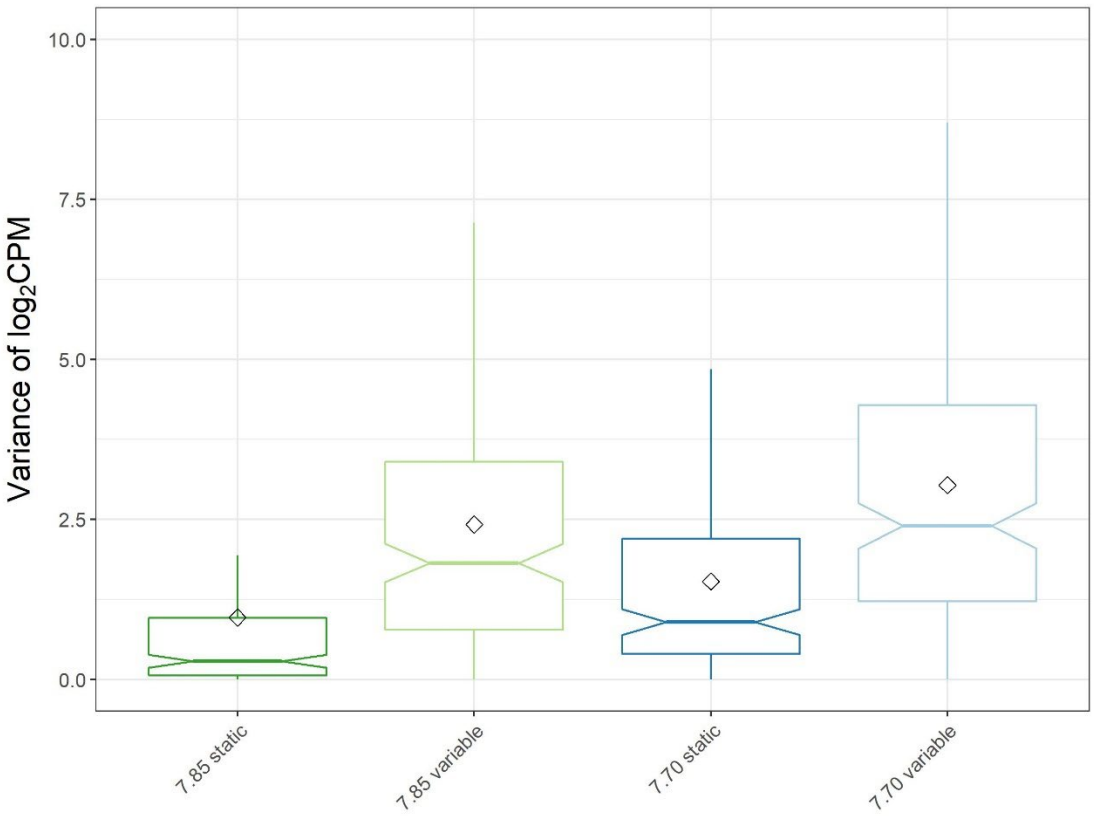


Figure 1.4 - Box plot of within-treatment variances in Experiment 2 (DEGs only, outliers removed for clarity). Diamonds mark the mean for each treatment. Notches represent a roughly 95% confidence interval around the median. Removed points lie outside of 1.5 times the IQR of each hinge.

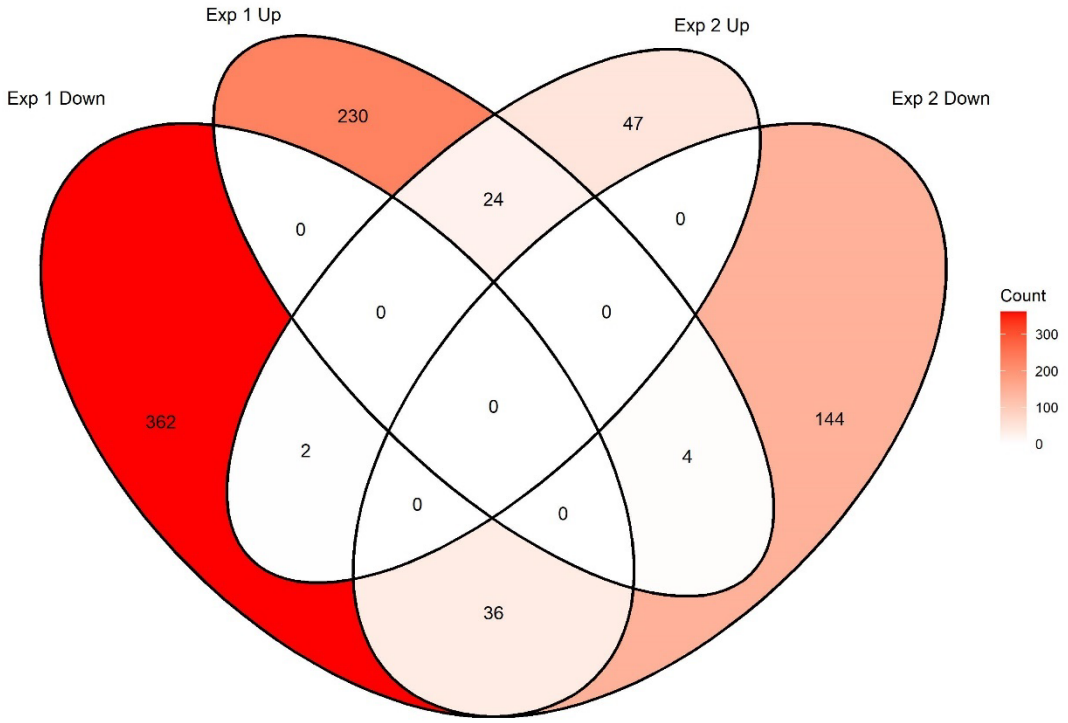


Figure 1.5 - Overlapping enriched gene sets across both experiments. “Up” and “Down” refer to gene sets that were upregulated or downregulated in the lower pH treatment relative to the higher pH treatment in each experiment (i.e., pH 7.85 treatments are treated as baseline in both cases). Only static treatments are included for Experiment 2.

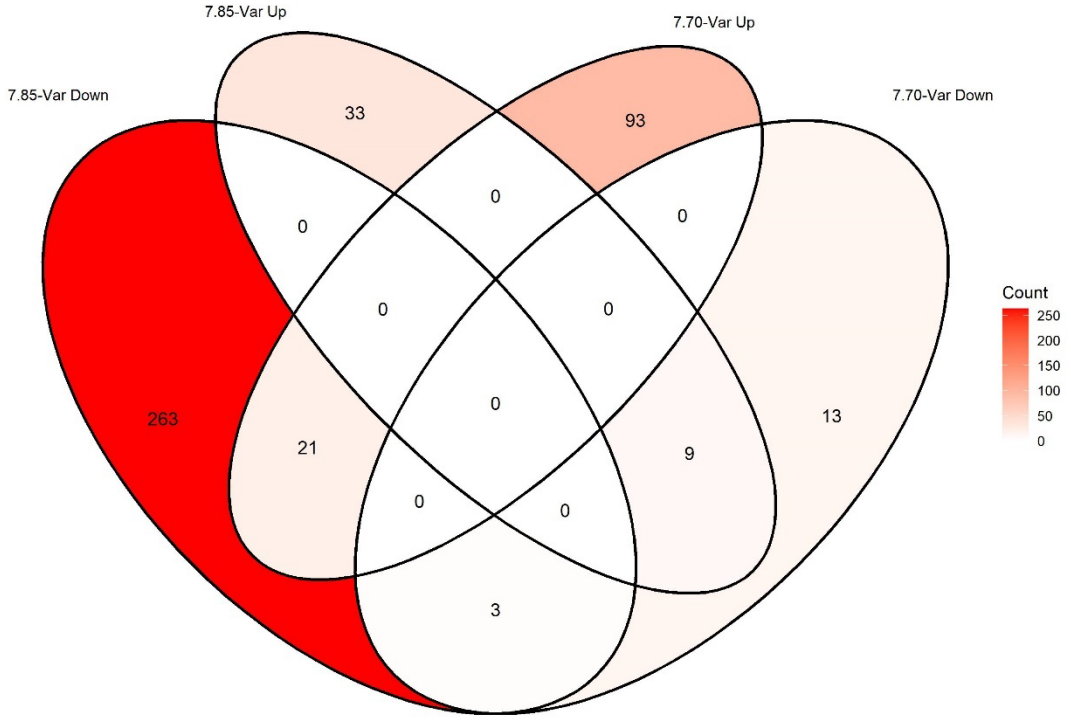


Figure 1.6 - Overlapping enriched gene sets across static vs. variable comparisons in Experiment 2. “Up” and “Down” refer to gene sets that were upregulated or downregulated in the “variable” treatment relative to the “static” treatment for a given pH level (7.85 or 7.70) (i.e., static treatments are treated as baseline).

TABLES

Table 1.1 - Carbonate chemistry and environmental parameters for treatment containers in Experiment 1. Aragonite saturation state (Ω) and pCO_2 were calculated with the R package seacarb (Gattuso et al. 2021) using the spectrophotometric pH and total alkalinity (TA) values from discrete bottle samples, and salinity and temperature values from YSI readings. All values are means ± SD. Mean pH_T (spec) and TA were calculated from bottle samples taken at 7 time points across the experiment. Mean pH_T (YSI) was calculated from daily readings that were calibrated using the discrete bottle samples.

Treatment	pH_T (spec)	pCO_2 (µatm)	Ω	TA (µmol/kg)	Temp (°C)	Salinity (ppt)	pH_T (YSI)
Target pH 7.85	7.88 ± 0.02	599 ± 36	1.50 ± 0.08	2193 ± 59	13.0 ± 0.5	33.8 ± 0.1	7.89 ± 0.04
Target pH 7.30	7.35 ± 0.06	2204 ± 333	0.48 ± 0.07	2212 ± 20	12.2 ± 0.8	33.8 ± 0.1	7.36 ± 0.14

Table 1.2 - Carbonate chemistry and environmental parameters for the headers of each treatment in Experiment 2. Aragonite saturation state (Ω) and pCO_2 were calculated with the R package seacarb (Gattuso et al. 2021) using the Durafet pH and temperature values, average TA values from discrete bottle samples, and salinity values from YSI readings. All values are means ± SD. Mean pH_T (Durafet) values were calculated using hourly averaged pH readings (from headers) that were calibrated using discrete (bottle) water samples and include only the time period of the first two pH cycles (September 1/2 - September 17/18). Mean pH_T (YSI) values were calculated using daily readings (from replicate containers) that were calibrated using bottle-calibrated Durafet values (taken simultaneously) from the September 1/2 - September 17/18 date range and include YSI readings from the entire length of the experiment (September 1/2 - September 23/24).

Treatment	pH_T (Durafet) (hourly)	pCO_2 (µatm)	Ω	TA (µmol/kg) (bottle samples)	Temp (°C) (Durafet)	Salinity (ppt) (YSI)	pH_T (YSI) (daily)
Ambient	8.00 ± 0.04	452 ± 55	2.36 ± 0.23	2266 ± 3	17.5 ± 0.06	34.3 ± 0.1	8.01 ± 0.08
Target pH 7.85 - Static	7.90 ± 0.01	586 ± 14	1.94 ± 0.07	2268 ± 3	17.6 ± 0.06	34.3 ± 0.1	7.90 ± 0.05
Target pH 7.85 - Variable	7.89 ± 0.08	614 ± 136	1.93 ± 0.34	2268 ± 4	17.5 ± 0.06	34.3 ± 0.1	7.88 ± 0.11
Target pH 7.70 - Static	7.76 ± 0.04	848 ± 64	1.46 ± 0.17	2268 ± 5	17.6 ± 0.06	34.3 ± 0.1	7.75 ± 0.08
Target pH 7.70 - Variable	7.76 ± 0.09	870 ± 199	1.47 ± 0.30	2267 ± 4	17.6 ± 0.06	34.3 ± 0.1	7.74 ± 0.10

Table 1.3 - Number of DEGs detected across all treatment comparisons in Experiment 2.

	Ambient	**pH 7.70 static**	**pH 7.70 variable**	**pH 7.85 static**
Ambient				
pH 7.70 static	5			
pH 7.70 variable	8	6		
pH 7.85 static	3	159	11	
pH 7.85 variable	6	11	9	6

Table 1.4 - Summary of Upregulated and Downregulated Gene Set Clusters in Experiment 1. Enriched gene sets (GO, KEGG, hallmark) were clustered by similarity using the AutoAnnotate and clusterMaker2 applications for the Cytoscape software platform. Clusters were then manually examined and named. See Table S1.12 for the full list of enriched gene sets in this experiment.

Upregulated in pH 7.30 Treatment		Downregulated in pH 7.30 Treatment	
Categorical Cluster	*Number of Gene Sets in each Cluster*	*Categorical Cluster*	*Number of Gene Sets in each Cluster*
mitochondrion, aerobic respiration, mRNA export from nucleus	44	transmembrane ion transport, regulation of synaptic signaling, ligand-gated ion channel activity, behavior, cognition & sensory perception	90
RNA metabolism, processing, splicing, modification, tRNA biosynthesis; ribosome biogenesis	41	regulation of nervous system development & growth	60
translation & protein localization	39	synaptic vesicle membrane, regulation of clathrin-dependent endocytosis	22
muscle development	22	axo-dendritic transport	20
organic acid catabolism	15	synaptic membrane & synapse	19
muscle contraction & adaptation, myogenesis	14	G protein-coupled receptor signaling	15
energy reserve & carbohydrate metabolic process	10	exocytosis and secretion	14
proteolysis, mRNA catabolism, negative regulation of cell cycle G2/M phase transition	10	central nervous system development	12
peroxisomal organization & transport, protein localization to organelle	8	regulation of pH & iron ion transport	9
innate immune response	6	aminoglycan & glycoprotein metabolic process	8
telomere maintenance via lengthening & organization	6	calcium-dependent phospholipid binding & cell-cell adhesion	8
RNA polymerase II	5	dopamine secretion & transport	7
protein modification by small protein conjugation or removal	3	axon, distal axon & terminal bouton	6
actin filament binding	2	dendritic tree & neuron spine	6
alpha actinin binding	2	GTPase activator activity	6
cytoplasmic stress granule	2	positive regulation of MAPK cascade	6
DNA polymerase activity	2	receptor localization to synapse	6
mitochondrial matrix & nucleoid	2	regulation of vesicle fusion	6
ribosome binding	2	dendrite membrane	5
RNA helicase activity	2	ephrin receptor signaling pathway	5
adipogenesis	1	extrinsic component of cytoplasmic side of plasma membrane	5
ADP binding	1	microtubule polymerization	5
allograft rejection	1	regulation of protein localization to membrane	5
androgen response	1	synaptic vesicle transport & localization	5
cell substrate junction	1	glycosphingolipid biosynthetic process	4
cysteine-type endopeptidase activity	1	cortical actin cytoskeleton	3
fatty acid metabolism	1	regulation of cell shape	3
ficolin-1-rich granule lumen	1	vascular transport	3

general transcription initiation factor binding	1	intrinsic component of Golgi membrane	2
interferon alpha response	1	long term depression & vascular smooth muscle contraction	2
lysine degradation	1	negative regulation of secretion & transport	2
MYC targets v1 (Hallmark)	1	neuron apoptotic process	2
MYC targets v2 (Hallmark)	1	regulation of amyloid precursor protein catabolic process	2
platelet morphogenesis	1	regulation of neurotransmitter receptor activity	2
positive regulation mitotic cell cycle	1	regulation of small GTPase-mediated signal transduction	2
receptor signaling pathway via STAT	1	response to catecholamine	2
rRNA binding	1	synaptic vesicle recycling	2
sarcolemma	1	vesicle docking	2
sarcoplasm	1	amyotrophic lateral sclerosis	1
starch & sucrose metabolism	1	anchored component of membrane	1
viral myocarditis	1	cyclic nucleotide-mediated signaling	1
		developmental maturation	1
		endocytosis	1
		gap junction	1
		genes upregulated by KRAS activation	1
		kinesin binding	1
		long-term potentiation	1
		neuron migration	1
		perinuclear region of cytoplasm	1
		phosphoprotein binding	1
		phosphoric diester hydrolase activity	1
		protein serine threonine kinase inhibitor activity	1
		regulation of neuron differentiation	1
		renal system process	1
		tau protein binding	1

Table 1.5 - Summary of Upregulated and Downregulated Gene Set Clusters in Experiment 2 (comparison of static treatments). Enriched gene sets (GO, KEGG, hallmark) were clustered by similarity using the AutoAnnotate and clusterMaker2 applications for the Cytoscape software platform. Clusters were then manually examined and named. See Table S1.13 for the full list of enriched gene sets in this experiment.

Upregulated in pH 7.70 Treatment		Downregulated in pH 7.70 Treatment	
Categorical Cluster	*# Gene Sets*	*Categorical Cluster*	*# Gene Sets*
RNA processing & splicing, histone methyltransferase complex	26	immune response	34
epigenetic regulation of gene expression & chromatin organization	8	lymphocyte proliferation, differentiation & activation	26
DNA repair, recombination & replication	7	secretory granule & myeloid leukocyte mediated immunity	13
mRNA export from nucleus	6	endothelial cell migration & blood vessel morphogenesis	10
E-box binding	4	JAK-STAT signaling pathway	9
ubiquitin-mediated proteolysis	4	neuropeptide/G protein-coupled receptor signaling pathway	7
ubiquitin ligase complex	4	cellular ion homeostasis	6
RNA phosphodiester bond hydrolysis	3	positive regulation of MAPK cascade	6
gene silencing	2	cell-cell junction assembly	5
nuclear speck	2	developmental growth involved in morphogenesis	5
A band	1	regulation of cytoskeleton & supramolecular fiber organization	5
cell cortex region	1	wound healing & regulation of body fluid levels	5
inositol phosphate-mediated signaling	1	leukocyte migration & regulation of chemotaxis	4
regulation of long-term synaptic potentiation	1	external side of plasma membrane	3
single-stranded RNA binding	1	leading edge membrane	3
structural constituent of cytoskeleton	1	plasma membrane signaling receptor complex	3
transcription coregulator activity	1	positive regulation of phagocytosis	3
		protein complex involved in cell adhesion & integrin-mediated signaling pathway	3
		regulation of cytokine production	3
		cilium movement & cell motility	2
		collagen-containing extracellular matrix	2
		endocytic vesicle	2
		positive regulation of cell-substrate adhesion	2
		receptor-mediated endocytosis	2
		regulation of peptidyl-tyrosine phosphorylation	2
		guanyl nucleotide binding	1
		calcium ion binding	1
		allograft rejection	1
		membrane microdomain	1
		complement system	1
		positive regulation of cell population proliferation	1
		smooth muscle contraction	1
		superoxide metabolic process	1
		response to organophosphorus	1
		Ras protein signal transduction	1
		response to dopamine	1

		inflammatory response	1
		odontogenesis	1
		coagulation	1
		leukocyte transendothelial migration	1
		pigment granule	1
		cell adhesion molecule binding	1
		ciliary plasm	1

Table 1.6 - Overlapping enriched gene sets between high pH vs. low pH comparisons in Experiment 1 and Experiment 2 (upregulated vs. downregulated gene sets in both experiments).

Upregulated in Both Experiments	Downregulated in Both Experiments
GOBP ribonucleoprotein complex biogenesis	GOBP MAPK cascade
GOBP ncRNA processing	GOCC side of membrane
GOCC ribonucleoprotein complex	GOBP receptor mediated endocytosis
GOMF catalytic activity acting on RNA	GOCC cell surface
GOBP translational termination	GOBP positive regulation of protein kinase activity
GOBP RNA export from nucleus	GOBP positive regulation of MAPK cascade
GOBP RNA processing	GOCC cell leading edge
GOBP RNA phosphodiester bond hydrolysis	GOBP cell-cell junction organization
GOBP mRNA export from nucleus	GOBP cell-cell adhesion
GOBP nuclear export	GOBP endocytosis
GOBP mRNA metabolic process	GOBP exocytosis
GOCC U2 type spliceosomal complex	GOBP cell-cell junction assembly
GOBP RNA 3'-end processing	GOBP cell growth
GOBP nucleic acid phosphodiester bond hydrolysis	GOBP taxis
KEGG spliceosome	GOCC secretory granule membrane
GOCC transferase complex	GOBP regulation of anatomical structure morphogenesis
GOBP protein modification by small protein conjugation	GOCC secretory vesicle
GOBP RNA localization	GOMF neuropeptide receptor activity
GOCC spliceosomal complex	GOBP cell junction assembly
GOBP protein modification by small protein conjugation or removal	KEGG cell adhesion molecules cams
GOBP RNA splicing	GOCC plasma membrane protein complex
GOBP mRNA processing	GOMF calcium ion binding
GOCC nuclear protein-containing complex	GOCC cell projection membrane
GOCC intracellular protein-containing complex	GOCC plasma membrane signaling receptor complex
	GOBP cell-cell adhesion via plasma membrane adhesion molecules
	GOBP developmental growth involved in morphogenesis
	GOBP developmental cell growth
	GOBP neuropeptide signaling pathway
	GOBP adenylate cyclase inhibiting G protein-coupled receptor signaling pathway

	GOCC leading edge membrane
	GOCC vesicle membrane
	GOMF G protein-coupled receptor activity
	GOCC receptor complex
	KEGG neuroactive ligand receptor interaction
	GOBP G protein-coupled receptor signaling pathway
	GOMF molecular transducer activity

Chapter 2: A high-quality reference genome of the kelp surfperch, *Brachyistius frenatus* (Embiotocidae), a wide-ranging Eastern Pacific reef fish with no pelagic larval stage.

This chapter has been submitted for publication in a peer reviewed journal and is reproduced here for inclusion in this dissertation. The citation for the submitted manuscript is:

Toy, J.A., Bernardi, G. A high-quality reference genome of the kelp surfperch, *Brachyistius frenatus* (Embiotocidae), a wide-ranging Eastern Pacific reef fish with no pelagic larval stage.

ABSTRACT

The surfperches (family Embiotocidae) are a unique group of mostly marine fishes whose phylogenetic position within the Ovalentaria clade (Percomorpha) is still unresolved. As a result of their viviparity and lack of a dispersive larval stage, surfperches are an excellent model for the study of speciation, gene flow, and local adaptation in the ocean. They are also the target of an immensely popular recreational fishery. Very few high-quality molecular resources, however, are available for this group and only for a single species. Here we describe a highly complete reference genome for the kelp surfperch, *Brachyistius frenatus*, assembled using a combination of short read (Illumina, ~47x coverage) and long read (Oxford Nanopore Technologies, ~27x coverage) sequencing. The 596 Mb assembly has a completeness level of 98.1% (BUSCO), an N50 of 2.6 Mb (n = 56), and an N90 of 406.6 Kb (n = 293). Comparative analysis revealed a high level of synteny between *B. frenatus* and its close relative, *Embiotoca jacksoni*. This assembly will serve as a valuable molecular resource upon which future evolutionary dynamics research will build, such as the investigation of local adaptation and the genomic potential for climate adaptation in wild populations.

INTRODUCTION

The surfperches (Embiotocidae) comprise a family of mostly marine, mostly Eastern Pacific fishes whose phylogenetic position within the Ovalentaria clade (a percomorph group that also includes the damselfishes, cichlids, mullets, and others) is

still unresolved (Ghezelayagh et al., 2022; G. Longo & Bernardi, 2015). Among marine fishes, the family of 23 species exhibits unusual life history traits. Surfperches undergo internal fertilization and viviparous gestation, and lack the dispersive pelagic larval stage exhibited by most other marine groups (Leis, 1991). Surfperches are also notable within the Ovalentaria because although they comprise relatively few species, they have invaded a large diversity of nearshore habitats, including sandy surf zones, rocky reefs, kelp forests, seagrass beds, estuaries, and even coastal freshwater habitats (Tarp, 1952), and can make up a large proportion of the fish biomass within these habitats in the Eastern Pacific (Laur & Ebeling, 1983). This unique ecology makes the surfperches an excellent group in which to study a range of ecological and evolutionary topics including speciation, adaptive radiation, life history evolution, and local adaptation.

The kelp surfperch, *Brachyistius frenatus*, is a smaller species (max TL: 21.6 cm) within the Embiotocinae subfamily with a highly kelp-associated ecology (Love, 2011). It is one of the widest ranging of the surfperches, occurring along the Pacific Coast of North America from at least Bahia Tortugas, Baja California Sur, MX to north of Sitka, Alaska, USA (personal observation; Love, 2011). This makes it an ideal species in which to study local adaptation and gene flow in coastal marine environments. Here we present a highly complete, highly contiguous de novo genome assembly for *B. frenatus* constructed with a combination of long read nanopore and short read shotgun sequencing data. In addition to its application to forthcoming continent-scale population genomics studies, this genome will serve as an important

resource for future studies of comparative genomics and evolutionary dynamics in the Eastern Pacific.

METHODS

DNA sampling and sequencing

We collected an adult male *B. frenatus* from Stillwater Cove in Pebble Beach, CA, USA (36.564821, -121.943556) on January 18, 2021 (CDFW permit S-193170005-19318-001) via spear and kept it on ice during transport back to the UC Santa Cruz Coastal Science Campus (Santa Cruz, CA, 95060). We sexed the specimen by dissection and sampled liver tissue for DNA extraction, after which the specimen was preserved in 95% ethanol. Genomic DNA was extracted from liver tissue using chloroform methods (adapted from Sambrook et al., 1989) and high molecular weight confirmed on a 0.7% agarose gel. Genomic DNA was then used to prepare libraries for both Nanopore long read sequencing and Illumina short read sequencing.

We prepared two libraries for two separate Nanopore sequencing runs using the SQK-LSK109 chemistry kit and protocol. Prior to preparation of the second library, we spun genomic DNA through a Covaris g-TUBE™ shearing tube to increase sequencer output. Libraries were sequenced on a Nanopore MinION device using two R9.4.1 flow cells. In total, we obtained 20.05 Gb of raw sequence data (6,782,325 raw reads).

Preparation of an Illumina sequencing library was done by Novogene Corporation Inc. (Sacramento, CA, USA) using fragmentation by sonication and the NEBNext® Ultra™ DNA Library Prep Kit for Illumina®. Library size distribution was evaluated on an Agilent 2100 Bioanalyzer (Agilent Technologies, CA, USA) and real-time PCR was used for quantification. Short read data was obtained in two sequencing runs on an Illumina NovaSeq 6000 (2 x 150 bp reads). In total, we obtained 28.7 Gb of raw sequence data (191,338,318 raw reads).

Read processing and genome assembly

Software versions used in the assembly are listed in Table 2.1. Base-calling of Nanopore reads was done using Oxford Nanopore Technologies' Guppy base caller software (v5.0.15) and the dna_r9.4.1_450bps_sup ("super accurate") model with default quality filtering parameters, resulting in only reads with >Q7 quality scores (nearly all reads >Q10). Sequencing adapters were then trimmed using Porechop (v2.4; https://github.com/rrwick/Porechop) and trimmed reads quality-checked with FastQC (v0.11.7; Andrews, 2010). Finally, reads less than 500 bp were removed. This filtered and trimmed dataset contained 4,966,516 reads with a mean length of 3,268.7 bp and an N50 of 4,886 bp. Additional Nanopore sequencing stats are listed in Table S2.1. Illumina reads were trimmed with Trimmomatic (v0.39; parameters: LEADING:2 TRAILING:2 MINLEN:25, Bolger et al., 2014) and quality-checked with FastQC.

The trimmed and filtered Nanopore reads were aligned into large contigs using the fuzzy-Bruijn graph-based assembler, wtdbg2 (v0.0, Ruan & Li, 2020), and twice-polished with the same reads using minimap2 (v2.17-r941, Li, 2018) for alignment and Racon (v1.4.13, Vaser et al., 2017) for consensus generation. The resulting assembly was then twice-polished with the trimmed Illumina short read data using BWA (v0.7.17-r1188; Li & Durbin, 2009) for alignment and Pilon (v1.23; Walker et al., 2014) for consensus generation. A blastn (v2.12.0) search of the polished assembly was then run against the nt database (NCBI) for use in contaminant detection via Blobtools2 (v3.0.0; Challis et al., 2020). In total, only five small contigs (less than 7,200 bp each; 27,511 bp total) were identified as contaminants (Proteobacteria) and removed from the assembly.

Following contamination removal, the mitochondrial genome was assembled and then mapped to the nuclear assembly to remove mitochondrial sequences. First, long and short reads were mapped to a reference mitogenome from the black surfperch, *Embiotoca jacksoni* (GenBank accession JAKOON010000230.1; Bernardi et al., 2022) using minimap2. The reads that successfully mapped were then imported into Geneious Prime (v2022.1.1, http://www.geneious.com/) for mitogenome assembly. In Geneious Prime, the long reads were mapped to the *E. jacksoni* reference mitogenome using the "Map to Reference" function and the default Geneious iterative mapper (Medium Sensitivity/Fast, default parameters). The generated consensus sequence was then used as a scaffold on which to map the paired short read data for polishing, again using the Geneious mapper. This consensus

sequence was then annotated using *E. jacksoni* reference annotations from GenBank accession NC_029362.1 (G. C. Longo et al., 2016) to identify potential spurious frameshift mutations and stop codons introduced by sequencing and/or assembly errors. We identified a total of three stop codons, one each in the coding sequences of the ND2, COI, and ND4 genes, which were all caused by incorrect calling of the length of mononucleotide repeats. After manual inspection of read data at each of these sites, an additional nucleotide was added to each repeat, which in each case resulted in the elimination of the stop codon through the re-shifting of the reading frame. A similar process was also used to change the only "N" in the sequence to an additional "C" at the end of a multi-C repeat (see supplementary methods). The final, corrected mitogenome assembly was then annotated using the MitoAnnotator pipeline (v3.74; Iwasaki et al., 2013).

In Geneious Prime, we ran a megablast search (NCBI, Altschul et al., 1990) with the completed mitogenome as query and the nuclear assembly as the database to identify regions of the nuclear assembly where mitochondrial sequence was incorrectly included. We filtered the list of contigs with BLAST hits by considering for removal only those contigs where BLAST hits to mitochondrial sequence made up >20% of the contig length. All other hits were assumed to be potential NUMTs (nuclear DNA segments of mitochondrial origin). We removed mitochondrial sequence from a total of four small contigs (largest: 11,036 bp). In each case, the validity of the remaining contig fragments was assessed by mapping long read data to each fragment. Fragments or portions of fragments with clear support from long read

data were kept as new contigs. In total, this process led to the complete removal of two contigs from the nuclear assembly, the splitting of one contig into two, and the trimming of another. For details, see supplementary methods.

Using RepeatMasker (v4.1.2-p1), we determined the repeat content of the final genome assembly by running a slow search (-s parameter) with the -species parameter set to “actinopterygii”.

Synteny Analysis

To compare the genome structure of *B. frenatus* to that of a close relative, an initial version of the assembly was aligned to an existing high-quality assembly for the genome of a close relative, the black surfperch, *Embiotoca jacksoni* (GCA_022577435.1; Bernardi and Toy et al. 2022), using minimap2 (H. Li, 2018) via the D-GENIES application (Cabanettes & Klopp, 2018). This mapping led to the identification of an apparent “translocation”, which upon further investigation, had little read support and was therefore determined to be a mis-assembly of two separate contigs. This contig was therefore split at a clear break in read support (see supplementary methods for details). The new assembly was again mapped to the *E. jacksoni* reference using D-GENIES.

An alternate version of the assembly was also created by further scaffolding the draft *B. frenatus* assembly with RagTag (v2.1.0; Alonge et al., 2021), using the *E. jacksoni* assembly as a reference. This scaffolding reduces the number of contigs from 1004 to 355.

RESULTS

A highly complete B. frenatus reference genome assembly

After contamination screening and mitogenome assembly, the final genome assembly has a size of 595,951,806 bp distributed across 1004 contigs (including the mitogenome). This indicates an average long read coverage of ~27x and an average short read coverage of ~47x. The largest contig is 12.3 Mb in length, and 50% of the genome is contained in 56 contigs that are ~2.6 Mb or greater in length (N50, Table 2.2). Ninety percent of the genome is contained in 293 contigs of length 406.6 Kb or greater (N90). The inner circle of Figure 2.1 shows the size distribution of all contigs in the genome. BUSCO analysis of the genome revealed a high level of completeness, with 98.1% of orthologs identified as present and complete. See Table 2.2 for additional assembly statistics.

The final mitochondrial assembly is 16,545 bp in length, 30 bp longer than the *E. jacksoni* reference used for assembly. Pairwise alignment of the *B. frenatus* mitogenome and reference sequence revealed a 90.2% pairwise identity between the two species. The base composition of the *B. frenatus* mitogenome assembly is A=28.0%, C=27.6%, G=16.3%, and T=28.1%. Annotation by MitoAnnotator identified 22 tRNAs, the 12S and 16S rRNAs, and 13 protein coding genes.

In total, RepeatMasker identified 50,258,669 bp of repeat sequence representing 8.43% of the genome. Retroelements accounted for 1.38% of the genome and DNA transposons 2.03%. Simple repeats were the largest major repeat

group, making up 4.34% of the genome, while low complexity regions, satellites, and small RNA (rRNA, snRNA, tRNA) accounted for 0.42%, 0.04%, and 0.08%, respectively (Table S2.2).

DISCUSSION

In this study, we present the second complete reference genome for the family Embiotocidae, providing new opportunities for comparative genomic and phylogenetic analyses within and outside of this group of uncertain placement. The assembly is of high quality, contiguity (contig N50 = 2.6 Mb), and completeness (98.1%) as a result of our combination of high accuracy shotgun and long read nanopore sequencing. With 90% of the genome contained within only 293 contiguous sequences of length 406.6 Kb or greater (N90), this assembly will serve as a reliable resource for comparative genomic studies and the estimation of population genomic parameters using whole genome resequencing data.

Overall, we found that the genome of *B. frenatus* is very similar to that of its close relative, *E. jacksoni.* The dot plot created with D-GENIES reveals a high level of synteny and sequence similarity between the two species, indicating speciation between the two occurred without a major structural rearrangement (Figure 2.2). The genome assemblies are also similar in size (596 Mb vs 635 Mb), repeat content (~8.4% of the genome for both species), and GC content (41.9% vs 41.6%), as would be expected for closely related species (Bernardi et al., 2022). Moving forward, this molecular resource will serve as the critical foundation for future resequencing

studies focusing on the characterization of genomic diversity and gene flow, as well as the evaluation of past and potential future climate adaptation, in reef fishes along the Pacific Coast. This work will ultimately inform management and conservation efforts, such as the classification of fisheries stocks and the evaluation of MPA efficacy.

ACKNOWLEDGEMENTS

We would like to thank Kristy Kroeker for support during the planning and execution of the project, and Remy Gatins and Merly Escalona for valuable assistance and insight during the assembly of the genome. We also thank Rion Parsons for technical assistance, Celine de Jong for support during field collection and dissection of the sample, and Dan Wright for assistance in the lab. Financial support for this project was provided by the Steven Berkeley Marine Conservation Fellowship (American Fisheries Society) awarded to JT.

DATA ACCESSIBILITY

Raw read data for NCBI BioProject PRJNA862534, BioSample SAMN29982938 are deposited in the NCBI Short Read Archive (SRA) under SRR21095639 and SRR21095640 for Nanopore data and SRR20680795 for short read Illumina data. The GenBank accession number for the full assembly (including the mitochondrial assembly) is JANHZZ000000000 and the annotated mitochondrial genome is accessioned as OP238470. Assembly scripts can be found at https://github.com/jtoy7 .

FIGURES

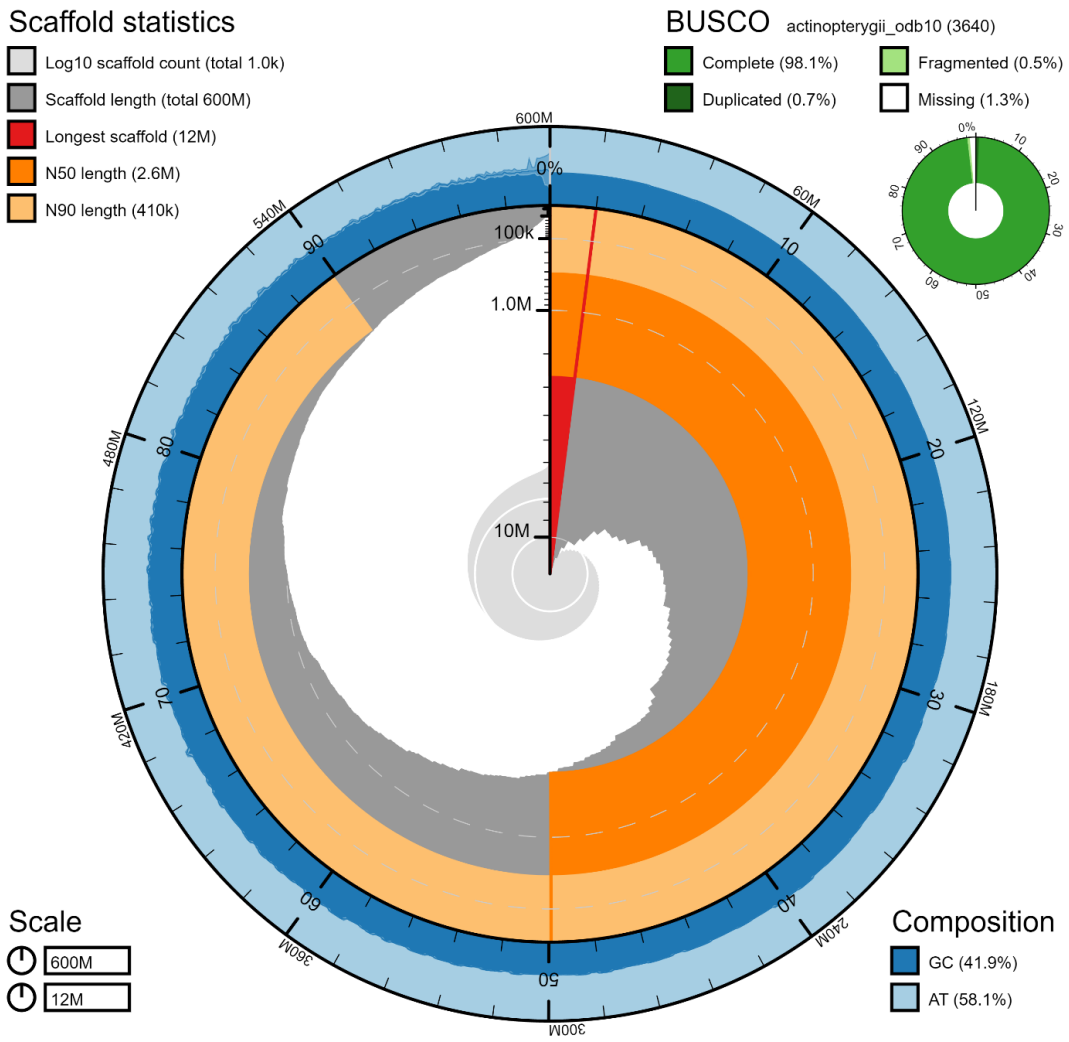


Figure 2.1 - Snail summary plot of the complete *B. frenatus* genome assembly produced using Blobtools2. The inner radial axis (gray) shows the length of each contig in descending order, with dark orange and light orange portions representing the N50 and N90 lengths, respectively. The dark/light blue ring shows the GC content along the length of the genome. The notched outer ring denotes the position (bp/proportion) along the genome. The inset at the top right shows the assembly completeness as assessed by BUSCO.

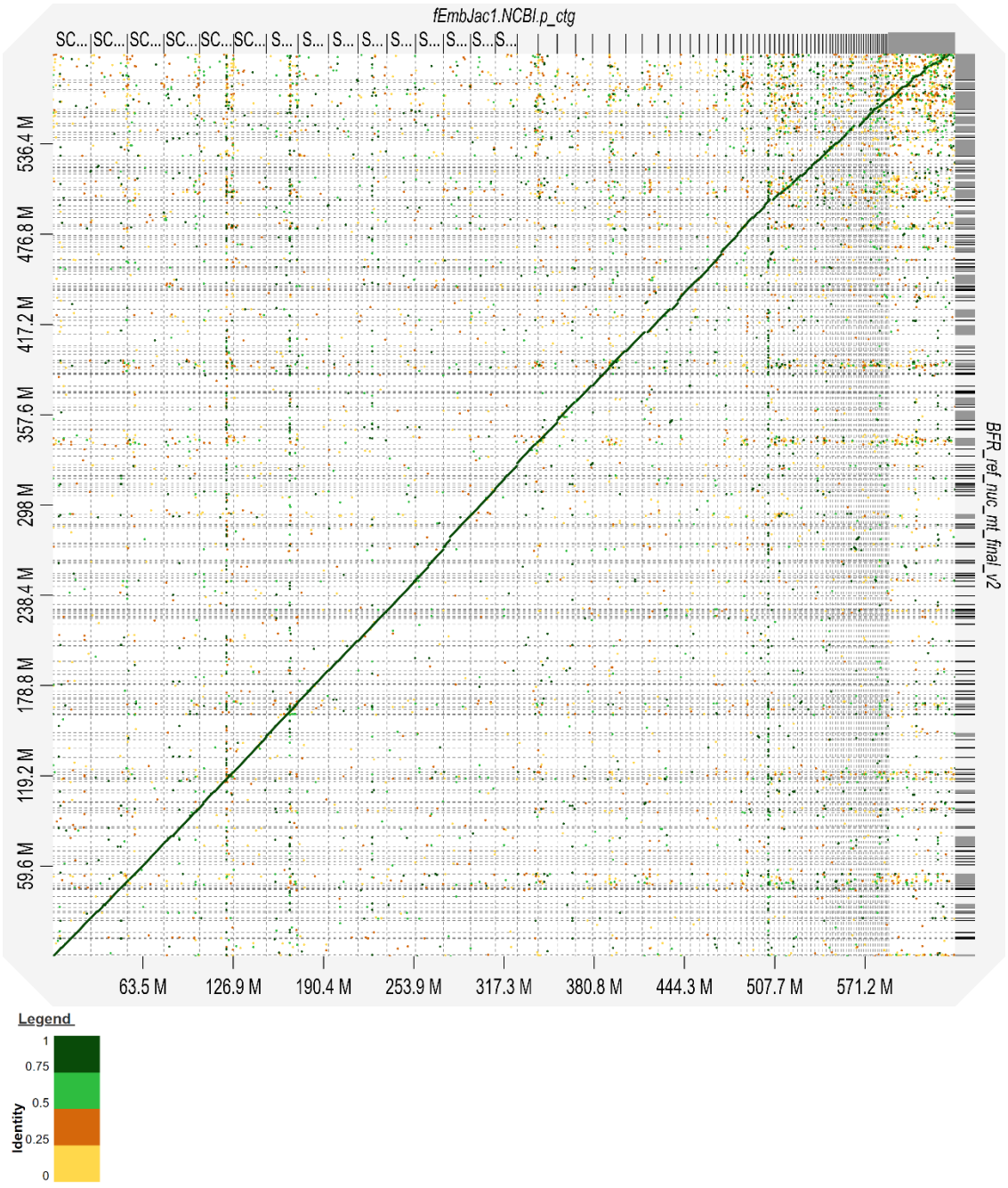


Figure 2.2 - Dot plot produced using D-GENIES of the alignment of the *B. frenatus* genome (vertical axis) to a reference *E. jacksoni* genome (horizontal axis; accession GCA_022577435.1). Dotted gridlines represent scaffold/contig boundaries. Darker/greener colors indicate greater similarity between the query and reference sequences. *E. jacksoni* scaffolds are ordered numerically in ascending order from left to right.

TABLES

Table 2.1 - Software names and versions used in the *de novo* genome assembly

Assembly Step	Software	Version
Nanopore base calling	Guppy	5.0.15
Nanopore adaptor trimming	Porechop	2.4
Illumina read trimming	Trimmomatic	0.39
De novo long read contig assembly	Wtdgb2	0.0
Long read contig polishing mapping	Minimap2	2.17-r941
Long read contig polishing consensus	Racon	1.4.13
Short read contig polishing mapping	BWA	0.7.17-r1188
Short read contig polishing consensus	Pilon	1.23
Contaminant detection	Blobtools2	3.0.0
Completeness evaluation	BUSCO	5.2.2
Repeat identification	RepeatMasker	4.1.2-p1
Filtering of mitochondrial reads	Minimap2	2.17-r941
Alignment of mitochondrial reads to reference	Geneious Prime	2022.1.1
Mitochondrial sequence consensus generation	Geneious Prime	2022.1.1
Mitochondrial genome annotation	MitoAnnotator	3.74

Table 2.2 - Assembly statistics and BUSCO completeness assessment for the *B. frenatus* genome

Assembly Statistics	
Assembly	nuclear + mitochondrial genome
Size (bp)	595,951,806
n scaffolds	1004
Average scaffold length	593,577.50
Largest scaffold	12,326,090
N50 (bp, n)	2,589,815 (56)
N60	1,896,895 (83)
N70	1,185,759 (123)
N80	758,445 (185)
N90	406,612 (293)
N100	597 (1004)
N count	0
Gaps	0
BUSCO Results	
Complete	98.1% (3572)
Complete and single-copy	97.4% (3547)
Complete and duplicated	0.7% (25)
Fragmented	0.5% (20)
Missing	1.4% (48)
Total BUSCOs searched	3640

Chapter 3: Range-wide whole genome resequencing reveals strong genetic structure and substantial adaptive variation to climate variables in the temperate reef fish, *Brachyistius frenatus*

ABSTRACT

Understanding the capacity of species to adapt genetically to a changing climate is a central question in global change biology that is still poorly understood. The characterization of the patterns of genetic diversity present within species and associations between this genetic variation and environmental variables can provide critical insight into the potential for evolutionary rescue. Population genetic structure, when combined with spatial variation in environmental conditions, can lead to local adaptation that maintains genetic variation within a species. Surfperches (family: Embiotocidae) are a unique group of mostly marine fishes in the temperate Pacific. Their viviparity and lack of pelagic larval stage reduces their dispersal and increases their likelihood of showing genetic structure across populations. Some species, like the kelp surfperch (*Brachyistius frenatus*), inhabit a latitudinal range that spans a large temperature gradient and a diverse mosaic of upwelling intensity. Here, I resequenced the genomes of kelp surfperch collected from Southeast Alaska to San Diego, California and, using medium-coverage whole genome resequencing techniques, estimated the levels of overall diversity and the extent of genetic structure among populations. I then looked for evidence of selection and local adaptation across the genome using two separate selection scan methods. Data provided evidence of very high levels of population structure along the latitudinal range, with individuals clustered into 5 major groups. The data also provide strong support for a pattern of isolation-by-distance among the California locations. Individuals collected from Alaska, British Columbia, and Puget Sound each had unique genetic

characteristics and low genetic diversity, suggesting a strong influence of founding effects in these regions. Selection scans revealed a large number of SNPs and genes under putative selection across all populations. Environmental association analysis confirmed the correlation of allele frequency differences in many of these candidate SNPs with environmental variables, most notably minimum sea surface temperature and mean annual pH. This study provides important insight into the patterns of adaptive genomic diversity present in temperate marine species, as well as the prevalence of local adaptation, both of which have implications for the resilience of marine species in the face of ongoing ocean change.

INTRODUCTION

In the face of rapidly changing environmental conditions, a species or population has three potential fates. A population may go extinct, shift its spatial distribution to fit that of its preferred climate envelope, and/or it may genetically adapt over generations to better survive and reproduce within its current range in a process often described as evolutionary rescue (Gomulkiewicz & Holt, 1995). While there is now a substantial body of research documenting climate-driven range shifts in terrestrial and marine systems, the potential role of adaptation in the persistence of species is still poorly understood. This is particularly true in marine species, where population dynamics are often more difficult to study (Selkoe et al., 2008).

Although genetic adaptation through de novo mutation is possible, rapid adaptation and evolutionary rescue on time scales relevant to modern conservation

concerns are generally thought to require standing genetic variation (Carlson et al., 2014). The level of genetic variation in a population is determined by a suite of evolutionary and ecological processes, including migration (gene flow), mutation rate, life history (e.g., mating patterns and reproductive strategies), spatial and temporal patterns of selection, and population size. A high level of genetic diversity increases the potential diversity in a population's phenotypic response to environmental change, which in turn increases the likelihood of evolutionary rescue (Carlson et al., 2014).

One potential source of genetic diversity for a species may be the local adaptation of subpopulations to their respective environmental conditions. Local adaptation in species that span a large range of environmental conditions may provide especially relevant genetic variation in the face of global change (Hofmann et al., 2014), potentially allowing species as a whole to be partially preadapted to future conditions. Local adaptation, however, may arise from a spectrum of evolutionary scenarios involving strong selection and/or limited gene flow (migration) (Sanford & Kelly, 2011). Local adaptation driven by extremely limited gene flow (e.g., low dispersal) may maintain adaptive genetic variation at the species level, but the restricted gene flow between subpopulations may limit its relevance to evolutionary rescue, and instead lead to subpopulations with relatively low effective genetic variation due to the diversity-depleting effects of selective sweeps. If, however, dispersal in a species is more moderate and selection relatively strong, local

adaptation may maintain high levels of adaptive diversity within a species without removing the possibility of adaptive gene flow between locations.

Until relatively recently, the extent of local adaptation in marine species was not well understood (compared to terrestrial or freshwater systems), in large part because of the traditional view that large dispersal distances and an open environment should result in low levels of local adaptation in the ocean (Palumbi, 1992, 2003). Numerous recent studies, however, have now documented evidence of local adaptation even within large, high-gene flow populations (Hofmann et al., 2014; Sanford & Kelly, 2011). Marine species with relatively low dispersal may therefore show even greater levels of local adaptation, further increasing the potential for species-level genotypic (and therefore phenotypic) diversity. Conversely, low dispersal may restrict the spread of adaptive alleles from one location to another, limiting the potential for evolutionary rescue in subpopulations with low immigration rates. In this study, I assess the effect of low dispersal on local adaptation along the west coast of North America using the kelp perch (*Brachyistius frenatus*), a viviparous, kelp-associated reef fish with no dispersive larval stage, as a model system. *B. frenatus* spans 29° of latitude from Baja California Sur, Mexico to Southeast Alaska, USA, a range that exhibits significant spatial heterogeneity in environmental conditions due to a combination of variation in upwelling occurrence and a latitudinal gradient in average sea-surface temperatures. These characteristics make *B. frenatus* in the NE Pacific an ideal system in which to study the capacity of

nearshore species to adapt to a changing coastal environment, and to gain insight into the prevailing drivers and mechanisms of selection within this range.

From Washington State, USA to Baja California Sur, MX, a major driver of environmental variation is a mosaic in the prevalence of coastal upwelling driven by northwesterly winds (Huyer, 1983). This translates into a mosaic of spatially variable temperature, pH/pCO_2, dissolved oxygen concentrations, and primary productivity as upwelling brings cold, acidic, deoxygenated, and nutrient-rich waters to the surface. Peaks in upwelling occur at various locations within this range, but the greatest levels occur in Northern California (between Eureka and Point Arena) and Southern Oregon (near Port Orford)(Chan, Barth, Blanchette, Byrne, Chavez, Cheriton, Feely, Friederich, Gaylord, Gouhier, Hacker, et al., 2017; Feely et al., 2008; Kroeker et al., n.d.). South of Point Conception the prevalence of upwelling decreases significantly, but increases again near Ensenada, MX (Huyer, 1983). North of Vancouver Island, BC, oceanographic conditions are driven instead by the Alaska Current and predominantly downwelling winds, and coastal ecosystems exhibit highly seasonal dynamics in abiotic and biotic conditions (Kroeker et al., 2021). This variation in upwelling, combined with a general latitudinal gradient in sea surface temperature (SST), result in a wide range of environmental conditions within the geographic range of *B. frenatus.* In Sitka Sound, AK, for example, subtidal winter temperatures are as low as 5 °C and summer pH can reach 8.6 (Kroeker et al., 2021), while summer temperatures in Southern California can exceed 24 °C and spring pH in Northern California can drop below 7.5 (Donham et al., n.d.).

Given these highly divergent environments and the limited dispersal of this species, one may expect substantial divergence between subpopulations in coding DNA sequences (as a consequence of selection on the thermal stability of proteins), as well as signals of selection in genes related to metabolism. Differences in temperature range/variability between regions may also lead to selection on genes related to transcript/protein turnover. Moreover, divergent pH regimes may lead one to anticipate differences in genes related to acid/base balance, homeostasis, and cell signaling (Toy et al., 2022). In this study, I use medium-coverage whole genome resequencing to compare the genomic diversity of *B. frenatus* subpopulations across this range. I assess the level of population genetic structure between regions, identify genomic signals of selection between subpopulations, and test for significant associations between genetic differences and specific environmental variables to better understand the mechanisms of climate adaptation in coastal populations.

METHODS

Sample collection

From 2018 to 2021, I collected *B. frenatus* from 13 sites along the coast of North America from Southeast Alaska to San Diego, California, spanning >24° of latitude (Table 3.1, Figure 3.1). Due to a combination of the inaccessibility of dive sites, poor diving conditions, and limited (or non-existent) sightings of *B. frenatus* on the outer coasts of Oregon and Washington, I was unable to sample this region (Figure 3.1). I collected 15 individuals from each location, except for Malcom Island, BC, where I

collected 16 because one individual had an anomalous appearance, and Tomales Bay, CA, Santa Cruz, CA, and Point Buchon, CA, where my sampling efforts yielded only 2, 2, and 3 individuals, respectively. Specimens were collected by spear and either immediately dissected or stored on ice until dissection. Gill tissue was dissected from each individual and stored in 100% ethanol in 2 mL screw-cap tubes. Preserved samples were kept as cool as possible (room temperature, on ice, in a refrigerator) until they could be stored at -20°C in the laboratory.

DNA library preparation, sequencing, and read processing

I extracted genomic DNA from all samples using a chloroform-based protocol adapted from Sambrook et al. (1989) and confirmed DNA quality on a 0.7% agarose gel. Genomic DNA was then used to prepare sequencing libraries using half reactions of the NEBNext® Ultra™ II FS DNA Library Prep Kit (New England Biolabs). Prepared libraries were quantified via Qubit, fragment analyzed, pooled, and then sequenced (150 bp paired end) on an Illumina NovaSeq 6000 (S4) at the Vincent J. Coates Genomics Sequencing Lab at the University of California, Berkeley. Sequences were adapter- and quality-trimmed using fastp v0.23.2 (Chen et al., 2018) with default parameters and a minimum length cutoff of 40 (-l parameter). Read quality before and after trimming was visually inspected using fastp, FASTQC v0.11.7 (Andrews, 2010), and multiQC v1.6 (Ewels et al., 2016). I mapped the trimmed reads to a *B. frenatus* reference genome (GenBank accession JANHZZ000000000)(Toy & Bernardi, n.d.) using Bowtie2 v2.4.1 (Langmead &

Salzberg, 2012). I then removed duplicates from the aligned reads on a per sample basis using the Picard (v2.27.1) MarkDuplicates function ("Picard Toolkit," 2019) and clipped overlapping read pairs using the clipOverlap function of BamUtil (v1.0.15)(Jun et al., 2015). Finally, reads were realigned around indels using the IndelRealigner tool from GATK (v3.8)(McKenna et al., 2010). Sequencing depth statistics were calculated for each sample using the Samtools v1.14 depth function (Danecek et al., 2021).

Reference genome annotation

To enable functional enrichment analysis, I annotated the *B. frenatus* reference genome using the AUGUSTUS v3.4.0 gene predictor (Stanke et al., 2008) with default settings (–species = zebrafish). I then identified predicted genes by running a blastp (Altschul et al., 1990) search of amino acid sequences against UniProt's SwissProt database (The UniProt Consortium, 2021) with the e-value parameter set to 1e-5. Genes with multiple hits were thinned to keep only the hit with the lowest e-value. UniProt accession numbers were converted to gene names using UniProt's ID mapping tool (https://www.uniprot.org/id-mapping).

Genotype and SNP calling

Using the ANGSD (v0.937)(Korneliussen et al., 2014) software package, I calculated genotype likelihoods across the genome (-GL 1), using only uniquely mapping paired reads (-uniqueOnly 1 -only_proper_pairs 1) with a mapping quality score of 20 (-

minMapQ 20). Sites were included only if data was present for at least 79 individuals (-minInd 79) at a given site and total depth across all individuals was greater than 316 (-setMinDepth) and less than 2294 (-setMaxDepth). Allele frequencies (-doMaf 1) and the determination of major and minor alleles were called using genotype likelihoods (-doMajorMinor 1).

Single nucleotide polymorphisms (SNPs) were called using the calculated allele frequencies with a minimum minor frequency filter of 0.01 (-minMaf 0.01) and a p-value cutoff (minor allele frequency significantly different from 0) of 1e-6 (-SNP_pval 1e-6). I then estimated linkage disequilibrium (LD) between SNPs using genotype likelihoods and ngsLD v1.1.1 (Fox et al., 2019) and used the prune_ngsLD.py script included with ngsLD to create an LD-pruned SNP list for downstream analysis (--max_dist = 10000, --min_weight = 0.5).

Principal Components Analysis & Genetic Differentiation

To assess genetic variation within and between sampling locations, I calculated from the genotype likelihoods a covariance matrix of allele frequencies across individuals using PCAngsd (Meisner & Albrechtsen, 2018) with default settings (-minMaf 0.05). Using the eigen function in R v4.1.0 (R Core Team, 2021), I then performed an individual-level principal components analysis (PCA). I simultaneously estimated admixture proportions for all individuals using the –admix option to visualize genetic ancestry. At this point, the anomalous individual from Malcom Island stood out as an outlier from the PCA grouping and ancestry of its sampling location. Given its

anomalous appearance, it was removed from the dataset for all further analyses, and the PCA and admixture analyses were rerun without it. The general results of these analyses were unaffected by this removal.

I removed low sample size populations (Tomales Bay, Santa Cruz, and Point Buchon) from the dataset, as further analyses are sensitive to low sample sizes. I then calculated sample allele frequency likelihoods at each genomic site for each population using the -doSaf option in ANGSD with -setMaxDepth set to 218, which were used to generate a 2D-site frequency spectrum (SFS) for each pairwise comparison of the 10 remaining sampling locations. To estimate genetic differentiation between sampling locations, the 2D-SFS were then used as input into the realSFS program in ANGSD to calculate site-level pairwise weighted F_{ST} statistics for each comparison. Site-level F_{ST} values were summarized at the contig level, normalized by contig length, and then averaged across contigs to calculate mean genome-wide F_{ST} values for each comparison. Finally, to visualize the extent of isolation by distance (IBD) exhibited by *B. frenatus*, linearized weighted F_{ST} for each comparison, as defined by Rousset (1997), was plotted against the straight-line geographic distance (great circle distance) between sampling locations. Distances were calculated using the distHaversine function of the geosphere R package (Hijmans, 2021).

Identifying Genes Under Selection

During the above run of PCAngsd, I also performed a PC-based scan for SNPs under selection using the –selection option (Meisner et al., 2021). This analysis is an extension of the method implemented in FastPCA (Galinsky et al., 2016) that detects unusual allele frequency differences along the principal components (PCs) inferred by PCAngsd. It produces a list of outlier SNPs for each PC (SNPs with differentiation along a given PC that is significantly greater than the null distribution under genetic drift).

In addition to the PC-based scan for SNPs under selection described above, I also identified candidate SNPs using an outlier scan method implemented in the "core model" of BayPass v2.31 (Gautier, 2015), using allele counts calculated from the allele frequencies estimated in ANGSD. The core model calculates the XtX statistic introduced by Günther and Coop (2013), which is an F_{ST}-like statistic of genetic differentiation that explicitly accounts for population structure and shared demographic history. For this model, I calibrated the XtX threshold used to determine candidate SNPs by creating pseudo-observed datasets (PODs) from the original data under the null model of no selection. I used the value of XtX corresponding to the 99^{th} percentile of the POD null distribution as the calibrated selection/neutrality threshold.

I compared the list of candidate SNPs from the XtX analysis to the combined list derived from the PC-based selection scan described above and created a final list of

SNPs putatively under selection by keeping only those identified in both analyses. To generate a list of genes putatively under selection, the final list of SNPs was mapped to genes in the reference genome annotation file and then thinned so that SNPs were only kept if they were inside or up to 1000 bp upstream of a predicted and annotated gene. I then performed a gene set enrichment analysis with g:profiler (Raudvere et al., 2019), querying the list of annotated candidate genes against a background annotation from the spiny damselfish (*Acanthochromis polyacanthus*; Ensembl ID ASM210954v1) (Cunningham et al., 2022). Gene sets tested included those from the Gene Ontology (GO) Biological Process (BP), Molecular Function (MF), and Cellular Component (CC) databases (The Gene Ontology Consortium, 2020). GO gene sets were filtered to include only those containing between 3 and 500 genes and a false discovery rate (FDR) of 0.05 was used to determine significance. To elucidate biological trends, the enriched gene sets were clustered by similarity using AutoAnnotate and clusterMaker2 applications for the Cytoscape platform (Reimand et al., 2019). Clusters were manually summarized and renamed using QuickGO (Binns et al., 2009) to reference the hierarchical relationship between clustered GO terms. To determine if there were different biological themes underlying the outlier SNPs along each PC in the PCAngsd analysis, I also repeated this process separately for each PC, using only those SNPs which were also identified as outliers by the XtX analysis.

Environmental Association Analysis

To better understand the environmental factors underlying the signatures of selection seen in the selection scan analyses, I performed a genotype-environment association (GEA) analysis using BayPass's Markov chain Monte Carlo (MCMC)-based "auxiliary model". This model identifies SNPs with allele frequencies that are significantly associated with environmental variables while accounting for the covariance between sampled locations that results from shared history. The environmental variables used in this analysis were annual mean sea surface temperature (SST), SST range, maximum SST, minimum SST, mean dissolved oxygen concentration, and mean pH. Environmental data corresponding to the GPS coordinates of each sampling location (Table 3.2) was extracted from the Bio-ORACLE climatology dataset (Assis et al., 2018) using the sdmpredictors package in R (Bosch & Fernandez, 2022). Allele frequencies were tested against each variable separately, with each analysis consisting of a burn-in of 5,000 MCMC iterations and 25,000 post-burn-in iterations, sampled every 25 iterations. I considered SNPs with Bayes factors (BF) > 20 deciban as significantly associated with the tested environmental variable, as recommended by Gautier (2015).

Estimates of Diversity

Finally, to estimate and compare overall levels of nucleotide diversity within each population, a separate list of genomic sites was created using sequence data from all individuals from the original 13 locations in ANGSD with no SNP significance

threshold (including both variable and non-variable sites) and no minimum allele frequency threshold. The parameters used were -minInd 79, -setMinDepth 316, and -setMaxDepth 2294. This generated a list of 574,065,898 sites, representing 96.9% of all sites in the reference genome. Using the limited list of samples from the 10 well-sampled locations mentioned above, I again calculated sample allele frequency likelihoods, this time to produce a single dimensional SFS for each population. Both the sample allele frequency likelihoods and SFS for each population were then used as input for the ThetaStat program in ANGSD to calculate the estimates of genetic diversity, Watterson's estimator (θ_W) and Tajima's theta (θ_π).

RESULTS

Sequencing, SNP calling & LD-pruning

Across all 158 samples sequenced, the median of the median genome-wide depth was 8x with a standard deviation of 6.52 after deduplicating, overlap clipping, and realignment of reads around indels (mean of means = 10.22x, SD = 6.26). The mean proportion of the 596 Mb reference sequence that was covered by at least one read was 97.4%. Additional information on sequencing coverage and depth is provided in the supplementary materials (Figures S3.1 & S3.2). Variant calling in ANGSD resulted in 9,020,431 SNPs that met the filtering criteria. Among these, the mean depth summed across all individuals was 1200x with a mode of 1500x, corresponding to an average of 7.6x per individual. LD-pruning with ngsLD yielded a final list of 6,711,262 unlinked SNPs.

Population genomic analyses

The principal components analysis resulted in 4 significant axes of variation (PCs, MAP test), explaining 11.4%, 6.7%, 3.4%, and 2.4% of the genetic variance, respectively. Overall, I observed strong structure between geographic regions. At the continental scale, individuals clustered into 5 major groups composed of the Southeast Alaska, British Colombia, Puget Sound, Northern California, and Southern California sampling regions, although most sampling locations within the Northern and Southern California regions also exhibit separation from each other along the major axes. In general, the placement of populations along PC 1 correlated with latitude, although the relative distances between locations along this axis did not always reflect relative geographic distances (Figure 3.3). The 5 major clusters were substantially differentiated along PC 1, except for British Columbia and Puget Sound, which held similar positions along this axis. Along PC 2, however, Puget Sound was highly differentiated from all other clusters, and British Columbia instead grouped together with the northern California and, to a lesser extent, southern California clusters. PC 3 mostly separated British Columbia from all other locations, although the other major clusters remained distinct (Figure 3.4). The Puget Sound and Alaska clusters grouped closely along PC 3 and PC 4. With respect to the Northern and Southern California locations, placement along PC 4 also strongly reflected latitude, but with greater separation among sampling locations, relative to their distances from the other three clusters, than was seen along PC 1. Along PC 4, the Alaska and Puget Sound clusters grouped closely with the sampling locations from Central California

(Santa Cruz, Pacific Grove, Carmel Bay, Big Sur), while British Columbia grouped closely with the Santa Barbara location. Most sampling locations exhibited relatively low levels of within-location variance, with most forming tight clusters along each set of PCs. One individual from Santa Barbara, however, broke with this pattern, grouping more closely to the Northern California cluster along PCs 1 and 2 than to the Southern California cluster.

As expected, pairwise weighted F_{ST} calculations confirmed the strong structure visualized in the PCA. Puget Sound exhibited the strongest differentiation of all the sampling locations, with all pairwise comparisons including this location resulting in F_{ST} values greater than 0.26 (Table 3.3). Surprisingly, the greatest differentiation seen across all comparisons was between Puget Sound and Sitka Sound (F_{ST} = 0.40), locations separated by only 1365 km. By contrast, San Diego and Salisbury Sound are separated by over 3000 km, but the divergence between them was 0.25. The lowest level of differentiation was seen between the two Southeast Alaska locations (Sitka Sound and Salisbury Sound), which also represent one of the smallest geographic distances between sampling locations (43 km).

To understand the extent to which population structuring might be driven by geographic distance, i.e., isolation by distance (IBD), and to highlight deviations from expected IBD patterns, I plotted linearized weighted F_{ST} for each comparison as a function of the shortest geographic distance between sampling locations (Figure 3.5). This resulted in a highly linear relationship for the majority of comparisons (r^2 = 0.96,

$p < 2.2 \times 10^{-16}$), with those involving Puget Sound, as well as comparisons between British Columbia and Alaska, deviating strongly from this pattern with comparatively greater ratios between F_{ST} and geographic distance.

PCAngsd infers admixture proportions by assigning the number of ancestral populations (K) as K = D + 1, where D is the number of significant principal components (Meisner & Albrechtsen, 2018). K was therefore set to 5 populations, which correspond to the 5 major clusters described above. The analysis revealed little admixture between the four northernmost populations (Alaska, British Columbia, and Puget Sound, and Northern California). Within the Southern California population, individuals from San Diego had little to no ancestry derived from the other four populations, but there appears to be a trend of increasing Northern California ancestry moving north from Santa Barbara (~10% admixture, n = 15) to Santa Cruz (~30% admixture, n = 2). Patterns in ancestry were generally consistent within sampling locations. Two individuals deviated notably from this general pattern. In addition to Northern California ancestry, an individual from Santa Barbara also had ancestry corresponding to the Puget Sound and British Columbia populations, up to a total of approximately 40% admixture. Another individual from Pacific Grove had a roughly twice the proportion of the Northern California ancestry (~40%) than the other individuals from that location (~20%).

Estimates of Diversity

Estimates of the site frequency spectra revealed substantial differences in spectrum shape between regions. The two Alaska locations, British Columbia, and Puget Sound all showed a relative lack of low frequency alleles, while the California locations exhibited a relative excess of these alleles (Figure 3.7). In general, there was an increasing trend of low frequency alleles moving south from Bodega Bay to San Diego.

Estimates of mean nucleotide diversity (θ_π) revealed substantial differences between sampling locations. In general, diversity increased from the south to the north (Table 3.4). The highest estimated diversity was observed in the San Diego and Carmel Bay locations (0.00260), though estimates for the other Southern California locations were only marginally smaller (> 0.00257). Bodega Bay, which was the only Northern California location included in this analysis, had an estimated nucleotide diversity of 0.00234. The lowest diversity estimates were observed in the four northernmost locations. Puget Sound had an estimated diversity of 0.00114 and the two Alaska populations were only slightly more diverse (Sitka Sound = 0.00125, Salisbury Sound = 0.00135). Interestingly, British Columbia had the highest diversity of the non-California locations (0.00176). Estimates of Watterson's θ (θ_W) were also calculated for each population and yielded similar general patterns (Table 3.4).

Evidence of Selection

The PC-based test for selection did not identify any outlier SNPs along PC 1, but found 10,247 outliers along PC 2, 16,798 along PC 3, and 7,528 along PC 4 (chi-squared test, FDR-adjusted p-value < 0.05). The four lists of outlier SNPs were combined (34,566 unique SNPs) and compared with the list of 120,339 outliers derived from the BayPass core model analysis (XtX > 99% quantile of the POD distribution). Only those SNPs present in both outlier lists were considered to be under selection, yielding a final list of 12,670 candidate SNPs. These SNPs mapped to a total of 6,473 genes on the annotated reference genome.

Enrichment analysis of the candidate gene list yielded 160 enriched gene sets (GO terms) with FDR-adjusted p-value < 0.05. Of the 6,473 genes queried, 4,006 were included in the analysis, 1,320 were duplicates, and 1,147 had gene names that were not recognized against the *A. polyacanthus* background. The enriched GO terms were clustered by similarity to form 23 functional gene set clusters (Table 3.5). In general, the two major biological themes among the enriched terms were embryonic development and metabolic processes. The three largest gene set clusters were associated with tissue, organ, and skeletal system development (30 gene sets), axonogenesis and neuron development (25 gene sets), and vasculature development and ameboidal-type cell migration (25 gene sets).

The PC-level enrichment analysis resulted in similar biological themes across PCs with substantial overlap of enriched terms among them. There were, however,

some differences in the lower-level, more specific GO terms (Table 3.6). For example, both PC 2 (which mainly separates Puget Sound from all other populations) and PC 3 (which does the same for British Columbia) have clusters related to development/morphogenesis, but PC 3 includes specific systems that aren't enriched along PC 2, such as vasculature development and sensory organ development. PC 4 (which mostly separates out California populations by latitude) gene set clusters are also mainly related to development and cell signaling, with a specific focus on eye, blood vessel, and nephron development.

GEA analysis yielded associated candidate SNPs (BF > 20) for all 6 environmental variables tested, which were filtered to include only those that were also identified in the XtX and PC-based analyses. The greatest number of filtered SNPs were associated with mean pH (4413 SNPs) and minimum annual sea surface temperature (SST, 1523 SNPs), respectively (Table 3.7). These SNPs mapped to 1813 unique genes for mean pH, 907 genes for minimum annual SST, 762 for mean annual SST, 592 for dissolved oxygen, 230 for maximum annual SST, and 83 for SST range.

Filtering greatly reduced the number of candidate SNPs and genes associate with each variable and therefore the number of enriched gene sets. Clustering of enriched gene sets for each environmental variable resulted in 7 enriched gene set clusters for mean pH, 4 clusters for minimum annual SST, 2 clusters for dissolved oxygen, 1 cluster for mean annual SST, 1 cluster for SST range, and no enriched gene sets or clusters for maximum annual SST (Table 3.8). The main difference in

enrichment results across environmental variables was the number of enriched gene sets. pH, min SST, mean SST, and dissolved oxygen were all associated with gene sets related to development, especially of the nervous system.

DISCUSSION

In this study, I demonstrated strong genetic structuring within *B. frenatus* across the majority of the species range. Genetic diversity generally decreased from south to north, suggesting a general recent history of northward expansion. I also revealed that a significant amount of the genetic structuring in this species is likely caused by differential selection and local adaptation. In particular, my results suggest the presence of a substantial amount of adaptive genetic variation to climate stressors, including varying levels of temperature and acidification. However, they also suggest low gene flow across the sampled range, which may limit the movement (and therefore availability) of adaptive genetic variation between subpopulations.

Patterns of differentiation & diversity

The observed population structure in *B. frenatus* seems to be the product of a combination of isolation-by-distance and vicariance (geographic/physical separation). In California, the overwhelming signal of divergence is that of isolation by distance, with pairwise weighted F_{ST} values (linearized) forming a strongly linear relationship with geographic distance ($r^2 = 0.96$) (Figure 3.5). Even within this region, however, there are indications of potential geographic boundaries. Most notably, individuals

from Bodega Bay are more diverged from the Central California populations than would be expected based on the regression of linearized F_{ST} and geographic distance. This can be seen visually in both the IBD plot (Figure 3.5), where these points are above the regression line and outside the standard error range, and in the PC plots (Figures 3.2-3.4). This divergence likely represents the San Francisco Bay as an ecological and geographic barrier. Kelp perch are highly associated with kelp forests, especially *Macrocystis pyrifera*, though they can also be found within other species of kelp (e.g., *Nereocystis luetkeana* and *Pterygophora californica*), seagrass beds, and other shallow, structured habitats (Love, 2011; Toy, personal obs.). The San Francisco Bay does not host appreciable amounts of canopy forming kelps (https://kelp.codefornature.org/) and may therefore represent an uninhabitable stretch of coast for *B. frenatus*. Additionally, kelp perch are rarely found at depths greater than 30 m (Hubbs & Hubbs, 1954; Love, 2011), so the depth and flow dynamics of the bay likely inhibit dispersal across its mouth. The Monterey Bay may represent a similar barrier to gene flow, but inference is limited due to a small sample size from Santa Cruz. Although there is abundant kelp habitat at the northern and southern edges, the interior of Monterey Bay is characterized by a low-structure, sandy coastal benthos, as well as a submarine canyon that divides it at the middle (Eittreim et al., 2002). This lack of suitable habitat may explain the gap seen between the Santa Cruz and Pacific Grove locations (Figure 3.3).

Admixture analysis supports the existence of these two biogeographic breaks by revealing low levels of admixture between California populations north of San

Francisco Bay and those south of it, with the two Santa Cruz individuals showing notably higher levels of Northern California ancestry than most other individuals from the Southern California cluster. Conversely, admixture analysis shows very similar ancestry proportions (~20% Northern California and 80% Southern California) for nearly all individuals from Pacific Grove to Point Buchon, with perhaps a slight increase in Southern California ancestry in the three individuals from Point Buchon. This pattern indicates relatively high levels of gene flow in this region, which is characterized by a relatively continuous distribution of large canopy-forming kelps (i.e., *M. pyrifera* and *N. luetkeana*; Nicholson et al., 2018). One individual from Pacific Grove, however, exhibits twice the proportion of Northern California ancestry (~40%) compared to others from that sampling location, likely reflecting a more recent admixture event (e.g., a fist generation backcross) in that individual lineage. Moving south from Point Buchon, the proportion of Southern California ancestry increases again to ~90% in individuals from Santa Barbara. Notably, I found little additional evidence of the traditionally hypothesized phylogeographic break at Point Conception (summarized by Burton, 1998), in concordance with predictions by Dawson (2001). Indeed, one individual from Santa Barbara displays a more admixed ancestry than most of the Central California locations (combined ~40% non-Southern California) with apparent additional ancestry from both the Puget Sound and British Columbia clusters, in addition to Northern California ancestry. The lack of strong evidence for this hypothesized phylogenetic gap may be due, in part, to the adult-centric mode of dispersal in *B. frenatus*. Structure between individuals within and

north of the Southern California Bite may be more prominent in species with pelagic larvae, whose dispersal is determined in large part by prevailing currents. Because kelp perch lack a pelagic larval stage, their connectivity may be less impacted by the potential effects of the divergent currents north and south of Point Conception. I also found little evidence of the well-documented Palos Verde break between Santa Barbara and San Diego (F_{ST} =0.03), as this point is very close to the regression line of linearized F_{ST} and geographic distance (Figure 3.5). There is, however, substantial separation between individuals from San Diego and Santa Barbara along PC 4 in Figure 3.4. Individuals from the San Diego location also exhibit a complete lack of admixture with locations north of San Francisco, while all other locations in the Southern California cluster showed some level of mixed ancestry/gene flow. This high structuring of individuals within the San Diego location may be the result of the effects of the Palos Verdes peninsula on local flow and deposition regimes (Dawson, 2001), and therefore habitat connectivity, possibly in combination with selection on divergent temperature regimes in the two locations.

Locations north of California are highly differentiated from the more southern locations, but also show strong differentiation from each other, apart from the two geographically close locations in Southeast Alaska (Figures 3.2 & 3.4). Comparisons between Puget Sound, British Columbia, and Alaska all stand out as outliers from the IBD regression, with high values of F_{ST} relative to geographic distance (Figure 3.5, red and green points). Even British Columbia and Puget Sound, which are separated by only 486 km, exhibit an F_{ST} of 0.31. For comparison, the divergence between San

Diego and Big Sur, which are separated by 535 km, is only 0.04. Interestingly, Puget Sound and British Columbia cluster relatively closely along PC 1, but are highly divergent along PC 2, indicating some level of genetic similarity relative to the Alaska locations. Similarly, Puget Sound and Alaska are highly divergent along PC 2 (and to a lesser extent along PC 1), but cluster closely along both PC 3 and PC 4. This strong genome-wide divergence (F_{ST}), but close similarity along some principal components is consistent with a scenario of a single initial colonization event of the more northern locations from California, followed by subsequent colonization events (and founder effects) within this region, with little recent gene flow between them (Figure 3.6).

This idea is further supported by estimates of nucleotide diversity within each location. Overall genetic diversity is highest in the southernmost locations, and generally declines with increasing latitude (Table 3.4), suggesting a generally poleward-expanding history of the species. Among the northern locations, British Columbia has the highest diversity and Puget Sound the lowest, indicating that British Columbia may have been the first of these locations to be colonized, and that Puget Sound and Southeast Alaska may have been subsequently colonized by individuals from British Columbia. Further, British Columbia and the two Alaska locations each show relatively lower levels of divergence from the two locations within the Southern California Bite (Santa Barbara and San Diego) than from the more northern California locations (these points fall below the regression line and outside of the standard error range in Figure 3.5). This may indicate that the individuals involved in

the initial colonization of this northernmost region migrated from Southern California, rather than Northern California. A possible mechanism of dispersal in this scenario would be through the association of migrants with kelp rafts, detached masses of tangled and buoyant kelp, which have been documented to travel great distances with their associated communities of algae and animals (Hobday, 2000a, 2000b; Mitchell & Hunter, 1970). Although the California Current pushes nearshore waters southward during the spring and summer, the Davidson current surfaces during the winter and runs south to north, close to shore (Checkley & Barth, 2009). The prevalence of kelp rafts in the Southern California Bite is also highest in winter (Hobday, 2000a), and the Southern California Countercurrent and Eddy, which drive water movement along the Southern California Bite, may have pushed kelp rafts hosting *B. frenatus* northwest and into the Davidson Current, where they were transported up towards British Columbia.

At the northernmost extent of this range, historical data also seems to support a more modern expansion north from British Columbia, though the remoteness of these northern areas may introduce historical sampling bias. Early records place the northern range limit of *B. frenatus* at Puget Sound (Jordan & Starks, 1895; Starks & Morris, 1907), but this limit was extended to Vancouver Island by 1928 at the latest (Ulrey & Greeley, 1928), and seems to have remained unchanged through the end of the century (Clemens & Wilby, 1961; Miller & Lea, 1976). The most reliable evidence of a modern northern expansion, however, comes from the first record of *B. frenatus* in Alaska. In 1998, 36 specimens were collected near Craig, AK (Csepp &

Wing, 1999). In this same study, several sites near Sitka Sound, farther north of Craig, were also sampled, but no *B. frenatus* individuals were recorded at these sites, indicating that the region around Craig may have been the true northernmost extent of the species range. However, three years later, several specimens were collected by Csepp from the southern end of Sitka Sound (Mecklenburg et al., 2002). During my own sampling, I observed large numbers of *B. frenatus* at sites throughout Sitka Sound, and my collection site in Salisbury Sound reflects the northernmost observation yet recorded for the species.

The shape of the site frequency spectrum in each location can also provide insight into the demographic history of this species. Population expansions and contractions can have strong effects on the frequency of singletons (alleles present in only one individual in a population) across the genome. Population expansions will tend to create a relative excess of singletons in a population compared to a neutrally evolving Wright-Fisher population. In contrast, population contractions will tend to remove singletons from the population, creating a skew in the distribution towards alleles of intermediate frequencies (Tajima, 1989). All four northern populations exhibit the lack of singletons that one would expect from a strong population contraction. The British Columbia spectrum is less skewed than the other three, further supporting the idea that British Columbia was colonized first by a small number of individuals and that this subpopulation then facilitated the colonization of locations to the north (Alaska) and south (Puget Sound). Conversely, the California populations exhibit site frequency spectra much closer in shape to what would be

expected from a constant-size Wright-Fisher population or an expanding population, with singletons by far representing the greatest proportion of polymorphic sites.

From both PC analysis and F_{ST} estimates, it is clear that the subpopulation within Puget Sound exhibits exceptional divergence from all other locations, even those closest to it. This likely results from a mix of the processes discussed above as well as strong local selection. As discussed previously, Puget Sound was likely colonized relatively recently. As such, this subpopulation may exhibit genetic divergence associated with a population bottleneck (founder effect). Subsequent physical separation (vicariance) may also have contributed to the divergence of this subpopulation. The Sound is a sheltered body of water with few outlets to the north and none to the south. Its main connection to the outer coast is the deep Strait of Juan de Fuca, which exhibits strong east-west tidal flows that may act as a barrier to dispersal (Holbrook et al., 1980). Selection may also play a role in this divergence, as the oceanographic conditions and terrestrial influences (freshwater flow, nutrient and pollutant concentrations) within the Sound can differ greatly from those of the outer coast (Holbrook et al., 1980; Moore et al., 2008; West et al., 2017). Along PC 2, which mainly separates Puget Sound from all other locations, I identified 10,247 SNPs as outliers potentially under selection. Thus, a combination of neutral and selective forces may contribute to the substantial divergence of this subpopulation from the others.

Given the seemingly high overall divergence between Puget Sound and British Columbia, it is possible that the Puget Sound location represents an independent colonization event from California migrants rather than colonization from nearby Vancouver Island. The lowest levels of divergence from Puget Sound are seen in comparison to the Pacific Grove location (F_{ST} = 0.26), but this value is only 0.05 less than the divergence between Puget Sound and British Columbia. Each of the four northern locations also exhibit nearly identical patterns of relative differentiation from the 6 southern locations (Figure 3.5), suggesting a common ancestry of the northern subpopulations. Furthermore, along the first two principal components (which together explain ~20% of the observed variance), British Columbia is the most similar location to Puget Sound, with the two grouping especially closely along PC 1. Ultimately, however, further demographic analysis is needed to determine which of these alternative histories is more likely.

Evidence of selection and local adaptation across the genome

Overall, I found strong evidence of selection across the *B. frenatus* genome (Figure S3.3). Between the PC and XtX analyses, 12,670 SNPs were overlapping. These SNPs mapped to 6,473 unique genes, which account for 19% of the 33,503 predicted genes within the genome, or 28% of the 22,936 genes with annotations. These proportions indicate a strong influence of selection in shaping the genetic diversity of this species.

Enrichment analysis of all candidate genes together identified 160 enriched GO terms, which clustered together into 23 general biological categories (Table 3.5). The largest cluster contained gene sets related to tissue, organ, and skeletal system morphogenesis and development, suggesting that the differences in environment across the species range have strong impacts on the development of embryos and juveniles. This makes intuitive sense, given that temperature is expected to be one of the largest differences in the abiotic environment between the northern and southern range extents, and that developmental rates are directly correlated with temperature in most ectotherms (Precht et al., 1973). In fact, this relationship has been directly documented in *Cymatogaster aggregata*, another Eastern Pacific surfperch species (Wiebe, 1968). Similarly, one would expect genes related to metabolism and transcript/protein turnover to be under selection under divergent temperature regimes. Indeed, I saw enrichment of gene sets consistent with this expectation, such as those related to protein catabolism, positive regulation of metabolic processes, and transcription coregulator activity.

The idea that temperature is an important selective factor was further corroborated by the environmental association analysis, as minimum sea surface temperature had the second highest number of candidate SNPs and candidate genes among all environmental variables tested (Table 3.7). The result that minimum SST, rather than mean or maximum SST, had the greatest number of associated SNPs is also consistent with theoretical expectations, given that the SST difference between the north and south range edges is greatest when using minimum SST (Table 3.2),

and that extreme temperature events can have strong impacts on species abundances and range boundaries (Kroeker et al., 2020; Wethey et al., 2011). Previous work on yellowtail clownfish populations in the Western Pacific also found more SNPs associated with minimum SST than maximum SST, although totals for minimum and mean SST were similar (Clark et al. 2021).

The strength of selection on gene sets related to development may in part reflect the viviparous reproduction of *B. frenatus*. In colder waters, slower rates of development would translate to longer gestation times for pregnant females. Prolonged gestation would likely have a substantial effect on fitness through direct impacts on both the mother and offspring. In elasmobranchs, for example, it has been suggested that longer gestation times may give mothers less time after parturition to restore energy reserves prior to the onset of winter conditions and the next reproductive cycle (Wallman & Bennett, 2006). Concomitantly, newborn offspring would also begin feeding later in the season, providing less opportunity for growth while resources are still abundant (Wallman & Bennett, 2006). Furthermore, pregnancy is often associated with increased risk of predation due to a decrease in mobility (Magnhagen, 1991), so prolonged gestation could further decrease the odds of female survival. The signals of selection observed in development-related genes may therefore reflect physiological adaptations that decrease gestation time, though common garden experiments would be necessary to confirm this.

Interestingly, there appeared to be a particular emphasis among the development-related gene sets on those related to the development of the central nervous system. Selection on these gene sets in particular may reflect spatially divergent pH regimes in the Northeast Pacific. Previous work in black surfperch (*Embiotoca jacksoni*) has demonstrated that even modest changes in pH can elicit changes in brain gene expression that may significantly impact neurological functioning (Toy et al., 2022). Mean pH was significantly associated with the second greatest number of candidate SNPs and genes among the environmental variables tested and, in fact, had the greatest totals of both when including only SNPs that were commonly identified as putatively under selection by the PC, XtX, and environmental association analyses (Table 3.7). The enrichment of gene sets related to neurological development may therefore reflect local adaptation to divergent pH regimes.

It is important to note that although enrichment analysis helps to interpret broader themes in adaptive differences between populations, it is also inherently biased against SNPs of large effect, since an ontology term with a single (or very few) genes under selection will likely never be recognized as significantly enriched, even if it is under strong selection. Therefore, this list of enriched gene set clusters is not necessarily inclusive of all important adaptive differences that may exist between populations. These analyses do, however, provide a useful overview of the general patterns of adaptation across the range of the species, and allow us to generate more specific hypotheses for mechanisms of adaptation.

Conclusions

In this study, I demonstrated strong population genetic structure in marine populations at the regional scale, and moderate levels at relatively small spatial scales (e.g., 0.017 over the <20 km between Pacific Grove and Carmel Bay). I found strong evidence that genetic variation in *B. frenatus* is shaped mainly by the pattern of isolation-by-distance in the south and by founder effects and vicariance in the north. I additionally identified evidence of phylogeographic breaks along the coast of California that are consistent with previous studies in other species. The signal of these phylogeographic breaks in divergence estimates stresses the importance of habitat continuity in maintaining connectivity in low-dispersal species. In the context of kelp forest conservation and restoration, my results demonstrate that large distances without suitable habitat (patchiness) may make recolonization of deforested areas difficult for direct-developing marine organisms. These are also important factors to consider in the design of MPAs, especially when setting and evaluating goals for connectivity between them (Palumbi, 2004).

I also found strong signals of selection across the range of *B. frenatus*, despite relatively low levels of genetic diversity in the northern populations. The strongest signal of selection seemed to be among genes related to development, with an emphasis on the central nervous system, and the two strongest selective pressures among the environmental variables tested were minimum SST and mean pH. These results provide evidence of local adaptation to climate stressors, and therefore a

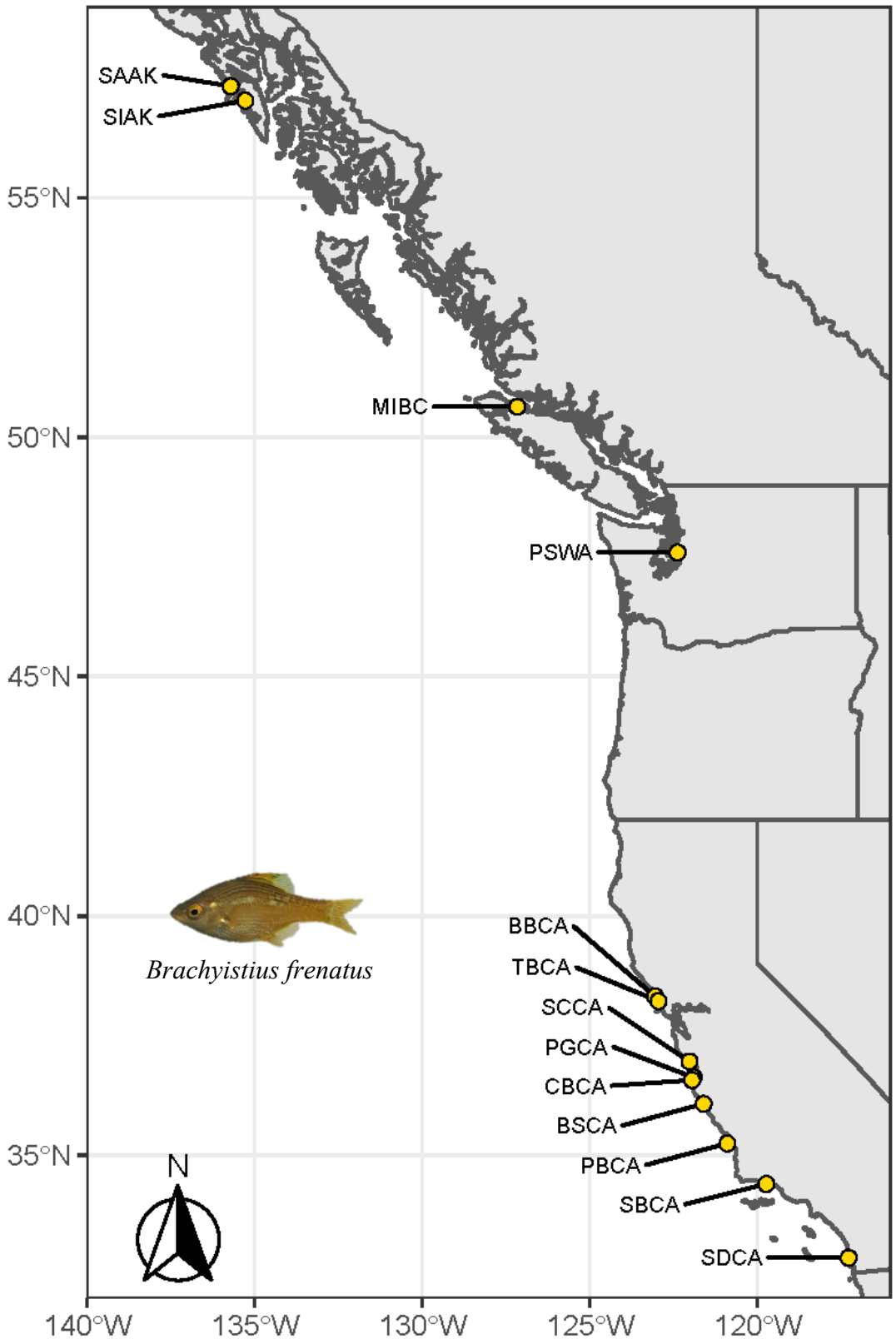
potential standing stock of adaptive variation that may increase the adaptive capacity of the species to future environmental changes. However, the high levels of population structure in the species indicates low levels of gene flow. Admixture analysis revealed little to no gene flow between each of the northern subpopulations and the other sampled locations, indicating a low likelihood that adaptive alleles from the south could reach the northern populations on a time-scale relevant to ongoing environmental change. The California populations, however, may fare better in this respect. Though these populations exhibited moderate levels of structuring, admixture analysis indicated significant gene flow occurs between them, likely due to relatively high habitat continuity in this region. Additional work is necessary to quantitatively estimate the rate at which adaptive alleles may be expected to flow between regions as well as the rate of gene flow that would be expected to significantly increase the likelihood of evolutionary rescue in this species.

Finally, my findings indicate a general history of south-to-north expansion in *B. frenatus*, with genetic diversity generally decreasing with latitude. Additionally, I have provided further evidence of a more recent range expansion in the North, as I note the northernmost observation of this species yet recorded. Both the genetic similarity of the two Alaska populations and their comparable levels of genetic diversity also indicate that my northernmost sites in Alaska may not represent the true current range edge, and that *B. frenatus* may in fact be present further north. I suggest that *B. frenatu*s would therefore provide an excellent system in which to empirically test hypotheses regarding the evolutionary dynamics of range edge populations during

marine range expansion, a pressing area of study in the face of a rapidly changing ocean environment.

FIGURES

Figure 3.1 - *Brachyistius frenatus* sampling locations. SAAK = Salisbury Sound, AK; SIAK = Sitka Sound, AK; MIBC = Malcolm Island, BC; PSWA = Puget Sound, WA; BBCA = Bodega Bay, CA; TBCA = Tomales Bay, CA; SCCA = Santa Cruz, CA; PGCA = Pacific Grove, CA; CBCA = Carmel Bay, CA; BSCA = Big Sur, CA; PBCA = Point Buchon, CA; SBCA = Santa Barbara, CA; SDCA = San Diego, CA. AK = Alaska, USA; BC = British Columbia, Canada; WA = Washington State, USA; CA = California, USA.


SAAK
SIAK
MIBC
PSWA
BBCA
TBCA
SCCA
PGCA
CBCA
BSCA
PBCA
SBCA
SDCA
55°N
50°N
45°N
40°N
35°N
140°W
135°W
130°W
125°W
120°W
N
Brachyistius frenatus

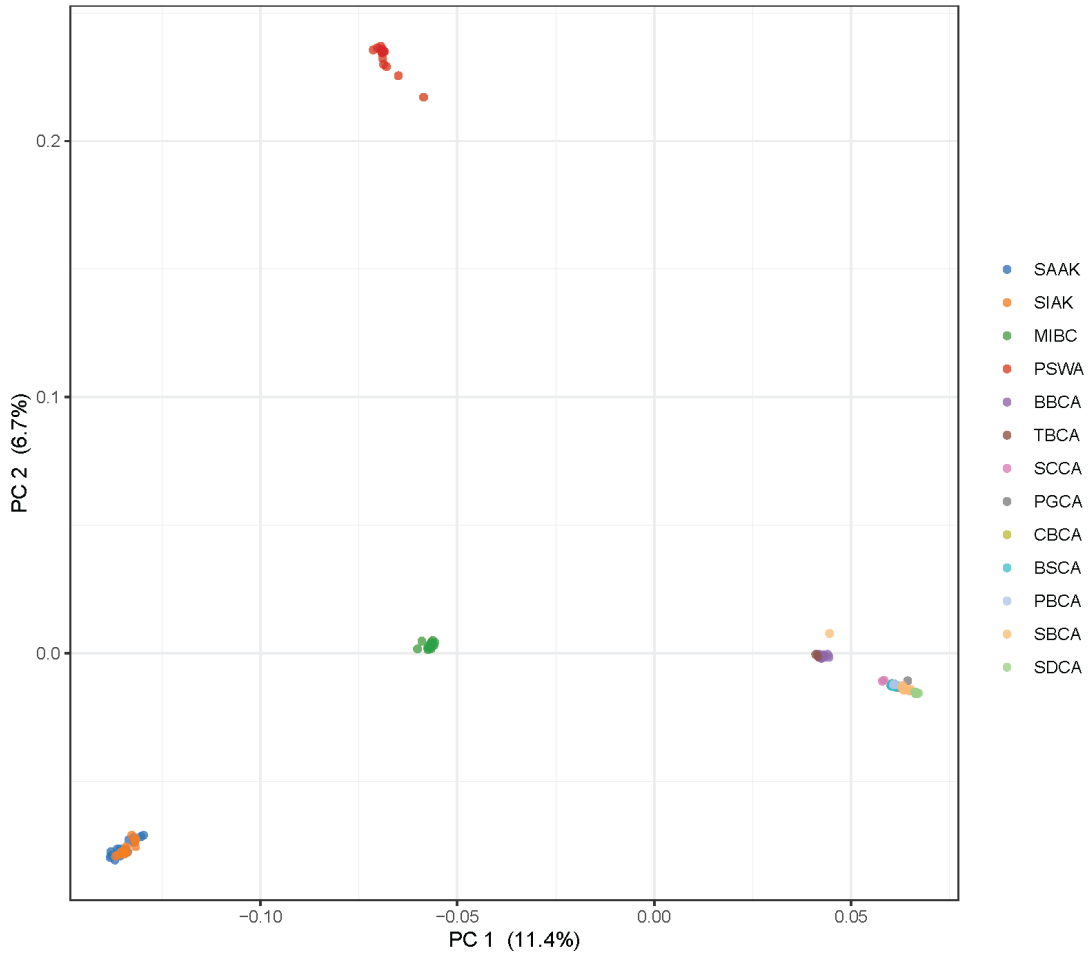


Figure 3.2 - Individual-level principal components plot for all sampling locations along the first two principal components. Axes show in parentheses the percent of the total variation explained by each principal component. Location codes as in Figure 3.1 and Table 3.1.

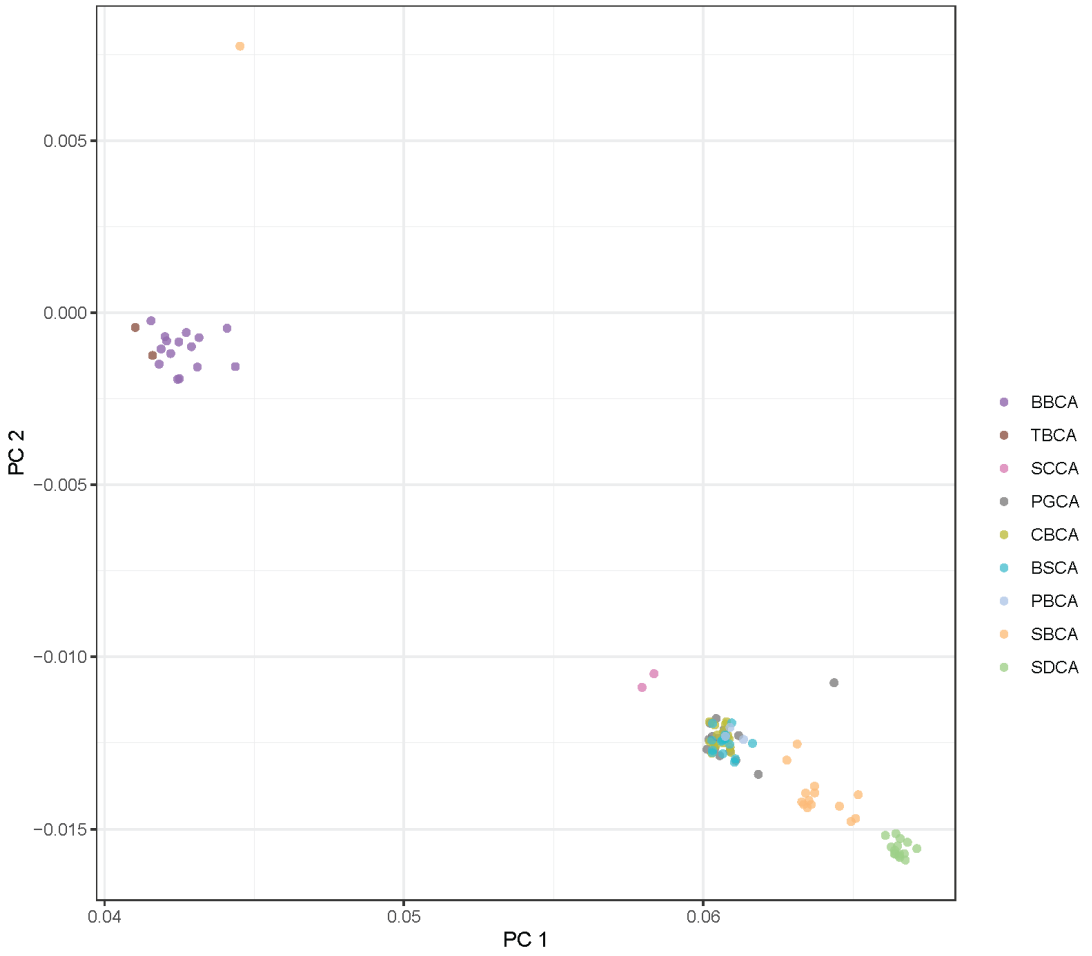


Figure 3.3 - Individual-level principal components plot for California locations along the first two principal components. Location codes as in Figure 3.1 and Table 3.1.

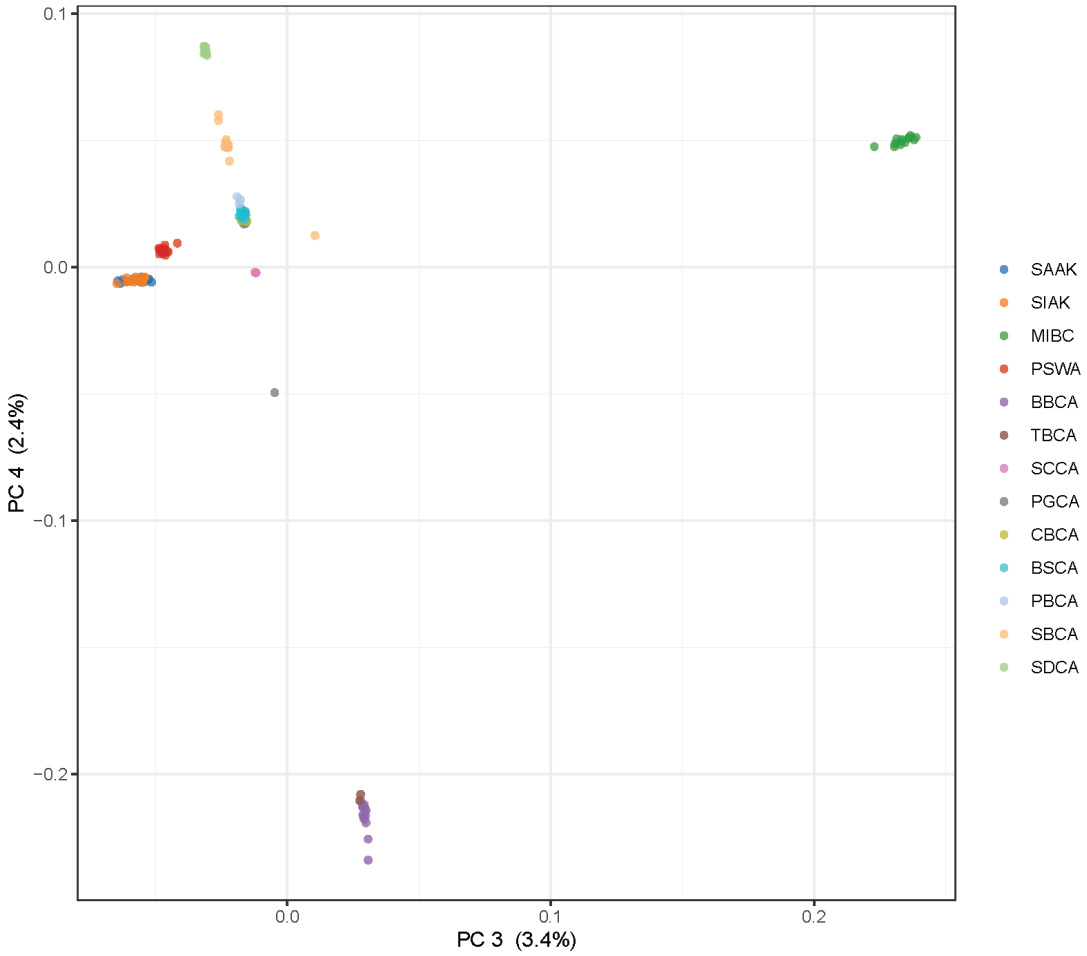


Figure 3.4 - Individual-level principal components plot for all sampling locations along the third and fourth principal components. Axes show in parentheses the percent of the total variation explained by each principal component. Location codes as in Figure 3.1 and Table 3.1.

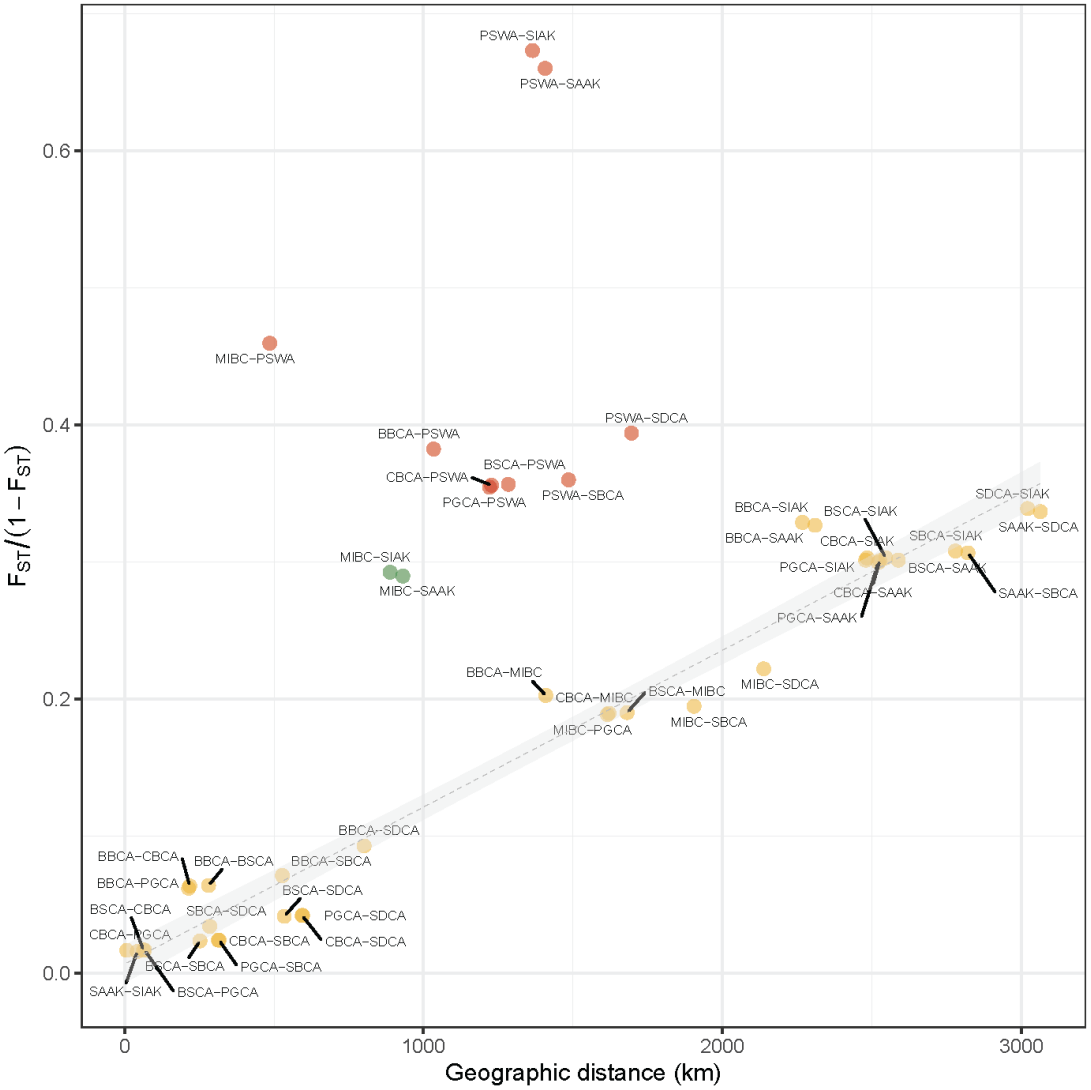


Figure 3.5 - Plot of isolation-by-distance (IBD) for all sampling locations. Linearized F_{ST} was regressed against the straight-line distance between sampling locations, excluding comparisons that included Puget Sound (red) and comparisons between the northern locations (green), which are likely to be driven by founder effects and vicariance.

Figure 3.6 - Plot of admixture proportions for all individuals. K=5 was selected based on the number of significant principal components.

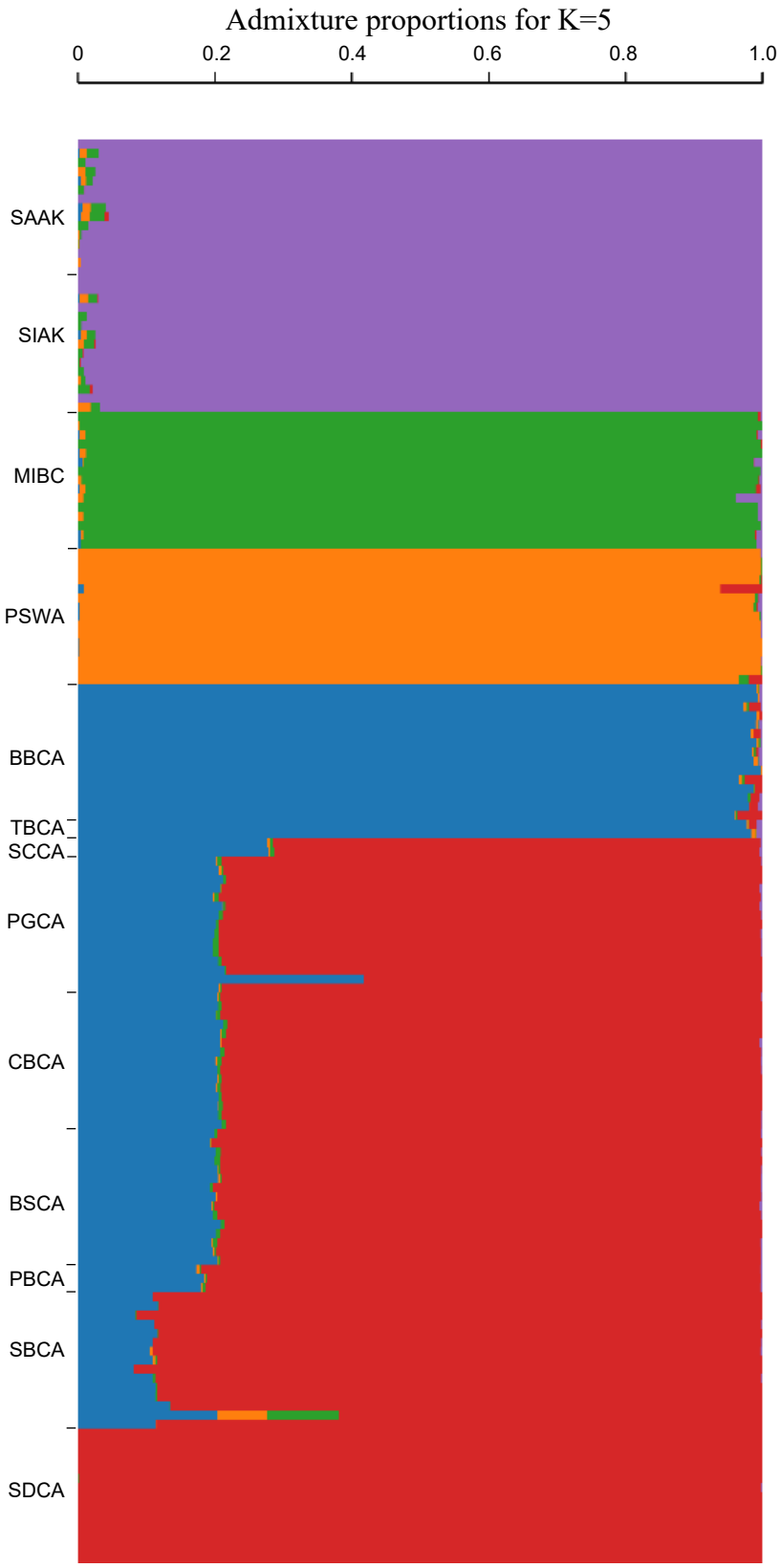
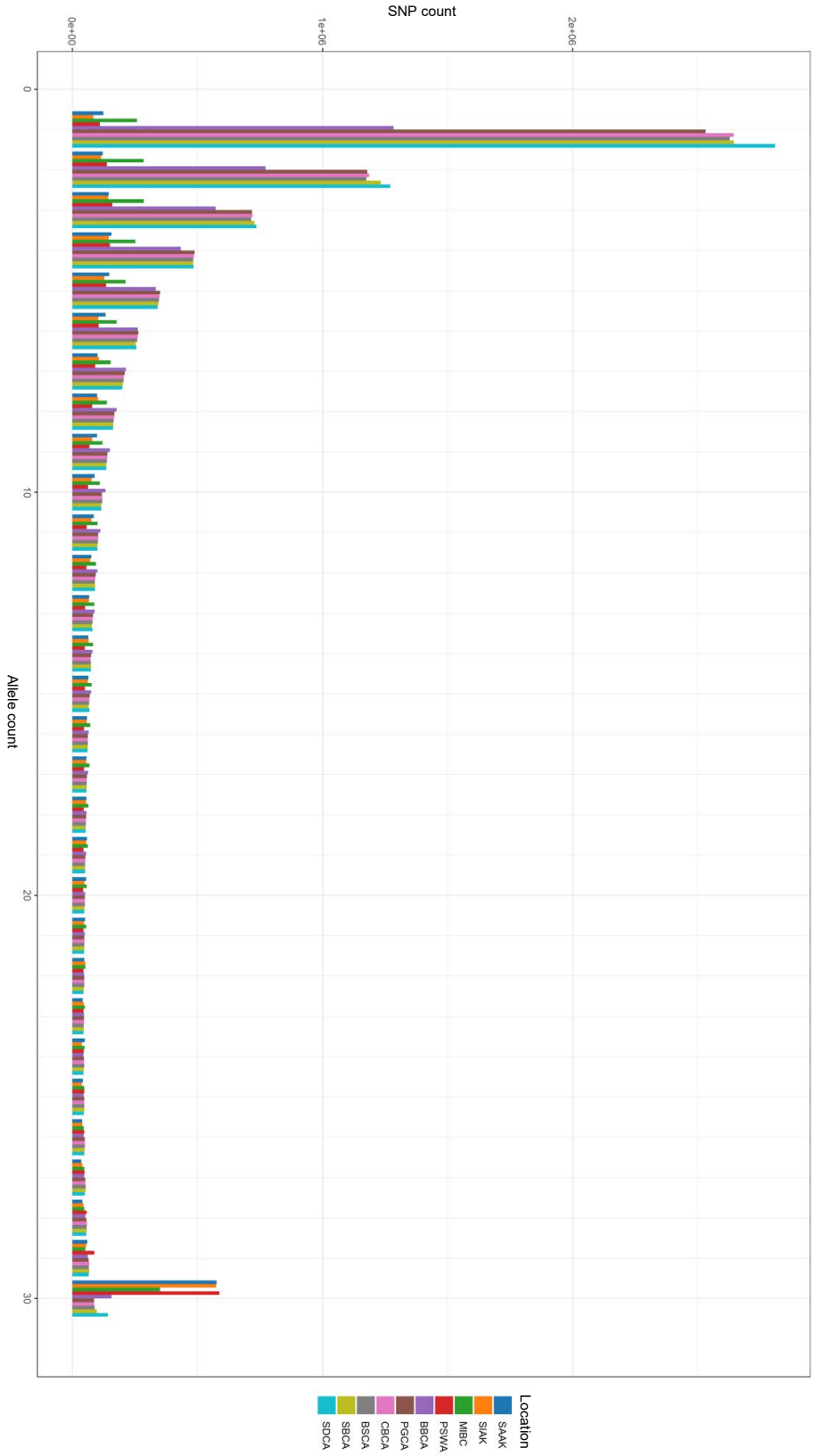

Admixture proportions for K=5
0
0.2
0.4
0.6
0.8
1.0
SAAK
SIAK
MIBC
PSWA
BBCA
TBCA
SCCA
PGCA
CBCA
BSCA
PBCA
SBCA
SDCA

Figure 3.7 - Site frequency spectra (SFS) for each sampling location. Major and minor alleles were defined based on genotype likelihoods from all 158 sequenced individuals. An allele count of 30 indicates a fixed minor allele.


SNP count
0e+00
1e+06
2e+06
Allele count
0
10
20
30
Location
SDCA
SBCA
BSCA
CBCA
PGCA
BBCA
PSWA
MIBC
SIAK
SAAK

TABLES

Table 3.1 - Collection date, name, code, coordinates, and sample size for each location.

Collection date	Location name	Location code	Latitude	Longitude	n
2019-06-21	Salisbury Sound, AK	SAAK	57.332	-135.712	15
2018-06-20	Sitka Sound, AK (Harris Island)	SIAK	57.031	-135.277	13
2018-06-20	Sitka Sound, AK (Sandy Cove)	SIAK	56.985	-135.321	2
2019-10-18	Malcolm Island, BC	MIBC	50.634	-127.159	16
2019-11-25	Puget Sound, WA (Seacrest Park)	PSWA	47.589	-122.379	15
2020-08-14	Bodega Bay, CA	BBCA	38.314	-123.052	15
2020-09-29	Tomales Bay, CA	TBCA	38.216	-122.95	2
2021-11-21	Santa Cruz, CA	SCCA	36.954	-122.023	2
2020-07-08	Pacific Grove, CA	PGCA	36.621	-121.901	15
2019-09-27	Carmel Bay, CA	CBCA	36.564	-121.944	15
2020-07-16	Big Sur, CA	BSCA	36.069	-121.601	15
2020-08-12	Pt. Buchon, CA	PBCA	35.241	-120.896	3
2020-08-31	Santa Barbara, CA	SBCA	34.395	-119.73	15
2020-08-25	San Diego, CA	SDCA	32.852	-117.276	15

Table 3.2 - Environmental data used in the BayPass environmental association analysis. Data are derived from the Bio-ORACLE database. SST = sea surface temperature in degrees Celsius. DO = dissolved oxygen concentration in mL/L.

Location	Mean SST	SST range	SST max	SST min	DO	pH	Latitude
SAAK	6.852	13.480	15.300	1.820	6.803	8.153	57.332
SIAK	8.541	12.605	15.544	2.939	6.921	8.151	57.031
MIBC	8.908	6.595	13.400	6.805	6.847	8.147	50.634
PSWA	10.918	7.686	14.797	7.111	6.094	7.833	47.589
BBCA	12.352	3.469	14.353	10.884	6.230	8.210	38.314
PGCA	13.865	3.884	16.239	12.355	6.345	8.207	36.621
CBCA	13.028	3.552	15.189	11.637	6.269	8.206	36.564
BSCA	13.237	3.456	14.935	11.479	5.987	8.198	36.069
SBCA	16.006	4.855	18.874	14.019	5.980	8.200	34.395
SDCA	18.155	7.311	22.136	14.825	5.798	8.221	32.852

Table 3.3 - Weighted F_{ST} values for all pairwise comparisons of sampling locations. Cells are colored according to F_{ST} value. Deeper red indicates greater differentiation.

	SAAK	SIAK	MIBC	PSWA	BBCA	PGCA	CBCA	BSCA	SBCA
SIAK	0.0156								
MIBC	0.2247	0.2264							
PSWA	0.3976	0.4023	0.3149						
BBCA	0.2463	0.2475	0.1685	0.2766					
PGCA	0.2306	0.2316	0.1587	0.2618	0.0585				
CBCA	0.2314	0.2325	0.1595	0.2626	0.0600	0.0165			
BSCA	0.2316	0.2326	0.1598	0.2629	0.0602	0.0167	0.0167		
SBCA	0.2347	0.2355	0.1630	0.2647	0.0666	0.0236	0.0237	0.0231	
SDCA	0.2520	0.2532	0.1818	0.2827	0.0850	0.0406	0.0405	0.0399	0.0331

Table 3.4 - Estimates of genetic diversity for each sampling location

Location	Mean θ_W	Mean θ_π
SAAK	0.00101	0.00135
SIAK	0.00093	0.00125
MIBC	0.00141	0.00176
PSWA	0.00091	0.00114
BBCA	0.00243	0.00234
PGCA	0.00322	0.00259
CBCA	0.00327	0.00260
BSCA	0.00325	0.00258
SBCA	0.00328	0.00258
SDCA	0.00336	0.00260

Table 3.5 - Enriched gene set clusters across genes under selection. Gene list included only genes putatively under selection according to all both the PC and XtX analyses. Gene sets are from the GO Biological Process, Molecular Function, and Cellular Component databases. Enriched gene sets were clustered by similarity and manually summarized. The second column indicates the number of enriched gene sets within a given cluster.

Gene set cluster	Num. gene sets
tissue, organ, & skeletal system morphogenesis/development	30
axonogenesis & neuron development	25
vasculature development & ameboidal-type cell migration	25
dendritic spine	12
protein catabolism	11
protein kinase activity	10
positive regulation of GTPase activity	8
renal system development	6
endosome/vesicle	5
Wnt signaling pathway	4
ATP-dependent chromatin remodeler activity	3
extracellular matrix organization	3
positive regulation of metabolic process	3
receptor complex & plasma membrane	3
Golgi vesicle transport	2
metallopeptidase activity	2
pattern specification process/regionalization	2
central nervous system development	1
flavin adenine dinucleotide binding	1
muscle structure development	1
response to mechanical stimulus	1
transcription coregulator activity	1
UDP glucosyltransferase activity	1

Table 3.6 - Clusters of enriched gene sets along each significant principal component. Gene sets are from the GO Biological Process, Molecular Function, and Cellular Component databases. Enriched gene sets were clustered by similarity and manually summarized. The second column indicates the number of enriched gene sets within a given cluster.

PC 2		PC 3		PC 4	
Gene Set Cluster	***Num. Gene Sets***	***Gene Set Cluster***	***Num. Gene Sets***	***Gene Set Cluster***	***Num. Gene Sets***
neuron differentiation, development & morphogenesis	14	embryonic skeletal system development & morphogenesis	11	neuron differentiation, development & morphogenesis	22
embryonic skeletal system development & morphogenesis	10	GTPase activator activity	8	embryonic eye/sensory organ development & morphogenesis	20
neuron projection/dendrite	6	neuron differentiation & development	6	epithelial cell migration & locomotion	8
renal system development & renal filtration	6	vasculature development	5	blood vessel & nephron tubule development	5
GTPase regulator activity	5	neural crest/mesenchymal cell differentiation & migration	5	transmembrane receptor protein tyrosine kinase activity	5
plasma membrane signaling receptor complex	4	deubiquitinase activity	3	central nervous system development	3
detection of/response to mechanical stimulus	4	embryonic medial fin morphogenesis	3	cell adhesion	1
protein kinase activity	2	renal filtration	2	phosphatidylserine metabolic process	1
synapse & postsynapse	2	sensory organ development & morphogenesis	2		
digestive tract development	1	cell-cell signaling	1		
enzyme linked receptor protein signaling	1	GABA-gated chloride ion channel activity	1		
epithelial cell differentiation	1	protein serine/threonine/tyrosine kinase activity	1		
Golgi vesicle transport	1				
protein catabolic process	1				
transcription elongation factor complex	1				

Table 3.7 - Total counts of SNPs and genes associated with each environmental variable. "Common" indicates the totals of associated SNPs and genes after filtering to include only those also identified in the two outlier analyses (PC and XtX).

Env. variable	Num candidate SNPs	Mean BF	Num genes w/ candidate SNPs	Mean BF	Num common candidate SNPs	Mean BF	Num genes w/ common candidate SNPs	Mean BF
Mean pH	109,216	24.42	7160	31.64	4413	34.08	1813	37.39
Min. SST	143,160	24.20	7844	30.49	1523	23.93	907	24.56
Mean SST	86,308	25.33	6978	31.70	1212	26.00	762	27.17
Diss. oxygen	63,985	25.93	6161	32.22	901	29.01	592	30.88
Max. SST	27,509	26.71	4406	31.43	281	26.50	230	27.09
SST Range	79,676	24.75	6625	30.31	103	24.05	83	23.88

Table 3.8 - Clusters of enriched gene sets for genes associated with each environmental variable. Gene sets are from the GO Biological Process, Molecular Function, and Cellular Component databases. Enriched gene sets were clustered by similarity and manually summarized. The number of enriched gene sets within each cluster is also indicated.

pH		SST min		SST mean		DO		SST range	
Gene Set Cluster	**Num. Enriched Gene Sets**	**Gene Set Cluster**	**Num. Enriched Gene Sets**	**Gene Set Cluster**	**Num. Enriched Gene Sets**	**Gene Set Cluster**	**Num. Enriched Gene Sets**	**Gene Set Cluster**	**Num. Enriched Gene Sets**
neuron differentiation, development & morphogenesis	14	embryonic morphogenesis	5	central nervous system neuron differentiation/ axonogenesis	2	central nervous system neuron differentiation/ development	4	vitamin binding	1
detection of/response to mechanical stimulus	3	cellular component morphogenesis	3			sensory organ morphogenesis	1		
animal organ development	2	gliogenesis	1						
GTPase regulator activity	2	Golgi vesicle transport	1						
protein transport & localization	2								
anatomical structure formation involved in morphogenesis	1								
gliogenesis	1								

Synthesis

To best understand, mitigate, and withstand the environmental disruptions caused by anthropogenic carbon emissions, the ecological framework of plastic, individual-level responses must be merged with an evolutionary framework of population-level responses. To this end, I conducted my dissertation research as an investigation of the relationships between population genomic diversity, local adaptation, and adaptive capacity in the face of global environmental change. The studies described in this dissertation have provided much-needed empirical insight into the role of local adaptation in mediating population persistence through rapid environmental change, as well as the biological mechanisms through which climate adaptation may occur.

In Chapter 1, I provided evidence of acidification as a significant selective force in surfperches through impacts on neurological function. This work also highlighted the importance of environmental variability in organismal responses to changing environments, and I argue that natural variability must be considered in manipulative ocean change experiments to ensure accurate prediction of ecological impacts. In Chapter 2, I assembled a high-quality reference genome for the kelp perch, *Brachyistius frenatus*, providing an important molecular resource (and only the second reference genome) for the study of this unique family of fishes and for the field of comparative genomics. In Chapter 3, I provide the first whole genome-level assessment of the genomic diversity of a surfperch species, incorporating samples from almost the entire species range. I demonstrated that limited dispersal has indeed

resulted in striking genetic differentiation within this species. I also identified several thousand outlier genes, indicating that the observed differentiation is in part due to spatial environmental heterogeneity and the local adaptation of subpopulations. Finally, I showed that a substantial number of genes under selection are significantly associated with regional differences in climate-related variables, most notably minimum sea surface temperature and mean pH. Taken together, I have provided strong evidence of local adaptation to climate variables across the sampled range, indicating significant standing adaptive variation in this species. My results, however, also indicate that limited dispersal may limit the flow and therefore availability of adaptive variation in and out of geographically distant subpopulations.

By providing a better understanding of the adaptive genetic diversity present within marine populations, this work should prove valuable not only to ecologists and evolutionary biologists, but to conservation practitioners and managers as well. A more comprehensive framework of global change impacts that incorporates evolutionary processes will allow managers to more accurately anticipate demographic and genetic changes in natural populations and allow them to better prepare for potential negative outcomes of global change. Additionally, the patterns of diversity revealed by this work provide insight into likely "hot" and "cold" spots of genomic diversity in an ecologically important group of fishes. This type of data is of critical importance to the management of coastal species, especially those that may be impacted by other human activities (i.e., fisheries), and may facilitate more effective use of limited resources by conservation practitioners. Finally, this work has provided

critical and novel insight into the likelihood of evolutionary rescue in nearshore, Eastern Pacific fishes. It has demonstrated that local adaptation along this coast has likely maintained a standing stock of potentially adaptive variation, but that low levels of gene flow within a species may limit its availability to subpopulations when faced with rapid environmental change. Overall, this dissertation not only addresses important empirical gaps in the field of global change biology, but also a pertinent practical management need: a more comprehensive and accurate understanding of population responses to global change.

Appendix 1: Supplementary material for Chapter 1

REFERENCE GENOME INFORMATION

Sample collection and DNA extraction

We generated a de novo genome of E. jacksoni using gill and liver tissue from a single individual freshly collected in Monterey, California in September of 2014 (collected on September 13th, sacrificed September 14th. Total DNA was extracted with Qiagen Blood and Cell Midi Kit following the manufacturer's protocol.

Sequencing and assembly

Sequencing and assembly were carried out by Dovetail® genomics. Briefly, genomic DNA was sheared and used to make an Illumina sequencing library (Meyer and Kircher, 2010). First, the DNA was sonicated to approximately 300 bp and end repaired. Next, sequencing adapters were ligated to both ends of the DNA. The adapters were filled-in, and the DNA was amplified in an indexing PCR. After library preparation, the library size distribution was confirmed by agarose gel electrophoresis, and the library was size-selected with a Sage Science BluePippin with a 2% agarose gel cassette. The library was sequenced on an Illumina HiSeq 2500 with 2x150 PE rapid run chemistry.

Genome Assembly Statistics

***E. jacksoni* Genome Assembly Statistics (accession JALAZG000000000)**	
Genome Assembly Size	567,544,745 bp
Estimated Chicago Library physical coverage (1-50 Kb pairs)	52.4x
Number of Scaffolds	1452
Number of Scaffolds > 1 kb	1450
Average Scaffold Length	390,871.04 bp
Longest Scaffold	21,990,620 bp
N50	5,372,808 bp; n = 25 scaffolds
N60	3,570,986 bp; n = 39 scaffolds
N70	2,019,634 bp; n = 61 scaffolds
N80	1,404,917 bp; n = 95 scaffolds
N90	746,313 bp; n = 149 scaffolds
N100	1,000 bp; n = 1452 scaffolds
N count	8,862,637
Gaps	95,485
BUSCO Scores	
Summary	C:92.4%[S:91.7%,D:0.7%], F:2.6%, M:5.0%, n:3640
Complete BUSCOs (C)	3361
Complete and single-copy BUSCOs (S)	3337
Complete and duplicated BUSCOs (D)	24
Fragmented BUSCOs (F)	93
Missing BUSCOs (M)	186
Total BUSCO groups searched (n)	3640

Figure S1.1 - Calibrated pH time series data for the duration of Experiment 2. Solid lines represent hourly averaged Durafet data. Dashed lines represent daily YSI data. The two vertical grey lines indicate the dates where each group of treatments was sampled for brain tissue (September 23 and 24). Headers 1 and 6 correspond to the ambient treatment, 2 and 7 to the pH 7.85 static treatment, 3 and 8 to the pH 7.85 variable treatment, 4 and 9 to the pH 7.70 static treatment, and 5 and 10 to the pH 7.70 variable treatment.

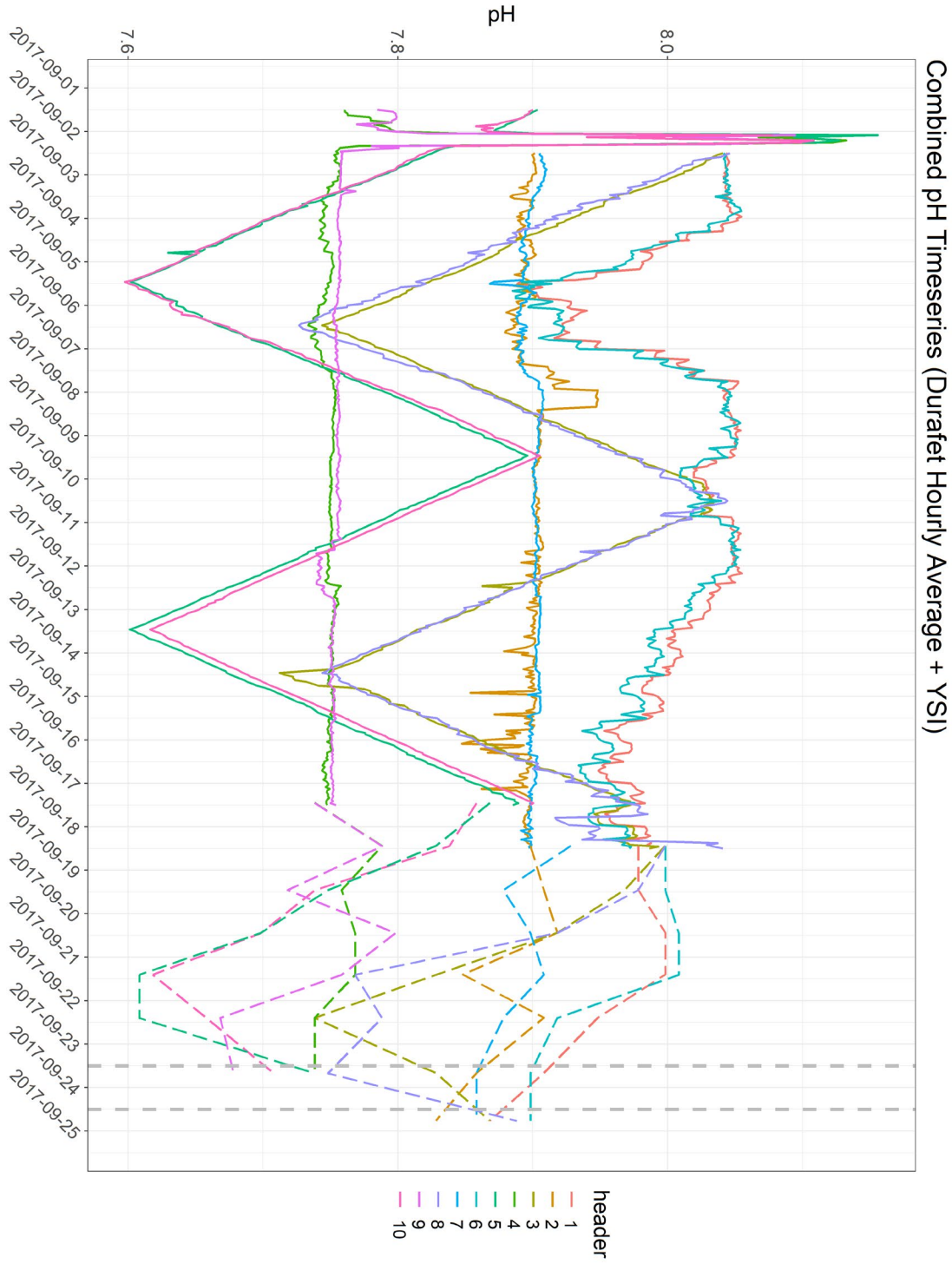
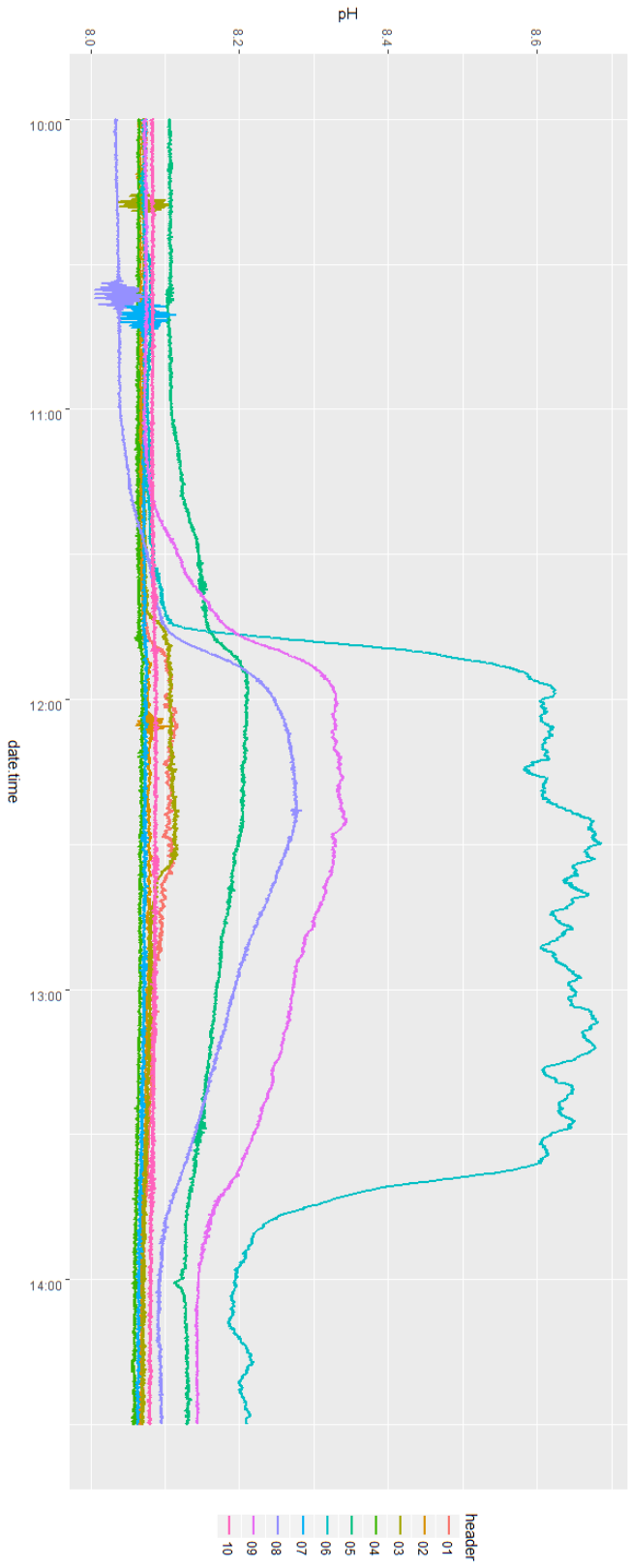

Combined pH Timeseries (Durafet Hourly Average + YSI)
pH
7.6
7.8
8.0
2017-09-01
2017-09-02
2017-09-03
2017-09-04
2017-09-05
2017-09-06
2017-09-07
2017-09-08
2017-09-09
2017-09-10
2017-09-11
2017-09-12
2017-09-13
2017-09-14
2017-09-15
2017-09-16
2017-09-17
2017-09-18
2017-09-19
2017-09-20
2017-09-21
2017-09-22
2017-09-23
2017-09-24
2017-09-25
header
1
2
3
4
5
6
7
8
9
10

Figure S1.2 - Post-hoc test of the effect of fouled Durafets on measured pH. The plot shows the pH (uncalibrated) recorded by each Durafet during a 4.5 hr time period around midday on October 5, 2017 (11 days after the experiment ended). Headers 1 and 6 correspond to the ambient treatment, 2 and 7 to the pH 7.85 static treatment, 3 and 8 to the pH 7.85 variable treatment, 4 and 9 to the pH 7.70 static treatment, and 5 and 10 to the pH 7.70 variable treatment.


pH
8.0
8.2
8.4
8.6
10:00
11:00
12:00
13:00
14:00
date.time
header
10
09
08
07
06
05
04
03
02
01

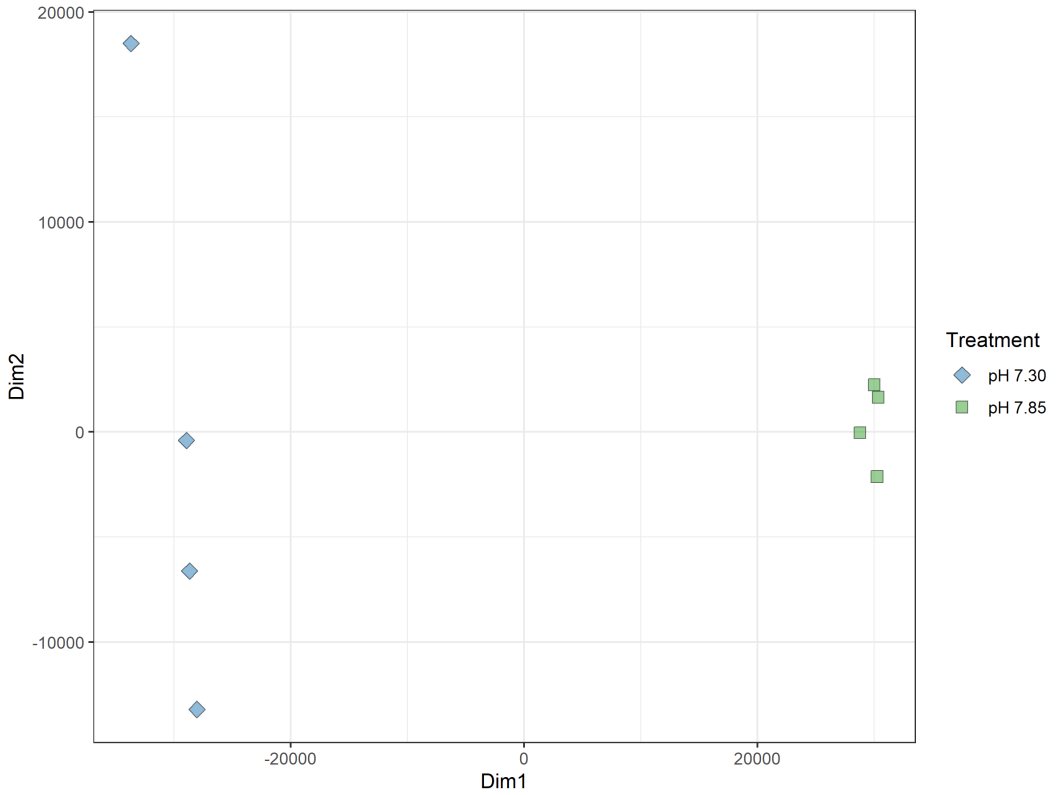


Figure S1.3 - Metric MDS plot of global gene expression (all genes) in Experiment 1. The distance matrix was created using Manhattan distances between points.

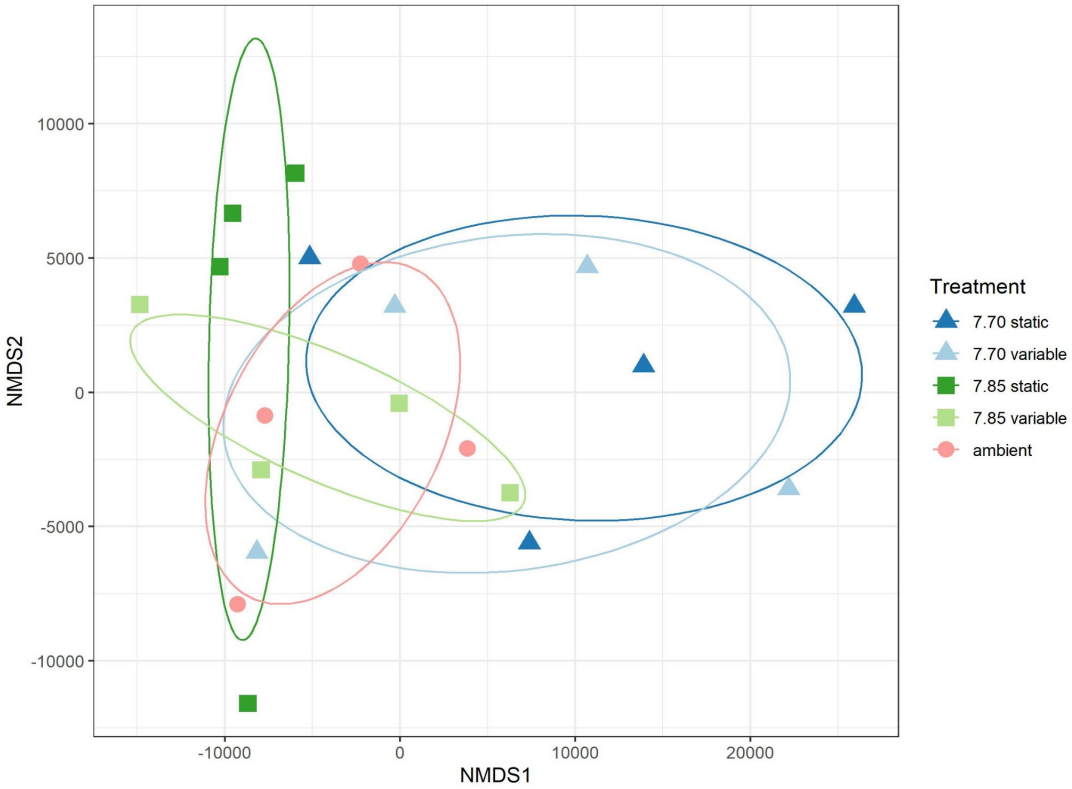


Figure S1.4 - nMDS plot of global gene expression (all genes) in Experiment 2. The distance matrix was created using Manhattan distances between points. Ellipses are 95% confidence ellipses.

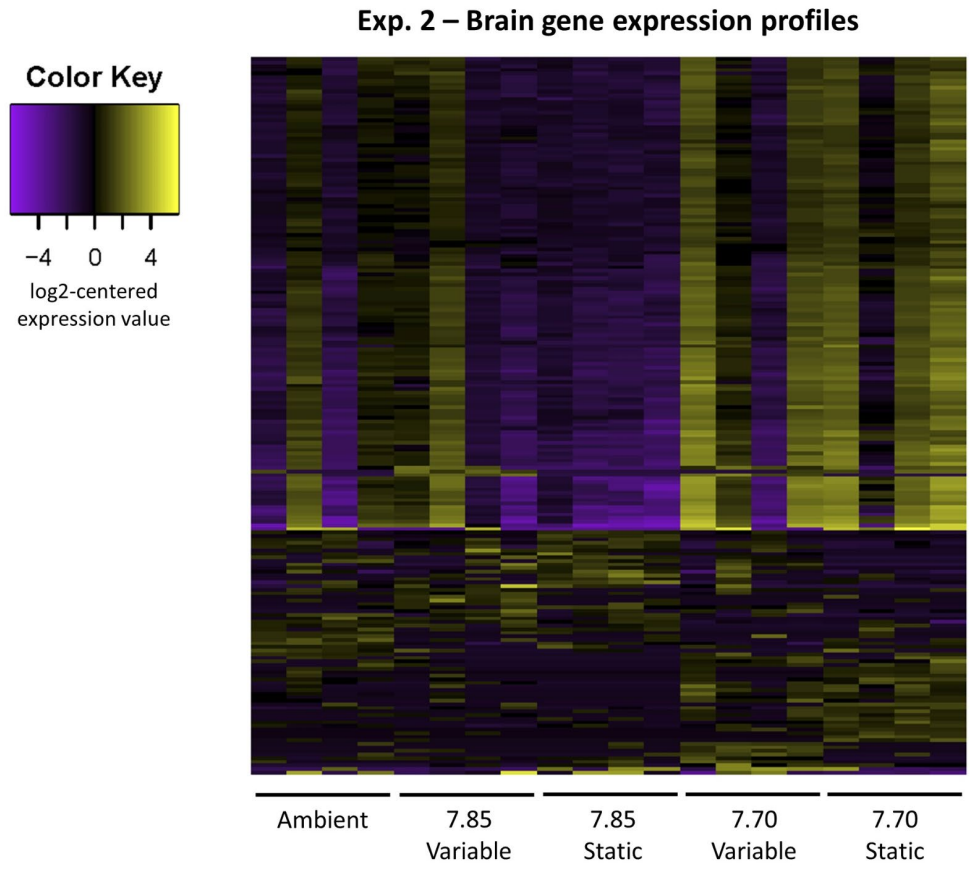


Figure S1.5 - Heatmap of gene expression profiles for each individual in Experiment 2. Each column represents an individual fish, and each row represents a gene. Relative expression is shown here for all genes found to be differentially expressed (DEGs) across all treatment comparisons. Yellow represents upregulation in a given treatment and purple represents downregulation.

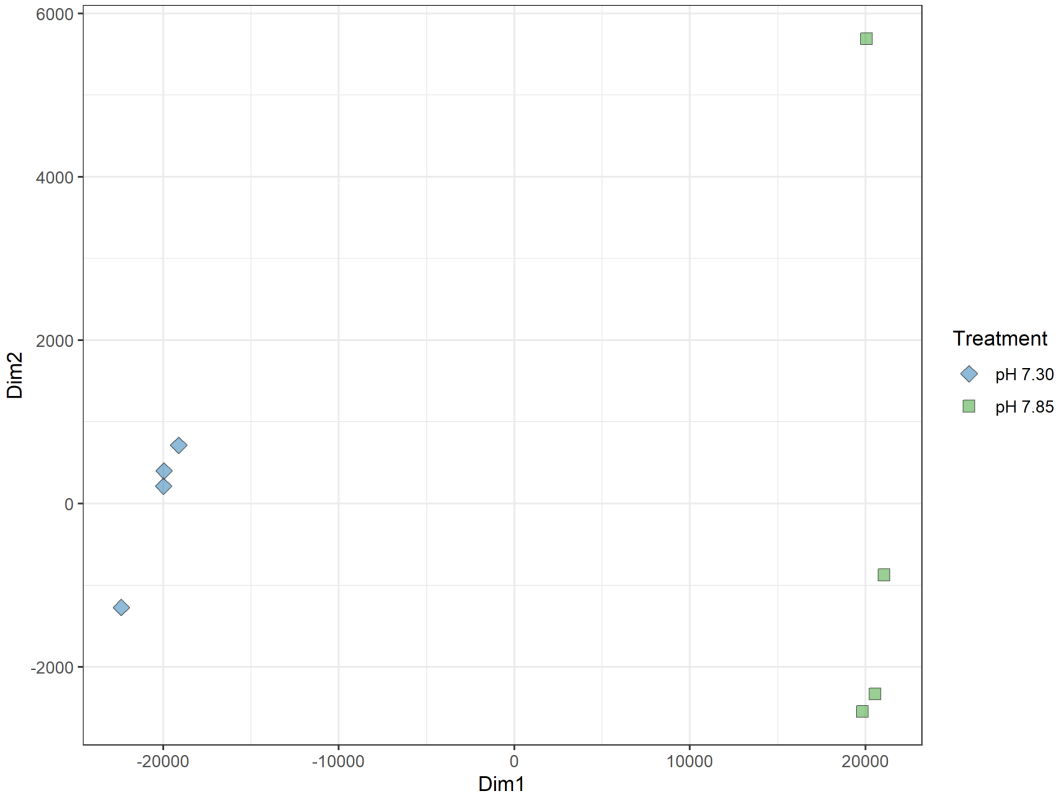


Figure S1.6 - Metric MDS plot of DEG expression in Experiment 1. The distance matrix was created using Manhattan distances between points.

SUPPLEMENTARY TABLES

Table S1.1 - Summary of relevant dates for Experiments 1 and 2.

	Experiment 1	Experiment 2
Collection Date(s)	4 Sep 2015	28 Jul - 14 Aug 2017
Tank Acclimation (start)	27 Oct 2015	23-24 Aug 2017
Ramp-up (start)	27 Oct 2015	26 Aug 2017
Ramp-up (end)	3 Nov 2015	28 Aug 2017
Experiment Start	3 Nov 2015	1-2 Sep 2017
Tissue Sampling	26 Nov 2015	23-24 Sep 2017

Table S1.2 - Summary statistics of the *E. jacksoni* transcriptome assembly calculated using the TransRate software.

	***E. jacksoni* Assembly**
n_seqs	71933
smallest	61
largest	40145
n_bases	2.11 E+08
mean_len	2933.171
n_under_200	4090
n_over_1k	51458
n_over_10k	1635
n_with_orf	46928
mean_orf_percent	43.72653
n90	1666
n70	3286
n50	4808
n30	6665
n10	10030
gc_percent	47.408
bases_n	142100
proportion_n	0.00067
score	NA
optimal_score	NA
cutoff	NA
weighted	NA

Table S1.3 - Number of reads per sample, mean reads per sample, and standard deviation for Experiment 1 samples after quality trimming with Trimmomatic.

Sample (Experiment 1)	Reads per Sample (trim)	Mean Reads per Sample	SD
pH 7.30 - 1	16,367,926	17,074,659	711,261
pH 7.30 - 2	17,797,231		
pH 7.30 - 3	17,564,720		
pH 7.30 - 4	16,568,758		
pH 7.85 - 1	16,200,642	13,453,330	1,977,285
pH 7.85 - 2	12,866,976		
pH 7.85 - 3	13,239,618		
pH 7.85 - 4	11,506,085		

Table S1.4 - Number of reads per sample, mean reads per sample, and standard deviation for Experiment 2 samples after quality trimming with Trimmomatic.

Sample (Experiment 2)	Reads per Sample (trim)	Mean Reads per Sample	SD
Ambient 1	13,935,486	16,371,070	2,952,497
Ambient 2	14,130,786		
Ambient 3	20,169,981		
Ambient 4	17,248,026		
pH 7.70 Static - 1	16,672,746	17,848,144	2,037,506
pH 7.70 Static - 2	15,589,007		
pH 7.70 Static - 3	19,376,973		
pH 7.70 Static - 4	19,753,849		
pH 7.70 Variable - 1	18,107,479	18,527,366	1,910,768
pH 7.70 Variable - 2	16,023,025		
pH 7.70 Variable - 3	20,328,241		
pH 7.70 Variable - 4	19,650,719		
pH 7.85 Static - 1	14,132,210	17,816,049	2,941,341
pH 7.85 Static - 2	16,785,087		
pH 7.85 Static - 3	19,852,330		
pH 7.85 Static - 4	20,494,568		
pH 7.85 Variable - 1	14,428,536	18,858,298	3,474,465
pH 7.85 Variable - 2	22,922,441		
pH 7.85 Variable - 3	18,974,671		
pH 7.85 Variable - 4	19,107,544		

Table S1.5 - PERMANOVA analysis of global gene expression for Experiment 1 (single factor).

1-way PERMANOVA (Experiment 1, factor = treatment)						
	df	**Sums of sqs**	**Mean sqs**	**F Model**	**r^2**	**Pr(>F)**
Treatment	1	7119107119	7119107119	25.766	0.81112	0.02857
Residuals	6	1657807605	276301268		0.18888	
Total	7	8776914724			1	

Table S1.6 - PERMANOVA analysis of global gene expression for Experiment 2 (2-factor).

2-Factor PERMANOVA (Experiment 2, ambient treatment excluded)						
	df	**Sums of sqs**	**Mean sqs**	**F Model**	**r^2**	**Pr(>F)**
pH mean (no ambient)	1	946119146	946119146	2.53414	0.15874	0.02137
pH var (no ambient)	1	221435079	221435079	0.59311	0.03715	0.88953
pH mean: pH var	1	312333151	312333151	0.83657	0.0524	0.52444
Residuals	12	4480184179	373348682	0.7517		
Total	15	5960071555	1			

Table S1.7 - Results of pairwise comparisons of global gene expression for all treatments in Experiment 2 using the pairwiseAdonis package for R.

Pairs	***df***	**Sums of sqs**	**F model**	**r^2**	**p value**
Ambient vs. pH 7.70 Static	1	632943714	1.7877803	0.229562	0.085714
Ambient vs. pH 7.70 Var	1	421245194	1.140972	0.159778	0.314286
Ambient vs. pH 7.85 Static	1	278412470	0.8893589	0.129092	0.542857
Ambient vs. pH 7.85 Var	1	194698128	0.6091689	0.09217	1
pH 7.70 Static vs. pH 7.70 Var	1	247959677	0.5923027	0.089848	0.942857
pH 7.70 Static vs. pH 7.85 Static	1	896000714	2.4718167	0.291769	0.057143
pH 7.70 Static vs. pH 7.85 Var	1	553236095	1.4990777	0.199902	0.142857
pH 7.70 Var vs. pH 7.85 Static	1	614318130	1.626702	0.21329	0.171429
pH 7.70 Var vs. pH 7.85 Var	1	362451583	0.9433668	0.135866	0.4
pH 7.85 Static vs. pH 7.85 Var	1	285808553	0.8712066	0.126791	0.571429

Table S1.8 - Multivariate analysis of DEG expression in Experiment 1. Single-factor PERMANOVA analysis of DEG expression in Experiment 1 yielded similar results to the analysis of global gene expression. There was a strong and significant effect of pH level ($r^2 = 0.953$, F = 122.26, p = 0.029), with pH explaining 95% of the observed variation.

1-way PERMANOVA (Experiment 1, factor = treatment)						
	df	**Sums of sqs**	**Mean sqs**	**F model**	**r^2**	**p value**
Treatment	1	3319506452	3319506452	122.26	0.95322	0.02857
Residuals	6	162905622	27150937		0.04678	
Total	7	3482412074			1	

Table S1.9 - Results of 2-factor PERMANOVA reveal significant differences in DEG expression levels between pH-level treatments, but not pH variability treatments (ambient treatment not included) in Experiment 2. The interaction, however, is marginally significant.

	df	**Sums of sqs**	**Mean sqs**	**F model**	**r^2**	**p value**
pH mean	1	590351	590351	10.7713	0.3878	0.003913
pH var	1	63093	63093	1.1512	0.04145	0.290594
pH mean:pH var	1	211192	211192	3.8533	0.13873	0.052156
Residuals	12	657692	54808	0.43203		
Total	15	1522328	1			

Table S1.10 - Results of pairwise comparisons of DEG expression for all treatments in Experiment 2 using the pairwiseAdonis package for R.

Pairs	***df***	**Sums of sqs**	**F model**	**r^2**	**p value**
Ambient vs. pH 7.70 Static	1	262744.4	4.690179	0.438737	0.085714
Ambient vs. pH 7.70 Var	1	159891.2	1.901314	0.240633	0.171429
Ambient vs. pH 7.85 Static	1	126518.2	2.911884	0.326742	0.057143
Ambient vs. pH 7.85 Var	1	58742.33	0.835613	0.122244	0.428571
pH 7.70 Static vs. pH 7.70 Var	1	116785.3	1.717644	0.222561	0.171429
pH 7.70 Static vs. pH 7.85 Static	1	663043.7	24.24698	0.801633	0.028571
pH 7.70 Static vs. pH 7.85 Var	1	262125.7	4.836718	0.446327	0.057143
pH 7.70 Var vs. pH 7.85 Static	1	391319	7.060927	0.540615	0.057143
pH 7.70 Var vs. pH 7.85 Var	1	138499.2	1.683474	0.219103	0.228571
pH 7.85 Static vs. pH 7.85 Var	1	157499.8	3.783892	0.386747	0.057143

Table S1.11 - Results of one-tailed t-tests of log(F-ratio) for each pH level in Experiment 2. F-ratios were calculated as the average variance of the variable treatment divided by the average variance of the static treatment. Alternative hypothesis: true mean is greater than 0.

One-tailed t-tests of log(F-ratio)			
	df	**t-statistic**	**p value**
Target pH 7.85	199	3.7205	0.000129
Target pH 7.70	199	1.8234	0.034869

Table S1.12 - See supplementary file "tables_s1.12-1.15_combined.xlsx". Enriched gene sets between the pH 7.85 and pH 7.30 treatments in Experiment 1 (identified using FGSEA analysis).

Table S1.13 - See supplementary file "tables_s1.12-1.15_combined.xlsx". Enriched gene sets between the static pH treatments (7.85 and 7.70) in Experiment 2 (identified using FGSEA analysis).

Table S1.14 - See supplementary file "tables_s1.12-1.15_combined.xlsx". Enriched gene sets between the pH 7.85 static and pH 7.85 variable treatments in Experiment 2 (identified using FGSEA analysis).

Table S1.15 - See supplementary file "tables_s1.12-1.15_combined.xlsx". Enriched gene sets between the pH 7.70 static and pH 7.70 variable treatments in Experiment 2 (identified using FGSEA analysis).

Table S1.16 - Overlapping enriched gene sets between the static vs. variable treatment comparisons (pH 7.85 and pH 7.70) in Experiment 2.

Upregulated in pH 7.85 Var / Downregulated in pH 7.70 Var	**Downregulated in pH 7.85 Var / Upregulated in pH 7.70 Var**
GOBP RNA splicing via transesterification reactions	GOBP anatomical structure formation involved in morphogenesis
GOBP RNA splicing	GOCC vesicle membrane
GOBP mRNA processing	GOBP supramolecular fiber organization
GOCC spliceosomal complex	GOBP positive regulation of cell differentiation
GOBP RNA processing	GOBP vasculature development
KEGG spliceosome	GOMF cell adhesion molecule binding
GOBP mRNA metabolic process	GOBP regulation of cellular component movement
GOCC nuclear protein-containing complex	GOBP positive regulation of locomotion
GOCC ribonucleoprotein complex	GOCC cell surface
	GOBP regulation of cell adhesion
	GOBP myeloid leukocyte differentiation
	GOBP exocytosis
	GOCC ficolin 1 rich granule membrane
	GOBP cell-cell adhesion
	GOBP integrin-mediated signaling pathway
	GOMF molecular transducer activity
	GOBP phagocytosis
	GOBP positive regulation of cell adhesion
	GOBP regulation of leukocyte differentiation
	GOBP leukocyte mediated immunity
	GOBP leukocyte migration

Table S1.17 - Up- and Downregulated genes related to GABA signaling (GO - *gamma-aminobutyric acid signaling pathway*, KEGG - *GABAergic synapse*) in Experiment 1 and 2. “U” and orange coloration indicate upregulation in the lower pH treatment. “D” and blue coloration indicate downregulation in the lower pH treatment. “C” and grey coloration indicate conflicting directions of expression change between gene IDs with the same gene symbol annotation. The “Exp 2” column refers to the comparison of the two static pH treatments only.

Gene	Description	Exp 1	Exp 2
adcy1	Adenylate Cyclase 1	D	
adcy2	Adenylate Cyclase 2	D	
adcy3	Adenylate Cyclase 3	D	
adcy5	Adenylate Cyclase 5	D	
adcy6	Adenylate Cyclase 6	D	
adcy8	Adenylate Cyclase 8	D	
atf4	Activating Transcription Factor 4	U	
bdnf	Brain Derived Neurotrophic Factor	D	
cacna1a	Calcium Voltage-Gated Channel Subunit Alpha1 A	D	
cacna1b	Calcium Voltage-Gated Channel Subunit Alpha1 B	D	
cacna1c	Calcium Voltage-Gated Channel Subunit Alpha1 C	D	
cacna1d	Calcium Voltage-Gated Channel Subunit Alpha1 D	D	
cacnb4	Calcium Voltage-Gated Channel Auxiliary Subunit Beta 4	D	
gabarapl1	GABA Type A Receptor Associated Protein Like 1	U	
gabarapl2	GABA Type A Receptor Associated Protein Like 2	D	
gabbr1	Gamma-Aminobutyric Acid Type B Receptor Subunit 1	D	
gabbr2	Gamma-Aminobutyric Acid Type B Receptor Subunit 2	D	
gabra1	Gamma-Aminobutyric Acid Type A Receptor Subunit Alpha1	D	
gabra2	Gamma-Aminobutyric Acid Type A Receptor Subunit Alpha2	D	
gabra3	Gamma-Aminobutyric Acid Type A Receptor Subunit Alpha3	D	
gabra4	Gamma-Aminobutyric Acid Type A Receptor Subunit Alpha4	D	
gabra5	Gamma-Aminobutyric Acid Type A Receptor Subunit Alpha5	D	
gabra6	Gamma-Aminobutyric Acid Type A Receptor Subunit Alpha6	D	U
gabrb1	Gamma-Aminobutyric Acid Type A Receptor Subunit Beta1	D	
gabrb2	Gamma-Aminobutyric Acid Type A Receptor Subunit Beta2	D	
gabrb3	Gamma-Aminobutyric Acid Type A Receptor Subunit Beta3	D	U
gabrg1	Gamma-Aminobutyric Acid Type A Receptor Subunit Gamma1	D	
gabrg2	Gamma-Aminobutyric Acid Type A Receptor Subunit Gamma2	D	
gabrg3	Gamma-Aminobutyric Acid Type A Receptor Subunit Gamma3	D	
gabrp	Gamma-Aminobutyric Acid Type A Receptor Subunit Pi	D	
gabrr2	Gamma-Aminobutyric Acid Type A Receptor Subunit Rho2	D	
gad1	Glutamate Decarboxylase 1	D	
gad2	Glutamate Decarboxylase 2	D	
gls	Glutaminase	C	
glul	Glutamate-Ammonia Ligase	D	
gnai1	G Protein Subunit Alpha I1	D	
gnai2	G Protein Subunit Alpha I2	D	
gnao1	G Protein Subunit Alpha O1	D	
gnb1	G Protein Subunit Beta 1	D	
gnb2	G Protein Subunit Beta 2	D	

gnb5	G Protein Subunit Beta 5	D	
gng12	G Protein Subunit Gamma 12	D	
gng13	G Protein Subunit Gamma 13	D	
gng2	G Protein Subunit Gamma 2	D	
gng4	G Protein Subunit Gamma 4	D	
gng7	G Protein Subunit Gamma 7	D	
gphn	Gephyrin	D	
gpr156	G Protein-Coupled Receptor 156	U	
htr4	5-Hydroxytryptamine Receptor 4	D	
kcnj6	Potassium Inwardly Rectifying Channel Subfamily J Member 6	D	
nsf	N-Ethylmaleimide Sensitive Factor, Vesicle Fusing ATPase	D	
phf24	PHD Finger Protein 24	D	
plcl1	Phospholipase C Like 1 (Inactive)	D	
plcl2	Phospholipase C Like 2	D	
prkaca	Protein Kinase CAMP-Activated Catalytic Subunit Alpha	D	
prkca	Protein Kinase C Alpha	D	
prkcb	Protein Kinase C Beta	D	
slc12a2	Solute Carrier Family 12 Member 2	D	
shisa7	Shisa Family Member 7	D	
slc12a5	Solute Carrier Family 12 Member 5	D	
slc32a1	Solute Carrier Family 32 Member 1	D	
slc38a3	Solute Carrier Family 38 Member 3	C	
slc6a1	Solute Carrier Family 6 Member 1	D	
slc6a11	Solute Carrier Family 6 Member 11	D	
slc6a13	Solute Carrier Family 6 Member 13	D	
src	SRC Proto-Oncogene, Non-Receptor Tyrosine Kinase	D	

SUPPLEMENTARY MATERIAL REFERENCES

TransRate: reference free quality assessment of de-novo transcriptome assemblies (2016). Richard D Smith-Unna, Chris Boursnell, Rob Patro, Julian M Hibberd, Steven Kelly. Genome Research doi: http://dx.doi.org/10.1101/gr.196469.115

Appendix 2: Supplementary material for Chapter 2

SUPPLEMENTARY METHODS

See supplementary file “supplementary_methods_ch2.docx”.

SUPPLEMENTARY TABLES

Table S2.1 - Summary table of MinION sequencing reads after quality filtering, adapter trimming, and length filtering (>500 bp). Produced using NanoStat (v1.5.0).

General Summary	
Mean read length	3,268.7
Mean read quality	14.6
Median read length	2,418.0
Median read quality	14.7
Number of reads	4,966,516.0
Read length N50	4,886.0
STDEV read length	2,973.0
Total bases	16,234,043,863
Number, percentage and megabases of reads above quality cutoffs	
>Q5	4,966,515 **(100.0%)** 16,234.0 Mb
>Q7	4,966,503 **(100.0%)** 16,234.0 Mb
>Q10	4,965,319 **(100.0%)** 16,231.7 Mb
>Q12	4,276,732 **(86.1%)** 14,000.6 Mb
>Q15	2,250,354 **(45.3%)** 7,299.6 Mb
Top 5 highest mean base call quality scores and their read lengths	
1	37.7 (2036)
2	35.0 (662)
3	28.6 (675)
4	27.7 (506)
5	27.4 (651)
Top 5 longest reads and their mean base call quality score	
1	8,7449 (14.5)
2	8,3790 (11.0)
3	8,0001 (15.0)
4	7,7865 (14.8)
5	7,7769 (13.9)

Table S2.2 - See supplementary file “table_s2.2.xlsx”. Summary table of repeats identified in the genome assembly using RepeatMasker.

Appendix 3: Supplementary material for Chapter 3

Figure S3.1 - Plot of mean sequencing depth across all 158 individuals sampled. The black dotted line indicates the mean of the mean depths, and the dotted red line indicates the median of the mean depths.

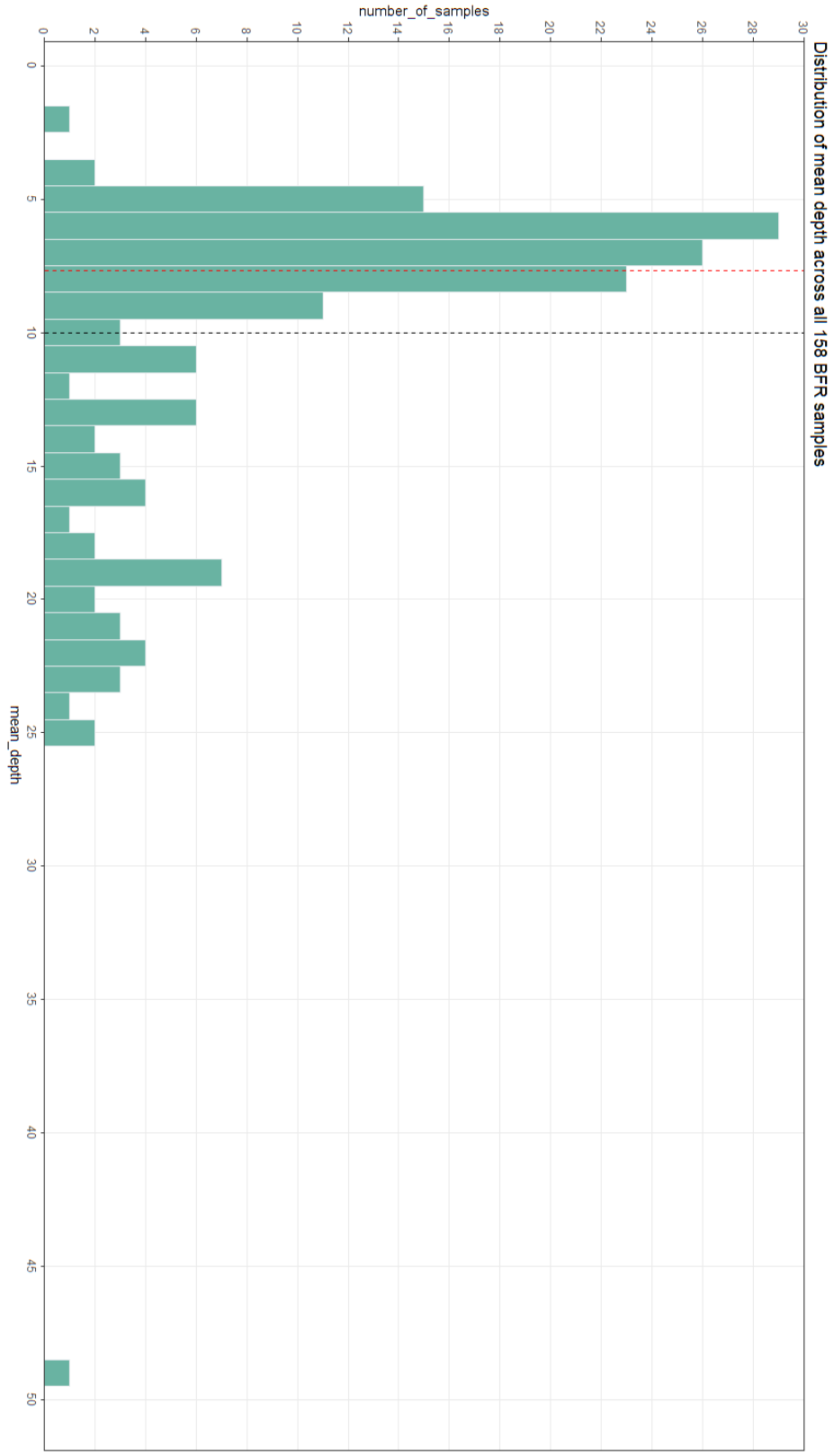

Distribution of mean depth across all 158 BFR samples
number_of_samples
mean_depth
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
0
5
10
15
20
25
30
35
40
45
50

Figure S3.2 - Plot of the distribution of the proportion of the reference sequence covered for each of the 158 individuals sampled. The mean proportion covered across all individuals is 0.974.

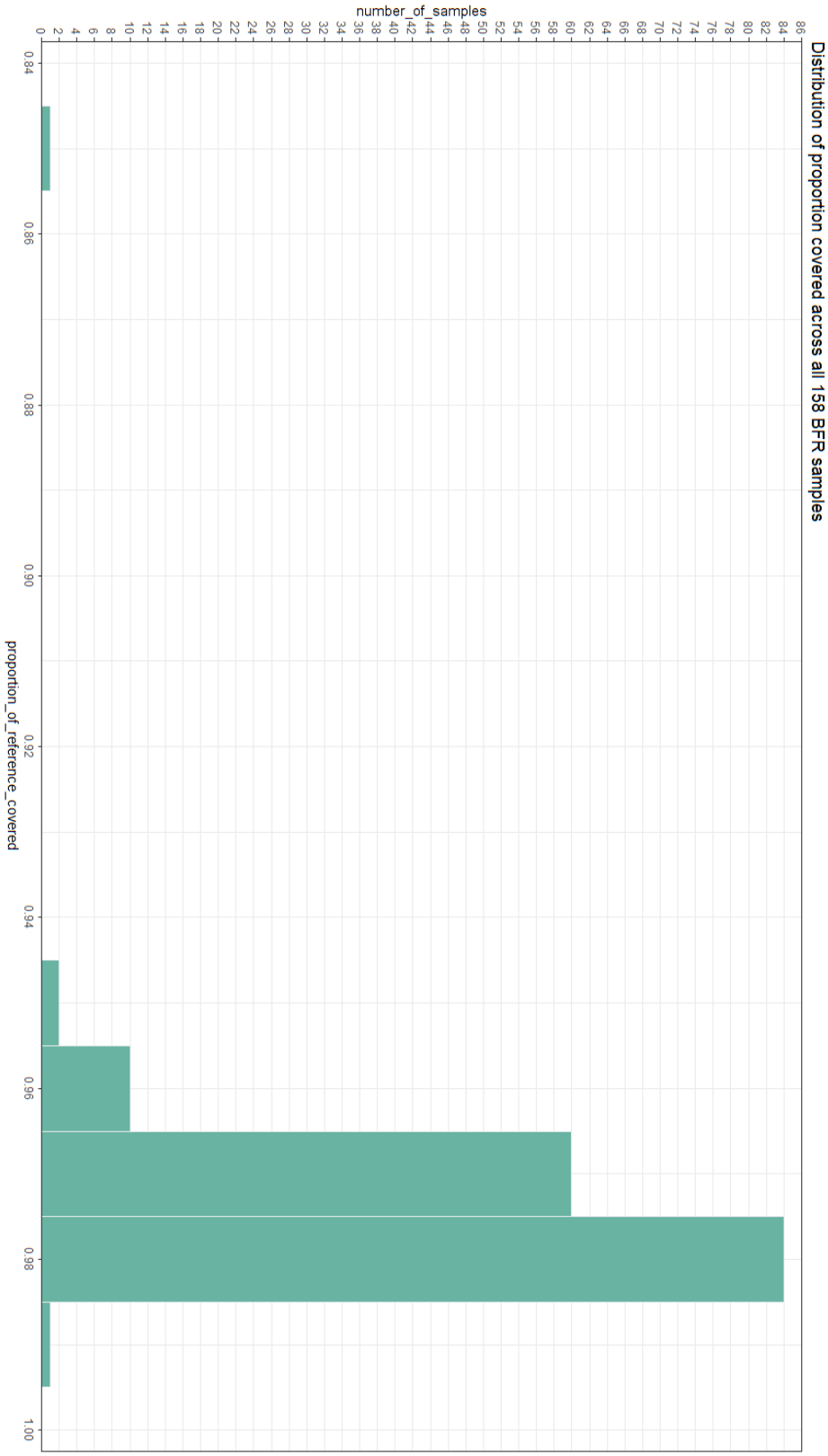
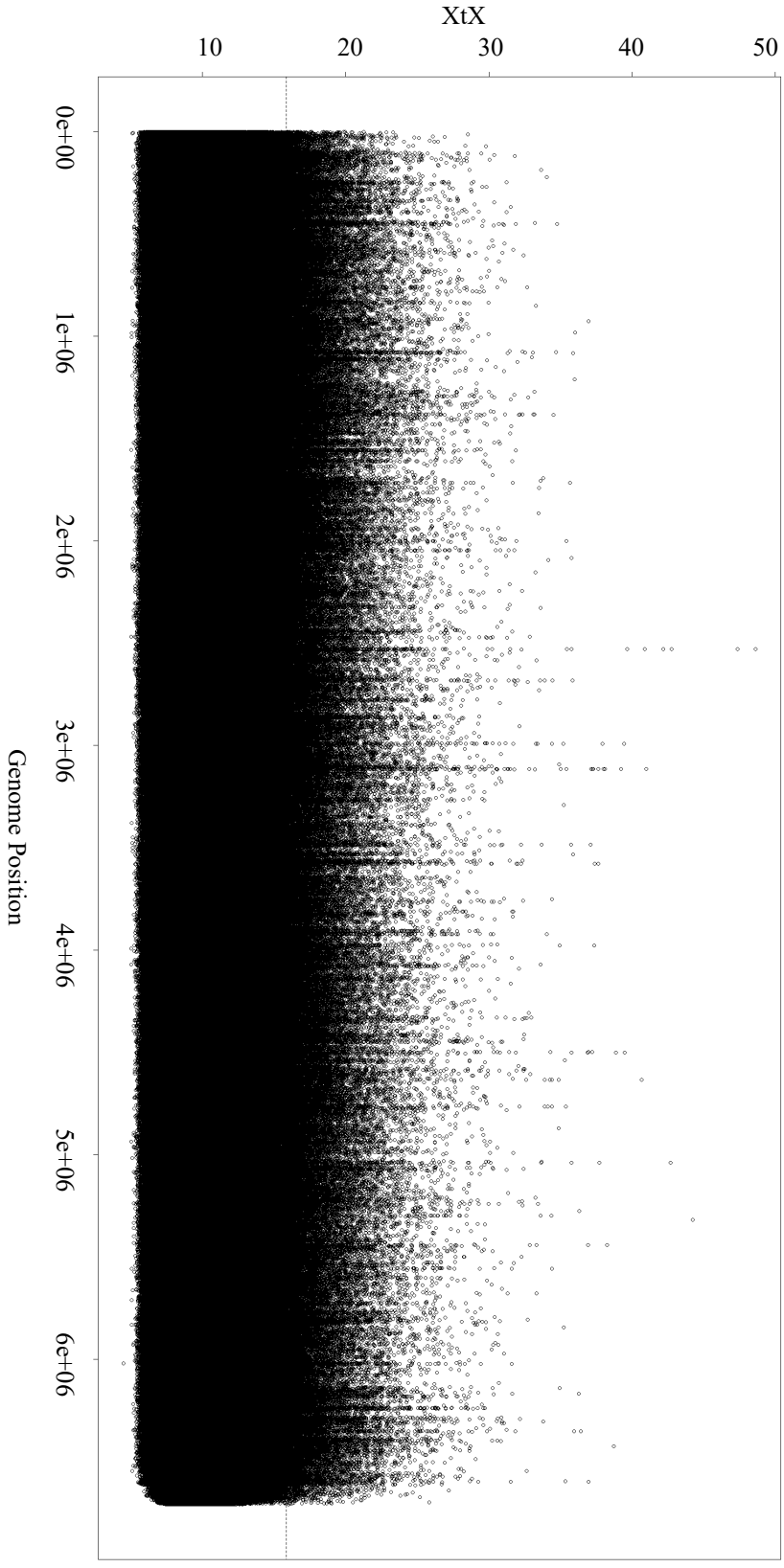

Distribution of proportion covered across all 158 BFR samples
number_of_samples
proportion_of_reference_covered
0.84
0.86
0.88
0.90
0.92
0.94
0.96
0.98
1.00

Figure S3.3 - XtX scores for variable sites across the genome. Points above the dotted line are considered putatively under selection. This threshold was determined by creating pseudo-observed datasets (PODs) from the data under the null model of no selection, using the value corresponding to the 99th percentile of the POD null distribution as the calibrated selection/neutrality threshold.


XtX
10
20
30
40
50
0e+00
1e+06
2e+06
3e+06
4e+06
5e+06
6e+06
Genome Position

List of Supplemental Files

The following supplementary files are included electronically:

"tables_s1.12-1.15_combined.xlsx"

This file contains the following:

- **Table S1.12** – Enriched gene sets between the pH 7.85 and pH 7.30 treatments in Experiment 1 (identified using FGSEA analysis).
- **Table S1.13** – Enriched gene sets between the static pH treatments (7.85 and 7.70) in Experiment 2 (identified using FGSEA analysis).
- **Table S1.14** – Enriched gene sets between the pH 7.85 static and pH 7.85 variable treatments in Experiment 2 (identified using FGSEA analysis).
- **Table S1.15** – Enriched gene sets between the pH 7.70 static and pH 7.70 variable treatments in Experiment 2 (identified using FGSEA analysis).

"supplementary_methods_ch2.docx"

This file contains the following:

- Detailed methods for the mitogenome assembly, removal of mitochondrial sequences from the nuclear genome assembly, and the splitting of a mis-join in the nuclear assembly.
- Accompanying images

"table_s2.2.xlsx"

This file contains the following:

- **Table S2.2** – Summary table of repeats identified in the genome assembly using RepeatMasker.

Literature Cited

Alonge, M., Lebeigle, L., Kirsche, M., Aganezov, S., Wang, X., Lippman, Z. B., Schatz, M. C., & Soyk, S. (2021). Automated assembly scaffolding elevates a new tomato system for high-throughput genome editing. *BioRxiv*, 2021.11.18.469135. https://doi.org/10.1101/2021.11.18.469135

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410.

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, *26*(1), 32–46.

Anderson, M. J. (2017). Permutational Multivariate Analysis of Variance (PERMANOVA). *Wiley StatsRef: Statistics Reference Online*. https://doi.org/10.1002/9781118445112.stat07841

Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]*. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrão, E. A., & De Clerck, O. (2018). Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. *Global Ecology and Biogeography*, *27*(3), 277–284. https://doi.org/https://doi.org/10.1111/geb.12693

Barrett, R. D. H., & Schluter, D. (2008). Adaptation from standing genetic variation. *Trends in Ecology and Evolution*, *23*(1), 38–44. https://doi.org/10.1016/j.tree.2007.09.008

Bednaršek, N., Calosi, P., Feely, R. A., Ambrose, R., Byrne, M., Chan, K. Y. K., Dupont, S., Padilla-Gamiño, J. L., Spicer, J. I., Kessouri, F., Roethler, M., Sutula, M., & Weisberg, S. B. (2021). Synthesis of Thresholds of Ocean Acidification Impacts on Echinoderms. *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.602601

Beever, E. A., Hall, L. E., Varner, J., Loosen, A. E., Dunham, J. B., Gahl, M. K., Smith, F. A., & Lawler, J. J. (2017). Behavioral flexibility as a mechanism for coping with climate change. *Frontiers in Ecology and the Environment*, *15*(6), 299–308. https://doi.org/10.1002/fee.1502

Bell, G. (2013). *Evolutionary rescue and the limits of adaptation*. *368*(1610), 1–6.

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., & Courchamp, F. (2012). Impacts of climate change on the future of biodiversity. *Ecology Letters*, *15*(4), 365–377. https://doi.org/10.1111/j.1461-0248.2011.01736.x

Berg, M. P., Toby Kiers, E., Driessen, G., van der Heijden, M., Kooi, B. W., Kuenen, F., Liefting, M., Verhoef, H. A., & Ellers, J. (2010). Adapt or disperse: Understanding species persistence in a changing world. *Global Change Biology*, *16*(2), 587–598. https://doi.org/10.1111/j.1365-2486.2009.02014.x

Bernardi, G. (2000). Barriers To Gene Flow in Embiotoca Jacksoni, a Marine Fish Lacking a Pelagic Larval Stage. *Evolution*, *54*(1), 226. https://doi.org/10.1554/0014-3820(2000)054[0226:BTGFIE]2.0.CO;2

Bernardi, G. (2005). Phylogeography and Demography of Sympatric Sister Surfperch Species, Embiotoca Jacksoni and E. Lateralis Along the California Coast: Historical Versus Ecological Factors. *Evolution*, *59*(2), 386–394. https://doi.org/10.1554/04-367

Bernardi, G., Toy, J. A., Escalona, M., Marimuthu, M. P. A., Sahasrabudhe, R., Nguyen, O., Sacco, S., Beraut, E., Toffelmier, E., Miller, C., & Shaffer, H. B. (2022). Reference genome of the Black Surfperch, Embiotoca jacksoni (Embiotocidae, Perciformes), a California kelp forest fish that lacks a pelagic larval stage. *Journal of Heredity*. https://doi.org/10.1093/jhered/esac034

Binns, D., Dimmer, E., Huntley, R., Barrell, D., O'Donovan, C., & Apweiler, R. (2009). QuickGO: a web-based tool for Gene Ontology searching. *Bioinformatics*, *25*(22), 3045–3046. https://doi.org/10.1093/bioinformatics/btp536

Bitter, M. C., Wong, J. M., Dam, H. G., Donelan, S. C., Kenkel, C. D., Komoroske, L. M., Nickols, K. J., Rivest, E. B., Salinas, S., Burgess, S. C., & Lotterhos, K. E. (2021). Fluctuating selection and global change: A synthesis and review on disentangling the roles of climate amplitude, predictability and novelty. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1957). https://doi.org/10.1098/rspb.2021.0727

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120.

Bosch, S., & Fernandez, S. (2022). *sdmpredictors: Species Distribution Modelling Predictor Datasets*. https://cran.r-project.org/package=sdmpredictors

Breitburg, D. L., Loher, T., Pacey, C. A., & Gerstein, A. (1997). Varying effects of low dissolved oxygen on trophic interactions in an estuarine food web. *Ecological Monographs*, *67*(4), 489–507. https://doi.org/10.1890/0012-9615(1997)067[0489:VEOLDO]2.0.CO;2

Buoro, M., & Carlson, S. M. (2014). Life-history syndromes: Integrating dispersal through space and time. *Ecology Letters*, *17*(6), 756–767.

https://doi.org/10.1111/ele.12275

Burton, R. S. (1998). Intraspecific phylogeography across the Point Conception biogeographic boundary. *Evolution*, *52*(3), 734–745. https://doi.org/https://doi.org/10.1111/j.1558-5646.1998.tb03698.x

Cabanettes, F., & Klopp, C. (2018). D-GENIES: dot plot large genomes in an interactive, efficient and simple way. *PeerJ*, *6*, e4958. https://doi.org/10.7717/peerj.4958

Carlson, S. M., Cunningham, C. J., & Westley, P. A. H. (2014). Evolutionary rescue in a changing world. *Trends in Ecology & Evolution*, *29*, 521–530. https://doi.org/10.1016/j.tree.2014.06.005

Castillo, K. D., Ries, J. B., Bruno, J. F., & Westfield, I. T. (2014). The reef-building coral *Siderastrea siderea* exhibits parabolic responses to ocean acidification and warming. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1797), 20141856. https://doi.org/10.1098/rspb.2014.1856

Challis, R., Richards, E., Rajan, J., Cochrane, G., & Blaxter, M. (2020). BlobToolKit – Interactive Quality Assessment of Genome Assemblies. *G3 Genes|Genomes|Genetics*, *10*(4), 1361–1374. https://doi.org/10.1534/g3.119.400908

Chan, F., Barth, J. A., Blanchette, C. A., Byrne, R. H., Chavez, F., Cheriton, O., Feely, R. A., Friederich, G., Gaylord, B., Gouhier, T., Hacker, S., Hill, T., Hofmann, G., McManus, M. A., Menge, B. A., Nielsen, K. J., Russell, A., Sanford, E., Sevadjian, J., & Washburn, L. (2017). Persistent spatial structuring of coastal ocean acidification in the California Current System. *Scientific Reports*, *7*(1), 1–7. https://doi.org/10.1038/s41598-017-02777-y

Chan, F., Barth, J. A., Blanchette, C. A., Byrne, R. H., Chavez, F., Cheriton, O., Feely, R. A., Friederich, G., Gaylord, B., Gouhier, T., & others. (2017). Persistent spatial structuring of coastal ocean acidification in the California Current System. *Scientific Reports*, *7*(1), 1–7.

Chavez, F. P., Sevadjian, J., Wahl, C., Friederich, J., & Friederich, G. E. (2018). Measurements of pCO2 and pH from an autonomous surface vehicle in a coastal upwelling system. *Deep Sea Research Part II: Topical Studies in Oceanography*, *151*, 137–146. https://doi.org/10.1016/J.DSR2.2017.01.001

Checkley, D. M., & Barth, J. A. (2009). Patterns and processes in the California Current System. *Progress in Oceanography*, *83*(1), 49–64. https://doi.org/https://doi.org/10.1016/j.pocean.2009.07.028

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Chevin, L. M. (2012). Genetic constraints on adaptation to a changing environment. *Evolution*, *67*(3), 708–721. https://doi.org/10.1111/j.1558-5646.2012.01809.x

Clark, T. D., Raby, G. D., Roche, D. G., Binning, S. A., Speers-Roesch, B., Jutfelt, F., & Sundin, J. (2020a). Ocean acidification does not impair the behaviour of coral reef fishes. *Nature*, *577*(7790), 370–375. https://doi.org/10.1038/s41586-019-1903-y

Clark, T. D., Raby, G. D., Roche, D. G., Binning, S. A., Speers-Roesch, B., Jutfelt, F., & Sundin, J. (2020b). Reply to: Methods matter in repeating ocean acidification studies. *Nature*, *586*(7830), E25–E27. https://doi.org/10.1038/s41586-020-2804-9

Clemens, W. A., & Wilby, G. V. (1961). *Fishes of the Pacific coast of Canada* (Vol. 68). Fisheries Research Board of Canada Ottawa.

Cline, A. J., Hamilton, S. L., & Logan, C. A. (2020). Effects of multiple climate change stressors on gene expression in blue rockfish (Sebastes mystinus). *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, *239*, 110580.

Conover, D. O., Clarke, L. M., Munch, S. B., & Wagner, G. N. (2006). Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. *Journal of Fish Biology*, *69*(SUPPL. C), 21–47. https://doi.org/10.1111/j.1095-8649.2006.01274.x

Csepp, D. J., & Wing, B. L. (1999). Northern Range Extensions and Habitat Observations for Blackeye Goby Rhinogobiops nicholsii and Kelp Perch Brachyistius frenatus in Southeastern Alaska. *Alaska Fishery Research Bulletin*, *6*(2), 78–84. http://www.adfg.alaska.gov/FedAidpdfs/AFRB.06.2.078-084.pdf

Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Bennett, R., Berry, A., Bhai, J., Bignell, A., Billis, K., Boddu, S., Brooks, L., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., … Flicek, P. (2022). Ensembl 2022. *Nucleic Acids Research*, *50*(D1), D988–D995. https://doi.org/10.1093/nar/gkab1049

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2), giab008. https://doi.org/10.1093/gigascience/giab008

Davenport, A. C., & Anderson, T. W. (2007). Positive Indirect Effects of Reef Fishes on Kelp Performance : The Importance of Mesograzers Published by : Ecological Society of America content in a trusted digital archive . We use information technology and tools to increase productivity and facilitat. *Ecology*, *88*(6), 1548–1561.

Dawson, M. N. (2001). Phylogeography in coastal marine animals: a solution from California? *Journal of Biogeography*, *28*(6), 723–736. https://doi.org/https://doi.org/10.1046/j.1365-2699.2001.00572.x

DelValls, T. A., & Dickson, A. G. (1998). The pH of buffers based on 2-amino-2-hydroxymethyl-1,3-propanediol ('tris') in synthetic sea water. *Deep Sea Research Part I: Oceanographic Research Papers*, *45*(9), 1541–1554. https://doi.org/10.1016/S0967-0637(98)00019-3

Dhabhar, F. S. (2014). Effects of stress on immune function: the good, the bad, and the beautiful. *Immunologic Research*, *58*(2), 193–210. https://doi.org/10.1007/s12026-014-8517-0

Dickson, A. G., Sabine, C. L., & Christian, J. R. (2007). Guide to best practices for ocean CO2 measurements. *PICES Special Publication 3*, *3*(8), 191. https://doi.org/10.1159/000331784

Domenici, P., Allan, B., McCormick, M. I., & Munday, P. L. (2012). Elevated carbon dioxide affects behavioural lateralization in a coral reef fish. *Biology Letters*, *8*(1), 78–81. https://doi.org/10.1098/rsbl.2011.0591

Donham, E. M., Flores, I., Hooper, A., O'Brien, E., Vylet, K., Takeshita, Y., Freiwald, J., & Kroeker, K. J. (n.d.). *Population-specific vulnerability to ocean change in a multi-stressor environment*.

Duarte, C. M., Hendriks, I. E., Moore, T. S., Olsen, Y. S., Steckbauer, A., Ramajo, L., Carstensen, J., Trotter, J. A., & McCulloch, M. (2013). Is Ocean Acidification an Open-Ocean Syndrome? Understanding Anthropogenic Impacts on Seawater pH. *Estuaries and Coasts*, *36*(2), 221–236. https://doi.org/10.1007/s12237-013-9594-3

Eirin-Lopez, J. M., & Putnam, H. M. (2018). Marine Environmental Epigenetics. *Annual Review of Marine Science*, *11*(1), 335–368. https://doi.org/10.1146/annurev-marine-010318-095114

Eittreim, S. L., Anima, R. J., & Stevenson, A. J. (2002). Seafloor geology of the Monterey Bay area continental shelf. *Marine Geology*, *181*(1), 3–34. https://doi.org/https://doi.org/10.1016/S0025-3227(01)00259-6

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Feely, R. A., Sabine, C. L., Hernandez-Ayon, J. M., Ianson, D., & Hales, B. (2008). Evidence for upwelling of corrosive "acidified" water onto the continental shelf. *Science*, *320*(5882), 1490–1492. https://doi.org/10.1126/science.1155676

Fox, E. A., Wright, A. E., Fumagalli, M., & Vieira, F. G. (2019). ngsLD: evaluating linkage disequilibrium using genotype likelihoods. *Bioinformatics*, *35*(19), 3855–3856. https://doi.org/10.1093/bioinformatics/btz200

Galinsky, K. J., Bhatia, G., Loh, P.-R., Georgiev, S., Mukherjee, S., Patterson, N. J., & Price, A. L. (2016). Fast Principal-Component Analysis Reveals Convergent Evolution of ADH1B in Europe and East Asia. *The American Journal of Human Genetics*, *98*(3), 456–472. https://doi.org/10.1016/j.ajhg.2015.12.022

Gautier, M. (2015). Genome-Wide Scan for Adaptive Divergence and Association with Population-Specific Covariates. *Genetics*, *201*(4), 1555–1579. https://doi.org/10.1534/genetics.115.181453

Ghalambor, C. K., McKay, J. K., Carroll, S. P., & Reznick, D. N. (2007). Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology*, *21*(3), 394–407. https://doi.org/10.1111/j.1365-2435.2007.01283.x

Ghezelayagh, A., Harrington, R. C., Burress, E. D., Campbell, M. A., Buckner, J. C., Chakrabarty, P., Glass, J. R., McCraney, W. T., Unmack, P. J., Thacker, C. E., Alfaro, M. E., Friedman, S. T., Ludt, W. B., Cowman, P. F., Friedman, M., Price, S. A., Dornburg, A., Faircloth, B. C., Wainwright, P. C., & Near, T. J. (2022). Prolonged morphological expansion of spiny-rayed fishes following the end-Cretaceous. *Nature Ecology & Evolution*. https://doi.org/10.1038/s41559-022-01801-3

Gomulkiewicz, R., & Holt, R. D. (1995). When does Evolution by Natural Selection Prevent Extinction? *Evolution*, *49*(1), 201–207. https://doi.org/10.2307/2410305

Goncalves, P., Anderson, K., Thompson, E. L., Melwani, A., Parker, L. m., Ross, P. M., & Raftos, D. A. (2016). Rapid transcriptional acclimation following transgenerational exposure of oysters to ocean acidification. *Molecular Ecology*, *25*(19), 4836–4849. https://doi.org/10.1111/mec.13808

Griffiths, J. S., Pan, T.-C. F., & Kelly, M. W. (2019). Differential responses to ocean acidification between populations of Balanophyllia elegans corals from high and

low upwelling environments. *Molecular Ecology*, *28*(11), 2715–2730.

Gruber, N., Hauri, C., Lachkar, Z., Loher, D., Frolicher, T. L., & Plattner, G.-K. (2012). Rapid Progression of Ocean Acidification in the California Current System. *Science*, *337*(6091), 220–223. https://doi.org/10.1126/science.1216773

Günther, T., & Coop, G. (2013). Robust identification of local adaptation from allele frequencies. *Genetics*, *195*(1), 205–220. https://doi.org/10.1534/genetics.113.152462

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., Macmanes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., … Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494–1512. https://doi.org/10.1038/nprot.2013.084

Hamilton, S. L., Logan, C. A., Fennie, H. W., Sogard, S. M., Barry, P., Makukhov, A. D., Tobosa, L. R., Boyer, K., Lovera, C. F., & Bernardi, G. (2017). Species-Specific Responses of Juvenile Rockfish to Elevated pCO2 : From Behavior to Genomics. *PLoS ONE*, *12*(1), e0169670. https://doi.org/10.1371/journal.pone.0169670

Hamilton, T. J., Holcombe, A., & Tresguerres, M. (2013). CO2-induced ocean acidification increases anxiety in rockfish via alteration of GABAA receptor functioning. *Proceedings of the Royal Society B: Biological Sciences*, *281*. https://doi.org/10.1121/1.4929899

Harms, L., Frickenhaus, S., Schiffer, M., Mark, F. C., Storch, D., Held, C., Pörtner, H.-O., & Lucassen, M. (2014). Gene expression profiling in gills of the great spider crab Hyas araneus in response to ocean acidification and warming. *BMC Genomics*, *15*(1), 789. https://doi.org/10.1186/1471-2164-15-789

Hauri, C., Gruber, N., Vogt, M., Doney, S. C., Feely, R. A., Lachkar, Z., Leinweber, A., McDonnell, A. M. P., & Munnich, M. (2013). Spatiotemporal variability and long-term trends of ocean acidification in the California Current System. *Biogeosciences*, *10*(1), 193–216. https://doi.org/10.5194/bg-10-193-2013

Heuer, R. M., & Grosell, M. (2014). Physiological impacts of elevated carbon dioxide and ocean acidification on fish. *American Journal of Physiology - Regulatory Integrative and Comparative Physiology*, *307*(9), R1061–R1084. https://doi.org/10.1152/ajpregu.00064.2014

Hijmans, R. J. (2021). *geosphere: Spherical Trigonometry*. https://cran.r-project.org/package=geosphere

Hirsh, H. K., Nickols, K. J., Takeshita, Y., Traiger, S. B., Mucciarone, D. A., Monismith, S., & Dunbar, R. B. (2020). Drivers of Biogeochemical Variability in a Central California Kelp Forest: Implications for Local Amelioration of Ocean Acidification. *Journal of Geophysical Research: Oceans*, *125*(11), 1–22. https://doi.org/10.1029/2020JC016320

Hixon, M. A. (1981). An Experimental Analysis of Territoriality in the California Reef Fish Embiotoca jacksoni (Embiotocidae). *Copeia*, *3*, 653–665. https://doi.org/10.2307/1444571

Hobday, A. J. (2000a). Abundance and dispersal of drifting kelp Macrocystis pyrifera rafts in the Southern California Bight. *Marine Ecology Progress Series*, *195*, 101–116. https://doi.org/doi:10.3354/meps195101

Hobday, A. J. (2000b). Persistence and transport of fauna on drifting kelp (Macrocystis pyrifera (L.) C. Agardh) rafts in the Southern California Bight. *Journal of Experimental Marine Biology and Ecology*, *253*(1), 75–96. https://doi.org/https://doi.org/10.1016/S0022-0981(00)00250-1

Hofmann, G. E., Evans, T. G., Kelly, M. W., Padilla-Gamiño, J. L., Blanchette, C. A., Washburn, L., Chan, F., McManus, M. A., Menge, B. A., Gaylord, B., Hill, T. M., Sanford, E., Lavigne, M., Rose, J. M., Kapsenberg, L., & Dutton, J. M. (2014). Exploring local adaptation and the ocean acidification seascape - Studies in the California Current Large Marine Ecosystem. *Biogeosciences*, *11*(4), 1053–1064. https://doi.org/10.5194/bg-11-1053-2014

Hofmann, Gretchen E. (2017). Ecological Epigenetics in Marine Metazoans. *Frontiers in Marine Science*, *4*(January), 1–7. https://doi.org/10.3389/fmars.2017.00004

Hofmann, G. E., Smith, J. E., Johnson, K. S., Send, U., Levin, L. A., Micheli, F., Paytan, A., Price, N. N., Peterson, B., Takeshita, Y., Matson, P. G., Crook, E. D., Kroeker, K. J., Gambi, M. C., Rivest, E. B., Frieder, C. A., Yu, P. C., & Martz, T. R. (2011). High-Frequency Dynamics of Ocean pH: A Multi-Ecosystem Comparison. *PLOS ONE*, *6*(12), 1–11. https://doi.org/10.1371/journal.pone.0028983

Holbrook, J. R., Muench, R. D., Kachel, D. G., & Wright, C. (1980). *Circulation in the Strait of Juan de Fuca: Recent Oceanographic Observations in the Eastern Basin*.

Holt, R. D. (1990). The microevolutionary consequences of climate change. *Trends in Ecology and Evolution*, *5*(9), 311–315. https://doi.org/10.1016/0169-5347(90)90088-U

Hoshijima, U., & Hofmann, G. E. (2019). Variability of Seawater Chemistry in a Kelp Forest Environment Is Linked to in situ Transgenerational Effects in the Purple Sea Urchin, Strongylocentrotus purpuratus. *Frontiers in Marine Science*, *6*(March), 1–18. https://doi.org/10.3389/fmars.2019.00062

Hubbs, C. L., & Hubbs, L. C. (1954). Data on the life history, variation, ecology, and relationships of the kelp perch, Brachyistius frenatus, an embiotocid fish of the Californias. *California Fish and Game*, *40*(2), 183–198.

Huyer, A. (1983). Coastal upwelling in the California current system. *Progress in Oceanography*, *12*(3), 259–284. https://doi.org/https://doi.org/10.1016/0079-6611(83)90010-1

Iwasaki, W., Fukunaga, T., Isagozawa, R., Yamada, K., Maeda, Y., Satoh, T. P., Sado, T., Mabuchi, K., Takeshima, H., Miya, M., & Nishida, M. (2013). MitoFish and MitoAnnotator: A Mitochondrial Genome Database of Fish with an Accurate and Automatic Annotation Pipeline. *Molecular Biology and Evolution*, *30*(11), 2531–2540. https://doi.org/10.1093/molbev/mst141

Jarrold, M. D., Humphrey, C., McCormick, M. I., & Munday, P. L. (2017). Diel CO2 cycles reduce severity of behavioural abnormalities in coral reef fish under ocean acidification. *Scientific Reports*, *7*(1), 10153. https://doi.org/10.1038/s41598-017-10378-y

Jarrold, M. D., & Munday, P. L. (2019). Diel CO2 cycles and parental effects have similar benefits to growth of a coral reef fish under ocean acidification. *Biology Letters*, *15*(2). https://doi.org/10.1098/rsbl.2018.0724

Jones, A. G. (2008). A theoretical quantitative genetic study of negative ecological interactions and extinction times in changing environments. *BMC Evolutionary Biology*, *8*(1), 1–10. https://doi.org/10.1186/1471-2148-8-119

Jordan, D. S., & Starks, E. C. (1895). *The fishes of Puget Sound* (Vol. 3). Leland Stanford Junior University. Hopkins Laboratory of Biology.

Jun, G., Wing, M. K., Abecasis, G. R., & Kang, H. M. (2015). An efficient and scalable analysis framework for variant extraction and refinement from population scale DNA sequence data. *Genome Research* . http://genome.cshlp.org/content/early/2015/04/14/gr.176552.114.abstract

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, *28*(1), 27–30. https://doi.org/10.1093/nar/28.1.27

Kang, J., Nagelkerken, I., Rummer, J. L., Rodolfo-Metalpa, R., Munday, P. L.,

Ravasi, T., & Schunter, C. (2022). Rapid evolution fuels transcriptional plasticity to ocean acidification. *Global Change Biology*, *28*(9), 3007–3022. https://doi.org/10.1111/gcb.16119

Kapsenberg, L., Bockmon, E. E., Bresnahan, P. J., Kroeker, K. J., Gattuso, J. P., & Martz, T. R. (2017). Advancing ocean acidification biology using Durafet® pH electrodes. *Frontiers in Marine Science*, *4*(OCT), 1–9. https://doi.org/10.3389/fmars.2017.00321

Kindinger, T. L., Toy, J. A., & Kroeker, K. J. (2022). Emergent effects of global change on consumption depend on consumers and their resources in marine systems. *Proceedings of the National Academy of Sciences*, *119*(18), e2108878119. https://doi.org/10.1073/pnas.2108878119

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, *15*(1), 356. https://doi.org/10.1186/s12859-014-0356-4

Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M. N., & Sergushichev, A. (2021). Fast gene set enrichment analysis. *BioRxiv*. https://doi.org/10.1101/060012

Kroeker, K. J., Bell, L. E., Donham, E. M., Hoshijima, U., Lummis, S., Toy, J. A., & Willis-Norton, E. (2020). Ecological change in dynamic environments: Accounting for temporal environmental variability in studies of ocean change biology. *Global Change Biology*, *26*(1), 54–67. https://doi.org/10.1111/gcb.14868

Kroeker, K. J., Donham, E. M., Vylet, K., Warren, J. K., Cheresh, J., Fiechter, J., Freiwald, J., & Takeshita, Y. (n.d.). *Exposure to extremes in multiple global change drivers: characterizing pH, dissolved oxygen and temperature variability in a dynamic, upwelling dominated ecosystem.*

Kroeker, K. J., Kordas, R. L., Crim, R., Hendriks, I. E., Ramajo, L., Singh, G. S., Duarte, C. M., & Gattuso, J.-P. (2013). Impacts of ocean acidification on marine organisms: quantifying sensitivities and interaction with warming. *Global Change Biology*, *19*(6), 1884–1896. https://doi.org/10.1111/gcb.12179

Kroeker, K. J., Powell, C., & Donham, E. M. (2021). Windows of vulnerability: Seasonal mismatches in exposure and resource identity determine ocean acidification's effect on a primary consumer at high latitude. *Global Change Biology*, *27*(5), 1042–1051. https://doi.org/https://doi.org/10.1111/gcb.15449

Kültz, D. (2005). Molecular and evolutionary basis of the cellular stress response. *Annual Review of Physiology*, *67*(1), 225–257.

https://doi.org/10.1146/annurev.physiol.67.040403.103635

Kwan, G. T., Hamilton, T. J., & Tresguerres, M. (2017). CO2-induced ocean acidification does not affect individual or group behaviour in a temperate damselfish CO 2 -induced ocean acidification does not affect individual or group behaviour in a temperate damselfish. *Royal Society Open Science*, *4*(7), 170283. https://doi.org/10.1098/rsos.170283

Lai, F., Fagernes, C. E., Jutfelt, F., & Nilsson, G. E. (2016). Expression of genes involved in brain GABAergic neurotransmission in three-spined stickleback exposed to near-future CO2. *Conservation Physiology*, *4*(1). https://doi.org/10.1093/conphys/cow068

Lande, R., & Shannon, S. (1996). The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*, *50*(1), 434–437.

Lanfear, R., Kokko, H., & Eyre-Walker, A. (2014). Population size and the rate of evolution. *Trends in Ecology and Evolution*, *29*(1), 33–41. https://doi.org/10.1016/j.tree.2013.09.009

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3), 1–10.

Laur, D. R., & Ebeling, A. W. (1983). Predator-prey relationships in surfperches. In D. L. G. Noakes, D. G. Lindquist, G. S. Helfman, & J. A. Ward (Eds.), *Predators and prey in fishes: Proceedings of the 3rd biennial conference on the ethology and behavioral ecology of fishes, held at Normal, Illinois, U.S.A., May 19--22, 1981* (pp. 55–67). Springer Netherlands. https://doi.org/10.1007/978-94-009-7296-4_7

Leis, J. M. (1991). The Pelagic Stage of Reef Fishes. In P. F. Sale (Ed.), *The Ecology of Fishes on Coral Reefs.* (pp. 182–229). Academic Press.

Li, B., & Dewey, C. N. (2011). RSEM: Accurate transcript quantification from RNA-seq data with or without a reference genome. *Bioinformatics: The Impact of Accurate Quantification on Proteomic and Genetic Analysis and Research*, *12*(323). https://doi.org/10.1201/b16589

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, *34*(18), 3094–3100. https://doi.org/10.1093/bioinformatics/bty191

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., & Tamayo, P. (2015). The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Systems*, *1*(6), 417–425. https://doi.org/10.1016/j.cels.2015.12.004

Longo, G., & Bernardi, G. (2015). The evolutionary history of the embiotocid surfperch radiation based on genome-wide RAD sequence data. *Molecular Phylogenetics and Evolution*, *88*, 55–63. https://doi.org/10.1016/j.ympev.2015.03.027

Longo, G. C., O'Connell, B., Green, R. E., & Bernardi, G. (2016). The complete mitochondrial genome of the black surfperch, Embiotoca jacksoni: Selection and substitution rates among surfperches (Embiotocidae). *Marine Genomics*. https://doi.org/10.1016/j.margen.2016.03.006

Love, M. (2011). *Certainly more than you want to know about the fishes of the Pacific Coast: a postmodern experience*. Really Big Press.

Lowe, A. T., Bos, J., & Ruesink, J. (2019). Ecosystem metabolism drives pH variability and modulates long-term ocean acidification in the Northeast Pacific coastal ocean. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-018-37764-4

Lynch, M., & Lande, R. (1993). Evolution and Extinction in response to environmental change. In P. M. Kareiva, J. G. Kingsolver, & R. B. Huey (Eds.), *Biotic Interactions and Global Change* (pp. 234–250). Sinauer Associates.

Magnhagen, C. (1991). Predation risk as a cost of reproduction. *Trends in Ecology & Evolution*, *6*(6), 183–186. https://doi.org/https://doi.org/10.1016/0169-5347(91)90210-O

Martinez Arbizu, P. (2020). *pairwiseAdonis: Pairwise Multilevel Comparison using Adonis*.

Martz, T. R., Connery, J. G., & Johnson, K. S. (2010). Testing the Honeywell Durafet® for seawater pH applications. *Limnology and Oceanography: Methods*, *8*(MAY), 172–184. https://doi.org/10.4319/lom.2010.8.172

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* , *20*(9), 1297–1303.

http://genome.cshlp.org/content/20/9/1297.abstract

McPherson, M. L., Finger, D. J. I., Houskeeper, H. F., Bell, T. W., Carr, M. H., Rogers-Bennett, L., & Kudela, R. M. (2021). Large-scale shift in the structure of a kelp forest ecosystem co-occurs with an epizootic and marine heatwave. *Communications Biology*, *4*(1), 298. https://doi.org/10.1038/s42003-021-01827-6

Mecklenburg, C. W., Mecklenburg, T. A., & Thorsteinson, L. K. (2002). *Fishes of Alaska*. American Fisheries Society.

Meisner, J., & Albrechtsen, A. (2018). Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data. *Genetics*, *210*(2), 719–731. https://doi.org/10.1534/genetics.118.301336

Meisner, J., Albrechtsen, A., & Hanghøj, K. (2021). Detecting selection in low-coverage high-throughput sequencing data using principal component analysis. *BMC Bioinformatics*, *22*(1), 470. https://doi.org/10.1186/s12859-021-04375-2

Miller, D. J., & Lea, R. N. (1976). *Guide to the coastal marine fishes of California* (Issues 154–158). UCANR Publications.

Mitchell, C. T., & Hunter, J. R. (1970). Fishes associated with drifting kelp, Macrocystis pyrifera, off the coast of Southern California and Northern Baja California. *California Fish and Game*, *56*(4), 288–297.

Moore, S. K., Mantua, N. J., Newton, J. A., Kawase, M., Warner, M. J., & Kellogg, J. P. (2008). A descriptive analysis of temporal and spatial patterns of variability in Puget Sound oceanographic properties. *Estuarine, Coastal and Shelf Science*, *80*(4), 545–554. https://doi.org/https://doi.org/10.1016/j.ecss.2008.09.016

Munday, P. L., Dixson, D. L., Mccormick, M. I., Meekan, M., Ferrari, M. C. O., Chivers, D. P., & Karl, D. (2010). Replenishment of fish populations is threatened by ocean acidification. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(29), 12930–12934. https://doi.org/10.1073/pnas.1004519107/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1004519107

Munday, P. L., Dixson, D. L., Welch, M. J., Chivers, D. P., Domenici, P., Grosell, M., Heuer, R. M., Jones, G. P., McCormick, M. I., Meekan, M., Nilsson, G. E., Ravasi, T., & Watson, S.-A. (2020). Methods matter in repeating ocean acidification studies. *Nature*, *586*(7830), E20–E24. https://doi.org/10.1038/s41586-020-2803-x

Muralidharan, S., & Mandrekar, P. (2013). Cellular stress response and innate

immune signaling: integrating pathways in host defense and inflammation. *Journal of Leukocyte Biology*, *94*(6), 1167–1184. https://doi.org/https://doi.org/10.1189/jlb.0313153

Nagelkerken, I., & Munday, P. L. (2016). Animal behaviour shapes the ecological effects of ocean acidification and warming: moving from individual to community-level responses. *Global Change Biology*, *22*(3), 974–989. https://doi.org/https://doi.org/10.1111/gcb.13167

Nardone, A., Ronchi, B., Lacetera, N., Ranieri, M. S., & Bernabucci, U. (2010). Effects of climate changes on animal production and sustainability of livestock systems. *Livestock Science*, *130*(1), 57–69. https://doi.org/https://doi.org/10.1016/j.livsci.2010.02.011

Nicholson, T. E., Mayer, K. A., Staedler, M. M., Fujii, J. A., Murray, M. J., Johnson, A. B., Tinker, M. T., & Van Houtan, K. S. (2018). Gaps in kelp cover may threaten the recovery of California sea otters. *Ecography*, *41*(11), 1751–1762. https://doi.org/https://doi.org/10.1111/ecog.03561

Nilsson, G. E., Dixson, D. L., Domenici, P., McCormick, M. I., Sørensen, C., Watson, S.-A., & Munday, P. L. (2012). Near-future carbon dioxide levels alter fish behaviour by interfering with neurotransmitter function. *Nature Climate Change*, *2*(3), 201–204. https://doi.org/10.1038/nclimate1352

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2020). *vegan: Community Ecology Package*. https://cran.r-project.org/package=vegan

Orr, H. A., & Unckless, R. L. (2008). Population Extinction and the Genetics of Adaptation. *The American Naturalist*, *172*(2), 160–169. https://doi.org/10.1086/589460

Osmond, M. M., & de Mazancourt, C. (2013). How competition affects evolutionary rescue. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1610). https://doi.org/10.1098/rstb.2012.0085

Otto, S. P. (2004). Two steps forward, one step back: The pleiotropic effects of favoured alleles. *Proceedings of the Royal Society B: Biological Sciences*, *271*(1540), 705–714. https://doi.org/10.1098/rspb.2003.2635

Palumbi, S. R. (1992). Marine speciation on a small planet. *Trends in Ecology & Evolution*, *7*(4), 114–118. https://doi.org/https://doi.org/10.1016/0169-5347(92)90144-Z

Palumbi, S. R. (2003). Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications*, *13*(sp1), 146–158. https://doi.org/https://doi.org/10.1890/1051-0761(2003)013[0146:PGDCAT]2.0.CO;2

Palumbi, S. R. (2004). Marine Reserves and Ocean Neighborhoods: The Spatial Scale of Marine Populations and Their Management. *Annual Review of Environment and Resources*, *29*(1), 31–68. https://doi.org/10.1146/annurev.energy.29.062403.102254

Picard toolkit. (2019). In *Broad Institute, GitHub repository*. Broad Institute.

Pistevos, J. C. A., Nagelkerken, I., Rossi, T., Olmos, M., & Connell, S. D. (2015). Ocean acidification and global warming impair shark hunting behaviour and growth. *Scientific Reports*, *5*(1), 16293. https://doi.org/10.1038/srep16293

Precht, H., Christophersen, J., Hensel, H., & Larcher, W. (1973). *Temperature and Life*. Springer, Berlin, Heidelberg. https://doi.org/https://doi.org/10.1007/978-3-642-65708-5

Queitsch, C., Sangster, T. A., & Lindquist, S. (2002). Hsp90 as a capacitor of phenotypic variation. *Nature*, *417*(6889), 618–624. https://doi.org/10.1038/nature749

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., & Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research*, *47*(W1), W191–W198. https://doi.org/10.1093/nar/gkz369

Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-lopes, C., Rostamianfar, A., Wadi, L., Meyer, M., Wong, J., Xu, C., Merico, D., & Bader, G. D. (2019). *Pathway enrichment analysis and visualization of omics data using g : Profiler, GSEA, Cytoscape and EnrichmentMap*. *14*(February).

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Root, T. L., Price, J. T., Hall, K. R., & Schneider, S. H. (2003). Fingerprints of global warming on wild animals and plants - Supplementary Information. *Nature*, *421*(tier 2), 1–35. https://doi.org/10.1038/nature01309.1.

Rousset, F. (1997). Genetic Differentiation and Estimation of Gene Flow from F-Statistics Under Isolation by Distance. *Genetics*, *145*(4), 1219–1228. https://doi.org/10.1093/genetics/145.4.1219

Ruan, J., & Li, H. (2020). Fast and accurate long-read assembly with wtdbg2. *Nature Methods*, *17*(2), 155–158. https://doi.org/10.1038/s41592-019-0669-3

Rutherford, S. L. (2000). From genotype to phenotype: buffering mechanisms and the storage of genetic information. *Bioessays*, *22*(12), 1095–1105.

Rutherford, S. L. (2003). Between genotype and phenotype: protein chaperones and evolvability. *Nature Reviews Genetics*, *4*(4), 263–274. https://doi.org/10.1038/nrg1041

Rutherford, S. L., & Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. *Nature*, *396*(6709), 336–342. https://doi.org/10.1038/24550

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual* (2nd ed.). Cold Spring Harbor Laboratory Press.

Sandblom, E., Gräns, A., Axelsson, M., & Seth, H. (2014). Temperature acclimation rate of aerobic scope and feeding metabolism in fishes: Implications in a thermally extreme future. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1794). https://doi.org/10.1098/rspb.2014.1490

Sanford, E., & Kelly, M. W. (2011). *Local Adaptation in Marine Invertebrates*. https://doi.org/10.1146/annurev-marine-120709-142756

Schoener, T. W. (2011). *The Newest Synthesis : Understanding Ecological Dynamics*. *426*. https://doi.org/10.1126/science.1193954

Schunter, C., Jarrold, M. D., Munday, P. L., & Ravasi, T. (2021). Diel pCO2 fluctuations alter the molecular response of coral reef fishes to ocean acidification conditions. *Molecular Ecology*, *30*(20), 5105–5118. https://doi.org/https://doi.org/10.1111/mec.16124

Schunter, C., Ravasi, T., Munday, P. L., & Nilsson, G. E. (2019). Neural effects of elevated CO2 in fish may be amplified by a vicious cycle. *Conservation Physiology*, *7*(1). https://doi.org/10.1093/conphys/coz100

Schunter, C., Welch, M. J., Nilsson, G. E., Rummer, J. L., Munday, P. L., & Ravasi, T. (2018). An interplay between plasticity and parental phenotype determines impacts of ocean acidification on a reef fish. *Nature Ecology & Evolution*, *2*(2), 334–342. https://doi.org/10.1038/s41559-017-0428-8

Schunter, C., Welch, M. J., Ryu, T., Zhang, H., Berumen, M. L., Nilsson, G. E., Munday, P. L., & Ravasi, T. (2016). Molecular signatures of transgenerational response to ocean acidification in a species of reef fish. *Nature Climate Change*, *6*(11), 1014–1018. https://doi.org/10.1038/nclimate3087

Seebacher, F., White, C. R., & Franklin, C. E. (2015). *Physiological plasticity increases resilience of ectothermic animals to climate change*. *5*(January), 61–66. https://doi.org/10.1038/NCLIMATE2457

Selkoe, K. A., Henzler, C. M., & Gaines, S. D. (2008). Seascape genetics and the spatial ecology of marine populations. *Fish and Fisheries*, *9*(4), 363–377. https://doi.org/https://doi.org/10.1111/j.1467-2979.2008.00300.x

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V, & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, *236*(4803), 787–792. https://doi.org/10.1126/science.3576198

Smith, J. M. (1989). The causes of extinction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences1*, *325*, 241–252.

Snell-Rood, E. C., Kobiela, M. E., Sikkink, K. L., & Shephard, A. M. (2018). Mechanisms of Plastic Rescue in Novel Environments. *Annual Review of Ecology, Evolution, and Systematics*, *49*(1), 331–354. https://doi.org/10.1146/annurev-ecolsys-110617-062622

Stanke, M., Diekhans, M., Baertsch, R., & Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics*, *24*(5), 637–644. https://doi.org/10.1093/bioinformatics/btn013

Starks, E. C., & Morris, E. L. (1907). *The marine fishes of southern California*. University of California Press.

Strader, M. E., Wong, J. M., & Hofmann, G. E. (2020). Ocean acidification promotes broad transcriptomic responses in marine metazoans: a literature survey. *Frontiers in Zoology*, *17*(1), 7. https://doi.org/10.1186/s12983-020-0350-9

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & others. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, *102*(43), 15545–15550.

Sunday, J. M., Calosi, P., Dupont, S., Munday, P. L., Stillman, J. H., & Reusch, T. B. H. (2014). Evolution in an acidifying ocean. *Trends in Ecology and Evolution*, *29*(2), 117–125. https://doi.org/10.1016/j.tree.2013.11.001

Tajima, F. (1989). The effect of change in population size on DNA polymorphism. *Genetics*, *123*(3), 597–601. https://doi.org/10.1093/genetics/123.3.597

Takeshita, Y., Frieder, C. A., Martz, T. R., Ballard, J. R., Feely, R. A., Kram, S., Nam, S., Navarro, M. O., Price, N. N., & Smith, J. E. (2015). *Including high-frequency variability in coastal ocean acidification projections*. 5853–5870. https://doi.org/10.5194/bg-12-5853-2015

Tarp, F. H. (1952). *Fish Bulletin No. 88. A Revision of the Family Embiotocidae (The Surfperches)*. https://escholarship.org/uc/item/3qx7s3cn

The Gene Ontology Consortium. (2020). The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Research*, *49*(D1), D325–D334. https://doi.org/10.1093/nar/gkaa1113

The UniProt Consortium. (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research*, *49*(D1), D480--D489.

Thompson, J. N. (1998). Rapid evolution as an ecological process. *Trends in Ecology & Evolution*, *13*(8), 329–332.

Toy, J. A., & Bernardi, G. (n.d.). *A high-quality reference genome of the kelp surfperch, Brachyistius frenatus (Embiotocidae), a wide-ranging Eastern Pacific reef fish with no pelagic larval stage.*

Toy, J. A., Kroeker, K. J., Logan, C. A., Takeshita, Y., Longo, G. C., & Bernardi, G. (2022). Upwelling-level acidification and pH/pCO2 variability moderate effects of ocean acidification on brain gene expression in the temperate surfperch, Embiotoca jacksoni. *Molecular Ecology*, *31*(18), 4707–4725. https://doi.org/https://doi.org/10.1111/mec.16611

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L., & Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, *7*(3), 562–578.

Travis, J. M. J., Delgado, M., Bocedi, G., Baguette, M., Bartoń, K., Bonte, D., Boulangeat, I., Hodgson, J. A., Kubisch, A., Penteriani, V., Saastamoinen, M., Stevens, V. M., & Bullock, J. M. (2013). Dispersal and species' responses to climate change. *Oikos*, *122*(11), 1532–1540. https://doi.org/10.1111/j.1600-0706.2013.00399.x

Tresguerres, M., & Hamilton, T. J. (2017). Acid–base physiology, neurobiology and behaviour in relation to CO2-induced ocean acidification. *Journal of Experimental Biology*, *220*(12), 2136–2148. https://doi.org/10.1242/jeb.144113

Ulrey, A. B., & Greeley, P. O. (1928). A list of the marine fishes (Teleostei) of southern California with their distribution. *Bulletin of the Southern California Academy of Sciences*, *27*(1), 1–53.

Vaser, R., Sović, I., Nagarajan, N., & Šikić, M. (2017). Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Research*, *27*(5), 737–746. https://doi.org/10.1101/gr.214270.116

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., Young, S. K., & Earl, A. M. (2014). Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE*, *9*(11). https://doi.org/10.1371/journal.pone.0112963

Wallman, H. L., & Bennett, W. A. (2006). Effects of Parturition and Feeding on Thermal Preference of Atlantic Stingray, Dasyatis sabina (Lesueur). *Environmental Biology of Fishes*, *75*(3), 259–267. https://doi.org/10.1007/s10641-006-0025-1

Warner, R. R. (1997). Evolutionary ecology: How to reconcile pelagic dispersal with local adaptation. *Coral Reefs*, *16*(Suppl. 1), S115–S120. https://doi.org/10.1007/s003380050247

West, J. E., O'Neill, S. M., & Ylitalo, G. M. (2017). Time Trends of Persistent Organic Pollutants in Benthic and Pelagic Indicator Fishes from Puget Sound, Washington, USA. *Archives of Environmental Contamination and Toxicology*, *73*(2), 207–229. https://doi.org/10.1007/s00244-017-0383-z

Wethey, D. S., Woodin, S. A., Hilbish, T. J., Jones, S. J., Lima, F. P., & Brannock, P. M. (2011). Response of intertidal populations to climate: Effects of extreme events versus long term change. *Journal of Experimental Marine Biology and Ecology*, *400*(1), 132–144. https://doi.org/https://doi.org/10.1016/j.jembe.2011.02.008

Wiebe, J. P. (1968). The effects of temperature and daylength on the reproductive physiology of the viviparous seaperch, Cymatogaster aggregata Gibbons. *Canadian Journal of Zoology*, *46*(6), 1207–1219. https://doi.org/10.1139/z68-171

Williams, C. R., Dittman, A. H., McElhany, P., Busch, D. S., Maher, M. T., Bammler, T. K., MacDonald, J. W., & Gallagher, E. P. (2019). Elevated CO 2

impairs olfactory-mediated neural and behavioral responses and gene expression in ocean-phase coho salmon (Oncorhynchus kisutch). *Global Change Biology*, *25*(3), 963–977. https://doi.org/10.1111/gcb.14532

Wittmann, A. C., & Pörtner, H.-O. (2013). Sensitivities of extant animal taxa to ocean acidification. *Nature Climate Change*, *3*(11), 995–1001. https://doi.org/10.1038/nclimate1982