THE UNIVERSITY OF CHICAGO

POPULATION STRUCTURE AND LOCAL ADAPTATION IN THE OLYMPIA OYSTER (*OSTREA LURIDA*)

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

COMMITTEE ON EVOLUTIONARY BIOLOGY

BY
KATHERINE ELIZABETH SILLIMAN

CHICAGO, ILLINOIS
JUNE 2019

DEDICATION

For my parents, who gave me Silent World by Jacques Cousteau, then provided unwavering support of my dreams all the years after.

I read a lot of literature

On oysters in my quest

Determined I would find a type

That outclassed all the rest.

. . .

At last my search was ended

Food of the Gods I'd found

The nearest place to Paradise—

"Olympia," Puget Sound.

Ode to the Olympia Oyster

Jay Bolster

TABLE OF CONTENTS

LIST OF FIGURES vii

LIST OF TABLES x

ACKNOWLEDGMENTS xii

ABSTRACT xiv

INTRODUCTION 1
Study System 2

1 POPULATION STRUCTURE, GENETIC CONNECTIVITY, AND ADAPTATION IN THE OLYMPIA OYSTER 4
1.1 Abstract 4
1.2 Introduction 5
1.3 Materials and Methods 7
1.3.1 Sample collection 7
1.3.2 Genotype-by-sequencing analysis 9
1.3.3 Detection of loci under putative selection 10
1.3.4 Summary statistics, population differentiation, and spatial structure . 12
1.3.5 Estimating connectivity and historical relationships 13
1.4 Results 14
1.4.1 GBS and outlier detection 14
1.4.2 Summary statistics, population differentiation, and spatial structure 15
1.4.3 Connectivity and historical relationships 19
1.4.4 Functional annotation of outliers 22
1.5 Discussion 24
1.5.1 Regional population structure and gene flow 24
1.5.2 Anthropogenic influences on population structure 28
1.5.3 Local adaptation 29
1.5.4 Potential Limitations 31

2 EVIDENCE FOR ADAPTIVE DIVERGENCE WITH GENE FLOW 33
2.1 Abstract 33
2.2 Introduction 33
2.3 Materials and Methods 36
2.3.1 Broodstock 36
2.3.2 Larval Rearing and Quantifying Reproductive Activity 37
2.3.3 Experimental Set-up: Larvae Growth 37
2.3.4 Experimental Set-up: Juvenile Growth 38
2.3.5 2b-RAD Genotyping 38
2.3.6 Analysis 39
2.4 Results 40

2.4.1 Reproductive Activity . . . 40
2.4.2 Larval Growth Experiment . . . 42
2.4.3 Juvenile Growth Experiment . . . 42
2.4.4 Genetic Diversity and Relatedness . . . 44
2.5 Discussion . . . 45

3 GENE EXPRESSION RESPONSES TO ACIDIFICATION IN TWO WIDESPREAD MARINE BIVALVES . . . 49
3.1 Introduction . . . 49
3.2 Materials and Methods . . . 53
3.2.1 Sampling and common garden experiment . . . 53
3.2.2 Transcriptome assembly . . . 55
3.2.3 Differential gene expression and functional enrichment . . . 57
3.3 Results . . . 60
3.3.1 Acidification experiment . . . 60
3.3.2 Differential gene expression in Olympia oysters . . . 61
3.3.3 Differential gene expression in rock scallops . . . 73
3.3.4 Comparing oysters and scallops . . . 78
3.3.5 Mapping candidate adaptive loci from Ch. 1 . . . 78
3.4 Discussion . . . 81
3.4.1 Conserved responses to acidification . . . 81
3.4.2 Constitutive differences in gene expression among populations . . . 84
3.4.3 Population-specific responses to low pH . . . 85
3.4.4 Candidate adaptive markers . . . 86

DATA ACCESSIBILITY . . . 88

APPENDIX: SUPPLEMENTARY MATERIAL FOR CH. 1 . . . 89
Sampling locations and population-specific summary statistics . . . 89
Summary statistics for phylogeographic regions . . . 90
Pairwise FST values for outlier and neutral datasets . . . 91
Additional results of outlier analyses . . . 93

REFERENCES . . . 95

LIST OF FIGURES

1.1 Map of 20 Olympia oyster (*Ostrea lurida*) collection sites from the west coast of North America. 8
1.2 Population structure results for 19 *O. lurida* populations using (A-B) neutral loci and (C-D) outlier loci. (A, C) Plots of individual admixture determined using the program STRUCTURE at the K recommended by the ΔK method ($K = 5$ neutral, $K = 2$ outlier), as well as at the value of K inferred from PCA ($K = 6$ neutral, $K = 5$ outlier). (B, D) Principal components analysis plots for PCs 1-5. PC1 is plotted against latitude of sampling site, then PC2 vs. PC3 and PC4 vs. PC5. Large transparent circles indicate the centroid of populations. Colors refer to the phylogeographic regions of each population. 17
1.3 Illustration depicting Olympia oyster sampling protocol at low tide. (ConcernedApe LLC, 2019) 18
1.4 A) Heatmap of pairwise F_{ST} values for 19 populations of *O. lurida* using 13,073 neutral SNPs. Populations are ordered from north to south, starting with Klaskino, BC. B) Isolation by distance plot of $F_{ST}/1-F_{ST}$ versus population pairwise coastal water distance. Neutral loci in red circles ($p < 0.001$) and outlier loci in blue diamonds ($p < 0.001$). 19
1.5 TreeMix results for 19 *O. lurida* populations using 1 SNP per neutral locus. Seven migration events are modeled, as this was the best value inferred by evaluating model fit. The tree is rooted by the southernmost populations, San Diego Bay, CA and Mugu Lagoon, CA, and ordered by latitude where possible. Populations are colored by their inferred phylogeographic region. 20
1.6 Model of effective migration rates as inferred by EEMS for neutral loci in *O. lurida*. Orange represents areas of low migration relative to the average and blue are areas of higher migration. Grey arrows indicate regions of significantly reduced or increased migration (posterior probability > 95%) 21
1.7 Diversity increases from north to south in *O. lurida*. (A) Effective diversity rates as inferred by EEMS, with orange representing areas of lower diversity and blue representing high diversity. (B) Expected heterozygosity (H_e) within each population versus population latitude. 22

2.1 Reproductive activity in F1 oysters from three Puget Sound populations. The left axis measures the number of larvae released on each sampling day, summed up across replicates and normalized by the number of adult oysters in each replicate. The right axis measures the cumulative number of larvae produced through time, normalized by the number of adult oysters in each population. 41

2.2 **a**: Cumulative number of larvae released within each replicate bucket over 7 weeks, normalized by the number of oysters in each bucket. Oysters from Fidalgo produced significantly fewer than those from Oyster Bay (Tukey post hoc test, $p = 0.0499$) and fewer, although not significantly, than those from Dabob (Tukey post hoc test, $p = 0.0954$). **b**: Calendar day of first observed larvae release. Fidalgo oysters released larvae 10 days later than Oyster Bay oysters (Tukey post hoc test, $p = 0.0434$) and 7 days later than Dabob oysters on average, although this was not statistically significant (Tukey post hoc test, $p = 0.156$). 42
2.3 Larval shell length of F2 oysters from three populations over 14 days. Data are means across replicates += s.e.m. Size varied significantly among populations at Day 7 (LMM,$p = 7.939e^{-5}$) and Day 14 (LMM, $p = 0.03573$). 43
2.4 Juvenile shell area growth rate of F2 oysters from three populations over 9 weeks. Juvenile shell growth of F2 oysters (growth rate = Δarea / # days). Growth rate between Day 0 - Day 48 differed significantly among populations (LMM, $p = 0.02236$) as well as between Day 48 - Day 68 (LMM, $p = 0.0012$). 44

3.1 Water chemistry observations for each tank replicate after the pH was lowered in the low pH treatment. Header 1 is the mixing header tank for the ambient treatment, Header 2 is the header tank for the low pH treatment. 61
3.2 PCA of gene expression in the 500 genes with the highest variance across all *O. lurida* samples. Color represents tissue and shape is pH treatment. 63
3.3 GO_MWU results for *O. lurida* genes with a conserved response to low pH across populations. a) ctenidia, b) mantle, 7 weeks. Blue are downregulated in response to treatment, red are upregulated. The font refers to significance of the FDR test for enrichment; the dendrograms depict the sharing of genes between categories; the fractions correspond to genes with an unadjusted p < 0.05, relative to the total number of genes within the category. 65
3.4 GO_MWU results for DEGs ctenidia genes between OR and CA *O. lurida* populations at 36 hours in the ambient treatment. Blue are downregulated in OR compared to CA, red are upregulated in OR. The font refers to significance of the FDR test for enrichment; the dendrograms depict the sharing of genes between categories; the fractions correspond to genes with an unadjusted p < 0.05, relative to the total number of genes within the category. 67
3.5 GO_MWU results for *O. lurida* genes with a significant change based on population (when controlling for treatment) in a) ctenidia samples and b) mantle samples. Font refers to significance of the FDR test for enrichment; fractions correspond to genes with an adjusted p< 0.05, relative to the total number of genes within the category. 68
3.6 GO_MWU results for ctenidia DEGs with a population-specific response to low pH at 36 hours. 70
3.7 GO_MWU results for ctenidia DEGs in OR at 36 hours. Blue are downregulated in response to treatment, red are upregulated. 71

3.8 KOG_MWU comparison for *O. lurida* genes that are differentially expressed in response to low pH, for all combinations of tissues, time points, and populations. Red indicates upregulated, blue is downregulated; * indicate significant enrichment after FDR correction, ***** represents the top upregulated and top downregulated classes for each dataset. Datasets are clustered by their correlation. . . . 72
3.9 PCA of gene expression for all *C. gigantea* samples, with ctenidia samples clustering on the left and mantle samples on the right. 74
3.10 Tissue-specific PCA of *C. gigantea* samples. A: PCA of ctenidia samples, with colors for population and shape for treatment. B: PCA of mantle samples, with colors for population and shape for treatment. C: PCA of mantle samples, with color for sampling time point and shape for sequencing batch. 74
3.11 GO_MWU results for *C. gigantea* genes with a conserved response to low pH across populations: a) ctenidia, b) mantle. Blue are downregulated in response to treatment, red are upregulated. 76
3.12 KOG_MWU comparison for *C. gigantea* genes that are differentially expressed in response to low pH, for all combinations of tissues, time points, and populations. Red indicates upregulated, blue is downregulated; * indicate significant enrichment after FDR correction, ***** represents the top upregulated and top downregulated classes for each dataset. Datasets are clustered by their correlation. . 77
3.13 Comparison of KOG_MWU enrichment among ctenidia and mantle datasets for both species at both time points, with the exception of *O. lurida* mantle genes at 36 hours, as no KOG classes were enriched. Red indicates upregulated, blue is downregulated; * indicate significant enrichment after FDR correction, ***** represents the top upregulated and top downregulated classes for each dataset. Datasets are clustered by their correlation. 79
S1 Venn diagram with number of SNPs and Genotype-by-Sequencing loci identified as outliers by three methods: *pcadapt*, *OutFLANK*, and BayeScan 93
S2 Outlier loci predominantly show clinal patterns in allele frequency. Allele frequency in 129 individual outlier loci plotted against latitude for 19 populations of *O. lurida*. One SNP is represented for each locus, except in the case where two outlier SNPs from the same locus showed different spatial patterns (e.g., locus_277490). Populations are colored by inferred phylogeographic regions. . . . 94

LIST OF TABLES

1.1 Overall summary statistics for the combined (13,424 SNPs), neutral (13,073 SNPs), and outlier (235 SNPs) datasets. H_o, observed heterozygosity averaged across loci; H_e, expected heterozygosity averaged across loci; F_{IS} & F_{ST}, Wright's F-statistics averaged across loci (Nei and Chesser, 1983); pairwise F_{ST}, average of all pairwise F_{ST} values (Weir and Cockerham, 1984) 15

1.2 BLASTx results for outlier loci. Only the 18 loci with positive BLAST hits are shown. 23

2.1 Genetic diversity and relatedness estimates for Puget Sound *O. lurida* populations. H_o, observed heterozygosity; H_e, expected heterozygosity; F_{is}, inbreeding coefficient; ϕ, kinship coefficient 45

3.1 Mean and SE of discrete water sampling during the experiment for the two treatments. * indicates parameters that were measured directly, others were estimated using the R package seacarb. 60

3.2 Transcriptome assemblies of *O. lurida* populations separately, *O. lurida* populations combined using OrthoFinder, and an assembly of *C. gigantea* using MetaDRAP. 62

3.3 Annotation results for *O. lurida* and *C. gigantea* transcriptomes. 62

3.4 Number of significantly differentially expressed genes for each tissue and time point, as determined by DeSeq2 (FDR < 0.1). Combined: samples are pooled across both time points, with "time" included as a factor in GLMs. 64

3.5 Number of differentially expressed genes for each tissue and time point in *C. gigantea*, as determined by DeSeq2. 75

3.6 Candidate adaptive markers in *O. lurida*, identified using both GBS and transcriptomics. DEG set refers to the gene set where the gene is significantly expressed (e.g., OR (CT1) is differentially expressed in OR ctenidia samples at T1-36 hours in response to treatment)). Abbreviations for KOG classes: C: Energy production and conversion, J: Translation, ribosomal structure and biogenesis, K: Transcription, L: Replication, recombination and repair, P: Inorganic ion transport and metabolism, T: Signal transduction mechanisms, U: Intracellular trafficking, secretion, and vesicular transport, Z: Cytoskeleton 80

S1 GPS coordinates of sampling sites and population-specific summary statistics averaged across markers using the combined dataset of 13,424 SNPs. H_e, expected heterozygosity; F_{IS}, inbreeding coefficient within the population, mean and 25%-75% confidence intervals (Nei and Chesser, 1983) 89

S2 Overall summary statistics for each phylogeographic region using the neutral dataset of 13,073 SNPs. H_o, observed heterozygosity averaged across loci; H_e, expected heterozygosity averaged across loci; F_{IS} & F_{ST}, Wright's F-statistics averaged across loci (Nei and Chesser, 1983). Note that F_{ST} may be skewed by variation in sampling strategy across regions. 90

S3 Pairwise F_{ST} values for all pairs of populations, using the neutral dataset of 13,073 SNPs. 91

S4 Pairwise F_{ST} values for all pairs of populations, using the outlier dataset of 235 SNPs. 92

ACKNOWLEDGMENTS

These projects would not have been possible without technical, financial, and emotional support from numerous individuals and organizations. Research funding for this work came from the National Science Foundation Doctoral Dissertation Improvement Grant (DEB1601281), National Geographic Young Explorers Grant, NOAA Washington Sea Grant, student awards through the American Fisheries Society and National Shellfisheries Association, Puget Sound Restoration Fund (PSRF), the Hinds Fund, and the Committee on Evolutionary Biology at the University of Chicago. Fellowship funding came from the NSF Graduate Research Fellowship Program and the U.S. Dept. of Education GAANN program.

For logistical assistance with oyster collections, I thank Brady Blake, Washington Dept. of Fish and Wildlife, Brian Kingzett, Jennifer Ruesink, Alan Trimble, Steven Rumrill, Danielle Zacherl, Dave Couch, Thomas McCormick, JoAnne Linnenbrink, Martin Ruane, Chela Zabin, Betsy Peabody, Daniel Silliman, and Stephen Barlow. Anna Deck, Geana Ayala, Jane Watson, Scott Russett, and Scott Groth collected and shipped oysters in December weather for Chapter 3, of which I am grateful. Data collection was facilitated by excellent research assistants—Sean Bennett and Tynan Bowyer.

All of the experimental manipulations were conducted at the K. K. Chew Center for Shellfish Research and Restoration, which is part of the NOAA Manchester Research Station and managed by Puget Sound Restoration Fund (PSRF). The staff at PSRF, notably Ryan Crim, Alice Helker, and Stuart Ryan, were instrumental to the success of these projects and I thank them for generously sharing their expertise, facilities, and time. Molecular work was primarily carried out at the Pritzker Laboratory for Molecular Systematics and Evolution at the Field Museum of Natural History. For assistance in the Pritzker Lab, I thank Brian Wray and Kevin Feldheim. I would also like to thank Trevor Price, Bettina Harr, Alex Gileta, and Marcus Kronforst for providing additional lab facilities at the University of Chicago.

Finally, I would like to thank my committee for all of their help—Cathy Pfister, John Novembre, Marcus Kronforst, and Rick Ree. I especially need to acknowledge my advisor

Cathy Pfister. Throughout my Ph.D., she enthusiastically encouraged my independence, pushed me to think about my results in the broader context, and was always available if I needed support (e.g., talking me off the proverbial ledge when all of my controls died one summer). Many other people in the Darwinian Sciences and Division of Biological Sciences provided support (educational, administrative, logistical, emotional) during my Ph.D. This includes but is not limited to Max Winston, Ben Winger, Stefano Allesina, Michael Coates, Carolyn Johnson, Libby Eakin, Dave Steele, Caitlin Meadows, Yoseop Yoon, Tim Wootton, Courtney Stepien, Amy Henry, Orissa Moulton, Sara Jackrel, Nicole Bitler, Simon Lax, Mark Bitter, John Park, Natasha Ershova, and Daniel Smith.

My collaborators at University of Washington provided crucial financial and technical support, without which this dissertation actually would not have been possible. In the Roberts Lab: Steven Roberts, Brent Vadopalas, Sam White, and Laura Spencer, and in the Hauser Lab: Lorenz Hauser, Natalie Lowell, and Isadora Jimenez. I'd especially like to acknowledge Steven Roberts, who welcomed me into his lab group as if I were one of his own students, and has been an invaluable mentor. Natalie Lowell planned and oversaw the execution of the ocean acidification experiment, and constantly inspires me to be more meticulous in both my science and my simulated farming.

I would like to thank my husband, Michael A. Alcorn, for editing more grants and papers than he would've liked, statistical discussions, code debugging, and teaching me the proper grammar of em-dashes.

ABSTRACT

Effective management of threatened species requires an understanding of both the genetic connectivity among populations and adaptive population divergence. For the numerous coastal marine species with planktonic dispersal, high connectivity can obscure population boundaries and oppose the diversifying effects of natural selection through homogenizing gene flow. Using an ecologically and commercially important marine bivalve as a model system, my dissertation aimed to characterize the spatial scales of neutral and adaptive differentiation in the face of gene flow and identify candidate loci under selection. The Olympia oyster (*Ostrea lurida*) is native from Baja California to the central coast of Canada and distributed over strong environmental gradients. Following devastating commercial exploitation by the early 20th century, recovery of *O. lurida* populations has faced other anthropogenic challenges, including ocean acidification. For my dissertation, I used high-throughput sequencing, bioinformatics, and mesocosm experiments to 1) describe the neutral and adaptive population genetic structure in *O. lurida*, 2) characterize adaptive phenotypic variation at a local scale, and 3) evaluate molecular responses to acidification stress across genetically diverged populations in two bivalve species.

Significant population structure in the Olympia oyster was observed using both neutral and putative adaptive genetic markers derived from genotype-by-sequencing of oysters across 20 sites. To determine if local adaptation can occur among populations with high inferred gene flow, I investigated genetic and phenotypic variation among three populations of oysters in Puget Sound, WA. Through a common garden experiment on oysters that had been reared for up to two generations in common conditions, I demonstrated that these three populations exhibit heritable differences in reproductive timing, larval growth rate, and juvenile growth rate. Adaptations to natural long-standing variation of ocean pH in widespread species along western North America may be informative for predicting resilience to projected conditions. Overlapping the Olympia oysters range, the purple-hinged rock scallop (*Crassadoma gigantea*) is found from southern California to the Aleutian Islands. To understand inter- and

intraspecific variation in response to reduced pH, I compared gene expression responses to two pH treatments (7.4 and 7.8) in adult oysters and rock scallops from multiple genetically diverged populations. Within species, genes were identified that exhibited a conserved response to pH across populations or a significant population-specific response—the latter are considered candidate genes involved in local adaptation.

INTRODUCTION

Natural environments exhibit spatial heterogeneity in both abiotic and biotic factors, oftentimes driving populations to evolve traits that confer a fitness advantage in their native habitat over foreign genotypes Kawecki and Ebert (2004). This process of local adaptation can be opposed by neutral processes, such as genetic drift within effectively small populations, or homogenizing gene flow from dispersal (Wakeley, 2005). However, if the scale of dispersal extends across strong selective gradients, adaptive population divergence can still occur through phenotype-environment mismatch–where strong purifying selection occurs each generation following dispersal (Marshall, 2008; Schmidt and Rand, 2001). Characterizing the spatial scales of both neutral and adaptive divergence is crucial for determining effective management strategies for taxa at risk from anthropogenic stressors (Baums, 2008; Carroll et al., 2014).

For the numerous coastal marine species with planktonic dispersal, high connectivity among populations can obscure population boundaries and oppose the diversifying effects of natural selection through gene flow (Lenormand, 2002). However, a growing body of evidence indicates that both limited effective dispersal and local adaptation may be more common in marine species than previously hypothesized (Cowen et al., 2000; Hauser and Carvalho, 2008; Sanford and Kelly, 2011; Weersing and Toonen, 2009; Levin, 2006). As direct measurement of dispersal is challenging or impossible for many such species, estimates of demographic population structure have increasingly come from genetic markers (Ashley, 2010). Adaptive divergence can be assessed through either experimental manipulations or by comparing allelic differences between populations at loci of adaptive importance (Sanford and Kelly, 2011). The rapid development of high-throughput sequencing has advanced studies of demography and adaptive variation in nonmodel organisms by facilitating genotyping at thousands of genomic markers, as well as cost effective assays of molecular responses to environmental variation (Matz, 2018).

For my dissertation, I have leveraged reduced-representation genomic and transcriptomic

sequencing, bioinformatics, and experimental manipulations in a marine invertebrate to determine the spatial scale of adaptive differentiation with gene flow and identify candidate loci under selection. In Chapter 1, I characterized rangewide neutral and adaptive population genetic structure using genotype-by-sequencing. For Chapters 2 and 3, I experimentally tested for local adaptation between populations at varying spatial scales and levels of inferred gene flow. In addition to addressing key evolutionary questions, my work has direct implications for ongoing restoration efforts of an ecologically and commercially important marine bivalve and understanding how this species will respond to rapid environmental change.

Study System

The Olympia oyster (*Ostrea lurida*, Carpenter 1864) is an estuarine bivalve native from Baja California to the central coast of Canada, patchily distributed over strong environmental gradients (Chan et al., 2017; Schoch et al., 2006). Whether this tolerance is primarily due to plasticity, local adaptation, or balanced polymorphism is largely unknown. A protandrous hermaphrodite, spawning events are synced between adult males and females based on environmental cues of temperature and seasonality (Coe, 1932). Unlike other oysters where both males and females spawn gametes (e.g., *Crassostrea*), the females fertilize eggs with sperm from the water column and initially brood larvae in the mantle cavity. After release, the larvae have been reported to be planktonic from seven days to eight weeks before settling on a hard substrate (Baker, 1995). This newly settled spat is around 300-340 um and will reach sexual maturity under good growing conditions in about a year. The impact of maternal brooding on population structure in Osterideae has not been examined.

As ecosystem engineers in estuaries, they provide structured habitat, remove suspended sediments, and limit eutrophication (zu Ermgassen et al., 2013; Coen et al., 2011). Following devastating commercial exploitation in the 19$^{\text{th}}$ and early 20$^{\text{th}}$ centuries, recovery of Olympia oyster populations has been stifled by other anthropogenic threats (e.g., water quality issues, habitat loss, and potentially ocean acidification (Blake and Bradbury, 2012; Hettinger et al.,

2013; Sanford et al., 2014)). The last 15 years has seen increased interest in the Olympia oyster, with restoration projects underway by both government and nongovernment agencies across its range (Pritchard et al., 2015). Current knowledge about the population genetic structure of *O. lurida* comes primarily from an unpublished 2011 dissertation (Stick, 2011) and two phylogeographic studies using two mitochondrial loci (Polson et al., 2009; Raith et al., 2016). It is critical that management plans utilize modern techniques to understand the genetic structure of this species and its adaptive potential in the face of climate change.

CHAPTER 1

POPULATION STRUCTURE, GENETIC CONNECTIVITY, AND ADAPTATION IN THE OLYMPIA OYSTER[1]

1.1 Abstract

Effective management of threatened and exploited species requires an understanding of both the genetic connectivity among populations and local adaptation. For such coastal marine species, population structure could follow a continuous isolation-by-distance model, contain regional blocks of genetic similarity separated by barriers to gene flow, or be consistent with a null model of no population structure. To distinguish between these hypotheses in *O. lurida*, 13,424 single-nucleotide polymorphisms (SNPs) were used to characterize rangewide population structure, genetic connectivity, and adaptive divergence. Samples were collected across the species range on the west coast of North America, from southern California to Vancouver Island. A conservative approach for detecting putative loci under selection identified 235 SNPs across 129 GBS loci, which were functionally annotated and analyzed separately from the remaining neutral loci. While strong population structure was observed on a regional scale in both neutral and outlier markers, neutral markers had greater power to detect fine-scale structure. Geographic regions of reduced gene flow aligned with known marine biogeographic barriers, such as Cape Mendocino, Monterey Bay, and the currents around Cape Flattery. The outlier loci identified as under putative selection included genes involved in developmental regulation, sensory information processing, energy metabolism, immune response, and muscle contraction. These loci are excellent candidates for future research and may provide targets for genetic monitoring programs. Beyond specific applications for restoration and management of the Olympia oyster, this study lends to the growing body of evidence for both population structure and adaptive differentiation across a range of marine

1. A version of this chapter has been published as: Silliman K. 2019. Population structure, genetic connectivity, and adaptation in the Olympia oyster. Evolutionary Applications. *In press.*

species exhibiting the potential for panmixia. Computational notebooks are available to facilitate reproducibility and future open-sourced research on the population structure of *O. lurida*.

1.2 Introduction

Coastal marine ecosystems provide important services such as carbon sequestration, food production, and recreation (Luisetti et al., 2014), yet contain some of the most exploited and threatened species on earth. As evidence for the direct impacts of human activities (e.g., over-harvesting, increasing atmospheric CO_2, and nutrient run-off) on these species grows, there has been increased focus on restoring depleted abundances, recovering ecosystem services, and determining which species are capable of adapting to environmental change (Granek et al., 2010). Effective management of threatened and exploited species requires an understanding of both the genetic connectivity among populations and adaptation across environmental gradients (Baums, 2008; Miller and Ayre, 2008; Palumbi, 2003). Anthropogenic movement of individuals between populations, either for aquaculture or restoration purposes, can confound signals of population structure and should be evaluated when drawing conclusions about genetic connectivity (David, 2018).

Neutral molecular markers (e.g., microsatellites) have traditionally been used exclusively to identify the geographic structure of subpopulations and estimate the genetic connectivity between them (Funk et al., 2012); however, these do not give insight into the scale or magnitude of adaptive divergence. For populations connected by dispersal, adaptation to local conditions can still occur if the strength of selection overcomes the homogenizing effect of gene flow (Hellberg, 2009). Advances in genomic and computational techniques, such as genotype-by-sequencing (GBS), have facilitated the detection of genomic regions that may be influenced by natural selection (outlier loci) (Stapley et al., 2010). Although often referred to as adaptive markers, these outliers may only be linked to loci that are under natural selection rather than confer an adaptive advantage themselves. Outlier loci have

provided increased spatial resolution of population structure compared to neutral loci alone for some marine species (Drinan et al., 2018; Van Wyngaarden et al., 2017; Milano et al., 2014), but not all (Moore et al., 2014). In addition to potentially resolving fine-scale genetic differentiation, outlier loci may be useful for characterizing the adaptive potential of a species or population to future environmental conditions (Eizaguirre and Baltazar-Soares, 2014). Inferences from both neutral and adaptive markers should be combined when making management recommendations (Funk et al., 2012).

Current knowledge about the population genetic structure of *O. lurida* comes primarily from an unpublished 2011 dissertation, which sampled from San Francisco, CA to Vancouver Island, BC and found regional population structure using microsatellites (Stick, 2011). Two phylogeographic studies using two mitochondrial loci identified a phylogeographic break north of Willapa Bay, WA and established the southern boundary divide between *O. lurida* and its sister species *Ostrea conchaphila* (Polson et al., 2009; Raith et al., 2016). Future and ongoing management plans would benefit greatly from thorough analysis of the fine-scale genetic structure using modern genomic techniques and rangewide sampling (Camara and Vadopalas, 2009).

The objective of this study was to characterize the spatial population structure of the Olympia oyster across the majority of its range using both neutral and adaptive markers derived from genome-wide single nucleotide polymorphisms (SNPs). I specifically tested whether patterns of genetic variation suggest a smooth continuum of allele frequency shifts consistent with isolation-by-distance (IBD) (Malécot, 1968), regional blocks of genetic similarity that correspond to physical barriers (Hare and Avise, 1996), or the null model of no significant genetic differentiation (Grosberg and Cunningham, 2001). SNPs produced from high-throughput sequencing have led to the identification of previously undetected population structure in a number of marine and terrestrial species (e.g., (Everett et al., 2016; Reitzel et al., 2013; Van Wyngaarden et al., 2017)). Compared to the Atlantic coast of North America (Hoey and Pinsky, 2018), studies utilizing genome-wide SNPs for marine

taxa from the Pacific coast are far fewer in number and have been limited to regional spatial scales (De Wit and Palumbi, 2013; Drinan et al., 2018; Gleason and Burton, 2016; Larson et al., 2014; Martinez et al., 2017) or in the number of sampling sites (Pespeni et al., 2010; Tepolt and Palumbi, 2015). This study is the first of my knowledge to utilize thousands of SNPs to extensively survey the rangewide population structure of a marine species along this coast.

A secondary aim of this study was to produce a reproducible computational pipeline to go from raw data to results and figures, using Jupyter Notebooks. Jupyter Notebooks are interactive documents that integrate text, code, and analysis results (Kluyver et al., 2016). A major issue for genomic analyses today is how to clearly explain the computational methods used in order to allow for reproducibility (Kanwal et al., 2017). This open access pipeline is intended to provide an example template to improve reproducibility in future studies and function as an instructional tool for biologists and early career scientists who wish to apply these methods to their own study organisms.

1.3 Materials and Methods

1.3.1 Sample collection

20-25 adult *Ostrea lurida* over 2cm in length were collected primarily by hand from the intertidal (approx. 0m to -1m tidal height) at 20 sites ranging from Klaskino Inlet, Vancouver Island (50°17 55) to San Diego Bay, CA (32°361 9) in 2014 (Fig. 1.1, Table S1). When possible, oysters were sampled randomly along 10m transects (Fig. 1.3). Each site represents a separate bay, except for Willapa Bay, WA and San Francisco Bay, CA which had two sampling sites each. In the case of one site from Willapa Bay, WA, oysters were collected subtidally (approx. 8m depth) through dredge harvesting of the Pacific oyster, *Crassostrea gigas*. Adductor tissue samples were preserved in RNALater, followed by storage in -80°C. DNA was isolated using DNeasy Blood & Tissue kits (Qiagen) and E.Z.N.A. Mollusc DNA

kits (Omega Bio-Tek) with RNAse-A treatment following manufacturer instructions and limiting tissue digestion time to no more than 90 minutes. All DNA samples were quantified using the Qubit dsDNA BR Assay kit (Life Technologies) on a Qubit v2.0 (Life Technologies). DNA quality was verified by agarose gel electrophoresis of 1-2 uL extracted DNA on an 0.8% TAE gel.

Figure 1.1: Map of 20 Olympia oyster (*Ostrea lurida*) collection sites from the west coast of North America.

1.3.2 Genotype-by-sequencing analysis

Library preparation for genotype-by-sequencing (GBS) followed the protocol by (Elshire et al., 2011) using the *ApeKI* restriction enzyme, with an additional size selection step and slight modifications to PCR amplification. A detailed protocol can be found at `https://github.com/ksil91/Ostrea_PopStructure`. All libraries were size-selected for fragment sizes between 200bp-450bp on a Blue PippinPrep (Sage Science) to reduce the number of loci sequenced and ensure adequate sequencing coverage. One pool of 90 samples was sequenced across two 100bp paired-end Illumina HiSeq 2500 lanes, with only the forward sequencing read used for analysis. Seven other pools with a maximum of 48 libraries each were sequenced on eight 100bp single-end lanes (246 different samples, 86 technical replicates, 332 libraries total). Sampling sites were spread out across all libraries in order to minimize batch effects from library preparation and sequencing. Raw sequencing reads were de-multiplexed, quality filtered, and de novo clustered using the API version of the seven step computational pipeline *ipyrad* v.0.7.24 (Eaton, 2014) and implemented in Python via a Jupyter Notebook running on a large computational cluster. Demultiplexing (Step 1) used sample-specific barcode sequences, allowing one mismatch in the barcode sequence. Base calls with a Phred quality score under 20 were converted to Ns, and reads containing more than 5 Ns were discarded. Adapter sequences, barcodes, and the cutsite sequences were trimmed from filtered reads, with only reads greater than 35bp retained (Step 2). Reads were then clustered using a sequence similarity threshold of 85% both within (Step 3) and between samples to genotype polymorphisms (Steps 4, 5) and identify orthologous loci (Step 6) with a minimum of 10x read coverage. Replicate samples were assembled separately and then compared using custom Perl scripts by Mikhail Matz (Wang et al., 2012). The replicate with the largest number of GBS loci after final filtering (Step 7) was retained. Samples were removed if they had fewer than 200,000 raw sequencing reads, fewer than 15,000 assembled clusters of at least 10x read depth, and were missing data for over 55% of loci assembled across at least 75% of samples, with Steps 4-7 rerun using the remaining individuals.

The final assembly was then filtered for excess heterozygosity based on deviations from Hardy-Weinberg equilibrium (HWE) in at least two populations, sample coverage of 75%, and an overall minor allele frequency (MAF) of 2.5%, retaining only GBS loci found in at least one individual from all populations. Preliminary analyses conducted on datasets allowing more or less missing data showed that the inferred population structure was robust to missing data up to 40%. Population genetic summary statistics, with the exception of F_{ST}, did change quantitatively due to missing data but not qualitatively (not shown) (Cariou et al., 2016). Filtering steps were conducted using VCFtools (Danecek et al., 2011), custom Python code, and code adapted from Jon Puritzs lab (Puritz et al., 2014). Input files and formats for subsequent analysis of population structure were created using a combination of custom Python code, custom R code, and the *radiator* R package (Gosselin, 2017). Every step of the assembly, filtering process, and creation of input files can be reproduced through Jupyter Notebooks.

1.3.3 Detection of loci under putative selection

Following recommendations to utilize multiple methods to detect loci under putative directional selection (Benestan et al., 2016; Rellstab et al., 2015), three approaches were used on the filtered SNP dataset: BayeScan v.2.1, *OutFLANK* v.0.2, and *pcadapt* v.4.0.2. For BayeScan and *OutFLANK*, individuals were grouped into populations by sampling site. GBS loci which had SNPs identified as outliers in at least two of the approaches were classified as putative adaptive GBS loci. From these GBS loci, any SNP that had been identified as an outlier by at least one approach was separated from the full SNP dataset to create an "outlier" SNP dataset. Subsequent analyses of population structure were conducted on three SNP datasets: all SNPs (combined), outlier SNPs, and neutral SNPs—which excluded any SNP found on a putative adaptive GBS locus.

BayeScan uses a Bayesian approach to apply linear regression to decompose F_{ST} coefficients into population and locusspecific components and estimates the posterior probability

of a locus showing deviation from HardyWeinberg proportions (Foll and Gaggiotti, 2008). BayeScan analysis was based on 1:100 prior odds, with 100,000 iterations, a burn-in length of 50,000, a false discovery rate (FDR) of 10%, and default parameters. Results were visualized in R. *OutFLANK* is an R package that identifies F_{ST} outliers by inferring a distribution of neutral F_{ST} using likelihood on a trimmed distribution of F_{ST} values. Because of its likelihood method, *OutFLANK* calculates F_{ST} without sample size correction when inferring the neutral distribution. Simulation studies have shown that this approach has lower false positive rates compared to other F_{ST} outlier methods (Whitlock and Lotterhos, 2015). *OutFLANK* was run using default parameters and a q-value threshold of 0.1, which can be considered a false discovery rate (FDR) of 10%. For the R package *pcadapt*, individuals are not sorted into predefined populations. Instead, *pcadapt* ascertains population structure using principal component analysis (PCA), then identifies markers under putative selection as those that are excessively correlated with population structure. When compared to BayeScan, *pcadapt* was shown to have greater power in the presence of admixed individuals and when population structure is continuous (Luu et al., 2017)—both scenarios which are likely in *O. lurida*. A scree plot representing the percentage of variance explained by each PC was used to choose the number of principal components (K) for *pcadapt*, and SNPs with a q-value threshold of 0.1 were categorized as outliers.

Putative adaptive GBS loci were functionally annotated through Blast2GO. Sequences were compared against molluscan sequences in GenBanks *nr* database using the BLASTx algorithm with default parameters and a e-value hit filter of 10^{-3}, and against EMBL-EBI InterPro signatures. Gene ontology terms were mapped to annotations with default parameters except for an e-value hit filter of 10^{-3} (Götz et al., 2008). Minor allele frequency was plotted against latitude individually for each outlier SNP in order to identify clinal patterns of allele frequency shifts.

1.3.4 Summary statistics, population differentiation, and spatial structure

Population genetic summary statistics were calculated on the combined, neutral, and outlier datasets to describe and compare overall and population-specific genetic diversity. Observed heterozygosity (H_o), expected heterozygosity (H_e), overall F_{ST}, and F_{IS} were calculated using the *basic.stats* function in the R package *hierfstat* (Goudet and Jombart, 2015). Confidence intervals for population-specific F_{IS} were determined using the *boot.ppfis* function in *hierfstat* with 1,000 bootstrap replicates. Pairwise F_{ST} following (Weir and Cockerham, 1984) was calculated using the *genet.dist* function in *hierfstat*. Heatmaps of pairwise F_{ST} values were created using *ggplot2* (Wickham, 2016). A Mantel test of coastal water distance (calculated by drawing routes between all sites on Google Earth) and $F_{ST}/1 - F_{ST}$ as implemented in *adegenet* tested for evidence of isolation-by-distance (Sokal, 1979; Jombart and Ahmed, 2011).

Rangewide population structure of *O. lurida* was characterized using a combination of Bayesian clustering and multivariate ordination approaches. These methods were applied to both the outlier and neutral datasets. The model-based Bayesian clustering method STRUCTURE v.2.2.4 (Pickrell and Pritchard, 2012) as implemented in the *ipyrad* API was used to determine the number of distinct genetic clusters (K) with a burn-in period of 50,000 repetitions followed by 200,000 repetitions. Five replicate analyses were performed for each dataset with values of K = 1-10, with each replicate independently subsampling one SNP per GBS locus and using a different random seed. Replicates were summarized and visualized using the CLUMPAK server (Kopelman et al., 2015). The ΔK method implemented in STRUCTURE HARVESTER was used to determine an optimal K (Earl and vonHoldt, 2012). PCA was implemented in the R package *adegenet* (Jombart and Ahmed, 2011) using unlinked datasets, where a single SNP with the least missing data across samples was chosen for each GBS locus (or the first SNP in the locus in the case of a tie). Missing data was filled by randomly drawing an allele based on the overall allele frequency across all individuals. The R package *PCAviz* was used to visualize PCA results and correlate PC loadings with

latitude (Novembre et al., 2018). Results from STRUCTURE, PCA, and pairwise F_{ST} were used to identify phylogeographic "regions". Summary statistics, including H_o, H_e, F_{IS}, and F_{ST} were calculated for each region using using the *basic.stats* function in the R package *hierfstat* (Goudet and Jombart, 2015).

1.3.5 Estimating connectivity and historical relationships

Spatial variation in gene flow and genetic diversity was calculated and visualized using the program EEMS (Estimated Effective Migration Surfaces) (Petkova et al., 2016). This method identifies geographic regions where genetic similarity decays more quickly than expected under isolation-by-distance based on sampling localities and a pairwise genetic dissimilarity matrix derived from SNP data. These regions may be interpreted as having reduced gene flow. A dissimilarity matrix was calculated for the neutral dataset using a variant of the *bed2diffs* R code included in the EEMS package that takes input from a *genind* R object. An outer coordinate file for defining the potential habitat of *O. lurida* was produced using the polyline method in the Google Maps API v3 tool (`http://www.birdtheme.org/useful/v3tool.html`). The habitat shape followed the shape of the coastline and excluded land regions that *O. lurida* larvae would not naturally be able to cross (e.g., the Olympic peninsula separating outer coast populations and those in Puget Sound, WA). The EEMS model is recommended to be run for various numbers of demes, which establishes the geographic grid size and resulting potential migration routes. Three independent analyses were run for each deme size (200, 250, 300, 350, 400, 500, 600, 700) for a total of 24 runs, with a burn-in of 1,000,000 and MCMC length of 5,000,000 iterations. The convergence of runs was visually assessed and results were combined across all analyses and visualized using the *Reemsplots* R package—producing maps of the effective diversity (q) and effective migration rate (m).

To infer the evolutionary relationship among sampling sites, including population splits and migration events, I reconstructed population graph trees using the software TreeMix (Pickrell and Pritchard, 2012). This method uses the covariance of allele frequencies between

populations to build a maximum likelihood graph relating populations to their common ancestor, taking admixture events ("migration") into account to improve the fit to the inferred tree. The population graph was rooted with the two southernmost *O. lurida* population (San Diego, CA and Mugu Lagoon, CA), then run allowing between 0-10 migration events. For each value of migration events, I calculated the proportion of variance in relatedness between populations that is explained by the population graph to evaluate model fit (Wang, 2017).

1.4 Results

1.4.1 GBS and outlier detection

117 samples remained after removal of 14 samples with $<$ 200,000 raw sequencing reads, 49 samples with $<$ 15,000 clusters, and 65 samples missing data for over 55% of loci assembled across at least 75% of samples. One of the sampling sites for Willapa Bay, WA had a low number of individuals after filtering, so individuals from these two sites were combined into one population, for 19 total populations (4–9 individuals per population, mean = 6.2). 41,159 biallelic SNPs across 9,696 GBS loci were genotyped in greater than 75% of these individuals (2.8% of prefiltered loci assembled by *ipyrad*). Average read depth per individual per GBS locus ranged from 21 to 120 (mean = 32$\pm$14). Further filtering by HWE and MAF $>$ 2.5% reduced the dataset to 13,424 SNPs across 6,187 GBS loci (the "combined" dataset).

Three different methods were employed to identify putative SNPs under selection. The number of outliers detected by each program and the overlap between programs is illustrated in Figure S1. *OutFLANK*, as the most conservative of the programs used (Whitlock and Lotterhos, 2015), had the lowest number of outlier markers detected with 31 SNPs across 16 GBS loci. 29 SNPs found across 16 GBS loci were identified as outliers by all three programs. 129 GBS loci contained SNPs identified as outliers by at least two approaches, with 235 SNPs included in the outlier dataset for subsequent population structure analyses.

The neutral dataset, with 13,073 SNPs across 6,057 GBS loci, excluded any SNP found on a GBS locus with an outlier SNP.

1.4.2 Summary statistics, population differentiation, and spatial structure

Summary statistics

Global F_{ST} for outliers (F_{ST} = 0.417) was almost four five times greater than for the combined and neutral SNPs (F_{ST}= 0.105 (combined), 0.097 (neutral)). The outlier dataset had the lowest H_o, but the highest H_e (1.1). Average F_{IS} within populations for the combined dataset was 0.0424, with all populations having a significantly positive F_{IS} value except Ladysmith, BC, Tomales Bay, CA, and South San Francisco Bay, CA which had small, yet significantly negative F_{IS} values. Mugu Lagoon had the highest F_{IS} value (S1). Summary statistics for the six phylogeographic regions identified in the following section are show in Table B1. Summary statistics were quantitatively very similar for the combined and neutral datasets, so that only the results for the outlier and neutral datasets are reported for all subsequent analyses.

Dataset	H_o	H_e	F_{IS}	F_{ST}	Avg. pairwise F_{ST}	Avg. within-pop F_{IS} (Min - Max)
Combined	0.191	0.225	0.051	0.105	0.1026 ($1.83e^{-3}$ - 0.1902)	0.0424 (-0.0960 - 0.1327)
Neutral	0.192	0.224	0.051	0.097	0.095 ($1.38e^{-3}$ - 0.177)	0.0425 (-0.0953 - 0.1288)
Outlier	0.167	0.294	0.030	0.417	0.372 (0.009 - 0.632)	0.0150 (-0.3006 - 0.3332)

Table 1.1: Overall summary statistics for the combined (13,424 SNPs), neutral (13,073 SNPs), and outlier (235 SNPs) datasets. H_o, observed heterozygosity averaged across loci; H_e, expected heterozygosity averaged across loci; F_{IS} & F_{ST}, Wright's F-statistics averaged across loci (Nei and Chesser, 1983); pairwise F_{ST}, average of all pairwise F_{ST} values (Weir and Cockerham, 1984)

Spatial structure

Bayesian population structure analysis from STRUCTURE differed slightly between the neutral and outlier datasets. For neutral SNPs, $K = 5$ had the strongest support based on the Evanno method. Visual inspection of STRUCTURE admixture plots for $K = 6$ included an additional population grouping of Willapa Bay, WA and Coos Bay, WA that was further supported by PCA and F_{ST} results (1.2). I refer to these six groupings as phylogeographic regions: Northwest Vancouver Island, BC (*NWBC*), Puget Sound, WA plus Victoria, BC and Ladysmith Harbour, BC (*Puget+BC*), Willapa Bay, WA plus Coos Bay, OR (*Willapa*), the other two Oregon sites (*Oregon*), Northern California (*NoCal*) from Humboldt Bay, CA to San Francisco Bay, CA, and Southern California (*SoCal*) from Elkhorn Slough, CA to San Diego Bay, CA. STRUCTURE results for the outlier SNPs supported $K = 2$, but visually results were similar between the outlier and neutral SNPs at $K = 5$ with the exception of Discovery Bay, WA in Puget Sound showing higher admixture with *NWBC* (Fig. 1.2). The separation of *Willapa* sites from *Oregon* was not observed using the outlier dataset until $K = 8$.

PCA on both neutral and outlier SNP datasets demonstrated a strong relationship between latitude and the first principal component (PC) (neutral: $R^2 = 0.858$, outlier: $R^2 = 0.826$) (1.2). PCs 2-5 in the neutral dataset separated out individuals by phylogeographic region, with PC2 separating (*Puget+BC*, *NWBC*) and (*Willapa*, *Oregon*), PC3 separating *NoCal* and *SoCal*, PC4 separating *NWBC* and *Puget+BC*, and PC5 separating *Oregon* from *Willapa* (Fig. 1.2). PC1 of the neutral dataset represented 5.8% of the total variance and PCs 2-5 represented 2.5%-1.5% of the variance. The outlier dataset showed similar regional spatial structure for PCs 2-4, but only showed slight separation of *Oregon* from *Willapa* (1.2). PC1 of the outlier dataset represented 23.6% of the total variance, and PCs 2-5 represented 10.3%-2.8% of the total variance.

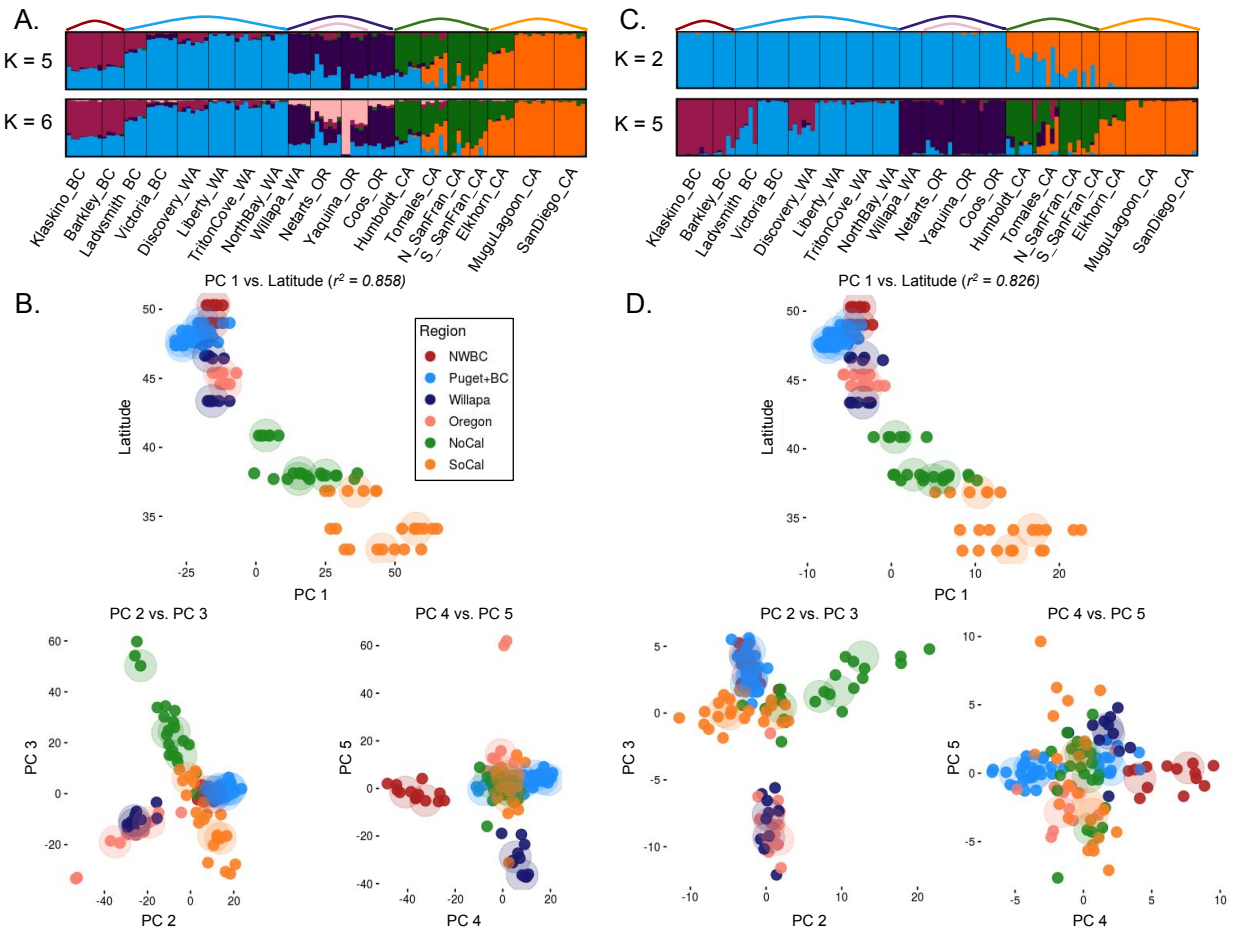


Figure 1.2: Population structure results for 19 *O. lurida* populations using (A-B) neutral loci and (C-D) outlier loci. (A, C) Plots of individual admixture determined using the program STRUCTURE at the K recommended by the ΔK method ($K = 5$ neutral, $K = 2$ outlier), as well as at the value of K inferred from PCA ($K = 6$ neutral, $K = 5$ outlier). (B, D) Principal components analysis plots for PCs 1-5. PC1 is plotted against latitude of sampling site, then PC2 vs. PC3 and PC4 vs. PC5. Large transparent circles indicate the centroid of populations. Colors refer to the phylogeographic regions of each population.

Population differentiation and isolation-by-distance

Pairwise population-specific F_{ST} was higher for outlier SNPs, but both datasets qualitatively illustrated roughly six geographic regions of genetically similar populations and an overall trend of isolation-by-distance, where F_{ST} values were higher between sites that were farther away from each other (1.4, S3, S4).

The three comparisons with the lowest pairwise F_{ST} using the neutral dataset were

Figure 1.3: Illustration depicting Olympia oyster sampling protocol at low tide. (ConcernedApe LLC, 2019)

Willapa Bay, WA/Coos Bay, OR (0.0015), Mugu Lagoon, CA/San Diego Bay, CA (0.0017), and North/South San Francisco Bay, CA (0.0035). Victoria, BC showed higher pairwise F_{ST} with the other British Columbia sites than with sites from Puget Sound, WA. Mantel tests showed a significant correlation between pairwise F_{ST} and coastal water distance for both datasets, indicating a strong trend of isolation-by-distance (p-value = 0.001) (1.4).

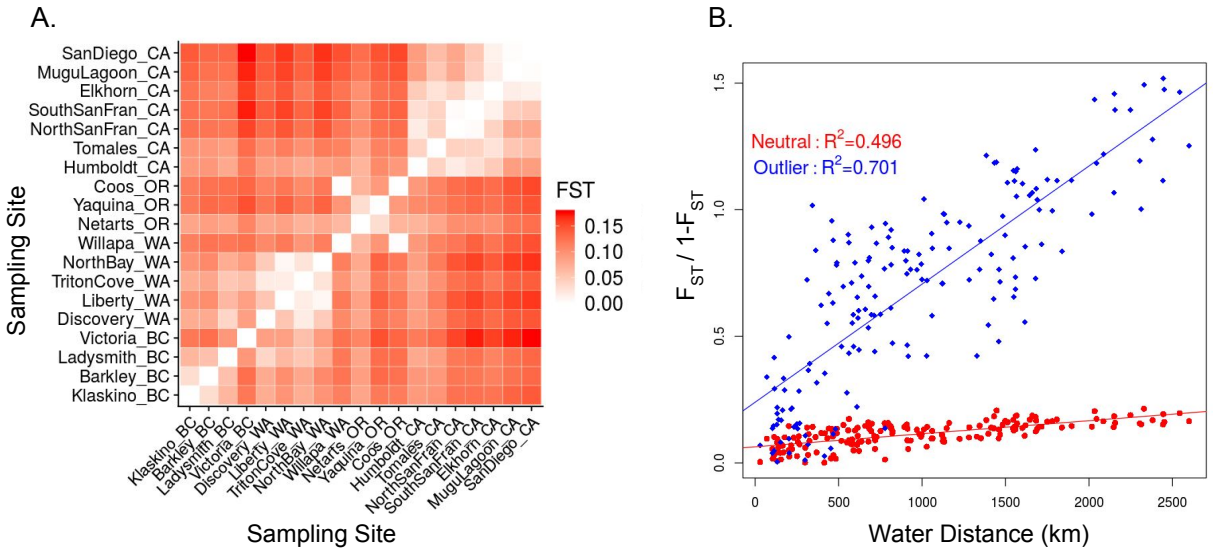


Figure 1.4: A) Heatmap of pairwise F_{ST} values for 19 populations of *O. lurida* using 13,073 neutral SNPs. Populations are ordered from north to south, starting with Klaskino, BC. B) Isolation by distance plot of $F_{ST}/1-F_{ST}$ versus population pairwise coastal water distance. Neutral loci in red circles ($p < 0.001$) and outlier loci in blue diamonds ($p < 0.001$).

1.4.3 Connectivity and historical relationships

TreeMix produced a population graph that supported a major phylogeographic split between the outer coast of Washington and Puget Sound. When allowing for an increasing number of migration events, the proportion of variance in relatedness between populations explained by the model began to asymptote at 0.994 for 7 migration edges (1.5). All of these migration events involve either the *Puget+BC* or *Willapa* regions, except for one from a *NoCal* population to a *SoCal* population.

Because the Coos Bay, OR population is likely a recent anthropogenic introduction from

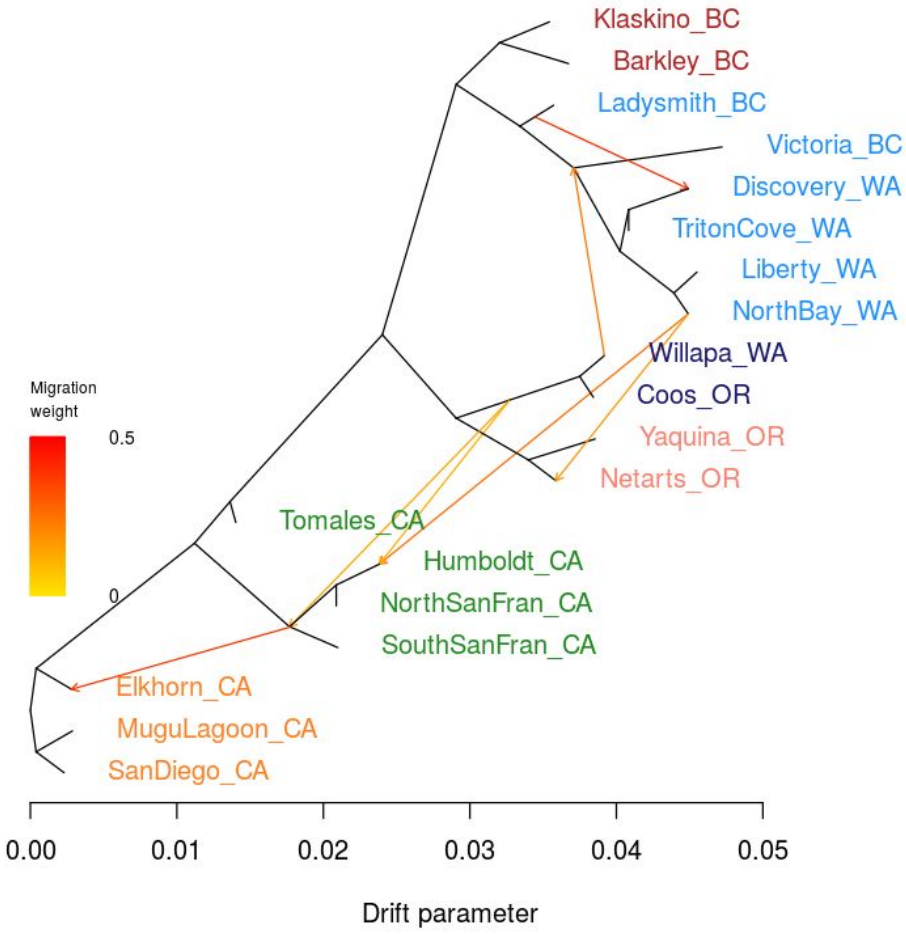


Figure 1.5: TreeMix results for 19 *O. lurida* populations using 1 SNP per neutral locus. Seven migration events are modeled, as this was the best value inferred by evaluating model fit. The tree is rooted by the southernmost populations, San Diego Bay, CA and Mugu Lagoon, CA, and ordered by latitude where possible. Populations are colored by their inferred phylogeographic region.

Willapa Bay, Coos Bay oysters were excluded from the EEMS analysis. The combined EEMS map for all runs identified four significant (posterior probability $> 95\%$) barriers to gene flow: 1) at the mouth of the Strait of Juan de Fuca, 2) around Victoria, BC, 3) extending from Willapa Bay, WA to southern Oregon, and 4) around San Francisco Bay, CA (1.6). These inferred barriers further lend credence to the 6 phylogeographic regions identified through other means. An area of significantly increased gene flow was inferred between Mugu Lagoon and San Diego. The EEMS method also estimated and mapped the genetic

diversity parameter q, which is an estimate of the expected within-deme coalescent time and is proportional to average heterozygosity (H_e). Populations from Oregon northwards had much lower genetic diversity than those in California. A linear regression of population-specific H_e and latitude using the neutral dataset shows a strong relationship between genetic diversity and latitude ($R^2 = 0.862$)(1.7), as did the outlier dataset ($R^2 = 0.834$).

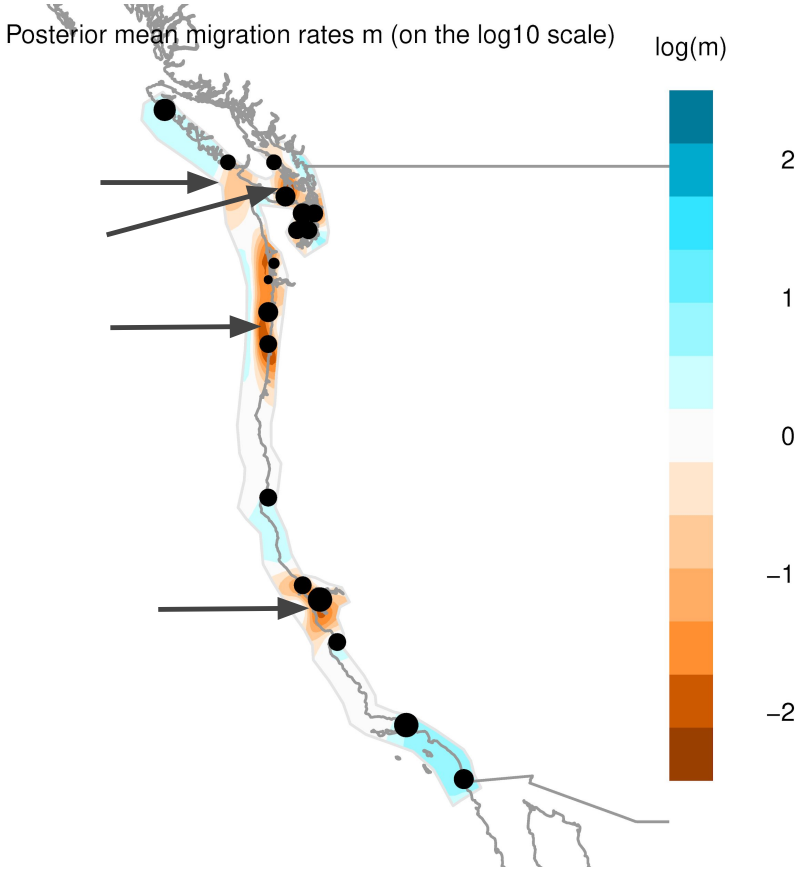


Figure 1.6: Model of effective migration rates as inferred by EEMS for neutral loci in *O. lurida*. Orange represents areas of low migration relative to the average and blue are areas of higher migration. Grey arrows indicate regions of significantly reduced or increased migration (posterior probability $> 95\%$)

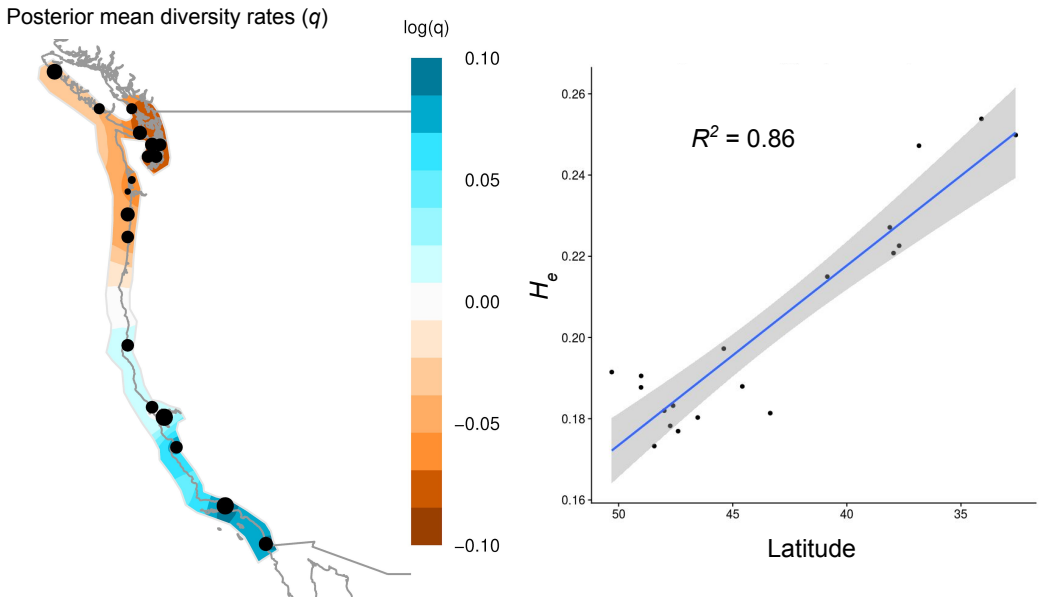


Figure 1.7: Diversity increases from north to south in *O. lurida*. (A) Effective diversity rates as inferred by EEMS, with orange representing areas of lower diversity and blue representing high diversity. (B) Expected heterozygosity (H_e) within each population versus population latitude.

1.4.4 Functional annotation of outliers

The 129 GBS loci containing outlier SNPs were functionally annotated using Blast2GO. Eighteen of these mapped to protein coding genes in the GenBank database, primarily from *Crassostrea virginica* and *Crassostrea gigas*. One mapped to the *O. lurida* mitochondrial NADH dehydrogenase subunit 5 gene (*nad5*), which exhibits high variability in oyster species and is commonly used for metazoan phylogenetics (Xiao et al., 2015). Annotated genes have potential roles in developmental regulation (glyoxalase 3, DNA N6-methyl adenine demethylase-like, transcriptional regulator ERG, serine/threonine-protein kinase), sensory information processing (serine/threonine-protein kinase, sodium-dependent phosphate transport, vesicular glutamate transporter), immune or stress response (*nad5*, E3 ubiquitin-protein ligase, Ty3-G Gag-Pol, helicase domino), energy metabolism (carnitine palmitoyltransferase, glucose dehydrogenase), heavy metal binding (Heavy metal-binding HIP), and

muscle contraction (myosin heavy chain-striated muscle, myosin-XVIIIa) (Table 1.2) (Epelboin et al., 2016; Szent-Györgyi et al., 1999; Li et al., 2017; Cheng et al., 2016; Anderson et al., 2015; Riviere et al., 2013; Pauletto et al., 2017; Wang et al., 2018; Shiel et al., 2017; Pan et al., 2015; de Lorgeril et al., 2005). 21 additional outlier GBS loci had positive matches to InterPro signatures without any BLASTx hits or gene ontology annotation. Plotting minor allele frequency against latitude for outlier SNPs demonstrates that the majority of outliers show a clinal pattern, where one allele is fixed from either Coos Bay, OR or San Francisco Bay, CA to the north, and the other alelle increases in frequency towards the south (Fig. S2).

Locus ID	Gene description	Top hit species
locus_5648	DNA N6-methyl adenine demethylase	*C. gigas*
locus_6412	glucose dehydrogenase [FAD, quinone]	*C. gigas*
locus_7299	transcriptional regulator ERG	*C. gigas*
locus_10670	Fez family zinc finger protein 1	*C. gigas*
locus_44811	sodium-dependent phosphate transport protein 2B	*C. gigas*
locus_50945	glyoxalase 3-like	*C. virginica*
locus_57217	uncharacterized protein	*C. virginica*
locus_98257	uncharacterized protein	*C. virginica*
locus_121489	E3 ubiquitin-protein ligase TRIM9	*C. virginica*
locus_123004	Transposon Ty3-G Gag-Pol polyprotein	*Mizuhopecten yessoensis*
locus_170867	carnitine O-palmitoyltransferase 2, mitochondrial	*C. gigas*
locus_196263	myosin-XVIIIa	*C. gigas*
locus_251628	myosin heavy chain, striated muscle	*C. gigas*
locus_252560	helicase domino-like	*C. virginica*
locus_276278	heavy metal-binding protein HIP	*C. gigas*
locus_277490	NADH dehydrogenase subunit 5, mitochondrion	*O. lurida*
locus_339584	serine/threonine-protein kinase B-raf	*C. virginica*
locus_339916	vesicular glutamate transporter 2.1	*C. gigas*

Table 1.2: BLASTx results for outlier loci. Only the 18 loci with positive BLAST hits are shown.

.

1.5 Discussion

Reduced-representation genomic methods, such as GBS, can greatly inform reintroduction efforts for threatened and exploited species by resolving fine-scaled population structure, providing estimates of genetic connectivity, and identifying informative markers for characterizing adaptive variation (Allendorf et al., 2010; Gagnaire et al., 2015). Using 13,424 GBS-derived SNPs, I characterized the rangewide population structure of the Olympia oyster from southern California to British Columbia and further identified 235 SNPs across 129 GBS loci potentially associated with local adaptation. Contrary to studies in some other marine species, neutral markers had greater power to detect fine-scale population structure compared to outliers. However, outlier loci did provide evidence for adaptive divergence among some populations with high inferred admixture, and are informative as candidate loci involved in local adaptation. This study highlights the importance of using both neutral and outlier markers for conservation and management applications.

1.5.1 Regional population structure and gene flow

Significant population structure was observed across the range of *O. lurida* in both the neutral and outlier markers, with sampling locations structured into six distinct regions. Notably, most of these regions fit well within previously described biogeographical provinces based on marine species distributions (Hall, 1964; Valentine, 1966; Fenberg et al., 2015). In addition to describing the rangewide population structure of *O. lurida*, the large geographic sampling of this study can facilitate the identification of oceanographic features along the eastern Pacific coast that may be important for structuring populations of marine species with similar life histories. Most of the inferred phylogeographic regions are bounded by areas of reduced gene flow, many of which align to oceanographic features that may be acting as barriers to dispersal. Below I discuss these phylogeographic regions and potential barriers in more detail, as well as provide some recommendations for management at local scales.

Southern California (*SoCal*)

The *SoCal* region, containing San Diego Bay, CA, Mugu Lagoon, CA, and Elkhorn Slough, CA, extends across both the Southern Californian and the Montereyan biogeographic provinces as defined by (Hall, 1964), with Monterey Bay as the northern boundary. Monterey Bay is a known biogeographic barrier for some marine algae (Abbott and Hollenberg, 1976), and has been proposed as a potentially important barrier to gene flow in marine invertebrates as well (Dawson, 2001). This region extends across Point Conception, which is a well-known site of species turnover (Valentine, 1966) and a phylogeographic barrier for some taxa (Marko, 1998; Wares et al., 2001). This finding is consistent with meta-analyses demonstrating that strong population structure across Point Conception is the exception rather than the rule for many marine invertebrates (Kelly and Palumbi, 2010; Dawson, 2001). Finer-scaled sampling of populations on either side of Point Conception may provide evidence for slight genetic clines undetected by the current study.

SoCal exhibits the highest genetic diversity of any region, for which I propose three nonexclusive mechanisms. 1) The southward direction of the California Current results in asymmetric gene flow and an accumulation of genotypes in the south (Wares et al., 2001). This hypothesis is supported by the inferred directionality of migration events in TreeMix (1.5). 2) Northern populations exhibit lower genetic diversity due to repeated extirpation or population bottlenecks from glaciation cycles (see *Puget+BC*) (Marko, 2004). 3) Ongoing or historical admixture from the southern sister species *O. conchaphila* has increased genetic diversity in these populations. Sampling and genotyping of *O. conchaphila* is underway to test this hypothesis. The low F_{ST} between Mugu Lagoon and San Diego suggests either a recent transplantation between sites or high gene flow. If the former, I hypothesize that Mugu Lagoon is the recent transplant due to a high inbreeding coefficient (F_{IS}). Nevertheless, three outlier loci exhibited allele frequency shifts of at least 50% between these two populations, suggesting some potential adaptive population divergence.

Northern California (*NoCal*)

San Francisco Bay, Tomales Bay, and Humboldt Bay constitute the *NoCal* region, which is encompassed by the northern half of the Montereyan biogeographic province as identified by (Fenberg et al., 2015) and delineated by Cape Mendocino to the north. Cape Mendocino, located 46 km south of Humboldt Bay, is an established phylogeographic break for multiple marine species (Kelly and Palumbi, 2010). EEMS identifies an area of significantly reduced gene flow surrounding San Francisco Bay, which may correspond to Monterey Bay (Dawson, 2001), or anthropogenic introductions (see *Anthropogenic influences on population structure*). The two sites within San Francisco Bay (Candlestick Park and Point Orient), exhibit potential adaptive divergence at some outlier GBS loci despite high potential for gene flow. This result supports evidence for local adaptation from reciprocal transplant studies within San Francisco Bay (Bible and Sanford, 2016), and highlights the importance of taking individual pairwise F_{ST} values (S3, S4) into account when making reintroduction decisions. TreeMix inferred significant migration between San Francisco Bay and Elkhorn Slough, however migration is likely not consistent between these populations based on synchrony of recruitment dynamics (Wasson et al., 2016).

Oregon and *Willapa*

Both the *Oregon* region, comprised of Netarts Bay and Yaquina Bay, and the *Willapa* region with Willapa Bay, WA and Coos Bay, OR, fall within the Mendocinian biogeographic province, which is usually demarcated by either Cape Flattery (Blanchette et al., 2008) or Vancouver Island (Fenberg et al., 2015) to the north. Evidence from TreeMix and Structure indicate that these two regions have a shared phylogeographic history—likely a combination of evolutionary and anthropogenic processes. EEMS robustly infers an area of significantly reduced migration from Willapa Bay, WA to southern Oregon, which I hypothesize is partly due to the high retention of oyster larvae within Willapa Bay during the summer reproductive season (Banas et al., 2009). EEMS also infers an area of slightly increased migration

to the west of the sampling sites. This result may be the EEMS model attempting to incorporate evidence for long-range migration events, likely anthropogenic in nature, between *Puget+BC* sites and sites in *Oregon* and *NoCal*, or it may be an artifact of the model. To my knowledge, this is the first application of EEMS to a marine invertebrate—simulations and additional empirical studies are necessary to evaluate the behavior of EEMS in linear habitats. Currently the protections against importing shellfish from outside of the state are higher than moving shellfish within the state. The strong phylogeographic divide between Willapa Bay, WA and Puget Sound, WA presented here indicates that transfer of Olympia oysters or *Crassostrea* shells between the outer coast of WA and Puget Sound should be considered equivalent to importing oysters from out of state.

Puget Sound, WA and British Columbia (*Puget+BC* and *NWBC*)

The *NWBC* region, comprised of Klaskino Inlet, BC and Barkeley Sound, BC, is significantly differentiated from other sites on Vancouver Island and shows evidence for decreased migration out of the region. The *Puget+BC* region is comprised of Ladysmith Harbour, BC, Victoria Gorge, BC, and all four sites in Puget Sound, WA. Strong evidence suggests that Victoria Gorge, BC has a shared evolutionary history with Puget Sound, WA, although EEMS indicates that migration is reduced between these sites. Ladysmith Harbour may belong to a separate phylogeographic region all together, as this site was intermediate between *NWBC* and *Puget+BC* regions in the STRUCTURE, PCA, and TreeMix analyses. Genetic sampling from additional sites on the central coast of British Columbia and eastern coast of Vancouver Island could test this hypothesis.

The separation of these two regions from those to the south corroborates previous evidence from mitochondrial loci of a strong phylogeographic divide (Polson et al., 2009). Although Cape Flattery and Puget Sound itself have both been classified as biogeographic barriers due to a bifurcation in ocean currents (Valentine, 1966; Kelly and Palumbi, 2010), there are surprisingly few studies evaluating the genetic structure of species found both

within Puget Sound and on the outer coast of Washington. Those that do focus on species with much longer dispersal times than *O. lurida* (Buonaccorsi et al., 2002; Cunningham et al., 2009; Iwamoto et al., 2015; Siegle et al., 2013; Jackson and O'Malley, 2017). To my knowledge, this is the first study in a marine mollusc to evaluate and identify significant population differentiation among Puget Sound populations and the outer coast. More studies are required to fully characterize the importance of this barrier across marine taxa.

Genetic differentiation within Puget Sound is relatively low at both neutral and outlier markers, with the exception of the northernmost site, Discovery Bay. The weak population structure within Puget Sound and the overall low genetic diversity in northern sites is likely due to recent genetic bottlenecks and range expansion after the last glacial maximum, which reached just north of Willipa Bay, WA (49°N latitude) until 12-13 kya (Dyke and Prest, 1987). Despite such low genetic differentiation, experimental assessments of local adaptation for populations within Puget Sound have detected heritable differences in fitness traits such as reproductive timing, growth rate, and gene expression in response to stress (Heare et al., 2017, 2018; Silliman et al., 2018). These results, coupled with experimental evidence for local adaptation to salinity among Northern California populations (Bible and Sanford, 2016), suggest that adaptive divergence in this species can occur in the face of high gene flow.

1.5.2 Anthropogenic influences on population structure

The evidence for reduced effective migration, low differentiation within most of the phylogeographic regions, and external estimates of effective dispersal (Carson, 2010), suggests that long distance dispersal is not a significant force in shaping population structure in this species. However, TreeMix inferred a few such migration events that cross aforementioned barriers to gene flow. To explain this evidence, I investigated the history of Olympia oyster exploitation and aquaculture through literature reviews, technical reports, grey literature, historical first-person accounts, and discussions with current restoration practitioners. The historical impact of human take and transportation on the Olympia oyster is substantial.

Beginning in 1850, oysters were shipped from Willapa Bay to northern California by the millions, including shipments of juvenile or seed oysters to be raised in local waters until reaching commercial size. The inferred migration events from North Bay, WA to California sites may be reflecting the historical transplantation of seed oysters from Oakland Bay, WA, about 30 km from North Bay (Woelke, 1959; Baker, 1995). After the crash of the Olympia oyster industry, the non-native oysters *Crassostrea virginica* and *C. gigas* were brought to the west coast for commercial aquaculture. The shells of these species are excellent substrate for Olympia oysters, and the movement of *Crassostrea* oysters between bays for culturing purposes (e.g., San Francisco to Humboldt in 1910, Willapa to Humboldt in 1950s (Barrett, 1963)) may have resulted in the accidental transfer of *O. lurida* (Townsend, 1895). The low F_{ST} of Willapa Bay, WA and Coos Bay, OR, despite being separated by 415 km, corroborates the theory of an accidental introduction of Olympia oysters from Willapa Bay on *C. gigas* shells in the 1980s (Baker et al., 1999). The low inbreeding coefficient for Coos Bay suggests a potentially large founding population. Future movement of *Crassostrea* for aquaculture purposes should be carefully monitored to prevent the accidental migration of nonnative *O. lurida* genotypes.

While it is encouraging that programs such as TreeMix can recover known human-mediated migration events, such artificial movement of individuals can complicate the determination of natural connectivity patterns. For example, the area of low effective migration inferred around San Francisco Bay may be due to the introgression of Washington genotypes rather than actual physical barriers to gene flow. Fortunately, intentional movement of Olympia oysters between regions ceased over 80 years ago with the exception of restoration efforts in Netarts Bay from 2005-2012, which utilized some broodstock from Willapa Bay.

1.5.3 Local adaptation

Detection of outlier loci using three different methods conservatively identified 129 GBS loci as under putative selection. Only 18 GBS loci mapped to protein-coding regions, 3 of

which were identified as outliers by all three approaches. Mapping of outlier GBS loci to the forthcoming *O. lurida* genome will aid in detecting loci that may be tightly linked to a gene or regulatory region. Direct information about the function of genes or proteins in oysters is sparse, but rapidly increasing with transcriptomic and physiological studies on the commercially important *Crassostrea* species. While these 129 GBS loci are likely only a fraction of all loci under divergent selection across the *O. lurida* genome (Lowry et al., 2017) and their functional associations are strictly hypotheses, they are nonetheless excellent candidates for future directed studies.

Plotting minor allele frequency against latitude demonstrates that the majority of outlier SNPs show a clinal pattern, where one allele is fixed north of Coos Bay, OR and the other increases in frequency towards the south (Figure (S2). Many clinal GBS loci with functional annotations are associated either directly or indirectly with development. Glyoxalase 3 expression has been linked to developmental competence in female oyster gametes (Pauletto et al., 2017), and DNA N6-methyl adenine demethylase has been linked to developmental timing in oysters (Riviere et al., 2013). ERG transcriptional regulator, kinase B-raf, and Fez family zinc finger protein are likely to be directly involved with developmental regulation (Gaitán-Espitia and Hofmann, 2017; Epelboin et al., 2016). Olympia oysters exhibit latitudinal variation in gonad development and spawning, with California oysters initiating spawning up to 6°C warmer than those from Puget Sound, WA (Coe, 1932; Hopkins, 1937). Recent evidence suggests that there is heritable, adaptive variation in reproductive timing, even among populations of oysters within the same phylogeographic region (Silliman et al., 2018; Barber et al., 2016). Another clinal locus of interest mapped to the mitochondrial gene carnitine Ol-palmitoyltransferase, which has been strongly associated with regulation of glycogen content (and therefore, tastiness) in *C. gigas* (Li et al., 2017). Other clinal genes have putative functions in sensory information processing and muscle contraction.

Some outlier loci exhibit the opposite of a clinal pattern, where populations in the middle of the range predominantly have a different allele than the northern and southern popula-

tions. E3 ubiquitin-protein ligase (locus_121489) is diverged in *Oregon* and Willapa Bay, WA compared to the other populations. An E3 uniquitin-protein ligase was recently identified as an important component of the neuroendicrine-immune response in *C. gigas* and is primarily expressed in the gonads. Heavy metal-binding protein (HIP) also exhibits this hump-shaped distribution of allele frequencies.

1.5.4 Potential Limitations

Although genomic methods such as GBS have proven useful for evolutionary biology and conservation genetic studies (Andrews et al., 2016), several potential limitations of GBS and the present study should be addressed. Non-random missing data due to polymorphisms in the restriction enzyme cut site ("allelic dropout") can bias population genetic analyses by underestimating genomic diversity and overestimating F_{ST}, however the impact of these biases on F_{ST} may be limited if effective population size (N_e) is small and if loci with large amounts of missing data are removed from analyses (Gautier et al., 2013; Cariou et al., 2016). Due to large variation in reproductive success every generation, N_e is likely small for *Ostrea* species (Hedgecock, 1994; Lallias et al., 2010). Loci with $> 25\%$ missing data were removed from population genetic analyses, and preliminary analyses allowing 40%–10% missing data still resulted in the same regional population structure and relative values of pairwise F_{ST}, although absolute values of F_{ST} changed slightly. Two reasons may underlie the large number of individuals (128) removed during filtering. First, too many individuals may have been pooled per sequencing lane given the number of loci targeted, resulting in low sequencing depth for some individuals (Andrews et al., 2016). Second, these libraries were made and sequenced in-house as opposed to a dedicated commercial GBS facility. The protocol learning curve may be why a disproportionate number of individuals failed or had low sequencing depth in the first few prepared libraries. This filtering resulted in 4–9 individuals per population in the final dataset, which is sufficient for estimating F_{ST} when $>$ 1,000 SNPs are used (Willing et al., 2012). While these small population sizes may limit

the power to detect outlier loci (Foll and Gaggiotti, 2008), the probability of false positives is reduced by comparing across multiple outlier methods (Rellstab et al., 2015). Lastly, while methods like EEMS and PCA can characterize genetic differentiation, they cannot distinguish between the different demographic scenarios that may result in these patterns (Petkova et al., 2016).

CHAPTER 2

EVIDENCE FOR ADAPTIVE DIVERGENCE WITH GENE FLOW[2]

2.1 Abstract

Adaptive evolution and plasticity are two mechanisms that facilitate phenotypic differences between populations living in different environments. Understanding which mechanism underlies variation in fitness-related traits is a crucial step in designing conservation and restoration management strategies for taxa at risk from anthropogenic stressors. Olympia oysters (*Ostrea lurida*) have received considerable attention with regard to restoration, however there is limited information on adaptive population structure. Using oysters raised under common conditions for up to two generations (F1s and F2s), we tested for evidence of divergence in reproduction, larval growth, and juvenile growth among three populations in Puget Sound, Washington. We found that the population with the fastest growth rate also exhibited delayed and reduced reproductive activity, indicating a potential adaptive tradeoff. Our results corroborate and extend upon a previous reciprocal transplant study on F1 oysters from the same populations, indicating that variation in growth rate and differences in reproductive timing are consistent across both natural and laboratory environments and have a strongly heritable component that cannot be entirely attributed to plasticity.

2.2 Introduction

Natural environments exhibit spatial heterogeneity in both abiotic and biotic factors, oftentimes driving populations to evolve traits that confer a fitness advantage in their native habitat over foreign genotypes (Kawecki and Ebert, 2004). Conclusively demonstrating

2. A version of this chapter has been published as: Silliman KE, Bowyer TK, and Roberts SB. 2018. Consistent differences in fitness traits across multiple generations of Olympia oysters. *Scientific Reports*. 8:6080.

adaptive divergence is complicated by phenotypic plasticity, where individuals adjust their phenotype according to the conditions they experience West-Eberhard (2003), which may confound inferences of local adaptation (Kawecki and Ebert, 2004; Teplitsky et al., 2008). Phenotypic plasticity is widespread in marine species (Conover et al., 2006; Pepin, 1991; Padilla and Savedo, 2013), and for marine invertebrates the most common trigger for plasticity appears to be the abiotic environment (Padilla and Savedo, 2013). Organisms can be raised their entire lives in common conditions in order to minimize the effects of phenotypic plasticity. However this approach may fail with strong transgenerational plasticity (TGP)—defined here as when the environment or phenotype of the parent affects the phenotype of the offspring (Kawecki and Ebert, 2004; Guillaume et al., 2016). The ideal experimental design to distinguish TGP from genetic change involves raising and breeding individuals for at least two generations in a common setting (Kawecki and Ebert, 2004), although TGP has persisted for more than two generations in some laboratory studies (Hercus and Hoffmann, 2000). Whenever breeding organisms for multiple generations, care should be taken to evaluate and reduce the influences of artificial selection and genetic drift (McClure et al., 2008). In a recent review of experimental evidence for local adaptation in marine invertebrates, only 11 out of 59 studies utilized 2 or more generations (Sanford and Kelly, 2011). Distinguishing between plastic responses and adaptive evolution to the environment is key to understanding the potential for marine species to acclimate or adapt to changing environmental conditions Salinas et al. (2013).

Marine molluscs, and bivalves in particular, constitute some of our most economically and ecologically important marine invertebrates. Like many other marine invertebrates, they exhibit complex life cycles which include both planktonic larval stages as well as benthic juvenile and adult stages. The larval stage of many marine molluscs has been shown to be particularly sensitive to risks from ocean acidification and warming (Parker et al., 2013; Byrne, 2011), resulting in an increasing need to understand the relative importance of adaptive and plastic processes in shaping phenotypic variation. Evidence that TGP might

be common for marine molluscs is growing (see Ross et al. (2016) for a thorough review), however only a handful of studies investigating adaptive differentiation in this clade have compared organisms raised in common conditions for at least 2 generations, and of those most involved Gastropoda (see Kuo and Sanford (2009); Palmer (1994); Sanford and Worth (2010); Dittman et al. (1998); Bible and Sanford (2016)).

There is political and economic pressure to restore abundance and recover ecosystem services offered by *O. lurida*, which has spurred increasing interest in understanding the genetic, phenotypic, and epigenetic variation at both local and regional scales Camara and Vadopalas (2009). A recent study on Olympia oysters in central California provided evidence for spatial adaptive differentiation through a reciprocal transplant experiment with first-generation laboratory-reared (F1) oysters. The authors also found suggestive, although not statistically significant, evidence of population-level differences in low salinity tolerance in second-generation, laboratory-reared (F2) oysters, and hypothesized adaptive divergence may occur in Olympia oysters over distances as short as 20-100 km Bible and Sanford (2016). In Puget Sound, WA, Heare et al. 2017 Heare et al. (2017) conducted a reciprocal transplant experiment with F1 Olympia oysters from three distinct populations (Dabob Bay, Oyster Bay, and Fidalgo Bay) in 2014. Variation in survival, growth rate, and reproductive activity was observed among populations and the four transplant sites. In particular, oysters from Fidalgo Bay had faster growth rates and reduced or delayed reproductive activity at most sites, while oysters from Dabob Bay had better survival yet slower growth rates, indicating potential adaptive trade-offsHeare et al. (2017).

Although both of these studies controlled for environmental exposure of broodstock for up to 5 months prior to producing F1 oysters and attempted to maximize genetic diversity, they did not sufficiently minimize concerns about TGP or evaluate the relatedness of their laboratory-reared populations. This study aims to mitigate the influence of TGP on inferring adaptive differentiation by testing if phenotypic differences among populations of Olympia oysters are consistent across generations. In the summer of 2015, we conducted a common

garden experiment on F1 and F2 oysters derived from the same three Puget Sound populations as Heare. Three fitness-related traits were measured across populations- reproductive activity, larval growth rate, and juvenile growth rate. We also estimated the relatedness and genetic diversity of the F1 generation using SNP data in order to demonstrate that the population-specific traits described here are not primarily due to family-level variation from few effective breeders.

2.3 Materials and Methods

2.3.1 Broodstock

Adult oysters were collected from three locations in Puget Sound, Washington; Fidalgo Bay (N 48.478252, W 122.574845), Oyster Bay (N 47.131465, W 123.021450), Dabob Bay (N 47.850948, W 122.805694) during November and December 2012. Oysters were held for 5 months in common conditions in Port Gamble, Washington and spawned in June 2013. Unlike many other oyster genera that broadcast spawn both eggs and sperm (e.g., *Crassostrea*), Olympia oyster females are fertilized with spermatozeugmata ('sperm packets') from the water column and brood larvae for approximately 10-12 days. After being released into the water column, larvae are planktonic for approximately two weeks before attaching to a hard substrate ('settlement'). To ensure genetic diversity, each population was allowed to spawn in 24 separate groups of 20-25 oysters. Larvae produced from each population were reared in tanks based on spawning group, settled on very small pieces of oyster shell, then fed ad libitum. In August 2013, 480 juvenile oysters (5-10 mm) from each source population were outplanted at Clam Bay located in central Puget Sound (N 47.571839, W 122.550813), a different site than any of the source populations. For the purposes of this study we will refer to the cohort outplanted at Clam Bay as F1s. Reproductive and growth characteristics of F1 oysters at Clam Bay have been described by Heare et. al. 2017. In June 2015, F1 oysters were moved into NOAA's Kenneth K. Chew Center for Shellfish Research and Restoration

in Manchester, WA and maintained in mesh bags suspended in separate 18.9 L buckets with a diet of mixed live algae in flowing seawater.

2.3.2 Larval Rearing and Quantifying Reproductive Activity

Spawning of broodstock (F1) was induced by elevating temperature to 18-20 °C in May 2015 and maintaining algae supplementation at 60,000-80,000 cells/mL. To ensure genetic diversity, each population (Fidalgo Bay = 101, Oyster Bay = 100, Dabob Bay = 94) was divided into 5 groups of 16-21 oysters. In addition to ensuring that multiple females were involved in reproduction, these replicate groups allowed us to statistically test population-level differences in reproductive activity. Larval release was checked and quantified every one to three days with larvae filtered (100 μM) and counted with triplicate drop counts. A cohort of these F2 larvae were used in the larval growth trials with the remainder raised in tanks (100 L) for the juvenile growth trial.

2.3.3 Experimental Set-up: Larvae Growth

To investigate differences in growth rate among populations at the larval stage, we set up a larval growth rate experiment starting on a day when all three populations were producing at least 25,000 total larvae across multiple buckets (i.e. multiple females). Larvae were pooled by population and mixed well in 1 L beakers. The concentration of larvae was estimated using triplicate drop counts, and 900 larvae per population were added to each of three replicate plastic beakers (1 L) in order to account for random effects due to the beaker environment. These beakers were each outfitted with a "silo", a 7.62cm (3") diameter section of conditioned PVC pipe covered in 100 μM mesh on the bottom. Larvae remained in the silo, which allowed for daily water changes by lifting up the silo and inserting into a new beaker with premixed fresh seawater (800 ml) containing a 50/50 mix of live *T. isochrysis* and diatom algae (final concentration 60,000-80,000 cells/mL). Approximately 20 larvae were sampled haphazardly by pipette from the initial larval pools at Day 0 and each replicate at Day 7 and Day 14

of the experiment, fixed in 10% buffered formalin, and photographed under a microscope for analysis using ImageJ v1.51 Schneider et al. (2012) software to determine shell area and length. We report results based on shell length, although shell area gave qualitatively similar results (not shown).

2.3.4 Experimental Set-up: Juvenile Growth

Larvae from each population greater than 224 μM (n=30,000) were moved to new tanks (100 L) where air was bubbled to maintain oxygen levels and stimulate water movement. These tanks were lined with PVC tiles (10x10cm) roughed on one side using coarse sandpaper to promote oyster settlement. After four weeks, settled oysters were culled to fewer than 30 oysters/tile to avoid overgrowth interactions (Fidalgo tiles = 10, Oyster Bay tiles = 7, Dabob tiles = 8). Tiles were randomized and attached to four protected outplanting trays that were suspended from the dock at NOAA Manchester Research Station (depth = 6m). Photos were taken of oysters on tiles prior to outplanting (Day 0), after 48 days, and after 68 days for analysis using ImageJ software to determine shell area.

2.3.5 2b-RAD Genotyping

Using a 2b-RAD reduced-representation sequencing approach Wang et al. (2012), we sequenced 279 individuals and 18 technical replicates from the F1 generation for a total of 297 samples across 5 lanes of Illumina HiSeq. The technical replicates are necessary for quality assessment and genotyping recalibration of downstream analyses. For each population, we mapped reads from the 50 individuals with the highest read depth as well as 4 -5 technical replicates to a draft *O. lurida* genome of 8,733 scaffolds over 10kb in length. For each sample in this subset, approximately 30% of reads mapped to the draft genome using Bowtie2. After genotyping single nucleotide polymorphisms (SNPs) using the UnifiedGenotyper tool in GATK v3.6 McKenna et al. (2010), we filtered SNPs for quality and excess heterozygosity based on Hardy-Weinberg equilibrium within populations using VCFTools Danecek et al.

(2011) and custom scripts by Mikhail Matz (`https://github.com/z0on/2bRAD_GATK`). For subsequent analyses of relatedness and genetic diversity, we thinned our dataset to one SNP per 2b-RAD tag for 677 SNPs confidently identified in at least 75% of individuals.

2.3.6 Analysis

To compare differences in reproductive activity among populations in the F1 generation, the five separate broodstock buckets per population were considered as independent replicates. The number of larvae released in each bucket on each sampling day was normalized by the number of adults in that bucket. To determine if there was a difference in total reproductive output among populations, the cumulative number of larvae produced throughout the 7 weeks for each bucket was analyzed using a one-way analysis of variance (ANOVA; R base) with post hoc Tukey's tests. In order to determine if populations differed in their timing of reproductive activity, one-way ANOVAs were also conducted on the number of days until the first observed larval release in each bucket after spawning conditions were induced, as well as the number of days until spawning peaked in each bucket.

Linear mixed models (LMMs) were used to measure the effect of source population on oyster growth using the R package *lme4* Bates et al. (2015). For the larval growth experiment, randomly selected live larvae were measured on Day 0, Day 7 and Day 14 with 10-12 larvae (mean=11.8, std=0.62) larvae per replicate. Dead oysters are easily distinguished from living oysters by having an entirely clear protoshell and no observable tissue. As *O. lurida* larvae grow prior to maternal release, a one-way ANOVA was used to test if size varied among populations on Day 0 prior to separating out into replicate beakers. For the LMM analysis, population was a fixed effect and replicate beaker was a random effect. Prior to running the LMM, size distributions were tested for normality using the Shapiro-Wilkes test with the stats R package. Significance of the LMM results was established using a Likelihood Ratio Test against a null model based only on random effects. Shell length was compared at each time point using pairwise t-tests with a Bonferroni correction for multiple testing.

For the juvenile experiment, growth was tracked on an individual basis (growth rate = Δarea/# days). All oysters that were alive (as determined by a healthy shell color and response to prodding) and not extending >50% off the tile at each time point were measured for shell area. For the linear mixed model, population was a fixed effect and tray containing the tile was a random effect. Growth rate was natural log-transformed based on indications of non-normal distributions from the Shapiro-Wilkes test. Pairwise comparisons for populations at each time point were performed with the Nemenyi post hoc test using information from the Kruskal-Wallis test (*PMCMR* package) Pohlert (2014).

Estimates of genetic diversity and relatedness in the F1 generation were calculated in R v3.4.1 using 677 high-quality SNPs obtained by 2b-RAD sequencing. Mean expected and observed heterozygosity was calculated for each population using hierfstat Goudet and Jombart (2015) and compared with Bartlett's test of homogeneity of variances Bartlett (1937). Pairwise estimates of relatedness were calculated using the KING algorithm Manichaikul et al. (2010) as implemented in VCFtoolsDanecek et al. (2011). Full sib pairs are classified by having a kinship coefficient over $\frac{1}{2^{\frac{3}{2}}}$. Sibship was estimated in each population using a full-maximum likelihood model as implemented by Colony v. 2.0.6.4 Wang and Santure (2009); Wang (2004) with the following parameters: polygamous males and females, long run length, full-likelihood analysis, high-likelihood precision, update allele frequencies during run, no prior information, and allelic dropout rate of 0.001.

2.4 Results

2.4.1 Reproductive Activity

The timing and quantity of larvae produced varied across the three populations (Fig. 2.1). The cumulative number of larvae produced over a 7 week period differed significantly among population (one-way ANOVA, Df = 2, F = 4.174, $p = 0.0421$), with F1 oysters from Fidalgo producing the fewest (Fig. 2.2a). Combined across all replicates, Oyster Bay oysters pro-

duced 2.7 million larvae, Dabob oysters produced 2.4 million, and Fidalgo oysters produced 1.1 million.

The onset of larval release also differed significantly among populations (one-way ANOVA, Df = 2, F = 4.033, $p = 0.0457$), with Fidalgo oysters exhibiting delayed reproduction compared to the other populations and much higher variance in the date of initial larval release (Fig. 2.2b). The timing of peak larval production did not vary significantly among populations (ANOVA, Df = 2, F = 0.097, $p = 0.908$).

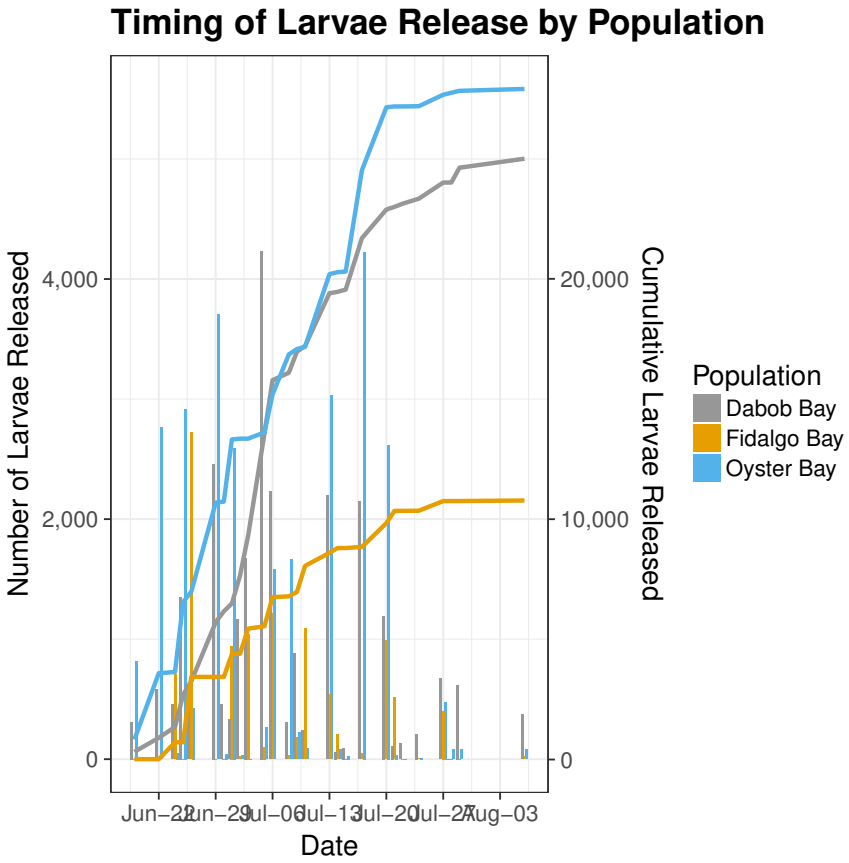


Figure 2.1: Reproductive activity in F1 oysters from three Puget Sound populations. The left axis measures the number of larvae released on each sampling day, summed up across replicates and normalized by the number of adult oysters in each replicate. The right axis measures the cumulative number of larvae produced through time, normalized by the number of adult oysters in each population.

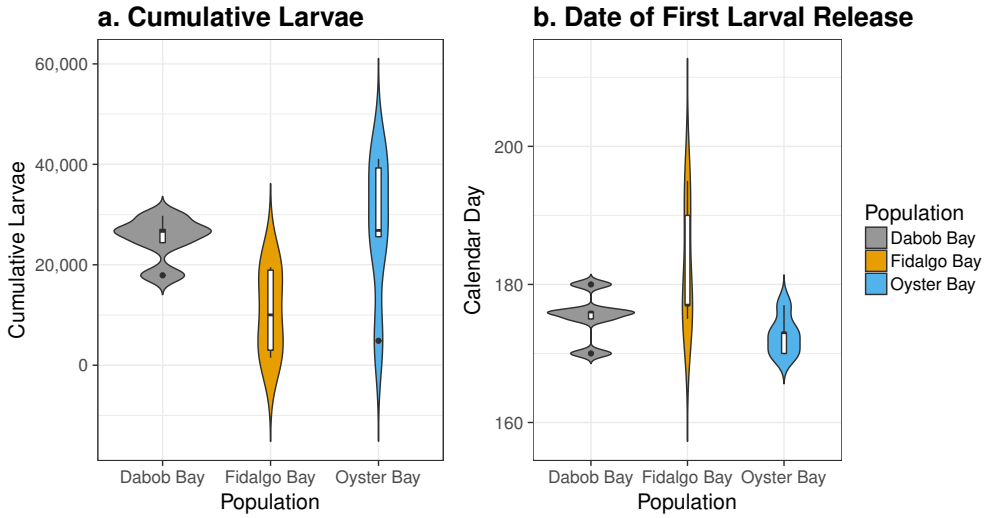


Figure 2.2: **a**: Cumulative number of larvae released within each replicate bucket over 7 weeks, normalized by the number of oysters in each bucket. Oysters from Fidalgo produced significantly fewer than those from Oyster Bay (Tukey post hoc test, $p = 0.0499$) and fewer, although not significantly, than those from Dabob (Tukey post hoc test, $p = 0.0954$). **b**: Calendar day of first observed larvae release. Fidalgo oysters released larvae 10 days later than Oyster Bay oysters (Tukey post hoc test, $p = 0.0434$) and 7 days later than Dabob oysters on average, although this was not statistically significant (Tukey post hoc test, $p = 0.156$).

2.4.2 Larval Growth Experiment

Significant differences in larvae size were not detected among populations on Day 0 of the larval growth experiment (one-way ANOVA, Df = 2, F = 0.939, p =0.401). By Day 7, size varied significantly among populations (linear mixed model(LMM), $p = 7.939e^{-5}$). Fidalgo larvae were 8% larger than Dabob larvae (t-test, $p = 1.8e^{-6}$) and 6% larger than Oyster Bay larvae (t-test, $p = 0.00026$). After 14 days, size still varied significantly among populations (LMM, $p = 0.03573$), but only the comparison between Fidalgo and Dabob larvae remained significant (9% larger; t-test, $p = 0.0017$) (Fig. 2.3).

2.4.3 Juvenile Growth Experiment

Significant differences in juvenile shell area at Day 0 were not detected (one-way ANOVA, Df = 2, F = 0.483, $p = 0.617$). Growth rate between Day 0 and Day 48 diverged among

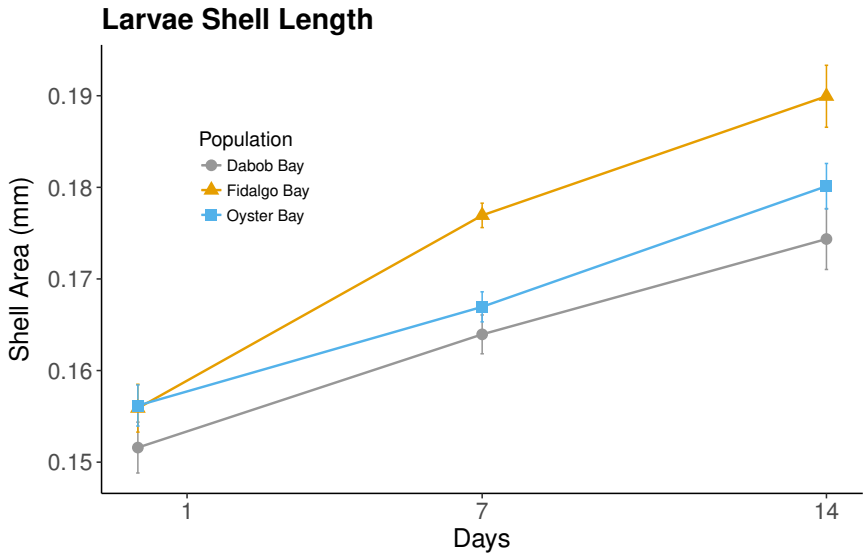


Figure 2.3: Larval shell length of F2 oysters from three populations over 14 days. Data are means across replicates += s.e.m. Size varied significantly among populations at Day 7 (LMM,$p = 7.939e^{-5}$) and Day 14 (LMM, $p = 0.03573$).

populations (LMM, $p = 0.02236$). Fidalgo oysters grew 46% faster than Dabob oysters (Kruskal-Wallis post hoc test, $p = 0.011$), but all other pairwise comparisons were not significant. Between 48 and 68 days, shell growth continued to differ among populations (LMM, $p = 0.0012$). Dabob oysters grew slower over this time period than Fidalgo oysters (Kruskal-Wallis post hoc test, $p = 0.0027$) and tended to grow slower than Oyster Bay oysters (Kruskal-Wallis post hoc test, $p = 0.0806$)(Fig. 2.4).

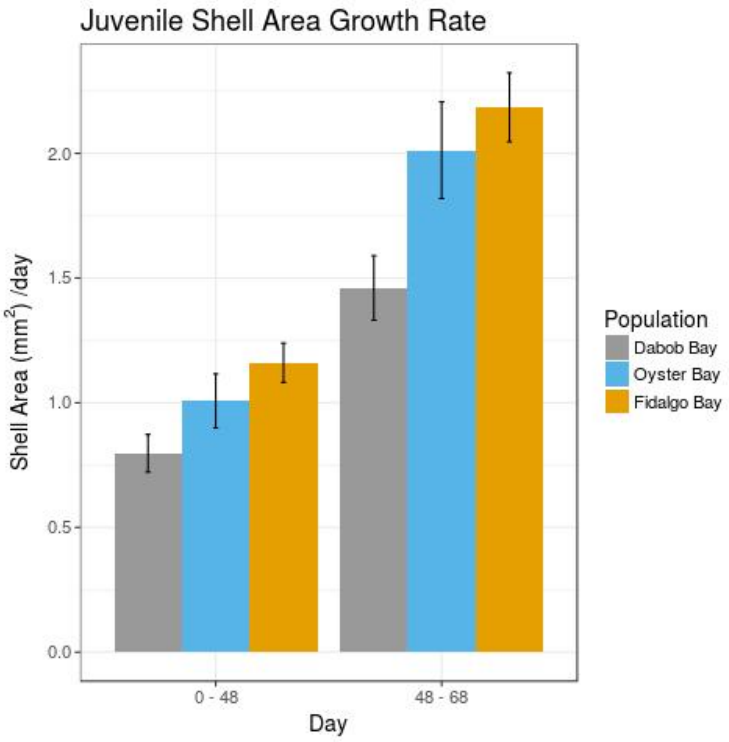


Figure 2.4: Juvenile shell area growth rate of F2 oysters from three populations over 9 weeks. Juvenile shell growth of F2 oysters (growth rate = Δarea / # days). Growth rate between Day 0 - Day 48 differed significantly among populations (LMM, $p = 0.02236$) as well as between Day 48 - Day 68 (LMM, $p = 0.0012$).

2.4.4 Genetic Diversity and Relatedness

Population-level estimates of genetic diversity and relatedness for F1 oysters are shown in Table 2.1. Observed heterozygosity (H_o), expected heterozygosity (H_e) and inbreeding coefficient (F_{is}) represent values averaged over loci that are polymorphic within the population, while kinship coefficient (ϕ) is averaged over all pairwise comparisons of individuals. Values of ϕ over 0.1768 are considered full-sibs. Observed heterozygosity varied across loci and did not significantly differ from expected heterozygosity except in Dabob oysters (Bartlett's test, $p = 0.02726$). Sibship reconstruction identified families of full-sib pairs within the 50 oysters sampled from each population. Fidalgo oysters comprised of 33 families (including "families" with a single individual), none of which included more than 8% of the total individuals. Dabob oysters had 21 full-sib families and Oyster Bay oysters had 17 families, the largest of

which contained 16% and 24% of individuals, respectively. Families for all populations were well mixed between spawning groups for production of F2 oysters.

	Fidalgo Bay	Dabob Bay	Oyster Bay
H_o	0.1957	0.2000	0.2151
H_e	0.2081	0.2064	0.2227
F_{is}	0.0593	0.0309	0.0341
ϕ	0.0155	0.0436	0.0372

Table 2.1: Genetic diversity and relatedness estimates for Puget Sound *O. lurida* populations. H_o, observed heterozygosity; H_e, expected heterozygosity; F_{is}, inbreeding coefficient; ϕ, kinship coefficient

2.5 Discussion

The results presented here provide evidence for a strong heritable component underlying phenotypic variation in growth and reproduction among three populations of Olympia oysters in Puget Sound, WA. By building on the Heare et al. 2017 study and showing consistent inter-population differences between first-generation (F1) and second-generation (F2), commonly reared oysters, this study is the first to demonstrate statistically significant phenotypic population structure extending through the F2 generation in *Ostrea lurida*, and one of the few to do so in a mollusc. Persistent population differences in growth rate and reproductive timing across several generations have also been documented for selectively-bred strains of the eastern oyster, *Crassostrea virginica* (Dittman et al., 1998; Barber et al., 1991), suggesting population divergence at these traits may be general across Ostreidae.

Reproductive activity was characterized for the adult F1 oysters, with Oyster Bay oysters producing the most larvae over the course of 7 weeks and the Fidalgo population demonstrating delayed reproductive activity, in accord with the reciprocal transplant study by Heare et al.. By showing consistent reproduction patterns in both variable natural environments and

controlled laboratory conditions, we demonstrated that reproductive timing in this species is not exclusively mediated by environmental conditions, but is also under genetic and/or epigenetic control. This result is in concordance with recent studies indicating that nearby populations of Olympia oysters vary in temperature thresholds for reproduction Barber et al. (2016); Seale and Zacherl (2009). The development of asynchronous reproduction on such small spatial scales has major implications for limiting gene flow and contributing to population divergence Palumbi (1994).

Fidalgo oysters exhibited the fastest growth during both the F2 larval and juvenile stages, while Dabob oysters exhibited the slowest growth, resulting in a significantly smaller size than Fidalgo oysters and smaller (although not significantly) than Oyster Bay oysters. This result is consistent with findings in F1 oysters by Heare et al., indicating a fixed underpinning for growth rate differences that manifests during both the larval and juvenile life stages. Interestingly, Fidalgo oysters also exhibited severely reduced and delayed reproductive activity. One explanation for this is an adaptive trade-off in energy allocation Perrin et al. (1993); Folkvord et al. (2014), where Fidalgo oysters are devoting more energy to shell growth and less to gonad development. Heare et al. also observed a potential adaptive trade-off in their reciprocal transplant study, where the population with slowest growth also had the highest survival across sites. Further investigation is required to fully resolve the link between growth, reproductive activity, and survival.

Given the characteristically high variance in reproductive success for *Ostrea* species Lallias et al. (2010); Launey and Hedgecock (2001), we estimated genetic diversity and relatedness for a subset of F1 oysters to ensure our approach of separating broodstock into smaller groups of 20-25 oysters produces offspring with many different families represented. Assuming the original wild population had kinship coefficients close to zero, our results do indicate that a reduced number of the original broodstock contributed to the F1 generation. However, by identifying between 17 and 33 different families within 50 sampled individuals, we are confident that the total pool of 94-101 F1 individuals per population represents upwards

of 40 wild genotypes and that our spawning method of mixing these families across smaller groups produced a genetically diverse F2 generation.

Environmental conditions in temperature, freshwater input, primary productivity, and anthropogenic waste effluent are known to vary among these sites Heare et al. (2017), supporting the possibility of selection driving the phenotypic variation observed in this study. In particular, up to 10 fold less chlorophyll a has been observed in Fidalgo Bay compared to the other sites. The proposed trade-off in energy allocation for this population could be driven by selection from such environmental differences in the quantity and timing of food supplyHeino and Kaitala (1999); Pontarp et al. (2015). As mentioned previously, selection could act through local adaptation or phenotype-environment mismatch, depending on the spatial scale of dispersal relative to the scale of environmental heterogeneity. Although dispersal information is not currently available for Olympia oyster larvae in Puget Sound, estimates using chemical fingerprinting in Southern California identified considerable larval exchange among estuaries separated by up to 75 km Carson (2010), which encompasses the distance between the populations used in this study. Further research using both chemical fingerprinting and genetic approaches are required to understand patterns of dispersal within Puget Sound in order to elucidate the mechanisms of adaptive population divergence.

Our results have implications for ongoing restoration efforts attempting to rebuild Olympia oyster populations. Current protocols for hatchery-based Olympia oyster restoration in Puget Sound involve using wild broodstock to produce hatchery-raised juveniles for outplanting in the same source population as the broodstock Blake and Bradbury (2012). These efforts have focused on multiple sites in Puget Sound, including those covered in this study. Our results support this practice, with a couple of caveats: 1) If local adaptation is indeed driving the observed phenotypic population structure, populations may be adapted to historical, rather than current, conditions. The convention of replenishing populations with only local broodstock sources may not provide the genetic diversity required to adapt in the face of rapid anthropogenic-induced environmental changes Jones (2013). 2) Hatchery

conditions vary dramatically from natural summer spawning conditions, and therefore artificial selective pressures are likely at play in the production of both the F1 and F2 oysters used in this studyMcClure et al. (2008). However, the observed differences in size among populations are corroborated by field observations of wild oysters (Brady Blake, WA Dept of Fish and Wildlife,7-10-2017, pers comm).

Despite the relatively close proximity of these populations and the potential for high gene flow, we observed significant phenotypic differences in fitness-related traits, even after multiple generations in the same environment. Of note, the population with the fastest growth rate also exhibited delayed and reduced reproductive activity, indicating a potential adaptive trade-off. By mitigating the influence of transgenerational plasticity, our results suggest these trait differences are heritable. Further research is now required to understand the mechanisms of inheritance underlying these observed differences, whether they be genetic or epigenetic.

CHAPTER 3

GENE EXPRESSION RESPONSES TO ACIDIFICATION IN TWO WIDESPREAD MARINE BIVALVES

3.1 Introduction

Ocean acidification (OA)—the reduction in ocean pH and associated changes in marine carbonate chemistry due to the uptake of anthropogenic CO_2, is a serious concern for many marine species that utilize $CaCO_3$ for building organic structures (Fabry et al., 2008). Physiological plasticity, local adaptation, or a combination of both can potentially facilitate resilience in species at risk from rapid environmental change (Bell and Gonzalez, 2009). One approach for understanding whether a species has the capacity to adapt to environmental change over time is to measure evolved differences along a current environmental gradient (i.e., "substituting space for time") (Blois et al., 2013). As the west coast of North America has long-standing natural variation in ocean pH, it is possible that some populations of widespread species can reveal adaptations to projected acidification conditions (Feely et al., 2008; De Wit and Palumbi, 2013).

Transcriptomics—the study of the RNA transcripts produced in specific environmental or developmental contexts, can provide insight into how an organism responds molecularly and physiologically to an environmental stress, particularly when external phenotypic changes are difficult to detect (DeBiasse and Kelly, 2016). For non-model species, transcriptomes often need to be assembled de novo and functionally annotated in order to link expressed transcripts with putative genes and gene products (Matz, 2018). High-throughput sequencing then provides estimates of transcript abundances, which can be used to identify differentially expressed genes (DEG) among factors or treatments of interest. Comparative transcriptomics across populations or different species can further characterize variation in molecular response to stress and provide insight into both plastic and evolutionary mechanisms underlying such responses (DeBiasse and Kelly, 2016). While multiple studies have

examined plastic responses to OA, fewer have compared these responses across populations and even fewer have evaluated interspecific variation (see Waldbusser et al. (2016) for an exception). Multi-species transcriptomic studies can facilitate identification of conserved responses to acidification stress.

The Olympia oyster (*Ostrea lurida*) and the purple-hinged rock scallop (*Crassadoma gigantea*) are native shellfish species with expansive, overlapping ranges on the Pacific coast of North America (Baker, 1995; Liao et al., 2018). Due to regional upwelling of deeper waters, some populations have already been exposed to pH as low as 7.4, whereas 7.82 is the projected mean for 2050 in the nearshore environment of the coast (Gruber et al., 2012). Considerable interest in developing sustainable aquaculture exists for both species (Baker, 1995; Olin et al., 2012), supporting the need to understand their adaptive potential in the face of environmental change. Studies of bivalve responses to OA have focused primarily on early life history stages, as disruption of calcification during early shell development has been shown to have severe developmental impacts and reductions in fitness (Waldbusser et al., 2015). Studies of adult bivalve response to ocean acidification are far fewer, as adults have been considered relatively robust to OA (Lemasson et al., 2017; Ross et al., 2011). Nevertheless, there is evidence for negative responses to OA in adults, such as reduced calcification, increased oxidative stress, decreased strength, altered immune systems, and increased mortality (Tomanek et al., 2011; Parker et al., 2013).

To understand inter- and intra-specific variation in response to reduced pH, this study compared gene expression responses to two pH treatments at two time points for adult Olympia oysters and rock scallops from three to four populations. Ctenidia and mantle were sampled from both species. Ctenidia are the respiratory organs of bivalves and are often sampled for examining metabolic rate and energy production (Timmins-Schiffman et al., 2014). The mantle is crucial to shell development in bivalves through secretion of calcium carbonate, which involves regulating intracellular pH and biomineralization (Hüning et al., 2013). As elevated pCO_2 and decreased pH negatively impact the availability of carbonate

ions for calcifying organisms, I hypothesized that genes expressed differentially in the mantle in response to low pH will functionally be involved in maintaining intracellular homeostasis and regulating biomineralization (Clark et al., 2013).

Olympia oysters were collected from three sites with distinct pH regimes: San Francisco Bay, CA, Coos Bay, OR, and Ladysmith Harbor, BC. My prior work has shown that each of these populations belongs to a genetically distinct phylogeographic region, delineated by areas of reduced gene flow (Silliman 2019). Oysters from OR experience the lowest average summer pH (7.77), followed by CA (7.81) and BC (7.94) (Chan et al., 2017). Based on this environmental variation, I hypothesized that oysters from BC would demonstrate increased physiological stress in response to low pH. The population in Coos Bay, OR was an accidental introduction from Willapa Bay, WA in the early 1900s and did not reach its current abundance until the 1980's (Baker et al., 1999). Examining the transcriptomics of a recently established population may provide insight into the physiological consequences of an anthropogenic range expansion.

The purple-hinged rock scallop (*Crassadoma gigantea*) is a marine bivalve distributed from Baja California, Mexico to Alaska. Little is currently known about the population structure of this species, and the only published genomic resources are a mitochondrial genome and a description of a non-publically available transcriptome assembly (Cao et al., 2018; Liao et al., 2018). *C. gigantea* is found subtidally down to 80m (Dijkstra and Dijkstra, 2010), and so has likely been exposed to more consistently lower pH and Ω_{arag} than the primarily intertidal *O. lurida* (Gruber et al., 2012). For this study, scallops were collected from Monterey Bay, CA, the outer coast of WA (Sekiu), Puget Sound, WA (Dabob Bay) and Seward, AK. While these sites were primarily chosen due to their importance to the developing scallop aquaculture industry, I anticipate that they also represent genetically distinct populations. An ongoing RADSeq study by collaborators at the University of Washington is testing for neutral and adaptive population structure of *C. gigantea* across these same sites plus four others (N. Lowell, pers. comm.).

For both species, I investigated four core questions: 1) What genes exhibit a consistent intraspecific change in expression in response to low pH? These genes may constitute a conserved plastic response to acidification stress (Maynard et al., 2018). 2) What are DEGs among populations, irrespective of treatment? Such constitutive differences among populations have been cited as a potential mechanism for evolved dissimilar environmental tolerances among populations (DeBiasse and Kelly, 2016; Dixon et al., 2015). 3) What genes respond to pH in a population-specific manner, and therefore exhibit genetic variation in plasticity? 4) Do patterns of gene expression differ after 36 hours of exposure compared to seven weeks? By evaluating changes in gene expression over time, we can identify differences in response to acute and prolonged stress. Many studies examining stress response at multiple time points have found an extreme transcriptomic response initially, followed by a return to gene expression baseline levels (e.g., (Franssen et al., 2011)). Following transcriptomics within each species, I compared gene expression patterns between *O. lurida* and *C. gigantea*. Specifically, I identified conserved functional groups present in both transcriptomes that exhibited similar expression patterns in response to low pH across species, and therefore had a deeply conserved functional response to acidification stress.

Evaluating the transcriptomic response of an organism to a stressor does not necessarily indicate whether the plastic response is adaptive (Grether, 2005). DEGs could underlie either an adaptive response or signal greater levels of stress. Studies comparing transcriptomic responses among locally adapted populations can help with this challenge, but are nonetheless correlative as transcriptomic data alone cannot inform on the actual fitness effects from expression of a particular transcript. The gene sets identified here provide excellent candidates for explicitly testing the connection between gene expression and phenotype, for example with directed qPCR (Heare et al., 2018) and fitness assays or CRISPR interference (Qi et al., 2013).

3.2 Materials and Methods

3.2.1 Sampling and common garden experiment

Oyster collection and husbandry

Approximately 120 adult oysters (at least 2cm in length) were collected from each of three sites in December 2015: San Francisco Bay, California (CA) (GPS), Coos Bay, Oregon (OR) (GPS), and Ladysmith Harbour, British Columbia (BC) (GPS). Both CA and OR are located in the central region of the California Current System, where upwelling of low pH and Ω_{arag} waters naturally occur and can even reach the surface in a mosaic of nearshore locations. Based on the SRES high-emissions scenario A2, nearshore surface pH along the central California Current System in 2050 is predicted to fall within 7.7±0.38 (Gruber et al., 2012; Busch and McElhany, 2016). Mean aragonite saturation (Ω_{arag}), representing the availability of $CaCO_3$ in seawater, is also predicted to fall from an annual mean of 1.67±0.16 in 2005 to 1.26±0.12 in 2050, with some bays reaching undersaturation ($\Omega_{\text{arag}} < 1.0$) in the summer.

Oysters were collected by hand and shipped overnight wrapped in wet newspaper to a quarantine facility at the K. K. Chew Center for Shellfish Restoration and Research in Manchester, WA. For five months, the three populations were maintained in mesh bags in a flow through tank at conditions similar to the adjacent Clam Bay in southwest Puget Sound and fed a diet of mixed live algae culture. In May 2016, the population groups were split up into 5 replicate buckets each and conditioned to spawn in common conditions of water temperature and food availability for a separate experiment. In September 2016, the oysters were transferred to the experimental system, with 16-20 oysters from each population suspended in a mesh bag into each tank replicate.

Scallop collection and conditioning

Approximately 26-45 adult rock scallops were collected from each of four sites: Annette Island, AK (AK) (GPS), Sekiu, WA on the Strait of Juan de Fuca (WS), Dabob Bay, WA in Puget Sound (WD), and Monterey Bay, CA (CS). Tags were attached to scallops with epoxy and scallops were maintained in either separate or mixed population tanks, provided with algae and flow through water from the same source. Small mantle tissue samples were taken nonlethally from all scallops for a separate population genetic study. Scallops from CS initially exhibited significantly higher mortality than the other populations upon initially arriving at Manchester. In September 2016, scallops were transferred to the experimental system, with 1-2 scallops per population in each tank replicate.

Acidification experiment

Two pH treatments were chosen in order to evaluate transcriptomic responses at both a relevant "current" condition and a realistic future stressor (Busch and McElhany, 2016). One treatment level reflected ambient water at the research station ("ambient", pH: 7.8, pCO_2: 1000, Ω_{arag}: 1.5). This pH value was 0.01-0.1 higher than the lowest 5th percentile of 2011-2013 observations for any oyster collection site (based on nearest site proxies, (Chan et al., 2017)). The other treatment ("low pH", pH: 7.4, pCO_2: 2500, Ω_{arag}: 0.7) reflects potential conditions for 2050 during periods of strong upwelling. Coos Bay, OR was the only site where 7.4 was observed in 2011-2013 (Chan et al., 2017).

Oysters and scallops were maintained for one month in the experimental system with ambient water for all tank replicates, after which the pH and Ω_{arag} were lowered for half of the tank replicates. Ctenidia and mantle tissues were dissected destructively from both oysters and scallops 36 hours after the pH was lowered. Oyster tissues were transferred immediately to 2 mL tubes in liquid nitrogen, while scallop tissues were placed in tubes with RNALater. All samples were transferred to -80°C for long term storage. Ctenidia and mantle tissues were sampled from the remaining scallop individuals 6.5 weeks after the pH

was lowered in the low pH treatment, and after 7 weeks for oysters. Due to low numbers, scallops from the AK population were only sampled at 36 hours.

The flow-through experimental system involved one header tank for each treatment delivering water to four separate open air tanks (8 total), each constituting "psuedo"-tank replicates for the two treatments (Cornwall and Hurd, 2016). pCO_2 in the header tanks was controlled by two Mobile Ocean Acidification Treatment Systems (MOATS), each of which is equipped with independent computer controllers running custom Labview software (National Instruments) to monitor pH in real time using Durafet probes (Honeywell), and constantly adjusted by direct bubbling a controlled combination of air, CO_2-free air, and pure CO_2. Water temperature was not adjusted from that of the source water, and therefore fell from 15.5°C to 12.5°C throughout the course of the experiment. Reeds algal food paste mixed with UV-treated, ambient seawater was dispensed by two separate pumps directly into the line prior to delivering water to the tanks. Pumps were regularly calibrated to ensure equal distribution of algae. Discrete water samples for pH, salinity, temperature, and total alkalinity (TA) were collected for 3 days before the start of the experiment, every day of the first week, once a week for the remainder of the experiment, and then the two days prior to the last tissue sampling day of the experiment. Temperature and pH were monitored continuously in the header tanks and in two of the treatment tanks using Durafet and X probes. DIC, pCO_2, CO_3, and Ω_{arag} values were calculated using the R package *seacarb* (Gattuso et al., 2015) and custom R scripts by Hollie Putnman.

3.2.2 Transcriptome assembly

Sample processing

Total RNA was extracted using the PureLink RNA Mini Kit with Trizol for tissue lysis and on-column DNase treatment (ThermoFisher Scientific). Extracted RNA was quantified with a Qubit v3 kit and evaluated for quality on a Bioanalyzer. 500 ng of total RNA per sample

was used for 3' mRNA sequencing with the QuantSeq FWD kit (Lexogen) (Moll et al., 2014). Sequencing library quality was assessed using Bioanalyzer HS chip, and 237 libraries (111 oysters, 126 scallops) were multiplexed on five 50bp single-end Illumina HiSeq 4000 lanes. 3' RNAseq, also known as TagSeq, only sequences a single fragment for each transcript, allowing for shorter sequencing reads and lower sequencing depth that traditional RNAseq. By only sequencing a single fragment, this method also eliminates the need for length bias correction in estimating transcript abundances, and has been shown to provide more accurate estimates of transcript abundances than standard RNAseq (Lohman et al., 2016). QuantSeq reads must be mapped to an independently sequenced and assembled transcriptome or genome, so directional RNAseq libraries for ten individual *O. lurida* samples and nine *C. gigantea* sample pools were prepared and sequenced by the UChicago Functional Genomics Core.

Assembling and annotating transcriptomes

Trinity v2.8.4 was used to assemble three *O. lurida* transcriptomes separately by population (Grabherr et al., 2011). TransDecoder (`https://github.com/TransDecoder/TransDecoder`) was used to extract ORF predictions from transcripts for each of the three assemblies and translate them into peptide sequences, one sequence per transcript. OrthoFinder (Emms and Kelly, 2015) was used to identify orthologous gene sets between populations ("orthogroups"), based on a similar strategy used for assembling transcriptomes for multiple populations of copepods in (DeBiasse et al., 2018). Transcripts assigned to orthogroups with at least one transcript from each population were combined into a single assembly, where the orthogroup designation is considered a "gene" made up of multiple isoforms. For *C. gigantea*, first individual pooled samples were assembled using DRAP, which utilizes Oases, cd-hit-est, and stringent filtering by read support to reduce a transcriptome down to a single isoform per gene (Cabau et al., 2017). These 9 separate assemblies were then combined using metaDRAP. For this transcriptome, each transcript was considered a single "gene". Transcriptomes were annotated using Trinotate (`https://github.com/Trinotate/Trinotate.github.io`

and eggnog_mapper to derive gene ontology (GO), KOG, and KEGG functional predictions (Huerta-Cepas et al., 2017).

3.2.3 Differential gene expression and functional enrichment

Differential gene expression with DeSeq2

Filtered QuantSeq reads for each species were mapped to their respective transcriptome assembly using Bowtie2. The number of sequence reads mapping to each isogroup was counted using Perl scripts by Misha Matz (`https://github.com/z0on/tag-based_RNAseq`), where reads mapping equally well to multiple isogroups in *O. lurida* were discarded. Isogroup expression is referred to here as gene expression, although it is possible that an isogroup contains multiple genes through chimerism or that a single gene is represented by multiple isogroups. Tests of differential gene expression in response to experimental factors was conducted separately for each species with the R package DeSeq2, which fits generalized linear models (GLMs) to raw read counts to estimate the overall expression strength of each gene and the log fold change (LFC) in expression in response to specified contrasts. Wald tests or likelihood ratio tests (LRT) (depending on the factor of interest) with Benjamini-Hochberg procedure are used to adjust for multiple testing (Love 2014). Significant differentially expressed genes were defined as those with an adjusted p-value (false discovery rate, FDR) $<$ 0.05. LFC values and adjusted and unadjusted p-values were retained from each contrast for gene ontology analysis. To control for batch effects from technical variation in sequencing and library preparation, I used the R package *sva*, which takes normalized count data and estimates variation that is not accounted for in a provided model design. Two continuous factors were used to represent hidden batch effects, and were included in all subsequent GLMs.

For each species, three different GLMs were fit to the count data in order to identify sets of differentially expressed genes (DEGs) of interest. 1) Genes that increased or decreased

consistently across populations in response to low pH were identified by controlling for variation due to population with model design = $\sim$ *population* + *treatment*, and model factors "population" (levels = BC, CA, and OR in oysters; AS, CS, WD, and WS in scallops) and "treatment" (levels=pH 7.8 and 7.4) DEGs of interest were extracted by contrasting the LFC at pH 7.4 compared to pH 7.8. 2) Genes with constitutive differences in expression among populations were identified in two ways. To identify genes with a significant effect of population on expression, irregardless of treatment, a LRT compared two models: $\sim$ population + treatment, and $\sim$ treatment. To identify DEGs between pairs of populations at the ambient pH, the "treatment" and "population" factors were combined into a single factor ("group") representing all combinations (e.g., for oysters, "group" levels = BClowpH, BCambient, CAlowpH, CAambient, ORlowpH, ORambient). The design $\sim$ group, is similar to modeling an interaction term between population and treatment, but allows for simpler extraction of desired gene sets. 3) Genes that responded to low pH in a population-specific matter (i.e. potentially involved in local adaptation) were identified with a LRT comparing $\sim$ population + treatment + population:treatment, and $\sim$ population + treatment. The $\sim$ group model was also used to contrast genes that were differentially expressed among populations at the low pH treatment. All models were run on tissue types separately. To compare differences among time points, all models were run three separate times: with samples combined across time points (and including "time" as a model factor, levels= 36hours, 7weeks), just the 36hours samples, and just the 7weeks samples.

To characterize broad patterns of variation between tissues, treatments, populations, and time points, a principal component analysis (PCA) of variance-stabilized transformed count data was performed on the 500 genes with the highest variance across samples using the plotPCA function in DeSeq2. To visualize expression patterns of DEG sets, heatmaps were created with variance-stabilized transformed counts and the pheatmap R package.

Gene ontology analysis

For gene ontology analysis of most DEG sets, I used GO_MWU in R (Wright et al., 2015), which ranks GO categories within a dataset based on relative gene expression in order to identify GO terms that are enriched in either up- or down-regulated genes, relative to the rest of the dataset. Enrichment significance is calculated using a two-sided Mann-Whitney U-test followed by an FDR correction. Gene expression was represented by the negative logarithm of a gene's unadjusted p-value from DeSeq2, then multiplied by -1 if the gene was down-regulated. GO terms are first clustered within the dataset, and then groups of GO categories are merged if the most dissimilar of terms in the group share $> 75\%$ of genes in the dataset. GO terms found in fewer than 3 genes or $> 15\%$ of genes were excluded. The GO terms passing an FDR of at least 10% are plotted as a dendogram based on the level of gene sharing between terms, then colored by the direction of gene expression change. For DeSeq2 analyses involving a LRT, the more traditional Fisher Exact test was performed using GO_MWU to test for GO enrichment in genes with an FDR of 0.10 from DeSeq2 relative to the rest of the dataset, following the same GO filtering and clustering thresholds. Dendograms from Fisher exact tests do not indicate direction of gene expression change, but significance of enrichment after FDR correction.

Eukaryotic Orthologous Groups (KOGs) and the R package KOGMWU were used to compare functional enrichment patterns between datasets of different tissue types or between the two species(Matz, 2019). KOGMWU works similarly to GO_MWU, but only operates over 23 KOG classes and facilitates a meta-analysis based on correlation of KOG delta-ranks.

Identifying candidate adaptive markers

The 129 putative adaptive GBS loci identified in Silliman (2019) were compared against the assembled *O. lurida* transcriptome using blastn. Transcripts that had been identified as a gene of interest using DeSeq2 and had a mapped adaptive GBS loci were characterized as candidate adaptive markers.

3.3 Results

3.3.1 Acidification experiment

All measured and estimated carbonate chemistry parameters (e.g., pH, total alkalinity, DIC, pCO_2, Ω_{arag}) varied significantly between treatments, while temperature and salinity did not (Fig. 3.1, Table 3.1).

In the 3 months prior to the start of the experiment, mortality across oyster populations was fairly even, with CA experiencing the most with 14.4% mortality and BC the least with 9.4%. During the experiment, CA experienced the highest mortality in both treatments (ambient: 6, low pH: 3), with BC and OR having the same mortality in both treatments (ambient: 4, low pH: 2). In the middle of the experiment, BC oysters from one of the tank replicates were accidentally left out of water for two days and experienced considerable mortality, which is not reflected in experimental mortality estimates. No scallop mortality was observed during the course of the experiment.

Water chemistry	Ambient pH	Low pH
Salinity*	29.24 (0.034)	29.22 (0.035)
Temperature*	13.52 (0.115)	13.51 (0.116)
pH*	7.774 (0.005)	7.427 (0.005)
CO2	42.46 (0.605)	102.99 (1.36)
pCO2	1052.59 (15.73)	2550.22 (33.41)
HCO3	2635.07 (10.25)	2874.52 (8.00)
CO3	95.69 (0.874)	46.95 (0.532)
DIC	2773.22 (10.30)	3024.46 (9.76)
TA*	2859.57 (9.18)	2984.03 (8.58)
Ω_{arag}	1.491 (0.014)	0.732 (0.008)

Table 3.1: Mean and SE of discrete water sampling during the experiment for the two treatments. * indicates parameters that were measured directly, others were estimated using the R package seacarb.

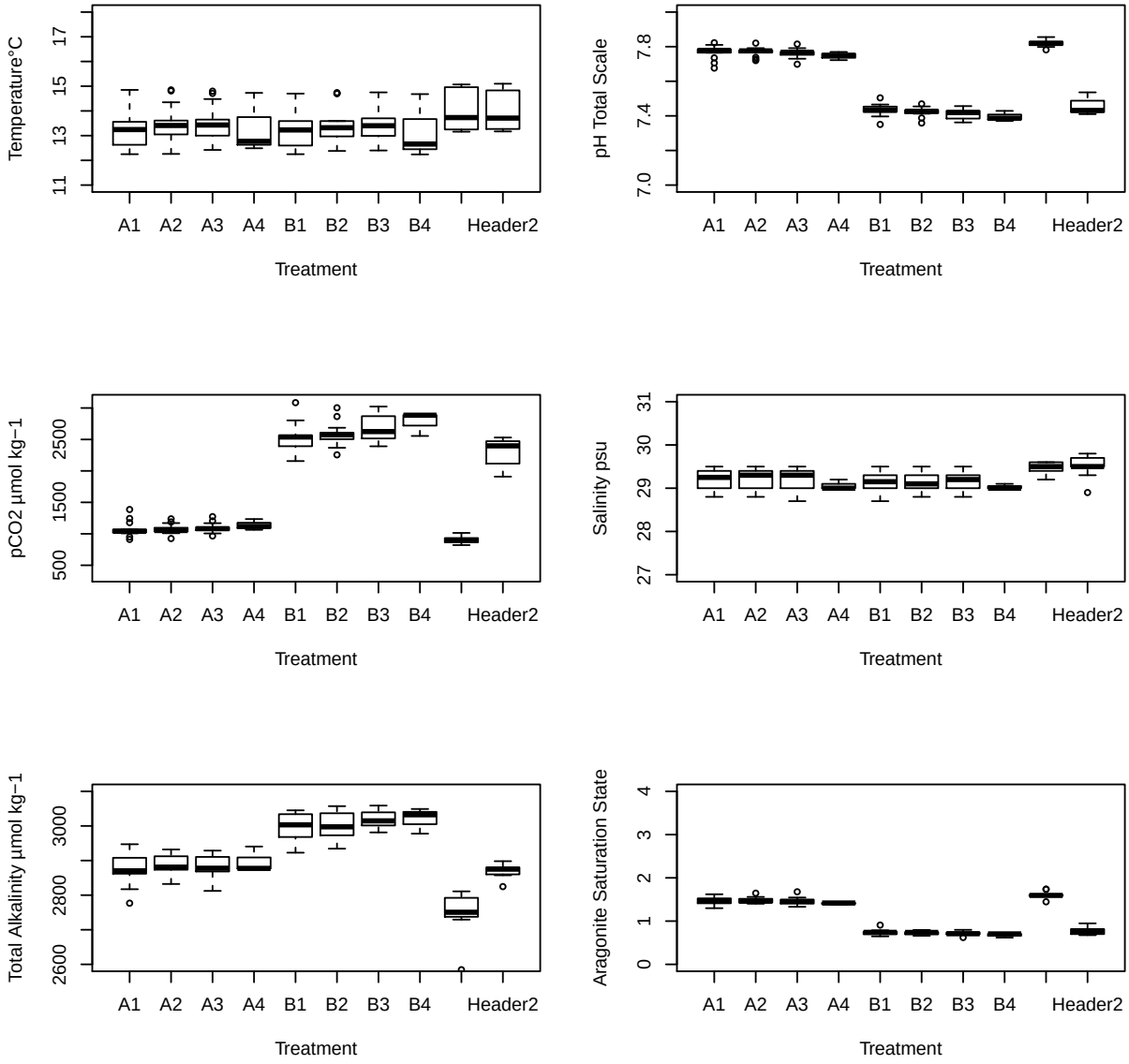


Figure 3.1: Water chemistry observations for each tank replicate after the pH was lowered in the low pH treatment. Header 1 is the mixing header tank for the ambient treatment, Header 2 is the header tank for the low pH treatment.

3.3.2 Differential gene expression in Olympia oysters

Transcriptome assembly, annotation, and QuantSeq mapping

Directional RNASeq of 10 *O. lurida* samples resulted in 1,295M total raw reads. After quality filtering, trimming, and removal of rRNA and overrepresented sequences, 1,210M

total reads remained. Non-orthologous sequences among populations were identified and removed, resulting in a combined transcriptome of 29,810 total orthogroups ("isogroups") with 182,046 transcripts (Table 3.2). This combined Ortho transcriptome was annotated with Trinotate, resulting in annotations for 161,080 transcripts (Table 3.3).

QuantSeq 3' mRNA sequencing of 110 samples resulted in mean 4,215,354 filtered reads per sample. 4 individuals were removed due to low sequencing depth. An average of 76.8% reads per sample mapped successfully to the Ortho transcriptome, which is within the range of $70\% - 90\%$ expected for high-quality transcriptomes (Conesa et al., 2016). An average of 74.2% reads mapped to multiple transcripts.

Assembly	# of individuals	# of contigs (of genes)	BUSCO
BC	2	240,026 (122,521)	99.0%
OR	2	249,664 (132,485)	99.1%
CA	6	335,877 (162,419)	99.1%
OrthoTrans	10	182,046 (29,810)	98.4%
MetaDRAP (*C. gigantea*)	8 pools	52,105	97.2%

Table 3.2: Transcriptome assemblies of *O. lurida* populations separately, *O. lurida* populations combined using OrthoFinder, and an assembly of *C. gigantea* using MetaDRAP.

Assembly	OrthoTran	MetaDRAP
SwissProt_blastx	140,737	17,804
SwissProt_blastp	124,997	14,705
KEGG	107,893	10,710
eggnog	103,811	15,650
PFAM	118,755	15,468
GO	126,605	17,258
signalP	16,618	2,879
rRNA	38	5
TmHMM	45,741	4,698
At least one	161,080	23,766
total contigs	182,046	52,105

Table 3.3: Annotation results for *O. lurida* and *C. gigantea* transcriptomes.

Overall patterns of gene expression

PCA of all oyster samples illustrated that tissue type was a major source of variation, so accordingly differential gene expression was analyzed for ctenidia and mantle tissues separately (Fig. 3.2). A distance matrix of all samples identified 6 outlier samples, which were removed from subsequent analyses. PCAs of ctenidia samples and mantle samples both showed some separation of CA samples from BC and OR along PC2. PCA of mantle samples had slight clustering of samples by Time Point along PC 1 (not shown). Treatment did not appear to be an obvious source of variation on PCs 1 and 2 for ctenidia or mantle.

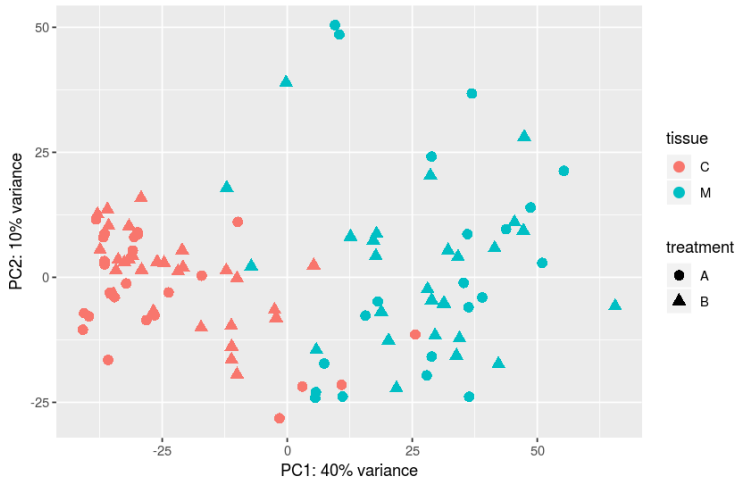


Figure 3.2: PCA of gene expression in the 500 genes with the highest variance across all *O. lurida* samples. Color represents tissue and shape is pH treatment.

Conserved response to low pH

Only 15 and 31 genes showed a consistent response across populations to pH treatment in ctenidia at Time 1 and 2, respectively, while 0 and 13 showed a consistent response at Time

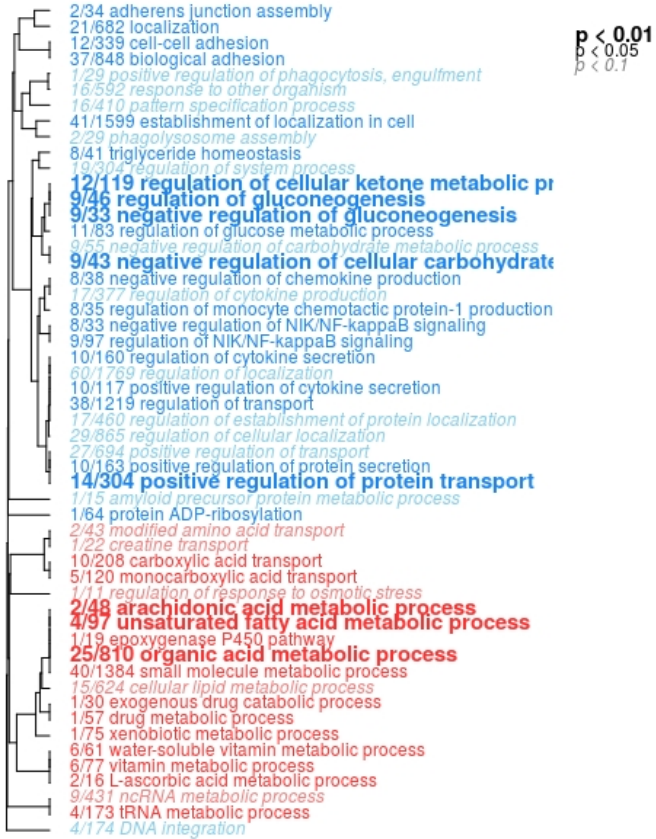
1 and Time 2 in mantle tissue. When Time 1 and 2 samples were analyzed together, 31 genes showed a consistent response in ctenidia and 1 gene showed a consistent response in the mantle (Table 3.4). Functional enrichment analysis for conserved response to acidification of all ctenidia genes using GO_MWU identified 53 significantly enriched GO terms. Upregulated genes are involved in environmental stress response (e.g., regulation of response to osmotic stress, carboxylic acid transport), organic acid metabolism (e.g., arachidonic acid metabolism, L-ascorbic acid metabolism), and tRNA metabolism. Downregulated genes are involved in negative regulation of metabolic pathways (e.g., gluconeogenesis), positive regulation of protein transport, cytokine secretion, and regulation of signaling pathways (e.g., NIK/NF-kappaB signaling) (Fig. 3.3a). In the mantle at 7 weeks, 15 GO terms were enriched, with downregulated genes including chromosome organization, RNA meteabolism, and cytoskeleton. Upregulated genes included regulation of peptidase and a toll-like receptor (Fig. 3.3b).

	Ctenidia 36 hours	Ctenidia 7 weeks	Ctenidia Combined	Mantle 36 hours	Mantle 7 weeks	Mantle combined
Number of samples	29	24	53	23	21	44.
Number of genes	22,196	20,232	21,834	20,191	19,523	20,414
Response to low pH	15	31	31	0	13	1
Population effect	821	287	1,035	822.	325	1,613
Population:pH interaction	4,406	86	3,186.	86	205	68
BC response to low pH	59	30	110	2	384	30
OR response to low pH	4,292	17	716	3	4	9
CA response to low pH	25	25	39	26	5	25

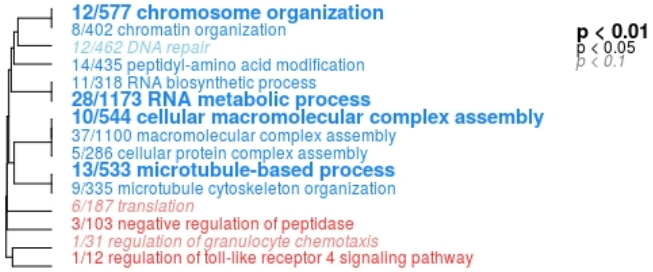
Table 3.4: Number of significantly differentially expressed genes for each tissue and time point, as determined by DeSeq2 (FDR < 0.1). Combined: samples are pooled across both time points, with "time" included as a factor in GLMs.

Constitutive differences in gene expression among populations

Numerous genes showed constitutive differences in gene expression among populations in both ctenidia and mantle, when controlling for time point and treatment. In the ctenidia,



(a) Ctenidia



(b) Mantle, 7 weeks

Figure 3.3: GO_MWU results for *O. lurida* genes with a conserved response to low pH across populations. a) ctenidia, b) mantle, 7 weeks. Blue are downregulated in response to treatment, red are upregulated. The font refers to significance of the FDR test for enrichment; the dendrograms depict the sharing of genes between categories; the fractions correspond to genes with an unadjusted $p < 0.05$, relative to the total number of genes within the category.

1,035 genes showed a significant population effect when both time points were analyzed together. 36 functional GO terms were enriched in these genes based on Fisher's exact test, with the majority involved in viral immune response (e.g., detection of virus, regulation of interferon-alpha production, response to cytokine stimulus), and the rest involved in reproduction (e.g., mating behavior, vulval location) or chromatin silencing (Fig. 3.5a).

Comparing OR and CA samples at the ambient treatment revealed that CA samples differed primarily at viral immune response genes, suggesting that this population had a viral infection not present in the others. Other genes that were upregulated in CA compared to OR included chromatin organization, male gamete generation, and cellular response to stress. OR had 814 and 606 DEGs at the ambient treatment compared to CA and BC, respectively, while only 110 genes were differentially expressed between CA and BC at ambient. A suite of genes were upregulated in OR at the ambient treatment compared to the other two populations, including some associated with the previously identified conserved intraspecific response to low pH (e.g., gluconeogenesis, organic acid metabolic process). Other uniquely upregulated genes in OR at the ambient treatment included cellular respiration, translation, and the electron transport chain (ETC) (Fig. 3.4).

In the mantle, 1,613 genes showed a significant population effect when both time points were analyzed together, indicating that they had a significant change in expression depending on source population, irrespective of pH treatment. 43 functional GO terms were enriched in these genes, most of which overlapped with those in the population-specific ctenidia genes and also involved immune response. Other enriched GO terms included regulation of inclusion body assembly, regulation of the Wnt signaling pathway, and detection of mechanical stimulus (Fig. 3.5b). Unlike in the ctenidia at 36 hours, the population with the most divergent mantle gene expression at 7 weeks was BC, with 679 DEGs than OR and 374 DEGs than CA. OR and CA only had 67 DEGs at ambient. Unfortunately, only two mantle BC samples were successfully sequenced for the ambient treatment at 7 weeks, which may be partially driving these results.

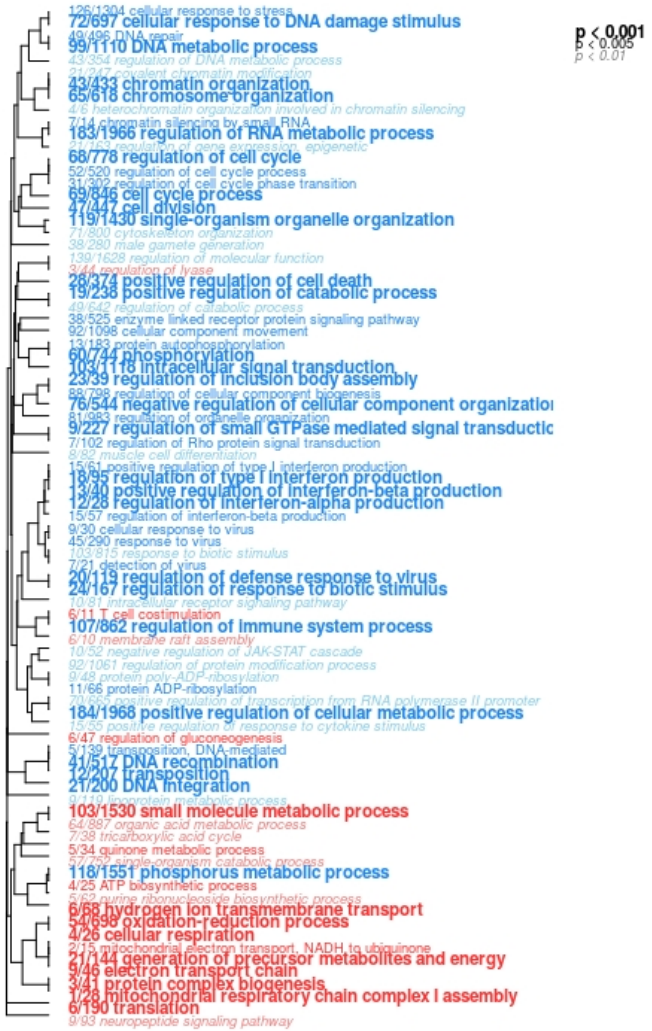
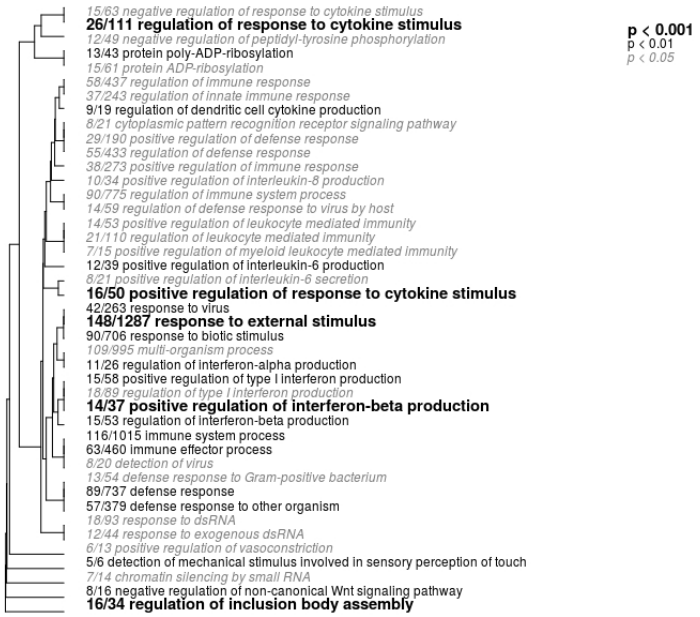


Figure 3.4: GO_MWU results for DEGs ctenidia genes between OR and CA *O. lurida* populations at 36 hours in the ambient treatment. Blue are downregulated in OR compared to CA, red are upregulated in OR. The font refers to significance of the FDR test for enrichment; the dendrograms depict the sharing of genes between categories; the fractions correspond to genes with an unadjusted $p < 0.05$, relative to the total number of genes within the category.

(a) Ctenidia



(b) Mantle

Figure 3.5: GO_MWU results for *O. lurida* genes with a significant change based on population (when controlling for treatment) in a) ctenidia samples and b) mantle samples. Font refers to significance of the FDR test for enrichment; fractions correspond to genes with an adjusted $p < 0.05$, relative to the total number of genes within the category.

Population-specific responses to low pH

The ctenidia and mantle tissue exhibited contrasting patterns of population-specific expression through time. At 36 hours, ctenidia had 4,406 genes with a population-specific response to low pH and the mantle had 28 genes. These ctenidia genes were functionally enriched for intracellular transport, cellular protein assembly, Golgi organization, and ATP production, while the population-specific mantle genes showed no functional enrichment (Fig. 3.6). When comparing individual responses to low pH within populations, OR had the highest number of DEGs in ctenidia at 36 hours (4,295), followed by BC (59), and CA (25). Three genes were shared between all populations, one of which was functionally annotated as "EGF-like domain", which is associated with neuron development. Another was predicted as a signal peptide, while the third was unannotated. OR had numerous uniquely upregulated gene functions including regulation of development (particularly muscle), locomotion (e.g., microtubule-based processes, phosphorylation, regulation of ion homeostasis, and neuron projection guidance). Uniquely downregulated genes in OR ctenidia included mRNA metabolism, ATP biosynthesis, and the ETC(Fig. 3.7).

At 7 weeks, the ctenidia had 86 genes with a population-specific response to the treatment, with DEGs ranging from 67 (OR) to 2 (BC), and no genes shared among populations. Conversely, the mantle had 205 DEGs, with 917 differentially expressed in BC, followed by 15 in CA and 4 in OR. There was no functional enrichment in the ctenidia genes at the later time point, and only "regulation of acute inflammatory response" was enriched for these mantle genes at 7 weeks. When comparing individual responses to low pH within populations, BC had the highest number of DEGs in mantle at 7 weeks (384), followed by 5 in CA and 3 in OR. No genes were shared between all populations.

KOG class enrichment was compared among 12 DEG sets representing each combination of population, tissue, and time point, and corresponded well with GO enrichment analysis comparisons among populations. After FDR correction, five KOG classes showed the most pronounced changes in expression and exhibited fairly distinct regulation patterns de-

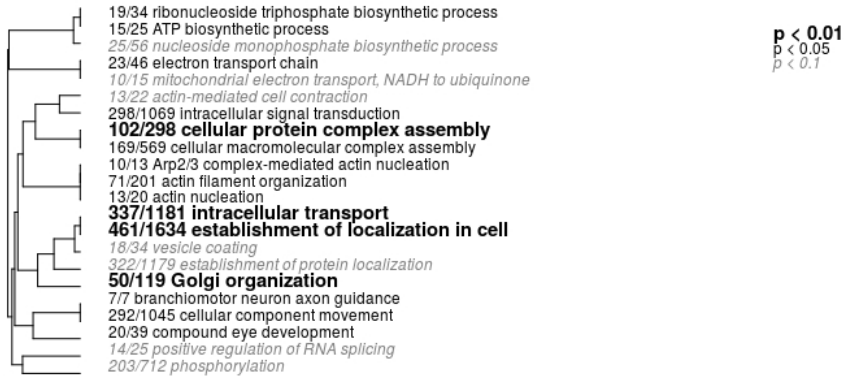


Figure 3.6: GO_MWU results for ctenidia DEGs with a population-specific response to low pH at 36 hours.

pending on population and time point. For CA and BC ctenidia at both time points and BC mantle tissue at 7 weeks, 'Energy production and conversion' and 'Translation, ribosomal structure and biogenesis' were significantly upregulated while 'Cytoskeleton', 'Signal transduction mechanisms', and 'Cell wall/membrane/envelope biogenesis' were largely downregulated. The opposite expression pattern was observed for all OR datasets and CA mantle at 7 weeks. BC and CA mantle tissues at 36 hours only 2-26 DEGs in response to treatment, and so their KOG enrichment results should be evaluated conservatively (Fig. 3.8).

Figure 3.7: GO_MWU results for ctenidia DEGs in OR at 36 hours. Blue are downregulated in response to treatment, red are upregulated.

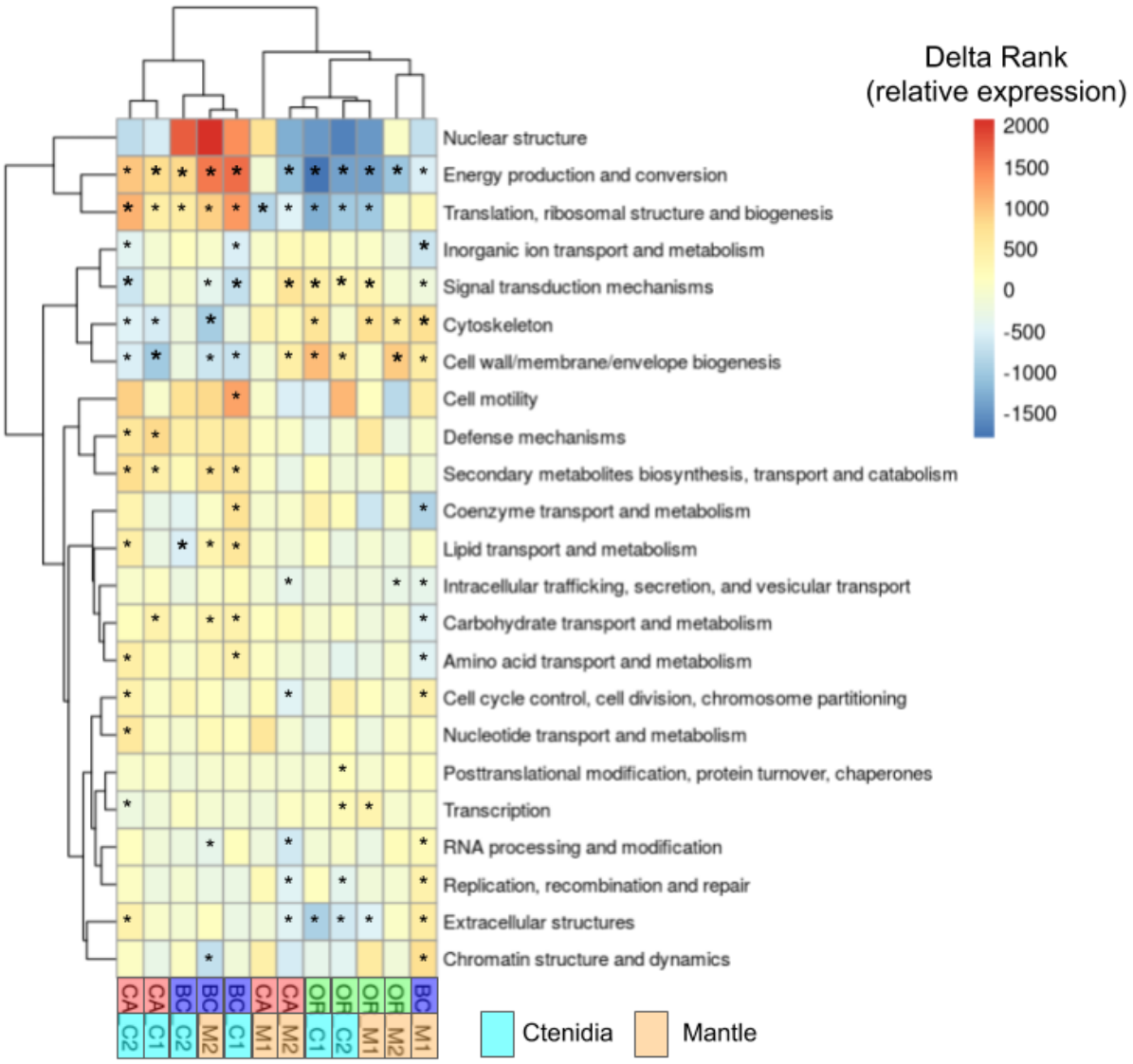


Figure 3.8: KOG_MWU comparison for *O. lurida* genes that are differentially expressed in response to low pH, for all combinations of tissues, time points, and populations. Red indicates upregulated, blue is downregulated; * indicate significant enrichment after FDR correction, ***** represents the top upregulated and top downregulated classes for each dataset. Datasets are clustered by their correlation.

3.3.3 Differential gene expression in rock scallops

Transcriptome assembly, annotation, and QuantSeq mapping

The 8 directional RNASeq pools were assembled separately with DRAP, and then merged using metaDRAP (Table 3.2) to produced 52,105 contigs, after filtering for FPKM of 1. This metaDRAP transcriptome was annotated with Trinotate, resulting in annotations for 23,766 transcripts (Table 3.3).

QuantSeq 3' mRNA sequencing of 126 samples resulted in an average of 4,716,547 filtered reads per sample. Three individuals were removed due to low sequencing coverage. An average of 84.5% reads per sample mapped successfully to the metaDRAP transcriptome. 35.7% reads mapped to multiple transcripts, all of were discarded when generating a read count table.

Overall patterns of gene expression

PCA of all scallop samples illustrated that tissue type was a major source of variation, so differential gene expression was analyzed for ctenidia and mantle tissues separately (Fig. 3.9). A distance matrix of all samples identified 8 outlier samples, which were removed from subsequent analyses. PCA of ctenidia samples showed no discernible clustering by population or treatment, with PCs 1 and 2 only capturing 8% and 5% of the total variation. PCA of mantle samples showed some structuring by population, with PC1 capturing variation due to sampling time point. Treatment did not appear to be an obvious source of variation on PCs 1 and 2 for ctenidia or mantle (Fig. 3.10).

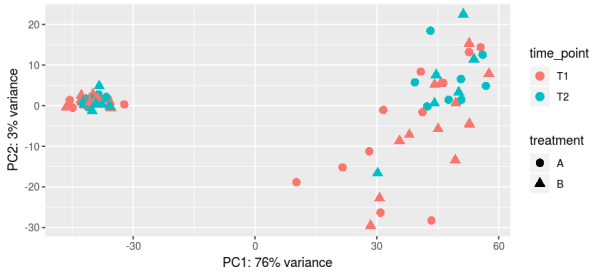


Figure 3.9: PCA of gene expression for all *C. gigantea* samples, with ctenidia samples clustering on the left and mantle samples on the right.

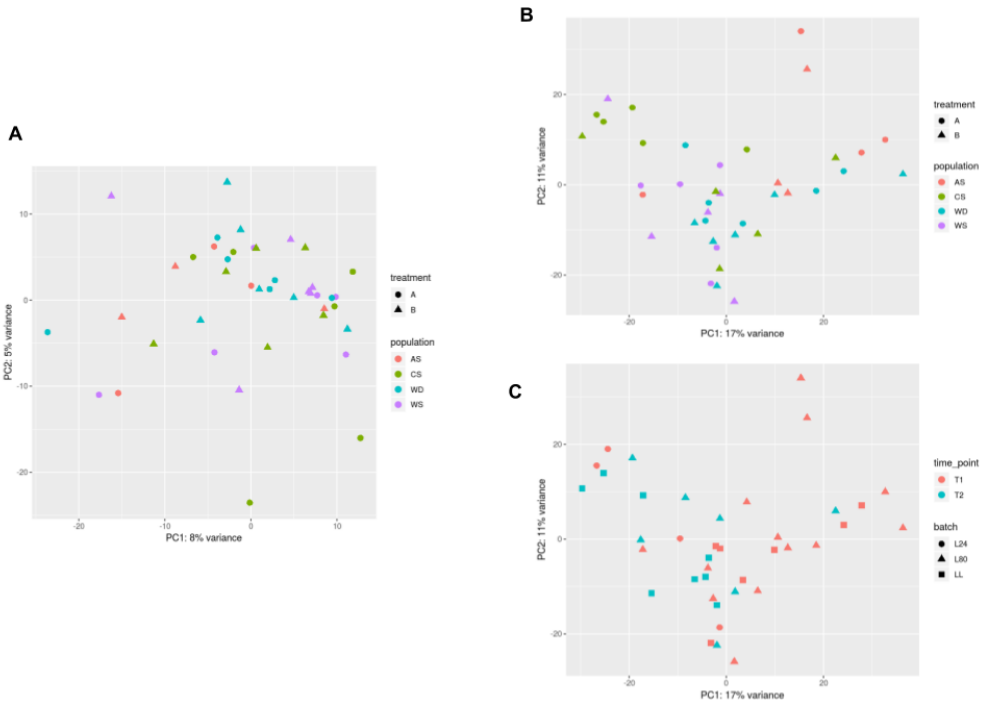


Figure 3.10: Tissue-specific PCA of *C. gigantea* samples. A: PCA of ctenidia samples, with colors for population and shape for treatment. B: PCA of mantle samples, with colors for population and shape for treatment. C: PCA of mantle samples, with color for sampling time point and shape for sequencing batch.

Conserved response to low pH

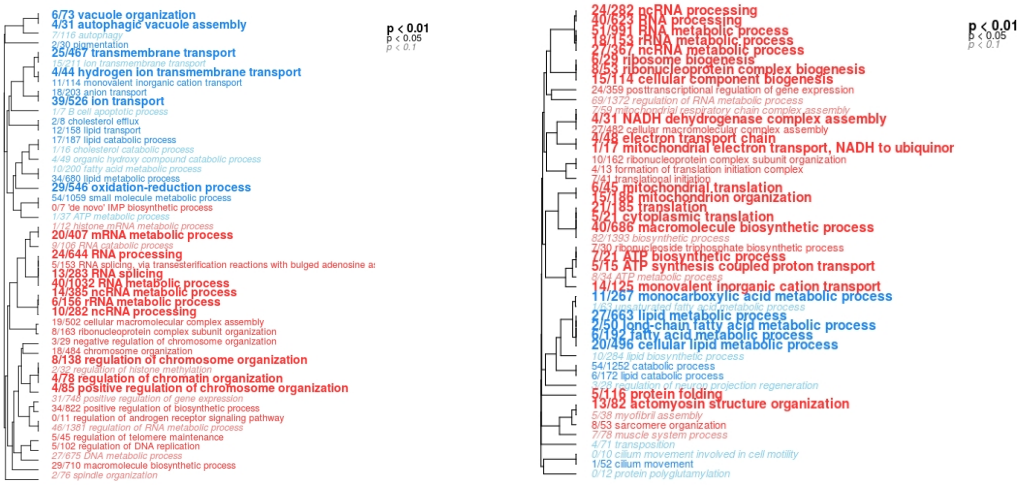
In scallop ctenidia, 68 genes showed a consistent response to low pH at 36 hours and 14 genes at 7 weeks. Functional enrichment indicated that upregulated genes were involved in RNA metabolic processes, chromatin reorganization, and regulation of DNA replication. Downregulated genes were involved in oxidation-reduction processes, lipid metabolism, and ion transmembrane transport (Fig. 3.11a). In the mantle tissue, there were 44 DEGs at 36 hours and 56 DEGs at 7 weeks. The top functional categories for upregulated DEGs in the mantle included RNA metabolic processes, ATP synthesis (e.g., electron transport chain), and translation in the mitochondria and cytoplasm. Downregulated DEGs in the mantle included lipid metabolism, protein folding, cilium movement, and muscle system processes (Fig. 3.11b).

	Ctenidia 36 hours	Ctenidia 7 weeks	Mantle 36 hours	Mantle 7 weeks
Number of samples	25	17	24	15
Number of genes	28,921	24,369	24,518	24,511
Response to low pH	68	14	44	56
Population effect	19	31	43	51
Population:pH interaction	38	26	19	50
AS response to low pH	148	NA	37	NA
WS response to low pH	184	337	123	84
WD response to low pH	191	225	87	59
CS response to low pH	259	162	117	104

Table 3.5: Number of differentially expressed genes for each tissue and time point in *C. gigantea*, as determined by DeSeq2.

Constitutive differences in gene expression among populations

Between 19 to 51 genes showed constitutive differences in gene expression among scallop population, when analyzing time point and tissue separately. However, none of these gene sets showed significant functional enrichment.



(a) Ctenidia (b) Mantle

Figure 3.11: GO_MWU results for *C. gigantea* genes with a conserved response to low pH across populations: a) ctenidia, b) mantle. Blue are downregulated in response to treatment, red are upregulated.

Population-specific responses to low pH

Between 19-50 genes across the two time points and tissues showed a significant interaction between population and treatment, however there were no GO categories functionally enriched in these gene sets. Within each population, 148-259 genes were differentially expressed in the ctenidia in response to low pH at 36 hours, with CS exhibiting the largest number of DEGs. At 7 weeks, Functionally, DEGs within each population were subsets of the conserved treatment response genes, with the exception of CS. For CS, genes involved in chromatin organization were downregulated instead of upregulated, along with novel functions of RNA metabolism and organ development (not shown).

In the mantle tissue, AS had the fewest DEGs with 37 at 36 hours, while WS had the greatest with 123 at 36 hours (Table 3.5). As in the ctenidia, most DEG sets within populations were subsets of those genes with a strong intraspecific trend in expression. Only WD exhibited novel functional groups, with some DEGs involved in calcium and metal ion transport (not shown).

KOG class enrichment was compared among 14 DEG sets representing each combination

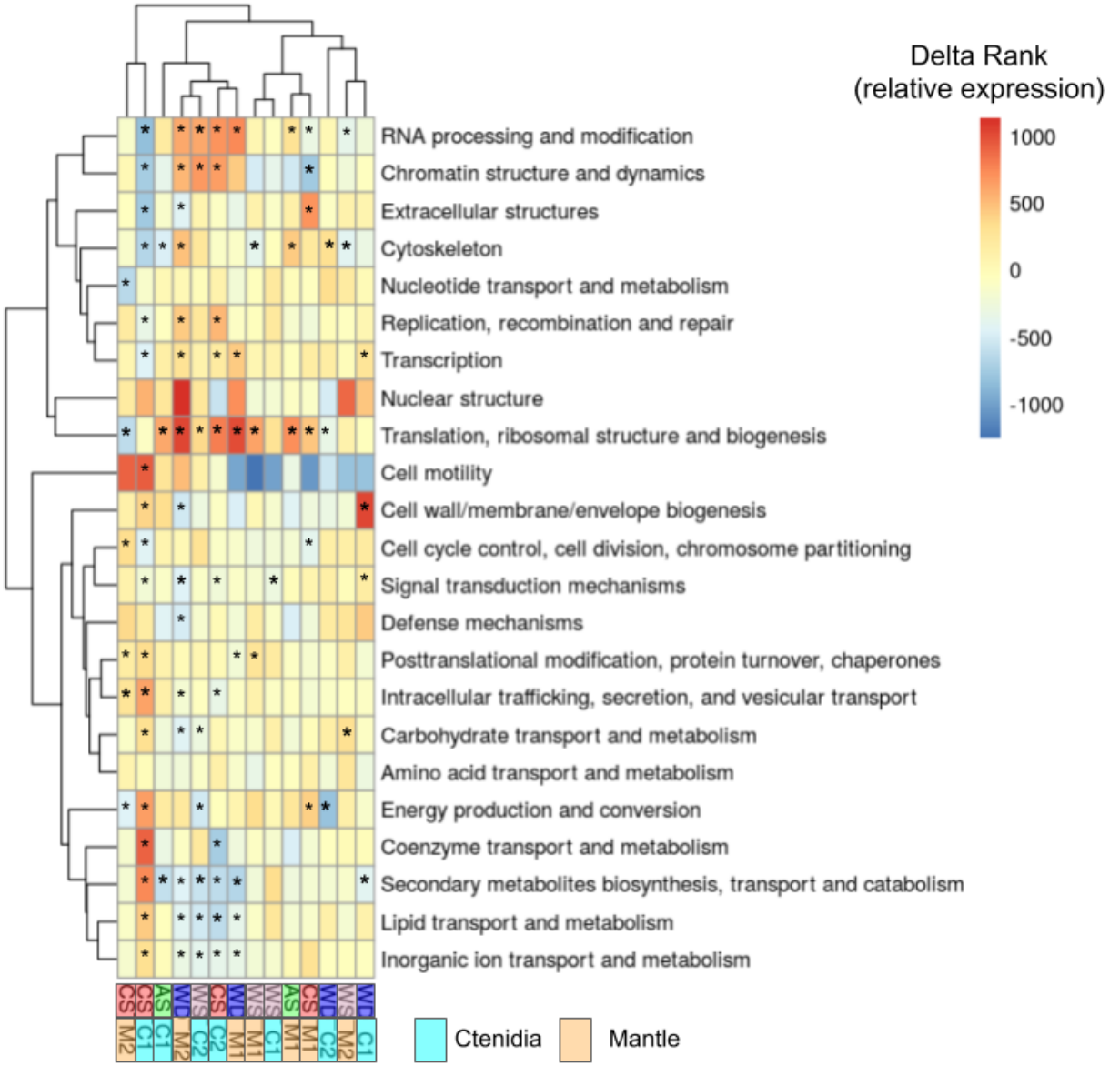


Figure 3.12: KOG_MWU comparison for *C. gigantea* genes that are differentially expressed in response to low pH, for all combinations of tissues, time points, and populations. Red indicates upregulated, blue is downregulated; * indicate significant enrichment after FDR correction, ***** represents the top upregulated and top downregulated classes for each dataset. Datasets are clustered by their correlation.

of population, tissue, and time ours were outliers in expression patterns of KOG classes compared to the other datasets. While 'Translation, ribosomal structure, and biogenesis' was the most pronounced upregulated class for most datasets (including CS mantle genes at 36 hours), it was the most significantly downregulated class for CS mantle genes at 7 weeks.

In CS ctenidia tissues at 36 hours, KOG classes classes involving certain metabolism and transport functions (e.g., lipids, inorganic ions) were upregulated, as opposed to downregulation in other datasets. Mantle tissues from WD for both time points exhibited consistent expression patters of KOG classes, as opposed to mantle tissues from CS and WS where opposite patterns of up- or downregulation were observed between time points (Fig. 3.12).

3.3.4 Comparing oysters and scallops

For most DEG sets, the Olympia oyster exhibited a larger number of differentially expressed genes compared to the rock scallop despite having a smaller transcriptome. Some of the GO functional groups with consistent intraspecific changes in respect to low pH were shared between species. In the ctenidia, both species had downregulation of genes involved in phagocytosis and upregulation of ncRNA metabolic processes. Oysters had upregulation of lipid metabolic processes, while scallops showed downregulation in response to low pH. Some of the DEGs that were only evident in the OR population showed the same expression in the scallops, such as downregulation of hydrogen ion transmembrane transport and ATP metabolic processes.

A comparison of KOG classes between oyster and scallop datasets for each tissue and time point showed that genes involved in 'Translation, ribosomal structure, and biogenesis' were significantly upregulated in almost all datasets. Both scallop tissues at 7 weeks exhibited downregulation of 'Secondary metabolites biosynthesis, transport, and catabolism', while all other datasets showed upregulation for this KOG class. Other interesting patterns included consistent downregulation of 'Extracellular structures' and 'Chromatin structure and dynamics' in the ctenidia of both species at 36 hours (Fig. 3.13).

3.3.5 Mapping candidate adaptive loci from Ch. 1

Of the 129 putative adaptive GBS loci identified in Silliman 2019, 26 mapped to the *O. lurida* transcriptome generated for this study across 25 different orthogroups. 14 of these

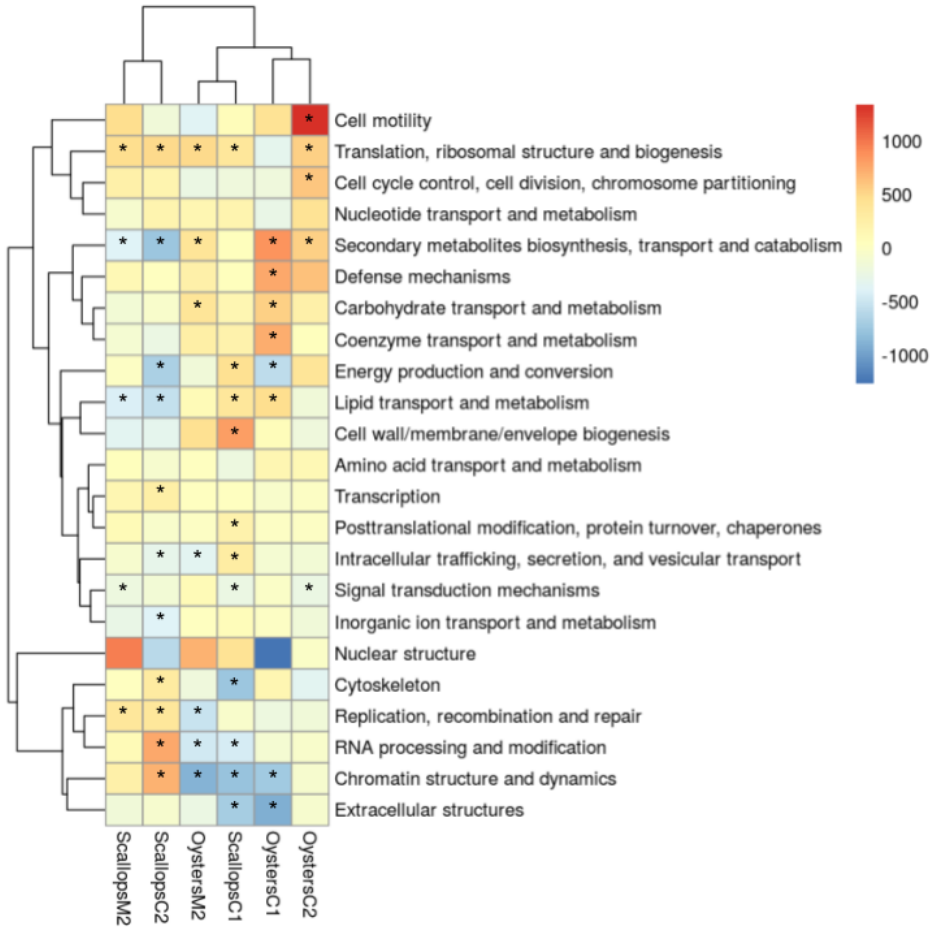


Figure 3.13: Comparison of KOG_MWU enrichment among ctenidia and mantle datasets for both species at both time points, with the exception of *O. lurida* mantle genes at 36 hours, as no KOG classes were enriched. Red indicates upregulated, blue is downregulated; * indicate significant enrichment after FDR correction, ***** represents the top upregulated and top downregulated classes for each dataset. Datasets are clustered by their correlation.

orthogroups belonged to at least one DEG set identified by DeSeq2. Functional annotations were similar for those that were functionally annotated in both the GBS assembly and transcriptome (Table 3.6).

Orthogroup	Gene name	KOG class	GBS locus	DEG set
OG0005445	None	None	Locus_42	pop:pH (CT1), ORlowpH (CT1)
OG0005274	NADH5	C	Locus_277490	pop:pH (MT2)
OG0012853	viral oncogene	T	Locus_339584	pop:pH (CT1), OR (CT1), BC (CT1, MT2)
OG0020784	None	None	Locus_42995	pop:pH (CT1), population (CT1), OR (CT1)
OG0009857	Tudor domain	K	Locus_324480	pop:pH (CT1), OR (CT1)
OG0036298	INO80 homolog	L	Locus_252560	pop:pH (CT1), population (MT1), OR (CT1)
OG0018595	polycystic kidney	P, T	Locus_170867	pop:pH (CT1), population (CT1), OR (CT1), BC (CT1, MT2)
OG0004314	phosphase	P	Locus_292806	pop:pH (CT1), BC (MT2)
OG0006037	None	None	Locus_185179	pop (CT1, MT1,T2)
OG0011776	Myosin XVIIIA	Z	Locus_196263	pop:pH (CT1), OR (CT1), BC(MT2)
OG0000156	Carbohydrate sulfotransferase 15	C	Locus_12991	population
OG0029433	Ribosomal protein	J	Locus_267489	pop:pH (CT1), OR(CT1)
OG0001632	TBC1 domain member	U	Locus_54082	population (MT2)
OG0036799	solute carrier family 34 (sodium phosphate)	P	Locus_18437, Locus_44811	BC (MT2)

Table 3.6: Candidate adaptive markers in *O. lurida*, identified using both GBS and transcriptomics. DEG set refers to the gene set where the gene is significantly expressed (e.g., OR (CT1) is differentially expressed in OR ctenidia samples at T1-36 hours in response to treatment)). Abbreviations for KOG classes: C: Energy production and conversion, J: Translation, ribosomal structure and biogenesis, K: Transcription, L: Replication, recombination and repair, P: Inorganic ion transport and metabolism, T: Signal transduction mechanisms, U: Intracellular trafficking, secretion, and vesicular transport, Z: Cytoskeleton

3.4 Discussion

Using a comparative transcriptomics approach, I have characterized and compared the molecular response of two marine bivalves to variable low pH conditions. Previous studies on the effects of acidification on transcriptomic or proteomic response in oysters (*Ostrea*, *Crassostrea*, *Saccostrea*, and *Pinctada fucata*) have primarily focused on the larval stage (Dineshram et al., 2016; De Wit et al., 2018), on a single population (Timmins-Schiffman et al., 2014; Liu et al., 2017; Zhang et al., 2015), or on a small subset of candidate genes (Wang et al., 2016; Heare et al., 2018). This study is the first to examine the molecular response of either an *Ostrea* species or *C. gigantea* to acidification, providing valuable molecular resources for both taxa.

3.4.1 Conserved responses to acidification

Species-wide response in Olympia oysters

Although both species exhibited considerable intraspecific variation in transcriptomic response to acidification, there were some functional responses that were consistent across populations within each species. Proteomics and transcriptomics of adult *C. gigas* and *C. virginica* in OA conditions (-0.4 to -0.5 pH relative to controls) have demonstrated that in the ctenidia, acidification enhances expression of genes regulating antioxidant processes, organization of the cytoskeleton, and redox balancing, as well as changes abundance of proteins involved in metabolic processes (Chapman et al., 2011; Goncalves et al., 2017a; Timmins-Schiffman et al., 2014). Many of the ctenidia genes showing a species-wide response to low pH in *O. lurida* were consistent with these results (Fig. 3.3). Unlike these other studies, genes explicitly involved in 'oxidation-reduction processes' were not included in this set.

The enriched functions for down-regulated ctenidia genes in response to pH were primarily involved with carbohydrate metabolism, immunity, signaling pathways, and protein transport. All of these are key components of the universal cellular stress response (CSR), and have been associated with oyster response to other environmental stressors (Anderson

et al., 2015). Interestingly, these down-regulated genes included GO terms involved in negative regulation of glucose metabolism, potentially resulting in increased glucose production. This result aligns with a proteomics study in *C. gigas*, where proteins involved in gluconeogenesis were detected at higher levels in oysters held at elevated pCO2 compared to a control (Timmins-Schiffman et al., 2014). Increased glucose metabolism may be required in order to maintain cellular homeostasis, with prolonged elevated metabolism leading to depletion of glucose stores and potential downstream fitness consequences. Down-regulation of CSR-related genes has been considered by some a sign of tolerance to the induced stress (Bailey et al., 2017). This result highlights the importance of considering the regulatory function of the gene product itself when inferring putative physiological responses.

Upregulated genes in *O. lurida* ctenidia included those involved in fatty acid metabolism and transport of organic acids. Lipid metabolism-associated genes have been found to be up-regulated in other marine invertebrates (Wong et al., 2011; Timmins-Schiffman et al., 2014), which may function to protect cellular membranes from reactive oxygen species damage (Pamplona et al., 2002). The coupling of changes in both lipid and carbohydrate metabolism suggest a significant shift in energetic resources during low pH stress. Upregulation of genes involved in organic acid transport (e.g., carboxylic acid) may be helping to maintain cellular pH homeostasis (Hurth et al., 2005).

Other studies of response to acidification in bivalve mantle tissue have shown enhanced expression of immune response genes, protease inhibitors, the production of antioxidants, and changes in the expression of cytoskeleton-related genes (Clark et al., 2013; Hüning et al., 2013). Notably, all of these studies assessed molecular changes after at least 30 days of exposure to lowered pH. While there were no significant DEGs or enriched GO terms for *O. lurida* mantle tissues at 36 hours, the 15 enriched GO terms at 7 weeks aligned with many of those found in previous studies. In particular, the decreased expression of DNA repair genes matches proteomic results from *C. gigas* , which has been hypothesized as evidence for a shift away from repairing DNA that has been damaged by increased oxidative stress

(Timmins-Schiffman et al., 2014). The downregulation of genes associated with cytoskeleton structuring aligns with another study of *C. gigas*, and may indicate the occurrence of cytoskeletal remodeling in order to minimize apoptosis due to intracellular stress (Goncalves et al., 2017a). One surprising result is the enrichment of genes involved in chromosome reorganization, which has not been noted in any previous oyster acidification studies. These genes merit further investigation of their putative function. The lack of significant changes in the mantle at 36 hours compared to the ctenidia suggests that these tissues play different roles in mediating acute pH stress response.

Comparison to rock scallops

In *C. gigantea* ctenidia, genes with a consistent response to OA included upregulation of RNA metabolic processes (mRNA, ncRNA, and rRNA), chromosome organization, and signaling pathways. While RNA processing or metabolism was mentioned in passing in some previous transcriptomic studies of invertebrates in response to OA (Timmins-Schiffman et al., 2014; Bailey et al., 2017), to my knowledge there is not a satisfactory explanation as why these genes would be differentially expressed in response to an environmental stress. As they are significantly enriched in DEGs from *O. lurida* ctenidia and both *C. gigantea* tissues, these genes merit further investigation. Downregulated genes in *C. gigantea* ctenidia included those involved in homeostasis, metabolism (lipid, fatty acids), and oxidation-reduction. Unlike in *O. lurida*, genes involved in gluconeogenesis did not change consistently in response to OA, however lipid metabolism was affected (but in the opposite direction). Similar to *O. lurida*, more *C. gigantea* ctenidia genes showed a conserved differential response to acidification at 36 hours than at 7 weeks, suggesting that physiological changes in the ctenidia vary between an acute and prolonged acidification response.

In *C. gigantea* mantle tissues, upregulated genes were involved in RNA processing, the electron transport chain (ETC) and associated pathways, and translation. NADH dehydrogenases have been shown to increase in oysters in response to a range of stressors, including

response to low pH in *C. gigas* (Timmins-Schiffman et al., 2014; Chapman et al., 2011). In one of the only comparative transcriptomic studies in oysters, researchers found that a low pH resistant strain of *Saccrostrea* (B2) had increased transcription of genes involved in the mitochondrial electron transport chain than wild type oysters, suggesting that B2 oysters had evolved an altered metabolic and oxidative stress response to low pH (Goncalves et al., 2017b). The results presented here suggest that *O. lurida* and *C. gigantea* also have different oxidative stress responses to this pH level.

3.4.2 Constitutive differences in gene expression among populations

Constitutive differences in gene expression, or how populations vary even without exposure to a stressor, have been proposed as important for determining environmental tolerance limits (DeBiasse and Kelly, 2016; Maynard et al., 2018). In *O. lurida*, both tissues had numerous genes exhibiting constitutive differences in expression among populations, with considerable overlap in function among genes sets. Enriched GO terms that differed among populations were overwhelmingly involved in immune function (Fig. 3.5), and particularly in antiviral defense (e.g., positive regulation of interferon-beta production). This result suggests that at least one of the populations either had a viral infection that was not present in the others, or that the infection was present in all populations but one population had a different molecular response to the infection due to no evolved immunity. *Vibrio* is a common viral disease in oysters, and is prevalent in many Oregon estuaries (Pfister et al., 2014).

A comparative study of response to low salinity in three *O. lurida* populations from California identified over 1,000 transcripts that varied in expression among populations, and suggested that modifying baseline expression can influence environmental tolerance and may contribute to adaptive differences among populations (Bible and Sanford, 2016). In this study, between 814 and 110 genes were differentially expressed among pairs of populations at the ambient pH level, with the most between Oregon and California (Table 3.4). The severe immune response observed in my study may be masking some population-specific responses

to the low pH treatment. With the exception of some general signaling pathways, the only GO terms without a connection to immune response are found in the ctenidia and involved in reproduction (e.g., multi-organism reproductive process) (Fig. 3.4). As demonstrated in Chapter 2, heritable variation in reproductive timing can develop even among populations with high connectivity.

3.4.3 Population-specific responses to low pH

Genes that exhibit population-specific responses to low pH may be the result of local adaptation to variable pH conditions along the west coast of North America. In *O. lurida*, over 4,000 genes showed a population:pH interaction in the ctenidia at 36 hours, and 205 in the mantle at 7 weeks. The ctenidia and mantle tissue exhibited contrasting patterns of population-specific expression through time. In particular, the number of ctenidia genes with a population-specific response decreased substantially between 36 hours and 7 week while the number of mantle genes increased. This result suggests that these two tissues have different roles during acclimatization or adaptation to low pH in *O. lurida*.

Comparison of KOG functional classes highlighted that genes involved in 'Energy production and conversion' had pronounced differences in upregulation or downregulation between populations (Fig. 3.8. One potential explanation for this pattern is that these populations stem from two diverged mitochondrial lineages, resulting in divergent functioning of the oxidative phosphorylation (OxPhos) pathway in response to low pH stress. The OxPhos pathway generates the majority of cellular ATP, and is encoded by both mitochondrial and nuclear genes. Due to elevated mutation rates in the mitochondria relative to the nuclear genome, populations can accumulate changes in the mitochondria quite rapidly, in some cases resulting in mitonuclear conflicts for genes in the OxPhos pathway (Hill, 2017; Barreto et al., 2018). Divergent mitochondrial haplotypes have been shown to generate gene-by-environment interactions in response to thermal stress for *Fundulus* fish species (Baris et al., 2016). Evidence from Chapter 1 and previous phylogenetic studies in *Ostrea spp.* suggest

that rapid mitochondrial evolution may contribute to adaptive divergence in Olympia oysters (Xiao et al., 2015).

C. gigantea did not provide as clear of a picture of adaptive divergence among populations. Only 19-50 genes showed a population:pH interaction across datasets, and none of these were functionally enriched for GO terms with using Fisher's exact test. However, comparing KOG class enrichment among datasets illustrated some specific cases of potential local adaptation. Ctenidia tissues in CS at 36 hours showed opposite expression patterns at 11 different KOG classes than most other datasets. 'Energy production and conversion' showed differential regulation between ctenidia samples from WD, CS, and WS (Fig. 3.12). This functional class may not have the same strong pattern of population divergence as was seen in *O. lurida*, because these genes have a conserved species-wide response to low pH (Fig. 3.11).

3.4.4 Candidate adaptive markers

Fourteen *O. lurida* loci were robustly identified as potentially involved in local adaptation through the intersection of reduced-representation sequencing, transcriptomics, and an acidification experiment. Four of these showed constitutive variation in expression across populations, and the rest had a significant population:pH interaction. The most frequent KOG annotations for those that were functionally annotated were 'Energy production and conversion' and 'Inorganic ion transport and metabolism'. Although these loci have demonstrated both geographical variation and population-specific response to pH, there are two major caveats for evaluating their importance to adaptive divergence in *O. lurida*. 1) Only functional assays can determine if variants in these genes have effects of fitness, 2) there are almost certainly many other genomic regions involved in local adaptation that are not represented here. Nonetheless, these 14 loci provide strong hypotheses of the gene regions involved in local adaptation to pH, and are excellent candidates for identifying cryptic variation in pH response among *O. lurida* populations (Eizaguirre and Baltazar-Soares, 2014). An

ongoing RADseq study of *C. gigantea* by collaborators at UWashington offers the potential for adaptive markers to be identified similiarly in that species.

DATA ACCESSIBILITY

Chapter 1

Genomic data (all filtered markers, putative neutral markers, and putative outliers) and sample metadata are available on Dryad (`https://doi.org/10.5061/dryad.114j8m1`). Raw demultiplexed DNA sequences for all sequenced individuals with > 200,000 raw sequencing reads are available on NCBI SRA (Project Accession Number: PRJNA511386). Reproducible Jupyter notebooks are available at

`https://github.com/ksil91/Ostrea_PopStructure`.

Chapter 2

The datasets generated during the Ch. 2 study are available on figshare, `DOI:10.6084/m9.figshare.5975452`. Reproducible R Markdown notebooks detailing the code used for statistical analyses can be found at `www.github.com/ksil91/PS-Oly-Larvae-Growth`.

APPENDIX: SUPPLEMENTARY MATERIAL FOR CH. 1

Sampling locations and population-specific summary statistics

Sampling Site	Latitude	Longitude	# of individuals used in analysis	H_e	F_{IS} (C.I.)
Klaskino Inlet, BC	50.29867	-127.72363	8	0.1903	0.0686 (0.0594 - 0.0786)
Barkley Sound, BC	49.01585	-125.31417	5	0.1865	0.0.0664 (0.0540 - 0.0795)
Ladysmith Harbour, BC	49.01138	-123.8357	5	0.1897	−0.0554 (−0.0700 - −0.0396)
Victoria Gorge, BC	48.43567	-123.37791	7	0.1717	0.0515 (0.0405 - 0.0631)
Discovery Bay, Puget Sound, WA	47.9978	-122.8824	7	0.1810	0.0593 (0.0483 - 0.0699)
Liberty Bay, Puget Sound, WA	47.7375	-122.6507	6	0.1768	0.0309 (0.0194 - 0.0426)
Triton Cove, Puget Sound, WA	47.6131	-122.982	6	0.1820	0.0336 (0.0219 - 0.0462)
North Bay, Puget Sound, WA	47.3925	-122.8138	6	0.1756	0.0524 (0.0404 - 0.0634)
Willapa Bay, WA (North & South)	46.62477 46.4400	-123.98879 -124.004	3 2	0.1798	0.0556 (0.0420 - 0.0685)
Netarts Bay, OR	45.39116	-123.95590	7	0.1968	0.0584 (0.0470 - 0.0698)
Yaquina Bay, OR	44.57954	-123.99577	6	0.1876	0.0143 (0.0018 - 0.0274)
Coos Bay, OR	43.35599	-124.19316	6	0.1809	0.0531 (0.0411 - 0.0654)
Humboldt Bay, CA	40.85580	-124.09746	6	0.2146	0.0327 (0.0209 - 0.0451)
Tomales Bay, CA	38.11755	-122.87450	6	0.2270	−0.0023 (−0.0133 - 0.0077)
Point Orient, San Francisco Bay, CA	37.95507	-122.42180	5	0.2209	0.0560 (0.0450 - 0.0668)
Candlestick Park, San Francisco Bay, CA	37.70867	-122.37761	4	0.2234	−0.0974 (−0.1181 - −0.0745)
Elkhorn Slough, CA	36.83982	-121.74278	6	0.2477	0.0859 (0.0745 - 0.0978)
Mugu Lagoon, CA	34.10191	-119.10434	9	0.2535	0.1327 (0.1239 - 0.1411)
San Diego Bay, CA	32.60250	-117.11889	7	0.2500	0.0948 (0.0851 - 0.1049)

Table S1: GPS coordinates of sampling sites and population-specific summary statistics averaged across markers using the combined dataset of 13,424 SNPs. H_e, expected heterozygosity; F_{IS}, inbreeding coefficient within the population, mean and 25%-75% confidence intervals (Nei and Chesser, 1983)

Summary statistics for phylogeographic regions

Region	H_o	H_e	F_{IS}	F_{ST}
NWBC	0.177	0.193	0.0821 (0.0738 - 0.0897)	0.016
Puget+BC	0.174	0.189	0.0814 (0.0758 - 0.0862)	0.046
Willapa	0.171	0.182	0.0583 (0.0495 - 0.0666)	0.001
Oregon	0.185	0.196	0.0556 (0.0474 - 0.0645)	0.016
NoCal	0.215	0.227	0.0536 (0.0472 - 0.0592)	0.022
SoCal	0.224	0.253	0.115 (0.1097 - 0.1209)	0.007

Table S2: Overall summary statistics for each phylogeographic region using the neutral dataset of 13,073 SNPs. H_o, observed heterozygosity averaged across loci; H_e, expected heterozygosity averaged across loci; F_{IS} & F_{ST}, Wright's F-statistics averaged across loci (Nei and Chesser, 1983). Note that F_{ST} may be skewed by variation in sampling strategy across regions.

Pairwise FST values for outlier and neutral datasets

	Klaskino_BC	Barkley_BC	Ladysmith_BC	Victoria_BC	Discovery_WA	Liberty_WA	TritonCove_WA	NorthBay_WA	Willapa_WA	Netarts_OR	Yaquina_OR	Coos_OR	Humboldt_CA	Tomales_CA	NorthSanFran_CA	SouthSanFran_CA	Elkhorn_CA	MuguLagoon_CA
Klaskino_BC	0																	
Barkley_BC	0.0306	0																
Ladysmith_BC	0.0663	0.0564	0															
Victoria_BC	0.1193	0.1270	0.0931	0														
Discovery_WA	0.0742	0.0768	0.0381	0.0873	0													
Liberty_WA	0.0965	0.1021	0.0650	0.0864	0.0526	0												
TritonCove_WA	0.0751	0.0800	0.0516	0.0575	0.0176	0.0166	0											
NorthBay_WA	0.0965	0.1050	0.0750	0.0904	0.0476	0.0075	0.0246	0										
Willapa_WA	0.1136	0.1228	0.1197	0.1247	0.1056	0.1184	0.0946	0.1141	0									
Netarts_OR	0.0825	0.0888	0.0824	0.1097	0.0804	0.0854	0.0703	0.0777	0.0642	0								
Yaquina_OR	0.1190	0.1304	0.1276	0.1480	0.1220	0.1379	0.1184	0.1359	0.0948	0.0310	0							
Coos_OR	0.1159	0.1245	0.1265	0.1338	0.1120	0.1217	0.1048	0.1189	0.00139	0.0687	0.0918	0						
Humboldt_CA	0.0915	0.0936	0.0789	0.1172	0.0884	0.0877	0.0767	0.0870	0.0927	0.0684	0.1045	0.0927	0					
Tomales_CA	0.0970	0.0940	0.0886	0.1282	0.1033	0.1165	0.0998	0.1171	0.1122	0.0821	0.1116	0.1117	0.0397	0				
NorthSanFran_CA	0.1248	0.1222	0.1198	0.1555	0.1309	0.1452	0.1252	0.1458	0.1207	0.1023	0.1244	0.1280	0.0176	0.0418	0			
SouthSanFran_CA	0.1256	0.1204	0.1258	0.1717	0.1424	0.1566	0.1347	0.1574	0.1364	0.1025	0.1318	0.1364	0.0295	0.0399	0.0039	0		
Elkhorn_CA	0.1228	0.1135	0.1108	0.1578	0.1292	0.1431	0.1182	0.1417	0.1231	0.1004	0.1247	0.1284	0.0470	0.0261	0.0398	0.0135	0	
MuguLagoon_CA	0.1351	0.1250	0.1241	0.1698	0.1414	0.1559	0.1366	0.1588	0.1402	0.1212	0.1386	0.1448	0.0880	0.0523	0.0779	0.0453	0.0164	0
SanDiego_CA	0.1417	0.1322	0.1320	0.1764	0.1445	0.1619	0.1392	0.1641	0.1481	0.1270	0.1467	0.1534	0.0892	0.0604	0.0811	0.0520	0.0132	0.0016

Table S3: Pairwise F_{ST} values for all pairs of populations, using the neutral dataset of 13,073 SNPs.

	Klaskino_BC	Barkley_BC	Ladysmith_BC	Victoria_BC	Discovery_WA	Liberty_WA	TritonCove_WA	NorthBay_WA	Willapa_WA	Netarts_OR	Yaquina_OR	Coos_OR	Humboldt_CA	Tomales_CA	NorthSanFran_CA	SouthSanFran_CA	Elkhorn_CA	MuguLagoon_CA
Barkley_BC	0.0660	0																
Ladysmith_BC	0.1848	0.0990	0															
Victoria_BC	0.4565	0.3553	0.3192	0														
Discovery_WA	0.2161	0.1151	0.0804	0.2707	0													
Liberty_WA	0.3705	0.2820	0.1894	0.1837	0.1306	0												
TritonCove_WA	0.3817	0.2981	0.1886	0.1616	0.1336	0.0398	0											
NorthBay_WA	0.3863	0.2874	0.2008	0.1777	0.1090	0.0109	0.0524	0										
Willapa_WA	0.4510	0.4372	0.4018	0.5178	0.3687	0.4915	0.4768	0.4871	0									
Netarts_OR	0.4119	0.3803	0.3214	0.3964	0.3247	0.3686	0.3832	0.3678	0.1496	0								
Yaquina_OR	0.4554	0.4401	0.3936	0.4739	0.4003	0.4548	0.4589	0.4515	0.2160	0.0453	0							
Coos_OR	0.4497	0.4396	0.3987	0.5085	0.3737	0.4934	0.4808	0.4839	0.0090	0.1328	0.2212	0						
Humboldt_CA	0.4659	0.4289	0.3807	0.5065	0.4316	0.4681	0.4676	0.4614	0.4397	0.3945	0.4271	0.4461	0					
Tomales_CA	0.3696	0.3283	0.3357	0.4652	0.3636	0.4571	0.4369	0.4443	0.3772	0.3299	0.3298	0.3935	0.2466	0				
NorthSanFran_CA	0.5057	0.4690	0.4456	0.5594	0.4951	0.5467	0.5366	0.5298	0.4958	0.4557	0.4642	0.4979	0.0475	0.2435	0			
SouthSanFran_CA	0.4822	0.4257	0.4303	0.5647	0.4762	0.5484	0.5446	0.5416	0.4800	0.4352	0.4515	0.4941	0.0830	0.1520	0.0637	0		
Elkhorn_CA	0.4782	0.4275	0.4407	0.5410	0.4716	0.5444	0.5343	0.5341	0.4457	0.4341	0.4413	0.4716	0.3276	0.1446	0.2733	0.1697	0	
MuguLagoon_CA	0.5749	0.5378	0.5490	0.6299	0.5760	0.6278	0.6190	0.6228	0.5608	0.5531	0.5555	0.5825	0.5130	0.3423	0.4705	0.3567	0.1320	0
SanDiego_CA	0.5744	0.5313	0.5471	0.6320	0.5758	0.6278	0.6194	0.6308	0.5700	0.5466	0.5507	0.5822	0.5059	0.3278	0.4753	0.3366	0.1271	0.0172

Table S4: Pairwise F_{ST} values for all pairs of populations, using the outlier dataset of 235 SNPs.

Additional results of outlier analyses

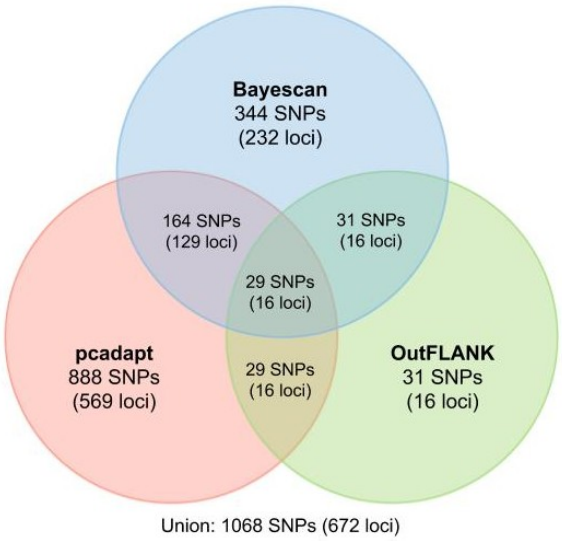


Figure S1: Venn diagram with number of SNPs and Genotype-by-Sequencing loci identified as outliers by three methods: *pcadapt*, *OutFLANK*, and BayeScan

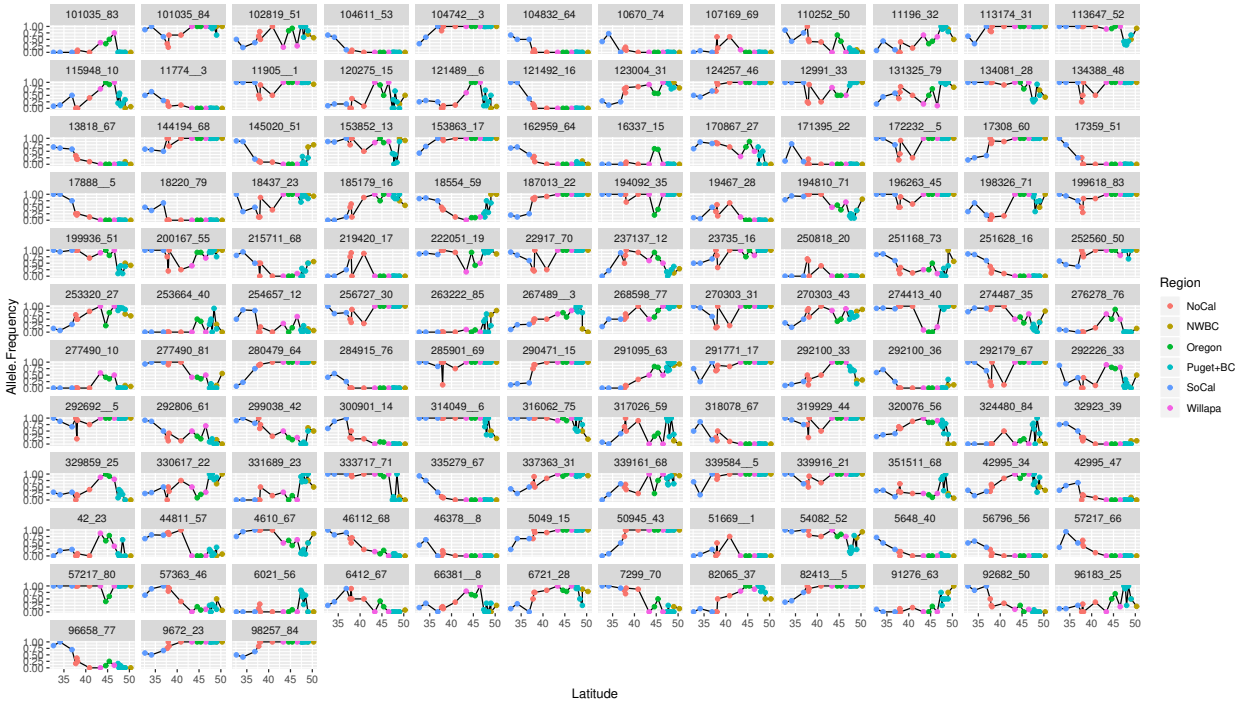


Figure S2: Outlier loci predominantly show clinal patterns in allele frequency. Allele frequency in 129 individual outlier loci plotted against latitude for 19 populations of *O. lurida*. One SNP is represented for each locus, except in the case where two outlier SNPs from the same locus showed different spatial patterns (e.g., locus_277490). Populations are colored by inferred phylogeographic regions.

REFERENCES

Abbott, L. A. and Hollenberg, G. J. (1976). *Marine algae of California*. Stanford University Press, Stanford, CA.

Allendorf, F. W., Hohenlohe, P. A., and Luikart, G. (2010). Genomics and the future of conservation genetics. *Nat. Rev. Genet.*, 11(10):697–709.

Anderson, K., Taylor, D. A., Thompson, E. L., Melwani, A. R., Nair, S. V., and Raftos, D. A. (2015). Meta-analysis of studies using suppression subtractive hybridization and microarrays to investigate the effects of environmental stress on gene transcription in oysters. *PLoS One*, 10(3):e0118839.

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., and Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.*, 17(2):81–92.

Ashley, M. (2010). Plant parentage, pollination, and dispersal: How DNA microsatellites have altered the landscape. *CRC Crit. Rev. Plant Sci.*, 29:148–161.

Bailey, A., De Wit, P., Thor, P., Browman, H. I., Bjelland, R., Shema, S., Fields, D. M., Runge, J. A., Thompson, C., and Hop, H. (2017). Regulation of gene expression is associated with tolerance of the Arctic copepod *Calanus glacialis* to CO_2-acidified sea water. *Ecol. Evol.*, 7(18):7145–7160.

Baker, P. (1995). Review of ecology and fishery of the Olympia oyster, *Ostrea lurida* with annotated bibliography. *J. Shellfish Res.*, 14(2):501–518.

Baker, P., Richmond, N., and Terwilliger, N. (1999). Reestablishment of a native oyster, *Ostrea conchaphila*, following a natural local extinction. In Pederson, J., editor, *Marine Bioinvasions*, pages 221–231.

Banas, N. S., McDonald, P. S., and Armstrong, D. A. (2009). Green crab larval retention in Willapa Bay, Washington: An intensive Lagrangian modeling approach. *Estuaries Coasts*, 32(5):893–905.

Barber, B. J., Ford, S. E., and Wargo, R. N. (1991). Genetic variation in the timing of gonadal maturation and spawning of the Eastern oyster, *Crassostrea virginica* (gmelin). *Biol. Bull.*, 181(2):216–221.

Barber, J. S., Dexter, J. E., Grossman, S. K., Greiner, C. M., and Mcardle, J. T. (2016). Low temperature brooding of Olympia oysters (*Ostrea lurida*) in northern Puget Sound. *J. Shellfish Res.*, 35(2):351–357.

Baris, T. Z., Blier, P. U., Pichaud, N., Crawford, D. L., and Oleksiak, M. F. (2016). Gene by environmental interactions affecting oxidative phosphorylation and thermal sensitivity. *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, 311(1):R157–65.

Barreto, F. S., Watson, E. T., Lima, T. G., Willett, C. S., Edmands, S., Li, W., and Burton, R. S. (2018). Genomic signatures of mitonuclear coevolution across populations of *Tigriopus californicus*. *Nat Ecol Evol*, 2(8):1250–1257.

Barrett, E. M. (1963). The California oyster industry. *Fish. Bull.*, 123.

Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proc. R. Soc. Lond. A Math. Phys. Sci.*, 160(901):268–282.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, Articles*, 67(1):1–48.

Baums, I. B. (2008). A restoration genetics guide for coral reef conservation. *Molecular Ecology*, 17(12):2796–2811.

Bell, G. and Gonzalez, A. (2009). Evolutionary rescue can prevent extinction following environmental change. *Ecol. Lett.*, 12(9):942–948.

Benestan, L., Quinn, B. K., Maaroufi, H., Laporte, M., Clark, F. K., Greenwood, S. J., Rochette, R., and Bernatchez, L. (2016). Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*). *Mol. Ecol.*, 25(20):5073–5092.

Bible, J. M. and Sanford, E. (2016). Local adaptation in an estuarine foundation species: Implications for restoration. *Biological Conservation*, 193:95–102.

Blake, B. and Bradbury, A. (2012). Washington Department of Fish and Wildlife plan for rebuilding Olympia oyster (*Ostrea lurida*) populations in Puget Sound with a historical and contemporary overview. Technical report, Washington Department of Fish and Wildlife, Brinnon, WA.

Blanchette, C. A., Melissa Miner, C., Raimondi, P. T., Lohse, D., Heady, K. E. K., and Broitman, B. R. (2008). Biogeographical patterns of rocky intertidal communities along the Pacific coast of North America. *J. Biogeogr.*, 35(9):1593–1607.

Blois, J. L., Williams, J. W., Fitzpatrick, M. C., Jackson, S. T., and Ferrier, S. (2013). Space can substitute for time in predicting climate-change effects on biodiversity. *Proc. Natl. Acad. Sci. U. S. A.*, 110(23):9374–9379.

Buonaccorsi, V. P., Kimbrell, C. A., Lynn, E. A., and Vetter, R. D. (2002). Population structure of copper rockfish (*Sebastes caurinus*) reflects postglacial colonization and contemporary patterns of larval dispersal. *Can. J. Fish. Aquat. Sci.*, 59(8):1374–1384.

Busch, D. S. and McElhany, P. (2016). Estimates of the direct effect of seawater pH on the survival rate of species groups in the California Current Ecosystem. *PLoS One*, 11(8):e0160669.

Byrne, M. (2011). Impact of ocean warming and ocean acidification on marine invertebrate life history stages. In *Oceanography and Marine Biology*, Oceanography and Marine Biology - An Annual Review. CRC Press.

Cabau, C., Escudié, F., Djari, A., Guiguen, Y., Bobe, J., and Klopp, C. (2017). Compacting and correcting trinity and oases RNA-Seq de novo assemblies. *PeerJ*, 5:e2988.

Camara, M. D. and Vadopalas, B. (2009). Genetic aspects of restoring Olympia oysters and other native bivalves: Balancing the need for action, good intentions, and the risks of making things worse. *Journal of Shellfish Research*, 28(1):121–145.

Cao, S., Zhu, L., Nie, H., Yin, M., Liu, G., and Yan, X. (2018). De novo assembly, gene annotation, and marker development using Illumina paired-end transcriptome sequencing in the *Crassadoma gigantea*. *Gene*, 658:54–62.

Cariou, M., Duret, L., and Charlat, S. (2016). How and how much does RAD-seq bias genetic diversity estimates? *BMC Evol. Biol.*, 16(1):240.

Carroll, S. P., Jorgensen, P. S., Kinnison, M. T., Bergstrom, C. T., Denison, R. F., Gluckman, P., Smith, T. B., Strauss, S. Y., Tabashnik, B. E., Jørgensen, P. S., Kinnison, M. T., Bergstrom, C. T., Denison, R. F., Gluckman, P., Smith, T. B., Strauss, S. Y., Tabashnik, B. E., Jorgensen, P. S., Kinnison, M. T., Bergstrom, C. T., Denison, R. F., Gluckman, P., Smith, T. B., Strauss, S. Y., and Tabashnik, B. E. (2014). Applying evolutionary biology to address global challenges. *Science*, 346(6207).

Carson, H. S. (2010). Population connectivity of the Olympia oyster in Southern California. *Limnology and Oceanography*, 55(1):134–148.

Chan, F., Barth, J. A., Bl, C. A., Byrne, R. H., C, F., and Cheriton, O. (2017). Persistent spatial structuring of coastal ocean acidification in the California Current System. *Scientific Reports*, 7:e2526.

Chapman, R. W., Mancia, A., Beal, M., Veloso, A., Rathburn, C., Blair, A., Holland, A. F., Warr, G. W., Didinato, G., Sokolova, I. M., Wirth, E. F., Duffy, E., and Sanger, D. (2011). The transcriptomic responses of the eastern oyster, *Crassostrea virginica*, to environmental conditions. *Mol. Ecol.*, 20(7):1431–1449.

Cheng, Q., Wang, H., Jiang, S., Wang, L., Xin, L., Liu, C., Jia, Z., Song, L., and Zhu, B. (2016). A novel ubiquitin-protein ligase E3 functions as a modulator of immune response against lipopolysaccharide in Pacific oyster, *Crassostrea gigas*. *Dev. Comp. Immunol.*, 60:180–190.

Clark, M. S., Thorne, M. A. S., Amaral, A., Vieira, F., Batista, F. M., Reis, J., and Power, D. M. (2013). Identification of molecular and physiological responses to chronic environmental challenge in an invasive species: the Pacific oyster, *Crassostrea gigas*. *Ecol. Evol.*, 3(10):3283–3297.

Coe, W. R. (1932). Development of the gonads and the sequence of the sexual phases in the California oyster (*Ostrea lurida*). *Scripps Institution of Oceanography*.

Coen, L. D., Dumbauld, B. R., and Judge, M. L. (2011). Expanding shellfish aquaculture: A review of the ecological services provided by and impacts of native and cultured bivalves

in shellfish-dominated ecosystems. In Shumway, S. E., editor, *Shellfish Aquaculture and the Environment*, chapter 9, pages 239–295. John Wiley & Sons, Inc.

ConcernedApe LLC (2019). Stardew Valley.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M. W., Gaffney, D. J., Elo, L. L., Zhang, X., and Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biol.*, 17:13.

Conover, D. O., Clarke, L. M., Munch, S. B., and Wagner, G. N. (2006). Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. *Journal of Fish Biology*, 69:21–47.

Cornwall, C. E. and Hurd, C. L. (2016). Experimental design in ocean acidification research: problems and solutions. *ICES J. Mar. Sci.*, 73(3):572–581.

Cowen, R. K., Lwiza, K. M., Sponaugle, S., Paris, C. B., and Olson, D. B. (2000). Connectivity of marine populations: Open or closed? *Science*, 287:857–859.

Cunningham, K. M., Canino, M. F., Spies, I. B., and Hauser, L. (2009). Genetic isolation by distance and localized fjord population structure in Pacific cod (*Gadus macrocephalus*): Limited effective dispersal in the northeastern Pacific Ocean. *Can. J. Fish. Aquat. Sci.*, 66(1):153–166.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., and 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15):2156–2158.

David, A. A. (2018). Reconsidering panmixia: The erosion of phylogeographic barriers due to anthropogenic transport and the incorporation of biophysical models as a solution. *Frontiers in Marine Science*, 5:280.

Dawson, M. N. (2001). Phylogeography in coastal marine animals: A solution from California? *Journal of Biogeography*, 28(6):723–736.

de Lorgeril, J., Saulnier, D., Janech, M. G., Gueguen, Y., and Bachère, E. (2005). Identification of genes that are differentially expressed in hemocytes of the Pacific blue shrimp (*Litopenaeus stylirostris*) surviving an infection with *Vibrio penaeicida*. *Physiol. Genomics*, 21(2):174–183.

De Wit, P., Durland, E., Ventura, A., and Langdon, C. J. (2018). Gene expression correlated with delay in shell formation in larval Pacific oysters (*Crassostrea gigas*) exposed to experimental ocean acidification provides insights into shell formation mechanisms. *BMC Genomics*, 19(1):160.

De Wit, P. and Palumbi, S. R. (2013). Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation. *Mol. Ecol.*, 22(11):2884–2897.

DeBiasse, M. B., Kawji, Y., and Kelly, M. W. (2018). Phenotypic and transcriptomic responses to salinity stress across genetically and geographically divergent *Tigriopus californicus* populations. *Mol. Ecol.*, 27(7):1621–1632.

DeBiasse, M. B. and Kelly, M. W. (2016). Plastic and evolved responses to global change: What can we learn from comparative transcriptomics? *J. Hered.*, 107(1):71–81.

Dijkstra, H. and Dijkstra, H. (2010). *Crassadoma Bernard*, 1986. `http://www.marinespecies.org/aphia.php?p=taxdetails&id=138316`.

Dineshram, R., Chandramouli, K., Ko, G. W. K., Zhang, H., Qian, P.-Y., Ravasi, T., and Thiyagarajan, V. (2016). Quantitative analysis of oyster larval proteome provides new insights into the effects of multiple climate change stressors. *Glob. Chang. Biol.*, 22(6):2054–2068.

Dittman, D., Ford, S. E., and Haskin, H. H. (1998). Growth patterns in oysters, *Crassostrea virginica*, from different estuaries. *Mar. Biol.*, 132(3):461–469.

Dixon, G. B., Davies, S. W., Aglyamova, G. A., Meyer, E., Bay, L. K., and Matz, M. V. (2015). Genomic determinants of coral heat tolerance across latitudes. *Science*, 348(6242):1460–1462.

Drinan, D. P., Gruenthal, K. M., Canino, M. F., Lowry, D., Fisher, M. C., and Hauser, L. (2018). Population assignment and local adaptation along an isolation-by-distance gradient in Pacific cod (*Gadus macrocephalus*). *Evol. Appl.*, 103:17302.

Dyke, A. S. and Prest, V. K. (1987). Late Wisconsinan and Holocene history of the Laurentide ice sheet. *Gographie physique et Quaternaire*, 41(2):237.

Earl, D. A. and vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.*, 4(2):359–361.

Eaton, D. A. R. (2014). PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics*, 30(13):1844–1849.

Eizaguirre, C. and Baltazar-Soares, M. (2014). Evolutionary conservation—evaluating the adaptive potential of species. *Evol. Appl.*, 7(9):963–967.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., and Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, 6(5):e19379.

Emms, D. M. and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.*, 16:157.

Epelboin, Y., Quintric, L., Guévélou, E., Boudry, P., Pichereau, V., and Corporeau, C. (2016). The kinome of Pacific oyster *Crassostrea gigas*, its expression during development and in response to environmental factors. *PLoS One*, 11(5):e0155435.

Everett, M. V., Park, L. K., Berntson, E. A., Elz, A. E., Whitmire, C. E., Keller, A. A., and Clarke, M. E. (2016). Large-scale genotyping-by-sequencing indicates high levels of gene flow in the deep-sea octocoral *Swiftia simplex* (Nutting 1909) on the west coast of the United States. *PLoS One*, 11(10):e0165279.

Fabry, V. J., Seibel, B. A., Feely, R. A., and Orr, J. C. (2008). Impacts of ocean acidification on marine fauna and ecosystem processes. *ICES J. Mar. Sci.*, 65(3):414–432.

Feely, R. A., Sabine, C. L., Martin Hernandez-Ayon, J., Ianson, D., and Hales, B. (2008). Evidence for upwelling of corrosive "acidified" water onto the continental shelf. *Science*, 320(5882):1490–1492.

Fenberg, P. B., Menge, B. A., Raimondi, P. T., and Rivadeneira, M. M. (2015). Biogeographic structure of the northeastern Pacific rocky intertidal: The role of upwelling and dispersal to drive patterns. *Ecography*, 38(1):83–95.

Folkvord, A., Jørgensen, C., Korsbrekke, K., Nash, R. D. M., Nilsen, T., and Skjæraasen, J. E. (2014). Trade-offs between growth and reproduction in wild Atlantic cod. *Canadian Journal Fisheries and Aquatic Sciences*, 71(7):1106–1112.

Foll, M. and Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180(2):977–993.

Franssen, S. U., Gu, J., Bergmann, N., Winters, G., Klostermeier, U. C., Rosenstiel, P., Bornberg-Bauer, E., and Reusch, T. B. H. (2011). Transcriptomic resilience to global warming in the seagrass *Zostera marina*, a marine foundation species. *Proc. Natl. Acad. Sci. U. S. A.*, 108(48):19276–19281.

Funk, W. C., McKay, J. K., Hohenlohe, P. A., and Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends Ecol. Evol.*, 27(9):489–496.

Gagnaire, P.-A., Broquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F., Arnaud-Haond, S., and Bierne, N. (2015). Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evol. Appl.*, 8(8):769–786.

Gaitán-Espitia, J. D. and Hofmann, G. E. (2017). Gene expression profiling during the embryo-to-larva transition in the giant red sea urchin *Mesocentrotus franciscanus*. *Ecol. Evol.*, 7(8):2798–2811.

Gattuso, J. P., Epitalon, J. M., Lavigne, H., Orr, J., Gentili, B., Hofmann, A., Proye, A., Soetaert, K., and Rae, J. (2015). seacarb: Seawater carbonate chemistry. *R package version*, 3(1).

Gautier, M., Gharbi, K., Cezard, T., Foucaud, J., Kerdelhué, C., Pudlo, P., Cornuet, J.-M., and Estoup, A. (2013). The effect of RAD allele dropout on the estimation of genetic variation within and between populations. *Mol. Ecol.*, 22(11):3165–3178.

Gleason, L. U. and Burton, R. S. (2016). Genomic evidence for ecological divergence against a background of population homogeneity in the marine snail *Chlorostoma funebralis*. *Mol. Ecol.*, 25:3557–3573.

Goncalves, P., Jones, D. B., Thompson, E. L., Parker, L. M., Ross, P. M., and Raftos, D. A. (2017a). Transcriptomic profiling of adaptive responses to ocean acidification. *Mol. Ecol.*, 26(21):5974–5988.

Goncalves, P., Thompson, E. L., and Raftos, D. A. (2017b). Contrasting impacts of ocean acidification and warming on the molecular responses of CO_2-resilient oysters. *BMC Genomics*, 18(1):431.

Gosselin, T. (2017). radiator: RADseq data exploration, manipulation and visualization using R.

Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., Robles, M., Talón, M., Dopazo, J., and Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.*, 36(10):3420–3435.

Goudet, J. and Jombart, T. (2015). *hierfstat: Estimation and Tests of Hierarchical F-Statistics*.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.*, 29(7):644–652.

Granek, E. F., Polasky, S., Kappel, C. V., Reed, D. J., Stoms, D. M., Koch, E. W., Kennedy, C. J., Cramer, L. A., Hacker, S. D., Barbier, E. B., Aswani, S., Ruckelshaus, M., Perillo, G. M. E., Silliman, B. R., Muthiga, N., Bael, D., and Wolanski, E. (2010). Ecosystem services as a common language for coastal ecosystem-based management. *Conserv. Biol.*, 24(1):207–216.

Grether, G. F. (2005). Environmental change, phenotypic plasticity, and genetic compensation. *Am. Nat.*, 166(4):E115–23.

Grosberg, R. and Cunningham, C. W. (2001). Genetic structure in the sea. In *Marine community ecology*, pages 61–84. Sinauer, Sunderland, MA.

Gruber, N., Hauri, C., Lachkar, Z., Loher, D., Frölicher, T. L., and Plattner, G.-K. (2012). Rapid progression of ocean acidification in the California Current System. *Science*, 337(6091):220–223.

Guillaume, A. S., Monro, K., and Marshall, D. J. (2016). Transgenerational plasticity and environmental stress: do paternal effects act as a conduit or a buffer? *Functional Ecology*, 30(7):1175–1184.

Hall, C. A. (1964). Shallow-water marine climates and molluscan provinces. *Ecology*, 45:226–234.

Hare, M. P. and Avise, J. C. (1996). Molecular genetic analysis of a stepped multilocus cline in the American oyster (*Crassostrea virginica*). *Evolution*, 50(6):2305–2315.

Hauser, L. and Carvalho, G. R. (2008). Paradigm shifts in marine fisheries genetics: Ugly hypotheses slain by beautiful facts. *Fish Fish*, 9(4):333–362.

Heare, J. E., Blake, B., Davis, J. P., Vadopalas, B., and Roberts, S. B. (2017). Evidence of *Ostrea lurida* Carpenter, 1864, population structure in Puget Sound, WA, USA. *Mar. Ecol.*, 38(5):e12458.

Heare, J. E., White, S. J., Vadopalas, B., and Roberts, S. B. (2018). Differential response to stress in *Ostrea lurida* as measured by gene expression. *PeerJ*, 6:e4261.

Hedgecock, D. (1994). Does variance in reproductive success limit effective population sizes of marine organisms. *Genetics and evolution of aquatic organisms*, 122.

Heino and Kaitala (1999). Evolution of resource allocation between growth and reproduction in animals with indeterminate growth. *J. Evol. Biol.*, 12(3):423–429.

Hellberg, M. E. (2009). Gene flow and isolation among populations of marine animals. *Annu. Rev. Ecol. Evol. Syst.*, 40:291–310.

Hercus, M. J. and Hoffmann, A. A. (2000). Maternal and grandmaternal age influence offspring fitness in *Drosophila*. *Proc. Biol. Sci.*, 267(1457):2105–2110.

Hettinger, A., Sanford, E., Hill, T. M., Lenz, E. A., Russell, A. D., and Gaylord, B. (2013). Larval carry-over effects from ocean acidification persist in the natural environment. *Glob. Chang. Biol.*, 19(11):3317–3326.

Hill, G. E. (2017). The mitonuclear compatibility species concept. *Auk*, 134(2):393–409.

Hoey, J. A. and Pinsky, M. L. (2018). Genomic signatures of environmental selection despite near-panmixia in summer flounder. *Evol. Appl.*

Hopkins, A. E. (1937). Experimental observations on spawning, larval development, and setting in the Olympia oyster, *Ostrea lurida*. *Bulletin of the Bureau of Fisheries*, 68:437–502.

Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., von Mering, C., and Bork, P. (2017). Fast Genome-Wide functional annotation through orthology assignment by eggNOG-Mapper. *Mol. Biol. Evol.*, 34(8):2115–2122.

Hüning, A. K., Melzner, F., Thomsen, J., Gutowska, M. A., Krämer, L., Frickenhaus, S., Rosenstiel, P., Pörtner, H.-O., Philipp, E. E. R., and Lucassen, M. (2013). Impacts of seawater acidification on mantle gene expression patterns of the Baltic Sea blue mussel: implications for shell formation and energy metabolism. *Mar. Biol.*, 160(8):1845–1861.

Hurth, M. A., Suh, S. J., Kretzschmar, T., Geis, T., Bregante, M., Gambale, F., Martinoia, E., and Neuhaus, H. E. (2005). Impaired ph homeostasis in *Arabidopsis* lacking the vacuolar dicarboxylate transporter and analysis of carboxylic acid transport across the tonoplast. *Plant Physiol.*, 137(3):901–910.

Iwamoto, E. M., Elz, A. E., García-De León, F. J., Silva-Segundo, C. A., Ford, M. J., Palsson, W. A., and Gustafson, R. G. (2015). Microsatellite DNA analysis of Pacific hake *Merluccius productus* population structure in the Salish Sea. *ICES J. Mar. Sci.*, 72(9):2720–2731.

Jackson, T. M. and O'Malley, K. G. (2017). Comparing genetic connectivity among Dungeness crab (*Cancer magister*) inhabiting Puget Sound and coastal Washington. *Mar. Biol.*, 164(6):123.

Jombart, T. and Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27(21):3070–3071.

Jones, T. A. (2013). When local isn't best. *Evol. Appl.*, 6(7):1109–1118.

Kanwal, S., Khan, F. Z., Lonie, A., and Sinnott, R. O. (2017). Investigating reproducibility and tracking provenance—a genomic workflow case study. *BMC Bioinformatics*, 18(1):337.

Kawecki, T. J. and Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, 7(12):1225–1241.

Kelly, R. P. and Palumbi, S. R. (2010). Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. *PLoS One*, 5(1):e8594.

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., and Jupyter Development Team (2016). Jupyter Notebooks — a publishing framework for reproducible computational workflows. In Loizides, F. and Schmidt, B., editors, *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, pages 87–90. IOS Press.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., and Mayrose, I. (2015). CLUMPAK: A program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.*, 15(5):1179–1191.

Kuo, E. S. L. and Sanford, E. (2009). Geographic variation in the upper thermal limits of an intertidal snail: Implications for climate envelope models. *Mar. Ecol. Prog. Ser.*, 388:137–146.

Lallias, D., Taris, N., Boudry, P., Bonhomme, F., and Lapègue, S. (2010). Variance in the reproductive success of flat oyster *Ostrea edulis* assessed by parentage analyses in natural and experimental conditions. *Genet. Res.*, 92(3):175–187.

Larson, W. A., Seeb, L. W., Everett, M. V., Waples, R. K., Templin, W. D., and Seeb, J. E. (2014). Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). *Evol. Appl.*, 7(3):355–369.

Launey, S. and Hedgecock, D. (2001). High genetic load in the Pacific oyster *Crassostrea gigas*. *Genetics*, 159(1):255–265.

Lemasson, A. J., Fletcher, S., Hall-Spencer, J. M., and Knights, A. M. (2017). Linking the biological impacts of ocean acidification on oysters to changes in ecosystem services: A review. *J. Exp. Mar. Bio. Ecol.*, 492:49–62.

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends Ecol. Evol.*, 17(4):183–189.

Levin, L. A. (2006). Recent progress in understanding larval dispersal: new directions and digressions. *Integr. Comp. Biol.*, 46(3):282–297.

Li, B., Song, K., Meng, J., Li, L., and Zhang, G. (2017). Integrated application of transcriptomics and metabolomics provides insights into glycogen content regulation in the Pacific oyster *Crassostrea gigas*. *BMC Genomics*, 18(1):713.

Liao, D., Zhou, Y., Tong, J., Cao, S., Yu, X., Fu, B., and Yang, D. (2018). Characterization and phylogenetic analysis of the complete mitochondrial genome from Rock Scallop (*Crassadoma gigantea*) using next-generation sequencing. *Mitochondrial DNA Part B*, 3(2):827–828.

Liu, W., Yu, Z., Huang, X., Shi, Y., Lin, J., Zhang, H., Yi, X., and He, M. (2017). Effect of ocean acidification on growth, calcification, and gene expression in the pearl oyster, *Pinctada fucata*. *Mar. Environ. Res.*, 130:174–180.

Lohman, B. K., Weber, J. N., and Bolnick, D. I. (2016). Evaluation of TagSeq, a reliable low-cost alternative for RNAseq. *Mol. Ecol. Resour.*, 16(6):1315–1321.

Lowry, D. B., Hoban, S., Kelley, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F., and Storfer, A. (2017). Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Mol. Ecol. Resour.*, 17(2):142–152.

Luisetti, T., Turner, R. K., Jickells, T., Andrews, J., Elliott, M., Schaafsma, M., Beaumont, N., Malcolm, S., Burdon, D., Adams, C., and Watts, W. (2014). Coastal Zone Ecosystem Services: From science to values and decision making; a case study. *Sci. Total Environ.*, 493:682–693.

Luu, K., Bazin, E., and Blum, M. G. B. (2017). pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Mol. Ecol. Resour.*, 1:67–77.

Malécot, G. (1968). *The Mathematics of Heredity*. San Francisco.

Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., and Chen, W.-M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics*, 26(22):2867–2873.

Marko, P. B. (1998). Historical allopatry and the biogeography of speciation in the prosobranch snail genus *Nucella*. *Evolution*, 52(3):757–774.

Marko, P. B. (2004). 'What's larvae got to do with it?' Disparate patterns of post-glacial population structure in two benthic marine gastropods with identical dispersal potential. *Mol. Ecol.*, 13(3):597–611.

Marshall, D. J. (2008). Transgenerational plasticity in the sea: context-dependent maternal effects across the life history. *Ecology*, 89(2):418–427.

Martinez, E., Buonaccorsi, V., Hyde, J. R., and Aguilar, A. (2017). Population genomics reveals high gene flow in grass rockfish (*Sebastes rastrelliger*). *Mar. Genomics*, 33:57–63.

Matz, M. V. (2018). Fantastic beasts and how to sequence them: Ecological genomics for obscure model organisms. *Trends Genet.*, 34(2):121–132.

Matz, M. V. (2019). KOGMWU.

Maynard, A., Bible, J. M., Pespeni, M. H., Sanford, E., and Evans, T. G. (2018). Transcriptomic responses to extreme low salinity among locally adapted populations of Olympia oyster (*Ostrea lurida*). *Mol. Ecol.*, 27(21):4225–4240.

McClure, M. M., Utter, F. M., Baldwin, C., Carmichael, R. W., Hassemer, P. F., Howell, P. J., Spruell, P., Cooney, T. D., Schaller, H. A., and Petrosky, C. E. (2008). Evolutionary effects of alternative artificial propagation programs: implications for viability of endangered anadromous salmonids. *Evol. Appl.*, 1(2):356–375.

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

Milano, I., Babbucci, M., Cariani, A., Atanassova, M., Bekkevold, D., Carvalho, G. R., Espiñeira, M., Fiorentino, F., Garofalo, G., Geffen, A. J., Hansen, J. H., Helyar, S. J., Nielsen, E. E., Ogden, R., Patarnello, T., Stagioni, M., FishPopTrace Consortium, Tinti, F., and Bargelloni, L. (2014). Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (*Merluccius merluccius*). *Mol. Ecol.*, 23(1):118–135.

Miller, K. J. and Ayre, D. J. (2008). Protection of genetic diversity and maintenance of connectivity among reef corals within marine protected areas. *Conserv. Biol.*, 22(5):1245–1254.

Moll, P., Ante, M., Seitz, A., and Reda, T. (2014). QuantSeq 3' mRNA sequencing for RNA quantification. *Nat. Methods*, 11(12):1–3.

Moore, J.-S., Bourret, V., Dionne, M., Bradbury, I., O'Reilly, P., Kent, M., Chaput, G., and Bernatchez, L. (2014). Conservation genomics of anadromous Atlantic salmon across its North American range: outlier loci identify the same patterns of population structure as neutral loci. *Mol. Ecol.*, 23(23):5680–5697.

Nei, M. and Chesser, R. K. (1983). Estimation of fixation indices and gene diversities. *Ann. Hum. Genet.*, 47(3):253–259.

Novembre, J., Williams, R., Pourreza, H., Wang, Y., and Carbonetto, P. (2018). *PCAviz: Visualizing Principal Components Analysis*.

Olin, P., Culer, C., Davis, J., and Menard, K. (2012). Profitable and biosecure rock scallop culture for the West Coast. Technical report, Western Regional Aquaculture Center.

Padilla, D. K. and Savedo, M. M. (2013). A systematic review of phenotypic plasticity in marine invertebrate and plant systems. *Advances in Marine Biology*, 65:67–94.

Palmer, R. A. (1994). Temperature sensitivity, rate of development, and time to maturity: Geographic variation in laboratory-reared *Nucella* and a cross-phyletic overview. In Wilson, W H Stricker, S, editor, *Reproduction and development of marine invertebrates*, pages 177–194. Johns Hopkins University Press.

Palumbi, S. R. (1994). Genetic divergence, reproductive isolation, and marine speciation. *Annu. Rev. Ecol. Syst.*, 25(1):547–572.

Palumbi, S. R. (2003). Population genetics, demographic connectivity, and the design of marine reserves. *Ecol. Appl.*, 13(1):146–158.

Pamplona, R., Barja, G., and others (2002). Membrane fatty acid unsaturation, protection against oxidative stress, and maximum life span: a homeoviscous-longevity adaptation? *Ann. N. Y. Acad. Sci.*

Pan, P.-Y., Marrs, J., and Ryan, T. A. (2015). Vesicular glutamate transporter 1 orchestrates recruitment of other synaptic vesicle cargo proteins during synaptic vesicle recycling. *J. Biol. Chem.*, 290(37):22593–22601.

Parker, L. M., Ross, P. M., O'Connor, W. A., Pörtner, H. O., Scanes, E., and Wright, J. M. (2013). Predicting the response of molluscs to the impact of ocean acidification. *Biology*, 2(2):651–692.

Pauletto, M., Milan, M., Huvet, A., Corporeau, C., Suquet, M., Planas, J. V., Moreira, R., Figueras, A., Novoa, B., Patarnello, T., and Bargelloni, L. (2017). Transcriptomic features of *Pecten maximus* oocyte quality and maturation. *PLoS One*, 12(3):e0172805.

Pepin, P. (1991). Effect of temperature and size on development, mortality, and survival rates of the pelagic early life history stages of marine fish. *Can. J. Fish. Aquat. Sci.*, 48(3):503–518.

Perrin, N., Sibly, R. M., and Nichols, N. K. (1993). Optimal growth strategies when mortality and production rates are size-dependent. *Evol. Ecol.*, 7(6):576–592.

Pespeni, M. H., Oliver, T. A., Manier, M. K., and Palumbi, S. R. (2010). Restriction Site Tiling Analysis: accurate discovery and quantitative genotyping of genome-wide polymorphisms using nucleotide arrays. *Genome Biol.*, 11(4):R44.

Petkova, D., Novembre, J., and Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nat. Genet.*, 48(1):94–100.

Pfister, C. A., Gilbert, J. A., and Gibbons, S. M. (2014). The role of macrobiota in structuring microbial communities along rocky shores. *PeerJ*, 2:e631.

Pickrell, J. K. and Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.*, 8(11):e1002967.

Pohlert, T. (2014). The pairwise multiple comparison of mean ranks package (PMCMR).

Polson, M. P., Hewson, W. E., Eernisse, D. J., Baker, P. K., and Zacherl, D. C. (2009). You say *Conchaphila*, I say *Lurida*: Molecular evidence for restricting the Olympia oyster (*Ostrea lurida* Carpenter 1864) to temperate western North America. *J. Shellfish Res.*, 28(1, SI):11–21.

Pontarp, M., Johansson, J., Jonzén, N., and Lundberg, P. (2015). Adaptation of timing of life history traits and population dynamic responses to climate change in spatially structured populations. *Evol. Ecol.*, 29(4):565–579.

Pritchard, C., Shanks, A., Rimler, R., Oates, M., and Rumrill, S. (2015). The Olympia oyster *Ostrea lurida*: Recent advances in natural history, ecology, and restoration. *J. Shellfish Res.*, 34(2).

Puritz, J. B., Hollenbeck, C. M., and Gold, J. R. (2014). ddocent: A RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ*, 2:e431.

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., and Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, 152(5):1173–1183.

Raith, M., Zacherl, D. C., Pilgrim, E. M., and Eernisse, D. J. (2016). Phylogeny and species diversity of Gulf of California oysters (Ostreidae) inferred from mitochondrial DNA. *Am. Malacol. Bull.*, 33(2):263–283.

Reitzel, A. M., Herrera, S., Layden, M. J., Martindale, M. Q., and Shank, T. M. (2013). Going where traditional markers have not gone before: Utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. *Mol. Ecol.*, 22(11):2953–2970.

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., and Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Mol. Ecol.*, 24(17):4348–4370.

Riviere, G., Wu, G.-C., Fellous, A., Goux, D., Sourdaine, P., and Favrel, P. (2013). DNA methylation is crucial for the early development in the oyster *C. gigas*. *Mar. Biotechnol.*, 15(6):739–753.

Ross, P. M., Parker, L., and Byrne, M. (2016). Transgenerational responses of molluscs and echinoderms to changing ocean conditions. *ICES J. Mar. Sci.*, 73(3):537–549.

Ross, P. M., Parker, L., O'Connor, W. A., and Bailey, E. A. (2011). The impact of ocean acidification on reproduction, early development and settlement of marine organisms. *Water*, 3(4):1005–1030.

Salinas, S., Brown, S. C., Mangel, M., and Munch, S. B. (2013). Non-genetic inheritance and changing environments. *Non-Genetic Inheritance*, 1:38–50.

Sanford, E., Gaylord, B., Hettinger, A., Lenz, E. A., Meyer, K., and Hill, T. M. (2014). Ocean acidification increases the vulnerability of native oysters to predation by invasive snails. *Proc. Biol. Sci.*, 281(1778):20132681.

Sanford, E. and Kelly, M. W. (2011). Local adaptation in marine invertebrates. *Ann. Rev. Mar. Sci.*, 3(1):509–535.

Sanford, E. and Worth, D. J. (2010). Local adaptation along a continuous coastline: prey recruitment drives differentiation in a predatory snail. *Ecology*, 91(3):891–901.

Schmidt, P. S. and Rand, D. M. (2001). Adaptive maintenance of genetic polymorphism in an intertidal barnacle: habitat- and life-stage-specific survivorship of *Mpi* genotypes. *Evolution*, 55(7):1336–1344.

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH image to ImageJ: 25 years of image analysis. *Nat. Methods*, 9(7):671–675.

Schoch, G. C., Menge, B. A., Allison, G., Kavanaugh, M., Thompson, S. A., and A. Wood, S. (2006). Fifteen degrees of separation: Latitudinal gradients of rocky intertidal biota along the California Current. *Limnol. Oceanogr.*, 51(6):2564–2585.

Seale, E. M. and Zacherl, D. C. (2009). Seasonal settlement of Olympia oyster larvae, *Ostrea lurida* (Carpenter 1864) and its relationship to seawater temperature in two southern California estuaries. *J. Shellfish Res.*, 28(1):113–120.

Shiel, B. P., Hall, N. E., Cooke, I. R., Robinson, N. A., and Strugnell, J. M. (2017). Epipodial tentacle gene expression and predetermined resilience to summer mortality in the commercially important greenlip abalone, *Haliotis laevigata*. *Mar. Biotechnol.*, 19(2):191–205.

Siegle, M. R., Taylor, E. B., Miller, K. M., Withler, R. E., and Yamanaka, K. L. (2013). Subtle population genetic structure in yelloweye rockfish (*Sebastes ruberrimus*) is consistent with a major oceanographic division in British Columbia, Canada. *PLoS One*, 8(8):e71083.

Silliman, K. (2019). Population structure, genetic connectivity, and adaptation in the Olympia oyster (*Ostrea lurida*) along the west coast of North America. *Evol. Appl.*, 11:697.

Silliman, K. E., Bowyer, T. K., and Roberts, S. B. (2018). Consistent differences in fitness traits across multiple generations of Olympia oysters. *Sci. Rep.*, 8(1):6080.

Sokal, R. R. (1979). Testing statistical significance of geographic variation patterns. *Syst. Zool.*, 28(2):227–232.

Stapley, J., Reger, J., Feulner, P. G. D., Smadja, C., Galindo, J., Ekblom, R., Bennison, C., Ball, A. D., Beckerman, A. P., and Slate, J. (2010). Adaptation genomics: The next generation. *Trends Ecol. Evol.*, 25(12):705–712.

Stick, D. A. (2011). *Identification of optimal broodstock for Pacific Northwest oysters*. PhD thesis, Oregon State University.

Szent-Györgyi, A. G., Kalabokis, V. N., and Perreault-Micale, C. L. (1999). Regulation by molluscan myosins. In Imai, S., Ohtsuki, I., and Endo, M., editors, *Muscle Physiology and Biochemistry*, pages 55–62. Springer US, Boston, MA.

Teplitsky, C., Mills, J. A., Alho, J. S., Yarrall, J. W., and Merilä, J. (2008). Bergmann's rule and climate change revisited: disentangling environmental and genetic responses in a wild bird population. *Proceedings of the National Academy of Sciences of the United States of America*, 105(36):13492–13496.

Tepolt, C. K. and Palumbi, S. R. (2015). Transcriptome sequencing reveals both neutral and adaptive genome dynamics in a marine invader. *Mol. Ecol.*, 24(16):4145–4158.

Timmins-Schiffman, E., Coffey, W. D., Hua, W., Nunn, B. L., Dickinson, G. H., and Roberts, S. B. (2014). Shotgun proteomics reveals physiological response to ocean acidification in *Crassostrea gigas*. *BMC Genomics*, 15:951.

Tomanek, L., Zuzow, M. J., Ivanina, A. V., Beniash, E., and Sokolova, I. M. (2011). Proteomic response to elevated pCO_2 level in eastern oysters, *Crassostrea virginica*: evidence for oxidative stress. *J. Exp. Biol.*, 214(Pt 11):1836–1844.

Townsend, C. H. (1895). The transplanting of eastern oysters to Willapa Bay, Washington, with notes on the native oyster industry. Technical Report 334, U.S. Bureau of Fisheries.

Valentine, J. W. (1966). Numerical analysis of marine molluscan ranges on extratropical northeastern Pacific shelf. *Limnol. Oceanogr.*, 11:198–211.

Van Wyngaarden, M., Snelgrove, P. V. R., DiBacco, C., Hamilton, L. C., Rodríguez-Ezpeleta, N., Jeffery, N. W., Stanley, R. R. E., and Bradbury, I. R. (2017). Identifying patterns of dispersal, connectivity, and selection in the sea scallop, *Placopecten magellanicus,* using RAD-seq derived SNPs. *Evol. Appl.*, 10:102–117.

Wakeley, J. (2005). The limits of theoretical population genetics. *Genetics*, 169(1):1–7.

Waldbusser, G. G., Gray, M. W., Hales, B., Langdon, C. J., Haley, B. A., Gimenez, I., Smith, S. R., Brunner, E. L., and Hutchinson, G. (2016). Slow shell building, a possible trait for resistance to the effects of acute ocean acidification. *Limnol. Oceanogr.*, 61(6).

Waldbusser, G. G., Hales, B., Langdon, C. J., Haley, B. A., Schrader, P., Brunner, E. L., Gray, M. W., Miller, C. A., Gimenez, I., and Hutchinson, G. (2015). Ocean acidification has multiple modes of action on bivalve larvae. *PLoS One*, 10(6):e0128376.

Wang, J. (2004). Sibship reconstruction from genetic data with typing errors. *Genetics*, 166(4):1963–1979.

Wang, J. and Santure, A. W. (2009). Parentage and sibship inference from multilocus genotype data under polygamy. *Genetics*, 181(4):1579–1594.

Wang, L. (2017). Population_Genomics_Scripts. `https://github.com/wlz0726/Population\_Genomics\_Scripts/`.

Wang, M., Liu, M., Wang, B., Jiang, K., Jia, Z., Wang, L., and Wang, L. (2018). Transcriptomic analysis of exosomal shuttle mRNA in Pacific oyster *Crassostrea gigas* during bacterial stimulation. *Fish Shellfish Immunol.*, 74:540–550.

Wang, S., Meyer, E., McKay, J. K., and Matz, M. V. (2012). 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nat. Methods*, 9(8):808–810.

Wang, X., Wang, M., Jia, Z., Wang, H., Jiang, S., Chen, H., Wang, L., and Song, L. (2016). Ocean acidification stimulates alkali signal pathway: A bicarbonate sensing soluble adenylyl cyclase from oyster *Crassostrea gigas* mediates physiological changes induced by CO_2 exposure. *Aquat. Toxicol.*, 181:124–135.

Wares, J. P., Gaines, S. D., and Cunningham, C. W. (2001). A comparative study of asymmetric migration events across a marine biogeographic boundary. *Evolution*, 55(2):295–306.

Wasson, K., Hughes, B. B., Berriman, J. S., Chang, A. L., Deck, A., Dinnel, P. A., Endris, C., Espinoza, M., Dudas, S., Ferner, M. C., Grosholz, T., Kimbro, D., Ruesink, J. L., Trimble, A. C., Schaaf, D. V., Zabin, C., and Zacherl, D. C. (2016). Coast-wide recruitment dynamics of Olympia oysters reveal limited synchrony and multiple predictors of failure. *Ecology*, 97(12):3503–3516.

Weersing, K. and Toonen, R. J. (2009). Population genetics, larval dispersal, and connectivity in marine systems. *Mar. Ecol. Prog. Ser.*, 393:1–12.

Weir, B. S. and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6):1358–1370.

West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. Oxford University Press.

Whitlock, M. C. and Lotterhos, K. E. (2015). Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of F_{ST}. *Am. Nat.*, 186:S24–s36.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Willing, E.-M., Dreyer, C., and van Oosterhout, C. (2012). Estimates of genetic differentiation measured by F_{ST} do not necessarily require large sample sizes when using many SNP markers. *PLoS One*, 7(8):e42649.

Woelke, C. E. (1959). Pacific oyster *Crassostrea gigas* mortalities: With notes on common oyster predators in Washington waters. Technical report, Washington Dept. of Fisheries, Quilcene Shellfish Laboratory.

Wong, K. K. W., Lane, A. C., Leung, P. T. Y., and Thiyagarajan, V. (2011). Response of larval barnacle proteome to CO_2-driven seawater acidification. *Comp. Biochem. Physiol. Part D Genomics Proteomics*, 6(3):310–321.

Wright, R. M., Aglyamova, G. V., Meyer, E., and Matz, M. V. (2015). Gene expression associated with white syndromes in a reef building coral, *Acropora hyacinthus*. *BMC Genomics*, 16:371.

Xiao, S., Wu, X., Li, L., and Yu, Z. (2015). Complete mitochondrial genome of the Olympia oyster *Ostrea lurida* (Bivalvia, Ostreidae). *Mitochondrial DNA*, 26(3):471–472.

Zhang, Y., Sun, J., Mu, H., Li, J., Zhang, Y., Xu, F., Xiang, Z., Qian, P.-Y., Qiu, J.-W., and Yu, Z. (2015). Proteomic basis of stress responses in the gills of the Pacific oyster *Crassostrea gigas*. *J. Proteome Res.*, 14(1):304–317.

zu Ermgassen, P. S. E., Gray, M. W., Langdon, C. J., Spalding, M. D., and Brumbaugh, R. D. (2013). Quantifying the historic contribution of Olympia oysters to filtration in Pacific Coast (USA) estuaries and the implications for restoration objectives. *Aquat. Ecol.*, 47(2).