

Comparing Low Salinity Transcriptomic Profiles Among
Hard Clam, *Mercenaria mercenaria*, Lines



A Thesis

Presented to

The Faculty of the School of Marine Science

The College of William & Mary

In Partial Fulfillment

of the Requirements for the Degree of

Master of Science



by

Leslie Speight Youtsey

January 2023

APPROVAL PAGE

This thesis is submitted in partial fulfillment of
the requirements for the degree of
Master of Science

Leslie Speight Youtsey

Approved by the Committee, January 2023

Jan R. McDowell, Ph.D
Committee Chair / Advisor

Kimberly S. Reece, Ph.D
Committee Co-chair / Co-advisor

Emily B. Rivest, Ph.D

Richard A. Snyder, Ph.D

William C. Walton, Ph.D

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	v
LIST OF TABLES.....	vi
LIST OF FIGURES	vii
ABSTRACT	ix
INTRODUCTION	2
MATERIALS AND METHODS	9
Salinity Experiments.....	9
Sample Preparation and Sequencing	11
Filtering and Alignment of Sequences	12
Principal Component Analysis	14
Adult Gill & Juvenile Whole-Body Response to Low Salinity.....	15
KEGG Pathways.....	15
Clam Lines and Salinity Interaction	16
RESULTS.....	17
Sample Quality and Quantity	17
PCAs.....	18
Response to Low Salinity in Adult Gills	19
Response to Low Salinity in Whole Bodies of Juveniles.....	19

The Overlapping DEGS Between Adults and Juveniles	20
KEGG Pathways in Adults.....	20
KEGG Pathways in Juveniles.....	21
Clam Line Response to Low Salinity	22
Comparing Clam Line Responses to Low Salinity	23
DISCUSSION.....	24
The Response of Adult Hard Clam Gills to Low Salinity Stress	25
The Role of Heat Shock Proteins in the Gills of Adult Hard Clams	26
The Response of Juvenile Clams to Low Salinity Stress	29
Comparing the Response of Adults and Juveniles to Low Salinity Stress.....	31
Low Salinity Response Compared Among Adult Clam Lines.....	32
Low Salinity Response Compared Among Juvenile Clam Lines.....	34
CONCLUSION	36
NEXT STEPS.....	37
POTENTIAL LIMITATIONS OF THIS STUDY	38
TABLES.....	41
FIGURES.....	44
LITERATURE CITED.....	67
APPENDIX: SUPPLEMENTARY TABLES AND FIGURES.....	83

ACKNOWLEDGEMENTS

First and foremost, I would like to praise and thank God, the Almighty, who has granted countless blessings, knowledge, and opportunities to me, so that I have was able to accomplish this thesis.

The successful completion of this thesis depended largely on the encouragement and guidance of many others. I would like to take this opportunity to express my gratitude to the people who have been instrumental in the successful completion of this thesis. It would be impossible to list all the names, but several people deserve my sincere and special thanks. I would like to show my greatest appreciation to Dr. Jan McDowell and Dr. Kimberly Reece, as the thesis supervisors, my research advisors, and my professional mentors. I cannot say thank you enough for this opportunity and the tremendous support and help.

My deepest gratitude to Dr. Emily Rivest, Dr. Richard Snyder, and Dr. William Walton for sitting on my committee and for the valuable inputs included in this thesis and the generous time spent giving guidance, encouragement, suggestions, and comments for this thesis. My enormous appreciation to my lab mates, colleagues, and peers at VIMS for their help and wisdom along the way. A special thanks to Elizabeth Weatherup and Heather Corson for their friendships and support throughout these last two and half years.

Finally, I will forever be indebted to my beloved parents Jeffrey and Gina Speight and my in-laws Dave and Christine Youtsey, for giving me unfailing support during the completion of this thesis. Last, but certainly not least, I would like to thank my beloved husband, Emerson Youtsey, for his unwavering love and encouragement throughout this process.

LIST OF TABLES

1. Sample Quality Scores From Novogene.....	41
2. Top Ten Degs For Adult Clams	41
3. Top Ten Degs For Juvenile Clams	41

LIST OF FIGURES

1. The Lifecycle of the Hard Clam - Adapted from Ethan Nedeau.....	44
2. Experimental Set Up.....	45
3. Experimental Design	45
4. Variance Stabilizing Transformation PCA Adults	46
5. Regularized-Logarithm Transformation PCA Adults	47
6. Variance Stabilizing Transformation PCA Juveniles.....	48
7. Regularized-Logarithm Transformation PCA Juveniles	49
8. Volcano Plot for Adult Clam Lines Comparing 15 ppt vs 35 ppt	50
9. Volcano Plot for Juvenile Clam Lines Comparing 15 ppt vs 35 ppt.....	51
10. Venn Diagram of DEGs Between Adults and Juveniles	52
11. Adult Clam 15 v 35 DEGs Annotated by BlastKOALA & KEGG Database.....	53
12. Pathways With Six or More Adult 15 v 35 DEGs.....	54
13. Protein Processing in Endoplasmic Reticulum Pathway With Adult 15 v 35 DEGs..	55
14. Rap1 Signaling Pathway With Adult 15 v 35 DEGs.....	56
15. Juvenile Clam 15 v 35 DEGs Annotated by BlastKOALA & KEGG Database.....	57
16. Pathways With Four or More Juvenile 15 v 35 DEGs	58
17. Pancreatic Secretion Pathway With Juvenile 15 v 35 DEGs.....	59
18. Rap1 Signaling Pathway With Juvenile 15 v 35 DEGs.....	60
19. MA Plot for Adult Clam Lines 15 ppt vs 35 ppt	61
20. Venn Diagram of DEGs Overlap Among Adult Lines.....	62
21. MA Plot for Juvenile Clam Lines 15 ppt vs 35 ppt.....	63
22. Venn Diagram of DEGS Overlap Among Juvenile Lines.....	64

23. MA Plot for Adult Clam Lines Interaction Terms	65
24. MA Plot for Juvenile Clam Lines Interaction Terms	66

Abstract

The hard clam is an important ecological and economic resource along the U.S. Eastern Seaboard. In Virginia alone, farm gate sales were estimated at \$38.8 million in 2018 and \$57.8 million in 2021, making Virginia the largest producer of hard clams in the U.S. This industry is primarily limited to higher salinity habitats on the seaside of the Eastern Shore of Virginia and lower Chesapeake Bay. Although the hard clam can be found in lower salinity habitats, they do not grow or survive at rates that are practical for productive aquaculture. Even in areas of higher salinity, hard clams are vulnerable to extreme precipitation events, which can lead to hyposaline stress and threaten natural and aquacultured populations. Transcriptomic analysis is a powerful tool for exploring the altered gene expression and cell cycle events that occur under osmotic stress. In the spring of 2019 and 2021, clam lines were created at the VIMS Eastern Shore Laboratory. Salinity exposures were conducted in the summer of 2021 with eight clam lines at four different salinities (35, 20, 15, and 12 ppt). RNA sequencing (RNA-Seq) data from either the gill (2019) or pooled whole bodies (2021) of exposed clams at 35 and 15 ppt were used to assess the transcriptomic response to low salinity stress. This study found 545 genes in the gills of adult hard clams and 465 genes in the whole bodies of juvenile hard clams that were significantly differentially expressed between 15 and 35 ppt. Some of the top genes by significance and log₂ fold change were those in the categories of heat shock proteins, apoptosis, and cellular polarity. The amino acid sequences of these differentially expressed genes (DEGs) were assigned to key pathways like protein processing in the endoplasmic reticulum and Rap1 signaling. The response to 15 ppt was compared among clam lines, resulting in some interesting expression profiles. Some clam lines from the same population had a large number of genes differentially expressed in response to low salinity (15ppt), while other clam lines derived from different populations showed few differences. Some of the genes differentially expressed by different clam lines included heat shock proteins, inorganic ion regulators, and free amino acid isomerase enzymes. The observed difference between clam lines could indicate different tolerance to low salinity and different adapted molecular approaches to combat osmotic stress. This study identified key genes and pathways that could be the focus of future studies. Results from this study allow for a better understanding of how hard clams respond to low salinity stress and begin to lead to a determination of whether all populations of clams respond the same, which could benefit the aquaculture industry and lead to strategic breeding programs for enhanced low salinity tolerance in hard clams.

Comparing Low Salinity Transcriptomic Profiles Among Hard Clam, *Mercenaria mercenaria*,
Lines

Introduction

The hard clam, *Mercenaria mercenaria* (Linnaeus, 1758), is an infaunal species that plays a key ecological role as a filter feeder and prey source for crabs, fishes, and rays, and is an important economic resource along the U.S. Eastern Seaboard. Its life history (Figure 1) has allowed it to persist since the Upper Miocene, 5-11 Mya, through multiple major climate change events including temperature and sea level fluctuations (Palmer, 1927). The hard clam exhibits consecutive hermaphroditism, where it can change sex once in its lifetime, but is strictly dioecious (Loosanoff 1936, 1937; Dalton & Menzel 1983; Eversole, 1989; Otto 1973). About 98% of all hard clams experience a juvenile male sexual phase after reaching 6-7 mm in shell length (SL). Half of these juvenile males will undergo a protandric sexual change into females at maturity. The other 2% of juveniles develop directly into females without the juvenile male phase (Loosanoff 1936, 1937). The hard clam reaches maturity around 20 - 35 mm SL, which usually occurs between 1 to 3 years of age depending on latitudinal position. (Belding, 1931; Loosanoff, 1936; Bricelj and Malouf, 1980; Eversole et al., 1980; Stanley and DeWitt, 1983; Knaub and Eversole, 1988; Hesselman et al., 1989; Walker, 1994; Hadley, 1997). There is no sexual dimorphism and what triggers the change from male to female is still not fully understood, but after reaching maturity wild populations have an approximately 1:1 male to female ratio (Eversole et al., 1980).

Adult clams produce gametes throughout their life with no apparent decrease in fecundity with age and are thought to have an average life span of 12-20 years (Peterson, 1983; Peterson, 1986). Spawning patterns vary among populations and latitudes. Belding (1931) suggested that the reproductive cycle of the hard clam is timed with favorable environmental conditions for larval survival like temperature, food availability, and salinity. The hard clam inhabits estuarine and coastal marine environments within its native geographic range from the Gulf of St. Lawrence to

the Gulf of Mexico, exhibiting an ability to survive in a wide range of habitats (Palmer, 1927; Roegner & Mann, 1990; Harte, 2001). Hard clam aquaculture can be found throughout the species' native range and has been introduced in other areas of the globe (Heppell, 1961; Hanna, 1966; Harte, 2001; Chew, 2001). In an aquaculture setting, hatcheries will extend the spawning period by collecting mid-sized adult clams from natural stocks or planted beds in January or February (within mid-latitudes of the Northern Hemisphere) and place them in a tank of seawater between 20-23°C. As ambient water temperature and food levels rise, clam gonads ripen naturally (Castagna & Kraeuter, 1981; Castagna, 2001). By manipulating the gonadal cycle and spawning period of broodstock, producers optimize use of larval and nursery systems and get an early start on seed production (Castagna, 2001). Broodstock may be selected based on desired traits such as fast growth, shell markings, and environmental adaptability (Chanley, 1961). Selecting adult clams based on sex is not possible unless clams have been previously spawned, and their sex identified. Alternatively, taking gonad samples prior to spawning is possible, but usually not worth the effort since adult populations have roughly equal numbers of both sexes, and hatcheries utilize mass spawning (Castagna, 2001). Clams with ripened gonads will usually spawn when subjected to a temperature shock, temperature cycling (heating and cooling of seawater), or stimuli (serotonin (5HT) or sperm) (Castagna, 2001). It is estimated that only 15-20% of competent individuals spawn in each spawning event (Ansell, 1967; Bricelj & Malouf, 1980). Individual females require several spawning events over several weeks to release all their ova (Loosanoff 1937; Davis & Chanley, 1956).

Eggs are fertilized externally in the water and will develop into embryos in 4-8 hrs. at 24°C. Within the first 16 hours post fertilization, the ovum develops into a blastula, then gastrula, then trochophore and to a shelled veliger stage. A shelled, straight-hinged veliger stage occurs for 1 to

3 days post fertilization. From days 3 to 20 post fertilization, a umbonal veliger stage finishes the planktonic stage (Carriker, 2001). At an age varying from 6 to 20 days post fertilization, veligers develop a foot and begin the swimming-crawling stage termed pediveliger. When the velum is lost and shell length reaches 200-300 μm , pediveliger larvae or plantigrades are limited to crawling on the bottom (Carriker, 2001). At this stage, most culturists move larvae harvested on a 130 or larger mesh size sieve to the nursery. In the wild, byssal plantigrades alternate between attachment to substrate and or crawling along fine sediments for several weeks until reaching 9 mm shell length (SL). After 9 mm SL, juvenile plantigrades remain beneath the sediment (Carriker, 2001). While recruitment to the benthos can occur in as little as 20 days, the final planktonic stage of umbonal veliger can last for weeks. In the hatchery, the period with highest observed mortality in clam seed production occurs at the time larvae metamorphose. Once metamorphosis or settlement (the full development of the siphon and transition into a juvenile) has occurred, clams must be grown to a larger size before they can be planted into beds within the bay or lagoon (Castagna, 2001). Clam seed smaller than 8 to 10 mm in shell height (SH) can experience high predation even under predator exclusion nets. Most producers will wait till seed reaches 10-15 mm or 12 mm average SH before field planting.

Throughout the process of spawning and raising clam seed, good water quality is a requirement, including appropriate temperature, dissolved oxygen, and salinity. Most hatchery systems rely on the water pumped from a direct source nearby. Post-settlement or juvenile clams are much more tolerant than embryos and larvae but can still be lost due to a lack of dissolved oxygen (DO), toxins, poor food, overcrowding, and rapid salinity changes (Castagna, 2001). Short-term exposures to low DO, low and/or changing salinity, or other poor water quality conditions are usually tolerated, but the length and timing of some suboptimal conditions are out of the control

of the hatchery. As the hard clam grows, its ability to tolerate stressors, like low salinity, increases. Hard clam larvae have a range of salinity tolerance between 20 and 35 parts per thousand (ppt) (Roegner & Mann, 1990). Juvenile clams can survive salinities as low as 15 ppt (Roegner & Mann, 1990; Baker, et al., 2005). Adult hard clams can tolerate a salinity range from 12 to 45 ppt; however, reproduction is inhibited below 15 ppt, and the optimal salinity for survival and growth is around 26-27 ppt (Davis, 1958; Loosanoff & Davis, 1963; Davis & Calabrese, 1964; Castagna & Chanley 1973, Roegner & Mann 1990). Hard clams are not abundant in waters below 18 ppt and are restricted to salinities above 12 ppt (Roegner & Mann, 1990). Survival at acute or low salinities has negative consequences for overall health, growth rate and reproduction of marine bivalves and although adult hard clams have been known to survive at 12 ppt, they do not grow (Carregosa et al., 2014; Roegner & Mann, 1990).

Assessment of how the hard clam responds to salinity fluctuations is an existing need. Within the Chesapeake Bay, wild hard clams can be found as far north as Pocomoke and Tangier Sounds, but the hard clam aquaculture industry is primarily limited to higher salinity habitats on the seaside of the Eastern Shore and lower Bay where clams grow at an economically viable rate. Even in areas of higher salinity, hard clams are vulnerable to extreme precipitation events, which can lead to hyposaline (low salinity) stress and threaten natural and aquacultured hard clam populations. The effects of climate change are predicted to bring an increase in tropical storms and hurricanes to the U.S. east coast, which can cause rapid decreases in salinity, destabilizing coastal ecosystems (Karl & Trenberth, 2003; Durack et al., 2012). It is unknown how these predicted climatological changes will affect the hard clam and impact the aquaculture industry, but steps can be taken now to investigate the potential impacts and prepare for increased precipitation events, including breeding hard clams to survive and grow to market size in low salinity.

Osmotic stress can occur when a cell or organism experiences a sudden change in solute concentration, like a drop in salinity, and can lead to altered metabolic processes, gene expression, and cell cycle events (Ktitz, 2000). To protect against osmotic stress, hard clams may immediately close their valves when salinity fluctuates outside their optimal range (Shumway, 1977). However, coastal bivalve species that live in environments that experience tidal and salinity shifts are considered osmoconformers, meaning their internal medium fluctuates according to the osmolality of the environment (McClary, 2014). If the shell valve remains open, the hemolymph and pallial fluid in the hard clam is isosmotic with external seawater medium (Anderson and Prosser, 1953). The gills of marine bivalves are important organs as they regulate oxygen consumption and food collection simultaneously. Salinity affects the pumping rate of the gills of hard clams, and rates can fall from 8 L h^{-1} at optimal salinities down to $< 1 \text{ L h}^{-1}$ at 15 ppt (Hamwi, 1969). Lower growth rates and mortality under low salinities are consequences of reduced pumping rates and the clam's inability to collect enough calories to support metabolic functions. However, the gill pump should not be viewed at either being on or off, but as a complex organ with many control mechanisms (Bayne, 1993).

Osmoconformers are mostly stenohaline and are restricted to more marine waters, so when they do experience hyposaline conditions, they rely on isosmotic intracellular regulation (IIR) to maintain cell volume and do not regulate extracellular fluid (Florkin, 1962; Rivera-Ingraham & Lignot, 2017). Under a hypotonic state, water will diffuse across the cell membrane and into the cell to dilute the intracellular osmolality leading to the expansion and lysis of cells. To prevent this from happening, osmoconformers regulate cell volume by decreasing membrane permeability to water, changing the concentration of osmotic effectors such as inorganic ions and free amino acids (Pierce, 1982), or changing the expression of membrane-bound transporters (Gilles, 1987). Small

inorganic ions such as potassium, sodium, and chloride make up a major portion of solutes in the cells and body fluids of organism and their distribution is critical for osmotic homeostasis (Dawson & Liu, 2008). However, because ion concentrations are key for proper cellular function, non-essential free amino acids (FAA) are transported across cell membranes to prevent ion depletion or augmentation (Grizzle et al., 2001). DuPaul and Webb (1974) observed a rapid initial increase in the concentration of FAA taurine and a slower delayed increase in the FAA alanine in the gills when hard clams were subjected to a sudden increase in salinity. Shumway et al. (1977) observed an increase in FAA concentrations in the adductor muscle of the hard clam during abrupt and gradual salinity decreases, in contrast to the decrease expected if FAAs were being transported to the hemolymph for intracellular isosmotic regulation. Estuarine species, such as *M. mercenaria*, likely experience fluctuating salinities, rather than prolonged reductions, and in response to experiencing a drop in salinity below 50% full-strength seawater, will typically close their valves and regulate hemolymph osmotic pressure (Shumway et al., 1977).

The regulation of cellular process and signaling is primarily controlled by proteins or polypeptides, which are products of protein coding messenger RNA (mRNA) transcribed from genes. mRNA is part of the transcriptome, which also includes non-coding RNA [ncRNA: ribosomal RNA (rRNA), transfer RNA (tRNA), and other ncRNAs] (Qian, et al., 2014). Transcriptomics is the study of transcripts in a specific cell, tissue, or organism for a given developmental stage, physiological condition, or external environment. Transcriptomic analysis is a powerful tool for exploring the relationship between phenotype and genotype, enabling a better understanding of how an organism responds to a specific condition. Transcriptomic studies are not limited to the quantification of expression level change for each gene among different samples but also include the mapping and annotation of the transcriptome and the determination of the

functional structure of each gene in the genome (Qian, et al., 2014). Due to the advances in next generation sequencing (NGS) technologies, whole transcriptome shotgun sequencing (WTSS), also known as RNA sequencing (RNA-Seq), is possible. Along with identifying differential gene expression and functional pathways, RNA-Seq analysis can be used to identify genotypic differences in the form of single nucleotide polymorphisms (SNPs) (Qian, et al., 2014; Quinn, et al., 2013; Salem, et al., 2012). SNPs directly associated with a gene, or the expression of a gene, can be powerful predictive molecular markers (Quinn, et al., 2013; Salem, et al., 2012).

SNPs have been used to examine the population structure of the hard clam. A study by Ropp et al. (under review), identified six genetically distinct populations from 15 different sampling locations: Canada, Maine, Massachusetts, Mid-Atlantic, Chesapeake Bay, and the Carolinas. Although population structure was found between high and low salinity environments (Mid-Atlantic vs. Chesapeake Bay), whether this was a function of salinity is unknown. Starting in 2019, clams from these populations were secured to create unique broodstock lines. Some of these lines were created from parent clams from Mobjack Bay and Pocomoke Sound, Virginia, locations within the Chesapeake Bay that experience salinities on the low end of what the hard clam can tolerate (20-12 ppt). Clam lines from the Chesapeake Bay, Mid-Atlantic, and Massachusetts showed differences in respiration performance after 48 hours of exposure to salinities 31, 20, 15, and 10 psu. After the 15 psu exposure, juvenile clams from the Mobjack line had oxygen consumption rates measured at $\sim 90 \mu\text{mol O}_2 \text{ h}^{-1} \text{ g dry weight}^{-1}$, compared to juveniles from the Great Bay, NJ, and Cape Cod, MA, lines, which had oxygen consumption rates measured at $\sim 50 \mu\text{mol O}_2 \text{ h}^{-1} \text{ g dry weight}^{-1}$ (Himes & Rivest, unpublished).

A logical next step for further exploring the genetic and performance differences observed among hard clams in previous studies is a comparative transcriptomic study. An in-depth look into

the cellular response via gene expression of different genetic populations and different life history stages to low salinity could give insight into the mechanisms used by the hard clam to combat osmotic stress and if these groups respond differently. The first objective of this study was to identify key genes expressed in the gill of adult and whole bodies of juvenile hard clams and identify pathways involved in the response to low salinity stress. The second objective of this study was to determine if the response to low salinity differed among clam lines descended from genetically distinct eastern seaboard populations. It is hypothesized that clam lines derived from the same genetic population will have the most similar responses or the most shared differentially expressed genes (DEGs) in response to 15 ppt exposure.

Materials and Methods

Salinity Experiments

The clam lines used in salinity challenge experiments were produced at the VIMS Eastern Shore Laboratory (ESL) from broodstock groups collected from six geographic locations and included four genetically distinct populations: 1) Cape Cod, Massachusetts (CC), 2) Wachapreague Channel, Virginia (WC) and Great Bay, New Jersey (NJ), 3) Pocomoke Sound, Virginia (PS) and Mobjack Bay, Virginia (MB), and 4) Bogue Sound, North Carolina (NC). The MB and PS locations and WC and NJ locations were not found to be significantly genetically different within each pair (Ropp et al., under review). In April of 2019, F1 generations of WC, PS, MB, NJ, and CC were made, and F1 individuals will be referred to as “adults” since their age and size indicate they were at or near sexual maturity at the time of these experiments (Supp Table 1). In April of 2021, F1 lines were made (WC x WC, WC x PS, and NC x CC), and the resulting F1 individuals will hereafter be referred to as “juveniles” (Supp Fig. 1). These crosses were created to study how the genetic combination of populations from high and low salinity

environments (WC x PS), and warm and cold-water environments (NC x CC) respond to different stressors as part of a previously funded NOAA NMFS Saltonstall-Kennedy (SK) award to R. Snyder et al., VIMS.

In June of 2021, individuals from each of the five F1 clam lines produced in 2019 (adults) were taken from ESL clam beds in Wachapreague, Virginia, placed in bags of ~50 animals/bag, labeled, and housed in a floating cage off the ESL dock until salinity experiments began, where salinities range from 25-35 ppt in Wachapreague channel. Experimental salinity conditions included 35 ppt, 20 ppt, 15 ppt, and 12 ppt and were chosen based on the salinity tolerance range of adult (12 – 45 ppt) and juvenile (15 – 35 ppt) clams (Roegner & Mann, 1990). For each salinity condition, four clams (16 in total/bag) were randomly selected from each of the five bags and held in five cups containing 900 ml of 35 ppt filtered UV seawater for 1-6 hrs. and water quality was monitored before each experiment began. The four adult clams from each line were then placed into cups containing 900 ml of the experimental salinity condition UV filtered seawater (20 cups total). Each cup was placed in a water table of circulating, shallow fresh water that regulated temperature at 25 °C and given its own air stone, and water was changed every 12 hrs. (Figures 2 and 3). After 26 hrs. of exposure to each salinity condition, three clams were randomly selected out of the four from each cup, and shell height was measured using a caliper. No mortality was observed among the clams used in the experiments. A total of 30 mg of gill tissue was immediately harvested from each sampled clam and preserved separately in 300 µl of RNAprotect Reagent (QIAGEN, Valencia, CA). All samples were stored at -20 °C at the ESL, transported in a cooler back to the VIMS Gloucester Point campus and stored at -80 °C until RNA isolation.

In June of 2021, ~70-day old juvenile clams produced in 2021 were taken from upwellers (~32 ppt), placed in 35 ppt UV filtered seawater and cleaned. Four replicates of ten clams were made for all three lines for each salinity treatment (48 sets in total). Experimental salinity conditions included 35 ppt, 20 ppt, 15 ppt, and 12 ppt. Each clam in each replicate was imaged and measured using ImageJ software (Schneider, et al., 2012). For each clam line and salinity treatment, each replicate was placed into a PVC pipe with a screened bottom (48 pipes in total). A PVC container for each clam line was then placed in a tank containing 6400 ml of the experimental salinity condition with an air stone. There were four tanks and three PVC containers per tank for each salinity treatment. Treatment tanks were placed in a water table of circulating, shallow fresh water that regulated temperature at 25 °C (Figures 2 and 3). After 26 hrs of exposure, four clams were randomly selected out of the ten clams for each line x treatment x replicate. No mortality was observed among the clams used in the experiments. Due to their small size, gill tissue alone could not be selected from the juvenile clams, so all soft tissue was pooled from four clams to reach the minimum amount of tissue needed for RNA extraction (30mg). Pooled tissue (30mg) was preserved in 300 µl of RNeasy Protect Reagent (QIAGEN, Valencia, CA). All samples were stored at -20 °C at the ESL and transported in a cooler back to VIMS at Gloucester Point and stored at -80 °C.

Sample Preparation and Sequencing

Only 35 ppt and 15 ppt samples were selected for RNA sequencing. The control salinity of 35 ppt was selected based on the salinity (25-35ppt) at which the clams had been raised (experienced) in Wachapreague, Virginia. The experimental low salinity of 15 ppt was selected based on the growth and survival limits of adult and juvenile clams. A total of 48 samples were selected for RNA-Seq and included replicates (n=3) for each salinity (n=2) and each adult line

(n=5) and juvenile line (n=3) (Supp Tables 3 & 4). Tissue samples were placed in 2 ml RNase Free PowerBead Tubes with 1.4 mm ceramic beads and 350 μ l of Buffer RTL (QIAGEN, Valencia, CA). Tissue samples were lysed and homogenized using the FastPrep-24 set to 4.0 m/s for 2 rounds of 20 seconds (MP Biomedicals, Irvine, CA). Total RNA was extracted from the homogenate using the RNeasy extraction kit following the recommended protocol (QIAGEN, Valencia, CA). Thermo Scientific DNase I, RNase-free, was used to eliminate genomic DNA from the total RNA samples (Thermo Fisher Scientific, Waltham, MA). RNA purity was checked by assessing A260/280 and A260/230 values using a NanoDropTM 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA), and total RNA concentration was measured with the QubitTM 4.0 Fluorometer (Invitrogen, Waltham, MA). Approximately 30 μ l of total RNA of each sample was sent to Novogene Corporation Inc. (Sacramento, CA) for library preparation and sequencing. Upon arrival, a quality check (QC) was performed on the samples to check RNA concentration and integrity. Samples then underwent library preparations by mRNA enrichment or rRNA removal using poly-A tail capturing methods, mRNA fragmentation, first and second strand cDNA synthesis or reverse transcription, adapter ligation to each fragment, and PCR enrichment. Samples were sequenced on an Illumina NovaSeq PE150 (6 G raw data per sample).

Filtering and Alignment of Sequences

Raw RNA-Seq data was received from Novogene, transferred to a William & Mary high performance computing (HPC) account, and uploaded to the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA). The quality of the sequences in each sample was checked using FastQC v0.11.9 (Andrews, 2010) and summarized using MultiQC v1.4 (Ewels et al., 2016). Adapters and poly-A tails were trimmed using Trimmomatic v0.36 with specifications to use truseq3-pe-2.fa adapter sequences, remove the first 10 nucleotides of each

read, remove the first and last nucleotides of a read if they fell below a quality of 3, remove reads less than 36 base pairs long, and scan the read with a 4-base wide sliding window, cutting when the average quality per base dropped below 25 (`illuminaclip = truseq3-pe-2.fa:2:30:10:2:true`, `headcrop = 10`, `leading = 3`, `trailing = 3`, `minlen = 36`, `slidingwindow = 4:25`) (Bolger et al., 2014). After trimming, the quality was checked again using FastQC (Andrews, 2010) and summarized using MultiQC (Ewels et al., 2016). The hard clam reference genome and gene annotation files from Song et al. (2021) were downloaded from [NCBI](#). The genome and annotation files were indexed using the `faidx` command in Samtools v1.3.1 (Li et al., 2009). The resulting General Feature Format (GFF) annotation file was converted to a Gene Transfer Format (GTF) annotation file using AGAT v0.8.0 (Dainat, 2020). The genome was indexed, and sequenced reads (fastq files) were mapped to the indexed reference genome using Hisat2 v2.2.1 (Kim et al., 2019). Alignment quality was assessed using Qualimap v2.2.1 (García-Alcalde et al., 2012) and the `stats` command in Samtools v1.15.1 (Danecek, 2021) and summarized using MultiQC (Ewels et al., 2016). To count the number of reads mapped to each feature or gene using the sequenced read alignments from Hisat2 v2.2.1 and the gene annotation GFF file from [NCBI](#), HTSeq (Putri et al., 2021) was implemented using Anaconda v2021.11 (Anaconda Software Distribution, 2021). Read fragments were assigned to the exons of a gene, and because whole mRNA/genes are of interest for this study, exons with the same gene accession or locus tag were combined, and the counts were summed. For example, the structure of the gene IDs from the GFF file were “`exon-XM_045310248.1-1`”, “`exon-XM_045310248.1-2`”, and “`exon-XM_045310248.1-3`”. All counts belonging to exons with gene accession number “`XM_045310248.1`” were summed and assigned to this gene accession.

Principal Component Analysis

Revised count data from HTSeq was imported using the R package: DESeq2 v1.36.0 (Love et al., 2014) as implemented in Bioconductor v3.15. After importing read count data, adult and juvenile data were separated for analysis. Throughout data analysis, the R package tidyverse v1.3.2 was used (Wickham et al., 2019). Features/genes that had less than 10 counts across samples were excluded. For both adult and juvenile clam samples, separate variance stabilizing transformation (VST) (Anders and Huber 2010), regularized-logarithm transformation (RLT) (Love et al., 2014), and generalized linear model principal component analysis (PCAs) (GLMPCA) (Townes et al., 2019) were conducted using the following programs implemented in R: DESeq2 v1.36.0 (Love et al., 2014), glmPCA v0.2.0 (Townes et al., 2019) and ggplot2 v3.3.6 (Slowikowski, 2021). All three data transformations were used to assess if different transformations altered PCA outputs and to identify any outliers. RLT accounts for variation in sequencing depth and does not use covariate (treatment vs control) information, treating all samples equally. This contrasts with VST, which is effective at stabilizing variance but does not consider differences in size factors. However, with RLT, the ordering of genes within a sample will change if neighboring genes undergo a different shrinkage strength (Love et al., 2014). GLMPCA is a generalization of PCA to exponential family likelihoods and operates on raw counts, avoiding the pitfalls of normalization, but is dependent on an iterative algorithm and is ten times slower than PCA (Townes et al., 2019). One sample, 15_WC_G_3 (the third replicate for the 15 ppt WC experiment), displayed abnormal results within FastQC, Qualimap, Samtools, and GLMPCA outputs (but not VST or RLT) and was removed from all subsequent analysis.

Adult Gill & Juvenile Whole-Body Response to Low Salinity

DESeq2 was used to identify genes that were differentially expressed between the 15 and 35 ppt salinity experiments in the gill tissues of adults and tissues from the whole bodies of juveniles. DESeq2 uses a shrinkage estimator and a negative binomial distribution to calculate dispersion and fold change. The shrinkage estimator `apegglm` v1.18.0 (Zhu et al., 2018) in R was used for all analyses. `Apegglm` is a recommended package that takes genes with high dispersion or low counts and shrinks the log₂ fold change closer to 0. Differentially expressed genes (DEGs) were defined as those transcripts with an absolute log₂ fold change (LFC) > 1 and a false discovery rate (FDR) < 0.05 (Benjamini and Hochberg, 1995). The first two DESeq2 models compared the gill tissues of adult clams at 15 ppt and 35 ppt (control) and the whole bodies of juvenile clams at 15 ppt and 35 ppt (control). For both adults and juveniles, all clam lines and samples were included (except 15_Wach_G_3), but clam line as a variable was not included in the two models (`design = ~ salinity`). To visualize those genes with an absolute LFC > 1 and FDR < 0.05, volcano plots were created using `EnhancedVolcano` v1.14.0 (Blighe et al., 2022) in R. A Venn diagram was created to compare overlaps in DEGs among the gills of adult clams and the whole body of juveniles using `VennDiagram` v1.7.3 (Chen, 2022) in R.

KEGG Pathways

To identify possible pathways that the products of DEGs could be involved in, protein sequences translated from each gene that started with "XM_" and had an absolute LFC > 1 and FDR < 0.05 were obtained in fasta format using gene accession numbers and `rentrez` v1.2.3 (Winter, 2017) in R. Protein sequences were then uploaded to BlastKOALA and Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology (KO) identifiers or K numbers were assigned to genes encoding those proteins (Kanehisa et al., 2016). K numbers were then used to

identify KEGG pathways that gene products may belong to using KEGG Mapper- Reconstruct (Kanehisa et al., 2021). Different pathways and LFC visualizations were created using pathview v1.36.0 (Luo & Brouwer, 2013) in R. Bar plots of KEGG Pathways with DEGs were created using ggplot v3.3.6 (Slowikowski, 2021) in R.

Clam Lines and Salinity Interaction

To investigate if clam lines responded differently to low salinity stress (15 ppt), an interaction term was used in the DESeq2 models, where the design = ~ 'Clam Line' + 'Salinity' + 'Clam Line x Salinity'. Adult and juvenile samples were analyzed separately, and all clam lines and samples were included (except 15_Wach_G_3). Features/genes that had less than 30 counts across adult samples and those that had less than 10 counts across juvenile samples were excluded. For adult samples, five DESeq2 objects were created, with each of the five clam lines used as the reference line once and five models were run. For each model, the main effect of salinity on the reference clam line (Salinity_15_v_35) was used with the shrinkage estimator Apeglm and the thresholds of an absolute LFC > 1 and FDR < 0.05 to estimate the number of DEGs between 15 and 35 ppt within each clam line. Scatter plots (MA plots) of log₂ mean expression or normalized counts versus log₂ fold change for each clam line were created using ggpubr v4.2.1 (Kassambara, 2020) in R. A Venn diagram was created to compare overlaps in DEGs among clam lines using VennDiagram v1.7.3 (Chen, 2022) in R.

The interaction terms produced by each of the five models were used to test if the condition effect of 15 ppt was different between each adult clam line and the adult reference clam line in each model. The shrinkage estimator Apeglm and the thresholds of an absolute LFC > 1 and FDR < 0.05 were used to estimate the number of DEGs. MA plots were created for each interaction term using ggpubr v4.2.1 (Kassambara, 2020) in R. Since, for example, Pocomoke vs. New Jersey

and New Jersey vs. Pocomoke, where the first clam line is the reference, produced slightly different results, all possible combinations of adult clam lines were compared and a total of 20 MA plots were created. Juvenile samples were analyzed the same as above, where only three, instead of five, DESeq2 models were created.

Results

Sample Quality and Quantity

Prior to being sent to Novogene for sequencing, quality and quantity assessments were performed. Samples had a mean A260/280 of 2.08 with a range from 2.00 to 2.22, and a mean A260/230 of 1.24 with a range from 0.43 to 2.23. The mean RNA concentration was 88.33 ng/ μ l with a range from 57.7 to 118.5 ng/ μ l. All samples were within RNA concentration and A260/280 limits set by Novogene for sequencing. Novogene performed quality and quantity assessment upon receiving the samples and reported a mean RNA concentration of 112.9 ng/ μ l with a range from 74.13 to 173.89 ng/ μ l. The mean RNA integrity number (RIN) was 7.48 with a range from 3.4 to 9.1. All samples were of sufficient quality and quantity for library preparation and sequencing. After sequencing, Novogene reported the mean number and standard deviation of raw reads (for paired-end sequencing, it equals the sum of read1 and read2) to be $4.76E+07 \pm 5.69E+06$, the mean size and standard deviation of raw data ((raw reads) * (sequence length (150 bp), calculated in gigabases) to be 7.15 ± 0.85 Gb, the total size of raw data to be 340 Gb, the mean effective percent ((clean reads/ raw reads)*100%) to be 97.95 ± 0.44 , the mean error percent (base error rate * 100%) to be 0.0290 ± 0.0031 , the mean Q20 and Q30 percent ((base count phred > 20 or 30 / total base count) * 100%) to be 97.90 ± 0.12 and 93.89 ± 0.31 respectively, and the mean GC content ((G & C base count / total base count)* 100%) to be 38.26 ± 1.42 (Table 1).

Quality control checks of the raw sequence reads from Novogene, indicated that all samples contained traces of Illumina universal adapter sequences. After sequences were filtered and trimmed to remove adaptors and retain only high-quality sequence fragments, no samples had > 0.1% adapter sequence. After reads were aligned to the hard clam genome, alignment qualities were summarized. One sample, Wachapreague replicate 3 from the 15 ppt experiment (Wach_3_G_15 or A_3_15), was a consistent outlier in all quality checks. This sample had the highest GC content (46%), the second highest duplication of reads (77-79.2%), the lowest overrepresented sequences (0%), the highest 5'-3' bias (7.18), the highest mapping percent (85.3%), and the highest percent of reads mapped to exonic regions (88.2%) and was removed from all subsequent analysis. After this sample was removed, the average GC content was 37% with a range from 39% to 36%, the average read duplication was 60.79% with a range from 90% to 53%, and the average number of sequences per sample was 23.6 million which ranged from 30.60 to 19.40. From the alignment summaries, the average 5'-3' bias was 1.40 with a range from 1.60 to 1.24 and the average reads aligned per sample was 38.08 million which ranged from 51.20 to 30.80. The average error rate was 0.60% with a range from 0.67% to 0.53%, and the average mapped percent was 80.76% with a range from 85.30% to 77.60%.

PCAs

PCAs are useful tools for identifying the main axis of variance and spotting patterns and outliers among samples. In the variance stabilizing transformation (VST) PCA based on the adult clam samples, 12% of the variance was explained by PC1 and 10% of the variance was explained by PC2. (Figure 4). Similarly, in the regularized-logarithm transformation (RLT) PCA, 12% of the variance was explained by PC1 and 11% of the variance was explained by PC2 (Figure 5). In the VST PCA for the juvenile clam samples, 28% of the variance was explained by PC1 and 18% of

the variance was explained by PC2 (Figure 6). In the RLT PCA for juvenile clam samples, 28% of the variance was explained by PC1 and 18% of the variance was explained by PC2 (Figure 7). The VST and RLT PCA clustered the WCxWC and WCxPS juvenile samples differently. For the VST PCA, samples grouped by salinity treatment first, then by clam line, where in the RLT PCA samples grouped by clam line first, then by salinity treatment.

Response to Low Salinity in Adult Gills

In total, 545 genes were found to be significantly differentially expressed between the 15 and 35 ppt treatments within the gills of adult hard clams. Of the 545 DEGs, 408 were up-regulated and 137 were down-regulated in the 15ppt treatment (Figure 8). The top ten genes by significance, or those ten with the lowest FDR that were characterized, included alpha-crystallin A chain-like with an LFC of 2.38, phosphoenolpyruvate carboxykinase, cytosolic [GTP]-like, transcript variant X2 with an LFC of 5.22, heat shock protein beta-1-like with an LFC of 2.00, marginal zone B- and B1-cell-specific protein-like with an LFC of 2.04, serine-arginine protein 55-like with an LFC of -1.42, heat shock protein HSP 90-beta-like with an LFC of 1.75, steroid 17-alpha-hydroxylase/17,20 lyase-like with a n LFC of -5.36, protein disulfide-isomerase A3-like with an LFC of 1.59; inositol-3-phosphate synthase 1-A-like with an LFC of 4.24, and calreticulin-like with an LFC of 1.93 (Table 2).

Response to Low Salinity in Whole Bodies of Juveniles

Overall, 465 genes were significantly differentially expressed between the 15 and 35 ppt treatments within the whole bodies of juvenile hard clams. Of those, 303 were up-regulated and 162 were down-regulated in the 15 ppt treatment (Figure 9). The top ten characterized genes by significance included sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1-like with an LFC of 5.05, sodium-dependent phosphate transport protein 2B-like with an

LFC of 3.74, solute carrier family 12 member 3-like with an LFC of 2.75, protein PIF-like with an LFC of 4.97, homeobox protein engrailed-2b-like with an LFC of 4.37, immediate early response gene 5-like protein with an LFC of 1.62, DBH-like monooxygenase protein 2 homolog with an LFC of 2.16, retinol dehydrogenase 7-like with an LFC of -2.16, protein BTG2-like with an LFC of 1.04, and bypass of stop codon protein 1-like with an LFC of -4.44 (Table 3).

The Overlapping DEGS Between Adults and Juveniles

When DEGs from the gills of adult clams (n=545) and the whole bodies of juveniles (n=465) were compared, 71 genes were found to be differentially expressed in both experiments (Figure 10). Of these, 55 were up-regulated in 15 ppt and 15 were down-regulated in 15 ppt in both adults and juveniles. Only 1 gene was up-regulated in juveniles and down-regulated in adults when exposed to the 15 ppt treatment, which was an uncharacterized ncRNA.

KEGG Pathways in Adults

Out of the 545 genes identified as significantly differentially expressed between the 15 and 35 ppt treatments within the gills of adult hard clams, 498 were categorized as mRNA by having amino acid sequences deposited in the NCBI database. Of the 498 amino acid sequences, 283 (56.8%) were assigned a K number by BlastKOALA. These 283 protein sequences were assigned to 17 different functional categories (Figure 11) and were assigned to 283 different pathways using KEGG Mapper Reconstruction. Of the 283 different KEGG pathways, 84 were categorized as human disease pathways and their relation to and role in stress response of clams is not known. The remaining 199 pathways were narrowed down to those pathways that included six or more DEG protein sequences (Figure 12). A total of 27 KEGG pathways had six or more DEG protein sequences and belonged to three different levels of pathway categories. Category one, the broadest, sorted pathways into four different families: Organismal Systems (n=12), Cellular Processes

(n=3), Environmental Information Processing (n=11), and Genetic Information Processing (n=1). Category two sorted pathways into nine different families: Development and Regeneration (n=2), Circulatory System (n=1), Endocrine System (n=5), Immune System (n=4), Cell Motility (n=1), Cellular Community – Eukaryotes (n=1), Transport and Catabolism (n=1), Signal Transduction (n=11), and Folding, Sorting, and Degradation (n=1). The third and most specific category named each of the 27 pathways (Figure 12). The pathway with the highest number of DEGs (n=13) was Protein Processing in Endoplasmic Reticulum (Figure 13). The second highest contained nine DEG protein sequences and included the pathways of Axon Regeneration, cGMP-PKG Signaling, and Rap1 Signaling (Figure 14).

KEGG Pathways in Juveniles

Out of the 465 genes identified as significantly differentially expressed between the 15 and 35 ppt treatments within the whole bodies of juvenile hard clams, 415 were categorized as mRNA by having amino acid sequences deposited in the NCBI database. Of the 415 amino acid sequences, 175 (42.2%) were assigned a K number by BlastKOALA. These 175 genes were assigned to 17 different functional categories (Figure 15) and were assigned to 255 different pathways using KEGG Mapper Reconstruction. Of the 255 different KEGG pathways, 75 were part of human diseases and their relation to and role in stress response of clams is not known. The remaining 180 pathways were narrowed down to those pathways that had four or more DEG protein sequences assigned to them (Figure 16). A total of 15 KEGG pathways had four or more DEG protein sequences and belonged to three different levels of pathway categories. Category one, the broadest, sorted pathways into three different families: Organismal Systems (n=4), Cellular Processes (n=4), and Environmental Information Processing (n=7). Category two sorted pathways into seven different families: Development and Regeneration (n=1), Digestive System (n=1), Endocrine

System (n=1), Immune System (n=1), Cellular Community – Eukaryotes (n=2), Cell Growth and Death (n=2), and Signal Transduction (n=7). The third and most specific category named each of the 15 pathways (Figure 16). The pathway with the highest number of DEGs (n=9) was Pancreatic Secretion (Figure 17). The second highest was six DEG protein sequences and included the Rap1 Signaling Pathway (Figure 18).

Clam Line Response to Low Salinity

To understand how each adult clam line individually responds to low salinity, the main effect of salinity on the reference clam line for each model was used to identify DEGs between the 15 and 35 ppt treatments (Figure 19). Five Deseq2 objects were created where each clam line was made the reference clam line once. When WC was the reference line, 64 genes were up, and 28 genes were down-regulated in 15 ppt. When MB was the reference line, 42 genes were up, and 20 genes were down-regulated in 15 ppt. When PS was the reference line, 409 genes were up, and 328 genes were down-regulated in 15 ppt. When CC was the reference line, 66 genes were up, and 120 genes were down-regulated in 15 ppt. When NJ was the reference line, 197 genes were up, and 167 genes were down-regulated in 15 ppt. When each list of DEGs from each clam line's response to 15 ppt vs 35 ppt were compared, five genes were differentially expressed in all clam lines (Figure 20). Out of these five, only two were characterized (phosphoenolpyruvate carboxykinase, cytosolic [GTP] and alpha-crystallin A chain).

Similarly for juvenile clam lines, three Deseq2 objects were created with each clam line made the reference clam line once, and the main effect of salinity on the reference clam line for each model was used to identify DEGs between 15 and 35 ppt (Figure 21). When WC x WC was the reference line, 198 genes were up, and 275 genes were down-regulated between 15 and 35 ppt. When WC x PS was the reference line, 99 genes were up-regulated, and 173 genes were down-

regulated comparing the two salinity treatments. When NC x CC was the reference line, 236 genes were up, and 166 genes were down-regulated between 15 and 35 ppt. When each list of DEGs from each clam line's response to 15 ppt vs 35 ppt were compared, 71 genes were differentially expressed in all clam lines (Figure 22).

Comparing Clam Line Responses to Low Salinity

The interaction terms were used to understand if the response to low salinity was different among clam lines. For adult clams, the same five DESeq2 models as above were run where each clam line was made the reference once. Each interaction term from each model was used to identify genes significantly differentially expressed between two adult clam lines each exposed to 15 ppt (Figure 23). The pair that had the highest number of DEGs was MB and PS. When MB was the reference, 61 gene were up, and 66 genes were down-regulated in PS. When PS was the reference, 62 genes were up, and 68 genes were down-regulated in MB. The pair that had the second highest number of DEGs was PS and CC. When PS was the reference, 24 genes were up, and 40 genes were down-regulated in CC. When CC was the reference, 39 genes were up, and 26 genes were down-regulated in PS. The pair with the third highest number of DEGs was NJ and MB. When NJ was the reference, 26 genes were up-regulated, and 14 genes were down-regulated in MB. When MB was the reference, 12 gene were up, and 29 genes were down-regulated in NJ. The pair with the fourth highest number of DEGs was NJ and CC. When NJ was the reference, 10 genes were up-regulated, and five genes were down-regulated in CC. When CC was the reference, three genes were up, and 11 genes were down-regulated in NJ. All other comparisons only had one to nine genes differentially expressed between the lines under 15 ppt (see Figure 24).

For juvenile clams, the same three DESeq2 models as above were run where each clam line was made the reference once. Each interaction term from each model was used to identify

genes significantly differentially expressed between two juvenile clam lines each exposed to 15 ppt (Figure 24). Between WC x WC (reference) and WC x PS, four genes were down-regulated in WC x PS. Where when WC x PS was the reference, one gene was down-regulated in WC x WC. Between WC x WC (reference) and NC x CC, 17 genes were up-regulated, and seven genes were down-regulated in NC x CC. When NC x CC was the reference, six genes were up-regulated, and 18 genes were down regulated in WC x WC. Between WC x PS (reference) and NC x CC, two genes were up-regulated in NC x CC. When NC x CC was the reference, seven genes were down-regulated in WC x PS.

Discussion

Understanding how the hard clam responds to low salinity stress at adult and juvenile stages is important for understanding how predicted climatological changes could affect this species and the important aquaculture industry it supports. Virginia is the largest producer of hard clams in the U.S., with a farm gate value of \$38.8 million in 2018 and \$57.7 million in 2021, making it the most lucrative aquaculture industry in Virginia (Hudson, 2019; Snyder, 2021). The current study revealed how the gills of adult hard clams and the whole bodies of juvenile clams responded to low salinity stress (15 ppt) and if this response differed among clam lines. The genes differentially expressed by the two tissue types and developmental stages (adults and juveniles) showed similarities and differences to one another in the categories their DEGs belonged to, and the metabolic/biochemical pathways in which the gene products played a role. The comparison of different lines to one another in response to low salinity produced some expected and unexpected results. Throughout this discussion, mRNA or protein coding RNA transcribed from genes is the focus since only mRNA levels were measured. However, the functional roles of the proteins that most likely would have been translated from these mRNAs will be discussed as there have been

studies to support a strong correlation between mRNA, specifically differentially expressed mRNA, levels and their corresponding proteins (Koussounadis et al., 2015; Cheng et al., 2016).

This is the first study to primarily focus on the hard clam, *Mercenaria mercenaria*, transcriptome in response to low salinity in the gills of adults and the whole bodies of juveniles, and to compare this response among genetically distinct hard clam lines created from different populations. However, this is the second study to produce hard clam transcriptomic data after low salinity exposures. A study in 2021 by Song et al. introduced the completion of the hard clam genome and the revelation of a massive expansion and diversification of inhibitors of apoptosis proteins (IAPs) in the genome. Song et al. (2021), focused on the response of IAPs to environmental stressors like aerial exposure, temperature, oxygen levels, and salinity. Salinity experiments were conducted for 10 days at 5, 15, 30, and 40 ppt (Song et al., 2021). The transcriptomic data from the Song et al. (2021) experiments are publicly available but have only been used to understand IAP gene expression.

The Response of Adult Hard Clam Gills to Low Salinity Stress

The genes differentially expressed in response to 15 ppt give insight into the molecular reaction of the gills when they are subjected to osmotic stress. The ciliary action on the gills of hard clams produces a current that pumps water in and out of the mantle cavity and across the gills, allowing for proper gas exchange and food capture (Grizzle, 2001). Salinity affects the rate at which pumping and feeding occurs within hard clams and at 15 ppt pumping rates may fall to 0.5-3.5 L h⁻¹ from 8 L h⁻¹ at more optimal salinities, reducing the clam's ability to support proper metabolic processes and energy demands (Hamwi, 1969). Within the current study, 545 genes in adult gill tissue were differentially expressed, where 408 were up-regulated and 137 were down-regulated in response to 15 ppt. A study by Nie et al. in 2017a compared two different lines of the

manila clam, *Ruditapes philippinarum*, and found 4,467 and 3,168 unigenes that were differentially expressed after low salinity challenges (30 ppt- control vs. 5 ppt- challenge) within two different clam lines. While the current study and that of Nie et al. (2017a) both used gill tissue, these were different species and the salinity experiments as well as the RNA-Seq analysis were conducted differently which may explain the observed differences in the number of genes found to be differentially expressed. The salinity experiments conducted by Nie et al. (2017a) exposed manila clams to decreasing salinities from 30ppt to 5ppt (5 ppt drop/day) for 7 days and bioinformatic analysis was performed de novo or without a genome, leading to a different pipeline of packages being used for the analyses. In both studies, the pathway category of signal transduction had the highest number of amino acid sequences from DEGs assigned to it, and pathways that were identified in both studies included the AMPK, MAPK, PI3K-Akt, Rap1, Ras, cAMP, and cGMP - PKG signaling pathways. These signaling cascades deal with cell polarity, cell adhesion, metabolism, cell death, cell survival, transcription, and more. Under osmotic stress, the regulation of solute concentrations or cell polarity, is imperative for maintaining proper cellular function and regulating cell death. These essential cellular processes will be discussed in greater detail later, but the findings suggest that these pathways play key roles in the sensing and intracellular transduction of low salinity stress signals in both clam species, *R. philippinarum* and *M. mercenaria*.

The Role of Heat Shock Proteins in the Gills of Adult Hard Clams

Several of the top genes by significance and LFC within gills of adult hard clams were genes that encode proteins in the heat shock family. Heat shock proteins (HSPs) are important proteins for physiological stress response in many organisms, including bivalves, and may contribute to bivalves' ability to adapt to changing environments (Fabbri, 2008). HSPs are best

known for their role in stress response but include a wide range of protein families that have specific functions. The gene and protein name terminology will be used in accordance with the following criteria: italics will be used for genes (e.g., *HSP70*, *Hsp72*, etc.). Non-italics will be used for proteins (e.g., HSP70, Hsp72, etc.). Capital letters will refer to the whole protein or gene family (i.e., HSP70, *HSP70*, etc.). Non-capital letters will refer to a specific protein or gene from a family (i.e., Hsp72, *Hsp72*, etc.).

All *HSP* genes identified in this study were up-regulated in response to 15 ppt. The most significant characterized genes based on FDR in this study was alpha-crystallin A chain, a gene in the small heat-shock protein family (sHSP) that encodes proteins with chaperone-like properties to prevent the aggregation of denatured proteins, increasing cellular tolerance to stress (Augusteyn, 2004). The gene encoding heat shock protein beta-1-like (*HspB1* or *Hsp27*), also a member of the sHSP family, was in the top ten most significant characterized genes. HspB1 is best known for counteracting the accumulation of oxidized proteins, protecting cells against oxidative stress (Fabbri, 2008). sHSPs are thought to be ATP-independent chaperones that are important in refolding proteins in combinations with other heat shock proteins (HSPs) (Narberhaus, 2002). Four heat shock protein 70 B2-like genes (*Hsp70 B2*) were in the top 15 genes by absolute LFC (Supp Table 4). Two other DEGs in the *HSP70* family were also identified and up-regulated; heat shock 70 kDa protein-like (*Hsp70*) and heat shock 70 kDa protein 4-like (*Hsp74*). The HSP70 family plays an important part in the defense against different environmental stressors by helping to maintain protein homeostasis (Fabbri, 2008). A study by Hu et al., 2022, identified a massive expansion of *HSP70* genes in the genome of *M. mercenaria* and found that most *HSP70* tandem duplication gene pairs were up-regulated under heat and hypoxia stress in the gills of adult hard clams. Upon encountering stressful conditions, hard clams may shut their valves, which could lead

to hypoxic conditions over time, however, hard clams have been known to sustain valve closure for up to 18 days at -1° to 6°C (Loosanoff, 1939) and survive. It is possible that since *HSP70* genes are expressed in response to heat and hypoxic conditions within the gills of adult hard clam, they are also expressed under hypoosmotic stress. Studies of other bivalves have found higher expression of *HSP70* genes in response to low salinity stress. *HSP70* genes were found to be up-regulated in the gills of the blue mussel, *Mytilus edulis*, after six days of exposure to low salinity stress (23 ppt (control) vs 5 ppt, & 15 ppt vs 5 ppt) (Barrett, 2022). A study by Nie et al., 2017b, exposed the manila clam, *R. philippinarum* to 32 ppt (control) and 15 ppt for 96 hrs and collected tissue samples at 0, 3, 6, 12, 24, 48, 72 and 96 h post-exposure; *HSP70* genes were most highly expressed in the gills as compared to other tissues and expression fluctuated throughout the 96hr and among different clam lines (Nie, 2017b).

Downstream of some *HSP70*/*HSP40* chaperone pathways, the *HSP90* family plays vital roles in protein regulation and cell signaling (Fabbri, 2008). The gene for heat shock protein 90-beta-like (*Hsp90B*) was in the top ten genes by significance in this current study. Like *HSP70* genes, *HSP90* genes were also found to be up-regulated in the blue mussel, *Mytilus edulis*, after exposure to low salinity stress (Barrett, 2022) and up-regulated in *M. mercenaria* after exposure to heat and hypoxic stress (Hu et al., 2022). Song et al. (2022) did a study comparing the transcriptomic response of *M. mercenaria* and its close relative, *M. campechiensis* to chronic heat exposure, and found that the genes encoding *Hsp70B2* and *HspB1* were up-regulated in both species and that the *Hsp90* gene was up-regulated in *M. campechiensis* following exposure to heat stress. Overall, these studies suggest that *HSP90* in conjunction with *HSP70* and *sHSPs* are important stress response genes for the hard clam and specific *HSP* genes may be expressed in response to low salinity stress.

The endoplasmic reticulum (ER) is the organelle within cells where protein homeostasis is maintained by guiding and preserving the proper folding of proteins, where heat shock proteins are a main regulator. The gene for the endoplasmic reticulum chaperone BiP-like (*BiP*), also known as heat shock protein family A, member 5 (*HspA5*), was identified in the top 15 genes by significance in this current study. BiP, along with HSP90 and HSP70 amino acid sequences, were assigned to the KEGG pathway with the highest number of amino acid sequences from DEGs, protein processing in endoplasmic reticulum (Figure 13). BiP is the sole chaperone of HSP70 and one of the most abundant proteins in the ER, indicating that it is very important for protein folding (Adams et al., 2019). These results indicate that when subjected to low salinity stress (15 ppt), the gills of adult hard clams may employ quality control machinery like molecular chaperones and HSPs to ensure proper protein folding and to detect the misfolding and degradation of proteins in the endoplasmic reticulum (ER) (Hartl et al., 2011).

The Response of Juvenile Clams to Low Salinity Stress

This is the first study to investigate the gene expression response in the whole-bodies of juvenile clams following low salinity stress. Juvenile clams are usually raised in up-wellers for multiple weeks until reaching an average SH of 12 mm. Hatcheries and growers usually rely on ambient water to flow through up-wellers and poor water quality, including DO, high or low temperature, sedimentation, and non-optimal (i.e., usually low) salinity, can cause stress and mass mortality events among juveniles. Even when juvenile clams are planted at grow out sites, a drop in salinity can affect growth and survival rates. However, juvenile clams can tolerate lower salinities for periods of time and have been known to survive 14 ppt for 6 days with limited mortality (Baker, 2005). In the event of osmotic environmental change, osmoconformers must reallocate their energy towards maintaining cell volume, possibly by regulating the solute

concentrations using channels and transporters located throughout their cell membranes. The top genes by FDR and LFC in this study included sodium-dependent phosphate transport protein 2B (*NPT2b*) and solute carrier family 12 member 3 (*SLC12A3*). *NPT2b* is a cotransporter of Na^+ and HPO_4^{2-} in the solute carrier family 34 (*SLC34*) (Virkki et al., 2007). Phosphorus (P) is an essential element for all living organisms, and in the form of inorganic phosphate (P_i), it can be used for ATP synthesis, phospholipids, genetic material and intracellular signaling (Knowles, 1980; Olsen et al., 2006). *SLC12A3* is a Na^+ Cl^- cotransporter (NCC), where both ions are transported into the cell and, along with phosphate, are essential small ions that make up most of the solutes in the cells and body fluids of an organism. Sodium and chloride distribution is critical for osmotic homeostasis, especially as Na^+ is needed by the Na^+ - K^+ ATPase. Other genes identified as significantly differentially expressed that code for solute transporters were all up-regulated under 15 ppt and included proteins responsible for transporting Gamma-aminobutyric acid (GABA), monocarboxylate, hydrogen ions, glycine, UDP-galactose, zinc, and multivitamins. Most research conducted on solute carrier genes and their encoded proteins have been with humans, mammals, or other vertebrates. Little research has been done to explore how solute carrier genes are expressed and how their proteins are used by molluscs under osmotic stress. In what organs and tissues these solute transporter genes are expressed within juvenile clams should also be further investigated.

Along with regulating cell volume and solute concentrations, controlling cell death or apoptosis under stressful conditions is crucial for the survival of an organism. Inhibitors of apoptosis (IAPs) primarily function to suppress caspases, the effector proteases in programmed cell death. IAPs have also been found to play a role in signaling cascades that are part of the innate immune response, cell migration, and cell-cycle regulation (Portt, 2011; Damaard, 2011; Gyrd-Hansen, 2010; Fulda, 2014). Multiple *IAPs* and *IAP* related genes were found to be differentially

expressed within juvenile clams after 15 ppt treatment in this current study. Baculoviral IAP repeat-containing protein 7-A, putative inhibitor of apoptosis, and three baculoviral IAP repeat-containing protein 3 genes were all up-regulated, while baculoviral IAP repeat-containing protein 2, another baculoviral IAP repeat-containing protein 3, procathepsin L, two E3 ubiquitin-protein ligase XIAP, and inhibitor of apoptosis protein genes were down-regulated. Sequencing of the hard clam genome revealed a massive expansion and diversification of *IAPs* and after exposure to air for 8 and 16 days, *IAPs* were found to be both significantly up- and down-regulated with expression highest in the hemolymph as compared to other organs (Song et al. 2021). That study also concluded that the expression of *IAPs* were specific for certain stressors; 134 *IAPs* were expressed under at least one environmental stressor (high temperature, low oxygen, and low salinity), but only 3 *IAPs* were found to be expressed under every stressor. The gene expression pattern of *IAPs* found within the whole-body of juvenile clams during this current study may be specific for dealing with low salinity stress and important for resilience against hypoosmotic conditions.

Comparing the Response of Adults and Juveniles to Low Salinity Stress

When the 545 DEGs from the adults and 465 DEGs from the juveniles were compared, 71 genes were commonly differently expressed in both gills of adults and the whole bodies of juveniles. Of these 71 genes, 55 were up-regulated and 15 were down-regulated after 26 hours in 15 ppt in both adults and juveniles (Figure 10). These overlapping genes included genes in the heat shock protein family, *IAP* related genes, and channels/transporters (Supp Table 6). These overlaps suggest that the process of regulating protein folding, apoptosis, and solute movement across cells membranes are important functions for the hard clam to combat low salinity stress. Along with overlapping genes, adults and juveniles shared pathways containing several proteins of the DEGs,

one of which was the Rap1 signaling pathway (Figure 14 & 18). Gene products belonging to the Rap1 pathway that were expressed by adults and juveniles included ras-like GTP-binding protein rhoA, insulin-like peptide receptor, guanine nucleotide-binding protein G(s) subunit alpha, and actin, cytoplasmic-like. Pathways that are linked to the Rap1 pathway include regulation of actin cytoskeleton, focal adhesion, adherens junction, MAPK and PI3K-Akt signaling pathway, all of which had proteins from DEGs belonging to them in either adults, juveniles, or both (Figure 12, 14, 16, & 18). Rap1 and downstream signaling cascades play a role in cellular communication, cell adhesion, migration, polarity, and gene activation. These pathways and genes shared by adults and juveniles may be key for low salinity stress response.

Low Salinity Response Compared Among Adult Clam Lines

Among the five adult clam lines, only five genes were differentially expressed across all lines between 15 and 35 ppt. These included phosphoenolpyruvate carboxykinase, cytosolic [GTP] and alpha-crystallin A chain. It was hypothesized that clam lines derived from the same genetic population would show similar responses when exposed to 15 ppt. However, when clam lines were compared, the expression profiles did not align with the known population structure of the hard clam. When each clam line's response to low salinity was examined and compared, the PS line had the most and NJ line had the second most DEGs in response to 15 ppt compared to 35 ppt. The PS and NJ lines also had the highest number of overlapping DEGs (Figure 19 & 20). It is unknown if this relatively large overlap between PS and NJ is a product of similar responses by these two clam lines or an artifact of these clam lines producing the largest number of DEGs among all the clam lines. Interestingly, the clam lines that were thought to originate from the same populations (WC/NJ & MB/PS) had no overlapping genes (Figure 20). Whether the observed lack

of consistency in DEGs among the clam lines is indicative of differences in response to low salinity or is an artefact of other factors, like low sample size, requires further investigation.

When comparing the responses of clam lines to one another under 15 ppt, the results that aligned with the known population structure of the hard clam were the number of DEGs in the comparisons of NJ/CC (n=14-15), MS/NJ (n=40-41), and PS/CC (n=62-65). These clam line pairs were from different genetic populations and the number of DEGs increased with geographic distance. However, the number of DEGs in the comparisons of PS/MB (n=127-128), PS/NJ (n=5-8), MB/CC (n=6-9) were unexpected as these clam line pairs were either from the same population but expressed a large number of DEGs (i.e., PS/MB) or from different populations that are geographically distant and expressed a limited number of DEGs (i.e., PS/NJ and MB/CC).

MB and PS had the highest number of DEGs in pairwise comparisons in response to exposure to 15 ppt (Figure 23). This result was unexpected as these two lines were created from parents that were considered to be from the same genetic population (Ropp et al., under review), and fewer differences were expected between these two clam lines relative to comparisons between lines from different populations. A possible explanation for this result could be a bottleneck effect that occurred when parents were sampled from these locations and/or a low number of parents that contributed to the two clam lines during the hatchery spawns. Extra samples taken at the time of the experiments in 2021 could be used to further investigate the genetic relationship between these two clam lines to assess whether increasing the sample size results in a significant genetic difference between these two locations. However, if these two clam lines inherited more than just genetic material in the form of differential methylation, otherwise known as epigenetics, this could affect which genes are able to be expressed and would not have been observed by Ropp, et al. (under review) when studying these two locations. The observed difference between these two

clam lines could also be an effect of reaction time to the 15 ppt exposure, as gene expression can change rapidly, and one clam line may have responded faster than the other. However, if this were the case, it may be expected that one clam line (PS or MB) would show comparable response to the other clam lines. The observation that PS and MB clam lines, which were chosen to represent lower salinity clam lines, showed drastically different responses to 15 ppt will require further investigation to understand the cause(s) of these differences.

While the number of DEGs among pairs of clam lines can reveal how similar or different the responses of any two clam lines are, which genes and their possible protein function may be most important when examining DEGs. Which genes, rather than how many, are differently expressed between clam lines could give insight into whether there is a difference in tolerance to low salinity. Clam lines PS and CC had the second highest number of DEGs when comparing their responses to 15 ppt (Figure 23) and in both the PS/CC and the NJ/CC comparisons, *HspB1* was up-regulated in PS and NJ in response to 15 ppt. This up-regulation of a heat shock protein could signify a difference in tolerance to low salinity by PS and NJ compared to CC, but if this up-regulation is indicative or not of better low salinity tolerance would need further investigation. Another *HSP* expressed differentially across clam lines was *HSP90B*, which was up-regulated by WC compared to MB in response to 15 ppt. However, WC had the smallest number of DEGs (1-9) when compared to all other clam lines, which could be a function of having one less sample because of the outlier that was removed.

Low Salinity Response Compared Among Juvenile Clam Lines

Juvenile lines were created to see if crosses between different genetic populations would alter the response of juvenile clams to different stressors, including low salinity. The responses of juvenile lines had more overlapping genes among them (n=71 genes) and that overlap was more

consistent with the known population structure and spawning events, relative to the adult lines. This could be a function of having two fewer lines to compare or the fact that each juvenile sample was derived from four individuals, reducing the effect of variation in the response between individuals. When comparing the responses of each line to 15 ppt vs. 35 ppt, WC x WC and WC x PS had the most DEGs overlapping, which was expected as WC parent clams were used in both crosses (Supp Figure 1). Both WC x WC and WC x PS also had a similar number of overlapping genes with NC x CC (Figure 22).

When the responses of the juvenile lines to 15 ppt were compared, the WC x WC and NC x CC lines had the highest number of DEGs, and the WC x WC and WC x PS lines had the fewest DEGs (Figure 24). Interestingly, when WC x PS and NC x CC were compared, all DEGs were up-regulated in NC x CC. Among the 24 DEGs between WC x WC and NC x CC, the majority were also up-regulated in NC x CC. Some of these up-regulated genes in the NC x CC line were *SLC12A3* and alanine racemase-like. Like mentioned earlier, *SLC12A3* is a cotransporter of sodium and chloride ions, and while inorganic ions are essential, augmentation or depletion of these ions to maintain osmotic homeostasis can be detrimental. Many marine bivalves use free amino acids (FAAs) as osmolytes in regulating intracellular osmolarity as there is virtually no detriment to having high concentrations within the cell (Ballantyne & Chamberlin, 1994; Yancey et al., 1982; Yancey, 1994). Alanine racemase catalyzes racemization between D- and L-alanine and was thought to be unique to prokaryotes until it was discovered within the genomes of eukaryotes, specifically invertebrates (Matsushima et al., 1984; Yamada and Matsushima, 1992; Low et al., 1996; Fujita et al., 1997a, b). Free D- amino acids have been found within the tissue of invertebrates as osmolytes to regulate cell volume (Matsushima et al., 1984; Low et al., 1996; Okuma et al., 1998), and a study by Nomura et al., 2001 found D- and L- alanine to be major

osmolytes in the intracellular FAA pool in the Japanese blue clam, *Corbicula japonica*. Normura et al. (2001) observed the fluctuation of the D-/L-alanine ratio in response to external salinity changes, concluding that alanine racemase had significant contributions to the physiological response of this organism. Why NC x CC clam line displayed higher concentration of mRNA for *SLC12A3* and alanine racemase genes and if this enhances tolerance to low salinity needs further investigation.

Conclusion

As the largest producers of hard clams in the U.S., Virginia clam growers, along with producers from other states, look to the future for ways to maintain and expand production. Potential threats to wild and aquacultured clams, whether they be disease or environmental, need to be assessed and understood so steps can be taken to devise mitigation strategies. An ongoing and worsening threat is the predicted increase in long-lasting precipitation events that lower salinities in coastal areas and can kill or stunt the growth of hard clams. Hatcheries and clam growers rely on ambient water to feed and support clams throughout their lives, with little ability to combat drops in salinity. Genetics and selective breeding have been used as successful tools in many aquaculture settings to enhance favorable characteristics like growth, survival, and stress response in animals (Abdelrahman et al., 2017; Sae-Lim et al., 2017; Hollenbeck & Johnston, 2018; Houston et al., 2020). Efforts to start and maintain hard clam lines with selected traits have been ongoing for the past three years (Evaluating hard clam (*Mercenaria mercenaria*) broodstock to sustain and expand VA hard clam aquaculture in a changing world, 2020; Sea Grant Hard Clam Selective Breeding Collaborative, 2022). With advancements in transcriptomic analysis, we can now begin to better understand how animals respond to stress, what genes may be important in stress response, and if response differs between lines and populations. Transcriptomics have been

used in conjunction with breeding programs to enhance selection methods and better understand differences between lines in other organisms including the intertidal clam (*Sinonovacula constricta*), the Pacific oyster (*Crassostrea gigas*), and the rainbow trout (*Oncorhynchus mykiss*) (Niu et al., 2013; Zhang et al., 2019; Cleveland et al., 2020). The findings of the current study revealed the importance of heat shock proteins, inhibitors of apoptosis, and osmotic effector molecule regulators in the response to low salinity stress by adult and juvenile clams and assessed how the expression of some of these genes differed between pairs of clam lines. While this study did not identify a clam line that is more tolerant to low salinity, the genes that were differentially expressed between lines in response to 15 ppt may be key genes that influence tolerance to low salinity. This study provided insight into the processes that occur within the gills of adult and whole bodies of juveniles under low salinity stress, identifying key genes and pathways that could be the foundation for future studies.

Next Steps

While the current study revealed the shifts in gene expression and the possible pathways involved in the clam's response after 26 hours of low salinity stress, future studies should explore how this response changes over time. Long term salinity experiments could be conducted with adult and juvenile clams over multiple days with multiple sampling intervals to test how gene expression changes through time. Also, the impact of immediate versus gradual decreases in salinity could be assessed, as a gradual change in salinity is more reflective of environmental changes. The results of the current study could also be compared to the low salinity transcriptomic data produced by Song et al. (2021), who conducted experiments for 10 days at 15 ppt. The samples collected from 20 and 12 ppt experiments in this current study could also be analyzed to assess how gene expression may change across decreasing salinities.

The observed variability in response to osmotic stress among clam lines should be further investigated. The large differences between PS and MB clam lines could be an artifact of a bottleneck effect, where the individuals that were collected and/or that contributed to each line are not truly representative of the genetic makeup of the Chesapeake Bay population. Very few individuals may have contributed to each clam line (Supp Table 1 & Supp Figure 1), leading to a hatchery effect that may have skewed results. Extra samples taken from the clam lines at the time of the experiments could be used to investigate if the clam lines used in this study align with the finding of Ropp et al. (under review). The possible downstream effects of the genes found to be up- or down-regulated in relation to low salinity tolerance is unknown. Resolution of these effects may involve a multiple year and generational study that monitors the growth, survival, and reproduction of different clam lines at different salinities and monitors the level of gene expression of specific genes, like HSPs, IAPs, and osmotic effector molecule regulators. Genetic markers in the form of SNPs linked to different expression of genes could also be identified.

Investigation into how response differs among tissue types could help better understand the findings based on the juvenile clams. Juvenile samples were taken from the whole bodies of four individuals, and it is unknown in which tissues and organs those genes were differentially expressed. It is also unknown if the genes that overlap between the juvenile and adults DEGs are all commonly expressed in the gills or are also expressed in other tissues. Understanding how gene expression changes with age within a particular tissue type for the hard clam is also lacking. At the time of the experiments in this current study, mantle tissue was also sampled for all adult clam lines and salinities and could be used to compare gene expression with the gills and juvenile samples.

Potential Limitations of this Study

The limitations of this study should be acknowledged as they may have contributed to the results. First, a limited number of parents were used to create each clam line, and, in some instances, it is unknown how many females and males contributed to each line (Supp Table 1). The relationship among clams within each line is unknown. In some cases, individuals from the same line could have been whole siblings, half siblings, or unrelated. The genetic makeup and variation within each clam line could have influenced results, and therefore results may not be representative of population-level responses. Second, due to the timing of this study and funding limitations, clam beds kept at a low salinity location for long-term low salinity exposures were not conducted in parallel as was originally planned, and the clam lines were only held in higher salinity beds. Subsets of each clam line were given to local growers to raise and maintain, and a portion was kept by VIMS for future breeding efforts. One of these growers kept clams under nets in a location off the Poquoson River (~20ppt). Efforts were made to sample these adult clams in October of 2021 as they would have been exposed to lower salinities for over two years; however, husbandry and shifting sands caused increased and non-proportional mortality among clam lines, necessitating abandonment of this portion of the study. Thirdly, the 26-hour salinity exposures used in this study represent a quick and drastic drop in salinity reflective of what may occur in the event of a freshet. Clams may respond differently to low salinity as time progresses, and more studies will be needed to understand the effect of long-term chronic stress. Fourthly, due to limited funding, 12 and 20 ppt tissue samples did not undergo RNA-Seq. This prevented the study from looking at how clams respond across a gradient of decreasing salinities. Future studies may be able to include those samples, which are preserved at -80 °C. Lastly, this study relied on the construction and annotation of the hard clam genome of Song et al., 2021 on NCBI. Many of the

DEGs came up as non-characterized based on this genome annotation, so aspects of how the hard clams respond to low salinity may have been missed.

# of Samples	Avg Raw Reads	Avg Raw Data	Avg Effective (%)
48	4.76E+07 ± 5.69E+06	7.15 ± 0.85	97.95 ± 0.44
Avg Error (%)	Avg Q20 (%)	Avg Q30 (%)	Avg GC (%)
0.0290 ± 0.0031	97.90 ± 0.12	93.89 ± 0.31	38.26 ± 1.42

Table 1. This table includes the mean and standard deviation of raw reads (for paired-end sequencing, it equals the number of read1 and read2), raw data ((raw reads) * (sequence length (150 bp), calculated in giga (G), where the total raw data calculated was 343.0 G), effective percent ((clean reads/ raw reads)*100%), error percent (base error rate * 100%), Q20 & Q30 percent ((base count phred > 20 or 30 / total base count) * 100%), and GC percent ((G & C base count / total base count)* 100%) for all 48 RNA sequencing samples.

Gene	Name	Base Mean	LFC	LFC SE	pvalue	padj
XM_045306916.1	alpha-crystallin A chain-like	967.5722041	2.384146675	0.270193385	5.79E-20	1.65E-16
XM_045335801.1	phosphoenolpyruvate carboxykinase, cytosolic [GTP]-like transcript variant X2	750.8122234	5.219203638	0.59436961	8.47E-20	1.81E-16
XM_045360456.1	heat shock protein beta-1-like	313.6829135	1.997567209	0.231494767	2.97E-19	5.63E-16
XM_045358251.1	marginal zone B- and B1-cell-specific protein-like	185.7804568	2.036149072	0.241911603	1.88E-18	3.20E-15
XM_045345110.1	serine-arginine protein 55-like	1261.482487	-1.417809782	0.175459315	2.87E-17	3.76E-14
XM_045329467.1	heat shock protein HSP 90-beta-like	7009.519588	1.752372951	0.215588701	2.47E-17	3.76E-14
XM_045344379.1	steroid 17-alpha-hydroxylase/17,20 lyase-like	1115.576467	-5.357123266	0.67366018	6.65E-17	8.09E-14
XM_045359836.1	protein disulfide-isomerase A3-like	3369.271394	1.586141711	0.198956546	8.77E-17	9.96E-14
XM_045316636.1	inositol-3-phosphate synthase 1-A-like	104.9794979	4.23730962	0.533941758	9.39E-17	1.00E-13
XM_045303941.1	alreticulin-like	4148.547277	1.934075756	0.245380251	1.71E-16	1.62E-13

Table 2. Top ten adult hard clam characterized genes whose expression differed between 35 and 15 ppt exposures, ranked by FDR or adjusted p-value (padj). Gene is the gene accession number on NCBI. Name is the name given for each feature on NCBI. Base mean is the average of the normalized count values, divided by the size factor, taken over all the samples. LFC is the log₂ fold change. LFC SE is the LFC standard error. The pvalue is the Wald test p-value, where the null hypothesis that there is no effect of treatment on the gene. The padj is the adjusted p-value based on the Benjamini-Hochberg method, or false discovery rate (FDR).

Gene	Name	Base Mean	LFC	LFC SE	pvalue	padj
XM_045318505.1	sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1-like	324.9447929	5.049886016	0.335112018	1.59E-52	2.76E-48
XM_045351851.1	sodium-dependent phosphate transport protein 2B-like	618.4400228	3.742419808	0.253368846	1.20E-50	1.04E-46
XM_045360561.1	solute carrier family 12 member 3-like	242.2277978	2.745805529	0.227974773	1.03E-34	5.95E-31
XM_045327875.1	protein PIF-like	390.2895119	4.968425591	0.414554956	2.35E-34	1.02E-30
XM_045354440.1	homeobox protein engrailed-2b-like	64.86771431	4.369365887	0.390775795	3.99E-30	1.38E-26
XM_045337089.1	immediate early response gene 5-like protein	3419.881254	1.618695542	0.15536069	1.05E-26	3.03E-23
XM_045333072.1	DBH-like monooxygenase protein 2 homolog	244.3909213	2.164166038	0.212763087	1.27E-25	2.76E-22
XM_045320073.1	retinol dehydrogenase 7-like	751.2496527	- 2.157683204	0.214645274	4.00E-25	7.70E-22
XM_045308535.1	protein BTG2-like	6383.71021	1.041444379	0.106403926	6.20E-24	9.77E-21
XM_045329160.1	bypass of stop codon protein 1-like	43.80081254	- 4.442274699	0.472267853	4.00E-22	5.33E-19

Table 15. Top ten juvenile hard clam characterized genes whose expression differed between 35 and 15 ppt exposures, ranked by FDR or adjusted p-value (padj). Gene is the gene accession number on NCBI. Name is the name given for each feature on NCBI. Base mean is the average of the normalized count values, divided by the size factor, taken over all the samples. LFC is the log₂ fold change. LFC SE is the LFC standard error. The pvalue is the Wald test p-value, where the null hypothesis that there is no effect of treatment on the gene. The padj is the adjusted p-value based on the Benjamini-Hochberg method, or false discovery rate (FDR).

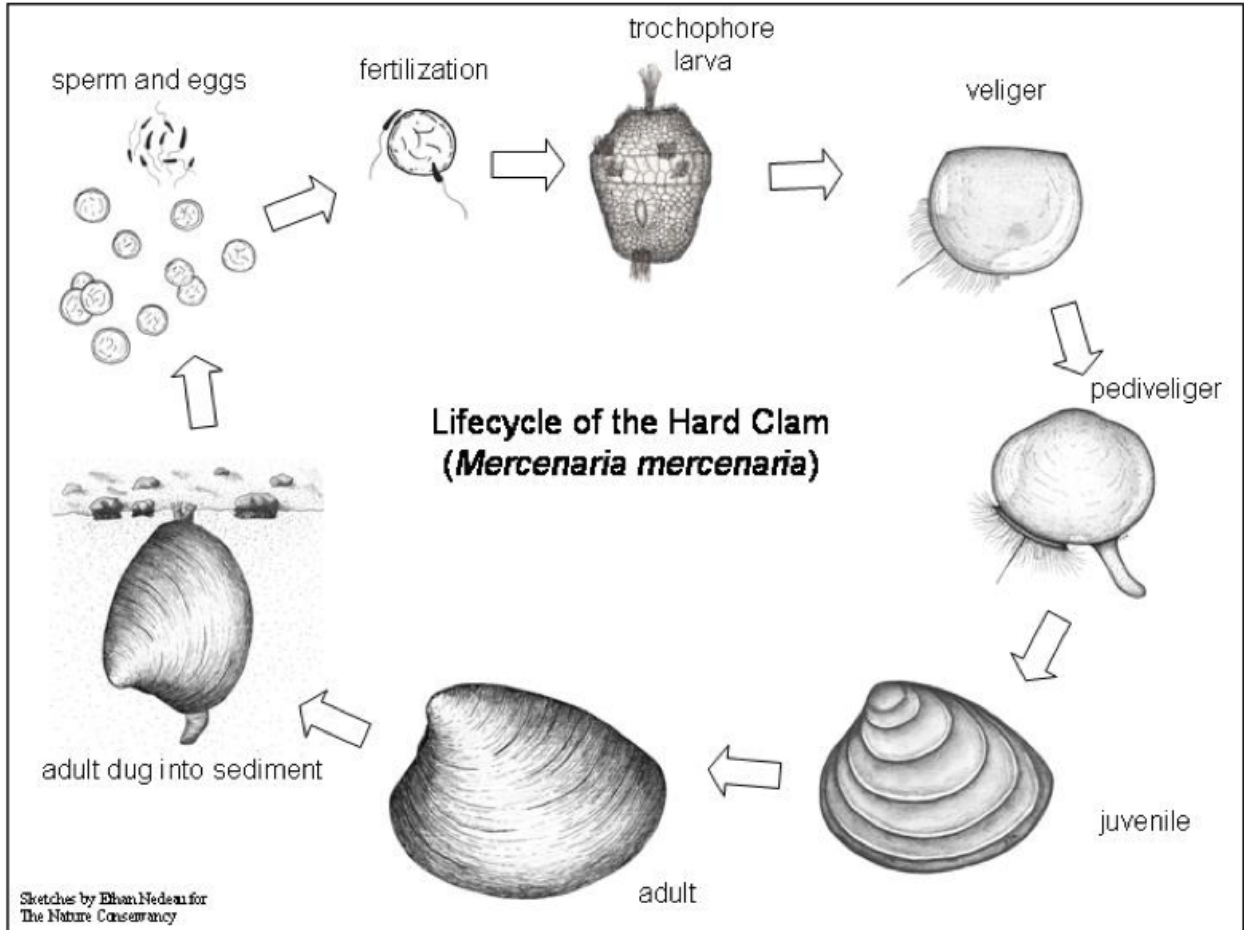


Figure 1. The Lifecycle of the Hard Clam - Adapted from Ethan Nedeau

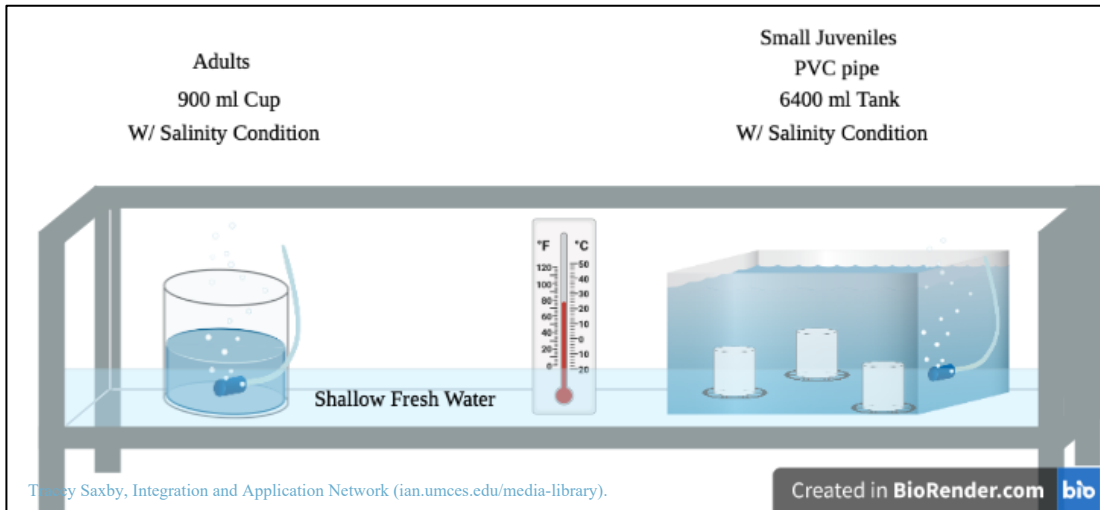


Figure 2. Experimental set up with cups of 900 ml treatment water that housed adult clams, and tanks of 6400 ml treatment water that housed small juveniles in PVC pipes, in shallow circulating freshwater that was used to regulate temperature at 25 °C.

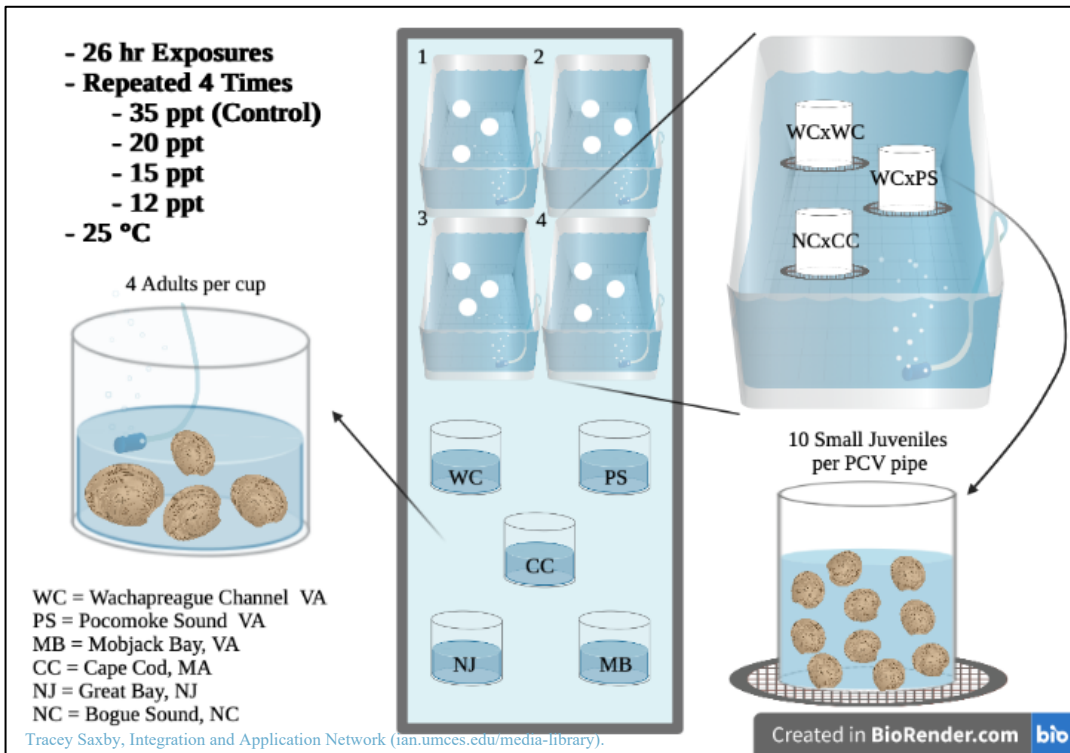


Figure 3. Experimental design with five cups of 900 ml treatment water that housed four clams of each 2019 line. Four tanks of 6400 ml treatment water housed a PVC pipe with a screened bottom with 10 clams from each 2021 clam lines, creating 4 replicates of 10 clams per line. This was repeated with each salinity condition for 26-hour exposures.

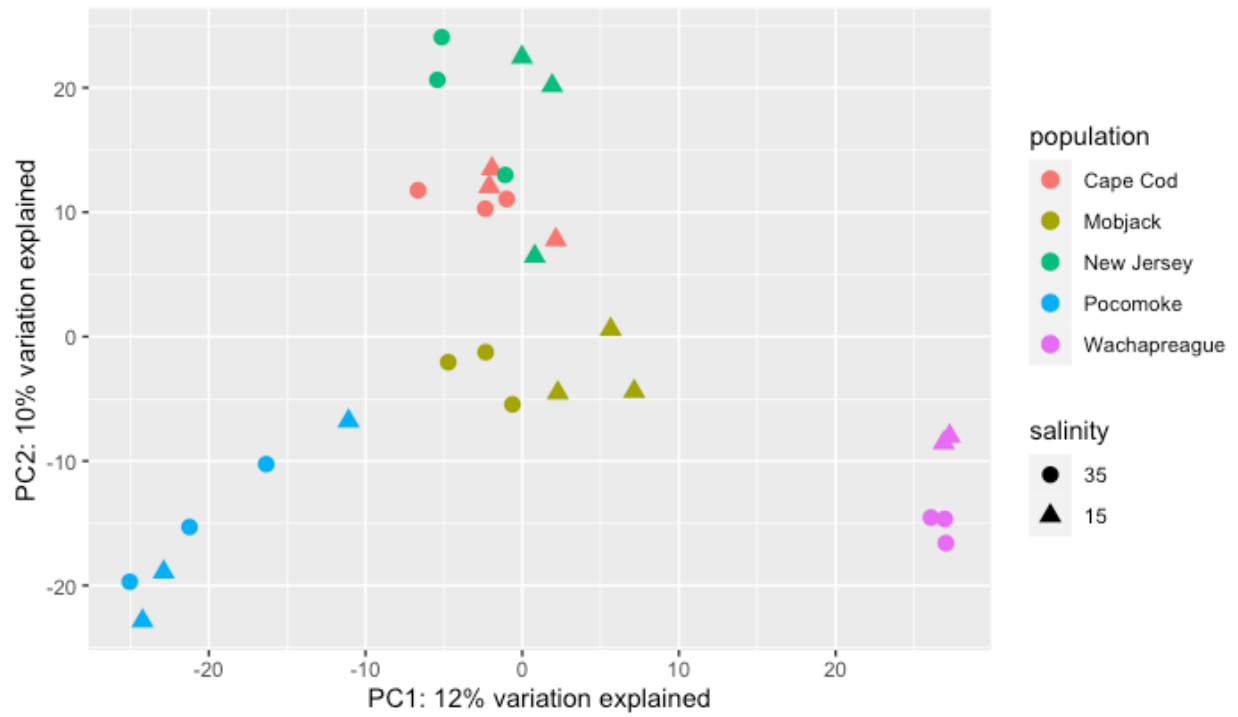


Figure 4. Variance Stabilizing Transformation PCA of PC1 & PC2 for adult clam samples lines totaling 22% of the variance explained. Sample 15_Wach_G_3 was removed as an outlier and is not included in this PCA.

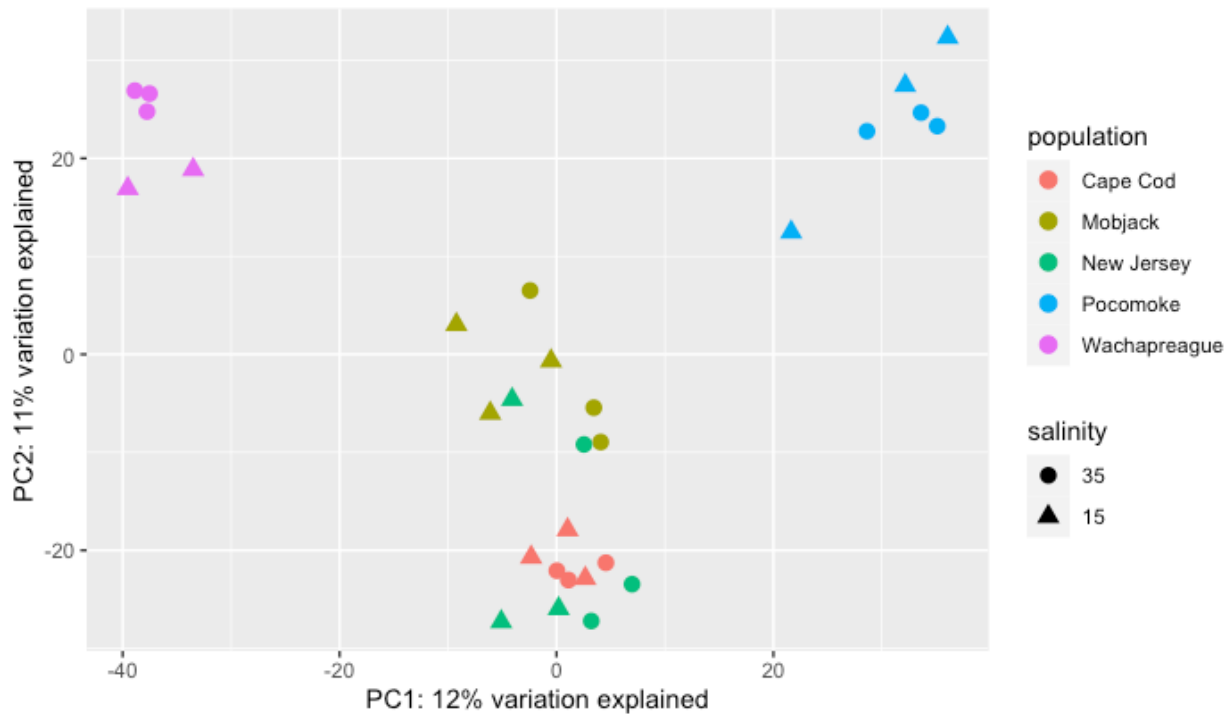


Figure 5. Regularized-Logarithm Transformation PCA of PC1 & PC2 for adult clam samples totaling 23% of the variance explained. Sample 15_Wach_G_3 was removed as an outlier and is not included in this PCA.

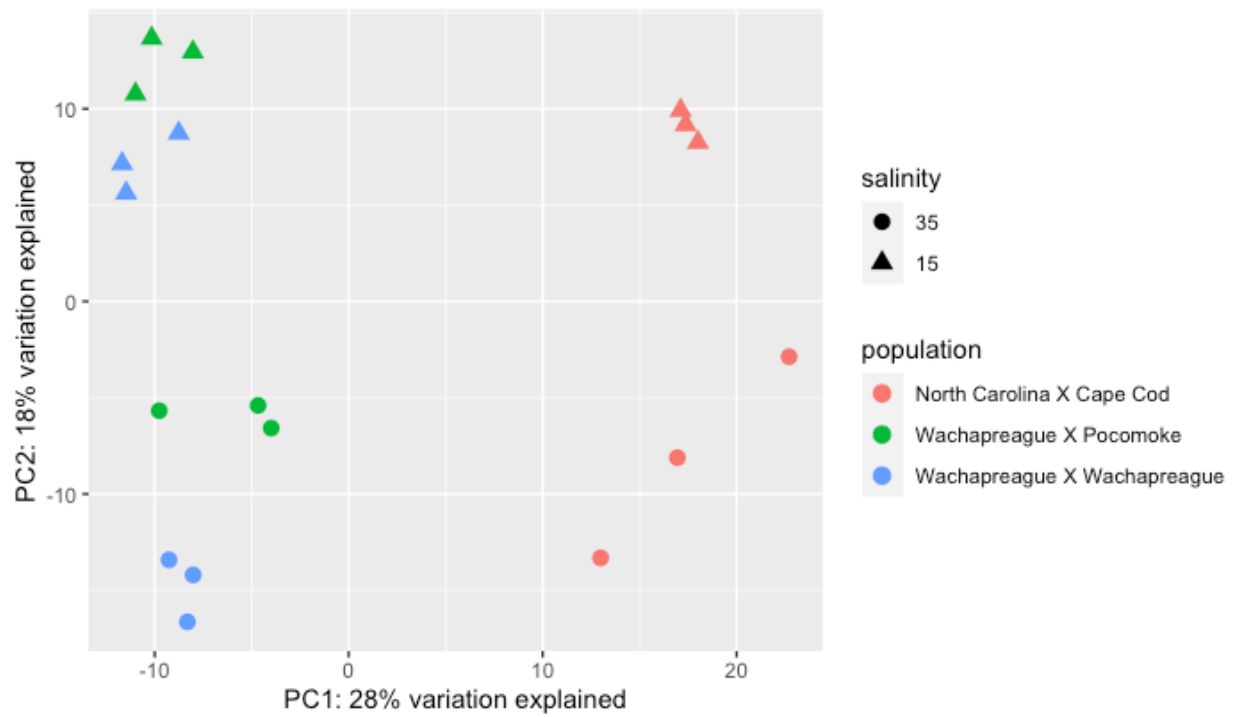


Figure 6. Variance Stabilizing Transformation PCA of PC1 & PC2 for juvenile clam samples lines totaling 46% of the variance explained.

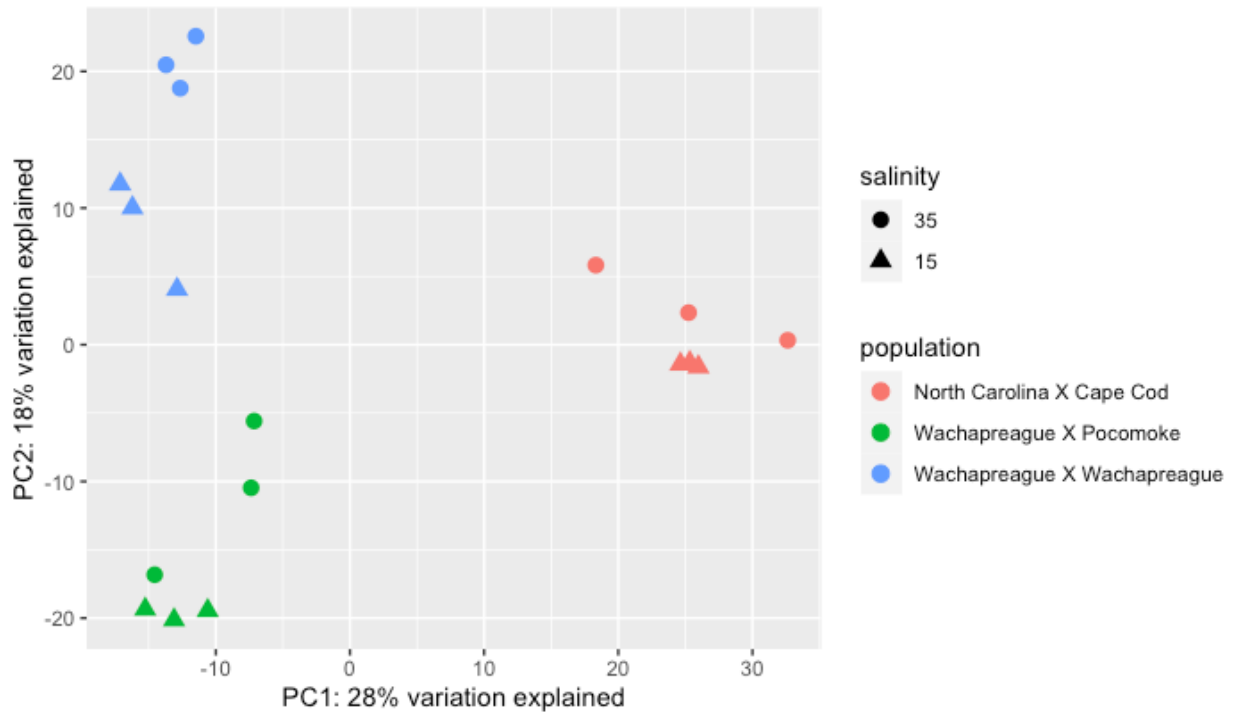
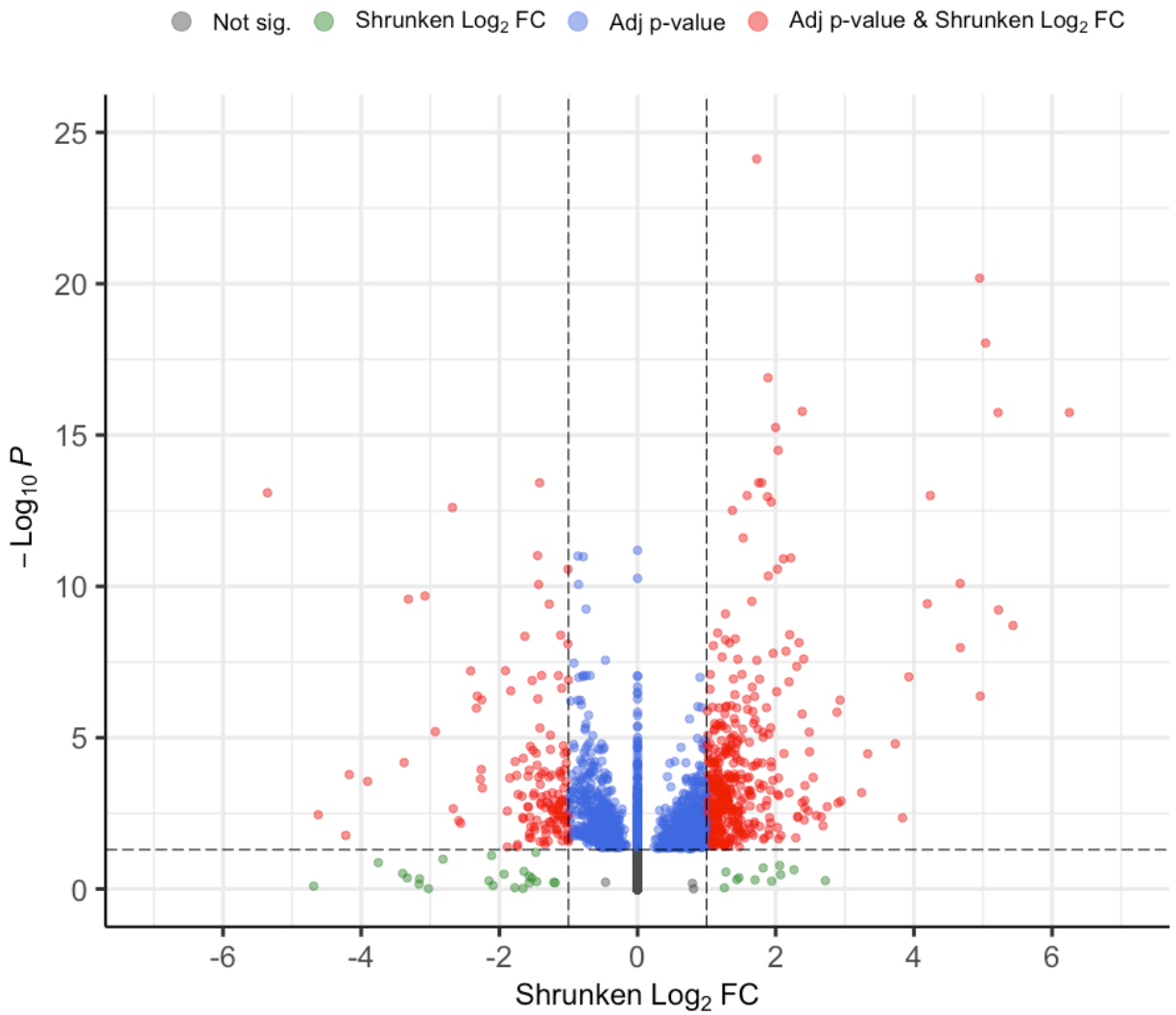


Figure 7. Regularized-Logarithm Transformation PCA of PC1 & PC2 for juvenile clam samples totaling 46% of the variance explained.



total = 25484 variables

Figure 8. Volcano plot of adult clam's gene expression between 15 ppt and 35 ppt. The $-\log(\text{base } 10)$ of the adjusted p-value is on the y-axis and the Shrunken LFC is on the x-axis. Those dots in grey represent genes that fall outside the FDR and LFC thresholds. Those dots in green represent genes that have an absolute LFC >1 but have an FDR >0.05 . Those dots in blue represent genes that have an FDR <0.05 but have an absolute LFC <1 . Those dots in red represent genes that are significantly differentially expressed between 15 and 35 ppt and make up a total of 545 genes. Of those, 408 are up-regulated in 15 ppt, and 137 are down-regulated in 15 ppt.

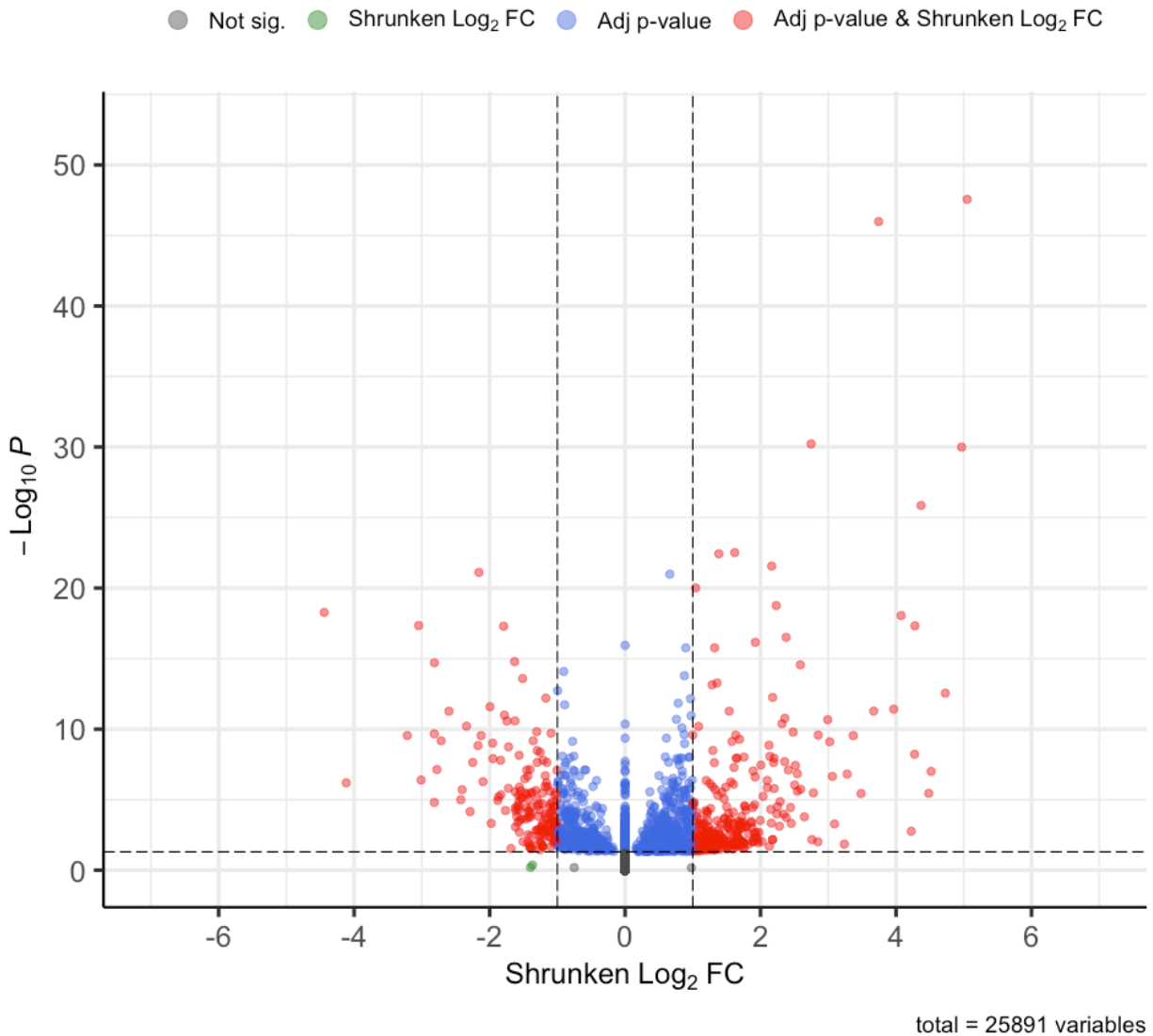


Figure 9. Volcano plot of juvenile clam's gene expression between 15 ppt and 35 ppt. The $-\log(\text{base } 10)$ of the adjusted p-value is on the y-axis and the Shrunken LFC is on the x-axis. Those dots in grey represent genes that fall outside the FDR and LFC thresholds. Those dots in green represent genes that have an absolute LFC >1 but have an FDR > 0.05 . Those dots in blue represent genes that have an FDR < 0.05 but have an absolute LFC < 1 . Those dots in red represent genes that are significantly differentially expressed between 15 and 35 ppt and make up a total of 465 genes. Of those, 303 are up-regulated in 15 ppt, and 162 are down-regulated in 15 ppt.

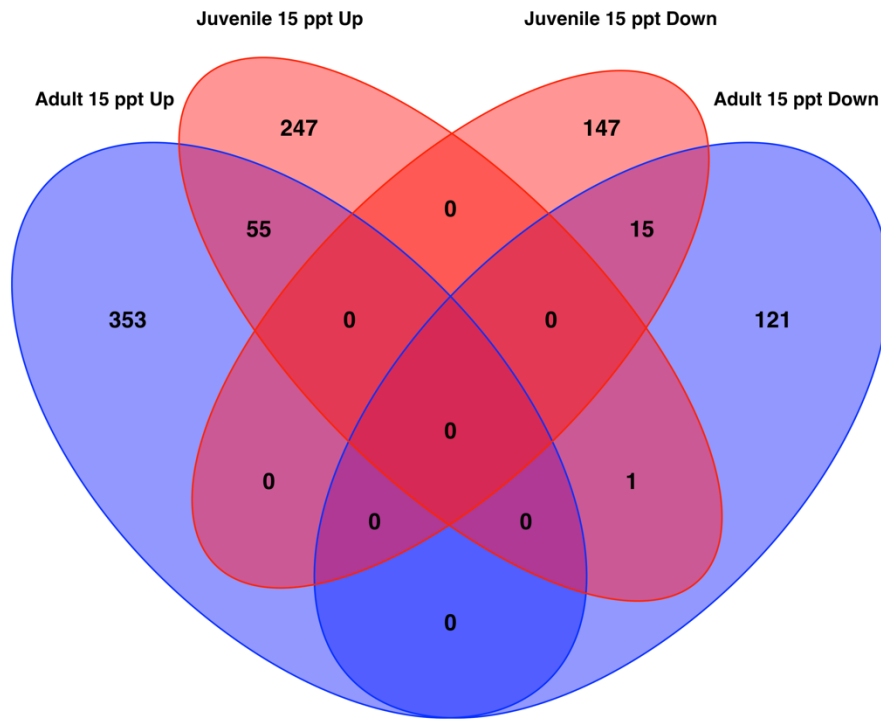


Figure 10. Venn diagram of the number of genes that are significantly differentially expressed (absolute LFC > 1 and FDR < 0.05) similarly or differently among the gills of adult clams and the whole body of juveniles. 71 genes were commonly differentially expressed in both tissues.

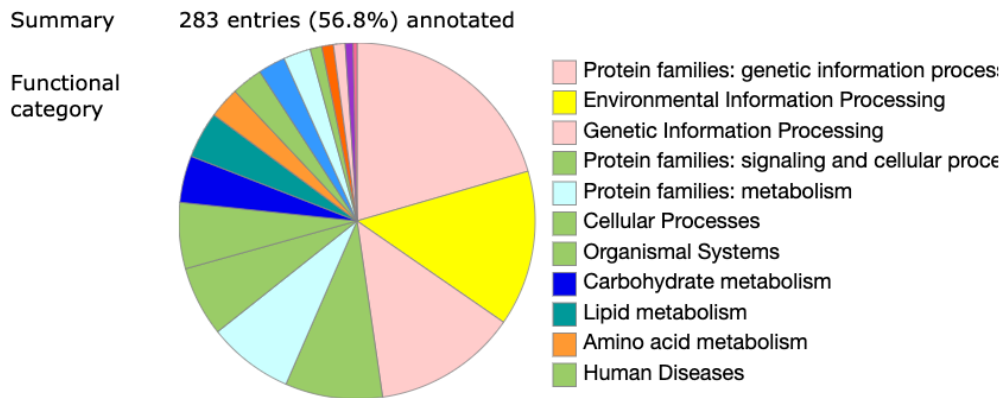


Figure 11. Adult clam 15 ppt vs. 35 ppt differentially expressed genes annotated by BlastKOALA & KEGG database. Out of the 545 genes identified as differentially expressed between 15 and 35 ppt within the gills of adult hard clams, 498 were categorized as mRNA/ had amino acid sequences on NCBI. Of the 498 amino acid sequences, 283 (56.8%) were assigned a K number by BlastKOALA. These 283 genes were assigned to 17 different functional categories and were assigned to 283 different pathways using KEGG Mapper Reconstruction.

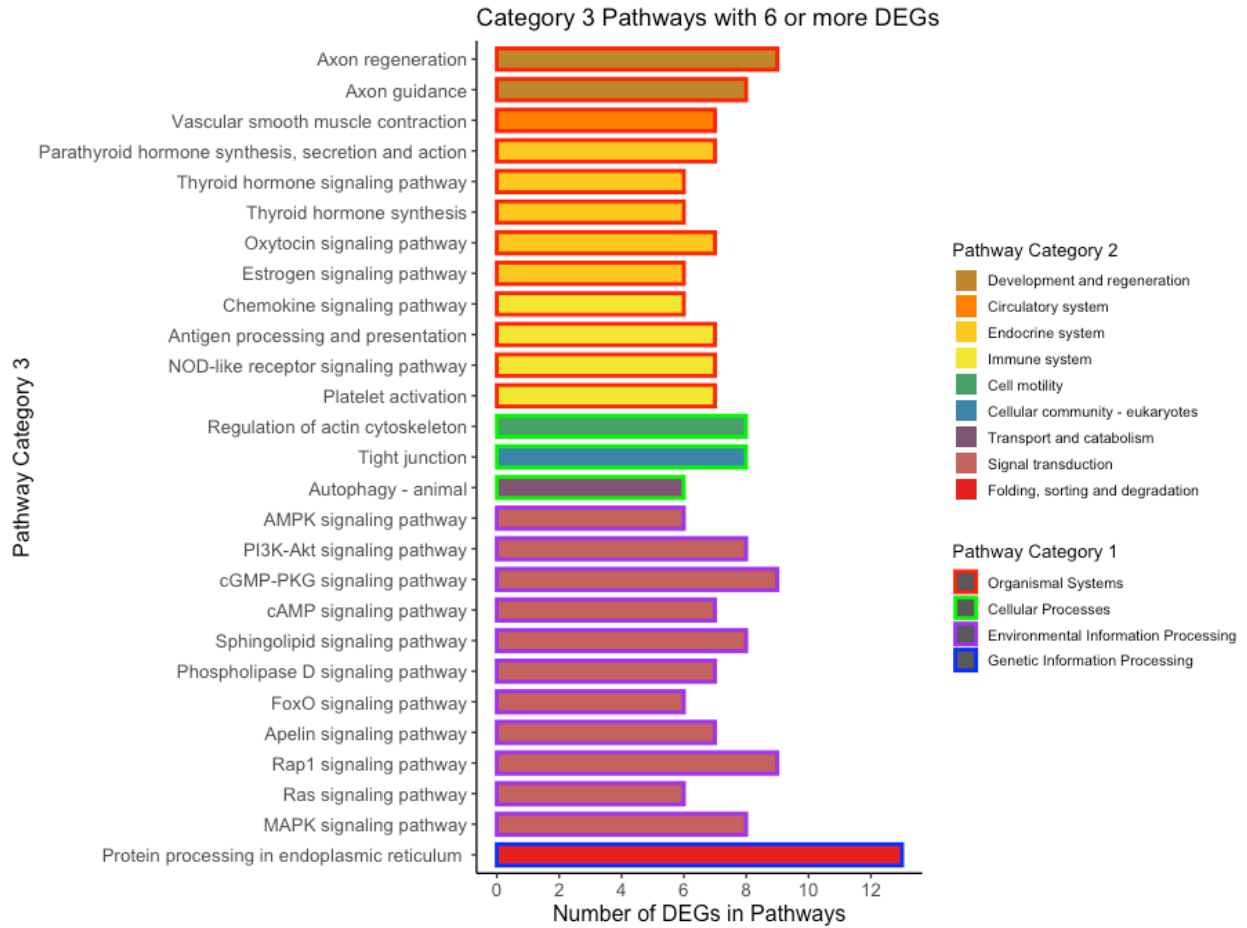


Figure 12. KEGG pathways with six or more adult 15 ppt vs. 35 ppt differentially expressed gene products. Pathway category one is the broadest and sorted the 27 pathways into four different categories, which is the color outline around each bar. Pathway category two sorted the 27 pathways into nine different categories, which is the color inside each bar. Pathway category three names the specific pathway on the y axis. The number of DEG products within that pathway is on the x-axis.

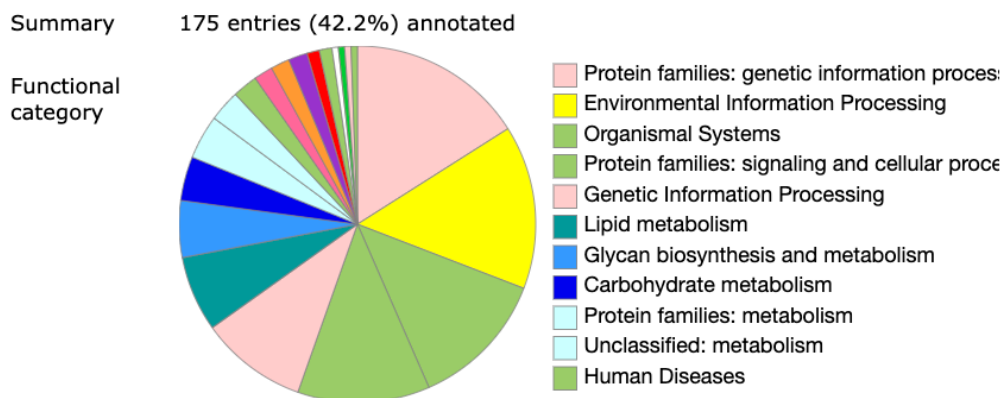


Figure 15. Juvenile clam 15 ppt vs. 35 ppt differentially expressed genes annotated by BlastKOALA & KEGG database. Out of the 465 genes identified as differentially expressed between 15 and 35 ppt within the whole bodies of juvenile hard clams, 415 were categorized as mRNA/ had amino acid sequences on NCBI. Of the 415 amino acid sequences, 175 (42.2%) were assigned a K number by BlastKOALA. These 175 genes were assigned to 17 different functional categories and were assigned to 255 different pathways using KEGG Mapper Reconstruction.

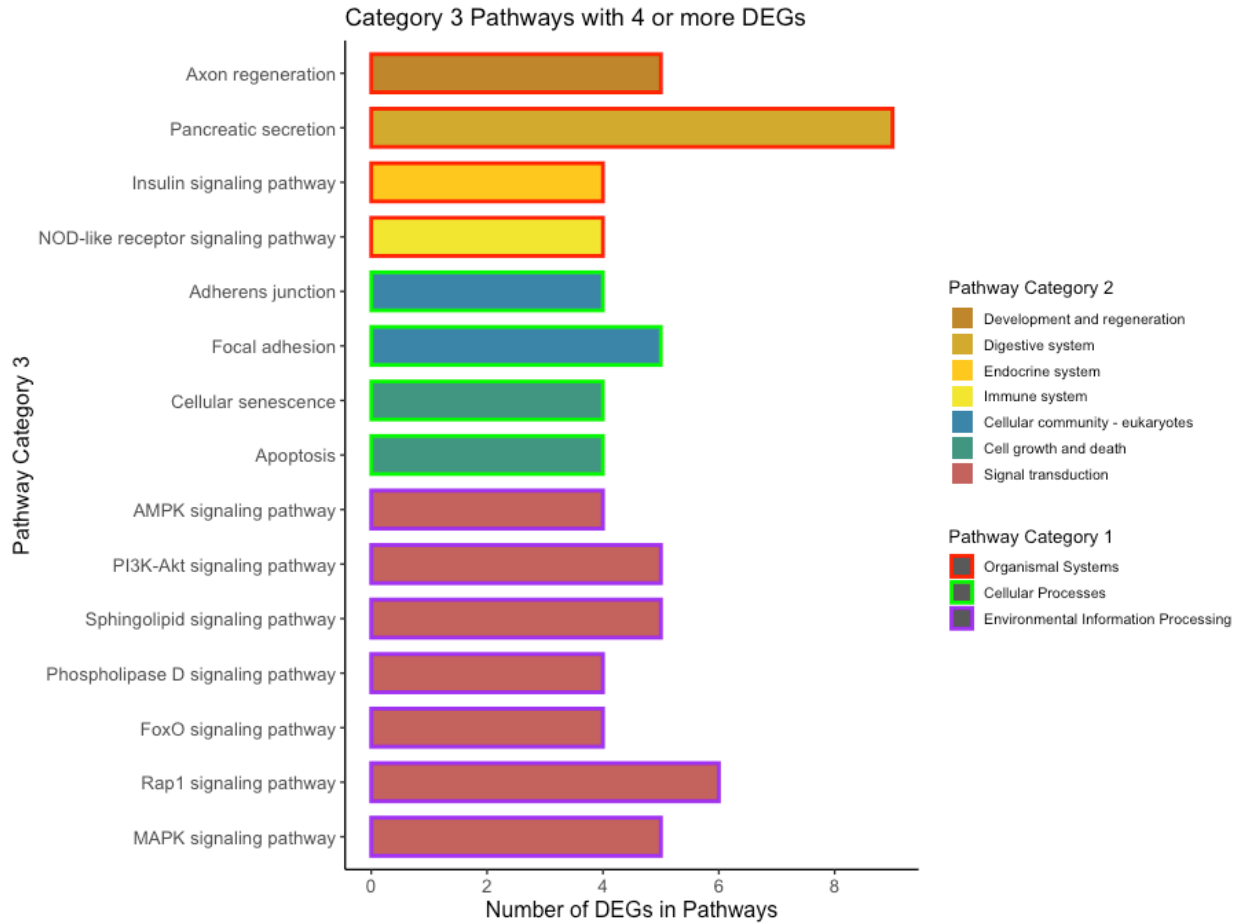


Figure 16. KEGG pathways with four or more juvenile 15 ppt vs. 35 ppt differentially expressed gene products. Pathway category one is the broadest and sorted the 15 pathways into three different categories, which is the color outline around each bar. Pathway category two sorted the 15 pathways into seven different categories, which is the color inside each bar. Pathway Category three names the specific pathway on the y-axis. The number of DEG products within that pathway is on the x-axis.

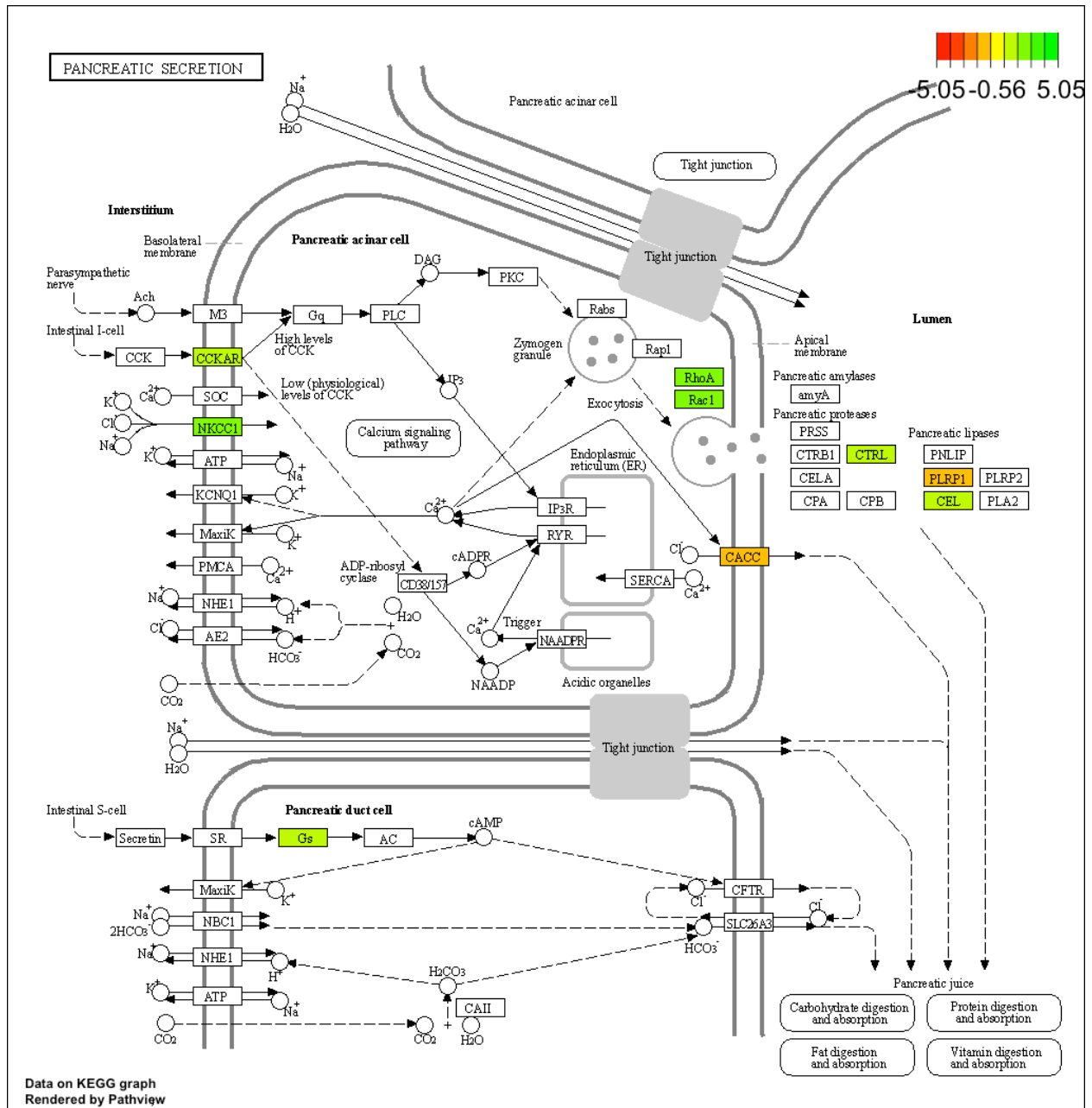


Figure 17. Pancreatic secretion pathway with juvenile 15 ppt vs. 35 ppt differentially expressed gene products. Nine DEG products within the whole bodies of juvenile clams between 15 and 35 ppt belong to this pathway. Those boxes in green are genes with positive LFC or were up-regulated in 15 ppt. Those boxes in yellow to red are genes with negative LFC and were down-regulated in 15 ppt.

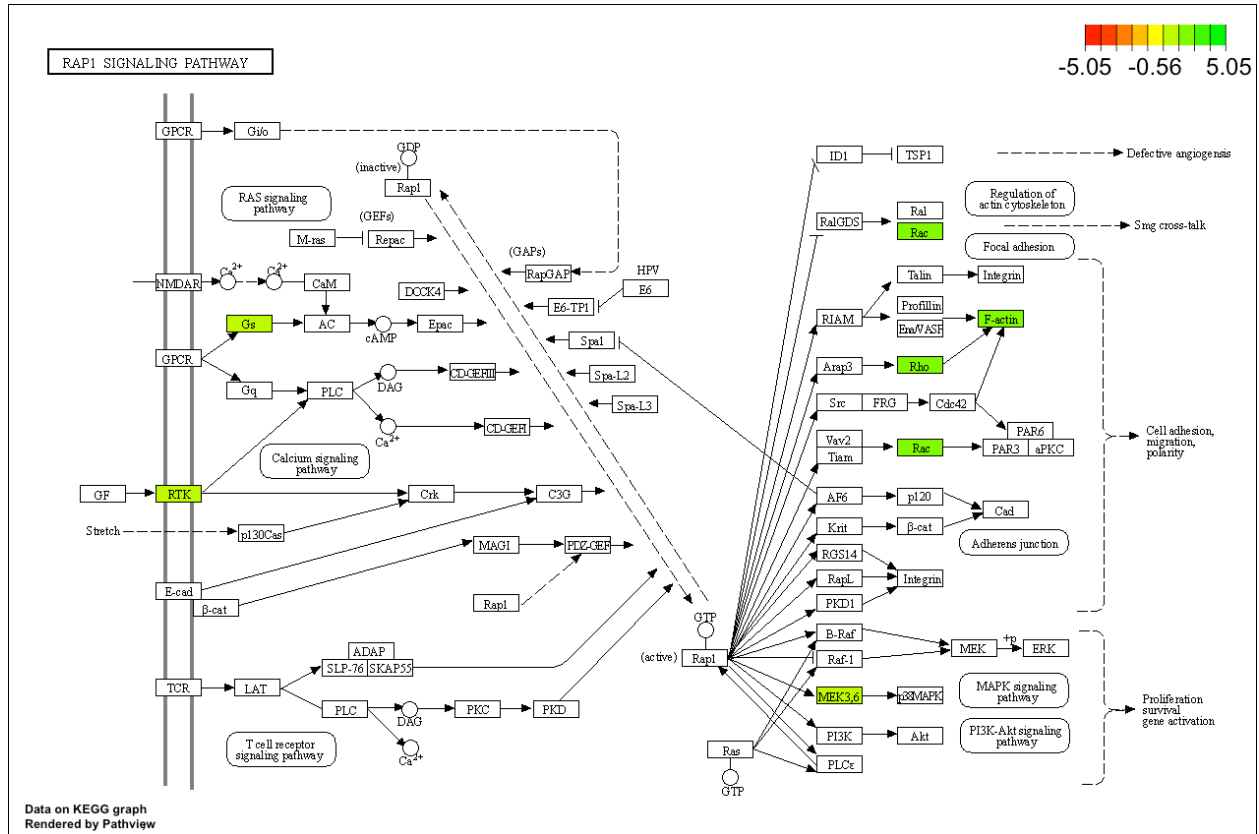


Figure 18. Rap1 signaling pathway with juvenile 15 ppt vs. 35 ppt differentially expressed gene products. Six DEG products within the whole bodies of juvenile clams between 15 and 35 ppt belong to this pathway. Those boxes in green are genes with positive LFC or were up-regulated in 15 ppt. Those boxes in yellow to red are genes with negative LFC and were down-regulated in 15 ppt.

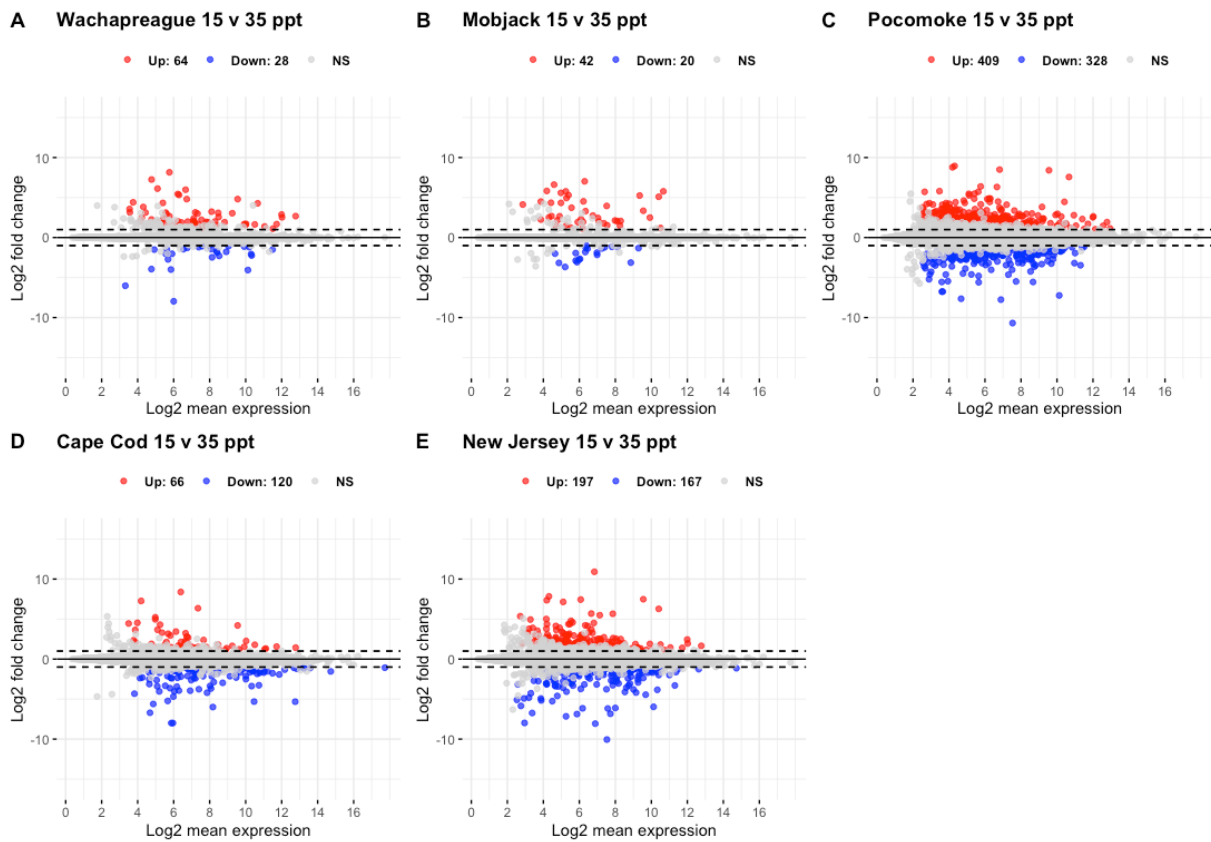


Figure 19. MA plot for adult clam lines 15 ppt vs. 35 ppt. The log₂ mean expression or log₂ mean of normalized counts are along the X-axis, and the log₂ fold changes are along the Y-axis for plots A-E. Those dots in red and blue represent genes that are significantly differentially expressed with an absolute LFC > 1 and an FDR < 0.05. Those dots in red are up-regulated in 15 ppt, and those dots in blue are down-regulated in 15 ppt. Those dots in grey are non-significant (NS). The Wachapreague clam line had 64 genes up and 28 genes down-regulated in 15 ppt. The Mobjack clam line had 42 genes up and 20 genes down-regulated in 15 ppt. The Pocomoke clam line had 409 genes up and 328 genes down-regulated in 15 ppt. The Cape Cod clam line had 66 genes up and 120 genes down-regulated in 15 ppt. The New Jersey clam line had 197 genes up and 167 genes down-regulated in 15 ppt.

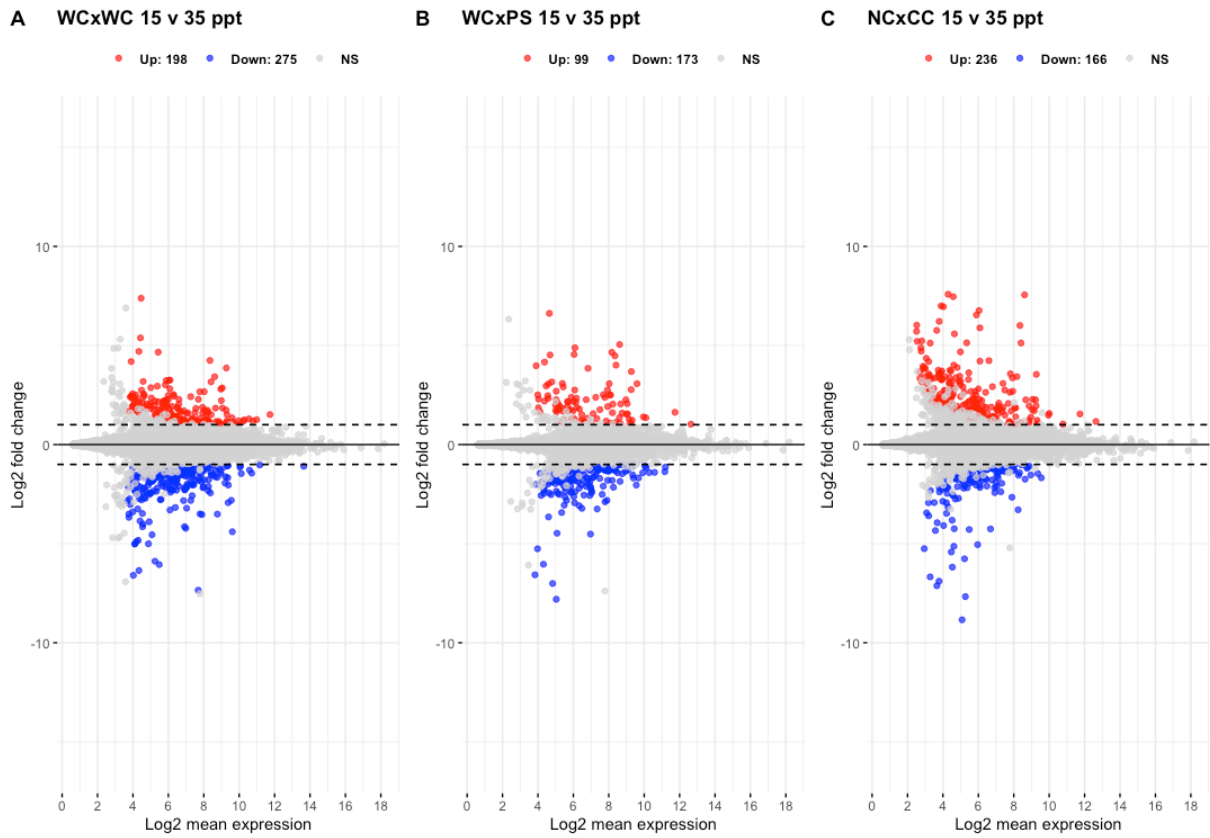


Figure 21. MA plot for juvenile clam lines 15 ppt vs. 35 ppt. The log₂ mean expression or log₂ mean of normalized counts are along the X-axis, and the log₂ fold changes are along the Y-axis for plots A-C. Those dots in red and blue represent genes that are significantly differentially expressed with an absolute LFC > 1 and an FDR < 0.05. Those dots in red are up-regulated in 15 ppt, and those dots in blue are down-regulated in 15 ppt. Those dots in grey are non-significant (NS). The WC x WC clam line had 198 genes up and 275 genes down-regulated in 15 ppt. The WC x PS clam line had 99 genes up and 173 genes down-regulated in 15 ppt. The NC x CC clam line had 236 genes up and 166 genes down-regulated in 15 ppt.



Figure 22. Venn diagram of the number of genes that are significantly differentially expressed (absolute LFC > 1 and FDR < 0.05) similarly or differently among the juvenile clam lines when each is run as the control line through Deseq2 (design = ~ Clam Line + Salinity + Clam Line * Salinity) and the main effect of salinity coefficient is used.

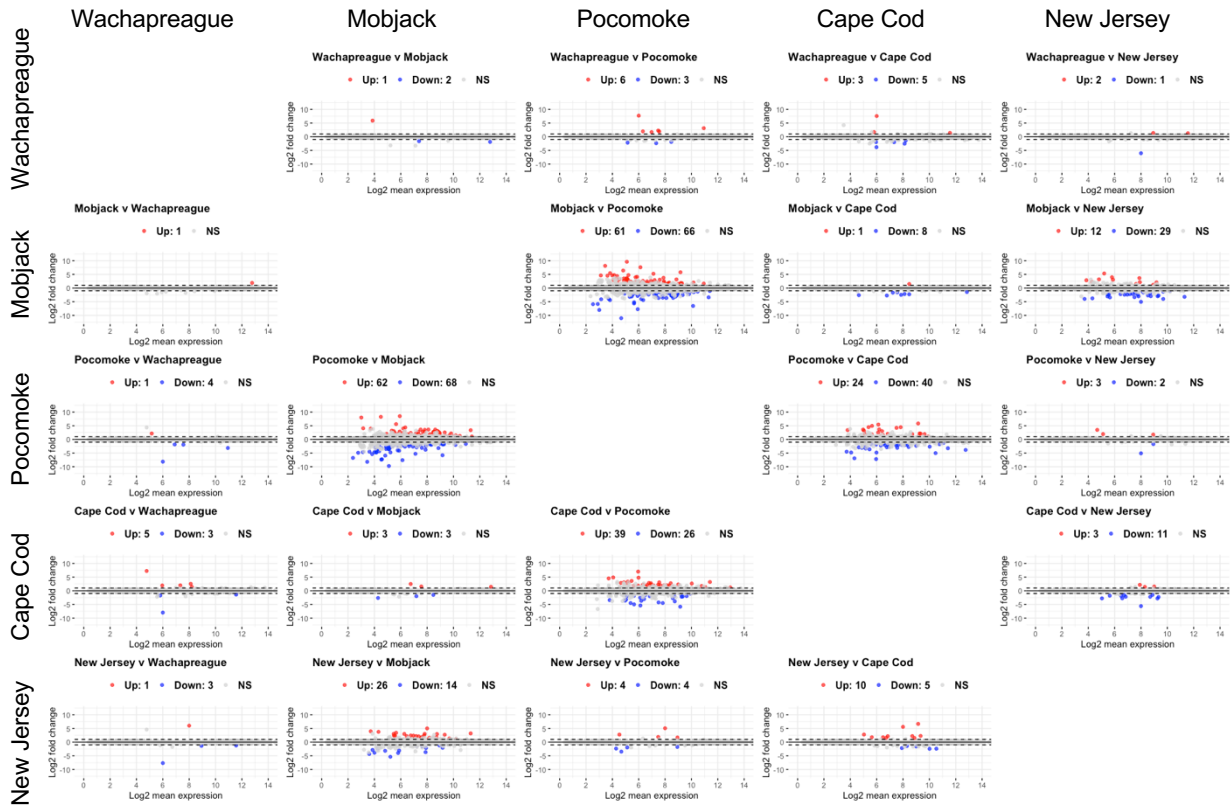


Figure 23. MA plot for adult clam lines interaction terms. The \log_2 mean expression or \log_2 mean of normalized counts are along X-axis and the \log_2 fold changes are along the Y-axis for each plot. Those dots in red and blue represent genes that are significantly differentially expressed with an absolute LFC > 1 and an FDR < 0.05. Those dots in red are up-regulated in the non-control clam line and those dots in blue are down-regulated in non-control clam line. Those dots in grey are non-significant (NS). The clam line that is listed first in each comparison is the control clam line for that pair.

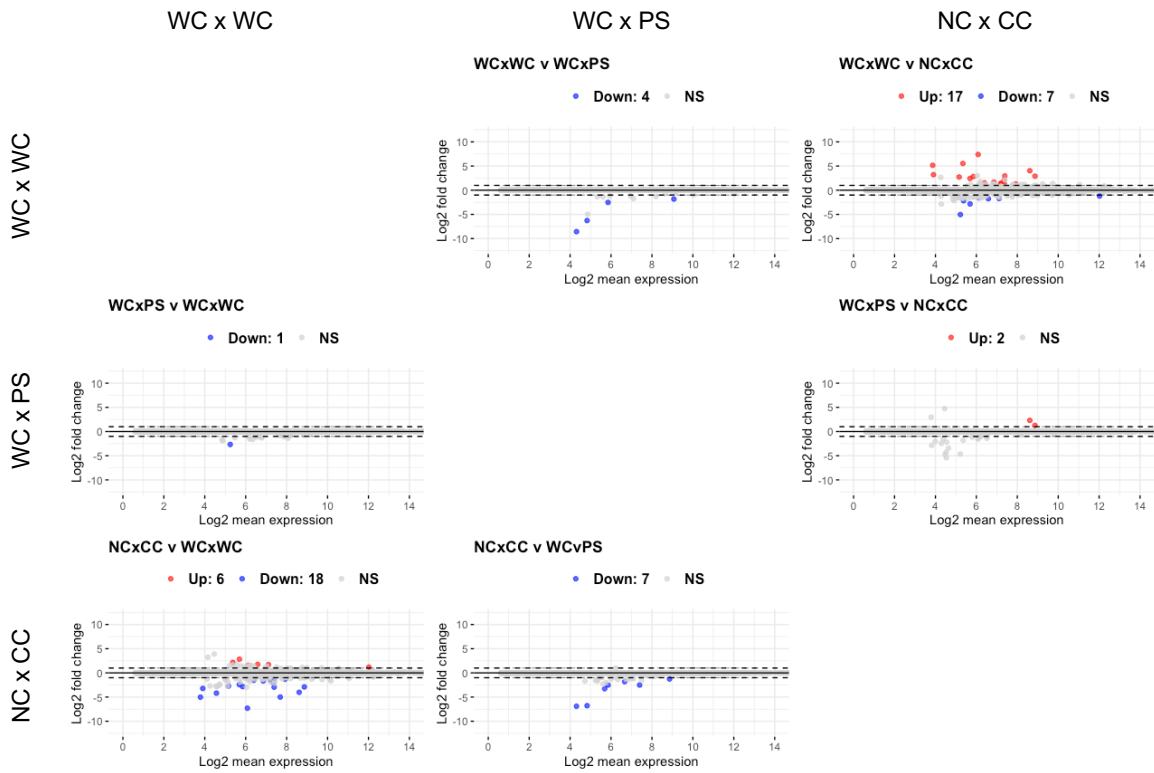


Figure 24. MA plot for juvenile clam lines interaction terms. The log₂ mean expression or log₂ mean of normalized counts are along the X-axis, and the log₂ fold changes are along the Y-axis for each plot. Those dots in red and blue represent genes that are significantly differentially expressed with an absolute LFC > 1 and an FDR < 0.05. Those dots in red are up-regulated in the non-control clam line, and those dots in blue are down-regulated in non-control clam. Those dots in grey are non-significant (NS). The clam line that is listed first in each comparison is the control clam line for that pair.

Literature Cited

- Abdelrahman, H., ElHady, M., Alcivar-Warren, A., Allen, S., Al-Tobasei, R., Bao, L., Beck, B., Blackburn, H., Bosworth, B., Buchanan, J., Chappell, J., Daniels, W., Dong, S., Dunham, R., Durland, E., Elaswad, A., Gomez-Chiarri, M., Gosh, K., Guo, X., ... Zhou, T. (2017). Aquaculture genomics, genetics and breeding in the United States: Current status, challenges, and priorities for future research. *BMC Genomics*, *18*(1).
<https://doi.org/10.1186/s12864-017-3557-1>
- Adams, C. J., Kopp, M. C., Larburu, N., Nowak, P. R., & Ali, M. M. (2019). Structure and molecular mechanism of ER stress signaling by the unfolded protein response signal activator IRE1. *Frontiers in Molecular Biosciences*, *6*.
<https://doi.org/10.3389/fmolb.2019.00011>
- Anaconda Software Distribution*. (2021). *Anaconda Documentation*. Anaconda Inc. Retrieved from <https://docs.anaconda.com/>
- Anderson, J.D., & Prosser, C.L. (1953). Osmoregulation capacity in populations occurring in different salinities (abstr.). *Biological Bulletin*, *105*, 369.
- Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Ansell, A.D. (1967). Egg production of *Mercenaria mercenaria*. *Limnology and Oceanography*, *12*, 172-176.
- Augusteyn, R. C. (2004). A-Crystallin: A review of its structure and function. *Clinical and Experimental Optometry*, *87*(6), 356–366. <https://doi.org/10.1111/j.1444-0938.2004.tb03095.x>

- Baker, S. M., Sturmer, L., Heuberger David, & Patrick Baker. (2005). Short-term effects of rapid salinity reduction on seed clams (*Mercenaria mercenaria*). *Journal of Shellfish Research*, 24(1), 29–33. [https://doi.org/10.2983/0730-8000\(2005\)24\[29:seorsr\]2.0.co;2](https://doi.org/10.2983/0730-8000(2005)24[29:seorsr]2.0.co;2)
- Ballantyne, J. S., & Chamberlin, M. E. (2020). Regulation of cellular amino acid levels. *Cellular and Molecular Physiology of Cell Volume Regulation*, 111–122. <https://doi.org/10.1201/9780367812140-8>
- Barrett, N. J., Thyrring, J., Harper, E. M., Sejr, M. K., Sørensen, J. G., Peck, L. S., & Clark, M. S. (2022). Molecular responses to thermal and osmotic stress in arctic intertidal mussels (*Mytilus edulis*): The limits of resilience. *Genes*, 13(1), 155. <https://doi.org/10.3390/genes13010155>
- Bayne, B. L. (1993). Feeding physiology of bivalves: Time-dependence and compensation for changes in food availability. *Bivalve Filter Feeders*, 1–24. https://doi.org/10.1007/978-3-642-78353-1_1
- Belding, D.L. (1931). The quahaug fishery of Massachusetts. Commonwealth of Massachusetts, Department of Conservation, Division of Fish and Game, Mar. Serv. 2. 41 pp.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Blighe, K., Rana, S., Lewis, M. (2022). EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. R package version 1.14.0, <<https://github.com/kevinblighe/EnhancedVolcano>>.

- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics*, btu170.
- Bricelj, V.M. & Malouf, R.E. (1980). Aspects of reproduction of hard clams (*Mercenaria mercenaria*) in Great South Bay, New York. *Proceedings of the National Shellfisheries Association*, 70, 216-229.
- Carregosa, V., Velez, C., Soares, A. M. V. M., Figueira, E., & Freitas, R. (2014). Physiological and biochemical responses of three Veneridae clams exposed to salinity changes. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 177–178, 1–9. <https://doi.org/10.1016/j.cbpb.2014.08.001>
- Carriker, M. R. (2001). Chapter 3 Embryogenesis and organogenesis of veligers and early juveniles. *Developments in Aquaculture and Fisheries Science*, 77–115. [https://doi.org/10.1016/s0167-9309\(01\)80031-4](https://doi.org/10.1016/s0167-9309(01)80031-4)
- Castagna, M. (2001). Chapter 15 Aquaculture of the hard clam, *Mercenaria mercenaria*. *Developments in Aquaculture and Fisheries Science*, 675–699. [https://doi.org/10.1016/s0167-9309\(01\)80043-0](https://doi.org/10.1016/s0167-9309(01)80043-0)
- Castagna M. & Chanley, P. (1973). Salinity tolerance of some marine bivalves from inshore and estuarine environments in Virginian waters on the western mid-Atlantic coast. *Malacologia*, 12, 47-96.
- Castagna, M., & Kraeuter, J. N. (1981) Manual for Growing the Hard Clam *Mercenaria*. Special Reports in Applied Marine Science and Ocean Engineering (SRAMSOE) No. 249. Virginia Institute of Marine Science, William & Mary. <https://doi.org/10.25773/8mh9-d326>

- Chanley, P.E. (1961). Inheritance of shell markings and growth in the hard clam, *Venus mercenaria*. *Proceedings of the National Shellfisheries Association*, 50, 163-169.
- Chen, H. (2022). VennDiagram: Generate High-Resolution Venn and Euler Plots. R package version 1.7.3, <<https://CRAN.R-project.org/package=VennDiagram>>.
- Cheng, Z., Teo, G., Krueger, S., Rock, T. M., Koh, H. W. L., Choi, H., & Vogel, C. (2016). Differential Dynamics of the mammalian mrna and protein expression response to misfolding stress. *Molecular Systems Biology*, 12(1), 855.
<https://doi.org/10.15252/msb.20156423>
- Chew, K. K. (2001). Chapter 16 introduction of the Hard Clam (*Mercenaria Mercenaria*) to the pacific coast of North America with notes on its introduction to Puerto Rico, England, and France. *Developments in Aquaculture and Fisheries Science*, 701–709.
[https://doi.org/10.1016/s0167-9309\(01\)80044-2](https://doi.org/10.1016/s0167-9309(01)80044-2)
- Cleveland, B. M., Gao, G., & Leeds, T. D. (2020). Transcriptomic response to selective breeding for fast growth in rainbow trout (*Oncorhynchus mykiss*). *Marine Biotechnology*, 22(4), 539–550. <https://doi.org/10.1007/s10126-020-09974-3>
- Dainat, J. (2020). AGAT: Another Gff Analysis Toolkit to handle annotations in any GTF/GFF format. (Version v0.8.0). Zenodo. <https://www.doi.org/10.5281/zenodo.3552717>
- Dalton, R. & Menzel, W. (1983). Seasonal gonadal development of young laboratory-spawned southern (*Mercenaria campechiensis*) and northern (*Mercenaria mercenaria*) quahogs and their reciprocal hybrids in northern Florida. *Journal of Shellfish Research*, 3, 11-17.

- Damgaard, R. B., & Gyrd-Hansen, M. (2011). Inhibitor of apoptosis (IAP) proteins in regulation of inflammation and innate immunity. *Discovery medicine*, *11*(58), 221–231.
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2). <https://doi.org/10.1093/gigascience/giab008>
- Davis, H.C. (1958). Survival and growth of clam and oyster at different salinities. *Biological Bulletin*, *114*, 296-307.
- Davis, H.C. & Calabrese, A. (1964). Combined effects of temperature and salinity on development of eggs and growth of larvae of *M. mercenaria* and *Q. virginica*. *U.S. Fish & Wildlife Service, Fishery Bulletin*, *63*, 643-655.
- Davis, H.C. & Chanley, E. E. (1956). Spawning and egg production of oysters and clams. *Biological Bulletin*, *110*, 117-128
- Dawson, D. & Liu, X. (2008). Osmoregulation. *Osmotic and Ionic Regulation*, 1–35. <https://doi.org/10.1201/9780849380525.ch1>
- DuPaul, W.D. & Webb, K.L. (1974). Salinity-induced changes in the alanine and aspartic aminotransferase activity in three marine bivalve molluscs. *Archives of Insect Biochemistry and Physiology*, *82*, 817-822.
- Durack, P. J., Wijffels, S. E., & Matear, R. J. (2012). Ocean salinities reveal strong global water cycle intensification during 1950 to 2000. *Science*, *336*(6080), 455–458. <https://doi.org/10.1126/science.1212222>

- Esri. (2022, September 12). Sea grant hard clam selective breeding collaborative. ArcGIS StoryMaps. Retrieved November 30, 2022, from <https://storymaps.arcgis.com/stories/3425623358164278bbe1ed7f7311a605>
- Eversole, A.G. (1989). Gametogenesis and spawning in North American clam populations: implications for culture. In: J.J. Manzi and M. Castagna, editors, *Clam Mariculture in North America*. Elsevier, 75-109.
- Eversole, A.G., Michener, W.K. & Eldridge, E.J. (1980). Reproductive cycle of *Mercenaria mercenaria* in a South Carolina estuary. *Proceedings of the National Shellfisheries Association*, 70, 20-30.
- Ewels, P., Magnusson, M., Lundin, S., & Källner, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32(19), 3047–3048. <https://doi.org/10.1093/bioinformatics/btw354>
- Fabbri, E., Valbonesi, P., Franzellitti, S. (2008). HSP expression in bivalves. *ISJ*, 5, 135-161
- Florkin, M. (1962). La régulation isosmotique intracellulaire chez les invertébrés marins euryhalins. *Academie Royale de Belgique. Bulletin de la Classe des Sciences*, 48, 687-694.
- Fujita, E., Okuma, E., & Abe, H. (1997a). Occurrence of alanine racemase in crustaceans and the changes of the properties during seawater acclimation of Crayfish. *Comparative Biochemistry and Physiology Part A: Physiology*, 116(1), 83–87. [https://doi.org/10.1016/s0300-9629\(96\)00120-x](https://doi.org/10.1016/s0300-9629(96)00120-x)
- Fujita, E., Okuma, E., & Abe, H. (1997b). Partial purification and properties of alanine racemase from the muscle of black tiger prawn *Penaeus monodon*. *Fisheries Science*, 63(3), 440–445. <https://doi.org/10.2331/fishsci.63.440>

- Fulda S. (2014). Regulation of cell migration, invasion and metastasis by IAP proteins and their antagonists. *Oncogene*, 33(6), 671–676. <https://doi.org/10.1038/onc.2013.63>
- García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L. M., Götz, S., Tarazona, S., Dopazo, J., Meyer, T. F., & Conesa, A. (2012). Qualimap: Evaluating next-generation sequencing alignment data. *Bioinformatics*, 28(20), 2678–2679.
<https://doi.org/10.1093/bioinformatics/bts503>
- Gilles, R. (1987). Volume regulation in cells of euryhaline invertebrates. *Current Topics in Membranes and Transport*, 30, 205-247.
- Grizzle, R. E., Bricelj, V. M., & Shumway, S. E. (2001). Chapter 8 Physiological Ecology of *Mercenaria Mercenaria*. *Developments in Aquaculture and Fisheries Science*, 305–382.
[https://doi.org/10.1016/s0167-9309\(01\)80036-3](https://doi.org/10.1016/s0167-9309(01)80036-3)
- Gyrd-Hansen, M., & Meier, P. (2010). IAPs: from caspase inhibitors to modulators of NF-kappaB, inflammation and cancer. *Nature reviews. Cancer*, 10(8), 561–574.
<https://doi.org/10.1038/nrc2889>
- Hadley, N. H., Manzi, J. J., Eversole, A. G., Dillon, R. T., Battey, C. E., & Peacock, N. M. (1997). A Manual for the culture of the hard clam *Mercenaria* spp. in South Carolina. S.C. Sea Grant Consortium.
- Hamwi, A. (1969). Oxygen consumption and pumping rate of the hard clam *Mercenaria mercenaria*. L. Ph.D. Dissertation, Rutgers University, New Brunswick, NJ, 177 pp.
- Hanna, G.D.(1966). Introduced mollusks of western North America. *Occasional Papers of the California Academy of Sciences*, 48, 108.

- Harte, M. E. (2001). Chapter 1 Systematics and taxonomy. *Developments in Aquaculture and Fisheries Science*, 3–51. [https://doi.org/10.1016/s0167-9309\(01\)80029-6](https://doi.org/10.1016/s0167-9309(01)80029-6)
- Hartl, F.U., Bracher, A., Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. *Nature*. 475 (7356), 324–332.
- Heppell, D. (1961). The naturalization in Europe of the quahog *Mercenaria mercenaria*. *Journal of Conchology*, 25, 21-34.
- Hesselman, D.M., Barber, B.J. & Blake, N.J. (1989). The reproductive cycle of adult hard clams, *Mercenaria* spp., the Indian River lagoon, Florida. *Journal of Shellfish Research*, 8, 43-49.
- Hollenbeck, C. M., & Johnston, I. A. (2018). Genomic tools and selective breeding in molluscs. *Frontiers in Genetics*, 9. <https://doi.org/10.3389/fgene.2018.00253>
- Houston, R. D., Bean, T. P., Macqueen, D. J., Gundappa, M. K., Jin, Y. H., Jenkins, T. L., Selly, S. L., Martin, S. A., Stevens, J. R., Santos, E. M., Davie, A., & Robledo, D. (2020). Harnessing genomics to fast-track genetic improvement in Aquaculture. *Nature Reviews Genetics*, 21(7), 389–409. <https://doi.org/10.1038/s41576-020-0227-y>
- Hu, Z., Song, H., Feng, J., Zhou, C., Yang, M.-J., Shi, P., Yu, Z.-L., Li, Y.-R., Guo, Y.-J., Li, H.-Z., & Zhang, T. (2022). Massive heat shock protein 70 genes expansion and transcriptional signatures uncover hard clam adaptations to heat and hypoxia. *Frontiers in Marine Science*, 9. <https://doi.org/10.3389/fmars.2022.898669>
- Hudson, K., & Virginia Sea Grant Marine Advisory Program. (2019) Virginia Shellfish Aquaculture Situation and Outlook Report: Results of the 2018 Virginia Shellfish Aquaculture Crop Reporting Survey. Marine Resource Report No. 2019-8; Virginia Sea

Grant VSG-19-03. Virginia Institute of Marine Science, William & Mary.

<https://doi.org/10.25773/jc19-y847>

Kanehisa, M., Sato, Y., & Kawashima, M. (2021). KEGG mapping tools for uncovering hidden features in biological data. *Protein Science*, 31(1), 47–53. <https://doi.org/10.1002/pro.4172>

Kanehisa, M., Sato, Y., & Morishima, K. (2016). Blastkoala and Ghostkoala: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of Molecular Biology*, 428(4), 726–731. <https://doi.org/10.1016/j.jmb.2015.11.006>

Karl, T.R. & Trenberth, K.E. (2003). Modern global climate change. *Science* 302, 1719–1723.

Kassambara A (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0, <<https://CRAN.R-project.org/package=ggpubr>>.

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with Hisat2 and HISAT-genotype. *Nature Biotechnology*, 37(8), 907–915. <https://doi.org/10.1038/s41587-019-0201-4>

Knaub, R.S. & Eversole, A.G. (1988). Reproduction of different stocks of *Mercenaria mercenaria*. *Journal of Shellfish Research*, 7, 371-376.

Koussounadis, A., Langdon, S. P., Um, I. H., Harrison, D. J., & Smith, V. A. (2015). Relationship between differentially expressed mrna and mrna-protein correlations in a xenograft model system. *Scientific Reports*, 5(1). <https://doi.org/10.1038/srep10775>

Knowles, J. R. (1980). Enzyme-catalyzed phosphoryl transfer reactions. *Annual Review of Biochemistry*, 49(1), 877–919. <https://doi.org/10.1146/annurev.bi.49.070180.004305>

- Ktitz, D. (2000). Cell and molecular response to stress : Environmental stressors and gene responses. (Vol. 1, pp. 157–179; K. B. Storey & J. Storey, Eds.). Retrieved from <https://www.sciencedirect.com/science/article/pii/S1568125400800145>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Loosanoff, V.L. (1936). Sexual phases in the quahog. *Science*, 83, 287-288.
- Loosanoff, V.L. (1937). Development of the primary gonad and sexual phases in *Venus mercenaria* Linneaus. *Biological Bulletin*, 72, 389-405.
- Loosanoff, V.L. & Davis, H.C. (1963). Rearing of bivalve mollusks. *Advances in Marine Biology*, 1, 1-136.
- Love, M.I., Huber, W., & Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15:550. [10.1186/s13059-014-0550-8](https://doi.org/10.1186/s13059-014-0550-8)
- Low, W. P., Ong, W. T., & Ip, Y. K. (1996). Different physiological functions of free D- and L-alanine in three body parts of the intertidal sipunculid *Phascolosoma arcuatum*. *Journal of Comparative Physiology B*, 165(7), 558–564. <https://doi.org/10.1007/bf00387517>
- Luo, W., & Brouwer, C. (2013). Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics*, 29(14): 1830-1831, doi: [10.1093/bioinformatics/btt285](https://doi.org/10.1093/bioinformatics/btt285)

- Matsushima, O., Katayama, H., Yamada, K., Kado, Y. (1984). Occurrence of free D-alanine and alanine racemase activity in bivalve molluscs with special reference to intracellular osmoregulation. *Marine Biology Letters*, 5, 217–225.
- McClary, M. (2014). Osmoconformer. Retrieved from [http://editors.eol.org/eoearth/wiki/Osmoconformer_\(Ecology\)](http://editors.eol.org/eoearth/wiki/Osmoconformer_(Ecology))
- Narberhaus, F. (2002). A-crystallin-type heat shock proteins: Socializing minichaperones in the context of a multichaperone network. *Microbiology and Molecular Biology Reviews*, 66(1), 64–93. <https://doi.org/10.1128/membr.66.1.64-93.2002>
- Nie, H., Jiang, L., Chen, P., Huo, Z., Yang, F., & Yan, X. (2017a). High throughput sequencing of RNA transcriptomes in *Ruditapes philippinarum* identifies genes involved in osmotic stress response. *Scientific Reports*, 7(1). <https://doi.org/10.1038/s41598-017-05397-8>
- Nie, H., Liu, L., Huo, Z., Chen, P., Ding, J., Yang, F., & Yan, X. (2017b). The HSP70 gene expression responses to thermal and salinity stress in wild and cultivated manila clam *Ruditapes philippinarum*. *Aquaculture*, 470, 149–156. <https://doi.org/10.1016/j.aquaculture.2016.12.016>
- Niu, D., Wang, L., Sun, F., Liu, Z., & Li, J. (2013). Development of molecular resources for an intertidal clam, *Sinonovacula constricta*, using 454 transcriptome sequencing. *PLoS ONE*, 8(7). <https://doi.org/10.1371/journal.pone.0067456>
- Nomura, T., Yamamoto, I., Morishita, F., Furukawa, Y., & Matsushima, O. (2001). Purification and some properties of alanine racemase from a bivalve *Molluscorbicula japonica*. *Journal of Experimental Zoology*, 289(1), 1–9. [https://doi.org/10.1002/1097-010x\(20010101/31\)289:1<1::aid-jez1>3.0.co;2-m](https://doi.org/10.1002/1097-010x(20010101/31)289:1<1::aid-jez1>3.0.co;2-m)

- Okuma, E., Watanabe, K., & Abe, H. (1998). Distribution of free D-amino acids in bivalve mollusks and the effects of physiological conditions on the levels of D- and L-alanine in the tissues of the hard clam, *meretrix lusoria*. *Fisheries Science*, *64*(4), 606–611.
<https://doi.org/10.2331/fishsci.64.606>
- Olsen, J. V., Blagoev, B., Gnäd, F., Macek, B., Kumar, C., Mortensen, P., & Mann, M. (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell*, *127*(3), 635–648. <https://doi.org/10.1016/j.cell.2006.09.026>
- Otto, S.V. (1973). Hermaphroditism in two species of pelecypod mollusks. *Proceedings of the National Shellfisheries Association*, *63*, 96-98.
- Palmer, K. V. W. (1927). In *The Veneridae of Eastern America; Cenozoic and recent*. (Vol. 1, pp. 209–522). essay, Palaeontographica Americana.
- Peterson, C.H. (1983). A concept of quantitative reproductive senility: application to the hard clam, *Mercenaria mercenaria* (L.). *Oecologia*, *58*, 164-168.
- Peterson, C.H. (1986.) Quantitative allometry of gamete production by *Mercenaria mercenaria* into old age. *The Marine Ecology Progress Series*, *29*, 93-97
- Pierce, S. K. (1982). Invertebrate cell volume control mechanisms: a coordinated use of intracellular amino acids and inorganic ions as osmotic solute. *Biological Bulletin*, *163*, 405-419.
- Portt, L., Norman, G., Clapp, C., Greenwood, M., & Greenwood, M. T. (2011). Anti-apoptosis and cell survival: A Review. *Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research*, *1813*(1), 238–259. <https://doi.org/10.1016/j.bbamcr.2010.10.010>

- Putri, G., Anders, S., Pyl, P.T., Pimanda, J.E., & Zanini, F. (2021). Analysing high-throughput sequencing data with HTSeq 2.0. [Online]. Available online at [arXiv:2112.00939](https://arxiv.org/abs/2112.00939)
- Qian, X., Ba, Y., Zhuang, Q., & Zhong, G. (2014). RNA-Seq Technology and Its Application in Fish Transcriptomics. *OMICS: A Journal of Integrative Biology*, 18(2), 98–110. <https://doi.org/10.1089/omi.2013.0110>
- Quinn, E. M., Cormican, P., Kenny, E. M., Hill, M., Anney, R., Gill, M., ... Morris, D. W. (2013). Development of strategies for SNP detection in RNA-Seq data: Application to lymphoblastoid cell lines and evaluation using 1000 genomes data. *PLoS ONE*, 8(3), e58815. <https://doi.org/10.1371/journal.pone.0058815>
- Rivera-Ingraham, G. A., & Lignot, J.-H. (2017). Osmoregulation, bioenergetics and oxidative stress in coastal marine invertebrates: raising the questions for future research. *The Journal of Experimental Biology*, 220(10), 1749–1760. <https://doi.org/10.1242/jeb.135624>
- Ropp, A. J., Reece, K. S., Snyder, R. A., Song, J., Biesack, E. E., McDowell, J. R. (under review). Fine-scale population structure of the northern hard clam (*Mercenaria mercenaria*) revealed by genome-wide SNP markers. *Evolutionary Applications*.
- Roegner, G., & Mann, R. L. (1990). Habitat requirements for the hard clam, *Mercenaria mercenaria*, in Chesapeake Bay. *Special scientific report* No. 126. Virginia Institute of Marine Science, College of William & Mary. <http://dx.doi.org/doi:10.21220/m2-1rar-wr63>
- Sae-Lim, P., Kause, A., Mulder, H. A., & Olesen, I. (2017). Breeding and genetics symposium: Climate change and selective breeding in aquaculture. *Journal of Animal Science*, 95(4), 1801. <https://doi.org/10.2527/jas2016.1066>

- Salem, M., Vallejo, R. L., Leeds, T. D., Palti, Y., Liu, S., Sabbagh, A., ... Yao, J. (2012). RNA-Seq identifies SNP markers for growth traits in rainbow trout. *PLoS ONE*, 7(5), e36264. <https://doi.org/10.1371/journal.pone.0036264>
- Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7), 671–675. [doi:10.1038/nmeth.2089](https://doi.org/10.1038/nmeth.2089)
- Shumway, S.E., Gabbott, P.A. & Youngson, A., 1977. The effect of fluctuating salinity on the concentrations of free amino acids and ninhydrin-positive substances in the adductor muscles of eight species of bivalve molluscs. *The Journal of Experimental Marine Biology and Ecology*, 29, 131-150.
- Slowikowski, K. (2021). ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'. R package version 0.9.1, <<https://CRAN.R-project.org/package=ggrepel>>.
- Snyder, R. A. (2021) Fisheries Landings for the Eastern Shore of Virginia (ESVA) 2021. VIMS Eastern Shore Laboratory Technical Report No. 10. Virginia Institute of Marine Science, William & Mary. <https://doi.org/10.25773/jxkh-b683>
- Song, H., Guo, X., Sun, L., Wang, Q., Han, F., Wang, H., Wray, G. A., Davidson, P., Wang, Q., Hu, Z., Zhou, C., Yu, Z., Yang, M., Feng, J., Shi, P., Zhou, Y., Zhang, L., & Zhang, T. (2021). The hard clam genome reveals massive expansion and diversification of inhibitors of apoptosis in Bivalvia. *BMC Biology*, 19(1). <https://doi.org/10.1186/s12915-020-00943-9>
- Song, J., Austin, J. D., & Yang, H. (2022). Comparative transcriptomics of the northern quahog *Mercenaria mercenaria* and southern Quahog *Mercenaria campechiensis* in response to

chronic heat stress. *Marine Biotechnology*, 24(2), 276–292.

<https://doi.org/10.1007/s10126-022-10101-7>

Stanley, J.G. & DeWitt, R. (1983). Species profiles: life histories and environmental requirements of coastal fishes and invertebrates (North Atlantic) -- hard clam. U.S. Fish Wildl. Serv., FWS/OBS-82/11.18. U.S. Army Corps of Engr., TP EL-82-4, 19 pp.

Townes, F.W., Hicks, S.C., Aryee, M.J., and Irizarry, R.A. (2019). Feature selection and dimension reduction for single cell RNA-Seq based on a multinomial model. *Genome Biology* <https://doi.org/10.1186/s13059-019-1861-6>

Virkki, L. V., Biber, J., Murer, H., & Forster, I. C. (2007). Phosphate transporters: A tale of two solute carrier families. *American Journal of Physiology-Renal Physiology*, 293(3).

<https://doi.org/10.1152/ajprenal.00228.2007>

Walker, R.L. (1994). Reproductive ecology of the northern quahog, *Mercenaria mercenaria* (Linnaeus, 1758) in coastal Georgia and its resource management implications. Ph.D. Dissertation, University of Georgia, Athens, 184 pp.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., ... Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), 1686.

<https://doi.org/10.21105/joss.01686>

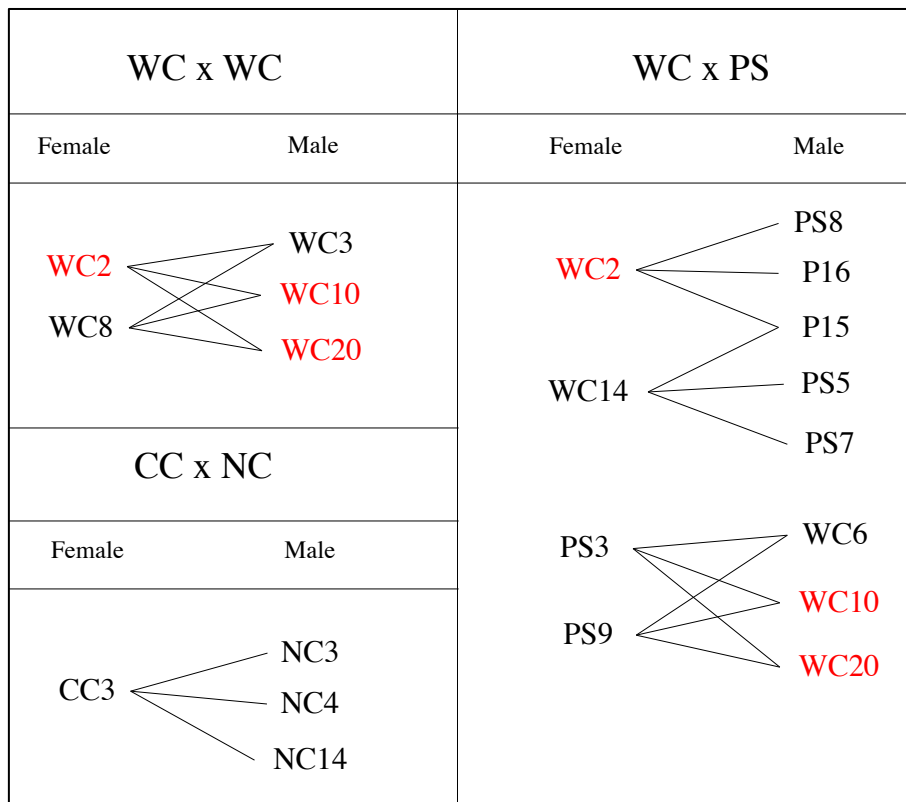
Winter, D. J. (2017). rentrez: an R package for the NCBI eUtils API. *The R Journal*, 9(2), 520-526

- Yamada, A., & Matsushima, O. (1992). The relation of D-alanine and alanine racemase activity in molluscs. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, 103(3), 617–621. [https://doi.org/10.1016/0305-0491\(92\)90379-6](https://doi.org/10.1016/0305-0491(92)90379-6)
- Yancey, P. H. (2020). Compatible and counteracting solutes. *Cellular and Molecular Physiology of Cell Volume Regulation*, 81–109. <https://doi.org/10.1201/9780367812140-7>
- Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D., & Somero, G. N. (1982). Living with water stress: *Evolution of Osmolyte Systems*. *Science*, 217(4566), 1214–1222. <https://doi.org/10.1126/science.7112124>
- Zhang, F., Hu, B., Fu, H., Jiao, Z., Li, Q., & Liu, S. (2019). Comparative transcriptome analysis reveals molecular basis underlying fast growth of the selectively bred Pacific oyster, *Crassostrea gigas*. *Frontiers in Genetics*, 10. <https://doi.org/10.3389/fgene.2019.0061>
- Zhu, A., Ibrahim, J. G., & Love, M. I. (2018). Heavy-tailed prior distributions for sequence count data: Removing the noise and preserving large differences. *Bioinformatics*, 35(12), 2084–2092. <https://doi.org/10.1093/bioinformatics/bty895>

Appendix: Supplementary Tables and Figures

Clam Line	# of clams induced	# of females observed	# of males observed
Wachapreague	16	1	2
Pocomoke	20	0	4
Cape Cod	26	3	3
Mobjack	21	1	5
New Jersey	27	2	5

Supplementary Table 1. 2019 Clam Spawn Parental Information: This table includes the number of individuals that were induced to spawn for each line and the number of female and male clams observed spawning. All spawns produced fertilized eggs, so observed number of spawning individuals may be under representative of the actual number of individuals that contributed gametes.



Supplementary Figure 1.2021 Clam Spawn Parental Information: This figure includes parental clams used in the 2021 spawns. Those clams in red were used in more than one cross. WC = Wachapreague Channel, VA, PS = Pocomoke Sound, VA, CC = Cape Cod, Massachusetts, and NC = Bogue Sound, NC.

No.	Clam Line	Year	Rep.	Salinity	SH (mm)	Sample Name	Conc. (ng/ul)	Vol. (ul)	Total (ug)	RIN	QC
1	Wachapreague	2019	1	35 ppt	32.90	A 1 32	131.53	27	3.55	7.1	Pass
2	Wachapreague	2019	2	35 ppt	31.10	A 2 32	123.29	27	3.33	6.9	Pass
3	Wachapreague	2019	3	35 ppt	23.80	A 3 32	113.09	28	3.17	7.5	Pass
4	Pocomoke	2019	1	35 ppt	24.90	B 1 32	115.01	28	3.22	7.6	Pass
5	Pocomoke	2019	2	35 ppt	29.80	B 2 32	118.12	28	3.31	8.4	Pass
6	Pocomoke	2019	3	35 ppt	22.00	B 3 32	125.41	28	3.51	6.9	Pass
7	Cape Cod	2019	1	35 ppt	32.60	C 1 32	117.35	28	3.29	5.4	Pass
8	Cape Cod	2019	2	35 ppt	34.20	C 2 32	102.79	28	2.88	7.6	Pass
9	Cape Cod	2019	3	35 ppt	25.60	C 3 32	139.33	26	3.62	7.5	Pass
10	Mobjack	2019	1	35 ppt	40.70	D 1 32	143.45	27	3.87	7.6	Pass
11	Mobjack	2019	2	35 ppt	36.60	D 2 32	129.40	28	3.62	6.6	Pass
12	Mobjack	2019	3	35 ppt	29.90	D 3 32	146.57	28	4.10	8.2	Pass
13	New Jersey	2019	1	35 ppt	31.30	E 1 32	95.53	27	2.58	5.8	Pass
14	New Jersey	2019	2	35 ppt	25.00	E 2 32	101.29	28	2.84	7.4	Pass
15	New Jersey	2019	3	35 ppt	19.10	E 3 32	122.31	27	3.30	8.1	Pass
16	WC x WC	2021	1	35 ppt	5.13 ± 0.59	F 1 32	94.43	28	2.64	4.2	Pass
17	WC x WC	2021	2	35 ppt	4.83 ± 0.40	F 2 32	98.03	27	2.65	7.1	Pass
18	WC x WC	2021	3	35 ppt	5.16 ± 0.48	F 3 32	112.30	27	3.03	7.5	Pass
19	WC x PS	2021	1	35 ppt	5.14 ± 0.29	G 1 32	96.21	28	2.69	8.3	Pass
20	WC x PS	2021	2	35 ppt	5.04 ± 0.36	G 2 32	106.99	27	2.89	8.5	Pass
21	WC x PS	2021	3	35 ppt	5.00 ± 0.26	G 3 32	131.19	27	3.54	4.3	Pass
22	NC x CC	2021	1	35 ppt	5.20 ± 0.60	H 1 32	141.17	27	3.81	6.5	Pass
23	NC x CC	2021	2	35 ppt	5.32 ± 0.58	H 2 32	173.89	27	4.69	8.3	Pass
24	NC x CC	2021	3	35 ppt	5.17 ± 0.64	H 3 32	164.24	27	4.43	8.6	Pass

Supplementary Table 2. 35 ppt hard clam sample information: This table includes each adult and juvenile line, the year they were created, the replicate number each represents, the salinity experiment that was performed with these samples, the SH or average SH, the sample name they were given when sent off to Novogene, the sample volume Novogene received, the total RNA concentration Novogene calculated, the RNA integrity number (RIN) that Novogene calculated, and whether each sample passed the initial quality checked (based on RIN and concentration) performed by Novogene.

No.	Clam Line	Year	Rep.	Salinity	SH (mm)	Sample Name	Conc. (ng/ul)	Vol. (ul)	Total (ug)	RIN	QC
25	Wachapreague	2019	1	15 ppt	27.30	A 1 15	85.98	28	2.40737	8.6	Pass
26	Wachapreague	2019	2	15 ppt	26.40	A 2 15	104.95	27	2.83353	8.4	Pass
27	Wachapreague	2019	3	15 ppt	26.10	A 3 15	102.1	27	2.75677	8.7	Pass
28	Pocomoke	2019	1	15 ppt	27.10	B 1 15	96.75	27	2.61233	8.6	Pass
29	Pocomoke	2019	2	15 ppt	24.80	B 2 15	94.45	26	2.45573	8.9	Pass
30	Pocomoke	2019	3	15 ppt	24.30	B 3 15	106.3	27	2.8701	9.1	Pass
31	Cape Cod	2019	1	15 ppt	38.00	C 1 15	109.78	27	2.96394	8.9	Pass
32	Cape Cod	2019	2	15 ppt	33.10	C 2 15	97.66	28	2.7346	9	Pass
33	Cape Cod	2019	3	15 ppt	28.10	C 3 15	122.9	27	3.31827	8.7	Pass
34	Mobjack	2019	1	15 ppt	41.80	D 1 15	127.47	28	3.56915	8.4	Pass
35	Mobjack	2019	2	15 ppt	44.70	D 2 15	136.55	27	3.68695	8.6	Pass
36	Mobjack	2019	3	15 ppt	34.80	D 3 15	155.9	27	4.20923	7.2	Pass
37	New Jersey	2019	1	15 ppt	30.70	E 1 15	87.35	28	2.44577	7.2	Pass
38	New Jersey	2019	2	15 ppt	29.50	E 2 15	99.51	27	2.68678	7.4	Pass
39	New Jersey	2019	3	15 ppt	25.60	E 3 15	88.25	27	2.38282	7.4	Pass
40	WC x WC	2021	1	15 ppt	5.38 ± 0.65	F 1 15	74.13	27	2.00156	7.7	Pass
41	WC x WC	2021	2	15 ppt	5.26 ± 0.58	F 2 15	78.54	27	2.12061	6.9	Pass
42	WC x WC	2021	3	15 ppt	5.61 ± 0.77	F 3 15	93.98	26	2.44335	7.2	Pass
43	WC x PS	2021	1	15 ppt	5.74 ± 0.84	G 1 15	87.23	28	2.44231	6.9	Pass
44	WC x PS	2021	2	15 ppt	5.18 ± 0.68	G 2 15	121.51	27	3.28074	7.7	Pass
45	WC x PS	2021	3	15 ppt	5.46 ± 0.51	G 3 15	100.79	27	2.72132	7.5	Pass
46	NC x CC	2021	1	15 ppt	5.77 ± 0.63	H 1 15	86.25	28	2.41489	3.4	Pass
47	NC x CC	2021	2	15 ppt	5.65 ± 0.50	H 2 15	104.65	28	2.93019	7.1	Pass
48	NC x CC	2021	3	15 ppt	5.44 ± 0.91	H 3 15	113.51	26	2.95117	7.4	Pass

Supplementary Table 3. 15 ppt hard clam sample information: This table includes each adult and juvenile line, the year they were created, the replicate number each represents, the salinity experiment that was performed with these samples, the SH or average SH, the sample name they were given when sent off to Novogene, the sample volume Novogene received, the total RNA concentration Novogene calculated, the RNA integrity number (RIN) that Novogene calculated, and whether each sample passed the initial quality checked (based on RIN and concentration) performed by Novogene.

Gene	Name	Base Mean	LFC	LFC SE	pvalue	padj
XM_045350540.1	uncharacterized LOC123558662	1454.546167	1.724807157	0.158667645	8.85E-29	7.54E-25
XM_045318306.1	uncharacterized LOC123535595, transcript variant X1	33.36191262	4.950983597	0.494955503	1.15E-24	6.51E-21
XM_045318307.1	uncharacterized LOC123535595, transcript variant X2	100.3642367	5.037707166	0.534501064	2.15E-22	9.14E-19
XM_045306594.1	uncharacterized LOC123527250	44.8123136	1.888198414	0.207108031	3.77E-21	1.28E-17
XM_045306916.1	alpha-crystallin A chain-like LOC123527454	967.5722041	2.384146675	0.270193385	5.79E-20	1.65E-16
XM_045359076.1	uncharacterized LOC123565067	17.28941208	6.251458035	0.743567593	7.40E-20	1.80E-16
XM_045335801.1	phosphoenolpyruvate carboxykinase, cytosolic [GTP]-like (LOC123548500), transcript variant X2	750.8122234	5.219203638	0.59436961	8.47E-20	1.81E-16
XM_045360456.1	heat shock protein beta-1-like (LOC123566409)	313.6829135	1.997567209	0.231494767	2.97E-19	5.63E-16
XM_045358251.1	marginal zone B- and B1-cell-specific protein-like (LOC123564575)	185.7804568	2.036149072	0.241911603	1.88E-18	3.20E-15
XM_045309959.1	uncharacterized LOC123529580	360.0719443	1.795565384	0.221864222	2.81E-17	3.76E-14
XM_045329467.1	heat shock protein HSP 90-beta-like (LOC123543395)	7009.519588	1.752372951	0.215588701	2.47E-17	3.76E-14
XM_045345110.1	serine-arginine protein 55-like (LOC123554775)	1261.482487	-1.417809782	0.175459315	2.87E-17	3.76E-14
XM_045344379.1	steroid 17-alpha-hydroxylase/17,20 lyase-like (LOC123554313)	1115.576467	-5.357123266	0.67366018	6.65E-17	8.09E-14
XM_045359836.1	protein disulfide-isomerase A3-like (LOC123566019)	3369.271394	1.586141711	0.198956546	8.77E-17	9.96E-14
XM_045316636.1	inositol-3-phosphate synthase 1-A-like (LOC123534401)	104.9794979	4.23730962	0.533941758	9.39E-17	1.00E-13
XM_045336317.1	uncharacterized LOC123548780	163.7805683	1.879883934	0.237059237	1.08E-16	1.09E-13
XM_045303941.1	alreticulin-like (LOC123525149)	4148.547277	1.934075756	0.245380251	1.71E-16	1.62E-13
XR_006683673.1	uncharacterized LOC123537950	230.7745968	-2.678636402	0.343884658	2.75E-16	2.47E-13
XM_045318661.1	uncharacterized LOC123535900	140.5006636	1.373440179	0.176649254	3.61E-16	3.07E-13
XM_045322411.1	CTD small phosphatase-like protein (LOC123538366)	49.30710418	1.529093227	0.203718873	3.08E-15	2.50E-12
XM_045304986.1	uncharacterized LOC123525831	2884.109855	-1.447008323	0.198298405	1.29E-14	9.59E-12
XM_045320928.1	endoplasmic reticulum chaperone BiP-like (LOC123537239)	4075.910935	2.218561432	0.304601356	1.76E-14	1.15E-11
XM_045307879.1	calcium-independent protein kinase C-like (LOC123528127)	26.21216795	2.111465093	0.291013799	1.95E-14	1.23E-11
XM_045312123.1	protein disulfide-isomerase A4-like (LOC123531297)	610.079902	2.024668903	0.28387345	4.59E-14	2.70E-11
XM_045335949.1	uncharacterized LOC123548578, transcript variant X3	168.1408532	-1.008469626	0.141517625	4.50E-14	2.70E-11
XM_045337220.1	glycogen-binding subunit 76A-like (LOC123549278)	223.096539	1.891950919	0.267973076	8.02E-14	4.56E-11
XM_045309308.1	heat shock protein 70 B2-like (LOC123529101)	41.15425451	4.668582579	0.669944314	1.51E-13	8.03E-11
XM_045353177.1	fibrillin-1-like (LOC123561027)	275.2406584	-1.431956532	0.206338043	1.72E-13	8.63E-11
XR_006685904.1	uncharacterized LOC123548540	368.7841425	-3.077279875	0.452858016	4.25E-13	2.07E-10
XM_045316046.1	uncharacterized LOC123534014	64.67545813	-3.315925778	0.490743991	5.59E-13	2.65E-10
XM_045315215.1	enolase-phosphatase E1-like (LOC123533550)	124.9643523	1.654862993	0.245236806	6.78E-13	3.12E-10
XM_045308816.1	haloacid dehalogenase-like hydrolase domain-containing 5 (LOC123528809), transcript variant X2	18.86171815	4.19367415	0.625433126	8.34E-13	3.74E-10
XM_045360275.1	serine-arginine protein 55-like (LOC123566309)	1174.364323	-1.277948028	0.191202325	8.87E-13	3.88E-10

XM_045309466.1	heat shock protein 70 B2-like LOC123529228	77.6202681	5.224210581	0.787847252	1.44E-12	6.00E-10
XM_045347940.1	rho guanine nucleotide exchange factor 17- like (LOC123556880),	23.73259684	1.272242414	0.192713138	1.99E-12	8.09E-10
XM_045309476.1	heat shock protein 70 B2-like LOC123529238	74.04678099	5.434475856	0.842673615	4.98E-12	1.97E-09
XM_045343782.1	dnaJ homolog subfamily C member 3-like (LOC123553978)	1020.026585	1.158814359	0.182291317	8.92E-12	3.45E-09
XM_045307354.1	uncharacterized LOC123527717	28.627238	2.202134025	0.346876331	1.05E-11	3.96E-09
XM_045312172.1	acidic mammalian chitinase-like (LOC123531325)	113.1065394	-1.113380435	0.176179984	1.10E-11	4.08E-09
XM_045315073.1	organic cation transporter protein-like (LOC123533435)	272.3846618	-1.630886106	0.258815016	1.23E-11	4.46E-09
XM_045314163.1	sarcoplasmic reticulum histidine-rich calcium-binding protein-like (LOC123532646)	2126.189993	1.413030892	0.22447104	1.52E-11	5.41E-09
XM_045300579.1	protein canopy homolog 2-like (LOC123523000)	453.6410943	1.271368233	0.202754874	1.68E-11	5.83E-09
XM_045339856.1	orkhead box protein O-like (LOC123551138)	67.27044447	2.337123787	0.375041433	2.13E-11	7.25E-09
XM_045310551.1	cysteine-rich with EGF-like domain protein 2 (LOC123529941)	31.78923478	1.331077862	0.213338478	2.23E-11	7.45E-09
XM_045356558.1	splicing factor U2AF 50 kDa subunit-like (LOC123536344)	427.0072759	-1.009861716	0.162909151	2.42E-11	7.92E-09
XM_045360845.1	alpha/beta hydrolase domain-containing protein 17B-like (LOC123566591)	137.6601561	1.09680588	0.177478641	2.84E-11	9.13E-09
XM_045346781.1	baculoviral IAP repeat-containing protein 7- A-like (LOC123556206)	1361.137578	4.673170096	0.758414075	3.35E-11	1.06E-08
XM_045351808.1	uncharacterized LOC123559747 (LOC123559747)	13.54854521	2.148817708	0.349779797	4.47E-11	1.39E-08
XM_045339804.1	headcase protein-like (LOC123551102)	1040.070317	1.962769813	0.322709516	5.36E-11	1.63E-08
XM_045303624.1	transmembrane prolyl 4-hydroxylase-like (LOC123524972)	44.65090488	1.223691389	0.202823749	7.23E-11	2.16E-08

Supplementary Table 4. Top 50 genes by ascending FDR or adjusted p-value (padj) between 35 and 15 ppt for adult samples. Those genes in red are in the top 15 for lowest FDR. Those genes in blue are in the top 15 for largest absolute LFC. Those genes in purple are in both the top 15 for FDR and LFC. Those genes in grey are uncharacterized. Gene is the gene accession number on NCBI. Name is the name given for each feature on NCBI. Base mean is the average of the normalized count values, divided by the size factor, taken over all the samples. LFC is the log₂ fold change. LFC SE is the LFC standard error. The pvalue is the Wald test p-value, where the null hypothesis that there is no effect of treatment on the gene. The padj is the adjusted p-value based on the Benjamini-Hochberg method, or false discovery rate (FDR).

Gene	Name	Base Mean	LFC	LFC SE	pvalue	padj
XM_045318505.1	sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1-like LOC123535755	324.9447929	5.049886016	0.335112018	1.59E-52	2.76E-48
XM_045351851.1	sodium-dependent phosphate transport protein 2B-like LOC123559775	618.4400228	3.742419808	0.253368846	1.20E-50	1.04E-46
XM_045360561.1	solute carrier family 12 member 3-like LOC123566450	242.2277978	2.745805529	0.227974773	1.03E-34	5.95E-31
XM_045327875.1	protein PIF-like LOC123542176	390.2895119	4.968425591	0.414554956	2.35E-34	1.02E-30
XM_045354440.1	homeobox protein engrailed-2b-like LOC123561826	64.86771431	4.369365887	0.390775795	3.99E-30	1.38E-26
XM_045337089.1	immediate early response gene 5-like protein LOC123549199	3419.881254	1.618695542	0.15536069	1.05E-26	3.03E-23
LOC123549369	LOC123549369	991.5596508	1.384430444	0.133415189	1.51E-26	3.75E-23
XM_045333072.1	DBH-like monooxygenase protein 2 homolog LOC123546627	244.3909213	2.164166038	0.212763087	1.27E-25	2.76E-22
XM_045320073.1	retinol dehydrogenase 7-like LOC123536703	751.2496527	-2.157683204	0.214645274	4.00E-25	7.70E-22
XM_045308535.1	protein BTG2-like LOC123528652	6383.71021	1.041444379	0.106403926	6.20E-24	9.77E-21
XM_045308407.1	uncharacterized LOC123528585	613.077675	2.23136787	0.235402386	1.20E-22	1.73E-19
XM_045329160.1	bypass of stop codon protein 1-like (LOC123543096)	43.80081254	-4.442274699	0.472267853	4.00E-22	5.33E-19
XM_045346781.1	baculoviral IAP repeat-containing protein 7-A-like (LOC123556206)	338.9549447	4.075575651	0.438367666	7.20E-22	8.92E-19
XM_045332701.1	cerebellin-2-like (LOC123546440)	237.0991526	-3.047605522	0.334831465	3.92E-21	4.53E-18
XM_045325267.1	streptavidin-like (LOC123540331)	67.09886255	4.278868977	0.468729451	4.37E-21	4.74E-18
XM_045336479.1	uncharacterized LOC123548868	85.97665564	-1.793853987	0.197471409	4.93E-21	5.03E-18
XM_045305739.1	uncharacterized LOC123526544, transcript variant X1	525.9854549	2.378088602	0.268191452	3.20E-20	3.08E-17
XM_045328390.1	hephaestin-like (LOC123542492)	72.28363535	1.923095329	0.21861244	7.60E-20	6.94E-17
XM_045354530.1	alternative oxidase, mitochondrial-like (LOC123561865)	340.0734137	1.321305251	0.152513387	2.09E-19	1.72E-16
XM_045321083.1	uncharacterized LOC123537385	653.1734557	-1.631076616	0.194326459	2.17E-18	1.63E-15
XM_045328854.1	complement C1q-like protein 2 (LOC123542837)	67.23946333	-2.81626013	0.336309944	2.73E-18	1.97E-15
XR_006681067.1	5.8S ribosomal RNA (LOC123523273)	218.7523857	2.586881596	0.311053285	4.02E-18	2.79E-15
XR_006685890.1	uncharacterized LOC123548481	433.0047152	-1.512458423	0.188141696	4.15E-17	2.57E-14
XM_045309941.1	acetylcholinesterase-like (LOC123529569)	275.5194972	1.35670531	0.170716313	8.91E-17	5.33E-14
XM_045302383.1	sulfide:quinone oxidoreductase, mitochondrial-like (LOC123524296)	635.2460256	1.28384176	0.162445392	1.25E-16	7.24E-14
XM_045318306.1	uncharacterized LOC123535595, transcript variant X1	13.90154115	4.728451979	0.597352	5.20E-16	2.82E-13
XR_006681070.1	small subunit ribosomal RNA (LOC123523276)	152.5877681	2.178494101	0.285646538	1.08E-15	5.66E-13
XM_045333469.1	uncharacterized LOC123546855	1534.739941	-1.170416033	0.153816304	1.25E-15	6.35E-13
XM_045344029.1	baculoviral IAP repeat-containing protein 2-like (LOC123554115)	129.7697632	-1.993517369	0.269014151	5.64E-15	2.57E-12
XM_045359076.1	uncharacterized LOC123565067	24.05215603	3.964809036	0.536271145	8.72E-15	3.87E-12
XM_045346799.1	putative inhibitor of apoptosis (LOC123556214), transcript variant X2	288.4671902	3.669608864	0.502533886	1.20E-14	5.21E-12
XR_006685566.1	uncharacterized LOC123546958	507.6308974	1.536540676	0.210414385	1.29E-14	5.32E-12
XM_045322555.1	uncharacterized LOC123538449 (LOC123538449), mRNA	120.8260992	-2.599451563	0.356098005	1.26E-14	5.32E-12
XM_045351610.1	estradiol 17-beta-dehydrogenase 2-like (LOC123559638)	68.74233317	-1.781690332	0.246971067	2.49E-14	1.00E-11

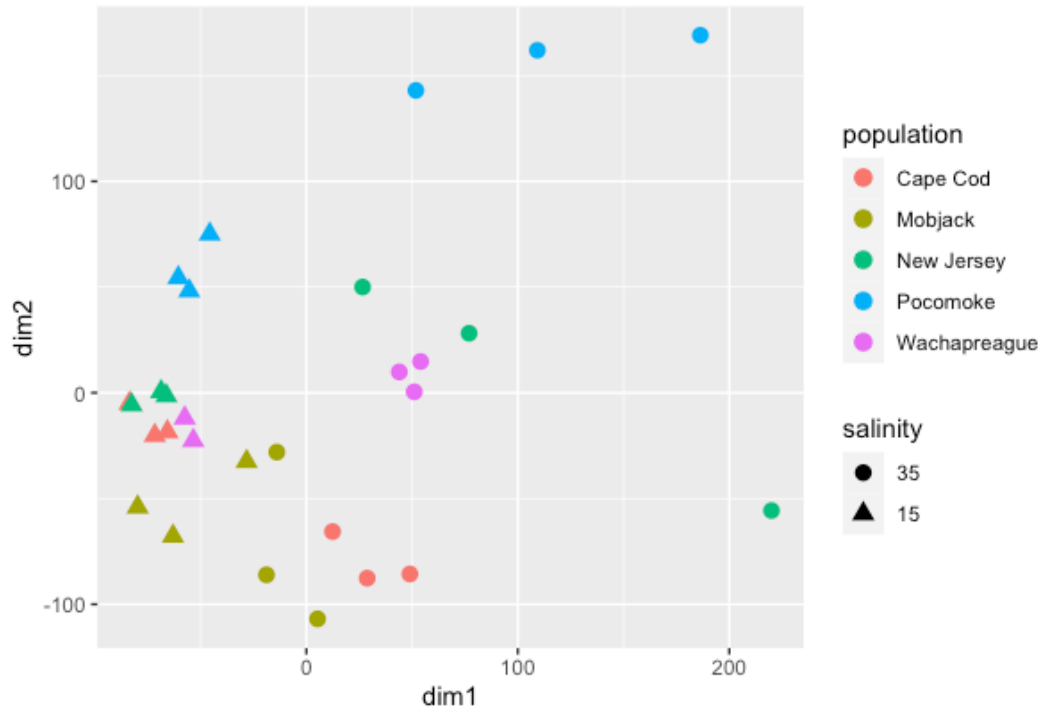
XM_045339737.1	uncharacterized LOC123551071	66.45652336	2.358098955	0.330228613	4.39E-14	1.69E-11
XM_045331480.1	uncharacterized LOC123545163	22.87738065	2.991632805	0.419750602	5.80E-14	2.14E-11
XR_006685904.1	uncharacterized LOC123548540	158.8379269	-1.62630593	0.230275037	7.26E-14	2.62E-11
XR_006687285.1	sulfotransferase 2A8-like (LOC123555432), transcript variant X2, misc RNA	107.8601845	-1.744258054	0.247161568	7.57E-14	2.68E-11
XM_045337220.1	glycogen-binding subunit 76A-like (LOC123549278)	498.3601109	2.315313712	0.330839351	1.19E-13	4.11E-11
XR_006686214.1	uncharacterized LOC123550221	152.9165523	-2.340936956	0.337877574	1.84E-13	6.13E-11
XM_045314954.1	uncharacterized LOC123533310	1238.805211	1.084184623	0.156443842	1.95E-13	6.38E-11
XM_045351832.1	uncharacterized protein DDB_G0274171- like	343.0920688	-1.304685689	0.191932642	4.74E-13	1.49E-10
XM_045308816.1	haloacid dehalogenase-like hydrolase domain-containing 5 (LOC123528809), transcript variant X2,	30.82705775	2.48070685	0.364355544	5.27E-13	1.63E-10
XM_045358618.1	uncharacterized LOC123564795	319.8599209	-1.091970969	0.161774928	6.53E-13	1.99E-10
XM_045309632.1	arylsulfatase B-like (LOC123529331)	129.9599995	-2.817292484	0.418670073	7.24E-13	2.16E-10
XM_045316901.1	clustered mitochondria protein homolog (LOC123534591)	765.5032974	1.000762225	0.149185478	8.94E-13	2.58E-10
XM_045343116.1	uncharacterized LOC123553388	57.9723746	2.851041352	0.425511363	9.38E-13	2.63E-10
XM_045313201.1	egl nine homolog 1-like (LOC123531899)	84.54352569	1.633053507	0.24361426	9.41E-13	2.63E-10
XM_045345682.1	tropomyosin-like (LOC123555103), transcript variant X1	44.90149397	-2.124810096	0.318107805	1.05E-12	2.84E-10
XM_045352589.1	uncharacterized LOC123560395	125.5718929	-3.215614756	0.481514273	1.04E-12	2.84E-10

Supplementary Table 5. Top 50 genes by ascending FDR or adjusted p-value (padj) between 35 and 15 ppt for juvenile samples. Those genes in red are in the top 15 for lowest FDR. Those genes in blue are in the top 15 for largest absolute LFC. Those genes in purple are in both the top 15 for FDR and LFC. Those genes in grey are uncharacterized. Gene is the gene accession number on NCBI. Name is the name given for each feature on NCBI. Base mean is the average of the normalized count values, divided by the size factor, taken over all the samples. LFC is the log2 fold change. LFC SE is the LFC standard error. The pvalue is the Wald test p-value, where the null hypothesis that there is no effect of treatment on the gene. The padj is the adjusted p-value based on the Benjamini-Hochberg method, or false discovery rate (FDR).

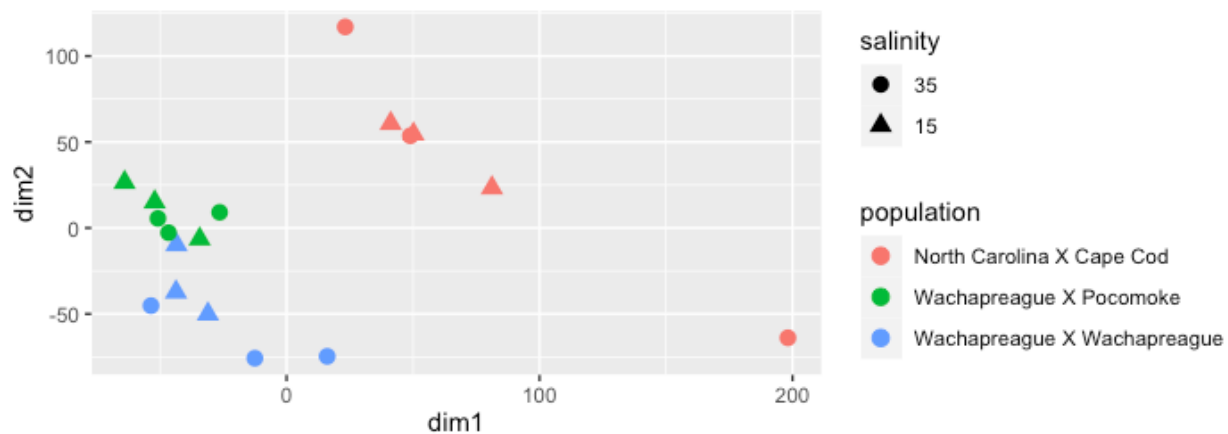
Genes	log2FoldChange.Adults	lfcSE.Adults	padj.Adults	log2FoldChange.Juveniles	lfcSE.Juveniles	padj.Juveniles	Names
XM_0453 46781.1	4.673170096	0.7584 14075	1.06E- 08	4.075575651	0.43836 7666	8.92E- 19	baculoviral IAP repeat-containing protein 7-A-like
XM_0453 46799.1	3.926131082	0.6845 44521	9.77E- 08	3.669608864	0.50253 3886	5.21E- 12	putative inhibitor of apoptosis, transcript variant X2
XR_0066 87285.1	-1.652783088	0.6245 61598	4.78E- 03	-1.744258054	0.24716 1568	2.68E- 11	sulfotransferase 2A8-like transcript variant X2, misc RNA
XM_0453 37220.1	1.891950919	0.2679 73076	4.56E- 11	2.315313712	0.33083 9351	4.11E- 11	glycogen-binding subunit 76A-like
XM_0453 08816.1	4.19367415	0.6254 33126	3.74E- 10	2.48070685	0.36435 5544	1.63E- 10	haloacid dehalogenase-like hydrolase domain-containing 5, transcript variant X2
XM_0453 13201.1	1.224828643	0.2803 59725	4.08E- 05	1.633053507	0.24361 426	2.63E- 10	egl nine homolog 1-like
XM_0453 45682.1	-1.45340421	0.3502 76339	8.12E- 05	-2.124810096	0.31810 7805	2.84E- 10	tropomyosin-like transcript variant X1
XM_0453 35801.1	5.219203638	0.5943 6961	1.81E- 16	3.368003859	0.50444 3833	2.88E- 10	phosphoenolpyruvate carboxykinase, cytosolic [GTP]-like, transcript variant X2
XM_0453 13711.1	3.243830015	0.9300 84157	0.0006 51	3.020841393	0.46144 3473	7.91E- 10	cholecystokinin receptor-like
XM_0453 46696.1	-1.514202826	0.3390 05226	2.60E- 05	-1.435047892	0.26154 8186	2.15E- 07	nuclear-pore anchor-like
XM_0453 60456.1	1.997567209	0.2314 94767	5.63E- 16	1.919039662	0.35117 3301	2.44E- 07	heat shock protein beta-1-like
XM_0453 19630.1	-1.729398293	0.5055 74879	7.52E- 04	-1.482002815	0.27455 3037	3.35E- 07	sulfotransferase 1B1-like
XM_0453 46770.1	1.87252886	0.5717 85094	0.0011 92885	2.504098161	0.48235 5594	9.15E- 07	E3 ubiquitin-protein ligase XIAP-like
XM_0453 18931.1	2.040376542	0.5235 71453	0.0002 16	1.761951395	0.39815 8857	2.67E- 05	potassium voltage-gated channel protein Shaker-like
XM_0453 53758.1	1.309072767	0.4659 60993	0.0035 79456	1.773289049	0.40755 4452	3.18E- 05	actin, cytoplasmic-like
XM_0453 15073.1	-1.630886106	0.2588 15016	4.46E- 09	-1.118530193	0.25859 8711	3.52E- 05	organic cation transporter protein-like
XM_0453 39804.1	1.962769813	0.3227 09516	1.63E- 08	1.189762578	0.28336 1905	6.07E- 05	headcase protein-like
XM_0453 44695.1	1.902616772	0.8695 64603	0.0109 08397	1.718818362	0.43911 019	0.00015 3	monocarboxylate transporter 13-like
XM_0453 09476.1	5.434475856	0.8426 73615	1.97E- 09	2.646806317	0.68155 0979	0.00016 5	heat shock protein 70 B2-like
XM_0453 35332.1	1.650808677	0.4834 20448	0.0008 23	1.995123724	0.53624 9029	0.00029 6	ras-like GTP-binding protein rhoA
XM_0453 01260.1	1.010199257	0.3271 54589	0.0019 81575	1.520074345	0.41411 8344	0.00035 5	dynein light chain Tctex-type 5-B-like
XM_0453 39856.1	2.337123787	0.3750 41433	7.25E- 09	1.747601498	0.48146 6985	0.00040 2	forkhead box protein O-like
XM_0453 41737.1	1.011472757	0.2675 22726	0.0002 86	1.097762551	0.31166 0673	0.00054 8	neuron navigator 3-like
XM_0453 18421.1	1.307577065	0.3827 53185	0.0008 39	1.409948423	0.39982 6165	0.00055 7	YLP motif-containing protein 1-like
XM_0453 07713.1	1.250922288	1.8358 16342	0.0398 99715	1.705934527	0.52164 7907	0.00112 5784	short-chain collagen C4-like
XM_0453 13300.1	-1.087482145	0.5680 08183	1.62E- 02	-1.000811714	0.30798 0434	1.20E- 03	acetylcholinesterase-like
XM_0453 49104.1	1.045767723	0.3602 25518	0.0028 88311	1.264855131	0.40910 1791	0.00185 0922	protocadherin Fat 3-like
XM_0453 38831.1	1.17935323	0.2247 41127	1.03E- 06	1.056655404	0.35369 8383	0.00245 49	nuclear hormone receptor FTZ-F1 beta-like
XM_0453 42359.1	2.40396229	0.3998 92887	2.51E- 08	1.610974048	0.57480 0538	0.00401 5183	zinc finger protein AEBP2-like
XM_0453 24159.1	1.27698809	0.2424 64939	9.81E- 07	1.284172643	0.46704 585	0.00412 2795	transcription factor Sp4-like
XM_0453 24643.1	1.765991394	0.3101 56797	1.17E- 07	1.306909602	0.52875 366	0.00721 7597	nuclear receptor coactivator 2-like
XM_0453 12580.1	1.740547486	0.3560 5039	5.19E- 06	1.201726272	0.49380 4085	0.00815 1485	sphingosine-1-phosphate phosphatase 2-like
XM_0453 60457.1	1.545674857	0.3647 78911	6.62E- 05	1.051458918	0.43754 4599	0.00830 2851	ADP-dependent glucokinase-like

XM_0453 36905.1	1.695729637	0.3119 77509	4.34E- 07	1.188058041	0.53430 0945	0.01117 7447	guanine nucleotide-binding protein G(s) subunit alpha
XM_0453 00723.1	1.67969198	0.3369 26428	3.38E- 06	1.070029662	0.53796 677	0.01603 7895	chondroitin sulfate proteoglycan 4-like
XM_0453 19886.1	1.478235236	0.5003 40373	0.0025 49148	1.24326125	0.63422 3416	0.01680 0365	nuclear receptor subfamily 1 group D member 1-like
XM_0453 09466.1	5.224210581	0.7878 47252	6.00E- 10	1.588620098	0.84493 3099	0.01752 2881	heat shock protein 70 B2-like
XM_0453 19245.1	1.294969013	0.3583 14819	0.0004 81	1.273385443	0.65701 4079	0.01787 726	UHRF1-binding protein 1-like
XM_0453 60899.1	1.192511586	0.4077 61113	0.0027 52625	1.118924586	0.60113 1113	0.01912 794	F-box only protein 33-like
XM_0453 52803.1	1.007461945	0.3362 16479	0.0023 93341	1.091923468	0.59171 5965	0.01945 0986	HMG box transcription factor BBX-like
XM_0453 46420.1	1.142664049	0.3795 71552	0.0023 55988	1.161300504	0.67416 1333	0.02234 4049	putative uncharacterized protein DDB G0271606
XM_0453 04130.1	1.357800919	0.4423 61433	0.0020 84686	1.072113756	0.78974 2431	0.03367 433	kelch-like ECH-associated protein 1B
XM_0453 23375.1	1.782999595	0.5829 30609	0.0020 84686	1.010030597	0.85549 5427	0.03805 2533	ADAM 17-like protease

Supplementary Table 6. Characterized genes that overlapped among adult and juvenile samples. Gene is the gene accession number on NCBI. Name is the name given for each feature on NCBI. Base mean is the average of the normalized count values, divided by the size factor, taken over all the samples. LFC is the log2 fold change. LFC SE is the LFC standard error. The padj is the adjusted p-value based on the Benjamini-Hochberg method, or false discovery rate (FDR).



Supplementary Figure 2. Generalized PCA of dimensions 1 & 2 for adult clam samples. Sample 15_Wach_G_3 was removed as an outlier and is not included in this PCA.



Supplementary Figure 3. Generalized PCA of dimensions 1 & 2 for juvenile clam samples.