GENETICS OF SHELL SHAPE AND COLOR IN THE EASTERN OYSTER, *CRASSOSTREA VIRGINICA*

By

Michael Whiteside

A thesis submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Master of Science

Graduate Program in Ecology and Evolution

Written under the direction of

Ximing Guo

And approved by

_____

_____

_____

New Brunswick, New Jersey

May 2021

ABSTRACT OF THE THESIS

Genetics of Shell Shape and Color in the Eastern Oyster, *Crassostrea virginica*

MICHAEL WHITESIDE

Thesis Director:
Ximing Guo

ABSTRACT

Because of the loss of wild populations, most oysters consumed in restaurants today are produced in hatcheries and grown to market size by extensive aquafarming. The demand for eastern oysters has been increasing and is poised for continued growth. As the industry grows, it will become increasingly important for breeding programs to broaden goals and address more traits that may be of interest to the commercial half-shell market. The appearance of the oyster's shell is the first thing consumers encounter and a better understanding of shell traits can provide an opportunity for product differentiation. The nature and inheritance of the shape and color of shells of the eastern oyster, *Crassostrea virginica*, remain largely unstudied. This thesis provides direct evidence that both traits have a strong genetic component in this species.

In the first chapter, the 'curl-back' hinge abnormality is addressed. Two cohorts of oysters selected for 'curl-back' and for 'flat' hinges were produced and the progeny from each of the two crosses were grown separately in a common garden experiment. In the following year, crosses were repeated using selected F1s, resulting in two generations of selection for 'flat' and for 'curled-back' oysters. Four and eleven-month-old F2s from each cohort were subsequently cut ventro-dorsally and the relative curvatures of right and left valves were compared using either relative ratios of basic measurements or following a theoretical morphologic approach. Utilizing three independent methodologies, the curl-back shell abnormality is demonstrated to have a strong genetic component in the eastern oyster.

The second chapter examines the modes of inheritance of color morphs in the eastern oyster. Although, color morphs have been described in other species, such as the Pacific oyster, *Crassostrea gigas*, to date, no studies have been reported in *C. virginica*. To better understand the nature of shell pigmentation in this species, four color morphs were described and crossed. The proportions of color morphs in the progeny indicate that color and pattern are under genetic control and that their inheritance involves a small number of genes.

The third and final chapter describes a comparative transcriptome study conducted in order to identify differentially expressed genes between dark and light regions of mantle tissue that correspond to shell pattens. Several genes were found to be differentially expressed, indicating a few gene families or pathways underlying the molecular basis of pigmentation that provides valuable insight for future investigations.

ACKNOWLEDGEMENTS

This thesis is the product of two summer seasons (2016 and 2017) of work in the oyster hatchery and follow up work completed in 2019. My project would not have been possible without the help and encouragement from many teachers, mentors, collaborators, and coworkers, as well as family (especially my parents) and friends.

I need to first thank Dr. Ximing Guo for taking me on as a lab technician in 2013 and for offering to serve as my master's thesis advisor. I cannot overstate my gratitude for his generosity, patience, and encouragement; for sharing his expertise; and for allowing me the freedom to learn in his lab and benefit from his wisdom and good judgement.

Patty Woodruff and Aaron Maffei helped ensure that animals in the hatchery and on the tidal flats at Rutgers University's Cape Shore Laboratory were well taken care of. I am happy to acknowledge the hard work and assistance provided to them and to me during the summers of 2016, 2017, and 2018 by David Lappin, James Hartley, Colin Hofmann, Mike DiLorenzo, Tiffany Burke, Shaun Wigglesworth, Lauren O'Neil, and Nicole Pray.

I am deeply grateful for the steadfast support, encouragement, and learning opportunities provided by the entire HSRL community for the duration of my time as a student. I feel indebted to everyone I have worked with at the HSRL, each of whom has contributed to my education. I'd especially like to extend a sincere thank you to my thesis committee members, Dr. Daphne Munroe and Dr. David Bushek.

I found the Ecology and Evolution community on main campus to be a great fit, and was lucky to be able to take courses ranging from Gaia to selfish genes. I am grateful for all of the opportunities the E&E program had to offer and only wish I could have taken more full advantage of all that was available to me in New Brunswick.

I was fortunate to receive some expert help on my thesis from Dr. George McGhee and Dr. Suzanne Williams; I could not have hoped for better assistance with studies of molluscan shell shape and color.

Dr. McGhee's books The Geometry of Evolution and Theoretical Morphology have served as my main guides for thinking about shape and the morphometric analysis of shells is fully his – I thank him for encouraging the analysis of oyster shells, for checking the math, and making sure I was doing it properly. I am very happy to have discovered his work and that he is also at Rutgers, which made him a convenient collaborator.

Dr. Williams's work on molluscan shell color was a great help and inspiration, and I was lucky to be able to meet her at the 110th National Shellfish Association meeting in Seattle, WA. She shares our excitement for oyster shell color and, in addition to offering suggestions for the transcriptome analysis, has helped to arrange follow up studies of our shells.

I finally would like to thank the educators in the biology department at the University of Tampa, whose openness and dedication to teaching left lasting impressions.

DEDICATION

This thesis is dedicated to my brothers and our family dogs

"I suppose that when the sapid and slippery morsel – which is and is gone, like a flash of gustatory summer lightning – glides along the palate, few people imagine that they are swallowing a piece of machinery (and going machinery too) greatly more complicated than a watch."

- T. H. Huxley

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INHERITANCE OF SHELL SHAPE AND MORPHOMETRIC ANALYSIS

INTRODUCTION

Background

Oyster reefs create critical habitats that support the maintenance of biodiversity in estuarine systems as well as the life history stages of many economically valuable species. Oysters also stabilize shorelines and remove excess quantities of phytoplankton, suspended sediments, and nitrogen from the water column. The current condition of wild oyster populations globally, however, is alarming; approximately 85% of oysters have been lost worldwide, rendering massive areas of reef habitat functionally extinct (Beck et al. 2011).

There are multiple reasons for the poor conditions of oyster reefs. A major contributor to loss was overharvesting (Beck et al., 2011), which peaked in New Jersey between 1870 and 1930 (Ford 1997), but, although some oyster fisheries are now well managed, threats are still present and are exasperated by the interactions and synergistic effects of sedimentation, disease, excess nutrients, hypoxia, and more. These processes lead to feedback loops and faster rates of loss (Rothschild et al. 1994; Maslo 2016).

Because of the loss of wild populations, many oysters consumed in restaurants today are produced in hatcheries and grown to market size by extensive aquafarming practices. Eastern oyster (*Crassostrea virginica*) farming is now one of the most important aquaculture industries in the northeastern United States. In 2015, the estimated value of the eastern oyster industry in the United States was $94 million, a 113% increase from 2005. The eastern oyster is also responsible for most of the recent increases in US oyster aquaculture production and the industry is in position for continued growth (FAO 2017).

Efforts to rear oysters from larvae for aquaculture began in the late nineteenth century (Kennedy 2014), and ongoing selective breeding programs were established beginning in the 1960's to create hardier stocks (Guo et al. 2008a). Technological advances have progressively allowed for more targeted selection and differentiation as a result of studies performed on the genetic level (Proestou et al. 2016; Hollenbeck and Johnston 2018).

Disease has been one of the major concerns for the industry as *C. virginica* is vulnerable to outbreak events, which can be unexpected and cause massive mortalities. At least 18 oyster disease-causing agents have been described, but the two that have had the most significant and chronic negative impacts historically have been *Perkinsus marinus* (Dermo) and *Haplosporidium nelson* (Multinucleated sphere X) (Ewart and Ford 1993; Ford and Tripp 1996). Naturally, surveys have shown that breeders have mostly been concerned with disease resistance, growth, and survival (Mahon 1982).

Like many other economically important aquaculture species, genomic resources have only recently been developed for the eastern oyster and have only begun to be applied to selective breeding (Guo et al. 2008b; Hedgecock 2011; Abdelrahman et al. 2017). As demand grows, and the genetics underlying many traits are better understood, it will become increasingly possible and important for breeding programs to broaden goals and address more traits that may be desirable to consumers making purchase decisions in the commercial half-shell market.

A few recent studies have examined preferences of consumers of eastern oysters in order to identify marketing implications and inform policy. For example, in 2017, the University of Delaware's Center for Experimental & Applied Economics published the results of surveys of more than 1,400 potential oyster consumers, local residents, and tourists conducted in order to identify preferences for oyster attributes and branding. They found that frequent oyster consumers

preferred farmed over wild-caught oysters and that both locals and tourists preferred oysters that were advertised as “local.” Consumers were willing to pay more for oysters accompanied by a logo and also if they were provided with basic information regarding how oysters filter water and have a positive effect on water quality and the environment. Fresh smell was the most important attribute to consumers, followed by saltiness, then meat size and meat color, and then shell size and shell color (Kecinski et al. 2017).

Although shell traits are apparently not the most important characteristics driving purchase decisions, the appearance of the shell is the first thing consumers encounter, and a consensus does exist among shellfish farmers and consumers regarding what constitutes a ‘good’ shell. Preferred shells are hard, robust and clean, and without dirt or epibionts; they are without blisters or chambers; and their shape resembles a teardrop with a length:width:depth ratio of about 3:2:1 (Mizuta and Wikfors 2018). Shape abnormalities can therefore affect marketability and value.

Curl-back Hinge Shell Abnormality

In 2016, a shell abnormality was found in a large number of hatchery-produced eastern oyster seed and brought to the attention of shellfish breeders at Rutgers University by local farmers. This undesirable trait is characterized by a sharp, downward curvature of the plane of commissure close to the hinge (Figure 1), making them difficult to shuck and to sell to consumers.

Figure 1. Comparison of cultured oysters with preferred shape (left), including a deeply cupped left valve, and a curl-back oyster (right) with curved valves and a shallow left valve.

Anecdotal observations of oysters available for purchase at seafood markets and sampled from Rutgers University's selective breeding program suggested that the abnormality was widespread in farm-raised oysters. It was hypothesized that this trait was likely genetically determined as researchers in Australia had described finding the same trait, which they called 'curl back', in the Pacific oyster, *Crassostrea gigas*, and reported that it may be caused by a single recessive gene (Boudy and Haffray 2008). However, the statement that curl-back is a deleterious recessive may have been based on unpublished results of many single pair crosses carried out at an early stage of the Pacific oyster improvement program. The inheritance of the trait in this and in other species of oysters is therefore considered to possibly be more complicated and potentially completely environmentally determined (B. Ward, personal communication, July 2016).

Morphological aberrations, such as curl-back, can be influenced by a combination of genetics, environmental conditions, and/or husbandry practices. Understanding the causes of such aberrations is important to determine how they might be rectified, but remains largely unstudied

in the eastern oyster. Generally, the form of a molluscan shell is related tightly and directly to function (Stanley 1970), and it is not immediately obvious why - if the curl-back trait has a genetic component - variation in oyster valve shape is maintained in populations.

Possible forms of a given evolutionary lineage are limited by a number of factors, including developmental, phylogenetic, functional, and geometric constraints. Together, these can be referred to as bauplans and they carry with them engineering limitations, which determine form (McGhee 1999). Iterative evolution and inherent adaptive limitations are apparent in Bivalvia (Stanley 1975), and a few basic morphologies are especially prominent throughout the fossil record. Mondal and Harries (2016), for example, recently described forty-four ecospaces that have been realized by taxa of bivalves throughout the Phanerozoic as different combinations of tiering, feeding, and motility.

Fitness Landscapes and Theoretical Morphology

Theoreticians modeling evolution will often employ the metaphor of the 'fitness landscape', first introduced by Sewall Wright (1932). This is also referred to as an adaptive landscape and represents phenotypes as a series of bivariate three-dimensional graphics. Genotype or morphocharacter combinations are plotted on the *x*-*y* axes, while their relative fitness or degree of adaptation is represented by the *z*-axis. Topographic highs within the space represent adaptive phenotypes or morphologies that function well in a given environment, and adaptive lows represent adaptive phenotypes or morphologies that will not function well in the same environment (Figure 2). The fitness landscape, although flawed when applied to many real systems (Pigliucci 2012; Plutynski 2008), is a helpful visual for thinking about evolvability – i.e., the capacity to evolve – as the space represents genetic potential (e.g. Simpson 1944; Dennett 1995).

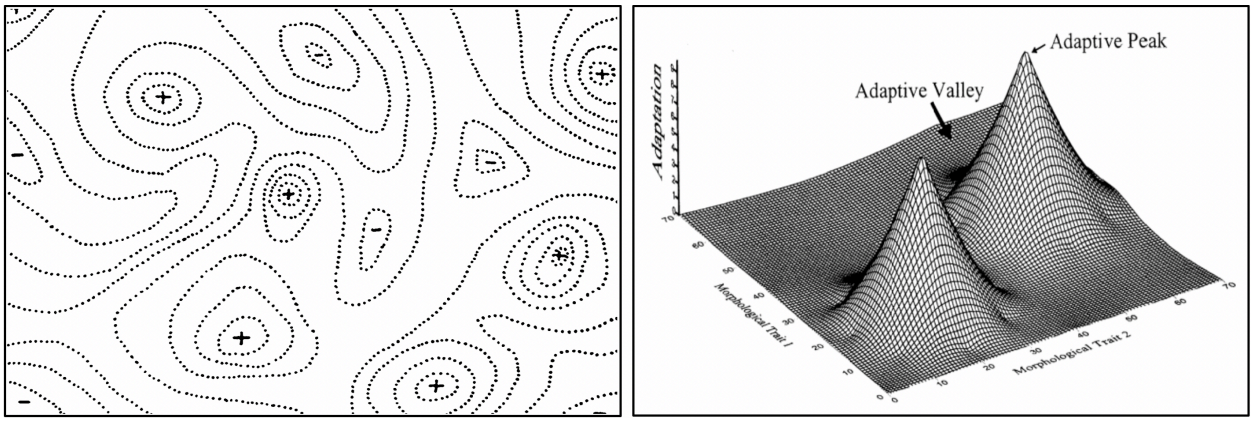


Figure 2. The adaptive landscape metaphor is a simulation of possible phenotypic or genotypic variation along the x and y axes and relative fitness topographically along the z-axis. Source: Wright 1932 (left) and McGhee 2007 (right).

The application of the adaptive landscape has been particularly successful in the examination of the geometry of molluscan shell morphologies because most of the variation exhibited can be expressed by just three parameters (and therefore modeled in three dimensions) and because a mollusc grows at shell margins and carries a complete record of accretionary growth through time (Moulton, Goriely, and Chirat 2012; Stone 1996). Raup (1966), following D'Arcy Thompson (1917), described a model of shell coiling that takes three geometric parameters of shell formation into account. Raup defined:

W.) Whorl expansion rate, which describes the rate of increase in size of the generating curve per revolution,

D.) Distance from central axis, which describes the distance between the generating curve and axis of coiling, and

T.) Translation rate, which refers to the angle between the geometric and biological generating curves

These parameters define a hypothetical 3-dimensional morphospace of molluscan morphology, and, by plotting the actual frequency of extant and extinct organisms in this space, a

4th dimension related to adaptive potential can effectively be added (Figure 3). Raup's cube has become a classic visual tool for illustrating ideas such as missing phenotypes, limits of organismal form, convergent evolution, and developmental constraints, and continues to be discussed widely and used in empirical studies (e.g., Stone 1996; Okabe and Yoshimura 2017).

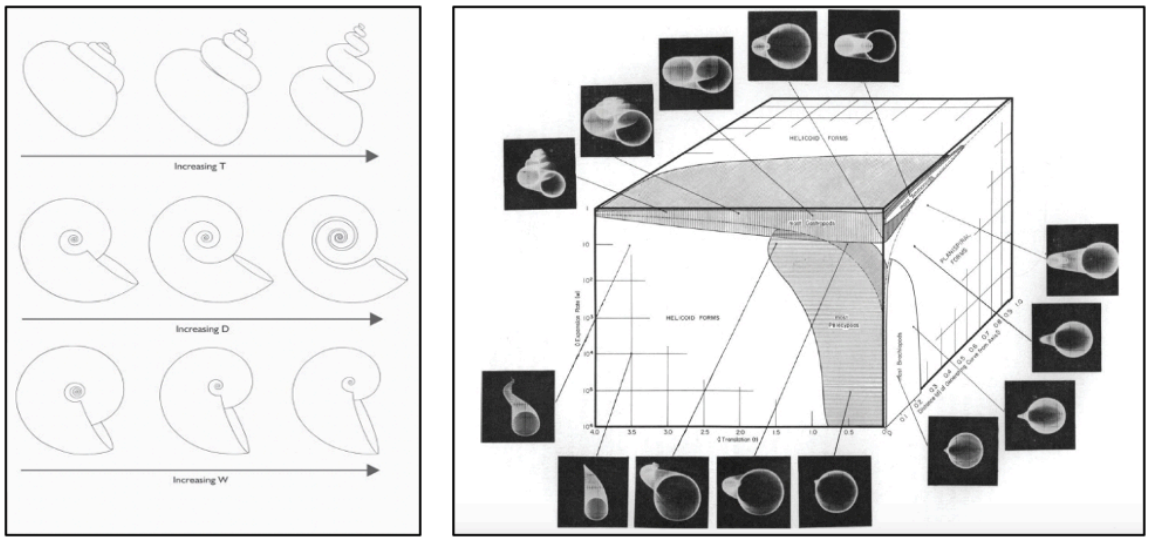


Figure 3. Left: Raup described three parameters, W, D, and T to model shell forms (Source: modified from Scales 2015). Right: The three parameters are used as axes to form a cube which defines a three-dimensional morphospace of all possible shells (Source: Raup 1966).

Raup originally considered ammonoid shells for his simulations, which included just two parameters, W and D, and the most frequent use of the model has also been for describing planispiral forms, where the T parameter equals zero (Gerber 2017). The absence of translation along the coiling axis reduces the morphospace to two dimensions as the generating curve stays in the same plane as coiling proceeds. The parameters W and D are derived from a two-dimensional logarithmic spiral model (McGhee 2007):

A logarithmic spiral has three geometric parameters (Figure 4): radius, coiling angle, and tangent angle,

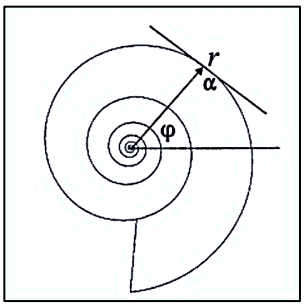


Figure 4. A logarithmic spiral with radius (r), coiling angle (ϕ), and tangent angle (α).

and is described by the following exponential equation

$$r_{\phi} = r_0 e^{c\phi}$$

where r_0 is the magnitude of the initial radius in the simulation, r_{ϕ} is the magnitude of the radius at coiling angle ϕ, and c is the "specific growth rate constant", which is defined as the derivative of the equation:

$$c = dr/r\, d\phi$$

This constant is related to the curvature of the spiral by

$$c = \cot \alpha$$

And the original equation can be rewritten as

$$r_{\phi} = r_0 e^{(\cot \alpha)\phi}$$

where r_0 is the magnitude of the initial radius in the simulation, and r_{ϕ} is the magnitude of the radius at coiling angle ϕ.

The tangent angle is replaced with ratios of radii to describe a new parameter (W):

$$W = r_{\phi} / r_0^{\,2\pi/\phi}$$

The equation is rewritten in terms of the introduced whorl expansion rate (W) as:

$$r_{\phi} = r_0 W^{\phi / 2\pi}$$

A typical coiled shell, such as an ammonoid, can be modeled by the relationship between two logarithmic spirals - one spiral describing the outermost edge of the shell and a second spiral describing the innermost edge. This relationship is how the parameter D is defined (Figure 5).

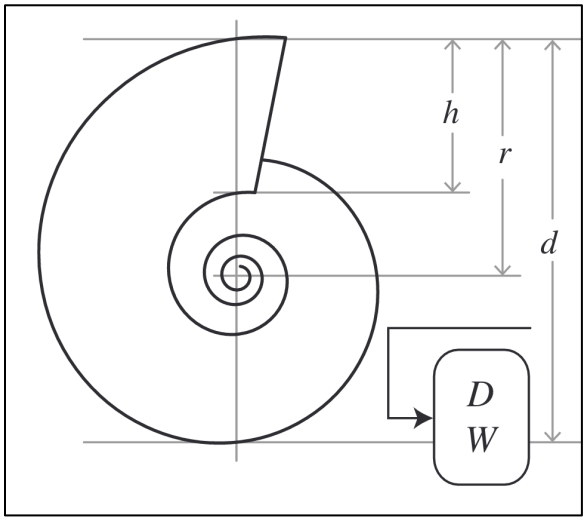


Figure 5. Raup's model of shell coiling characterizes planispiral forms using parameters D and W, which are based on three linear measurements (shell diameter d, whorl height h, and radius length r). Using these measurements, $W = r / (d - r)^2$ and $D = (r - h) / r$ Source: Modified from Gerber 2017.

The morphology of a coiled shell is therefore modeled (Figure 6) by the two equations:

$$\text{Outermost Spiral } r_{\phi} = r_0 W^{\phi / 2\pi}$$

$$\text{Innermost Spiral } r_{\phi} = D \left(\text{Outermost Spiral } r_{\phi}\right)$$

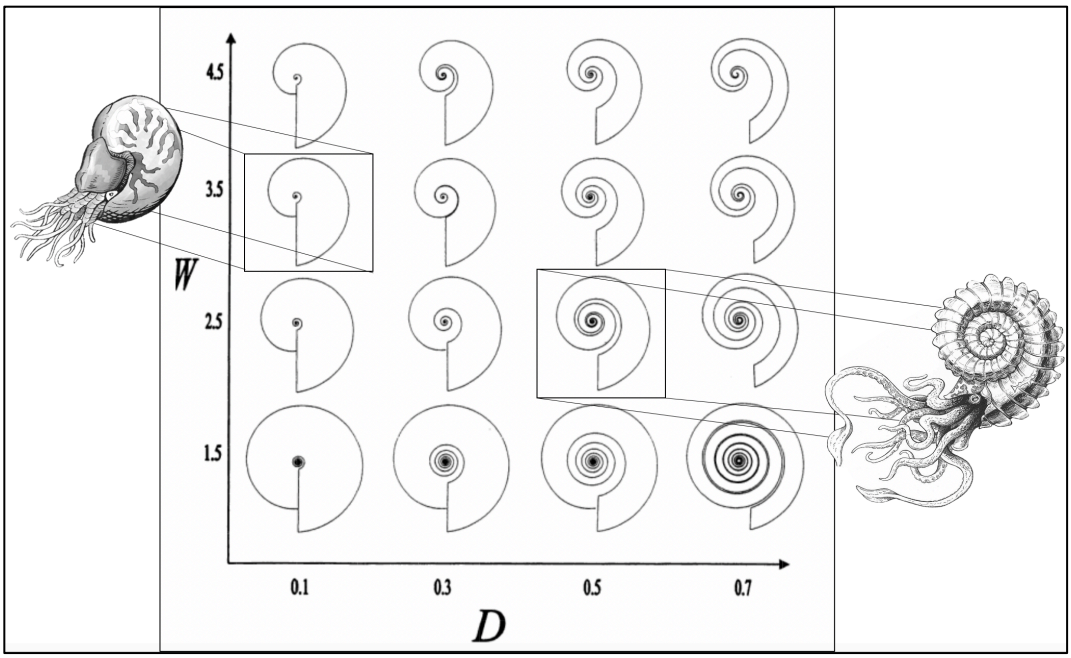


Figure 6. The theoretical morphospace of univalve form created with parameter dimensions W and D. Forms, such as typical ammonoid and nautilus shells, can be plotted in this space for intra- or interspecific comparisons. Source: Modified from McGhee (2007); illustrations of nautilus and ammonoid by Aaron John Gregory.

In order to create a theoretical morphospace of shell forms with two valves using these two geometric parameters and brachiopods as a model, McGhee (1980) specified separate W values for both the dorsal and ventral valves and kept D values constant so that the condition $W > 1/D$ is met. McGhee found that typical brachiopod shells have dorsal D values near 0.01 and ventral D values of near 0.1. Molluscs with higher values of D or low values of W, which violate the above condition, are impossible because the two hypothetical valves would interpenetrate one another. McGhee's theoretical morphospace is designed to simulate the spectrum of possible brachiopods that are biconvex (Figure 7).

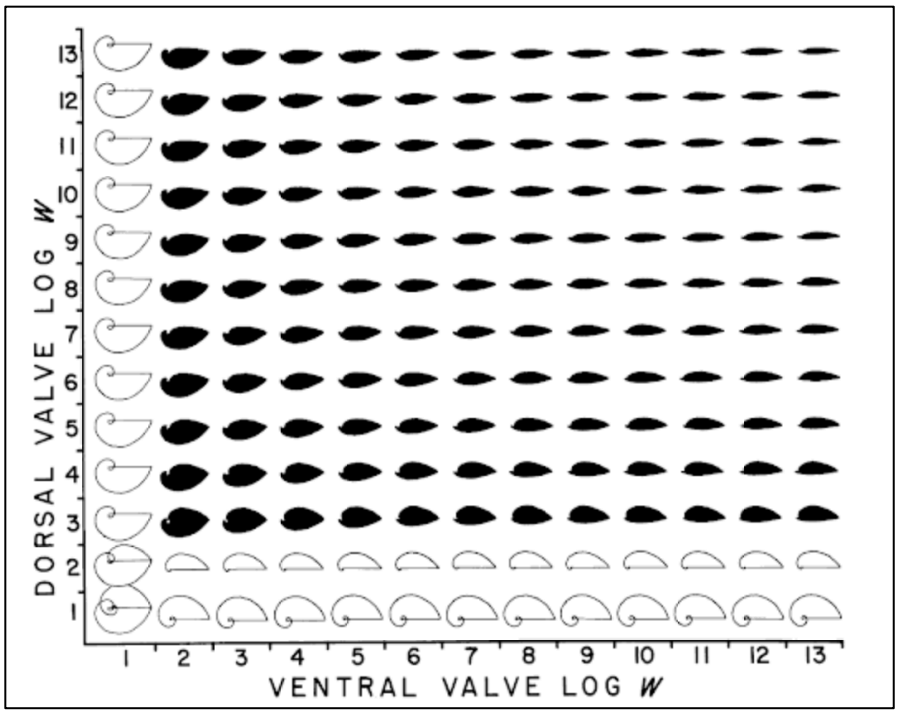


Figure 7. Theoretical morphospace of hypothetical brachiopod shells where separate W values are specified for each of the two valves. Possible forms are shaded black and impossible forms are white. Source*: McGhee (1980).*

To simulate more complicated, non-biconvex shell geometries with the logarithmic spiral model, McGhee (1999) abandoned the W parameter and used the tangent angle of the spiral (α) as a measure of curvature to create a theoretical morphospace that includes the total spectrum of hypothetical brachiopod shell form, from concavo-convex to plano-convex to biconvex to convexi-plane to convexi-concave geometries. This morphospace also holds values of D constant for dorsal and ventral valves at 0.01 and 0.1, respectively, and simulates the spectrum of forms with independently varying tangent angle (α) values for each of the valves (Figure 8).

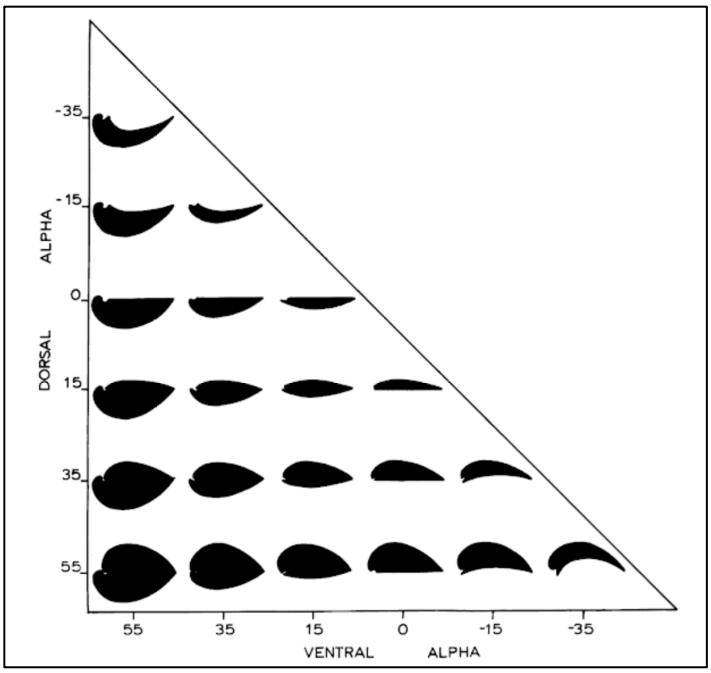


Figure 8. Theoretical morphospace of hypothetical non-biconvex brachiopod shells where separate alpha values are specified for each of the two valves. Source*: McGhee (1999).*

In this morphospace, growth is assumed to be isometric as the proportions of the logarithmic spiral remain the same and the tangent angle (α) does not change as the spiral grows (Figure 9).

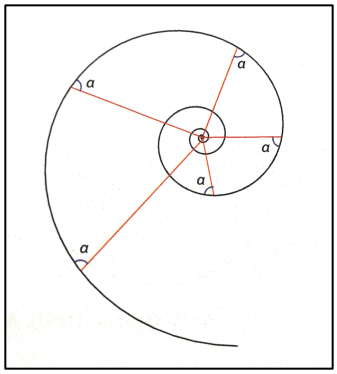


Figure 9. An equiangular spiral with tangent angle (α). Source: Hammer (2016).

Objective of Study

In order to determine whether the curl-back trait has a genetic component in the eastern oyster, two successive generations of selection for this trait were performed and the response was observed in the F2 progeny. The shapes of oysters were measured and compared and then mapped onto theoretical morphospaces by taking several radii and angle measurements to calculate averaged parameters.

METHODS

In June 2016, 3-year-old cultured wild stock oysters were collected from the tidal flats at Rutgers University's experimental oyster farm at the Cape Shore Laboratory research field site in Green Creek, New Jersey and brought into the hatchery. Oyster shell shape varied continuously from oysters with very cupped left valves and flat right valves (normal) to oysters with the curl-back shell abnormality (Figure 10). The oysters were sorted into two groups based on whether or not they exhibited the curl-back hinge trait for use as parents in experimental crosses. Available individuals with the most exaggerated shapes (normal or curl-back) were selected for use in experimental crosses.

Figure 10. Oyster shell shape varied continuously from very cupped (left in image) to curl-back (right in image). Available individuals with exaggerated shapes were selected for use in experimental crosses.

Oysters selected for use in crosses were shucked and each individual's sex was determined by microscopic examination of gonadal tissue. Gametes from females and males were obtained by strip-spawning (cite) and eggs were passed through an 80µm nylon mesh screen to remove debris and caught and washed on a 20µm nylon mesh screen before fertilization using several milliliters of dense sperm suspension.

Four by four crosses (16 families per culture) were produced using both curl-back and normal oysters crossed amongst themselves (curl-back crossed with curl-back and normal crossed with normal) and cultured in separate 200 liter tanks in parallel stocked on the same day.

The cultures were initially stocked with 50 larvae per mL and then thinned to 900,000 – 1,000,000 larvae per tank (about 5 larvae per mL) on day two. Larval culture tanks were kept at 20 – 25°C and complete seawater changes were performed with 1-micron (1 µm) filtered sweater (FSW) every other day. Each day, the oyster larvae were fed a mixed diet of cultured *Tisochrysis lutea* (T-iso)*, Isochrysis galbana* (C-iso)*, Pavlova lutheri* (P-mono)*,* and/or *Nannochloropsis sp.* (Nanno B) microalgae, as well as Shellfish Diet 1800® commercial algae paste (Reed Mariculture Inc.); the daily amount of algae the cultures were fed was based on the cultures' age, size, and density (Loosanoff and Davis 1963).

When larvae had visible eye spots and were ready to metamorphose, they were treated with epinephrine to stimulate metamorphosis without attachment to substrate (Coon, Bonard and Weiner 1986), and were then placed into downweller silos with 200 µm nylon mesh screened bottoms in the same two 200 liter larvae tanks. After a period of about two weeks in downwellers, oysters were placed into a nursery where each group was stocked in a separate upweller silo at roughly equal densities (10,000 to 15,000 per upweller silo, which each had approximately 0.5 meter diameter). When oysters were >5 mm in length, they were placed in separate ADPI OBC

bags with 4mm mesh and deployed on Rutgers University's intertidal experimental oyster farm (rack-and-bag culture).

In July 2017, the 13-month-old F1 progeny from 'curl-back' and 'normal' cultures were sorted based on whether or not they exhibited the curl-back hinge trait. Three by three crosses (9 full-sib families each) were performed to produce F2 generations of oysters selected for the 'curl-back' trait as well as oysters that were 'normal' (Figure 12). Oyster larvae were reared following the methods described above. On day four, counts for the curl-back group and normal cultures were 300,000 (1.5/mL) and 76,000 (0.38/mL), respectively, and stocking densities throughout the larval culture period were therefore considered to be much lower than is typical in shellfish hatcheries. The initial stocking densities in silos in the land-based upweller nursery system at the Cape Shore Laboratory field station for the curl-back and normal groups were 4,000 and 7,500 1mm seed oysters, respectively. This was also much lower than typical stocking densities and any density effects on shape were therefore minimized.

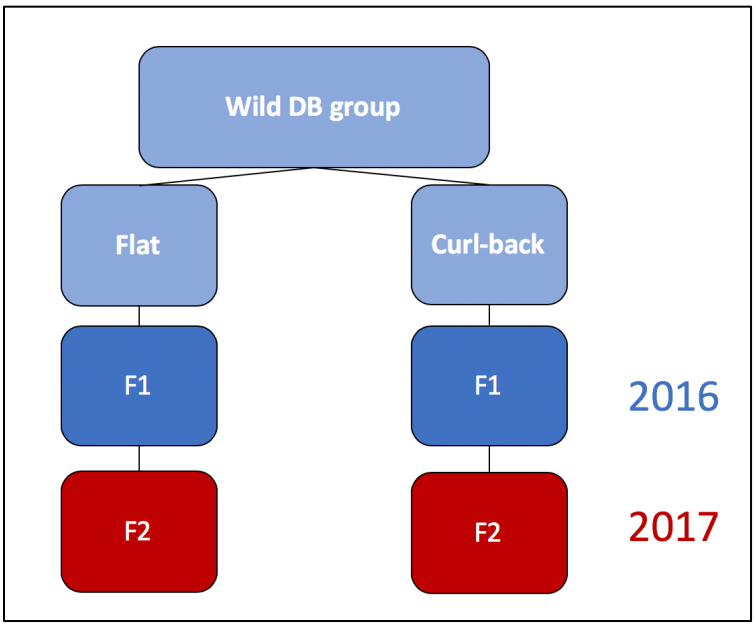


Figure 11. Schematic of crosses performed in 2016 after selection of wild DE bay parents for flat and curl-back oysters to produce F1 (blue). Crosses were repeated in 2017 using selected F1s as parents to produce F2 generation (red).

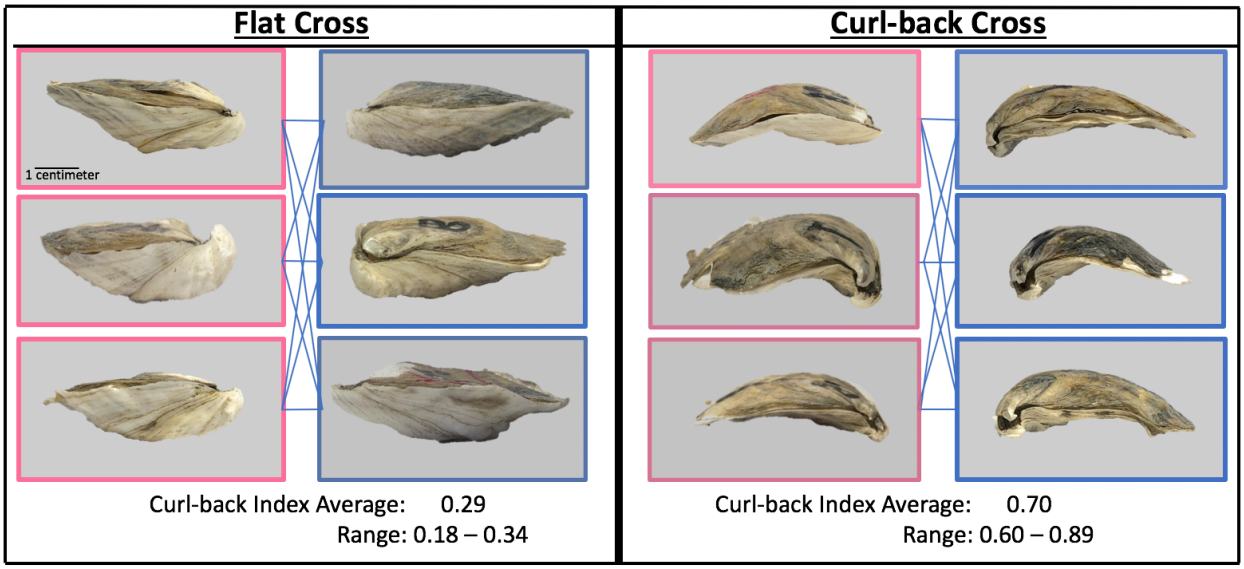


Figure 12. Photos of selected F1s used in 3x3 flat cross and in 3x3 curl-back cross in 2017 to produce F2s. Border colors indicate females (pink) and males (blue).

F2 oysters produced by these crosses were examined using two indices that compared relative lengths and heights of right and left valves, and a morphometric analysis described by McGhee (1980, 1999).

Indices

To measure the difference in degree of curl-back in the F2 progeny between the curl-back and flat selected lines, two simple indices were used. One hundred four-month-old oyster shells from each of the cohorts, Curl-back F2 and Normal (flat) F2, were cut ventro-dorsally using a Dremel with a diamond blade. Images of the cross-sectioned shells were digitized and the relative curvature of the right and left valves were measured using the open-source computer program FIJI.

The first index (called 'curl-back index') compared the length from the top of the right valve at the point of greatest shell inflation to an approximation of the plane of commissure (orange line in Figure 13) to the total height at the point of greatest shell inflation (yellow line in Figure 13).

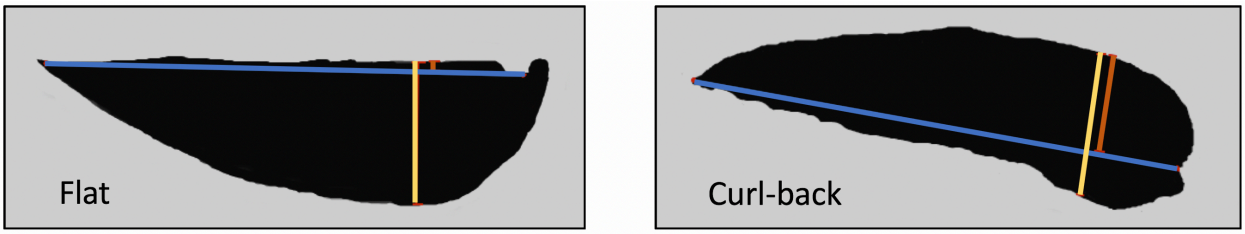


Figure 13. The 'curl-back' index is a comparison of the distance from the right valve to the plane of articulation relative to the total height, or distance from right to left valve, at the point of greatest shell inflation.

The second index (called 'one third valve length ratio') was a comparison of the actual right valve length to one third of the total length (one third of the total length was chosen because it lies near

the point of greatest shell inflation) (pink line in Figure 14) to one third of the measured total length following the approximated line of the plane of commissure (green line in Figure 14).

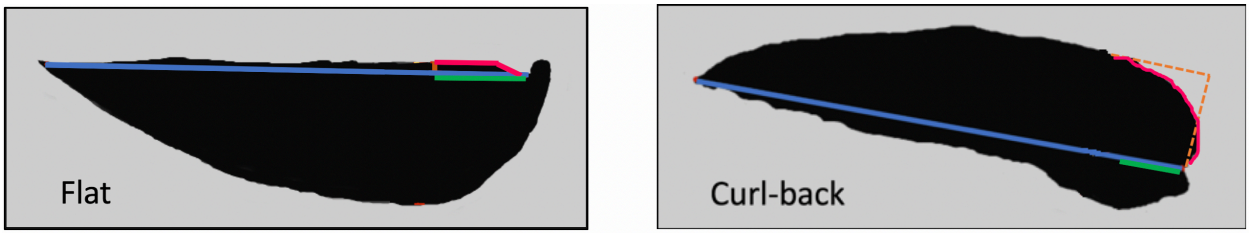


Figure 14. The 'one third valve length ratio' (not to scale) is a comparison of the actual right valve length to one third of the total length of shell relative to one third of the total length measured from the hinge to the bill.

For both indices, measured ratios for relatively flat oysters will be close to zero while values for relatively curled oysters will be larger.

Morphometric Analysis

Oysters from each of the F2 cohorts, were collected 7 months after initial sampling for measurements using the two indices described above. Eleven-month-old shells were cut ventro-dorsally and images of the cross-sectioned shells were digitized for measurement.

A line of reference was drawn between the hinge and the bill of each oyster to differentiate between the right and left valves. Radii length measurements were then made at regular angle increments of either 5 or 7.5 degrees along the shell margin for both valves (increments chosen depended on convexity / concavity of valve) for 37 individuals from the F2 generation selected for the curl-back and 41 individuals from the F2 generation selected for flat shells.

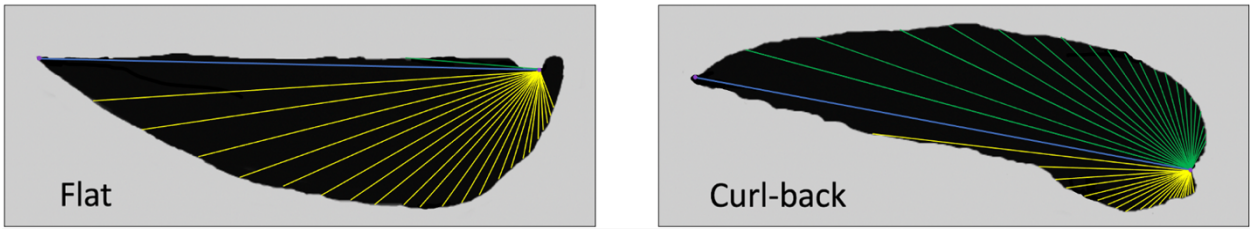


Figure 15. The open source computer program FIJI was used to draw lines of reference between the hinge and bill (blue) for each cross-sectioned shell and then to measure radii lengths at regular angle increments for each valve (green, yellow).

The measured degree data was converted to radians (arc lengths) and radii lengths (measured in millimeters) were converted to natural logarithms. Fitting linear regressions to the ln-radii and angle data allows for the calculation of c, the slope of the regression best-fit line, for each shell. This value is equal to cot(α), where α refers to the tangent angle of a logarithmic spiral.

To determine values for α, the tangent of α is first found by calculating 1/cot(α) or 1/c. Next, finding the arc tangent of the tangent of α results in α angles expressed in radians. Radian values are converted to degrees by multiplying by 180 and then dividing by 3.14.

$$cot\ \alpha = \text{slope of linear ln function}$$

$$tan\ \alpha = 1/cot\ \alpha$$

$$\boldsymbol{\alpha = tan^{-1}(tan\ \alpha)\ (180/\pi)}$$

To determine whorl expansion rate (W) values, the constant e is raised to the power of c, the slope from the linear regression, multiplied by 2π. Logarithms of whorl expansion rates are reported as the conversion allows for values to be plotted in McGhee's LogW morphospace.

$$W = e^{2\pi c}$$

$$lnW = 2\pi c$$

$$\boldsymbol{LogW = lnW / 2.3025}$$

RESULTS

The observed effect of two generations of selection for curl-back between the two cohorts was visually apparent (Figure 16).

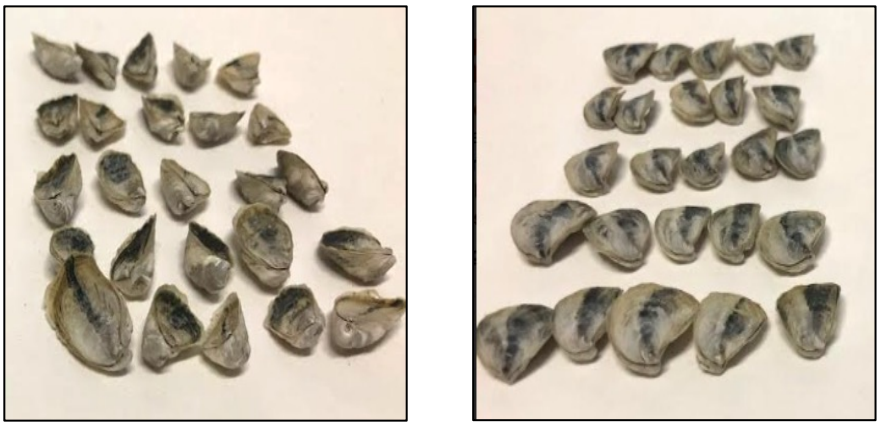


Figure 16. Photos of four month old F2 oysters selected for flat (left) and curl-back (right) shapes visually illustrate the effect of selection. One hundred individuals from each cohort were sampled and cross-sectioned for comparisons using curl-back index and one third valve length ratio.

A two tailed t-test comparing 'curl-back index' values between the two groups of 100 sampled individuals yielded a *p*-value of 1.26E-16, while a two tailed t-test comparing 'one third valve length ratio' values of the same individuals, was also significantly different (p = 2.15E-05) between the two cohorts.

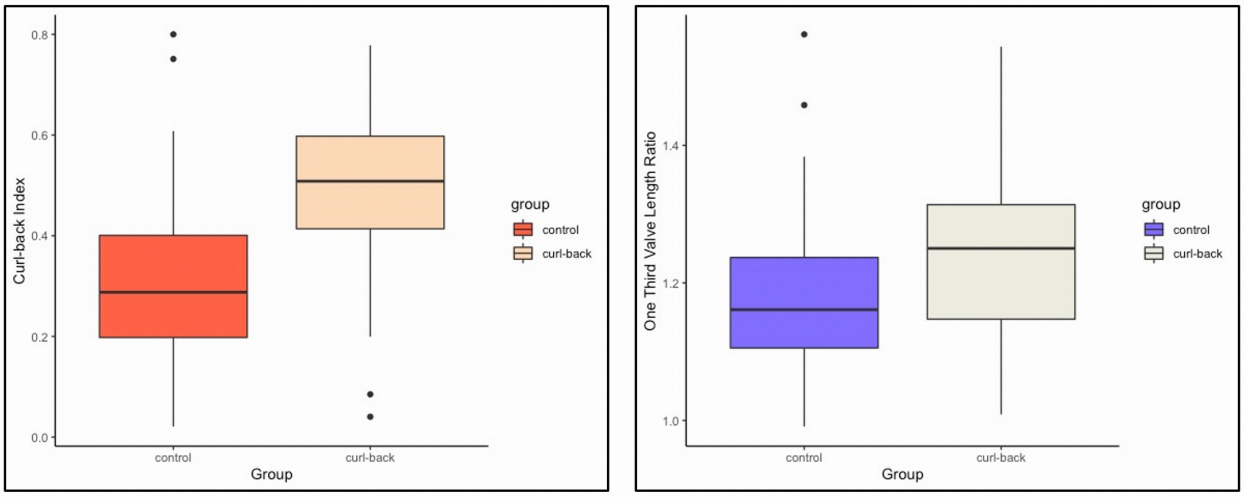


Figure 17. Boxplot comparisons of 'curl-back index' (left plot) and 'one third valve length ratio' (right plots) measured values for control (flat) and curl-back show there is a difference between cohorts.

Alpha and specific whorl expansion rates were calculated for shells of individual 11-month-old oysters in order to plot values in respective morphospaces (Figures 18 and 20).

Alpha Morphospace

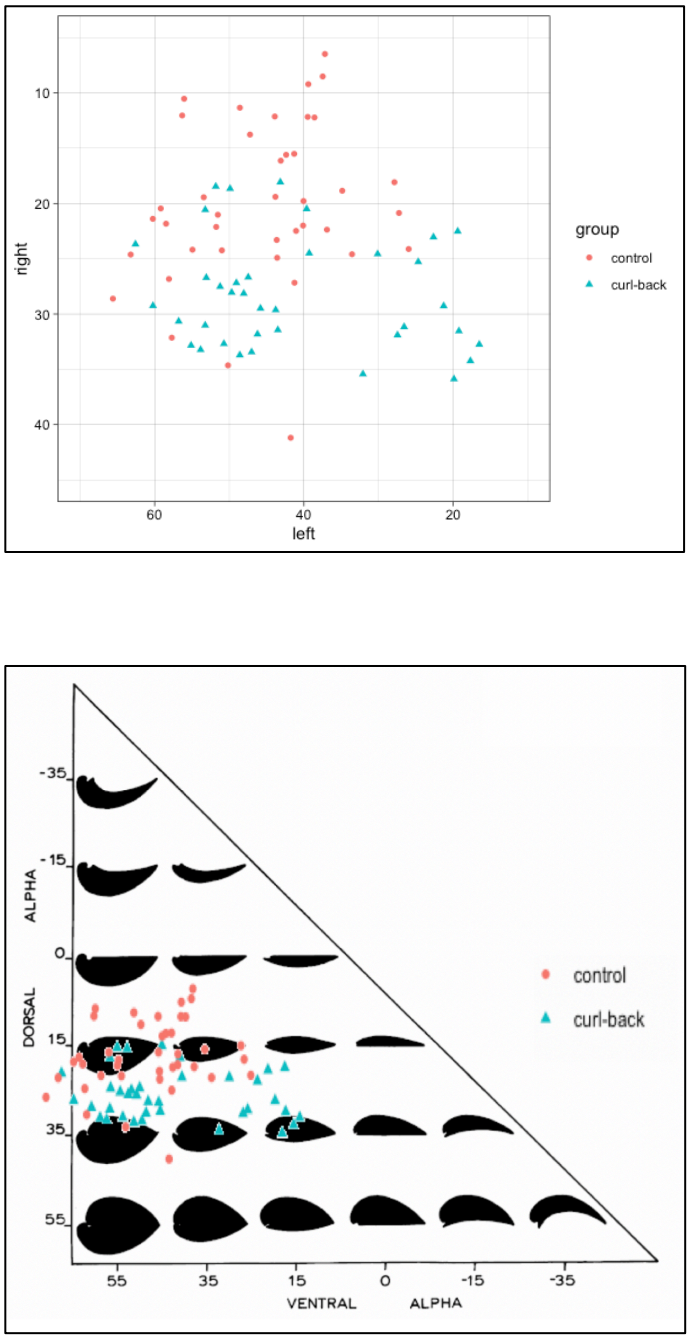


Figure 18. Calculated alpha angle values from measured eleven-month-old oysters from the two cohorts, control (red) and curl-back (green), were plotted and superimposed in McGhee's brachiopod alpha morphospace.

Comparisons of alpha values for right and left valves between the two cohorts using two tailed t-tests yielded *p*-values of 1.31E-07 and 0.0601, respectively.

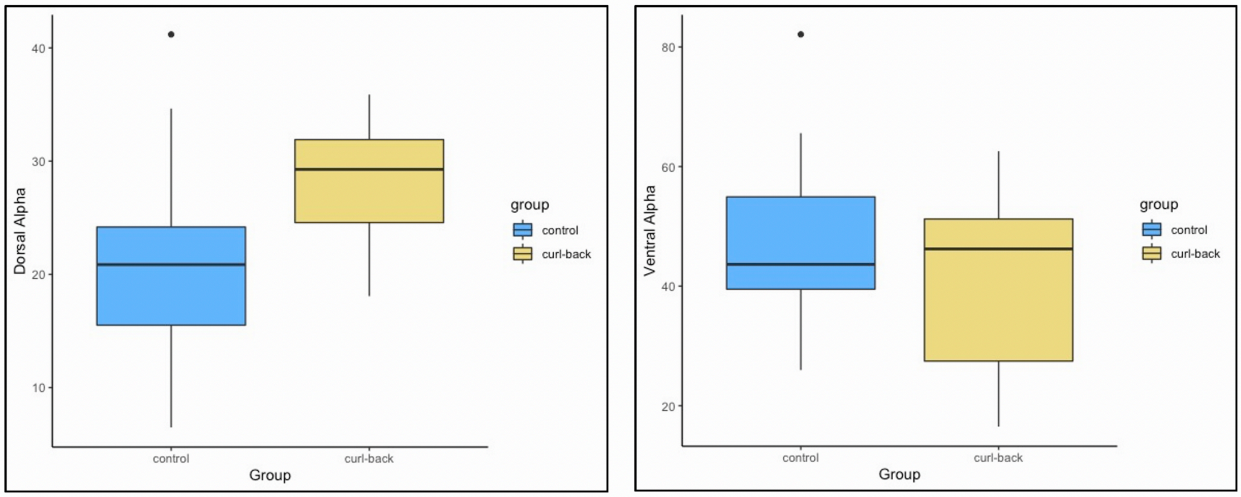
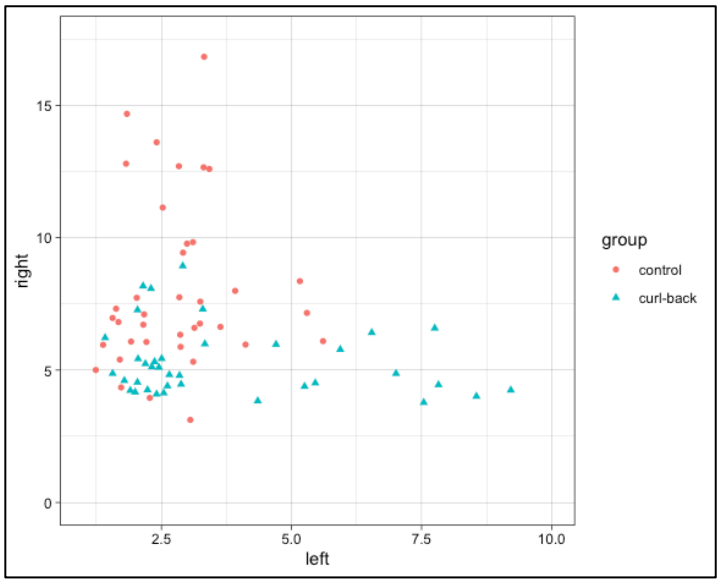


Figure 19. Boxplot comparisons of alpha values for control (blue) and curl-back (gold) cohorts. The plots on the left shows values for right (dorsal) valves and the plots on the right shows values for left (ventral) valves.

LogW Morphospace



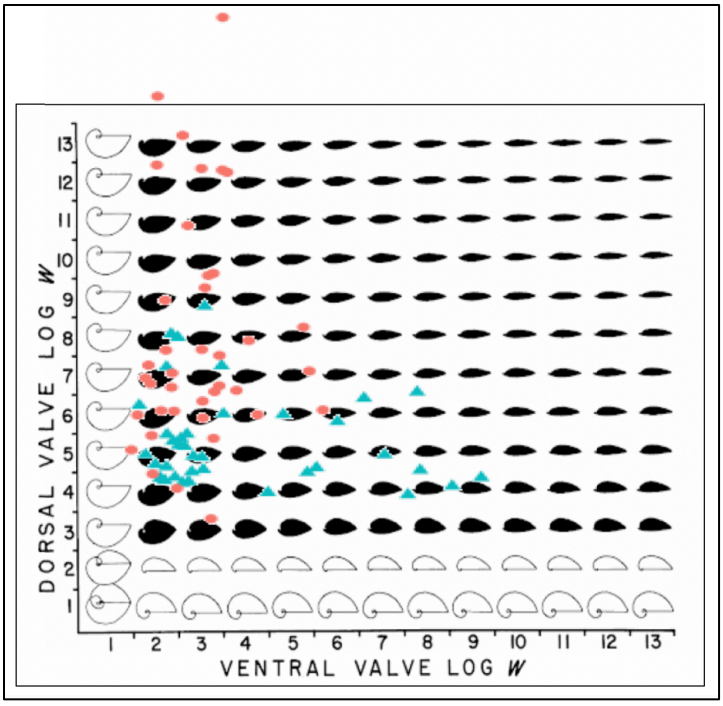


Figure 20. Calculated log whorl expansion rate (LogW) values from measured eleven-month-old oysters from the two cohorts, control (red) and curl-back (green), were plotted and superimposed in McGhee's LogW morphospace.

The plotted values, LogW, for right and left valves of individuals were then compared between the two cohorts using two tailed t-tests. There was a highly significant difference between LogW values of right valves (p = 8.947E-06), but, as in calculated alpha values, the difference in left valve LogW values was not as strong (p = 0.0206).

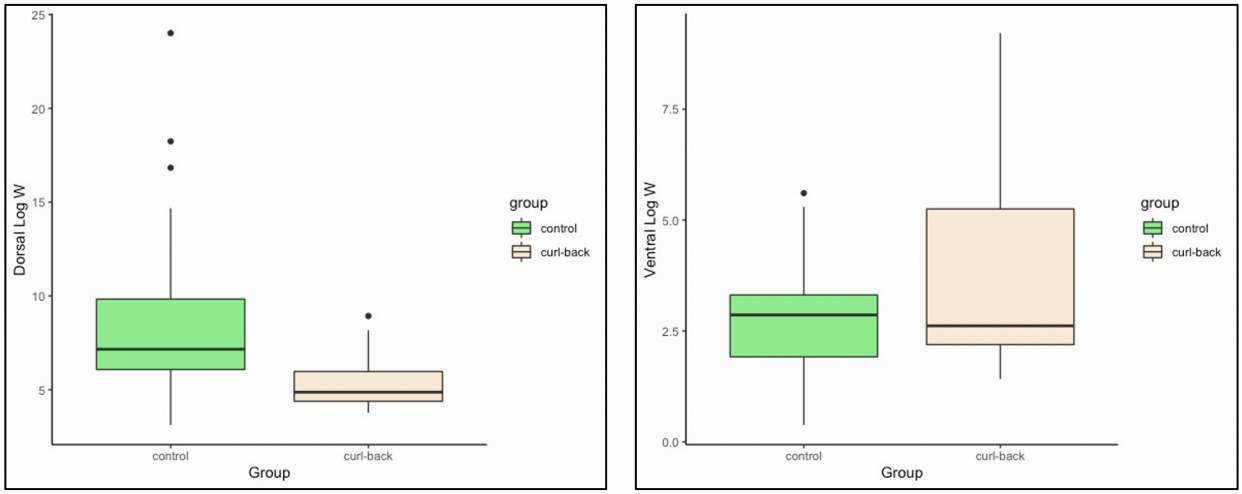


Figure 21. Boxplot comparisons of LogW values for control (green) and curl-back (tan) cohorts. The plots on the left shows values for right (dorsal) valves and the plots on the right shows values for left (ventral) valves.

DISCUSSION

The present study is the first to demonstrate that the curl-back shell abnormality has a strong genetic component in the eastern oyster, *Crassostrea virginica*. Although not calculated, heritability for this trait appears to be high as only two generations of selection for curl-back and normal (flat) hinge morphologies resulted in significant differences in phenotypes as measured in two cohorts of oyster seed using three independent methodologies: comparisons of relative heights of right and left valves at the point of greatest shell inflation, comparisons of relative lengths of right and left valves at the point of greatest shell inflation, and comparisons of the degrees to which

the whole right and whole left valves are curved based on a series of paired radii and angle measurements of each valve.

Selection for the curl-back trait was done by selecting oysters for crosses that had the most exaggerated right valve curvatures. This selection also had an effect, although not as strong, on shell morphologies of left valves, providing evidence that the valves of oysters grow together, influencing each other's form, rather than growing independently. The difference observed in left valves is therefore likely due to occlusion and this may suggest that morphologies are limited – i.e., all forms present in the alpha and LogW morphospaces may not be accessible to seed oysters. Personal observations of adult oysters suggest that adult forms exhibit more of the full range of possible morphologies, from concavo-convex to plano-convex to biconvex to convexi-plane to convexi-concave. However, the higher frequency of irregularities in shells of adult compared to seed oysters may make them less appropriate for morphometric analyses.

Although a morphometric analysis of the kind performed here has not previously been attempted for *C. virginica*, a similar methodology for comparing shells was described by Galtsoff (1964, pp 21-30). Here, a description can be found regarding how to compare forms by modelling valves as if they grew following a logarithmic spiral as outlined by Lison (1942).

Attempts to describe oyster shells using models where shells grow uniformly is inherently problematic because oysters do not always grow regularly and uniformly. Instead, oysters attach to substratum by cementation, and more carbonate is secreted relative to that by most other bivalved molluscs (Stanley 1975), making shells of oysters generally more irregularly shaped. In this study, this was ignored and shells were treated as though they curved smoothly and uniformly throughout their growth. This allows for their general shape to be discerned and is sufficient for the basic comparisons performed here.

Oysters that dramatically change their angle of coiling during early growth are not well suited for this analysis. For example, growth that begins extremely convex and then flattens, or vice versa, can have a marked effect on calculated values, however, this study did not correct for such effects. Instead, average alpha or LogW values were effectively recorded and plotted. A different problem occurs when measuring valves that are extremely flat, as it limits the number of paired angle and radii values that can be collected.

This study did not address environmental factors or growing techniques that may influence shell shape. Practices such as mechanical tumbling using rolling meshed cylinders are often used by growers to help promote harder, more rounded, and cupped shells (Robert et al. 1993; Cheney 2010). A few recent studies have also addressed how various aquaculture gear types may also influence oyster shell shape and growth environmentally. For example, Walton et al. (2013) compared the size and shape of diploid and triploid oysters deployed in four different gear types (LowPro bottom cages, adjustable long-line baskets, OysterGro floating cages, and floating bags) for five months. Triploids appeared to grow better in all gear types and growth was generally poorest in the bottom cages. Gear type did not influence observed fan ratios (shell length/shell height). Triploids had significantly higher cup ratios (shell width/shell height) than diploids, which suggests the difference was either related to genetic background or to growth rate.

Thomas, Allen, and Plough (2019) deployed sub-market sized triploid oysters in four gear treatments that varied in water column location and exposure to wave action for a four-month "finishing" period. Oysters reared in floating OysterGro™ gear were heavier than those in the subtidal bottom cage control gear, and the OysterGro™ floats also promoted shapes closest to the desired 3:2:1 height:length:width ratio.

Despite inherent imperfections of the present study, mapping oyster shell forms in morphospaces originally created for brachiopods is apt as bivalve molluscs and brachiopods, which occupy similar niches, offer an interesting comparative case study of the effects of bauplan and evolvability. Although they look similar, the relative orientations of bivalve mollusc and brachiopod anatomies are different: a brachiopod's two shells are top and bottom, whereas a bivalve mollusc's are left and right.

Brachiopods have depended upon the lochophore, through which they feed and respire, since they first evolved. The lochophore is a bilaterally symmetrical ciliary pump organ, which is arranged so the two brachia divide the mantle cavity into inhalant and exhalent chambers. The filaments of the lochophore are never fused and this feature restricts brachiopod morphology because, in order for this arrangement to be metabolically efficient, brachiopods cannot evolve conical forms and need to maintain a degree of symmetry, restricting their bauplan (Stanley 1968).

The bivalve molluscan bauplan, on the other hand, is not restricted in this way, and this difference might be a key reason for the relative success of bivalve molluscs today. The adaptive radiation of bivalves following the Permian extinction gave rise to 15 new infaunal superfamilies, all of which were siphon-feeding groups that probably descended from a non-siphonate Paleozoic group, such as the *Astartacea*. The radiation was probably a consequence of mantle fusion and siphon formation, which allowed the evolution of many new infaunal ways of life that had been inaccessible to Paleozoic bivalves and remained inaccessible to brachiopod groups (Stanley 1968; Clapham et al. 2006).

The presence of curl-back morphology and the general high variation of oyster shell form is likely an adaptation to life in different environments where wave action and water flow, sedimentation, and bottom types can vary dramatically. The fossil record of the extinct oyster,

Gryphaea, for example, indicates that significant changes to shape occurred in lineages in order to adapt to soft or hard substrates. *Gryphaea* oyster fossils appear to have generally become more coiled over time. John Maynard Smith (1958) summarizes the interpretation of general changes in the well-studied *Gryphaea* fossil record, reviewed by Westoll (1950), as follows:

> *"The relevant changes are an increase in size and in the degree of coiling of the lower valve of the shell, and a change from rigid attachment to a solid substrate to a form resting on and partly embedded in a soft substrate. The first two changes can also be understood as adaptations to life on a muddy rather than on a rocky bottom, since both tend to raise the opening of the shell, through which the animal feeds, clear of the mud."*

It might be therefore suggested that the curl-back trait, which is almost inverse in shape relative to the coiled *Gryphaea* fossils, is more adapted to rigid attachment. An educational publication on oyster communities in the Chesapeake Bay makes the distinction between "reef" and "seaside" oysters (see Harding, Mann, and Clark 1999, and references therein). Here, the shell of a "seaside" oyster is described as being "long and narrow with a slightly hooked hinge area." Although the frequency and distribution of the curl-back trait is not well documented, the streamlined shape appears to be an adaptation to high energy reefs, where the reduction of tidal force may help to keep cemented oysters from becoming dislodged.

REFERENCES

Abdelrahman, H., ElHady, M., Alcivar-Warren, A., Allen, S., Al-Tobasei5, 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., Hackett, P., Hanson, T., Hedgecock, D., Howard, T., Holland, L., Jackson, M., Jin, Y., Khalil, K., Kocher, T., Leeds, T., Li, N., Lindsey, L., Liu, S., Liu, Z., Martin, K., Novriadi, R., Odin, R., Palti, Y., Peatman, E., Proestou, D., Qin, G., Reading, B., Rexroad, C., Roberts, S., Salem, M., Severin, A., Shi, H., Shoemaker, C., Stiles, S., Tan, S., Tang, K. F. J., Thongda, W., Tiersch, T., Tomasso, J., Tri Prabowo, W., Vallejo, R., van der Steen, H., Vo, K., Waldbieser, G., Wang, H., Wang, X., Xiang, J., Yang, Y., Yant, R., Yuan, Z., Zeng, Q., and Zhou, T. 2017. Aquaculture Genomics, Genetics and Breeding in the United States: current status, challenges, and priorities for future research. *BMC Genomics,* (18): 191.

Beck, M.W., Brumbaugh, R.D., Airoldi, L., Carranza, A., Coen, L.D., Crawford, C., Defeo, O., Edgar, G.J., Hancock, B., Kay, M.C., Lenihan, H.S., Luckenbach, M.W., Toropova, C.L., Zhang, G., and Guo, X. 2011. Oyster Reefs at Risk and Recommendations for Conservation, Restoration, and Management. *BioScience*, 61(2):107-116.

Boudy, P., Degremont, L., Haffray, P. 2008. "The genetic basis of summer mortality in Pacific oyster spat and potential for improving survival by selective breeding in France". In: Summer mortality in the Pacific oyster: The Morest project. Ed. Samain, J. and McCombie, H. Ifremer; QUAE.

Cheney, D.P. 2010. Bivalve Shellfish Quality in the USA: from hatchery to the consumer. *Journal of the World Aquaculture Society*, 41(2): 192–206.

Clapham, M.E., Bottjer, D.J., Powers, C.M., Bonuso, N., Fraiser, M.L., Marenco P.J., Dornbos S.Q., and Pruss, S.B. 2006. Assessing the Ecological Dominance of Phanerozoic Marine Invertebrates. *Palaios*, 21: 431- 441.

Coon, S. L., Bonard, D.B., and Weiner, R.M. 1986. Chemical production of cultchless spat using epinephrine and norepinephrine. *Aquaculture*, 58: 255-262.

Dennett, D. 1995. Darwin's Dangerous Idea: Evolution and the Meanings of Life. New York: Simon & Schuster.

Ewart, J.W. and Ford, S.E. 1993. History and Impact of MSX and Dermo Diseases on Oyster Stocks in the Northeast Region. NRAC Fact Sheet 200 Northeastern Regional Aquaculture Center, North Dartmouth, MA, 1–8.

FAO Fisheries Department. 2017. Global Aquaculture Production 1950-2015. FAO – Rome http://www.fao.org/fishery/statistics/global-aquaculture-production/query/en

Ford, S.E. 1997. "History and Present Status of Molluscan Shellfishes from Barnegat Bay to Delaware Bay". In: The History, Present Condition, and Future of the Molluscan Fisheries of North and Central America and Europe. Ed. MacKenzie, Jr., C.L., Burrell, Jr., V.G., Rosenfield, A., Hobart, W.L. NOAA Technical Report NMFS 127. Seattle, Washington: U.S. Department of Commerce, 1: 119-140.

Ford, S.E. and Tripp, M.R. 1996. "Diseases and defense mechanisms". In: The Eastern oyster. Ed. Kennedy, V.S., R. I. E. Newell, R.I.E, and Eble, A. F. College Park, MD: Maryland Sea Grant College Program, Univ. of Maryland. pp. 581–660.

Galtsoff, P. 1964. The American oyster: *Crassostrea virginica* Gmelin. United States Department of the Interior, Fish and Wildlife Service, Bureau of Commercial Fisheries.

Gerber, S. 2017. The geometry of morphospaces: lessons from the classic Raup shell coiling

model. *Biological Reviews*, 92: 1142-1155.

Guo, X., Wang, Y., DeBrosse, G., Bushek, D. and Ford, S.E. 2008a. Building a Superior oyster For Aquaculture. *The Jersey Shoreline*, 25(1): 7-9.

Guo, X., Y. Wang, L. Wang and J-H Lee. 2008b. "Oysters." in: Genome Mapping and Genomics in Fishes and Aquatic Animals, Ed. Kocher, T.D. and Kole, C. Springer, Berlin, 163 – 175.

Hammer, O. 2016. The Perfect Shape: Spiral Stories. Cham, Switzerland: Springer Nature.

Harding, J., Mann, R.L., and Clark, V.P. 1999. Shell Games. Educational series No. 47. Virginia Institute of Marine Science, William & Mary.

Hedgecock, D. 2011. "Genetics of Shellfish on a Human-Dominated Planet." In Shellfish Aquaculture and the Environment. Ed. Shumway, S.E. John Wiley and Sons, Inc., 339-357.

Hollenbeck, C.M. and Johnston, I.A. 2018. Genomic Tools and Selective Breeding in Molluscs. *Frontiers in Genetics*, 9: 253.

Kecinski, M., Messer, K.D., Knapp, L., and Shirazi, Y. 2017. Consumer Preferences for Oyster Attributes: field experiments on brand, locality, and growing method. *Agricultural and Resource Economics Review*, 46 (2): 315–37.

Kennedy, V. S. 2014. Technological Constraints During the First 40 Years of Eastern Oyster *Crassostrea virginica* Aquaculture. *Reviews in Fisheries Science & Aquaculture*, 22(1): 55–72.

Lison, L. 1942. Caractéristiques Géométriques Naturelles des Coquilles de Lamellibranches. *Bulletin de la Academie Royale de Belgique*, Bruxelles, 5e sér., t. XXVIII.

Loosanoff, V.L. and Davis, H.C. 1963. Rearing of Bivalve Mollusks. Advances in Marine Biology, 1: 1-136.

Mahon, G.A.T. 1982. Selection Goals in Oyster Breeding. *Aquaculture*, 33(1-4): 141-148.

Maslo, B. 2014. "Cascading effects of global oyster reef loss on the health of estuaries." In Coastal Conservation. Ed. Maslo, B. and Lockwood, J. University Printing House, Cambridge, UK: Cambridge University Press. 131-160.

Maynard Smith, J. 1993. The Theory of Evolution. Cambridge, UK; New York, USA; and Melbourne, AU: Cambridge University Press.

McGhee, G.R. 1980. Shell Form in the Bioconvex Articulate Brachiopoda: a geometric analysis. *Paleobiology*, 6 (1): 57-76.

McGhee, G.R. 1999. Theoretical morphology: the concept and its applications. New York: Columbia University Press.

McGhee, G.R. 2007. The Geometry of Evolution: adaptive landscapes and theoretical morphospaces. Cambridge: Cambridge University Press.

Mizuta, D.D. and Wikfors, G.H. 2018. Seeking the Perfect Oyster Shell: a brief review of current knowledge. *Reviews in Aquaculture*. 11(3): 586-602.

Mondal, S. and Harries, P.J. 2016. Phanerazoic Trends in Ecospace Utilization: the bivalve perspective. *Earth Science Review*, 152: 106-118.

Moulton, D.E., Goriely, A., and Chirat, R. 2012. Mechanical Growth and Morphogenesis of Seashells. *Journal of Theoretical Biology*, 311: 69-79.

Okabe, T. and Yoshimura, J. 2017. Optimal Designs of Mollusk Shells from Bivalves to Snails. *Scientific Reports*, 7: 42445.

Pigliucci, M. 2012. "Landscapes, Surfaces, and Morphospaces: what are they good for?" In: The Adaptive Landscape in Evolutionary Biology. Ed. Svensson, E. and Calsbeek, R. Oxford,

UK: Oxford University Press.
Plutynski, A. 2008. The Rise and Fall of the Adaptive Landscape? *Biology and Philosophy*, 23: 605-623.
Proestou, D.A., Vinyard, B.T., Corbett, R.J., Piesz, J., Allen Jr, S.K., Small, J.M., Moss, J., Li, C., Liu, M., DeBrosse, G., Guo, X., Rawson, P., Gomez-Chiarri, M. 2016. Performance of Selectively-bred Lines of Eastern Oyster, *Crassostrea virginica*, Across Eastern US Estuaries. *Aquaculture*, 464: 17-27.
Raup, D.M. 1966. Geometric Analysis of Shell Coiling: general problems. *Journal of Paleontology*, 40: 1178-1190.
Robert, R., Trut, G., Borel, M., Maurer, D. 1993. Growth, Fatness and Gross Biochemical Composition of the Japanese Oyster*, Crassostrea gigas*, in Stanway Cylinders in the Bay of Arcachon, France. *Aquaculture*, 110, 249–261.
Rothschild, B.J., Ault, J.S., Goulletquer, P., and Heral, M. 1994. Decline of the Chesapeake Bay Oyster Population: a century of habitat destruction and overfishing. *Marine Ecology Progress Series*, 111: 29-39.
Scales, H. 2015. Spirals in Time: the secret life and curious afterlife of seashells. London and New York: Bloomsbury Sigma.
Simpson, G.G. 1944. Tempo and Mode in Evolution. New York: Columbia University Press.
Stanley, S.M. 1968. Post-Paleozoic Adaptive Radiation of Infaunal Bivalve Molluscs: a consequence of mantle fusion and siphon formation. *Journal of Paleontology*, 42 (1): 214-229.
Stanley, S.M. 1970. Shell Form and Life Habits in the Bivalvia (Mollusca). Boulder, CO: The Geological Society of America, Inc.
Stanley, S.M. 1975. Adaptive Themes in the Evolution of the Bivalvia (Mollusca). *Annual Review of Earth and Planetary Sciences*, 3(1): 361-385.
Stone, J.R. 1996. The Spirit of D'Arcy Thompson Dwells in Empirical Morphospace. *Mathematical Biosciences*, 142:13-30.
Thomas, L.L., Allen, S.K., and Plough, L.V. 2019. The Effect of Aquaculture Gear on the Growth and shape of the Oyster *Crassostrea virginica* During a "Finishing Period" in Chesapeake Bay, USA, *Aquaculture*, 508:1-9.
Thompson, D.W. 1917. On Growth and Form. Cambridge, UK: Cambridge University Press.
Walton, W.C., Rikard, F.S., Chaplin, G.I., Davis, J.E., Arias, C.R., and Supan, J.E. 2013. Effects of Ploidy and Gear on the Performance of Cultured Oysters, *Crassostrea virginica*: survival, growth, shape, condition index and *Vibrio* abundances. *Aquaculture*, 414-415: 260-266.
Westoll, T.S. 1950. Some Aspects of Growth Studies in Fossils. *Proceedings of the Royal Society of London Series B, Biological Sciences*, 137(889): 490-509.
Wright, S. 1932. The Roles of Mutation, Inbreeding, Crossbreeding, and Selection in evolution. *Proceedings of the Sixth International Congress on Genetics*, 1 (8): 355–66.

CHAPTER 2: INHERITANCE OF SHELL PIGMENTATION COLOR MORPHS

INTRODUCTION

Background

In contrast to shell shape, consumer preferences for shell color are not well known and may be heterogenous or regional. For example, Kang et al. (2013) reported that Pacific oysters with darkly pigmented mantle tissues are favored by consumers in South Korea, while Nell (2001) had reported that golden-shelled Pacific oysters are preferred in New Zealand. Nevertheless, a better understanding of the inheritance and molecular basis of shell color can provide farmers an additional opportunity for product differentiation, and several oyster farms already actively use the particular colors of their products to advertise or name their brand(s), e.g., Peconic Golds (Cutchogue, New York, https://www.peconicgoldoysters.com/), Black Pearls (Quadra Island, British Columbia, CA, https://outlandish-shellfish.com/), Carlsbad Blondes and Lunas (Carlsbad, California, https://carlsbadaquafarm.com/oysters/).

Review of Studies of Pigmentation in the Pacific Oyster

To date, no studies have reported on the inheritance of color morphs in *Crassostrea virginica*, however color in the Pacific oyster, *C. gigas*, is relatively well studied. Although the literature on pigmentation in *C. gigas* is vast and growing quickly, at this time there is not a comprehensive understanding of modes of inheritance. Some studies even appear to contradict others and there is still much to be discovered.

The first study to explicitly describe genetic control of mantle edge and shell pigmentation in the Pacific oyster, *Crassostrea gigas,* was by Brake et al. (2004), although Imai and Sakai (1961) had found differences in shell and mantle color by strain and Nell (2001) had previously reported

that true-breeding lines of Pacific oysters with golden mantles were successfully developed after just a few generations of selection. Brake et al.'s study investigated three cohorts of oysters consisting of 30-40 pedigreed families, each of which was created by crossing one sire from an existing selected group with four dams from another. For each of the three cohorts, a random sample of 50 oysters from each of the best performing 15 families (750 oysters per cohort) were scored for mantle edge and shell pigmentation on a continuous scale from 1 to 4. In all cohorts, family effects significantly affected both mantle edge and shell pigmentation, indicating a high degree of genetic control for both traits. This showed there was potential to select for mantle edge and shell color, and shell pigmentation was not correlated with performance, which suggested it could be selected for independently.

These initial breeding studies considered variation to likely be a quantitative trait controlled by a large number of genes. However, later studies suggested that only a few major genes controlled shell color. A study by Hedgecock, Grupe, and Voigt (2006) was initiated after observing what was considered to be possibly heritable variation in shells of an F2 family. The shape and color of one hundred and seventy-nine oysters from this family were photographed and measured. Principal component analyses were then used to consolidate the measured traits, and mapping, utilizing 59 microsatellite markers on 11 linkage groups was done on both principal components and some individual traits. A quantitative trait loci (QTL) for pigment saturation of the left valve was identified that explained 32% of shell color variance. Shell pigmentation appeared to be additive and QTL were suggested for use in selection for oysters with more desirable shell characteristics.

In 2009, Evans, Camara, and Langdon reported that a small number of genes likely control discrete color variants in *C. gigas*. The authors collected wild oysters, strip spawned 52

individuals, and paired them randomly to create 26 unrelated full-sib families. The total left-shell pigmentation was significantly affected by family and was mostly additive in nature. The offspring in two families segregated into phenotypically distinct "lighter" and "darker" shell groups in a 3:1 ratio, which suggested that a single major gene is segregating in these families with the "light" allele being dominant over the "dark" allele. The estimated broad-sense heritability was 0.9, which was in agreement with Brake et al. (2004), who found the broad-sense heritability to range from 0.67 to 0.90. Thirty-two percent of variation was attributed to dominance.

Kang et al. (2013) selectively bred for the presence (black) and absence (white) of shell and mantle pigmentation. Oysters were collected from six Korean farms and the darkest and lightest individuals from each were used to produce six full sibling F1 families by single pair mating. The degree of shell and mantle edge pigmentation of F1s was scored by eye on a continuous scale from one to six and one to ten, respectively. The siblings with the highest scores for these traits were then crossed together at the age of twelve months to generate two and four F2 families selected for black and white phenotypes, respectively. Comparisons of coloration of F0 and successive F1 and F2 generations had increasing proportions of the selected phenotype. After two generations of selection for these traits, estimates of heritability were 0.41 and 0.77 for black and white shell color and 0.27 and 0.08 for black and white mantle edge pigmentation. Shell and mantle pigmentation were found to be correlated.

The location of the quantitative trait locus (QTL) identified by Hedgecock, Grupe, and Voigt (2006) was confirmed in 2014 by a study conducted by Zhong et al. They reported sets of QTLs linked to shell pigmentation and glycogen content in an F1 full sibling family based on a sex-averaged linkage map. Pairwise comparisons using Pearson's correlation revealed glycogen content and shell pigmentation were positively correlated with growth-related traits. One QTL was

identified to be related to shell pigmentation on linkage group nine (LG9). At this locus, paternal and maternal alleles explained 6.75 and 17.44 percent of phenotypic variation. In agreement with previous studies, the authors concluded that selective breeding could be highly effective for shell pigmentation in Pacific oysters.

Also in 2014, Ge et al. identified quantitative trait loci (QTL) for shell pigmentation using bulked segregant analysis (BSA) along with amplified fragment length polymorphisms (AFLPs). To perform the BSA, 91 oysters were produced by crossing a single lightly pigmented female with a single darkly pigmented male obtained from a wild population. These 91 oysters showed a continuous, bimodal distribution and the 9 darkest and 9 lightest oysters were combined separately to produce two pools of DNA. DNA was digested with an ECORI/MseI combination and fragments were ligated to double-stranded adaptors. The sequence information associated with pigmentation was used to design locus specific primers and these were used to amplify genomic DNA from the two parents and 18 progeny.

Two hundred and twenty-five primer pairs were used to screen the AFLPs and a total of 7,425 bands were scored. Of these, 2,005 (27%) differed between the parents used in the cross and 56 primer pairs produced polymorphic bands between parents and between the two bulks. These primers together produced 77 polymorphic bands between parents and between the two bulks. Seven polymorphic bands, produced by 6 primer pairs, were found to be associated with color, and only two recombinants were found among the 18 individuals. Seven AFLP associated sequences were examined using BLAST; two were highly homologous with known proteins: integrin alpha-8 and RPTP (receptor-type tyrosine-protein phosphatase alpha).

All 91 F1s were examined using these seven markers and they were mostly specific to one group and they occurred or did not occur at a ratio of about 1:1. The markers were mapped onto

a single linkage group that spanned 11.4 cM. A significant QTL that explained 79.5% of the total phenotypic variance was detected. To develop SCAR markers, the seven markers were sequenced and used to design primer pairs for amplification. Of the seven markers, one showed differential amplification between the dark and light pools. The SCAR marker was called SP-170. They proposed a single gene mode of inheritance for producing color: AA white; Aa light; and aa dark.

Following up on this study, Ge et al. (2014b) then set out to determine if the genetic basis for golden shells was the same as that for black and white and, if not, what loci control golden shell color. They used one-year old oysters from nine full-sib families to conduct crosses. They used very golden, very white, and very dark oysters for the crosses. One x one crosses using each combination of the three colors were performed with replication (GxG, WxW, BxB, GxW, GxB, BxW). The progeny from each cross was described as either having a white or golden shell background color and the black foreground was rated categorically from 0-4. Twenty-three full-sib families were examined. Dark coloration was found on shells with both white and golden background coloration. Golden color, when present, was found to be distributed all over the shell, whereas the dark coloration existed on a continuum from mottled to full. Gold was dominant over white and each of the three families produced using golden x golden parents produced golden and white progeny at a ratio approaching 3:1, while no white x white crosses produced any golden offspring.

The authors concluded that the two background colors (white and golden) are controlled by one locus, with one allele for golden shells being dominant to the white allele. These results were consistent with the observation of Nell (2001), who obtained true breeding oysters with golden shell and mantle tissue after just a few generations of selection. The authors also suggested that the degree of black foreground pigmentation was related to background pigmentation: there

were more golden offspring with lighter foreground pigmentation and more white offspring with darker foreground pigmentation. This suggested there may be epistatic effects. The identified AFLPs linked to shell color were sequenced and one sequence-characterized amplified region (SCAR) was identified and mapped to the *C. gigas* genome.

Ge et al. (2015) found that the results from the bulked segregant analysis (BSA) and amplified fragment length polymorphism (AFLP) analyses were in agreement with their study of foreground and background pigmentation; both suggested that shell background color in the Pacific oyster is under a one locus system with the allele for golden color being dominant for the allele for white. The seven AFLP markers they produced were tightly linked and came from the golden dam that was involved in a 1x1 cross with a single white shelled male. The progeny for the cross was about 1:1 white/golden and it was proposed that the genotypes for the parents were Gw x ww. The single locus SCAR marker, SNP marker, and SSR marker were proposed to be later used in marker assisted selection (MAS) for color. The SCAR marker (named SP-170) they produced in their previous study (Ge et al. 2014b) was well correlated with light and dark pigmentation, but did not display linkage with the seven AFLP markers and three single locus markers. This suggests the golden and black shell colors may be controlled by different loci and therefore have independent modes of inheritance.

Song et al. (2016) surveyed one wild population and four pure breeding shell color variants of *C. gigas,* produced after four generations of successive selection for shell color, using amplified length polymorphisms (AFLPS). Eleven loci were identified by both hierarchical analysis and a Bayesian likelihood method. Of these, many were strongly associated with one of the four color variants, gold, white, black, purple or the wild population. The four shell color variants of *C. gigas* maintained genetic diversity similar to the wild population, but the white shell variant was

observed to have slightly lower genetic diversity than the others. This led the authors to hypothesize that the white shell genotype may be homozygous recessive, which was consistent with the previous observation of Ge et al. (2014) who found that the white shell color was controlled by a recessive allele.

In 2017, Xu et al. conducted a study to determine if growth and shell color of black shelled Pacific oyster can be improved by genetic selection. A black shell color strain was produced by four-generations of successive selective breeding. Oysters from six F4 families with darker shells were used for broodstock for a cross with a nested mating design including 30 dams and 10 sires to produce 30 full-sib families of black shelled oysters. The larvae were cultured together. After 24 months, 432 of the offspring from crosses were harvested randomly. Left shells were processed using two quantitative indexes of pigmentation - the first measured variation in pigmentation intensity and coverage of the entire shell, while the second described color in lightness, hue and saturation. Oyster shell color was treated as a as a quantitative trait. Next, four panels of multiplex primers containing 10 microsatellites were used to unambiguously assign parentage to 410 (94.9%) of the 432 progeny assayed. Heritability as well as genetic and phenotypic correlations for shell color and growth traits were estimated using linear mixed models. Shell color was determined to be highly heritable (0.52–0.69), however, shell color was also negatively correlated with growth, which was not in agreement with the findings of Brake et al. (2004) or of Zhong et al. (2014).

Song et al. (2018) produced two F1 full-sib mapping families of *C. gigas* with golden and white shell color variants from strains that had been successively selected for three generations for pure and stable shell color and fast growth. Photoshop CS6 was utilized to measure the values of red and yellow as well as degree of lightness of the shells of the progeny, and microsatellite markers were used to test genetic segregation in the two families. Polymorphic markers were

subsequently genotyped and identified QTLs associated with golden shell color, Mg and Zn contents, and growth traits were mapped to two constructed sex-averaged linkage maps. In total, seventeen QTLs and fourteen associated genes were found. Of these, two QTLs associated with golden shell color were identified, which could explain 20.2 and 10.5% of the observed phenotypic variation. QTLs correlated with shell color were found on LG8 in and on LG10. A couple of genes flanking the QTL markers were confirmed using RT-qPCR to be consistently up-regulated in golden *C. gigas*. These genes included PTPPK (receptor-type tyrosineprotein phosphatase kappa-like) and KMT2D (histone-lysine N-methyltransferase 2D).

Yu et al. (2015) had confirmed that melanin was present in shells and mantles of *C. gigas*, and Xing et al. (2018) then utilized this finding to quantity mantle edge pigmentation by melanin content using UV–visible spectrophotometry. The correlation between mantle edge and shell pigmentation was analyzed in twenty-four-month old individuals from a white-shell strain of C. gigas produced after four generations of selection. Values of absorbance with the wavelength setting at 500 nm were determined and were used to calculate total melanin content in pigmented mantle edge. Photographs of shells were taken and the optical density (OD) values of all images were measured by using macros as values of pigmentation. Multiplex PCR was then performed using 11 microsatellite loci to successfully assign parents to 460 progeny, and then estimate heritability and phenotypic and genetic correlations.

The heritability of mantle edge pigmentation (0.215 ± 0.092) was slightly higher than that of shell pigmentation (0.156 ± 0.078), and the genetic correlation between mantle edge and shell pigmentation was very high and positive (0.980 ± 0.094). The heritability estimate of mantle edge pigmentation was similar to the results of Kang et al. (2013), but the heritability estimate of shell pigmentation was lower than reported by both Evans et al. (2009) and Kang et al. (2013). In one

family, the ratio of light to darker individuals was 3:1, which the authors suggested could be explained by two alleles at a single locus with “lighter” allele being dominant over the “dark” allele. This result was very similar to what Evans et al. (2009) found in one family produced using wild type parents, and the authors hypothesized that a dominant allele may code for the lighter shell and a recessive allele may code for the darker shell.

Wang et al (2018) established strains of black and white shelled oysters after four generations of selection for white and black shells. Two reciprocal-cross F1 families were then generated (white shell ♀ × black shell ♂ and black shell ♀ × white shell ♂) by mating single pairs of parents. Twelve moth old progeny were sampled, and photographed to quantify lightness, redness, and yellowness. Microsatellite multiplex polymerase chain reaction (PCR) panels for 6 loci were then used for parental assignment. Next, two genetic linkage maps, one for each family, were constructed using 1061 single nucleotide polymorphism markers developed from expressed sequence tags, which together produced a map with 351 EST-SNP markers distributed across 10 linkage groups. One map detected 11 quantitative trait loci (QTLs) for shell color traits, explaining about 11.76% of the phenotypic variance, while the other map detected 9 QTLs for shell color traits that explained about 12.84% of the phenotypic variance. Three shared-QTLs associated with shell color traits were identified on the integrated map, and one of them was significantly homologous to *C. gigas* calmodulin-like protein.

Han et al. (2019) reported that a small number of Pacific oysters with solid orange shells were produced accidently by a cross using broodstock with either completely black or purple shells. They then compared the genetic variability of the orange shell variant after three generations of successive selection to four wild populations using 20 microsatellite loci and mitochondrial cytochrome oxidase I sequences (mtCOI). The orange-shell line was determined to

be a rare variant that exhibited very low genetic variability, but the variability was maintained through three generations of selection. Following this study, Han et al. (2020) investigated the inheritance of the orange shell color trait by conducting crosses of the orange variant with white and black color variants. The authors reported that the orange shell color trait is a recessive trait compared to black and white, and is likely determined by two independent recessive genes. In agreement with Ge (2014b, 2015), the authors also reported that some different shell colors are inherited independently; they found that the genetic loci controlling orange shell had no effects on the inheritance or formation of black foreground pigment.

Objective of Study of Color Morphs in the Eastern Oyster

The objective of this study is to understand the modes of inheritance of shell color in the Eastern oyster by studying segregation in genetic crosses.

METHODS

Color Morph Crosses

Four color morphs found in seed oysters were described based on patterns they exhibit on their right valves: positive stripe, negative stripe, light, and full (Figure 22). Crosses were attempted using oysters from each color morph with oysters with its same as well as each of the other color morphs in order to attempt several biparental crosses for each of 10 possible pairings (Figure 22).

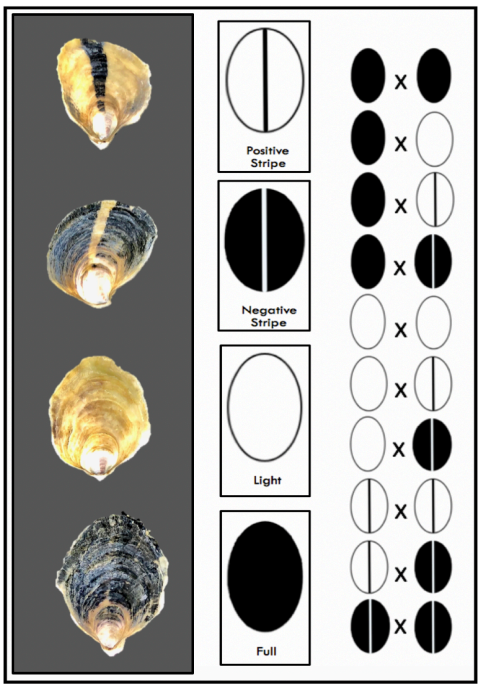


Figure 22. Four color morphs were described: positive stripe, negative stripe, light pigmentation, and full pigmentation

For each pair of color morphs, larvae from one 1x1 cross was cultured in a 15 liter bucket and larvae from several 1x1 crosses (either three or four families) were combined together to be cultured in a common 200 liter tank. Cultures were initially stocked with 50,000 - 100,000 larvae in each 15 liter bucket and 900,000 - 1,000,000 larvae in each 200 liter tank (about 5 larvae per mL) on day two. Larval culture tanks were kept at 20-25°C and complete seawater changes were performed with 1-micron (1 µm) filtered sweater (FSW) every other day (Figures 23 and 24).

Figure 23. For each of the ten possible paired combinations of color morphs, combined families were cultured in 200 liter tanks.

Figure 24. Single family crosses (1x1) were cultured in 15 liter buckets

Each day, the larvae were fed a mixed diet of cultured *Tisochrysis lutea* (T-iso)*, Isochrysis galbana* (C-iso)*, Pavlova lutheri* (P-mono)*,* and *Nannochloropsis sp.* (Nanno B) microalgae, as well as Shellfish Diet 1800® commercial algae paste (Reed Mariculture Inc.); the daily amount of algae the cultures were fed was based on the cultures age, size, and (Loosanoff and Davis 1963).

When larvae developed a foot and were ready to metamorphose, they were treated with epinephrine to induce pediveligers to metamorphose (Coon, Bonard and Weiner 1986), and were then placed into downweller silos with 200 μm nylon mesh screened bottoms in the 200 liter larvae tanks.

After about two weeks in downwellers, oysters were placed into a nursery where each group was stocked in separate upweller silos at roughly equal densities. When oysters were >5 mm in length, they were placed in separate ADPI OBC bags with 4mm mesh and deployed on Rutgers University's intertidal experimental oyster farm (rack-and-bag culture).

After about 3 months of growth in the culture bags, color morphs of individual oysters produced by each of the color crosses were determined and percentages of color morphs in the progeny were compared.

RESULTS

Proportions for all twenty proposed crosses were not obtained due to the unavailability of gametes, culture crashes, or cross contamination (Table 1).

Table 1. Twelve of the attempted twenty paired color morph crosses were successfully cultured in the hatchery

	Combined families	Single family
Dark x Dark		
Dark x Positive		
Dark x Negative	✓	
Dark x Light		✓
Positive x Positive	✓	✓
Positive x Negative	✓	✓
Positive x Light	✓	✓
Negative x Negative	✓	
Negative x Light	✓	✓
Light x Light		✓

The number and percentage of individual oysters assigned to each of the four color morphs as well as an ambiguous category (where the color morph was not obvious) are reported in Tables 2 through 13 for each of twenty attempted crosses. Where possible, the color morphs of at least one hundred oysters per cross were assigned.

Dark x Negative

3 families

Table 2. Number and percent of total counted of each of the 5 color morphs in the progeny of dark x negative cross (3 pooled families)

Progeny Color Morph	Number Counted	Percent of Total
Dark	54	13.70%
Positive	0	0.00%
Negative	250	63.40%
Light	90	22.80%
Ambiguous	0	0.00%

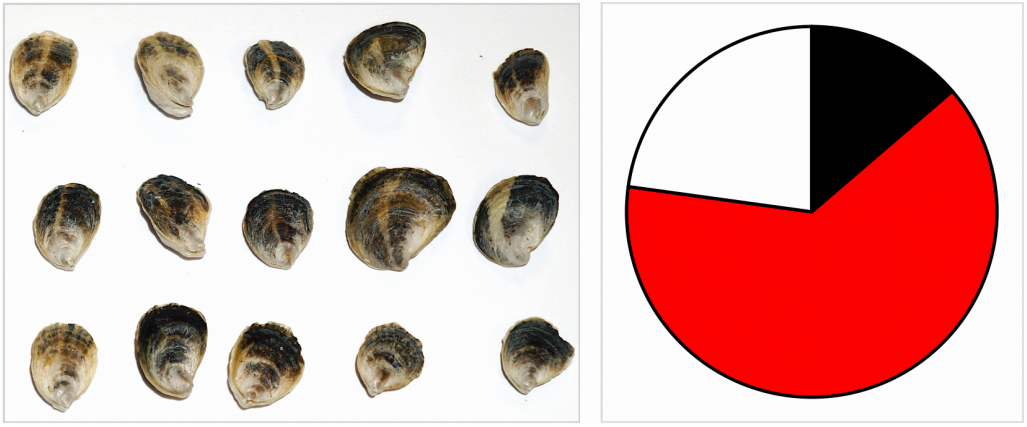


Figure 25. Left: Photo of representative shells from dark x negative cross (3 pooled families). Right: Percentages of color morphs found in the progeny are represented as parts of pie chart. Key to colors - black = dark; white = light; green = positive stripe; red = negative stripe; blue = ambiguous.

Dark x Light

1 family

Table 3. Number and percent of total counted of each of the 5 color morphs in the progeny of dark x light cross (1 family)

Progeny Color Morph	Number Counted	Percent of Total
Dark	25	25.80%
Positive	1	1.00%
Negative	3	3.10%
Light	68	70.10%
Ambiguous	0	0.00%

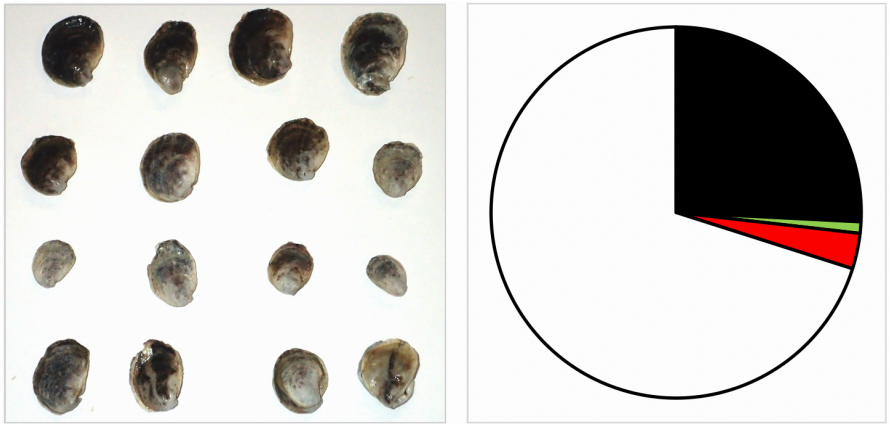


Figure 26. Left: Photo of representative shells from dark x light cross (1 family). Right: Percentages of color morphs found in the progeny are represented as parts of pie chart. Key to colors - black = dark; white = light; green = positive stripe; red = negative stripe; blue = ambiguous

Positive x Positive

4 families

Table 4. Number and percent of total counted of each of the 5 color morphs in the progeny of positive x positive cross (4 pooled families)

Progeny Color Morph	Number Counted	Percent of Total
Dark	0	0.00%
Positive	115	58.90%
Negative	30	15.40%
Light	17	8.70%
Ambiguous	33	16.90%

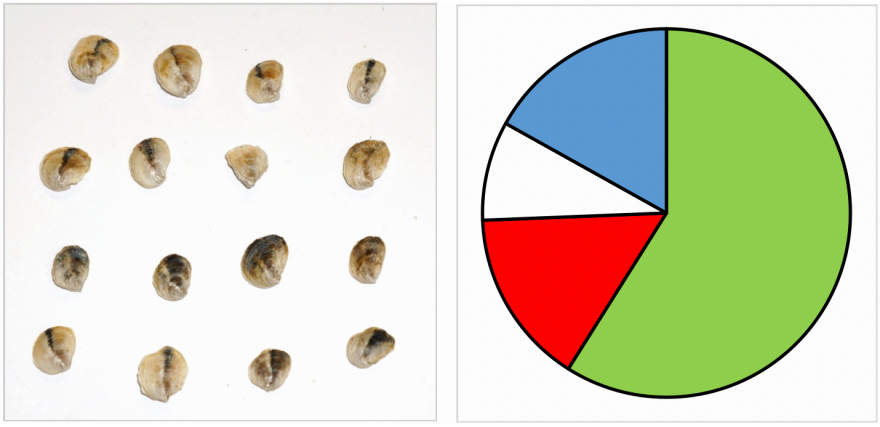


Figure 27. Left: Photo of representative shells from positive x positive cross (4 pooled families). Right: Percentages of color morphs found in the progeny are represented as parts of pie chart. Key to colors - black = dark; white = light; green = positive stripe; red = negative stripe; blue = ambiguous

Positive x Positive

1 family

Table 5. Number and percent of total counted of each of the 5 color morphs in the progeny of positive x positive cross (1 family)

Progeny Color Morph	Number Counted	Percent of Total
Dark	33	3.60%
Positive	582	63.80%
Negative	59	6.50%
Light	10	1.00%
Ambiguous	226	24.80%

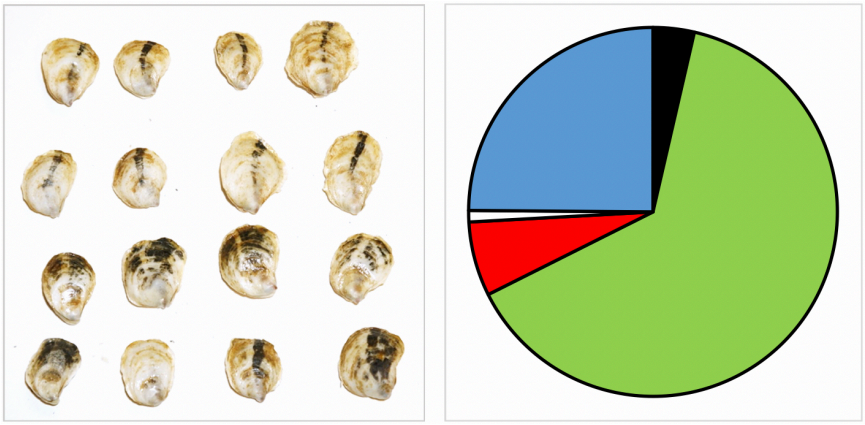


Figure 28. Left: Photo of representative shells from positive x positive cross (1 family). Right: Percentages of color morphs found in the progeny are represented as parts of pie chart. Key to colors - black = dark; white = light; green = positive stripe; red = negative stripe; blue = ambiguous

Positive x Negative

4 families

Table 6. Number and percent of total counted of each of the 5 color morphs in the progeny of positive x negative cross (4 pooled families

Progeny Color Morph	Number Counted	Percent of Total
Dark	4	2.20%
Positive	74	41.80%
Negative	63	35.60%
Light	36	20.30%
Ambiguous	0	0.00%

Figure 29. Left: Photo of representative shells from positive x negative cross (4 pooled families). Right: Percentages of color morphs found in the progeny are represented as parts of pie chart. Key to colors - black = dark; white = light; green = positive stripe; red = negative stripe; blue = ambiguous

Positive x Negative

1 family

Table 7. Number and percent of total counted of each of the 5 color morphs in the progeny of positive x negative cross (1 family)

Progeny Color Morph	Number Counted	Percent of Total
Dark	17	0.50%
Positive	175	57.80%
Negative	64	21.10%
Light	13	4.30%
Ambiguous	34	11.20%

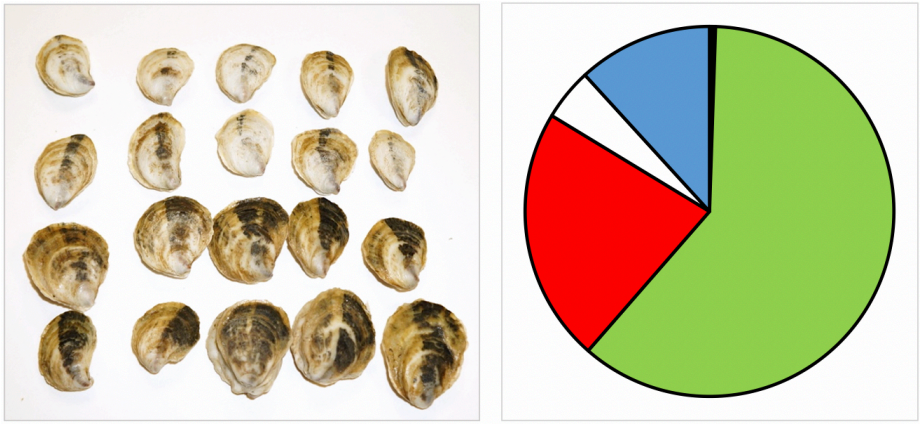


Figure 30. Left: Photo of representative shells from positive x negative cross (1 family). Right: Percentages of color morphs found in the progeny are represented as parts of pie chart. Key to colors - black = dark; white = light; green = positive stripe; red = negative stripe; blue = ambiguous

Positive x Light

3 families

Table 8. Number and percent of total counted of each of the 5 color morphs in the progeny of positive x light cross (3 pooled families

Progeny Color Morph	Number Counted	Percent of Total
Dark	21	7.00%
Positive	145	48.50%
Negative	50	16.70%
Light	83	27.80%
Ambiguous	0	0.00%

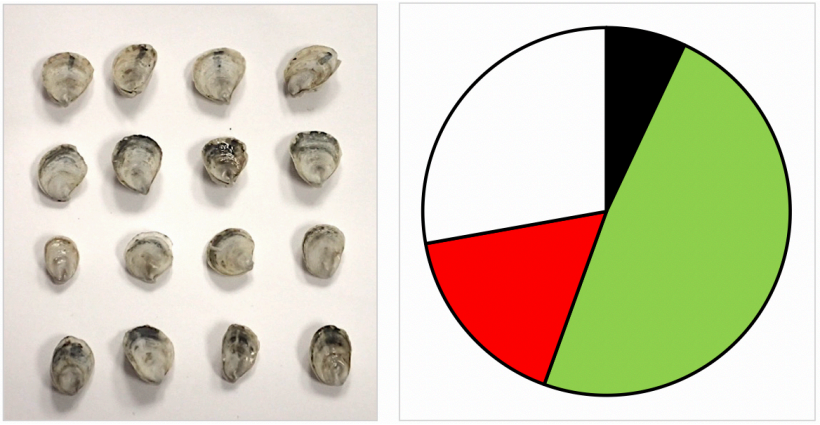


Figure 31. Left: Photo of representative shells from positive x light cross (3 pooled families). Right: Percentages of color morphs found in the progeny are represented as parts of pie chart. Key to colors - black = dark; white = light; green = positive stripe; red = negative stripe; blue = ambiguous

Positive x Light

1 family

Table 9. Number and percent of total counted of each of the 5 color morphs in the progeny of positive x light cross (1 family)

Progeny Color Morph	Number Counted	Percent of total
Dark	22	12.40%
Positive	32	18.10%
Negative	3	1.70%
Light	107	60.50%
Ambiguous	13	7.30%

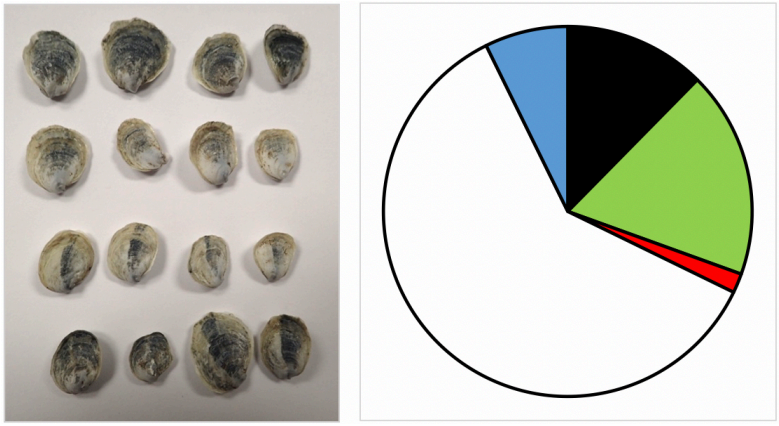


Figure 32. Left: Photo of representative shells from positive x light cross (1 family). Right: Percentages of color morphs found in the progeny are represented as parts of pie chart. Key to colors - black = dark; white = light; green = positive stripe; red = negative stripe; blue = ambiguous

Negative x Negative

3 families

Table 10. Number and percent of total counted of each of the 5 color morphs in the progeny of negative x negative cross (3 pooled families)

Progeny Color Morph	Number Counted	Percent of total
Dark	0	0%
Positive	0	0%
Negative	175	100%
Light	0	0%
Ambiguous	0	0%

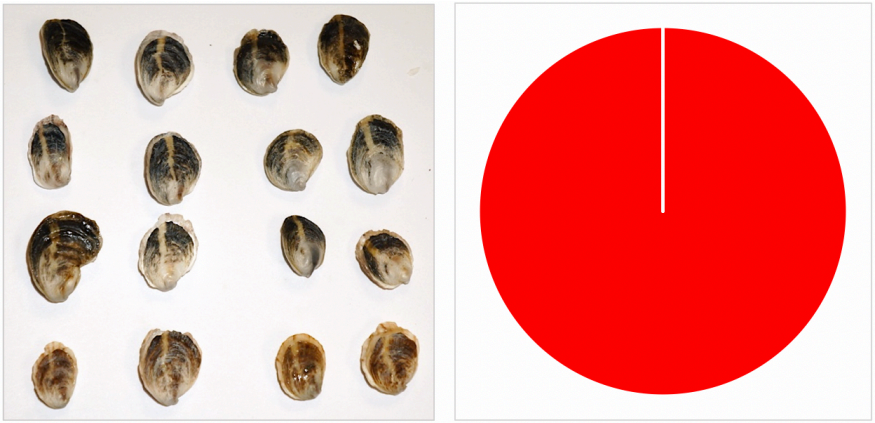


Figure 33. Left: Photo of representative shells from negative x negative cross (3 pooled families). Right: Percentages of color morphs found in the progeny are represented as parts of pie chart. Key to colors - black = dark; white = light; green = positive stripe; red = negative stripe; blue = ambiguous

Negative x Light

4 families

Table 11. Number and percent of total counted of each of the 5 color morphs in the progeny of negative x light cross (4 pooled families

Progeny Color Morph	Number Counted	Percent of total
Dark	5	4.50%
Positive	0	0.00%
Negative	78	70.90%
Light	27	24.50%
Ambiguous	0	0.00%

Figure 34. Left: Photo of representative shells from negative x light cross (4 pooled families). Right: Percentages of color morphs found in the progeny are represented as parts of pie chart. Key to colors - black = dark; white = light; green = positive stripe; red = negative stripe; blue = ambiguous

Negative x Light

1 family

Table 12. Number and percent of total counted of each of the 5 color morphs in the progeny of negative x light cross (1 family)

Progeny Color Morph	Number Counted	Percent of total counted
Dark	2	1.40%
Positive	0	0.00%
Negative	89	60.50%
Light	56	38.10%
Ambiguous	0	0.00%

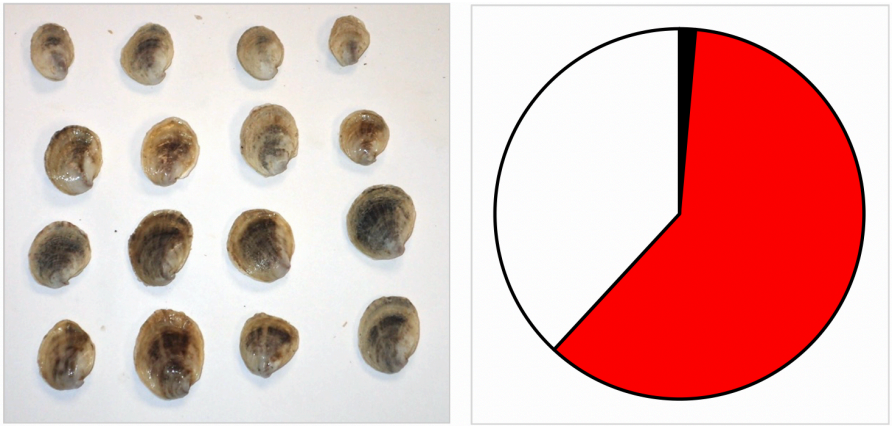


Figure 35. Left: Photo of representative shells from negative x light cross (1 family). Right: Percentages of color morphs found in the progeny are represented as parts of pie chart. Key to colors - black = dark; white = light; green = positive stripe; red = negative stripe; blue = ambiguous

Light x Light

1 family

Table 13. Number and percent of total counted of each of the 5 color morphs in the progeny of light x light cross (1 family)

Progeny Color Morph	Number Counted	Percent of total counted
Dark	0	0.00%
Positive	0	0.00%
Negative	5	7.20%
Light	65	92.80%
Ambiguous	0	0.00%

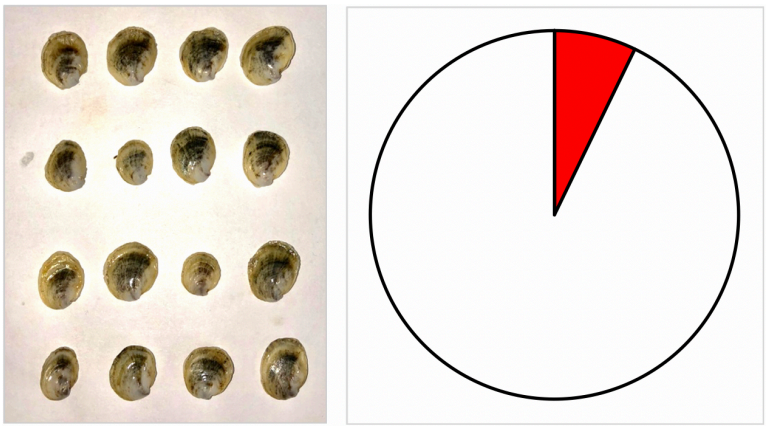


Figure 36. Left: Photo of representative shells from light x light cross (1 family). Right: Percentages of color morphs found in the progeny are represented as parts of pie chart. Key to colors - black = dark; white = light; green = positive stripe; red = negative stripe; blue = ambiguous

DISCUSSION

Consistent with the literature on the heritability of mantle edge and shell coloration in related species, such as *Crassostrea gigas*, color morphs in the eastern oyster, *Crassostrea virginica*, are under genetic control. The regular occurrence of the four morphs described above in wild and hatchery produced seed oysters as well as the apparent proportions of these discrete phenotypes in the progeny when crossed amongst each other indicates that the mode of inheritance is likely somewhat simple, meaning that pigmentation is probably controlled by a small number of genes. At this time, reported results are, at most, only suggestive of the manner in which the morphs are inherited due to the complicated nature of breeding studies and several limitations of the experimental design.

One of the major limitations of the present study was the number of crosses that were possible to culture in one spawning season. Twenty cultures were attempted to allow for each of the ten pairings to be cultured twice: one family in a 15 liter bucket and several families combined together in a 200 liter tank. Eight of the proposed twenty cultures either were not attempted due to the unavailability of appropriate ripe broodstock, did not survive the larval period, or became contaminated with animals from other groups. Ideally, all crosses would have been single family cultures and each pairing would have been replicated many times. In addition, the oysters used to conduct crosses had unknown genotypes. To obtain more robust results, they ideally would have been hybridized from true breeding lines.

Another limitation of this study was that the color morph of each individual was determined subjectively. Inherently, there were several occasions when it was difficult to categorically describe shell coloration. Oysters were assigned a color morph in all cases where possible and only labeled as ambiguous when color patterns deviated markedly from what was expected. When

examining an individual's color, oysters were described as being negatively or positively striped only if a well-defined stripe occurred approximately in the center of the right valve and nowhere else and extended continuously from the hinge to the bill. It is possible that very thin stripes on the center of shells that were not easily perceptible without the aid of instrumentation may not have accounted for.

The amount of pigmentation can be seen as being an additive trait, and the dark and light color morphs can therefore be described as existing on a continuum. Oysters with intermediate pigmentation were present, but were only described as being either dark or light. Many diffuse, lighter stripes were categorized as the light color morph, while oysters with many broad, darker stripes were described as the dark color morph.

A further notable complication is that on a small number of occasions it was observed that it is possible to have both dark and light stripes visible on the same shell. This is suggestive of the possibilities that dark and light stripes could be inherited independently (an oyster might have zero, one, or two colored central stripes) and/or that stripes present on a similarly colored background may go undetected.

Although more studies are required to elucidate the inheritance of color morphs in the eastern oyster, it is possible to draw some broad conclusions based on the results of crosses in order to provide direction and encouragement for future studies. With the limitations of this study in mind, and assuming minor cross contamination, the results of the twelve color morph crosses taken together suggest that

1. the amount of pigmentation is heritable
2. oysters without stripes (dark or light) crossed amongst each other produce oysters without stripes, suggesting that these color morphs are homozygotes in the parents.

3. the negative stripe color morph crossed with the negative stripe color morph produced only negative striped progeny, suggesting that the parents with negative stripe are homozygous.
4. the negative stripe color morph crossed with oysters without stripes produces either negative striped oysters or oysters without stripes, but not the positive stripe morph
5. positive stripe appears to be the most dominant color morph; each cross that involved the positive stripe color morph produced progeny of all four color morphs and also ambiguous individuals.

REFERENCES

Bandaranayake, W. M. The Nature and Role of Pigments in Marine Invertebrates. 2006. *Natural Product Reports*, 23: 223-255.

Brake, J., Evans, F., Langdon, C. 2004. Evidence for genetic control of pigmentation of shell and mantle edge in selected families of Pacific oysters, *Crassostrea gigas*. *Aquaculture*, 229:89–98.

Cook, L.M. 2017. Reflections On Molluscan Shell Polymorphisms. *Biological Journal of the Linnean Society*, 121: 717-730.

Coon, S. L., Bonard, D.B., and Weiner, R.M. 1986. Chemical production of cultchless spat using epinephrine and norepinephrine. *Aquaculture*, 58: 255-262.

Evans, S., Camara, M.D., Langdon, C.J. 2009. Heritability of shell pigmentation in the Pacific oyster, *Crassostrea gigas*. *Aquaculture*, 286: 211–216.

Ge, J., Li, Q., Yu, H., Kong, L. 2014. Identification and mapping of a SCAR marker linked to a locus involved in shell pigmentation of the Pacific oyster (*Crassostrea gigas*). *Aquaculture*, 434: 249–253.

Ge, J., Li, Q., Yu, H., Kong, L. 2015a. Identification of single-locus PCR based on markers linked to shell background color in the in the Pacific oyster *Crassostrea gigas*. *Marine Biotechnology*, 17: 655–662.

Ge J, Li Q, Yu H, Kong L. 2015b. Mendelian inheritance of golden shell color in the Pacific oyster *Crassostrea gigas*. *Aquaculture*, 441: 21–24.

Gunji, Y. 1990. Pigment Color Patterns of Molluscs as an Autonomous Process Generated by Asynchronous Automata. *BioSystems*, 23: 317-334.

Han, Z.Q., Li, Q., Liu, S.K., Yu, H., and Kong, L.F. 2019. Genetic variability of an orange-shell line of the Pacific oyster *Crassostrea gigas* during artificial selection inferred from microsatellites and mitochondrial COI sequences. *Aquaculture*, 508: 159–166.

Han, Z.Q. and Li, Q. 2020. Mendelian inheritance of orange shell color in the Pacific oyster *Crassostrea gigas*. *Aquaculture*, 516: 734616.

Hedgecock, D., Grupe, P., Voigt, M. 2006. Mapping genes affecting shell color and shape in the Pacific oyster, *Crassostrea gigas*. *Journal of Shellfish Research*, 25: 738.

Imai, T., Sakai, S. 1961. Study of breeding of Japanese oyster *Crassostrea gigas*. *Tohoku Journal of Agriculture Research*, 12: 125–171.

Kang, J.H, Kang, H.S, Lee J.M., An, C.M, Kim, S.Y., Lee, Y.M., Kim, J.J. 2013. Characterizations of shell and mantle edge pigmentation of a Pacific oyster, *Crassostrea gigas*, in Korean Peninsula. *Asian-Australasian Journal of Animal Science*, 26:1659–1664.

Loosanoff, V.L. and Davis, H.C. 1963. Rearing of Bivalve Mollusks. Advances in Marine Biology, 1: 1-136.

Meinhardt, M. 2009. The Algorithmic Beauty of Sea Shells, 4th Edition. Berlin, GER: Springer.

Nell, J.A. 2001. The history of oyster farming in Australia. *Marine Fisheries Review*, 63: 14–25.

Song, J., Li, Q., Yu, Y., Wan, S., Han, L. Du, S. 2018. Mapping Genetic Loci for Quantitative Traits of Golden Shell Color, Mineral Element Contents, and Growth-Related Traits in Pacific Oyster (*Crassostrea gigas*). *Marine Biotechnology*, 20: 666–675.

Xing, D., Li, Q., and Kong, L., and Yu, H. 2018. Heritability estimate for mantle

edge pigmentation and correlation with shell pigmentation in the white-shell strain of Pacific oyster, *Crassostrea gigas*. *Aquaculture*, 482: 73-77.

Xu, L., Li, Q., Yu, H., and Kong, L. 2017. Estimates of Heritability for Growth and Shell Color Traits and Their Genetic Correlations in the Black Shell Strain of Pacific Oyster *Crassostrea gigas*. Marine Biotechnology, 19: 421–429.

Wang, J., Li, Q., Zhong, X., Song, J., Kong, L., and Yu, H. 2018. An integrated genetic map based on EST-SNPs and QTL analysis of shell color traits in Pacific oyster *Crassostrea gigas*. *Aquaculture*, 492: 226-236.

Yu, W.C., He, C., Wu, C.L., Wang, J., Li, Z., Guo, T., Li, Y.C., Wang, X.T., 2015. Extraction and identification of melanin in shell and mantle of Paific oyster *Crassostrea gigas*. *Oceanologia et Limnologia Sinica*, 46: 909–914 (in Chinese).

CHAPTER 3: COMPARATIVE TRANSCRIPTOME ANALYSIS OF MANTLE EDGE PIGMENTATION

INTRODUCTION

Background

The chemical nature of molluscan pigments and varied molecular pathways involved in their synthesis remains largely unknown, however many of the most common classes of biological pigments are well studied, such as tetrapyrroles, melanins, and carotenoids (Comfort 1950; Williams 2017). This chapter aims to understand shell coloration in the eastern oyster utilizing ribonucleic acid (RNA) sequencing data to shed light on the molecular basis of pigment production. Messenger RNA carries the instructions from the DNA to ribosomes, which produce the proteins that enable the structure and function of living systems (Watson et al. 2014). The complete set of mRNAs (transcripts) of a cell type is called the transcriptome. RNA-seq, or whole-transcriptome shotgun sequencing, refers to the use of high-throughput sequencing technologies to characterize the RNA content of a biological sample in order to quantify broad-scale differences in gene expression patterns associated with a phenotype of interest. These differences in gene expression serve as a starting point for *a priori* biological inference (Wilhelm and Landry 2009).

Objective of Study

The objective of this study is to understand the molecular mechanisms of shell pigmentation in the eastern oyster by studying what genes are expressed in pigmented mantle tissue.

METHODS

Sample collection

In September 2017, 30 eastern oysters between one and three years of age, which exhibited pronounced dark stripes on their right valves, called ‘positive stripe color morph’ in the previous chapter, were selected from the tidal flats at Rutgers University’s experimental oyster farm at the Cape Shore Laboratory research site in Green Creek, New Jersey. Oysters were from both wild stock and established lines produced by Rutgers University’s selective breeding program and, consequently, had varied genetic backgrounds. They were harvested while the tide was receding to minimize time of exposure to air, and were immediately placed in an aerated bucket of seawater and transferred to the Haskin Shellfish Research Laboratory (HSRL) in Bivalve, New Jersey.

At the HSRL, all oysters were acclimated together in a single 200 liter raceway at 25°C in 24 parts per thousand (ppt) seawater and fed 30 mL (300,000 cells/mL) of Shellfish Diet 1800® commercial algae paste (Reed Mariculture, Inc.). After one hour, oysters were divided into three groups of 9 individuals. Individuals 1–9 made up Group A; 10–18 made up Group B; and 19–27 made up Group C. Right valves were removed and pigmented and nonpigmented mantle edge tissues were sampled and separated. Mantle edge tissues were used to examine shell pigmentation after it was observed that pigmentation patterns on shells correspond with patterns on mantle edges (Figure 37).

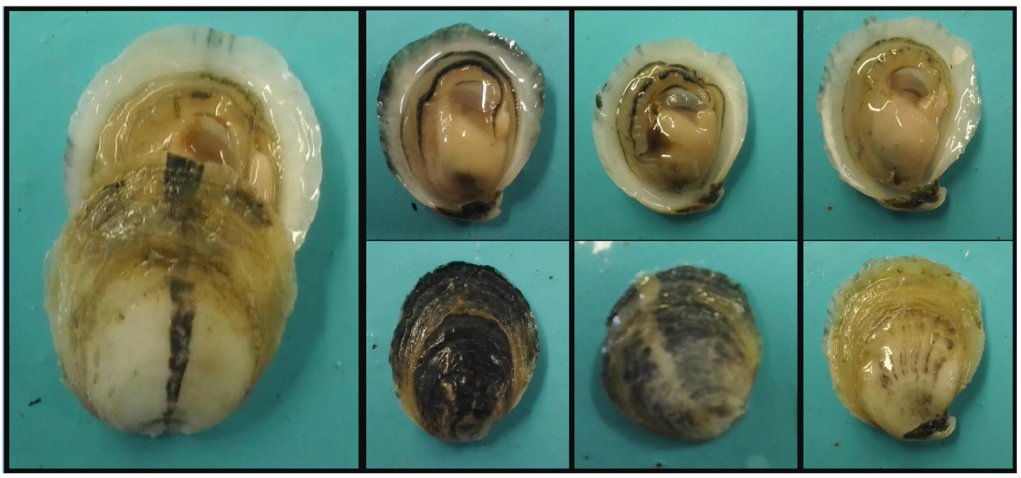


Figure 37. Pigmentation patterns on shells correspond with pigmentation atterns on mantle edges. From left to right, the positive stripe, full, negative stripe, and light color morphs are pictured with valves removed to show patterns on mantle tissues.

Pigmented and nonpigmented samples from the 9 oysters were pooled by group and placed in RNAlater Stabilization Solution (ThermoFisher) to replicate a pigmented ("B" = black) vs nonpigmented ("W" = white) paired comparison of mRNA levels three times (A: B1 vs. W1; B: B2 Vs. W2; C: B3 Vs. W3) (Figure 38).

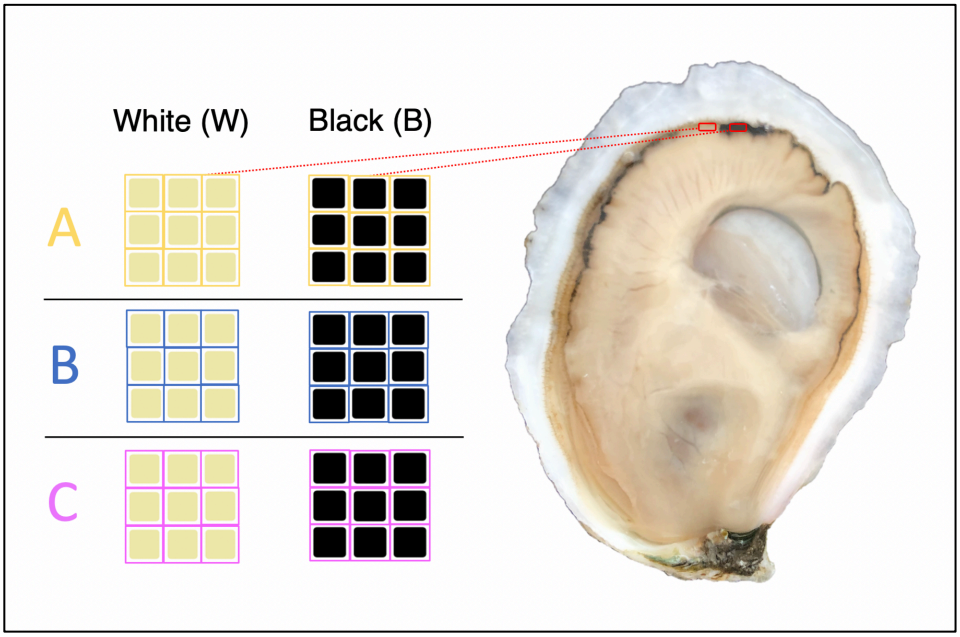


Figure 38. Sampling design. Biopsies of dark and light pigmented mantle edge tissue were pooled separately from three groups of nine oysters (27 oysters total) for paired comparisons as described in text.

Sample preparation, Illumina sequencing, library preparation, and initial bioinformatic analysis was carried out by Novogene Bioinformatics Technology Co., Ltd. Subsequent bioinformatic analysis was performed starting with raw read counts following Warnes et al. (2005); Gonzalez (2014); Love, Anders, and Huber (2014); Kopelainen et al. (2015); Wang and Sun (2018); and Dundar, Skrabanek, and Zumbo (2019).

RNA Extraction (Carried out by Novogene)

Total RNA was extracted using TRIzol™ Reagent (Invitrogen, Carlsbad, CA, USA). One milliliter of TRIzol™ Reagent per 50–100 mg of tissue was added to samples and the lysate was homogenized. Samples were then incubated for 5 minutes to permit complete dissociation of the nucleoprotein complex. Next, 0.2 mL of chloroform per 1 mL of TRIzol™ Reagent used for lysis was added and samples were incubated for 2–3 minutes. The samples were centrifuged for 15 minutes at 12,000 × g at 4°C to separate the mixture into a lower red phenol-chloroform, an interphase, and a colorless upper aqueous phase. The upper aqueous phase, which contained the RNA, was transferred to a new tube. One half milliliter of isopropanol per 1 mL of TRIzol™ Reagent used for lysis was added to the aqueous phase and incubated for 10 minutes. Samples were centrifuged for 10 minutes at 12,000 × g at 4°C and the supernatant was discarded. The precipitate pellet containing total RNA was resuspend in 1 mL of 75% ethanol per 1 mL of TRIzol™ Reagent used for lysis. The sample was vortexed briefly and then centrifuged for 5 minutes at 7500 × g at 4°C. The supernatant was discarded and RNA pellets were allowed to dry before resuspension in 20–50 μL of RNase-free water.

Samples were run on a 1% agarose gel to examine RNA degradation and contamination: 18S and 28S ribosomal RNA bands are clearly visible in intact total RNA samples, whereas degraded RNA appears as a lower molecular weight smear.

A NanoPhotometer® spectrophotometer (IMPLEN, Westlake Village, CA, USA) was used to examine RNA purity. An A260/A280 ratio of ~2 is considered pure, and the formula A260 × dilution × 40 yields the RNA concentration in μg RNA/mL.

RNA quantity was measured using the Qubit® RNA Assay Kit in a Qubit® 2.0 Flurometer (Life Technologies, CA, USA), and RNA integrity was assessed using an RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

Library Preparation and Sequencing (Carried out by Novogene)

Equal amounts of total RNA (3 ng) from each of the six pools were used for sample preparations. Messenger RNA (mRNA) was purified from total RNA using poly-T oligo-attached magnetic beads. Random fragmentation was carried out using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5X). First strand complementary DNA (cDNA) was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase (RNase H). Second strand cDNA synthesis was subsequently generated by nick-translation in a custom second strand synthesis buffer (Illumina) with dNTPs, *Escherichia coli* Polymerase I, and RNase H.

Double stranded cDNA was purified and remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. End repaired cDNA was purified on Solid-Phase Reversible Immobilization (SPRI) beads. Next, both 3′ ends of eluted end-repaired ds cDNA fragments were adenylated, and NEBNext library specific adaptors with hairpin loop structures were ligated to prepare for hybridization. In order to select cDNA fragments of preferentially

150~200 bp in length, the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, USA). Purified size-selected, adaptor-ligated cDNA was then treated with Uracil-Specific Excision Reagent (USER) Enzyme (NEB, USA) and incubated at 37°C for 15 min followed by 5 min at 95°C.

To enrich the library, a minimal number of cycles of PCR was performed using universal sequences from the adaptor as primers and Phusion High-Fidelity DNA polymerase. PCR products were then purified (AMPure XP system) and library quality was assessed by visualizing abundance and size distributions by capillary electrophoresis on an Agilent Bioanalyzer 2100 system. Libraries were normalized and pooled at equal volumes in preparation for cluster generation and sequencing.

Clustering of the index-coded samples was performed using the cBot Cluster Generation System using HiSeq PE Cluster Kit (Illumina). After cluster generation, sequence by synthesis of clusters was performed using the Illumina Hiseq 2000 platform to generate paired-end reads of 150 nucleotides in length (PE150).

Quality Control (Carried out by Novogene)

Raw data in FASTQ format were first processed through Novogene in-house Perl scripts. ASCII encoded quality scores were converted to and reported as Phred Q scores, which are related to the base calling error probabilities by

$$\mathbf{Q = -10 \log_{10} P}$$

where P is the probability that the base call is incorrect.

Clean reads were obtained by removing reads containing adapter sequences as well as uncertain (where N > 10%), and low-quality nucleotides (where base quality < 20 constitute more than 50% of read) from raw data.

Next, GC content was quantified in all samples. After reverse transcription and size selection, PCR amplification does not always uniformly amplify all fragments, and differences in GC content should be corrected for, especially when comparing the expression levels of different genes within samples. GC content distribution is also evaluated to detect potential AT/GC separation, which affects subsequent gene expression quantification.

Aligning and Counting Reads (Carried out by Novogene)

Reference genome and gene model annotation files were downloaded and an index of the reference genome was built using Bowtie2 (v2.2.3). Paired-end clean reads were aligned to the eastern oyster reference genome using TopHat2 (v2.0.12). HTSeq (v0.6.1) was used to count the number of reads mapped to each gene.

Normalization (Carried out by Novogene)

FPKM, the expected number of Fragments Per Kilobase of transcript sequence per Million base pairs sequenced, was used to report abundance, and differential expression analysis between the two conditions (dark and light mantle tissue) was performed using the DESeq2 R package (Love, Huber, and Anders 2014).

FPKM normalization accounts for sequencing depth by dividing the number of fragments mapped to each gene by the number (in millions) of total reads per sample and then normalizes for gene length by dividing by the number of kilobases in each gene. Because of the possibility of either one or two reads originating from a single fragment when doing paired-end RNA-seq,

FPKM also keeps track of fragments so that they are not counted twice. Transcript length is more important to consider when comparing the expression levels of genes within one sample.

The DESeq2 pipeline uses a relative log expression (RLE) method for normalizing count data. It assumes read counts are proportional to expression levels and to sequencing depth and that most genes are not differentially expressed. It then adjusts for differences in both sequencing depth and composition of the reads in each library by calculating and dividing by a scaling factor for each sample.

Mean-Variance Relationship

RNA-seq count data follow a negative binomial distribution, where the variance is not equal to the mean. Instead, as expression level increases, the data become over-dispersed. DESeq2 models gene-wise logistic regressions assuming that the read counts are distributed in this way. To verify and visualize this relationship, a plot of the mean count data vs. the variance was produced.

The dispersion of a given gene's expression between replicates was estimated by squaring gene-wise biological coefficients of variation (BCV), and within group variation was then adjusted by assuming genes that are expressed at similar levels also have similar dispersion.

Identifying Differentially Expressed Genes (Carried out by Novogene and also by the author beginning with raw read counts)

Dispersion estimates were used to model probabilities of the number of sequenced reads to be assigned to a gene using the negative binomial distribution. Statistical testing for differential expression of each gene between the two conditions (B vs. W) as well as pairwise comparisons of groups A–C (B1 vs. W1; B2 vs. W2; B3 vs. W3) based on how well the actual read counts for

each gene fit the model were performed using Wald tests. Resulting *p*-values were adjusted using the Benjamini and Hochberg's (BH) approach (Benjamini and Hochberg1995) for controlling the false discovery rate. Genes with a BH adjusted *p*-value < 0.05 and $\log_2$(fold change) ± 1 were considered differentially expressed (Anders and Huber 2010).

GO and KEGG Enrichment Analysis (Carried out by Novogene)

Gene Ontology (GO) enrichment analysis of differentially expressed genes was implemented by the GOseq R package. GO terms with corrected *p*-values < 0.05 were considered significantly differentially expressed. The Kyoto Encyclopedia of Genes and Genomes (KEGG) was used to examine higher level functions of biological systems using the molecular level information, and the KOBAS (2.0) software was used to detect the statistical enrichment of differentially expressed genes (DEGs) in KEGG pathways. The Cufflinks v2.1.1 Reference Annotation Based Transcript (RABT) assembly method was used to construct and identify known as well as predict novel transcripts from TopHat alignment results.

RESULTS

Read Counts, Coverage, and Mapping

Illumina paired-end sequencing generated an average of 41,828,451 clean reads per sample from an average of 44,211,878 raw reads. An average of 25,989,069 (62.22%) total reads per sample were mapped to the *C. virginica* reference genome. An average of 3,066,105.7 raw reads (7.34%) were mapped to multiple locations in the genome, while 22,922,963.5 (54.88%) were uniquely mapped to one site (Table 14). Reads were mapped to exons (average = 74.9%), introns (average = 9.2%), or to intergenic regions (average = 15.9%) (Figure 39).

Table 14. Output from Illumina sequencing and summary of reads mapped to reference genome

Sample	B1	B2	B3	W1	W2	W3
Raw reads	46010622	41726602	44013822	46432924	46847546	40239752
Clean reads	44292598	35220380	42594902	45014800	45287446	38560578
Clean bases	6.6 G	5.3 G	6.4 G	6.8 G	6.8 G	5.8 G
Total mapped	26271222 (59.31%)	22827062 (64.81%)	27112420 (63.65%)	28069769 (62.36%)	27908926 (61.63%)	23745016 (61.58%)
Multiple mapped	3309329 (7.47%)	2781337 (7.9%)	3038819 (7.13%)	3232016 (7.18%)	3312049 (7.31%)	2723084 (7.06%)
Uniquely mapped	22961893 (51.84%)	20045725 (56.92%)	24073601 (56.52%)	24837753 (55.18%)	24596877 (54.31%)	21021932 (54.52%)
Read-1	11906615 (26.88%)	10251586 (29.11%)	12473941 (29.29%)	12839350 (28.52%)	12751112 (28.16%)	10843452 (28.12%)
Read-2	11055278 (24.96%)	9794139 (27.81%)	11599660 (27.23%)	11998403 (26.65%)	11845765 (26.16%)	10178480 (26.4%)
Reads map to +	11382887 (25.7%)	9969603 (28.31%)	11942060 (28.04%)	12328700 (27.39%)	12198868 (26.94%)	10436790 (27.07%)
Reads map to -	11579006 (26.14%)	10076122 (28.61%)	12131541 (28.48%)	12509053 (27.79%)	12398009 (27.38%)	10585142 (27.45%)

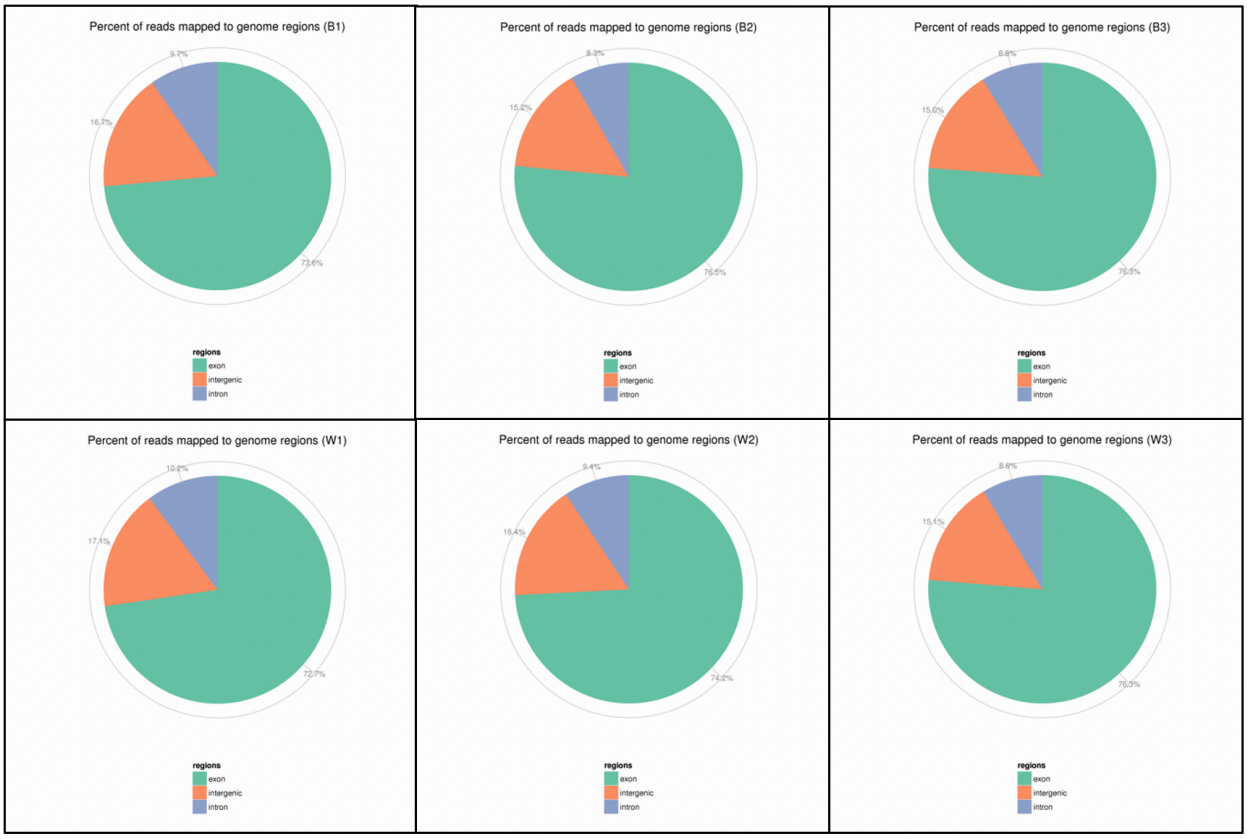


Figure 39. Percent of reads mapped to exons, intergenic regions, and introns in each of the six libraries

Error Rate and GC Content

The average GC content across all samples was 45.40%. The average error rate for a given base call across all samples was 0.0183. On average, 97.03% of bases had a Q score of at least 20 (Q20), which indicates an incorrect base call probability of 1 in 100. An average of 92.71% of bases had a Q score of 30 (Q30), which is equivalent to the probability of an incorrect base call 1 in 1000 times (base call accuracy is 99.9%) (Table 15).

Table 15. Error rate of sequencing per nucleotide and percent of guanine-cytosine content in each of the six libraries

Sample	Error rate (%)	Q20 (%)	Q30 (%)	GC content (%)
B1	0.02	96.58	91.71	45.49
B2	0.01	97.9	94.7	45.8
B3	0.02	96.94	92.51	45.48
W1	0.02	97.03	92.69	45.14
W2	0.02	96.89	92.46	45.06
W3	0.02	96.84	92.23	45.42

Normalization

The abundance of all genes (40,007) in fragments per kilobase per million (FPKM) intervals for each library is reported in Table 16 and values are displayed as a density plot in Figure 40. An average of 57.19% of genes were expressed at very low levels (FPKM interval 0–1), while an average of 2.64% were highly abundant in all libraries (FPKM > 60).

Table 16. Number and percent of transcripts in fragments per kilobase per million intervals for each of the six libraries

FPKM Interval	B1	B2	B3	W1	W2	W3
0-1	23156 (57.88%)	23702 (59.24%)	22702 (56.75%)	22627 (56.56%)	22688 (56.71%)	22391 (55.97%)
1-3	6103 (15.25%)	6019 (15.04%)	6294 (15.73%)	6073 (15.18%)	6191 (15.47%)	6357 (15.89%)
3-15	7272 (18.18%)	7006 (17.51%)	7441 (18.60%)	7653 (19.13%)	7576 (18.94%)	7685 (19.21%)
15-60	2403 (6.01%)	2266 (5.66%)	2515 (6.29%)	2590 (6.47%)	2491 (6.23%)	2508 (6.27%)
>60	1073 (2.68%)	1014 (2.53%)	1055 (2.64%)	1064 (2.66%)	1061 (2.65%)	1066 (2.66%)

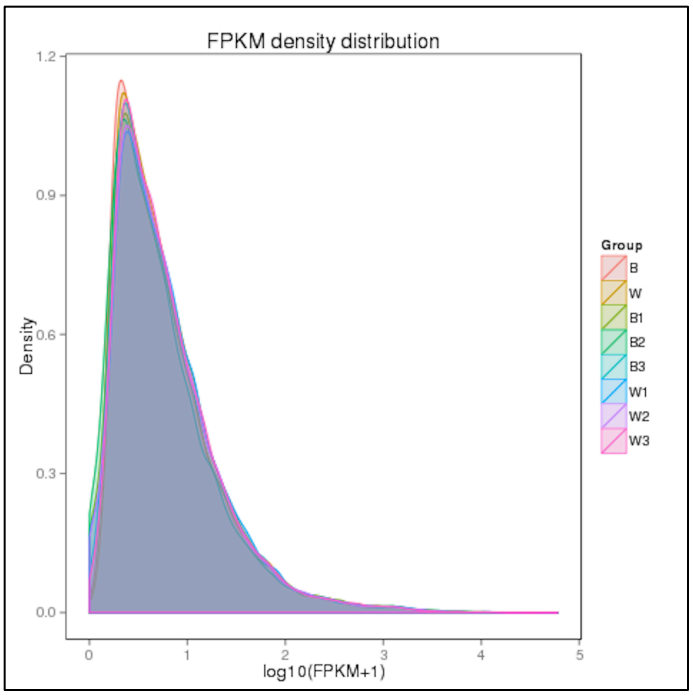


Figure 40. Density plot of FPKM values for each of the six libraries

The means and medians of count data before and after normalization as well as the size scaling factor for normalization of raw counts following the DESeq2 method for each of the four libraries are reported in Table 17.

Table 17. Means and medians of read counts per gene before (raw) and after (normalized) DESeq normalization by size scaling factors

Sample	B1	B2	B3	W1	W2	W3
Raw Count Mean	323.8	324.9	275.7	304.4	256.5	315.9
Raw Count Median	13	13	13	11	9	13
DESeq Normalized Mean	277.4	286.8	288.1	304.3	331.7	296.9
DESeq Normalized Median	11.1	11.5	13.6	11	11.6	12.2
Size Scaling Factor	1.001	0.773	1.064	1.167	1.327	0.957

Diagnostic Plots of Normalized Counts

Histograms of normalized pseudocounts ($\log_2$(count+1)) show that gene expression is similarly bimodally distributed across all samples (Figure 41), and box and density plots illustrate the effects of DESeq normalization (Figure 42).

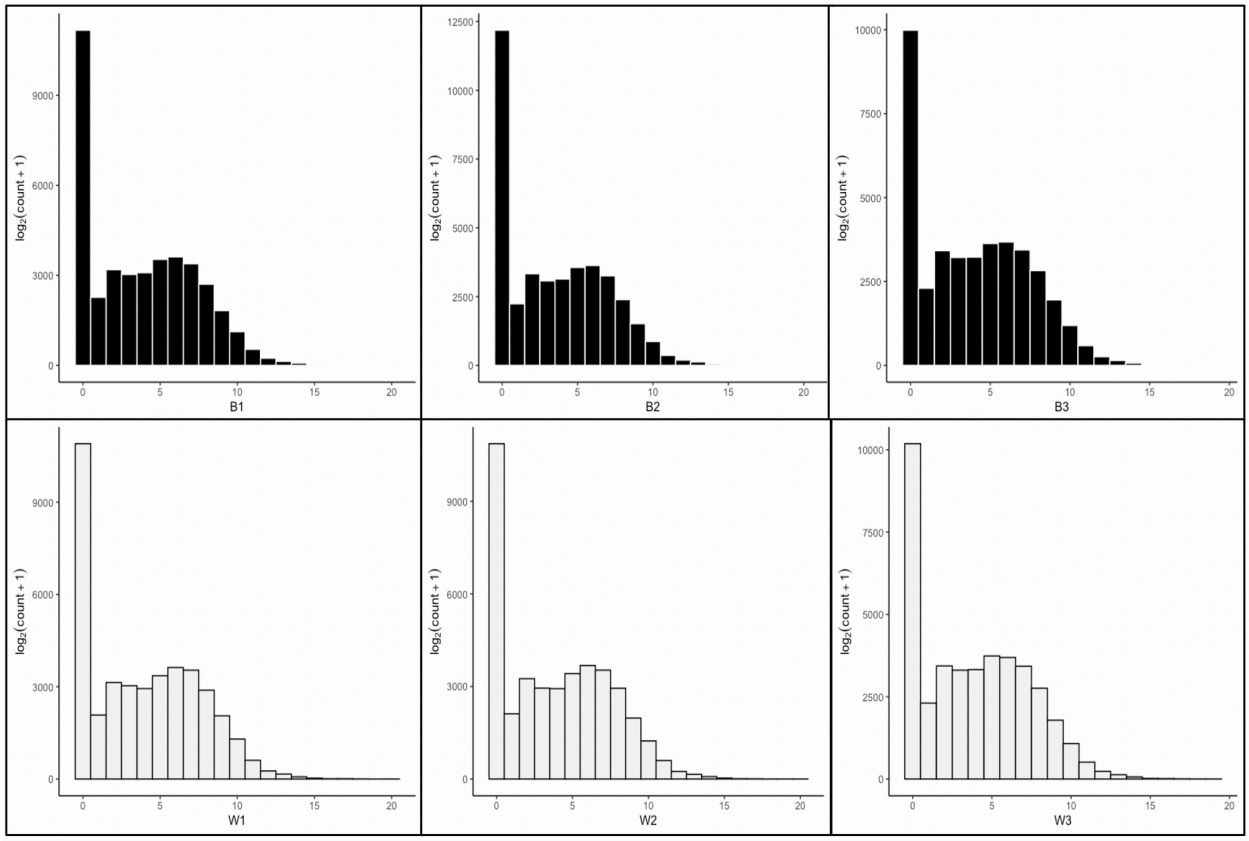


Figure 41. Histograms of normalized pseudocounts for each of the six libraries

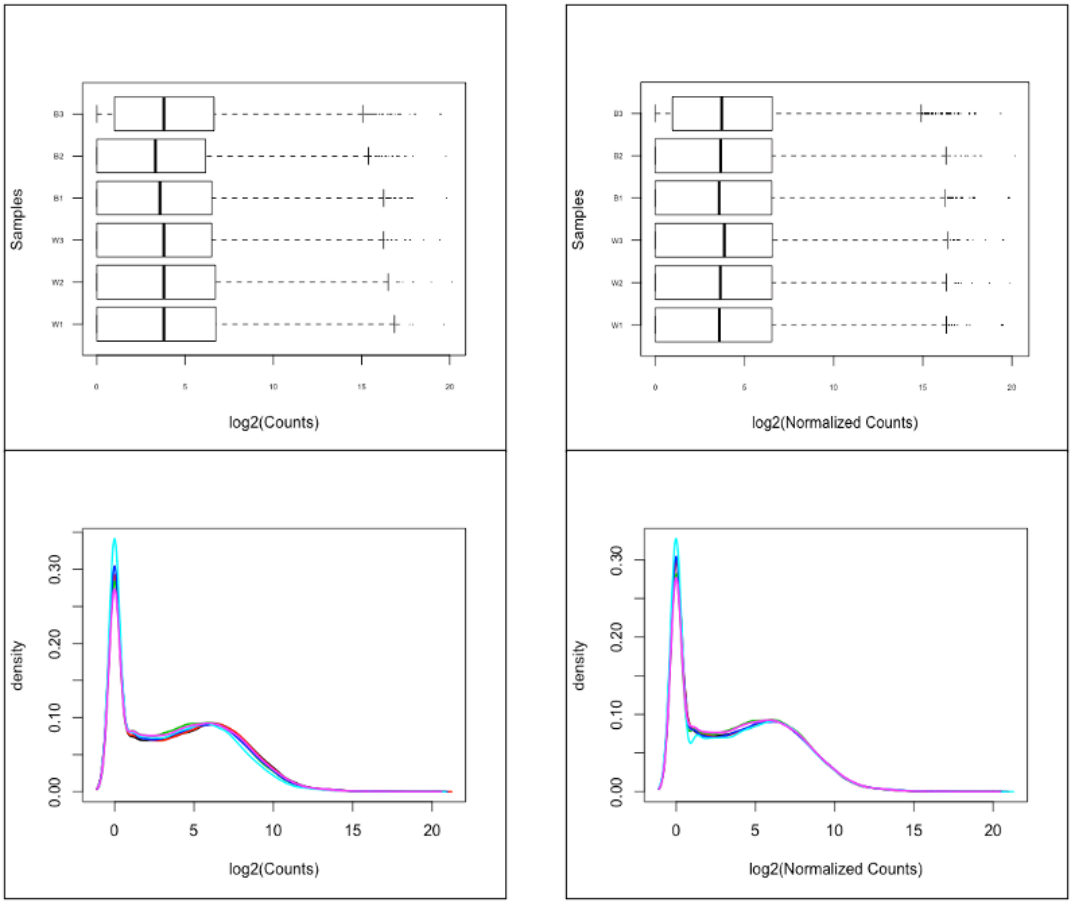


Figure 42. Box and density plots of raw and normalized read counts show the effect of normalization

An MA scatter plot visualizes M, the mean log-fold change per gene between the two tissue types (pooled samples of black vs. white), against A, the mean of normalized counts for each gene. In this study, M vs. A showed that normalization was properly executed by DESeq (Figure 43). The majority of points are centered around a log ratio of 0, which is expected as most genes will not be found to be differentially expressed, and there is higher variability of log ratios of genes with lower counts.

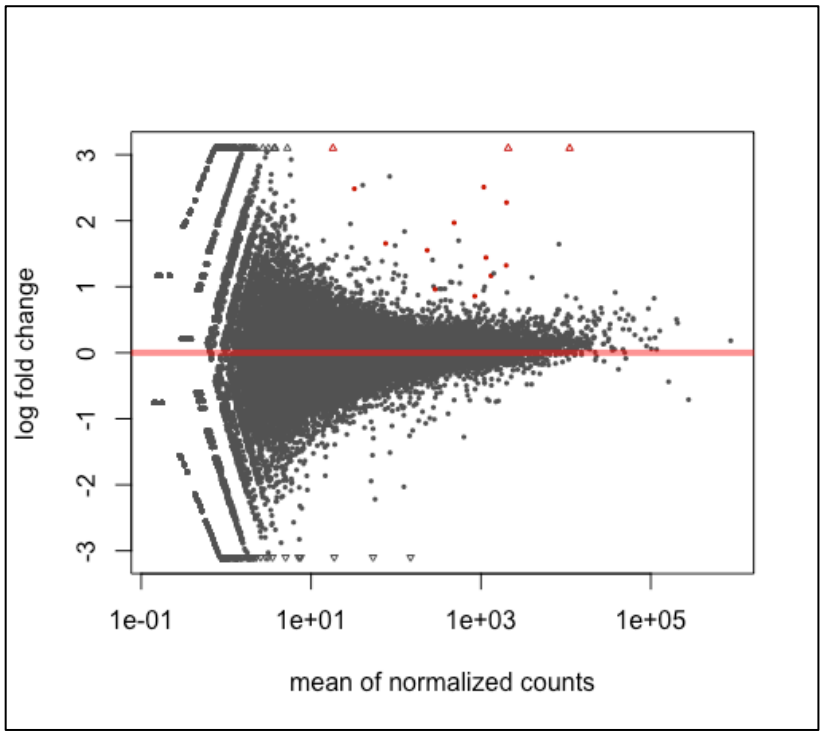


Figure 43. MA plot shows the mean log-fold change per gene (M) against the mean of normalized counts for each transcript (A)

The Pearson correlation coefficient (R), defined as the covariance of two variables divided by the product of their standard deviation, was used to assess the strength of linear relationships between all pairwise combinations of RNA library expression profiles. Coefficients of determination, or the square of the Pearson correlation coefficients (R^2), between each of the

expression profiles are reported below in Figure 44. Here, all pairwise correlation coefficients (as well as coefficients of determination) between the six libraries were above 0.93, indicating gene expression was extremely similar in all samples.

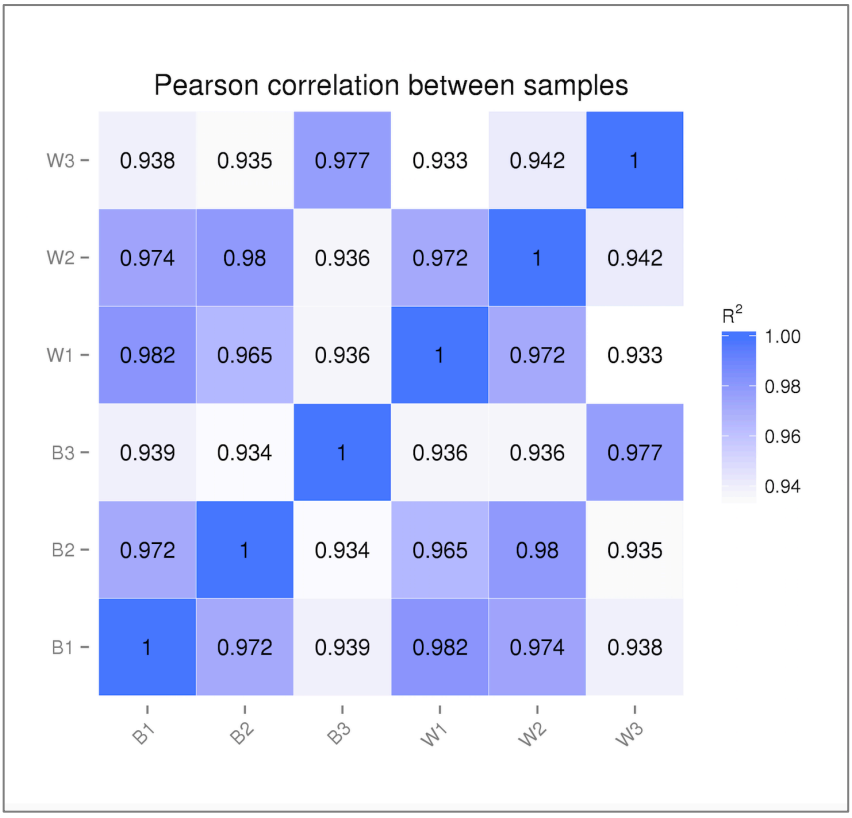


Figure 44. Pairwise Pearson correlation coefficients (R^2) between each of the expression profiles

Principal component analysis (PCA) was used to visualize the variation in the dataset and to determine whether outliers were present. PCA plots of RNA sequence data reduce the read count dataset to lower dimensions for analysis based on patterns of expression of the most significantly dysregulated genes. Samples are plotted along these dimensions (components) and cluster according to their similarity. In this study, a PCA plot showed that the six libraries clustered by replicate pair rather than by tissue pigmentation type sampled (Figure 45).

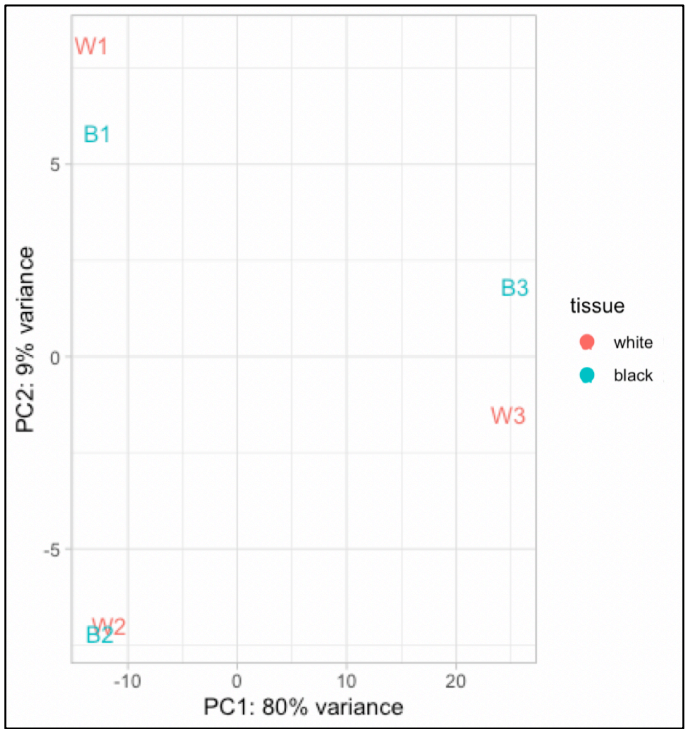


Figure 45. Principal component analysis plot of the six libraries shows that they cluster by replicate pair

The relationship between the mean and variance of normalized counts confirmed the data followed a negative binomial distribution (Figure 46).

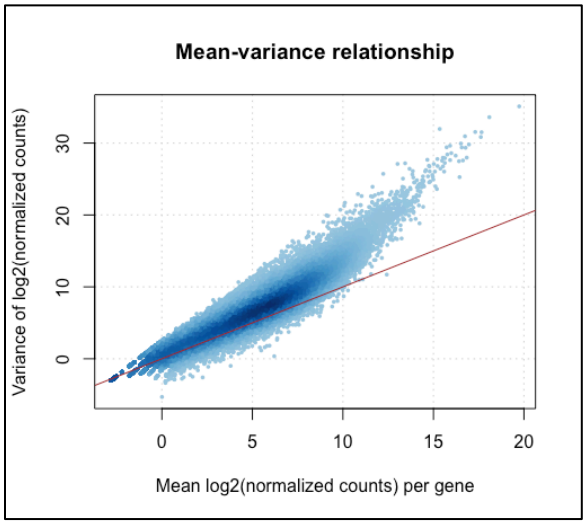


Figure 46. Plot of the variance against the mean of normalized read counts shows the data follow a negative binomial distribution. Reference line with slope of 1 is shown in red.

Differential Expression Analysis

Together, transcripts in the six libraries mapped to a total of 17,897 genes. Of these, 16,701 were coexpressed in both tissue types, while 909 genes were uniquely expressed in light mantle tissue and 287 genes were uniquely expressed in dark mantle tissue (Figure 47).

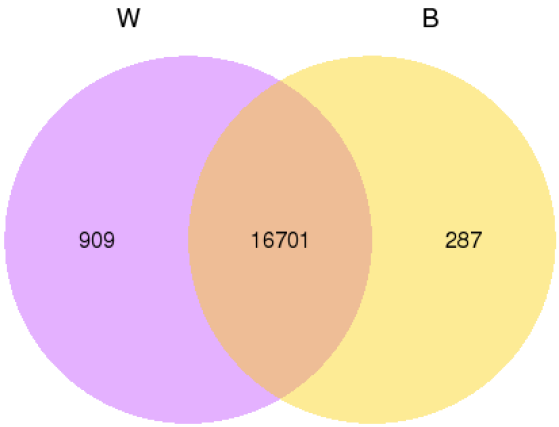


Figure 47. Venn diagram of the number of transcripts uniquely or co-expressed in lightly pigmented (W) and darkly pigmented (B) mantle tissue

After adjustment following the BH approach, the number of genes found to be differentially expressed (p-value < 0.05) across pooled replicates was 12. Genes that were not differentially expressed had p-values spread somewhat uniformly from zero to one. The majority of genes were expressed at similar levels in both tissue types or had very low read counts and therefore had high adjusted p-values (Figure 48).

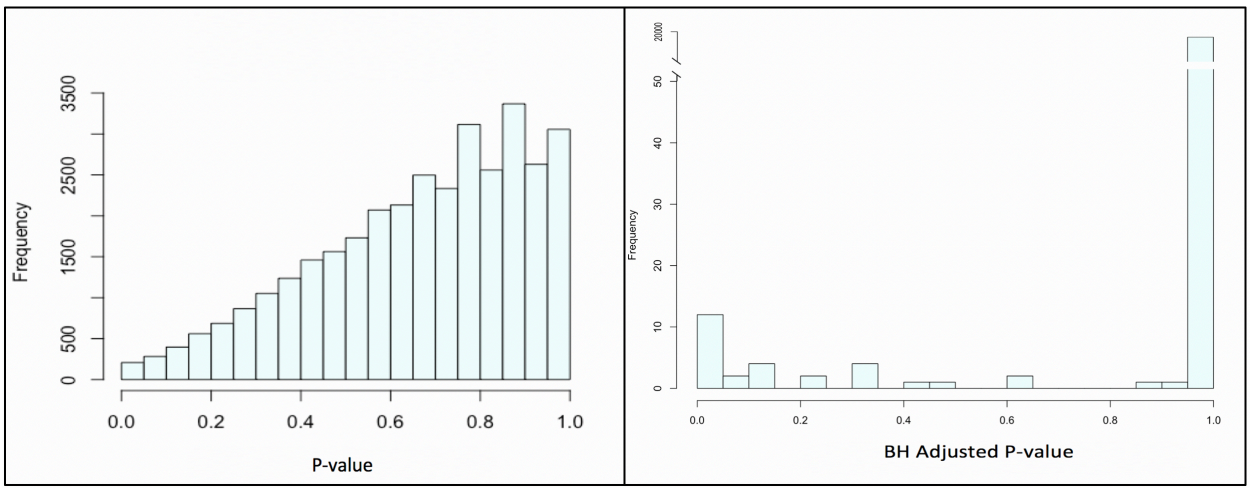


Figure 48. Histograms of the frequencies of p-values before (left) and after (right) Benjamini and Hochberg adjustment

All 12 transcripts found to be significantly differentially expressed across the three replicates were up-regulated in dark mantle tissue relative to white mantle tissue (Figure 49; Table 18).

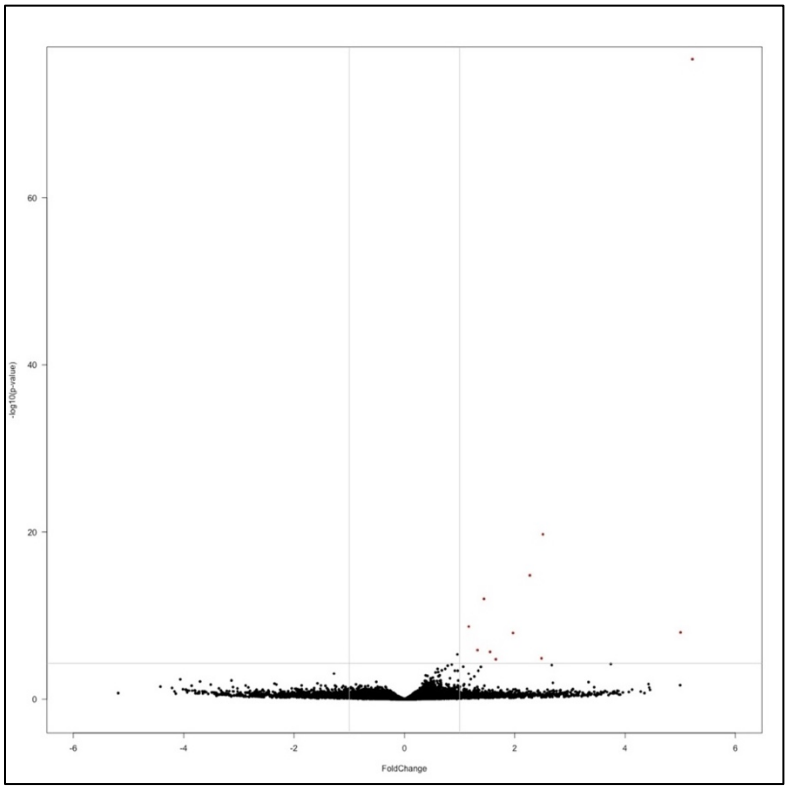


Figure 49. Volcano plot of pooled replicate comparison shows transcripts with either nonsignificant or significant adjusted p-value (<0.05) and fold change (+/- 1) in black and red, respectivel

Table 18. Summary data for the twelve transcripts found to be differentially expressed in the pooled replicate comparison

	Gene ID	Annotation	Log_2 (Fold Change)	B Average Normalized Read Count	W Average Normalized Read Count	Adjusted *p*-value (BH)
1	111131701	uncharacterized	5.22	21576.17	578.44	5.00E-73
2	111131416	uncharacterized	5.00	4055.53	126.26	3.27E-05
3	111116164	porphobilinogen deaminase-like	2.51	1837.47	322.39	1.79E-16
4	111105521	rho GTPase-activating protein gacII-like	2.48	54.76	9.86	2.27E-02
5	111110787	L-ascorbate oxidase-like	2.28	3318.18	685.26	9.30E-12
6	111119804	kynurenine--oxoglutarate transaminase 3-like	1.97	770.71	196.64	3.28E-05
7	111125957	ATP-binding cassette sub-family B member 6, mitochondrial-like	1.66	114.66	36.19	2.70E-02
8	111108232	5-aminolevulinate synthase, erythroid-specific, mitochondrial-like	1.55	347.35	118.58	4.60E-03
9	111108307	5-aminolevulinate synthase, nonspecific, mitochondrial-like	1.44	1676.50	616.96	4.67E-09
10	111126053	delta-aminolevulinic acid dehydratase-like	1.32	2835.26	1132.58	3.13E-03
11	111132904	multidrug resistance-associated protein 1-like	1.17	1815.76	809.56	7.80E-06
12	*111136004*	*uncharacterized*	*0.96*	*382.22*	*196.31*	*8.39E-03*

Replicate level comparisons yielded 31 (replicate A : B1 vs. W1), 27 (replicate B : B2 vs. W2), and 59 (replicate C : B3 vs. W3) DEGs.

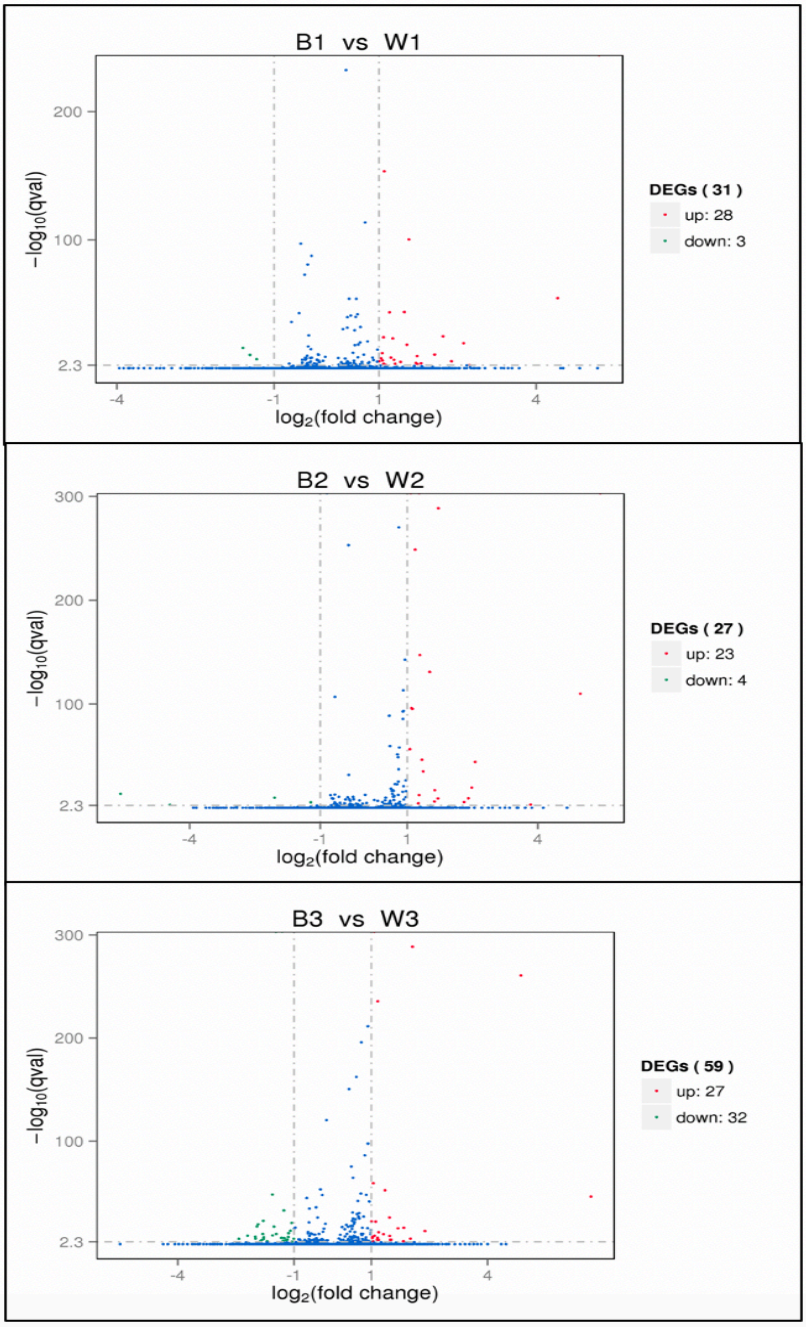


Figure 50. Volcano plots of each individual replicate comparison show transcripts with nonsignificant differential expression (blue), and up (red) and down (green) significant differential expression

Comparisons of DEGs found in each of the replicates revealed that 1, 1, and 4 transcripts were found to be differentially expressed in common between replicates A and B, A and C, and B and C, respectively. Seven transcripts were found to be differentially expressed in all three pairwise replicate comparisons (Figure 51; Table 19). After manual examination of sample read count data, 1 of 1 potential DEGs found in replicates A and B: 111126053; 0 of 1 potential DEGs found in replicates A and C; and 4 of 4 potential DEGs found in replicates B and C: 111109602, 111111270, 111130779, and 111132904 were considered to be of interest. One gene identified by DESeq (111138165) was rejected as a DEG because of overall low expression.

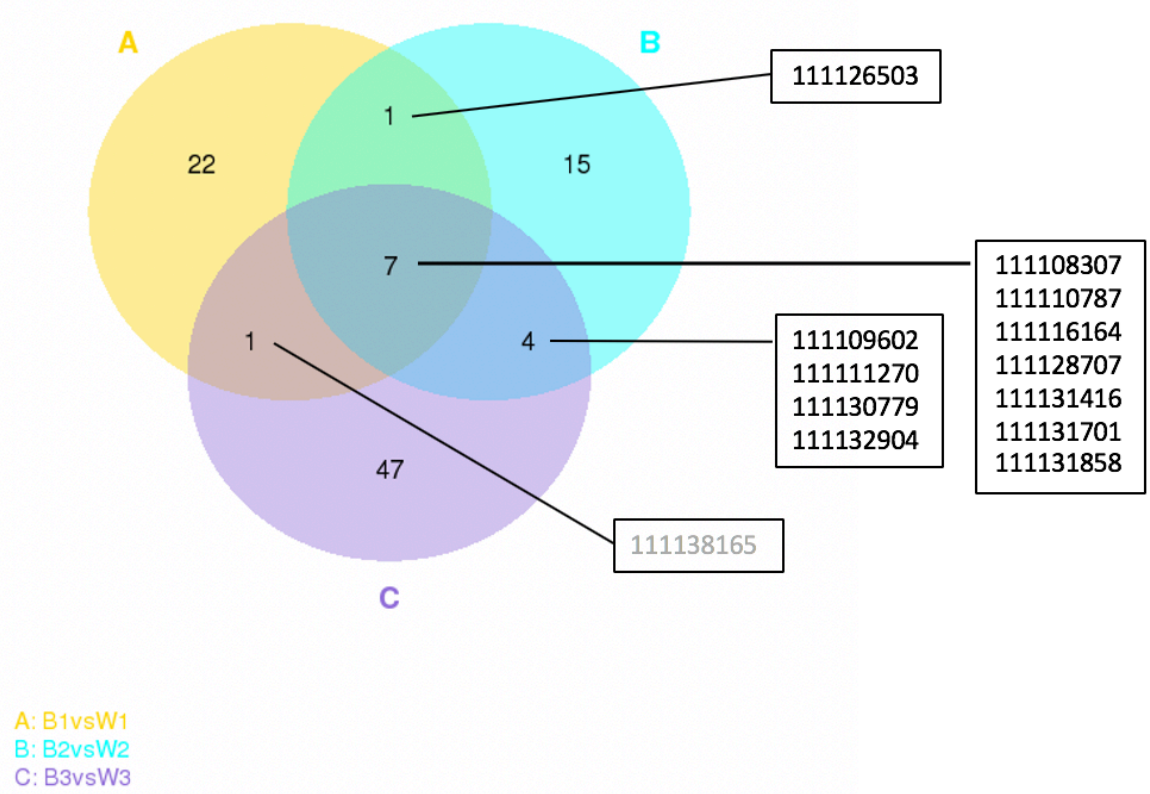


Figure 51. Venn diagram displaying the number of significantly expressed genes found to be uniquely enriched in or shared in common between each of the three individual replicate comparisons

Table 19. Summary data for the twelve transcripts found to be differentially expressed in at least two of the three individual replicate comparisons

	Gene ID	Description	Replicates Enriched	Log_2 (Fold Change)	B Normalized Read Count	W Normalized Read Count	Adjusted *p*-value (BH)
1	111108307	5-aminolevulinate synthase, nonspecific, mitochondrial-like	A (B1 v. W1)	1.41	1452.27	543.08	1.28E-05
			B (B2 v. W2)	1.71	2023.82	617.13	9.34E-10
			C (B3 v. W3)	1.15	1553.40	690.67	4.17E-05
2	111110787	L-ascorbate oxidase-like	A (B1 v. W1)	2.22	2930.52	624.46	1.40E-25
			B (B2 v. W2)	2.56	4739.48	798.12	6.41E-45
			C (B3 v. W3)	1.84	2284.53	633.20	1.23E-16
3	111116164	porphobilinogen deaminase-like	A (B1 v. W1)	2.62	1903.04	309.23	3.23E-20
			B (B2 v. W2)	2.48	2282.45	406.12	3.50E-20
			C (B3 v. W3)	2.38	1326.92	251.82	2.04E-13
4	111128707	shematrin-like protein 2	A (B1 v. W1)	1.01	2919.53	1446.78	3.80E-06
			B (B2 v. W2)	1.63	1586.72	511.18	8.33E-07
			C (B3 v. W3)	1.19	1410.56	612.31	7.53E-05
5	111131416	uncharacterized LOC111131416	A (B1 v. W1)	4.41	3182.39	149.05	2.34E-55
			B (B2 v. W2)	4.98	6509.83	205.71	1.18E-110
			C (B3 v. W3)	6.67	2474.36	24.03	6.91E-47
6	111131701	uncharacterized LOC111131701	A (B1 v. W1)	5.20	23480.14	636.45	0.00E+00
			B (B2 v. W2)	5.43	27576.92	633.90	0.00E+00
			C (B3 v. W3)	4.86	13671.46	464.98	2.00E-261
7	111131858	glycine-rich cell wall structural protein-like	A (B1 v. W1)	1.09	9167.37	4294.95	7.08E-25
			B (B2 v. W2)	1.27	4476.97	1845.20	6.37E-13
			C (B3 v. W3)	1.08	2854.96	1336.42	7.99E-09
8	111126053	delta-aminolevulinic acid dehydratase-like	A (B1 v. W1)	1.23	2858.56	1215.50	4.53E-09
			B (B2 v. W2)	1.63	3790.29	1215.71	1.38E-17
9	111109602	elastin-like	B (B2 v. W2)	1.27	189523.57	78139.40	0.00E+00
			C (B3 v. W3)	1.07	127703.32	60230.68	0.00E+00
10	111111270	uncharacterized LOC111111270	B (B2 v. W2)	-1.22	532.79	1230.72	5.81E-06
			C (B3 v. W3)	-2.00	240.58	949.81	6.31E-08
11	111130779	glycine-rich cell wall structural protein 1.0-like	B (B2 v. W2)	1.29	46511.56	18924.35	7.44E-148
			C (B3 v. W3)	1.17	61742.40	27221.51	2.39E-236
12	111132904	multidrug resistance associated protein 1-like (ABC a1)	B (B2 v. W2)	1.25	1955.28	814.89	5.23E-05
			C (B3 v. W3)	1.16	2015.76	893.38	7.27E-07

In summary, only five transcripts were found to be differentially expressed in common in both the pooled replicate (B1, B2, B3 vs. W1, W2, W3) and across all three replicate-level pairwise comparisons (B1 vs. W1, B2 vs. W2, B3 vs. W3): 111131701, 111131416, 111116164, 111110787, and 111108307. Three transcripts were found to be differentially expressed in both pooled replicate and at least one of the three replicate level pairwise comparisons: 111119804, 111126053, and 111132904. Three transcripts were found to be uniquely differentially expressed in the pooled replicate comparison only: 111105521, 111125957, and 111108232, while two transcripts were found to be uniquely differentially expressed in all three replicate level pairwise comparisons, but not pooled replicates: 111128707 and 111131858. Three transcripts were found to be differentially expressed in two of three replicate level pairwise comparisons, but not in the pooled replicate comparison: 111109692, 111130779, and 111111270. These sixteen transcripts are likely the most biologically relevant to pigment production in *C. virginica* (Figures 52 and 53).

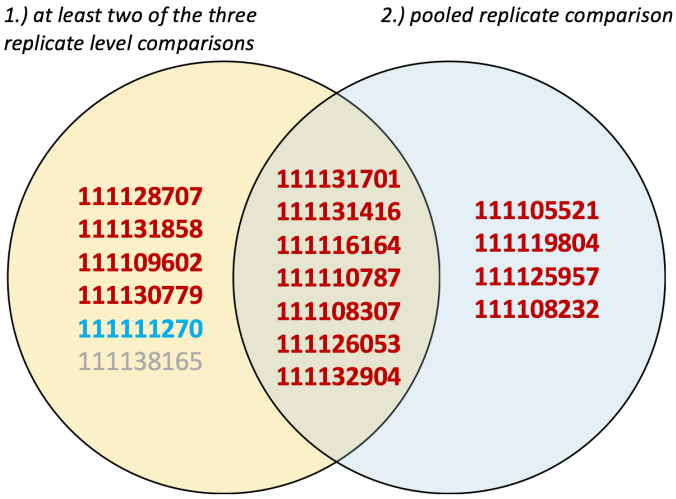


Figure 52. Venn diagram displaying transcripts found to be differentially expressed in either or in common between the pooled replicate and at least two of the three individual replicate level pairwise comparisons. Red text indicates the gene was upregulated in pigmented tissue and blue text indicates the gene was downregulated in pigmented tissue. Grey text indicates the gene had low expression in both pigmented and non-pigmented tissue and was rejected as a DEG.

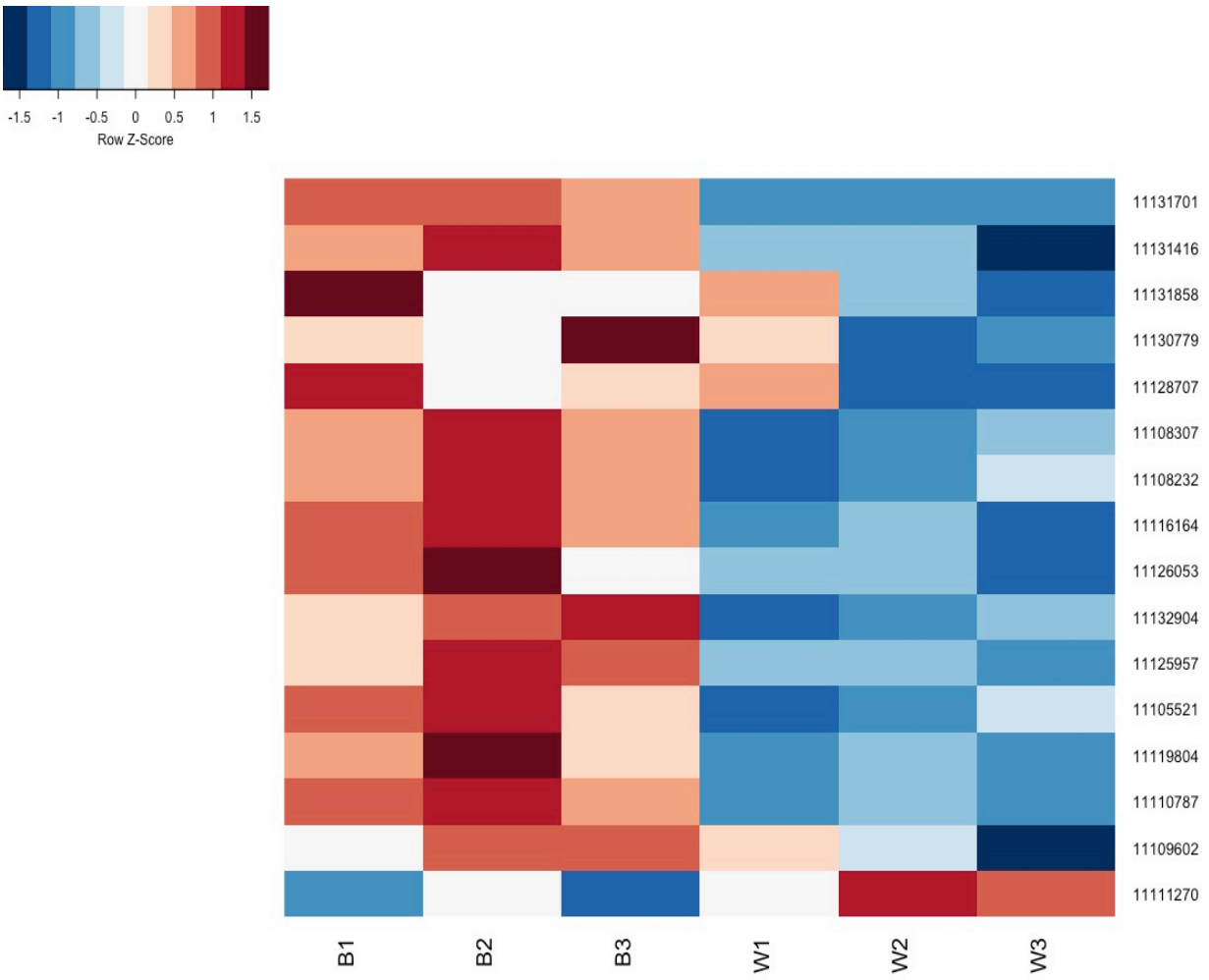


Figure 53. Heatmap showing relative expression of the sixteen genes found to be differentially expressed in either or in common between the pooled replicate and at least two of the three individual replicate level pairwise comparisons

After manual examination of read counts for genes found to be differentially expressed in single pairwise replicate comparisons, 11 of 22 potential DEGs found in replicate A only: 111102592, 111105796, 111106289, 111107493, 111113330, 111122337, 111123832, 111127417, 111130780, 111130793, and 111130993; 10 of 14 potential DEGs found in replicate B only: 111101279, 111101430, 111105231, 111107498, 111115840, 111119804 (also differentially expressed in pooled comparison), 111130134, 111133339, 111138182, and Novel00329, and 14 of 47 potential DEGs found in replicate C only: 111112352, 111112914, 111115637, 111119970, 111124576, 111128439, 111130983, 111131137, 111131323, 111132689, 111133361, 111133576, 111135614, and 111137444; were considered to possibly be linked to pigmentation. The above listed thirty-five genes had significant differential expression in one replicate as well as either non-significant differential expression in the same direction or near even expression in the two other replicate comparisons of tissue types. Expression data for the replicates that displayed significant enrichment are reported below (Table 20; Figure 54).

Table 20. Summary data for thirty-five genes found to be putatively differentially expressed in one individual replicate comparison only

	Gene ID	Annotation	Replicate Enriched	Log_2 (Fold Change)	B Normalized Read Count	W Normalized Read Count	Adjusted *p*-value (BH)
1	111101430	peroxidase-like protein	B (B2 v. W2)	1.37	10302.71	3970.27	7.60E-36
2	111128439	elastin-like	C (B3 v. W3)	1.03	2739.37	1330.15	1.10E-07
3	111131137	uncharacterized	C (B3 v. W3)	1.11	6672.21	3060.49	2.08E-25
4	111115637	uncharacterized	C (B3 v. W3)	-1.06	2928.26	6048.88	2.36E-21
5	111102592	uncharacterized	A (B1 v. W1)	1.04	3597.18	1739.74	2.29E-08
6	111106289	uncharacterized	A (B1 v. W1)	1.07	2808.58	1334.57	1.11E-06
7	111130793	ctenidin-3-like	A (B1 v. W1)	1.53	3868.05	1331.14	4.90E-19
8	111107493	spore coat protein YeeK-like	A (B1 v. W1)	1.09	1801.09	842.89	3.38E-04
9	111137444	uncharacterized	C (B3 v. W3)	-1.93	633.39	2382.35	5.41E-20
10	111124576	snake venom metalloproteinase ACLH-like	C (B3 v. W3)	-1.96	548.81	2108.59	5.52E-18
11	111130983	zonadhesin-like	C (B3 v. W3)	-1.32	598.62	1478.52	7.78E-07
12	111123832	uncharacterized	A (B1 v. W1)	1.29	2142.92	874.58	2.96E-07
13	111127417	low-density lipoprotein receptor-related prot. 1-like	A (B1 v. W1)	1.73	1783.10	535.37	3.38E-10
14	111133576	uncharacterized	C (B3 v. W3)	-1.80	887.12	3046.90	2.17E-23
15	111135614	uncharacterized	C (B3 v. W3)	1.04	1314.71	631.11	1.76E-03
16	111119970	60S ribosomal protein L10-like	C (B3 v. W3)	2.01	794.09	195.39	4.15E-06
17	111133361	proline-rich proteoglycan 2-like	C (B3 v. W3)	-1.28	397.51	956.08	5.33E-04
18	*111119804*	*kynurenine--oxoglutarate transaminase 3-like*	*B (B2 v. W2)*	*2.40*	*1222.05*	*229.55*	*6.21E-10*
19	111130993	neuropeptide-like protein 29	A (B1 v. W1)	-1.46	648.67	1776.57	4.49E-11
20	111107498	heavy metal-assoc. isoprenylated plant prot. 32-like	B (B2 v. W2)	2.30	850.91	171.28	3.23E-06
21	111105231	uncharacterized	B (B2 v. W2)	3.84	316.83	22.07	8.51E-04
22	111132689	uncharacterized	C (B3 v. W3)	1.83	561.97	156.73	1.42E-03
23	Novel00329	uncharacterized	B (B2 v. W2)	-4.46	10.35	226.02	1.39E-03
24	111131323	glycine-rich protein 2-like	C (B3 v. W3)	1.00	8023.57	3968.50	7.78E-23
25	111122337	scuwaprin-a-like	A (B1 v. W1)	1.57	18885.46	6319.08	3.88E-101
26	111101279	uncharacterized	B (B2 v. W2)	1.34	13979.20	5496.76	4.88E-47
27	111130780	circumsporozoite protein-like	A (B1 v. W1)	1.20	13356.25	5780.28	2.48E-44
28	111113330	extracellular superoxide dismutase--like	A (B1 v. W1)	1.48	9282.31	3303.02	1.65E-44
29	111133339	glycine-rich cell wall structural protein 1-like	B (B2 v. W2)	1.10	42171.67	19612.10	1.03E-96
30	111138182	glycine-rich cell wall structural protein-like	B (B2 v. W2)	1.12	40127.16	18370.79	5.48E-96
31	111115840	mantle protein-like	B (B2 v. W2)	1.06	27000.16	12888.15	3.76E-57
32	111105796	5'-3' exoribonuclease 2-like	A (B1 v. W1)	1.10	53514.97	24830.88	3.09E-154
33	111112352	loricrin-like	C (B3 v. W3)	-1.46	141100.37	385331.95	0.00E+00
34	111112914	uncharacterized protein DDB_G0271670-like	C (B3 v. W3)	-1.30	94570.62	230929.38	0.00E+00
35	111130134	spidroin-2-like	B (B2 v. W2)	1.18	92304.11	40424.03	2.00E-249

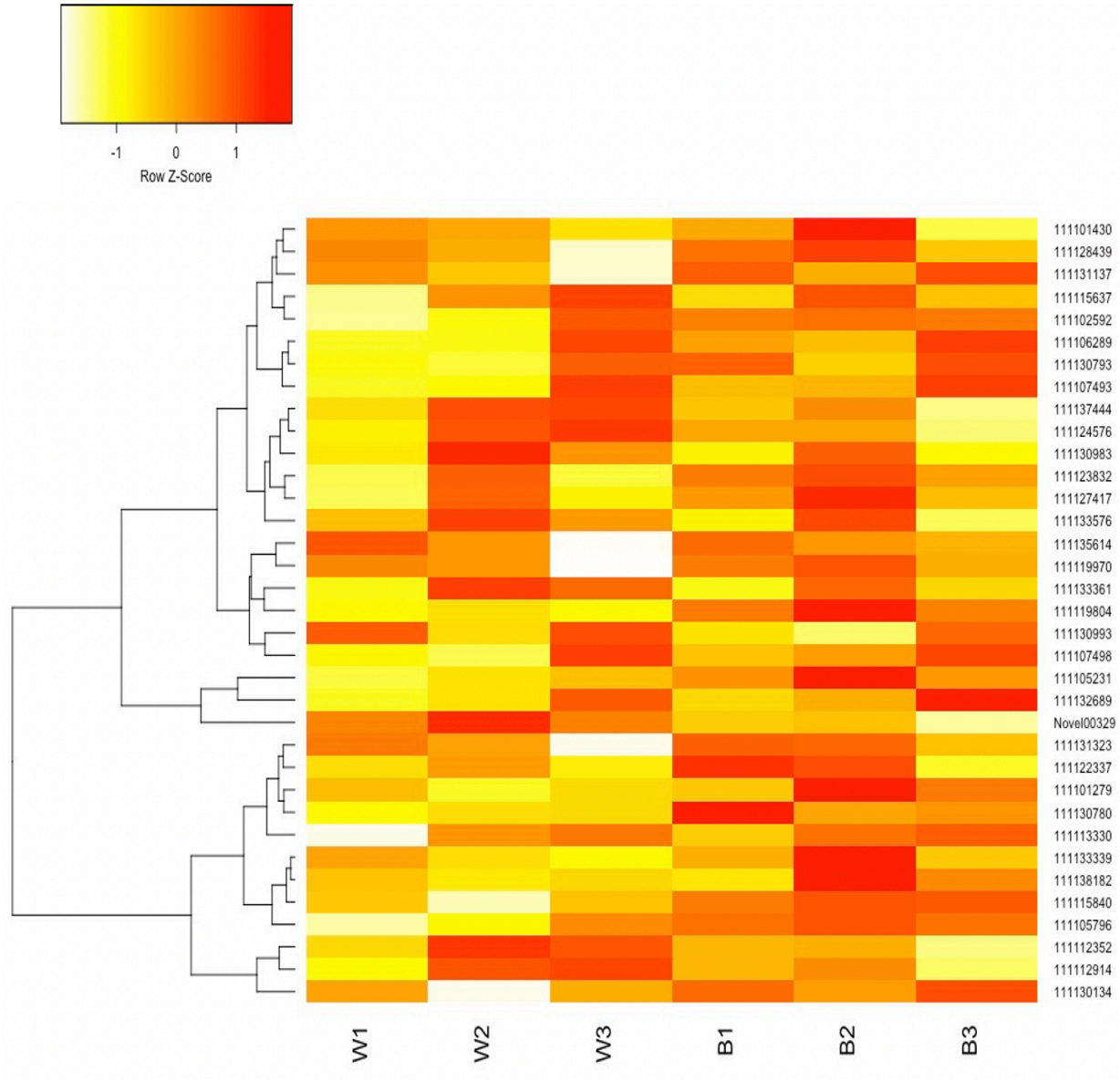


Figure 54. Heatmap showing relative expression of the thirty-five genes found to be putatively differentially expressed in one individual replicate comparison only

DISCUSSION

Transcriptome analysis of dark and light colored mantle tissue in *Crassostrea virginica* identified a small number of individual candidate genes that may potentially have a role in producing pigment. Together, pooled replicate and replicate level pairwise comparisons provide strong evidence of differential expression in just sixteen genes. Additional genes are considered to be of interest based on individual replicate pairwise comparisons as well as manual examination of sample read count data for genes of interest.

The low number of differentially expressed genes found can possibly be explained by the study's experimental design. First, this study examined differences in the expression of just one part of a single tissue, the mantle edge. Also, dark and light tissues were sampled at the same time with a small number of carefully chosen individual oysters contributing to each replicate pool, which were all collected, handled, and maintained under the same conditions. Additionally, each of the three replicate comparisons of gene expression were performed by sampling both dark and light mantle tissue from each individual oyster included in a replicate and then pooling by pigmentation type. In summary, all individuals sampled in this study exhibited pronounced striping patterns on their shells and mantles; they contributed similar amounts of tissue to both dark and light pools in a given replicate; and each individual selected was used in one of three replicates only. This manner of sampling therefore likely minimized the effects of differences in background expression due to environment, tissue type, and individual variation. Indeed, as is indicated by reported coefficients of determination, gene expression was extremely similar in each of the six libraries.

Principal component analysis shows that the libraries clustered by replicate pair rather than by tissue pigmentation type sampled; i.e., gene expression patterns varied more due to background

differences between individuals than by tissue pigmentation type. This provides evidence that genes identified as differentially expressed are strongly linked to differences in pigmentation type, however these genes should nevertheless be understood to be correlated with, but not necessarily the cause of, pigmentation.

In this study, fifteen of the sixteen genes whose differential expression have the most evidence for contributing to differences in pigmentation were up-regulated in dark relative to light mantle tissue. This suggests that upregulation of a number of genes, rather than suppression, leads to pigmentation. This result does not agree with what was found in a transcriptome study of four color morphs (white, black, mottled, and golden) in the Pacific oyster, *Crassostrea gigas*, where RNA transcripts were collected and quantified from samples of whole mantle tissue and white shelled individuals were found to exhibit the highest number of upregulated genes (432 uniquely upregulated genes in the white color morph compared to 91, 31, and 43 upregulated genes in black, mottled, and golden morphs, respectively). This result led the authors to suggest that the white variant is regulated by a molecular mechanism that requires more genes compared to the other color variants, and functional enrichment suggested these genes were related to endocytosis; specifically, enriched transcripts in the white variant were involved in 'nucleotide binding', 'small molecule binding', and 'nucleoside phosphate binding' (Feng et al. 2015).

While most of the differentially expressed genes identified were enriched in dark tissue in the present study, it is worth noting that more genes were uniquely expressed in light mantle tissue than in dark mantle tissue (909 compared to 287). This difference may indicate there is higher transcript diversity in light mantle tissue leading to suppression of pigment production, but it is also likely the difference is due to chance or the result of differences in sequencing depth.

The low number of differentially expressed genes identified led to limited results from GO term and KEGG enrichment analyses; these results mainly suggest enrichment of ion transmembrane transport biological processes, lyase activity molecular functions, as well as porphyrin and glycine, serine and/or threonine metabolic pathways. Further examination of gene set analyses was not performed. Instead, differentially expressed genes and genes of interest were examined individually and will be discussed in turn.

Uncharacterized genes 111131701 and 111131416

The most significantly enriched gene identified in this study was uncharacterized. This gene (111131701) stands out with both the largest average fold-change difference and the most significant read count difference between pigmentation types across all three replicates.

The National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (*BLAST*) yielded no sequences with high similarity to 111131701, except for another uncharacterized gene found in *C. virginica*, 111131416. This gene was also upregulated in dark mantle tissue and the difference between pigmentation types across all three replicates was significant.

These two genes are located close to each other on chromosome four, and a nucleotide sequence alignment of 111131701 and 111131416 showed the query cover was 100% and that the two sequences were 98% identical (E value = 0) (Figure 55). The sequence similarity suggests these transcripts are recent duplicates or misassembled haplotigs. The gene with id 111131416 is fourteen base pairs (5 amino acids) longer than 111131701, although the length of all three exons are the same size. There is a difference in the size of intron 2 between 111131701 and 111131416. The few differences in coding nucleotide sequences result in a difference of 2 amino acid residues.

```
98.7% identity in 157 residues overlap; Score: 787.0; Gap frequency: 0.0%

111131701       1 MNAVIGLLVVFCSLLLLEHSHQEDLQSLHGARDKRAVAVTTEAPEPTEAPEPTDPPRTSR
111131416       6 MNAVIGLLVVFCSLLLLEHSHQEDLQSLHGARDKRAVAVTTEAPEPTEAPEPTDPPRTSR
                  ************************************************************

111131701      61 RWSRKNRLRAAGNSTTGGRRQSRKNRIGGSTNSTTSGGRRRSNRKNRLASSGNSTTSRTS
111131416      66 RWSRKNRLGAAGNSTTGGRRQSRKNRIGGSTNSTTSGGRRRSSRKNRLASSGNSTTSRTS
                  ******** ******************************** ****************

111131701     121 RRNRRKNRRQQQTTTEAPKTTEDTLKPIDRILRNARQ
111131416     126 RRNRRKNRRQQQTTTEAPKTTEDTLKPIDRILRNARQ
                  ***********************************
```

Figure 55. Amino acid sequence alignment of genes 111131701 and 111131416 generated using Clustal Omega.

Utilizing NCBI BLAST, no conserved domain were found in the protein sequences of these gene(s), but a signal peptide with 15 hydrophobic residues (AVIGLLVVFCSLLLL) appears to be present at the N-terminus (Figure 56), suggesting it may be a secreted protein, and gene ontology analysis based on sequence similarity also suggests these gene(s) are involved in protein localization or ion transmembrane transport.

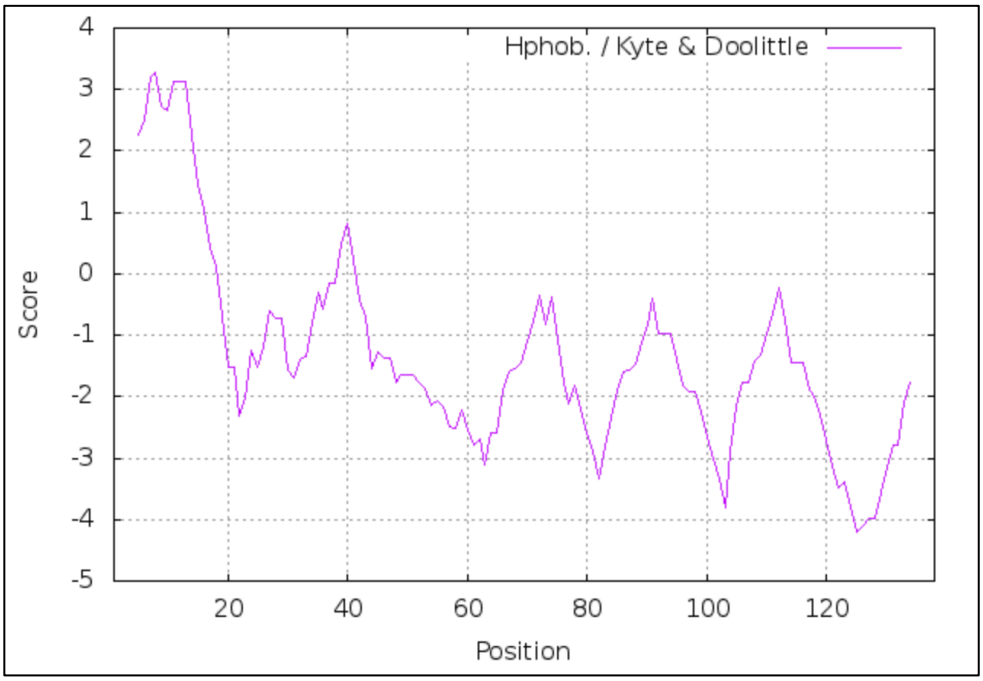


Figure 56. Kyte and Doolittle plot of the degree of hydrophobicity of amino acid sequence for gene 111131701 produced with ExPASy ProtScale tool.

Because this uncharacterized gene(s) is so highly enriched in dark mantle tissue relative to light mantle tissue, its expression is likely linked to pigment production. This gene was renamed "unknown color associated secreted protein from mantle", and should be addressed by future studies.

Additional uncharacterized genes

The uncharacterized gene 111111270 was the only gene found to be significantly downregulated in more than one replicate level comparison in dark relative to light mantle tissue. It was significantly differentially expressed in replicates B and C.

Another uncharacterized gene, 111136004, may have a role in pigment production as it was found to be upregulated in dark tissue by the pooled replicate comparison, however slightly nonsignificant differential expression was observed as the $\log_2$(fold change) for this gene was just 0.96. This gene is probably involved in protein binding based on gene ontology analysis.

Other genes with evidence for enrichment in dark pigmented mantle tissue include 111102592, 111106289, 111101279, 111105231, 111131137, 111123832, and Novel00329, while 111112914, 111115637, 111133576, and 111137444 were potentially enriched in light mantle tissue. Each of these 12 genes was found to be differentially expressed in only one of the three replicate comparison (A, B, or C) and the evidence for their influence on pigmentation is tentative.

It is of note that 111133576 may be similar to gigasin, a lactamase-related protein found in the calcifying shell matrix of *Crassostrea gigas* (Marie et al. 2011). NCBI BLAST results show that this gene has some homology with predicted mesenchyme-specific cell surface glycoprotein-like proteins in both *C. virginica* **(**query cover = 24%; percent identical = 78% ; E score = 1e-54) and *C. gigas* **(**query cover = 16% ; percent identical = 65% ; E score = 7e-13). Several additional

genes upregulated in dark mantle tissue were also similar to shell matrix, glycine-rich, or mantle specific genes.

Shematrins and Other Genes Involved in Biomineralization

Shell matrix proteins (shematrins) are proteins produced exclusively in the mantle and secreted into the extrapallial space where they interact with inorganic ions leading to the crystallization of calcium carbonate and formation of molluscan shells (Yano et al. 2006). To date, three studies of gene expression in pearl oysters, *Pinctada sp.,* have linked the expression of shematrins or shematrin-like proteins as well as mantle proteins involved in shell calcite layer biosynthesis and shell nacreous layer biosynthesis to shell pigment production or incorporation.

Guan, Huang, and He (2011) used suppression subtractive hybridization (SSH) to identify genes related to pigmentation in red shelled pearl oysters, *P. fucata*. Five red shelled and five non-red shelled oysters from the second generation of selection for red shells were used and whole mantle tissues were sampled. Both forward (red shell cDNA served as tester and non-red shell served as driver) and reverse (where non-red shell cDNA served as tester and red shell served as driver) subtractive libraries were constructed simultaneously. Ten genes, Nacrein, Nacrein II, a glycine-rich protein, Aspein, Shem-1, Shem-2, Shem-6, Shem-7, mantle gene 5, and 28s were confirmed to be enriched in red shelled pearl oysters and their expression was evaluated using quantitative PCR using a third generation of red selected oysters. Of these ten genes, variation in just three, She-2-F10, She-7-F10, and Nacrein II-F10, were suggested to be responsible for variation in shell color.

A study with similar methods, where suppression subtractive hybridization was followed by RT-qPCR, was conducted by Lemer, Saulnier, Gueguen, and Planes (2015) in order to identify

genes associated with color in the black-lipped pearl oyster, *Pinctada margaritifera*. After SSH, thirty seven genes were selected for expression analysis of the three studied phenotypes: full albino (FA, white shell and white mantle), half albino (HA, white shell and black mantle), and black phenotype (C, black shell and black mantle). Compared to the C phenotype, four genes were found to be significantly up-regulated in the FA phenotype: shematrin 1 (SHEM 1), peroxidase (PEROX), serine protease inhibitor (SERP), and zinc metalloprotease (ZINC), and two were found in the HA phenotype: lysine-rich mantle protein (KRMP) and serine protease inhibitor (SERP). Eight genes were downregulated in the FA and HA phenotypes: shematrin 9 (SHEM 9), PDZ domain protein (PDZ), Pif177-like protein (PIF), collagen alpha-1 XI chain precursor (COLL), shematrin 4 (SHEM 4), mantle protein 8 (MP8), aspein shell matrix protein (ASP), and prism uncharacterized shell protein (PRISM). Two additional genes were found to be downregulated in the HA phenotype only relative to the C phenotype: flavonol cinnamoyl COA reductase related protein (FLAV) and zinc metalloprotease (ZINC).

Based on this result, the authors selected 9 genes for subsequent individual RT-qPCR analysis based on their high differential expression level and their known or predicted role at different levels of the shell formation: SHEM 1, SHEM 4, PDZ, PIF, MP8, PRISM, SERP, ZINC and KRMP. In addition, although they did not show significant differential expression in the pooled analysis, two more genes, CHIT and TYR 2a, were included in the individual RT-qPCR analysis because of their known role in melanin biosynthesis. The authors concluded that the expression of SHEM 4, MP8, KRMP, CHIT, and SERP were most likely responsible for the variability of shell color among black colored pearl oysters.

A more recent transcriptomic and proteomic study by Xu et al. (2019) had some similar conclusions as the authors also found the expression of three nacre protein coding genes, two

uncharacterized shell protein coding genes, and four mantle genes, among others, including two calmodulin genes (which were differentially expressed in different directions), astacin, putative tyrosinase-like protein 3, sulfide:quinone oxidoreductase, and retinoic acid binding protein, to be associated with the difference in pigmentation of yellow and black shell pearl oysters, *Pinctada fucata martensii*. Among the fifty differentially expressed proteins were a glycine-rich protein and mantle gene 4.

In the present study, many mantle specific or shematrin-like genes were found to be enriched in dark mantle tissue relative to light mantle tissue. The shematrin-like gene 111128707 and the glycine-rich cell wall structural protein-like gene 111131858 were found to be uniquely differentially expressed in all three replicate level pairwise comparisons. A glycine-rich cell wall structural protein 1-like gene, 111130779, was significantly enriched in both replicates B and C. A mantle protein like gene, 111115840, as well as the glycine-rich cell wall structural protein 1-like gene, 111133339, a glycine-rich cell wall structural protein like gene, 111138182, and a glycine-rich cell wall structural protein 2 like gene, 111131323, were each found to be enriched in one pairwise replicate comparison. A gene coding for a shematrin-like protein 2, 111130982, was found to be upregulated (not significant) in dark mantle tissue. These differentially expressed shematrin-like genes are highly similar to each other and also share some sequence similarity with a few other genes found to be enriched in dark tissue in one replicate comparison, including a ctenidin-3 like gene, 111130793, circumsporozoite protein-like gene, 111130780, and spidroin-2-like gene, 111130134. The identities of these genes need to be confirmed in future studies.

Elastin and Loricrin

In this study, two genes that code for elastin were upregulated in dark relative to light tissue. 111109602 was up-regulated in replicates B and C, while 111128439 was up-regulated in replicate C only. Elastin is the dominant extracellular matrix protein found in many connective tissues, such as the epidermis, and helps to form a network of fibers which give these tissues flexibility. These proteins are highly hydrophobic and are similar to collagen as they are rich in proline and glycine. Unlike collagen though, they are not glycosylated and they contain some hydroxyproline, but no hydroxylysine (Alberts et al. 2015). A proline-rich proteoglycan 2-like gene, 111133361, has some sequence similarity with collagen and was found to be upregulated in dark mantle tissue in replicate C.

Loricrin is another gene involved in the formation and maintenance of the epidermis. This gene is especially important in the formation of the stratum corneum, the outer tissue that provides a tough barrier between the body and its environment. Each cell in the stratum corneum, called a corneocyte, is surrounded by a protein shell called a cornified envelope (CE), of which loricrin proteins make up about 70% (Nithya, Radhika, and Jeddy 2015). The loricrin-like gene, 111112352, was downregulated in dark relative to light mantle tissue in *C. virginica*.

Tetrapyrroles

Tetrapyrrole pigments occur as either cyclic structures known as porphyrins or as linear structures known as bilins and are synthesized as side products of the highly conserved eight step heme pathway (Figure 57). The biosynthesis of heme in animals begins with the Shemin pathway where the rate limiting enzyme 5-aminolevulinic acid synthase (ALAS) catalyzes the condensation of glycine with succinyl coenzyme A to form 5-aminolevulinic acid. Two molecules of 5-aminolevulinic acid are then condensed with each other by the enzyme delta-aminolevulinic acid

dehydratase (ALAD) to form porphobilinogen. Next, four molecules of porphobilinogen are condensed by the third enzyme in the pathway, porphobilinogen deaminase (PBGD), which yields hydroxymethylbilane, a highly unstable linear tetrapyrrole. Hydroxymethylbilane is converted to uroporphyrinogen III by the uroporphyrinogen III synthase (UROS).

Heme production proceeds as uroporphyrinogen decarboxylase converts the acetic acid groups of uroporphyrinogen III to methyl groups yielding coproporphyrinogen III. Oxidative decarboxylation of the two propionic acid groups of coproporphyrinogen III are then catalyzed by coproporphyrinogen oxidase (CPOX) to yield protoporphyrinogen IX. Protoporphyrinogen IX is then oxidized to protoporphyrin IX by protoporphyrinogen oxidase (PPOX), and ferrochelatase (FECH) catalyses the insertion of ferrous iron into protoporphyrin IX to produce heme (Lim 2010).

In the eastern oyster mantle samples, the first three enzymes in the heme pathway, but not the fourth, fifth, sixth, seventh, or eighth, were significantly upregulated in dark tissue.

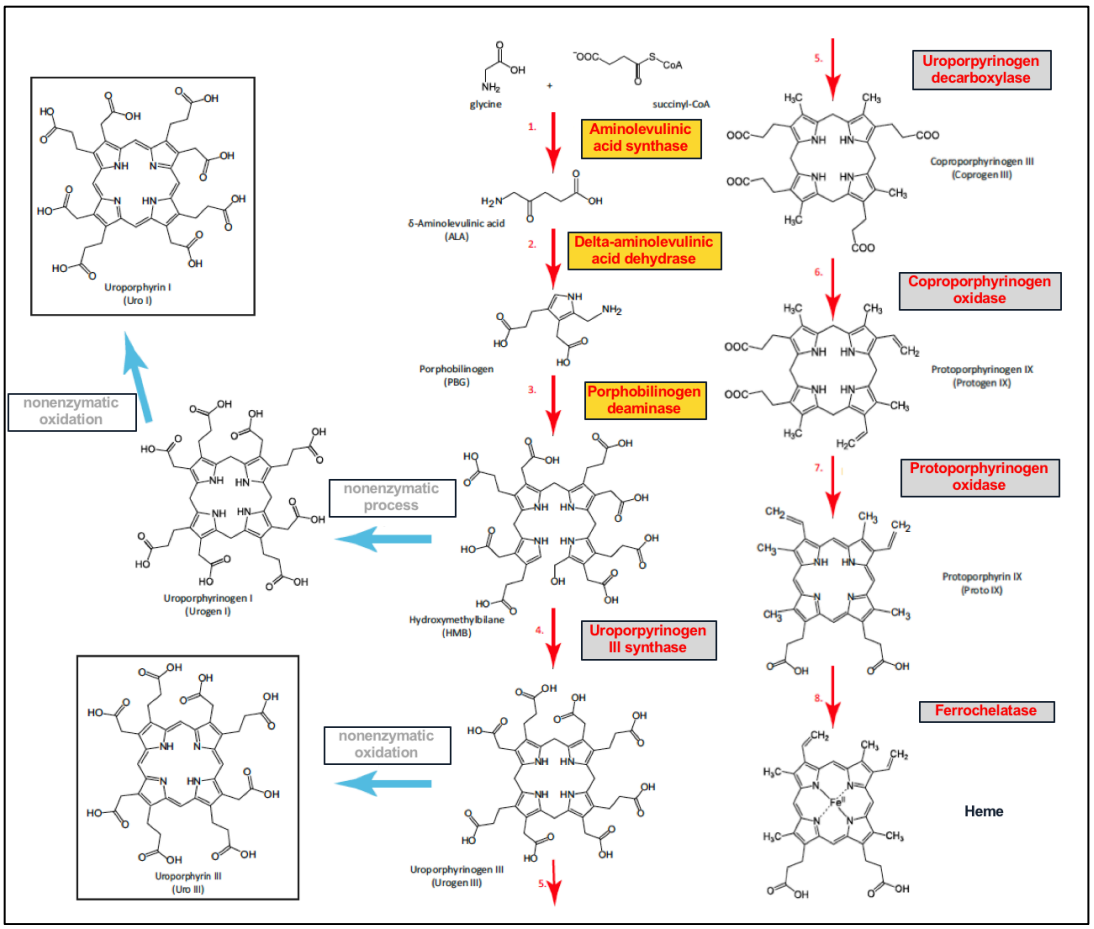


Figure 57. The first three enzymes in the heme pathway (highlighted in yellow) were significantly enriched in dark mantle tissue, while the other five enzymes in the pathway were not (in grey). Source: Modified from Williams et al. 2017.

111126157 had a high $\log_2$(fold change), but an adjusted *p*-value slightly higher than 0.05. This gene is one of two found that codes for porphobilinogen deaminase (PBGD).

Upregulation of these three enzymes in dark mantle tissue provides strong evidence that the heme pathway is involved in pigment production in the eastern oyster. The enrichment of ALAS, ALAD, and PBGD in dark mantle tissue, but not of UROS, UROD, CPOX, PPOX, and FECH suggests that hydroxymethylbilane may be accumulating in cells. The accumulation of hydroxymethylbilane or uroporphyrinogen III can lead to the production of the pigments uroporphyrin I or uroporphyrin III, respectively, through nonenzymatic processes. Uroporphyrin I is produced when the expression of uroporphyrinogen synthase (UROS) is decreased or the

expression of porphobilinogen deaminase (PBDB) is increased. Uroporphyrin III is produced when the fifth enzyme in the pathway, uroporphyrinogen decarboxylase (UROD), is decreased (Milgrom 1997; Williams et al. 2017).

It is possible that one or both of these porphyrin pigments could be incorporated in shells to produced color or that their ring structures could be broken down to linear tetrapyrrole pigments. Follow up studies are needed to confirm if porphyrins or linear tetrapyrroles are present in shells.

ATP Binding Cassette Transporters

In the heme pathway, ALAS is considered the rate limiting enzyme, and the upregulation of it, ALAD and PBGD leaves open the possibility that even the full heme pathway, rather than just part of it, could be enriched. One piece of evidence supporting this possibility is the upregulation of another rate limiting enzyme, ATP binding cassette transporter B6 (ABCB6), 111125957. This gene, which was significantly upregulated in the pooled replicate comparison, has been shown to be uniquely associated with the outer mitochondrial membrane and it interacts specifically with heme and other porphyrins, which it can transport into the mitochondria. Coporphyrinogen III is the most probable substrate of ABCB6 and its rate limiting transfer through the mitochondrial membrane facilitates heme biosynthesis (Donegan et al. 2018) (Figure 58).

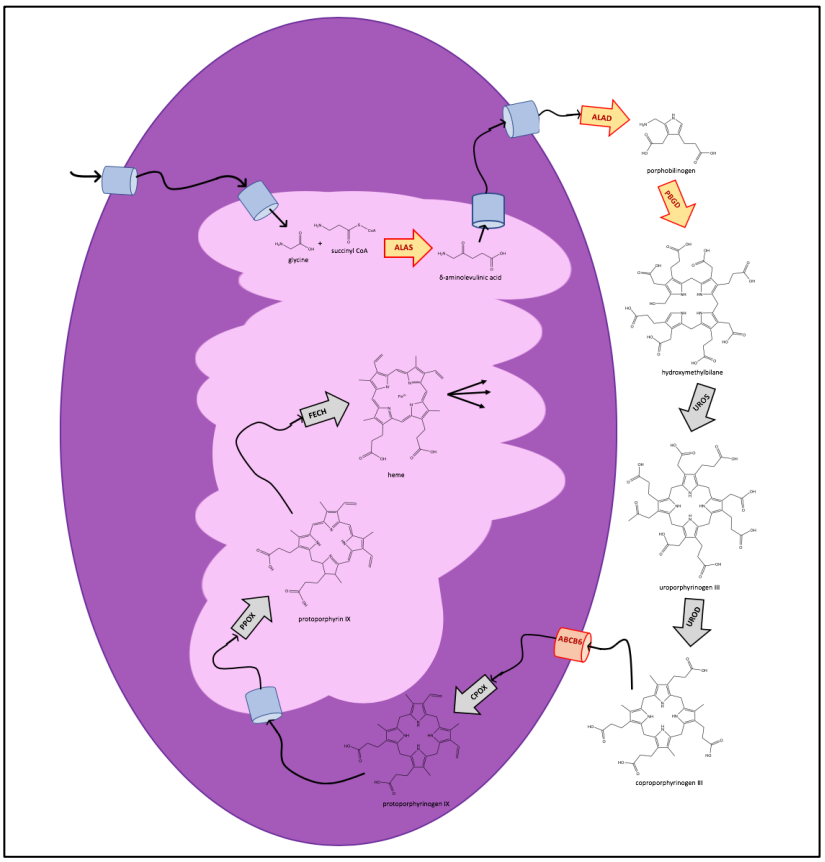


Figure 58. Four of the eight enzymatic reactions in the heme pathway occur in the mitochondria. The ATP binding cassette transporter B6 is likely involved in the transport of coporphyrinogen III into the mitochondria. Adapted from Donegan et al. 2018.

If the complete heme pathway is enriched, increased levels of heme can breakdown to produce a number of pigments (Rüdiger 1970) (Figure 59).

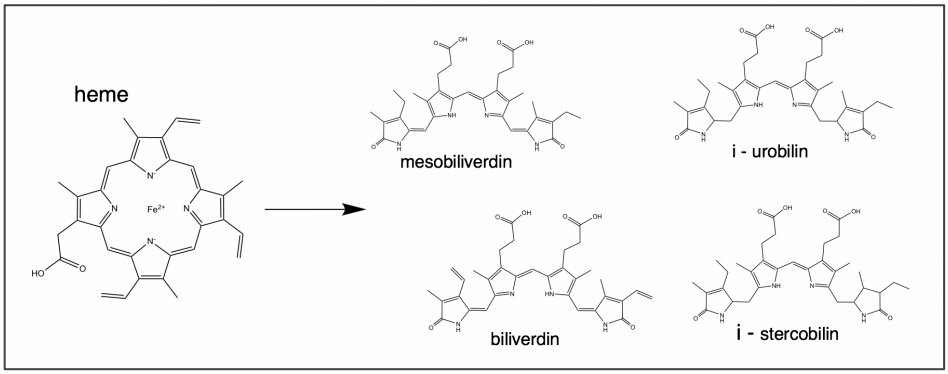


Figure 59. Examples of bile pigments formed by the breakdown of heme. Adapted from Rüdiger 1970.

Another ATP binding cassette transporter gene, ATP binding cassette transporter a1 (111132904) was also upregulated in dark tissue in replicates B and C as well as in the pooled replicate comparison in this study. The transcriptome study of shell pigmentation in the Pacific oyster, *Crassostrea gigas*, also observed differential expression in ABC transporters, among several other genes. The ATP-binding cassette genes Abca1, Abca3, and Abcb1 were associated with the white color variant in *C. gigas*. It was hypothesized that these ABC transporters may be responsible for the transportation of pigment related substrates and export cholesterol to decrease melanin production. Abca1 appeared to be especially associated with non-pigmented shells in *C. gigas* (Feng et al. 2015).

Rho GTPase activating protein

A gene that codes for rho GTPase-activating protein gacII, 111105521, was found to be significantly upregulated in dark tissue in the pooled replicate comparison. Rho GTPases are small G proteins of the Ras signaling protein superfamily that control several vital cellular processes, including cytoskeletal dynamics, cell cycle progression, and cell transformation. They are also involved in the regulation of gene expression and in growth-promoting and anti-apoptotic processes (Alberts et al. 2015; Li et al. 2015). The transcriptome study of shell pigmentation in *C. gigas* found that the three variants with pigmentation had genes with the GO term ‘endocytosis’ significantly enriched. At least five enriched genes were found that have an effect on this process, including three G proteins: Rab5, Rab7, and Rab11, and the authors suggest this process may influence pigment granule formation (Feng et al 2015). It is not immediately apparent how the upregulation of this G protein may be linked to pigment production in *C. viriginica*, but it is also possible it may play a regulatory role in melanin production (Amsen 2009) or, that through their

role in regulating actin cytoskeleton polymerization, branching and bundling, it may stimulate or inhibit melanocyte dendrite formation (Scott 2002).

Melanins

Melanins are polymers of indolequinone and dihydoxyindole produced by membrane bound cytoplasmic organelles that occur in specialized dendritic cells called melanocytes. Melanin granules are transferred on microtubules by dendrites to neighboring cells, where they accumulate. The regulatory pathways leading to melanin production are well studied in birds and mammals, but are not fully characterized in molluscs (Nordlund et al. 2006; Williams 2017).

Melanins exist in two forms: eumelanin and pheomelanin. Eumelanin is produced from precursors by a pathway involving three enzymes. In the absence of sulfhydryl compounds, the rate-limiting enzyme tyrosinase (TYR) catalyzes the hydroxylation of the amino acid tyrosine to dopa, the oxidation of DOPA to dopaquinone, and the oxidation of 5,6-dihydroxyindole (DHI) to indole-5,6-quinone. Tyrosinase related protein 1 (TRPl) functions as 5,6-dihydroxyindole-2-carboxylic acid (DHICA) oxidase whereas tyrosinase related protein 2 (TRP2) functions as dopachrome tautomerase. When the level of sulfhydryl compounds such as cysteine or glutathione is higher than the level of dopaquinone produced enzymically by tyrosinase, cysteinyldopas are produced exclusively. Subsequent oxidative cyclization and polymerization of cysteinyldopas results in the production of pheomelanins (Ito et al. 2011) (Figure 60).

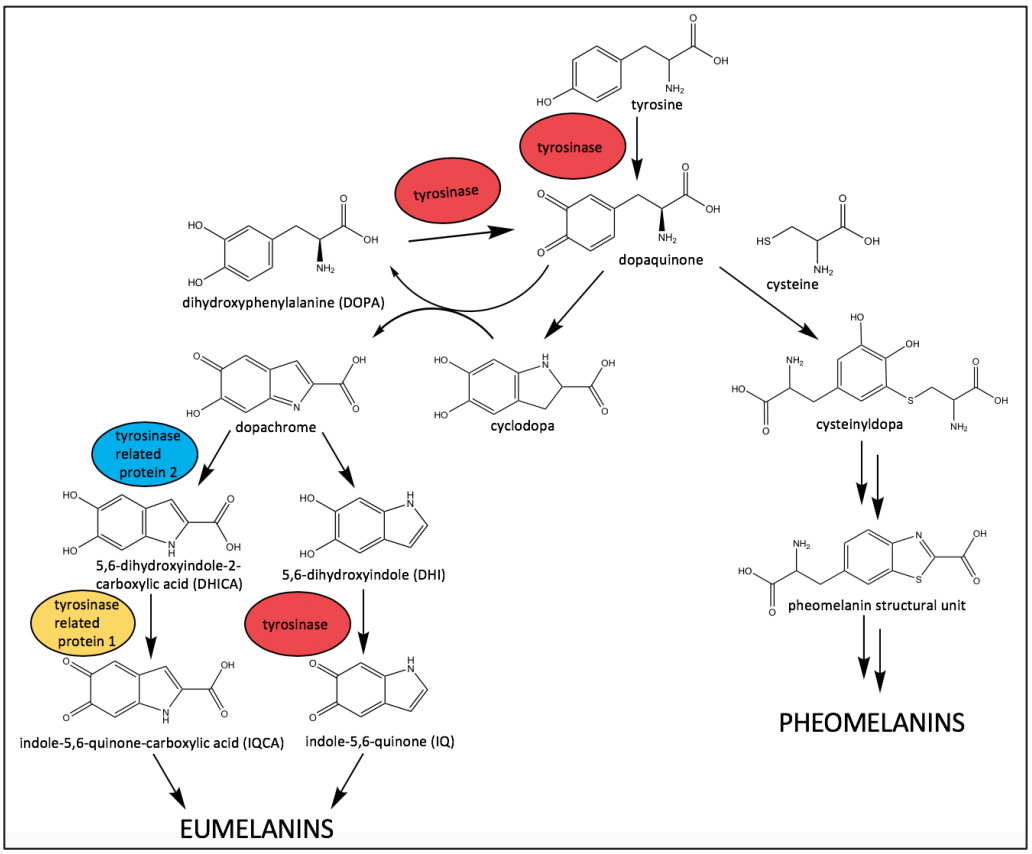


Figure 60. Schematic of the Raper-Mason pathway outlining melanogenesis. The amount of cysteine present determines how much eumelanin and pheomelanin is produced. Adapted from Ito et al. 2011.

In the present study, no tyrosinase or tyrosinase related genes were identified as being significantly differentially expression by the DESeq2 pipeline. However, after examination of expression data of 20 tyrosinase-like genes and an additional 20 uncharacterized transcripts with putative tyrosinase function, at least two were found to be slightly enriched in dark mantle tissue: 111103632 and 111137389 (Figure 61). 111103632 is somewhat similar to *C. gigas* tyrosinase-like protein 2 (query cover =44% ; percent identical = 64.89% ; E score = 8e-31), and 111137389 was somewhat similar to *Pinctada maxima* tyrosinase A3 (query cover = 44% ; percent identical = 70.49% ; E score = 9e-106) and to *Pinctada martensi* tyrosinase protein 1 (query cover =57% ; percent identical = 70.07% ; E score = 3e-100). Nucleotide sequences for these genes share some similarity.

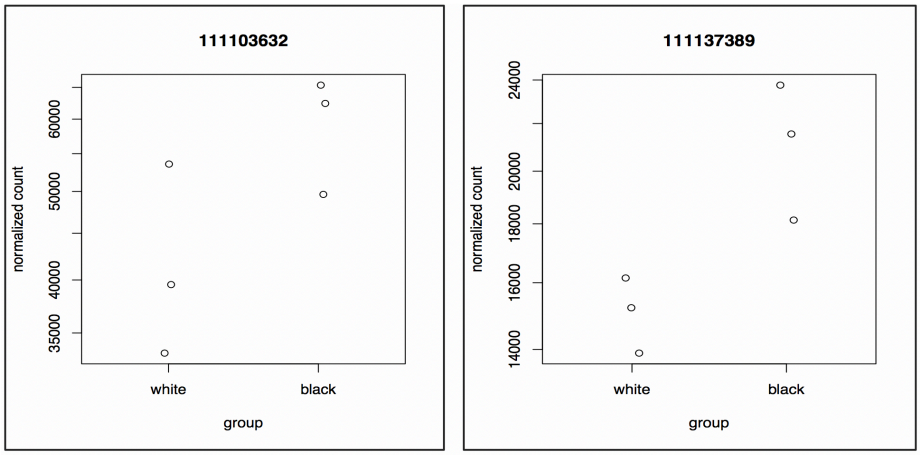


Figure 61. Although not significantly differentially expressed according to DESeq2, plots of the expression of genes 111103632 and 111137389 suggest that these tyrosinase-like genes may be enriched in dark relative to light mantle tissue.

This result, of little differential expression of tyrosinase genes, is in agreement with the findings of the transcriptome study of shell pigmentation in *C. gigas* where tyrosinase was also of special interest. The authors identified twelve transcripts coding for tyrosinase related genes from DEGs. Of these, none were identified that were differentially expressed between the BS and WS groups, however the BS and GS had six tyrosinase-related genes differentially expressed between them (Feng et al. 2015).

In the present study, the expression of 111107924, which codes for microphthalmia transcription factor (MITF), was examined because it is a transcription factor that regulates tyrosinase expression (Pillaiyar, Manickam, and Jung 2017) (Figure 62). This gene of interest was not differentially expressed in oyster mantle tissue.

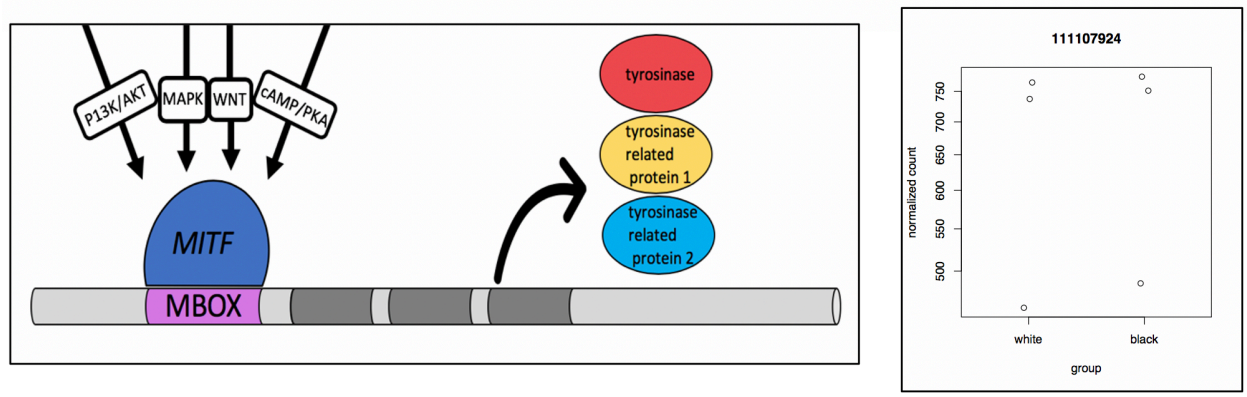


Figure 62. Left: At least four pathways regulate MIFT, a transcription factor which regulates the expression of tyrosinase and tyrosinase related proteins. Right: In this study, MIFT was not differentially expressed in any of the three replicate comparisons.

REFERENCES

Alberts, B., Johnson, A.D., Lewis, J., Morgan, D., Raff, M., Roberts, K., and Walter, P. 2015. Molecular Biology of the Cell, Sixth Edition. New York, NY: Garland Science.

Amsen, E.M. 2009. Studies of Proteins that Regulate Melanin Synthesis and Distribution (Doctoral dissertation, University of Toronto, Toronto, ON, CA). Retrieved from Library and Archives Canada, ISBN: 978-0-494-59041-6.

Anders, S., and Huber, W. 2010. Differential expression analysis for sequence count data. Genome biology, 11(10), R106.

Anders, S., Pyl, P. T., and Huber, W. 2015. HTSeq—a Python framework to work with highthroughput sequencing data. Bioinformatics, 31(2): 166-169.

Benjamini, Y. and Hochberg, Y. 1995. Controlling False Discovery Rate: A practical and powerful approach to multiple testing, *Journal of Royal Statistical Society*, 57: 289-300.

Comfort, A. 1950. The Pigmentation of Molluscan Shells. *Biological Reviews*, 26: 285–301.

Cufflinks : cufflinks.cbcb.umd.edu

Donegan, R.K., Moore, C.M., Hanna, D.A., and Reddi, A.R. 2018. Handling Heme: the mechanisms underlying the movement of heme within and between cells. *Free Radical Biology and Medicine*, 133: 88-100

Dündar, F., Skrabanek, L., and Zumbo, P. 2019. Introduction to differential gene expression analysis using RNA-seq. Applied Bioinformatics Core.

Feng, D., Li, Q., Yu, H., Zhao, X., and Kong, L.F. 2015. Comparative Transcriptome Analysis of the Pacific Oyster *Crassostrea gigas* Characterized by Shell Colors: identification of genetic bases potentially involved in pigmentation. *PLoS One*, 10: e0145257.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. 2005. *"Protein Identification and Analysis Tools on the ExPASy Server".* In: The Proteomics Protocols Handbook. Ed. Walker, J.M. Humana Press.

The Gene Ontology Consortium. 2000. Gene Ontology: tool for the unification of biology. 25(1): 25-29.

Gonzalez, I. 2014. Statistical analysis of RNA-seq data: tutorial. Plateforme Biostatistique – IMT Universite Toulouse III.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. 2010. A New Bioinformatics Analysis Tools Framework at EMBL-EBI. *Nucleic acids research,* Suppl: W695-9.

Guan, Y.Y., Huang, L.M., and He, M.X. 2011. Construction of cDNA Subtractive Library from Pearl Oyster (*Pinctada fucata* Gould) with Red Color Shell by SSH. *Chinese Journal of Oceanology and Limnology*, 29(3): 616–622.

HTSeq: Analysis high-throughput Sequencing data with Python: http://wwwhuber.embl.de/users/anders/HTSeq/doc/overview.html

Ito, S., Wakamatsu, K., d'Ischia, M., Napolitano, A., and Pezzella, A. 2011. "Structure of Melanins." In: Melanins and Melanosomes: Biogenesis, physiological, and pathological functions, 1st Edition. Ed. Borovansky, J. and Riley, P.A. Weinheim, GER: Wiley-VCH Verlag & Co. KGaA, Boschstr.

Nordlund, J.J., Boissy, R.E., Hearing, V.J., King, R.A., Oetting, W.S., Ortonne, J-P (Editors). 2006. The Pigmentary System: physiology and pathophysiology, 2nd Edition. Malden, MA, USA; Oxford, OX, UK; and Carlton, VIC, AU: Blackwell Publishing Ltd.

Kanehisa, M., and Goto, S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*(1), 27–30.

Kanehisa M, Goto S, Sato Y, Furumichi M, and Tanabe M. 2012. KEGG for Integration and Interpretation of Large-scale Molecular Data Sets. *Nucleic Acid Research*, 40, D109-D114.

Korpelainen, E., Tuimala, J., Somervuo, P., Huss, M., and Wong, G. 2015. RNA-seq Data Analysis: A Practical Approach. Boca Raton, Chapman & Hall/CRC Press.

Langmead, B. and Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4): 357–359.

Lemer, S., Saulnier, D., Gueguen, Y., and Planes, S. 2015. Identification of Genes Associated with Shell Color in the Black-lipped Pearl Oyster, *Pinctada margaritifera*. *BMC Genomics*, 16:14.

Li, X., Wang, R., Xun, X., Jiao, W., Zhang, M., Wang, S., Wang, S., Zhang, L., Huang, X., Hu, X., and Bao, Z. 2015. The Rho GTPase Family Genes in Bivalvia Genomes: Sequence, Evolution and Expression Analysis. *PLoS ONE* 10(12): e0143932.

Lim, C. H. 2010. High-Performance Liquid Chromatography and Mass Spectrometry of Porphyrins, Chlorophylls, and Bilins. Toh Tuck Link, Singapore; Hackensack, NJ; and London, UK: World Scientific Publishing Company.

Love, M., Anders, S., and Huber, W. 2014. Beginner's guide to using the DESeq2 package. *Genome Biology*, 15, 550.

Love, M., Huber, W., and Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Genome Biology, 15, 550.

Marie, B., Zanella-Cléon, I., Guichard, N., Becchi, M., and Marin, F. 2011. Novel Proteins from the Calcifying Shell Matrix of the Pacific oyster *Crassostrea gigas.* Marine Biotechnology, 13:1159–1168.

Marin, F., Le Roy, N., and Marie, B. 2012. The Formation and Mineralization of Mollusk Shell. *Frontiers in Bioscience*, S4: 1099-1125.

Milgrom, L.R. 1997. The Colours of Life: an introduction to the chemistry of porphyrins and related compounds. Oxford, UK; New York, NY; Tokyo, JPN: Oxford University Press.

Nithya, S., Radhika, T., and Jeddy, N. 2015. Loricrin: an overview. *Journal of Oral and Maxillofacial Pathology*, 19(1): 64-68.

Pillaiyar, T., Manickam, M. and Jung, S-H. 2017. Downregulation of Melanogenesis: drug discovery and therapeutic options. *Drug Discovery Today*, 22(2): 282-298.

Roberts, A., Pimentel, H., Trapnell, C., and Pachter, L. 2011. Identification of Novel Transcripts in Annotated Genomes Using RNA-Seq. *Bioinformatics*, 27 (17): 2325–2329.

Rüdiger, W. 1970. Recent Chemistry and Biochemistry of Bile Pigments. *Angewandte Chemie International Edition*, 9(7): 473-480.

Scott, G. 2002. Rac and Rho: the story behind melanocyte dendrite formation. *Pigment Cell Research*, 15: 322-330.

Sievers, F., Wilm, A., Dineen, D.G., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., and Higgins, D. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Molecular Systems Biology,* 7: 539.

Subramanian A., Tamayo P., Vamsi K. M., Mukherjee, S., Benjami, L.E, Paulovich, A., Pomeroy, S.L., Golub, T. R., Lander, E.S., and Mesirov, J.P. 2005. Gene Set Enrichment Analysis: a knowledge bases approach for interpreting genome-wide expression profiles

Proceedings of the National Academy of Sciences of the United States of America,102: 15545-15550.

TopHat : tophat.cbcb.umd.edu/manual.html

Wang, Y. Sun, M. (Eds). 2018. Transcriptome Data Analysis: methods and protocols. New York, NY: Springer Nature.

Warnes, G., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., Schwartz, M., Venables, B., and Galili, T. 2015. gplots: Various R Programming Tools for Plotting Data. R package version 2.17.0. 2(4): https://doi.org/10.1111/j.0022-3646.1997.00569.x.arXiv:1510.05677.

Watson, J., Baker, T., Bell, S., Gann, A., Levine, M., and Losick, R. 2014. Molecular Biology of the Gene, 7th Edition. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Wilhelm, B.T. and Landry, J.R. 2009. RNA-Seq: quantitative measurement of expression through massively parallel RNA-sequencing. *Methods*, 48(3): 249-257.

Williams, S.T. 2017. Molluscan Shell Colour. *Biological Reviews*, 92(2):1039–1058.

Williams, S.T., Lockyer, A.E., Dyat, P., Nakano, T., Churchill, C.K.C., and Speiser, D.I. 2017. Colorful Seashells: identification of haem pathway genes associated with the synthesis of porphyrin shell color in marine snails. *Ecology and Evolution*, 7: 10379-10397.

Wolf, J.B.W. 2013. Principles of Transcriptome Analysis and Gene Expression Quantification: an RNA-seq tutorial. *Molecular Ecology Resources*, 13: 559-572.

Xu, M., Huang, J., Shi, Y., Zhang, H., and He, M. 2019. Comparative transcriptomic and proteomic analysis of yellow shell and black shell pearl oysters, *Pinctada fucata martensii*. *BMC Genomics*, 20: 469.

Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., Kong, L., Gao, G., Li, C. Y., and Wei, L. 2011. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic acids research*, *39*(Web Server issue): W316–W322.

Yano, M., Nagai, K., Morimoto, K., and Miyamoto, H. 2006. Shematrin: a family of glycine rich structural proteins in the shell of the pearl oyster *Pinctada fucata*. *Comparative Biochemistry & Physiology Part B*, 144(2): 254–262.

Young, M.D., Wakefield, M.J., Smyth, G.K., and Oshlack, A. 2010. Gene Ontology Analysis for RNA-seq: accounting for selection bias. *Genome Biology*, 11, R14.