Summer 2020

An Assessment Of The Population Genetic Structure And Evidence Of Multiple Paternity In Channeled Whelk, Busycotypus Canaliculatus, Along The Us Atlantic Coast

Samantha E. Askin
College of William and Mary - Virginia Institute of Marine Science, samantha.e.askin@gmail.com

An Assessment of the Population Genetic Structure and Evidence of Multiple Paternity in Channeled Whelk, *Busycotypus canaliculatus*, Along the US Atlantic Coast

---

A Thesis

Presented to

The Faculty of the School of Marine Science

The College of William & Mary

In Partial Fulfillment

of the Requirements for the Degree of

Master of Science

---

by

Samantha Elaine Askin

August 2020

APPROVAL PAGE

This thesis is submitted in partial fulfillment of

the requirements for the degree of

Master of Science

Samantha Elaine Askin

Approved by the Committee, August 2020

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

Robert A. Fisher, M.S.

Mary C. Fabrizio, Ph.D

Emily B. Rivest, Ph.D

This Master's thesis is dedicated to my parents: Ron Askin and Janice Askin.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Thank you to my thesis committee for their contributions to the I would like to thank my advisor Dr. Jan McDowell for the mentorship and guidance throughout the entirety of my thesis. For the past three years Dr. Jan McDowell has had a profound impact on the way that I conduct science. She is an amazing example of a woman in science and has helped me develop the skills that I need to be successful in my future endeavors, wherever that takes me. I am forever grateful for the encouragement she has shown towards my career goals and professional development in environmental education. I will always consider her as my academic mother, and I am proud to be a part of the McDowell lab family.

I would also like to thank Bob Fisher for his insight and mentorship. Throughout this project Bob has provided me with valuable experiences and has facilitated important relationships with various groups of people including industry members and management agencies. I would also like to thank Bob for the wonderful sampling trips out to the Eastern Shore, I'll remember those days dearly.

In addition, I would like to thank all the members of the Fisheries Genetics Lab (McDowell lab), who assisted me throughout this project: Ellen Biesack, Heidi Brightman, Jingwei Song, Ann Ropp, Lindsey Nelson, and Jackson Martinez. Thank you to the graduate student community for providing me with emotional support throughout my time at VIMS.

design and execution of this project. I would also like to thank Rick Robbins, Massachusetts Division of Marine Fisheries, and South Carolina Department of Natural Resources for providing samples for this project. Thank you to my funding sources: Fisheries Resource Grant Program, International Women's Fishing Association, ASLO Multicultural Program, and VIMS Office of Academic Studies for providing fiscal support for this project.

Finally, I would like to thank my family for always providing the right amount of laughter and stability in my life. I love you.

LIST OF TABLES

LIST OF FIGURES

ABSTRACT PAGE

Channeled whelk (*Busycotypus canaliculatus* Linnaeus, 1758) are predatory marine gastropods that are found in intertidal regions of the continental slope along the U.S. Atlantic coast from Cape Cod, Massachusetts to Cape Canaveral, Florida. The channeled whelk is a commercially important species that supports a fishery along the Atlantic coast of the United States. The resource is managed at the state level with minimum landing size (MLS) varying by state. Biological assessments of channeled whelk in the mid-Atlantic and Massachusetts have indicated that females have a low probability of being mature when they enter the fishery. The life history characteristics of channeled whelk including slow growth, late maturation, and direct development paired with unsuitable MLS make this species vulnerable to overexploitation. Currently, the population genetic structure of channeled whelk is unknown, impeding the creation of appropriate management strategies.

This study aimed to delineate the population genetic structure of channeled whelk using genotyping-by-sequencing of 252 samples from 10 resource areas from Massachusetts to South Carolina, with fine-scale sampling in the mid-Atlantic region. A total of 2,570 single nucleotide polymorphisms (SNPs) were used to estimate genetic diversity and delineate population structure among resource areas. Analysis of all 10 resource areas revealed four major genetic groups: 1) North and mid-Atlantic resource areas, 2) Pamlico Sound, North Carolina, 3) Wilmington, North Carolina, and 4) Charleston, South Carolina with additional structuring within the north and mid-Atlantic. Analysis North and mid-Atlantic resource samples indicated that Virginia Beach, Virginia was significantly different from all other mid-Atlantic resource areas (F_{ST} = 0.001 – 0.016, $p < 0.001$). Massachusetts and Rhode Island were also significantly different from each other (F_{ST} = 0.084, $p < 0.001$) and from mid-Atlantic samples (F_{ST} = 0.031 – 0.128, $p < 0.001$). Overall, data indicate 7 genetically distinct populations along US East Coast. Estimates of genetic divergence (F_{ST}) spanned several orders of magnitude with elevated divergence levels observed when comparing samples from across known biogeographic barriers along the Atlantic coast. F_{ST} values were highest (0.150 – 0.465, $p <0.001$) when comparing samples off Pamlico Sound, North Carolina, off Wilmington, North Carolina, and off Charleston, South Carolina to samples from all other resource areas. The complex population genetic structure revealed by this study underscores the need for further study and for new management strategies for the channeled whelk fishery in resource areas along the Atlantic coast. Results from this study can be used to assess alternative decisions about the appropriate scale for management of the channeled whelk resource in the mid-Atlantic region, which is threatened by overharvest and removal of whelk before they are sexually mature.

An Assessment of the Population Genetic Structure and Evidence of Multiple Paternity in Channeled Whelk, Busycotypus canaliculatus, Along the US Atlantic Coast

INTRODUCTION

The phylum Mollusca is extremely diverse and is comprised of over 50,000 described living species, with approximately 35,000 described fossil species (Barnes 1987; Bunje 2003a). Molluscs have diversified to allow a wide range in size, structure, and habitat selection through adaptations of the mantle, foot, and nervous system (Barnes 1987; Huey 2002). For example, a study examining the relationship of shell morphology and habitat selection in 95 species from the class Bivalvia found that variations in flattened or rounded shells as well as shell thickness could be related to environmental pressures like exposure to the elements caused by tidal influence (Stanley 1970). The largest living mollusc is the giant squid (*Architeuthis dux* Steenstrup, 1857) reaching up to 18 meters in length. In contrast, one of the smallest living molluscs is a marine snail (*Ammonicera minortalis* Habe, 1972*)*, at 0.4 millimeters in diameter (Habe, 1972; Roper and Boss, 1982). Although molluscs can be found in marine, freshwater, and terrestrial habitats, the majority of molluscs are found in marine ecosystems from intertidal regions to the deep ocean, including deep-sea thermal vents (Bunje 2003a).

The largest class by number within the phylum Mollusca is the class Gastropoda with over 35,000 existing species, making up about 80% of all living molluscs (Barnes 1987; Bunje 2003b). The order Neogastropoda is comprised of more than 40 families with at least 16,000 identified species of predatory marine

snails (Cunha et al. 2009; WoRMS Editorial Board 2020). Some families within the order Neogastropoda exhibit direct development, which is comparatively rare among marine snails and is predicted to limit opportunities for dispersal (Bohonak 2004; Cunha et al. 2009; Sanford and Kelly 2011; WoRMS Editorial Board 2020). However, evolutionary advantages of direct development include acclimation to local environment, reduced levels of larval mortality, and regularity of recruitment to the local population (Strathmann 1990; Wray and Raff 1991; Teske et al. 2007).

The family Buccinidae, within the order Neogastropoda, represents “true whelks”, and is characterized by moderate to large shell size, predatory behavior, and eggcase production as these species are direct developers (Magalhaes 1948; Walker et al. 2008; Cunha et al. 2009; Power et al. 2009; Vaux et al. 2017). Whelk in the Buccinidae family primarily eat molluscs in the class Bivalvia including mussels, oysters, clams and occasionally other gastropods and conspecifics (Magalhaes 1948). The movement of Buccinidae whelk has been recorded in various studies by tracking tagged individuals in the field and measuring the direction and distance traveled (Magalhaes 1948; Shalack 2007; Edmundson 2016). The activity of Buccinidae whelk can vary greatly, with movement ranging from periods of inactivity to several days of activity expressed by crawling on the benthic floor (Magalhaes 1948; Shalack 2007). Buccinidae whelk have been observed burying during low tide, potentially reducing their activity to avoid exposure to high temperatures or predation (Shalack 2007; Edmundson 2016).

Some whelk in the family Buccinidae are known to exhibit multiple paternity, with females having the capability to store viable sperm for up to six months, with use in successive fertilization events occurring as near-random draws from a blended pool of gametes (Walker et al. 2007; Brante et al. 2011). Polyandry, when females mate with multiple partners within a reproductive period, is a phenomenon seen throughout many taxonomic groups in the animal kingdom (Jennions and Petrie 2000; Jones et al. 2010; Avise et al. 2011). Polyandry is thought to maximize a females reproductive potential during a reproductive period, with only one to few mating events required (Jennions and Petrie 2000). Even though there is a high energetic cost associated with mating events, there are various evolutionary advantages associated with polyandry including fertilization assurance, higher offspring diversity, and genetic bet-hedging (Jennions and Petrie 2000; Xue et al. 2014).

The genus *Busycotypus* belongs to the family Buccinidae and comprises a single species of whelk (WoRMS Editorial Board 2020), the channeled whelk (*Busycoptypus canaliculatus* Linnaeus, 1758). Channeled whelk can be distinguished from other whelk in the Buccinidae family by their more “globular” shape and narrow siphon (Magalhaes 1948; Walker et al. 2008). The channeled whelk shell is smooth and lacks spines, and the shell is also thinner than *Busycon* whelk (Magalhaes 1948; Shalack 2007; Walker et al. 2008; Fisher and Rudders 2017). Channeled whelk are found along the U.S. Atlantic coast from Cape Cod, MA to Cape Canaveral, FL (Davis and Sisson 1988; Edwards and Harasewych 1988).

The age of channeled whelk has been estimated using sectioned statoliths, a calcium carbonate structure used for equilibrium (Fisher and Rudders 2017). Statoliths gain annual depositional rings that can be counted for estimation of age (Rodhouse and Hatfield 1990; Barroso et al. 2005; Fisher and Rudders 2017). Growth rates for channeled whelk can be determined through measurement of the shell length (SL) from the apex to the end of the siphonal canal, or shell width (SW) the widest portion of the shell (Peemoeller and Stevens 2013; Fisher and Rudders 2017). Sexual dimorphism is present in channeled whelk with males tending to have slower growth rates and lower maximum size than females (Walker et al. 2008; Peemoeller and Stevens 2013; Fisher and Rudders 2017).The age and growth of channeled whelk varies depending upon sex and location (Walker et al. 2008; Fisher and Rudders 2017; Nelson et al. 2018). In New England, channeled whelk mature between ages 7 to 8.5 years, with the maximum recorded age estimated at 14 years (Peemoeller and Stevens 2013). In the mid-Atlantic region, channeled whelk mature between ages 5 to 7 years on average, with the maximum recorded age estimated at 16 years (Fisher and Rudders 2017). Female channeled whelk from the mid-Atlantic region have a mean SL of 146.43 mm and males are 131.73 mm SL. Size at maturity has not been assessed for channeled whelk in the South Atlantic; however female channeled whelk from Georgia have a mean SL of 105.8 mm and males are 96.10 mm SL, with a maximum SL of 175 mm for channeled whelk from Georgia (Walker et al. 2008).

Like other whelk in the family Buccinidae, channeled whelk lack a pelagic larval stage and instead develop directly into a free crawling benthic form (Edwards and Harasewych 1988). Channeled whelk exhibit internal fertilization with females producing eggcases with multiple capsules that contain whelk embryos, with upwards of 50 embryos contained within a capsule (Fisher and Rudders 2017). Eggcase deposition occurs in intertidal to subtidal regions where females anchor eggcases to hard substrates (Rawlings 2007; Walker et al. 2008; Harding 2012). Catch data suggests that during the spring channeled whelk migrate to the intertidal region in search of food, but the reverse direction (intertidal to subtidal) for mating and egg case deposition during fall spawning periods (Walker et al. 2008). The following spring channeled whelk juveniles hatch from egg cases (Walker et al. 2008; Harding 2012).

The commercial whelk fishery, also known as a conch fishery, is found throughout the species' range. The commercial fishery in the mid-Atlantic involves the use of baited pots and is primarily a supplemental winter fishery for mid-Atlantic watermen (Davis and Sisson, 1988; Fisher 2015). Since the 1970s channeled whelk have provided economic benefits to directed fisheries in the mid-Atlantic region (Davis and Sisson, 1988). Beginning in the mid 1980's an unregulated commercial fishery developed in New England and by the 1990s, landings from post-production processed whelk meat had reached 1.4 million pounds at $1.80 per pound, generating approximately $2.5 M in ex-vessel revenue (Fisher 2015). Between the 1990s and 2010s shellstock prices fluctuated, however at the end of the 2010s shellstock prices were three times

what they were in the 1990s (Fisher 2015). In 2016, Virginia's whelk industry estimated landings at 1.2 million pounds of shellstock valued at $2.8 million dollars (Robins 2018). The channeled whelk fishery has developed into an important source of diversity and income for mid-Atlantic commercial fishermen.

The channeled whelk resource is managed state by state in the mid-Atlantic region, with minimum landing sizes (MLS) varying by state. The current MLS is 127 millimeters (mm) in New Jersey (New Jersey Division of Fish and Wildlife 2020), 152.4 mm in Delaware and Maryland (Deleware Division of Fish and Wildlife 2020; Maryland Department of Natural Resources 2020), and 139.7 mm in Virginia (Virginia Marine Resource Commision 2020). In the mid-Atlantic region, males reach sexual maturity between 121 to 134 millimeters shell length (mm SL) as compared between 149 to 159 mm SL in females (Fisher and Rudders 2017). An assessment conducted in the mid-Atlantic region found that females harvested at the MLS of 139.7 mm SL in Virginia had a low probability (between 1 to 15%) of being sexually mature, making them extremely vulnerable to overexploitation (Fisher and Rudders, 2017). Based on shell width, another measurement used for determining MLS, an age at MLS of 69.9 mm shell width (SW) in Buzzards Bay, Massachusetts was calculated at 7.5 years for males and 6.3 years for females (Peemoeller and Stevens, 2013). At this MLS, females were estimated to be entering the fishery approximately 2 years before they are sexually mature (Peemoeller and Stevens, 2013). These data from biological assessments highlight a rising concern about the potential for channeled whelk

resources to collapse due to recruitment overfishing, the removal of females before they have a chance to reproduce (Fisher and Rudders, 2017).

Stocks can be defined as groups of individuals with similar population dynamics, genetic composition, and are self-sustained through random mating (Begg et al. 1999; Waples and Gaggiotti 2006). In 2018, the first stock assessment for channeled whelk was performed in Massachusetts. This assessment found that channeled whelk populations in the Nantucket Sound are likely overfished and overfishing is occurring (Nelson et al., 2018). In the early 1990s, Virginia began issuing experimental fishing permits in recognition of knowledge gaps surrounding stock structure and reproductive parameters of channeled whelk in the mid-Atlantic region (Virginia Marine Resource Commission, 2017). Despite these efforts, the amount of available data on the biology of channeled whelk for use in stock assessments remains limited. Current regulations for the channeled whelk commercial fishery in Virginia are based on the biology of knobbed whelk (*Busycon carica* Gmelin, 1791), a closely related species in the Buccinidae family that occupies the same geographic range as channeled whelk (Davis and Sisson 1988; Fisher 2015, Fisher and Rudders 2017).

Genetic studies can be useful for further understanding stock status by addressing questions related to spatiotemporal patterns of connectivity, genetic diversity, and realized dispersal (Baird et al. 2008; Helyar et al. 2011; Seeb et al. 2011; Bourret et al. 2013; Ackiss et al. 2018; Jeffery et al. 2018; Mamoozadeh et al. 2019). In recent years, there has been an increase in the number of published

papers utilizing low-cost high-throughput DNA sequencing technologies like RADseq (Baird et al. 2008) and DArTseq (Kilian et al. 2012), which have allowed for the characterization of thousands of genetic markers called single nucleotide polymorphisms (SNP) from multiple individuals across a target genome (Helyar et al. 2011; Leaché and Oaks 2017; Xu et al. 2017) at relatively low cost. A SNP is the result of a mutation or substitution of a single nucleotide at a specific genetic locus, and can vary among individuals, species, or populations (Krebs et al. 2011). SNPs are highly abundant and distributed throughout the genome, making them a valuable tool for identifying genetic differences among samples. For example, SNPs are estimated to occur every 5 base pairs in the direct developing marine snail *Littorina saxatilis*, every 40 base pairs in marine bivalve *Magallana gigas*, and every 47 base pairs in the marine bivalve *Ostrea edulis* (Sauvage et al. 2007; Galindo et al. 2010; Harrang et al. 2013; Leaché and Oaks 2017; Xu et al. 2017). Studies that have utilized SNP loci have demonstrated the versatility in these markers for answering questions related to population genetic structure (Baird et al. 2008; Helyar et al. 2011; Seeb et al. 2011; Bourret et al. 2013; Ackiss et al. 2018; Jeffery et al. 2018; Mamoozadeh et al. 2019).

Other genetic markers, including nuclear microsatellites, have also been widely used in population genetic studies (Pritchard et al. 2000; Kamel et al. 2014; McDowell and Brightman 2018; Underwood and Darden 2019). Microsatellite markers are simple tandem repeats of nucleotides (e.g. ACACACAC) and occur throughout the nuclear genome of most taxa (Selkoe and Toonen 2006). While microsatellites are useful and appropriate for

answering questions about contemporary gene flow or patterns of divergence over thousands of generations, they may not be well suited for range-wide assessments for marine molluscs that exhibit direct development, like channeled whelk, due to the high incidence of null alleles associated with the high mutation rates found in molluscan taxa (McInerney et al. 2011; Sanford and Kelly 2011; Mariani et al. 2012; Haymer 2015; Teske et al. 2015; Vieira et al. 2016; Dohner et al. 2018; Fleming et al. 2018; Underwood and Darden 2019). A null allele occurs when an allele fails to amplify consistently at a specific locus due to a mutation in the primer binding site (Dakin and Avise 2004). However, microsatellite markers have been the primary marker of choice in paternity studies because loci are highly polymorphic with multiple alleles per locus, and thus usually only a few loci are required for accurate paternal identification (Emery et al. 2001; Walker et al. 2007; Brante et al. 2011; Xue et al. 2014; Ylitalo et al. 2018).

Increased understanding of dispersal and gene flow in channeled whelk populations will provide information that can be used to address the impacts of current fisheries management actions. A baseline measure of genetic diversity can be used to understand how evolutionary processes determine species population genetic structure and be used in the quantification of realized dispersal. Additionally, it is crucial to understand what level of polyandry is occurring for channeled whelk and how multiple paternity may be providing an opportunity for increased genetic diversity and a reduction in inbreeding for a species with direct development and limited dispersal. The purpose of this thesis

was to use SNPs to investigate the population genetic structure of channeled whelk sampled from across the species' range with specific focus on delineating structure and estimating the limits of dispersal within the mid-Atlantic region. The following null hypothesis was tested: *Channeled whelk sampled along the US Atlantic coast comprise a single panmictic population.* Second, microsatellite markers were used to assess the presence of multiple paternity in channeled whelk by testing the following null hypothesis: *There is no evidence of multiple paternity in channeled whelk.*

MATERIAL AND METHODS

Sample Collection and DNA Isolation

Tissue sample collection was focused on the mid-Atlantic region; however, to put the mid-Atlantic region into context, sample collection also took place throughout the entire geographic range of channeled whelk (Figure 1, Table 1). Samples were collected from 10 resource areas (commercial fishing regions) with a goal of acquiring approximately 30 samples per location. Collection locations included: Buzzard Bay, Massachusetts in 2019 (MA, n = 30); Rhode Island in 2018 (RI, n = 12); Ocean City, Maryland in 2018 (OC, n = 50); Chincoteague, Virginia in 2018 (CT, n = 41); Hog Island, Virginia in 2018 (ES, n = 50); Light Tower, Virginia in 2018 (LT, n = 25); Sand Bridge, Virginia in 2018 (VB, n = 50); Pamlico Sound, North Carolina in 2018 (NCPS, n = 30); Wilmington, North Carolina in 2019 (NCW, n = 30); and Charleston, South Carolina from 2015 to 2018 (SC, n = 30) (Figure 1, Table 1). Channeled whelk were obtained through dredging or baited pots with the help of commercial fishers, South Carolina Department of Natural Resources, and Massachusetts Division of Marine Fisheries. For all sampled individuals, shell length and shell width were measured in millimeters, individuals were sexed, and a small piece of foot muscle was placed into 95% ethanol and stored in a freezer at - 20 °C until DNA extraction. Metadata for samples can be found in the Appendix (S. 1).

Channeled whelk egg cases were collected from Buckroe Beach, Virginia during 2013 and 2014 (Figure 1, Table 2). Egg cases were visually located during low tide and collected by hand. In lieu of sampling every embryo within an egg case string (>1000), genetic samples (n=280) were taken from four individual egg cases. The following process was completed for each eggcase: The number of unopened capsules per eggcase was assessed, and capsules were organized into 7 consecutive sections. Within each section, a random capsule was selected using a random number generator. From each randomly selected capsule, 10 embryos were selected. The sampled channeled whelk embryos were removed from the eggcase and rinsed with deionized water before being placed into a vial and stored in a freezer at - 20 °C until DNA extraction.

Total genomic DNA was isolated from 348 channeled whelk foot tissue samples and 280 channeled whelk embryos using the Macherey-Nagel NucleoSpin Tissue DNA Extraction Kit according to the manufacturer's protocol (Machery Nagle, Düren, Germany). Isolated DNA was visualized using a UV light after being size separated on a 1% agarose gel that included GelRed® Nucleic Acid Gel Stain (Biotium, Freemont, California, United States) following standard protocols. Briefly, DNA was loaded into the gel and separated at 120V for 60 minutes. A 1 kb plus size standard (Sigma-Aldrich, St. Louis, Missouri, United States) was used to infer the size of DNA isolated in kilobases. The purity and concentration (µg/ml) of isolated DNA was assessed using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, United States) to quantify the molecular weight of the DNA recovered from isolation. The

purity of the DNA was assessed using the ratio of absorbance at 260 nm to the absorbance at 280 nm, with good quality DNA falling within a 260/280 ratio of 1.7 – 2.0.

Genotyping by Sequencing for Assessment of Population Structure

A genotyping-by-sequencing method, DArTseq (Kilian et al. 2012), was completed at the Diversity Arrays Technology facility in Canberra, Australia (DArT PL). This process is a next generation sequencing (NGS) approach that uses a restriction enzyme digestion step to reduce genome complexity to target low copy sequences with high throughput sequencing allowing for the identification of thousands of SNPs (DArT PL, Canberra, Australia). DArTseq ensures high call rates and reproducibility, with reliability in calling heterozygotes (DArT PL, Canberra, Australia). A total of 252 samples from 10 resource areas along the US Atlantic coast was sent to Diversity Arrays Technology; MA (n = 21), RI (n = 12), OC (n = 34), CT (n = 25), ES (n = 16), LT (n = 23), VB (n = 34), NCPS (n = 30), NCW (n = 30), SC (n = 27) (Table 1). Metadata for samples sent to DArT can be found in the Appendix (S. 2). Only those samples with high molecular weight DNA (20µl of an aqueous solution of DNA at 50 – 100 ng per µl) were used to perform next generation sequencing for the identification of SNP loci.

SNP Filtering

Additional filtering of the DArT dataset aimed at preferentially retaining individuals over loci was performed on the SNP data provided by DArT PL in R version 3.6.1 (R Core Team 2020) using the program package “dartR” (Gruber et al. 2018). SNPs were initially filtered out of the dataset if locus genotype calls and individual genotype calls were missing in ≥ 10% and ≥ 20% of samples respectively. This was done to eliminate SNPs that were low in quality. After filtering individual genotypes, all monomorphic SNPs (loci with only one allele) were removed because they do not provide useful genetic information. The average counts of sequence tags recovered at a locus for an individual (the read depth) across all sequences for all individuals was 9.965. If read depths were greater or lesser than the default settings (< 5 and > 25) for coverage (depth) SNPs were removed. SNPs with low coverage can result in base pair miscall, while SNPs with high coverage can indicate that paralogous loci have been aligned. SNPs that had an average reproducibility of technical replicate assay pairs that fell below 100% were removed. A filter for pairwise Hamming distance, the number of base pair differences between two sequences, was performed to prevent inclusion of sequences that are very similar to one another (Hamming 1950). SNPs were filtered and removed if individual genotype calls were missing in ≥ 5% of samples. Monomorphic loci were again removed. SNPs were filtered and removed if locus genotype calls were missing in ≥ 4% of samples. If minor allele frequencies (MAF), the frequency at which the alternate allele occurs in a given population, were < 1% they were removed to prevent obtaining false

positives in detecting genetic differences due to genotyping error. In cases where there was more than one SNP in the same sequence fragment, known secondary SNPs, one of the SNPs was randomly removed from the dataset. Loci that did not in conform with the expectations of Hardy-Weinberg Equilibrium (HWE) were removed using the program package “radiator” (Gosselin et al. 2020) in R version 3.6.1. The p-value was set at the default (0.0001), and SNPs were removed if they were out of HWE in at least 2 populations (Gosselin et al. 2020). Loci used to assess genetic structure are assumed to be neutral, not subject to evolutionary constraints; however, not all loci are considered neutral. Outlier loci are putatively under selection, which can affect the genetic diversity and ultimately bias estimates of genetic structure (Narum and Hess 2011; Russello et al. 2012; Feng et al. 2015). Outlier loci were detected in R version 3.6.1 using the packages “dartR” and “pcadapt” (Luu et al. 2017). The outlier function in the package “dartR” was used to identify outlier loci per population using the Lotterhos and Whitlock (2014) outflank method, a method that infers genetic differentiation among populations without using spatially heterogenous or homogenous selection (Lotterhos and Whitlock 2014). The outlier function in the package “pcadapt” performed a principal component analysis and computed p-values to test for outlier loci, which is based on the correlation between genetic variation and the first K principal components. The false discovery rate threshold for calculating q-values (adjusted p-values) was 0.05 for both programs.

Population Identification

A principal component analyses (PCA) created in R version 3.6.1 using the program package “adegenet” (Jombart 2008) was used to visualize the genetic similarities and dissimilarities in the data and identify genetic clusters of individuals without using prior genetic information. An unbiased estimator of Wrights F-statistic (F_{ST}) (Wright 1943), a measure of population differentiation, was calculated between pairs of sampled populations to evaluate the presence of population structure using the program package “StAMPP” (Pembleton et al. 2013) in R version 3.6.1. The index of F_{ST} is bounded from 0 to 1, with 0 implying panmixia (lack of detectable population structure) and 1 implying maximal differentiation. The significance of F_{ST} values was calculated using 10,000 bootstrap iterations of the data with p-values set at 0.01. A discriminant analysis of principal components (DAPC) was completed in R version 3.6.1 using the program package “adegenet” to assess population structure. This method uses PCA analysis to examine the genetic similarities or dissimilarities between individuals to maximize variation between groups, while minimizing variation within groups to assign individuals to identified clusters using successive K-means. The optimal number of genetic clusters was assessed using a successive number of K-means (K= 1 to K = 10), with the optimal number of clusters selected using the Bayesian information criterion (BIC). Analyses of molecular variance (AMOVA) were conducted to test for the presence of hierarchical population structure (Excoffier et al. 1992) with the program GenAlEx 6.5 (Peakall and Smouse 2012). AMOVAs were used to identify how much

genetic differentiation is due to differences among groups (F_{RT}) or among resource areas within groups (F_{SR}) using an infinite allele model. The indices for F_{RT} and F_{SR} are bounded from 0 to 1, with 0 implying no detectable differentiation and 1 implying maximal differentiation. Significance was assessed using 999 permutations of the data. Alternate groupings were based on findings from DAPCs, PCAs, and pairwise F_{ST} values (Table 3a).

Isolation by Distance

Isolation by Distance (IBD) refers to the theory that populations closer in space are more genetically similar than populations that are father apart in the absence of hard barriers to dispersal (Slatkin 2006). Mantel tests were performed to identify isolation by distance using a matrix comparing geographic distance using the closest known coordinates for sampling locations and genetic distance using Wright's F_{ST} (Wright 1943). Mantel tests were conducted for various regions throughout the geographic range of channeled whelk. For each Mantel test, a total of 999 permutations were used to assess significance in R version 3.6.1 using the program package "dartR".

Population Summary Statistics

Summary statistics for the evaluation of genetic diversity including estimations of the inbreeding coefficient (G_{IS}) and observed (Ho), and expected (He) heterozygosity were calculated for genetic groups (populations) using the program GenoDive (Meirmans 2020). The inbreeding coefficient (G_{IS}) measures

the excess of homozygotes in an individual relative to the expectation under Hardy-Weinberg Equilibrium. The index of G_{IS} is generally bounded from 0 to 1, with 0 implying no inbreeding and a 1 implying no heterozygous individuals in a population. Occasionally, a G_{IS} value can be negative implying there are more heterozygous individuals than expected under Hardy-Weinberg Equilibrium. The observed (Ho) heterozygosity is the proportion of heterozygotes observed in the population of interest, while the expected (He) heterozygosity is the proportion of heterozygotes expected under Hardy-Weinberg Equilibrium. Genetic diversity summary statistics were also calculated separately for females and males and significant differences between sexes were evaluated using a paired t-test. Estimations of effective population size (Ne) were calculated for genetic groups (populations) of channeled whelk using the program NeEstimator (Do et al. 2014). Effective population size, the size of an idealized population with similar gene diversity to an actual population under consideration, is predicted to dictate genetic effects like inbreeding and the rate of loss of genetic variation (Wright 1938; J. et al. 1971). The Linkage Disequilibrium model, the correlation among alleles at different loci, with random mating was used for this analysis and critical values were set at 0.05, 0.02, and 0.01.

Development of Microsatellite Markers for Assessment of Multiple Paternity

There are no published microsatellite markers for *Busycotypus canaliculatus* or for any closely related species, so *de novo* markers were developed. Briefly, total genomic DNA was isolated from a single individual

following standard isolation protocols as above. The genomic DNA was randomly sheared and sequenced on a MiSeq System using the Miseq Reagent Kit V.3 (Illumina, San Diego, CA) at the Virginia Institute of Marine Science. The resulting sequences were transformed into the FASTQ format using the FASTQ groomer (v 1.0.4) in the Galaxy platform (Blankenberg 2010) and sequence quality was checked using the FASTQC software (Babraham Bionformatics, Cambridge UK). Sequences were filtered to remove low-quality sequences using the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). After processing, sequences were converted to the FASTA format and exported and screened for the presence of microsatellites using the MSATCOMMANDER 1.0.8 software (Faircloth 2008) and primers were designed using PRIMER3 (Misener et al. 2003). Primers were designed for a subset of the identified loci with 18 to 30 bp in primer length for specificity, 40 to 60% GC content, and a melting point (Tm) between 55 and 80 °C.

Primer amplification was tested on a subset of 24 adult channeled whelk sampled from MA (n = 8), VB (n = 8), and SC (n = 8). Amplification of each locus was performed with an initial denaturation at 94 °C for 3 minutes, followed by 35 cycles at 94 °C for 1 minute, a primer specific annealing temperature (Ta) (Table 8) for 1 minute, an extension at 72 °C for 1 minute, and a final extension step at 72 °C for 7 minutes. The resulting PCR products were labeled with a T3-labeled fluorescent probe (FAM, VIC, NED, or PET) and were separated on an ABI 3500*xl* Prism Genetic Analyzer (Applied Biosystems, Foster City, California, United States). All reactions included an internal GeneScan 500-Liz size

standard (Applied Biosystems, Foster City, California, United States). The chromatic peaks for each microsatellite locus were sized using the GeneMarker AFLP/Genotyping Software, ver. 1.75 (SoftGenetics, State College Pennsylvania, United States). The MICROCHECKER 2.2.3 (Van Oosterhout et al. 2004) software was used to check for the presence of null alleles and evidence of scoring errors using 10,000 permutations of the data and 95% confidence intervals. The GENEPOP'007 software package (Rousset 2008) was used to test for conformance to the expectations of HWE, linkage disequilibrium among loci and to calculate summary statistics including: number of alleles, and observed (Ho) and expected (He) heterozygosities (Guo and Thompson 2006). Primers that consistently amplified across the subset of samples were then selected for use in the analysis of multiple paternity.

Parentage Analysis

Parentage was reconstructed based the genotypes of offspring collected from eggcases using COLONY 2.0 (Wang 2004). This method uses the genotypes of full- or half-sibling groups to reconstruct parental genotypes under the assumption that offspring in a group share a common parent. Paternal genotype reconstruction consisted of identifying the genotype of the shared parent (in this case, the mother) and then determining the minimum number of additional parents that contributed alleles to offspring in each eggcase. The relationship of alleles across loci for the offspring determined the genotypes for the unknown parents (Sefc and Koblmüller 2009; Jones et al. 2010). Maximum

likelihood estimates for the paternal genotypes were calculated from simulations assuming Mendelian segregation to identify sibling relationships and infer parental genotypes. Using known biological characteristics of Buccinidae whelk, the species parameter was set as dioecious diploid and inbreeding was assumed. The inferred mating system for analysis was female polygamy and male monogamy under the assumption that female Buccinidae whelk mate with multiple males.

RESULTS

Genotyping by Sequencing and SNP Filtering

Of the 348 channeled whelk DNA extractions, 252 had a sufficient amount of high molecular weight DNA (20µl of an aqueous solution of DNA at 50 – 100 ng per µl) to perform DArTseq (Table 1). DArTseq yielded 27,344 SNPs that passed standard Diversity Array Technology quality filtering procedures. A total of 7,247 SNP loci were initially removed because locus genotype calls and individual genotype calls were missing in ≥ 10% and ≥ 20% of samples respectively. A further 1,597 SNP loci were considered monomorphic and were removed from the dataset. A total of 2,196 SNPs were removed for having average read depth < 5 and > 25, with another 7,791 SNPs removed for having average reproducibility below 100%. Additionally, 1,223 SNPs were removed after filtering by Hamming distance. No SNP loci were removed after filtering by individual genotype. A total of 493 SNP loci were monomorphic and removed. After filtering for locus genotype calls missing in ≥ 4% of samples, 329 SNPs were removed. Additionally, 3,896 SNP loci were removed because they had minor allele frequencies that were < 1%. No SNPs were found to be secondary SNPs. Two SNP loci were not in conformance with the expectations of HWE and were removed. A total of 227 outlier loci were detected using the package “pcadapt” and zero outlier loci were detected using the package “dartR”. The remaining loci were grouped into a dataset consisting of 2343 neutral loci. The

227 outlier loci detected with the package “pcadapt” were grouped into an outlier loci data set, and together with the main dataset comprised 2,570 SNPs from 227 individuals.

Population Identification

Patterns of genetic differentiation in PCAs were consistent among all (2570 SNPs), neutral (2343 SNPs), and outlier (227 SNPs) loci; therefore, the results for the all loci dataset are presented in the remainder of analyses. Metadata for PCAs can be found in the Appendix (S. 3). When comparing the 10 sampled resource areas, the first principal component explained 22.10% of the variation and the second principal component explained 4.02% of the variation (Figure 2). The PCA indicated that the southern-most resource areas (NCPS, NCW, SC) were separated from one another and from the remaining resource areas (Figure 2). The high level of separation among sample collections from the southern-most resource areas resulted in tight clustering of the remaining seven resource areas in the PCA. To better visualize the data, the three southern-most resource areas were excluded, and a second PCA of the 7 resource areas located north of Pamlico, Sound, NC (MA, RI, OC CT. ES, LT, VB) was performed (Figure 3). The first principal component explained 2.24% of the variation and the second principal component explained 1.66% of the variation (Figure 3). The visualization of the genetic relationship among these resource areas resulted in a visual ‘funneling effect’ with the widest spread of individuals occurring in the southern most resource area retained (VB) and a gradual

decrease in spread occurring northward (LT, ES, CT, OC, RI, MA) (Figure 3). This funneling effect was also present when visualizing data from the mid-Atlantic resource areas (OC, CT, ES, LT, VB), with 2.30% and 1.87% of the data explained by the first and second principal components respectively (Figure 4).

Estimates of genetic divergence, F_{ST}, among the 10 sampling locations ranged from 0.0001 to 0.465, with the highest level of divergence (F_{ST} = 0.465, $p < 0.001$) observed between OC and SC and the lowest level of divergence (F_{ST} = 0.0001, p = 0.487) observed between CT and ES (Table 4). F_{ST} values were highest (0.125 – 0.465, $p < 0.001$) when comparing samples from the southern-most resource areas; NCPS, NCW, and SC to samples from all other resource areas (Table 4). Comparisons of F_{ST} values for MA to samples from all other resource areas north of Cape Hatteras ranged from 0.081 – 0.128 ($p < 0.001$), which were an order of magnitude higher than F_{ST} values seen when samples in the mid-Atlantic (OC, CT, ES, LT, VB) were compared to MA and RI samples (0.031 – 0.047, $p < 0.001$) (Table 4). All but four pairwise comparisons of F_{ST} values, all within the mid-Atlantic region (OC vs. CT, OC vs. ES, CT vs. ES, and LT vs. ES), were significant ($p < 0.001$). Within the mid-Atlantic region, F_{ST} values were elevated for pairwise comparisons involving VB (0.001 – 0.016, $p < 0.001$) as compared other mid-Atlantic comparisons (0.0001 – 0.007, $p \leq 0.001$ - 0.487) (Table 4). Significant differences were also observed between LT and CT (0.005, p = 0.005) and LT and OC (0.007, $p < 0.001$) (Table 4).

A DAPC plot was created to visualize the posterior membership probability for the 227 channeled whelk sampled and cluster to individuals into genetically

distinct groups. This broad-scale visualization of population structure revealed 4 distinct genetic groups; a group of the north and mid-Atlantic individuals (MA, RI, OC, CT, ES, LT, VB), a group consisting of individuals from NCPS, a group of individuals from NCW, and a group of individuals from SC (Figure 5). A single individual sampled from NCW was assigned to the group comprised of individuals from the north and mid-Atlantic resource areas (Figure 5). To better visualize population structure in the north and mid-Atlantic, a second DAPC plot was created for the 148 channeled whelk sampled from the 7 northern-most resource areas (MA – VB). The DAPC revealed 3 distinct genetic groups: a group of individuals from MA and RI, a group of individuals from OC, CT, ES, and LT, and a group of individuals from VB (Figure 6). However, there was some overlap in population assignment with 12 individuals mixing equally between cluster 2 (VB) and cluster 3 (OC, CT, ES, LT) (Figure 6). A fine-scale DAPC was conducted to examine the population structure of 117 channeled whelk in the 5 mid-Atlantic resource areas. The DAPC revealed 2 distinct genetic groups a group of individuals from OC, CT, ES, LT, and a group consisting of individuals from VB (Figure 7). The overlap of the same 12 individuals was observed in population assignment between cluster 1 and cluster 2 (Figure 7).

The best grouping of individuals was also assessed in an AMOVA framework. Alternative groupings were conducted using all 227 channeled whelk from the 10 resource areas. Groupings were selected based on findings from PCA, DAPC, and pairwise comparisons of F_{ST} (Table 3a). All AMOVAs showed significant divergence among groups (populations) (F_{RT} = 0.279 – 0.326, p =

0.001) but not among resource areas within groups (F_{SR} = -0.024 – -0.014, p = 0.991 – 1.000) (Table 3b). The optimal grouping revealed by analyses was AMOVA 5, separating resource areas into 6 groups which included: 1) MA, 2) RI, 3) mid-Atlantic (OC, CT, ES, LT, VB), 4) NCPS, 5) NCW, and 6) SC (Table 3a), which maximized differentiation among resource areas (F_{RT} = 0.326, p = 0.001) explaining 58.5 % of the variation, while minimizing differentiation among groups of resource areas (F_{SR} = -0.011, p = 1.000) (Table 3b).

Isolation by Distance

A significant pattern of IBD was detected among the 10 resource areas (r = 0.5693, p = 0.009) (Figure 8). Barriers to gene flow can confound tests of IBD, so patterns of IBD were also examined for the 7 northern-most resource areas (MA, RI, OC, CT, ES, LT, VB). A significant pattern of IBD was detected among the 7 northern-most resource areas (r = 0.6111, p = 0.011) (Figure 9). Patterns of IBD were also tested among the 5 mid-Atlantic resource areas (OC, CT, ES, LT, VB). No significant pattern of IBD was detected among samples from the mid-Atlantic resource areas (r = 0.4686, p= 0.075) (Figure 10).

Population Summary Statistics

Population level summary statistics were calculated for each of the 7 genetically distinct groups delineated by the previous analyses. The average values of observed and expected heterozygosity across all 7 groups were 0.098 and 0.097 respectively (Table 5a). Observed and expected heterozygosity (gene

diversity) values were highest at the southern-most (NCPS, NCW, SC) resource areas (Ho = 0.124 – 0.253, He = 0.121 – 0.248) and lowest in the northern-most (MA, RI) resource areas (Ho = 0.011 – 0.013, He = 0.010 – 0.012) (Table 5a). The average inbreeding coefficient for the 7 groups were -0.005, with mid-Atlantic (OC, CT, ES, LT) and SC having the highest inbreeding coefficients (0.026) and RI having the lowest inbreeding coefficient (-0.035) (Table 5a). To evaluate whether there were differences between sexes, population summary statistics were also calculated separately for males (n = 97) and females (n =111). Overall, males and females had different observed heterozygosities (0.107 vs. 0.166) and the same expected heterozygosities (0.108) (Table 5a). Although the inbreeding coefficient was higher in males (-0.023) than females (-0.053), a paired t-test found no significant difference between GIS values (p = 0.257) (Table 5b).

Contemporary estimates of effective population size were calculated for each of the 7 genetically distinct groups delineated by the previous analyses. Overall, Ne estimates increased from north to south. The lowest Ne was found in MA at 32.4 (95% CI = 27.7 – 38.8) and the highest Ne was found in SC at 833.6 (95% CI = 664.2 – 1117.5) (Table 6). The total number of monomorphic loci also decreased (number of polymorphic loci increased) from north to south, with MA having the highest number of monomorphic loci (2373), and NCW having the lowest number of monomorphic loci (284) (Table 6).

Microsatellite Marker Development

MSATCOMMANDER 1.0.8 software (Faircloth 2008) identified 1,048,576 microsatellite loci. In total, 288,939 primer pairs were designed using PRIMER3 (Misener et al. 2003). Of these, primer pairs for 62 loci were selected for testing and optimization on a subset of channeled whelk samples including primers for 54 tetra-nucleotide loci (a 4-nucleotide repeat motif) and 8 di-nucleotide loci (a 2-nucleotide repeat motif). Metadata for microsatellite markers can be found in the Appendix (S. 4). Four of these microsatellite markers amplified consistently across test samples and had multiple alleles, while the remainder (16 markers) were either monomorphic or failed to amplify across the subset of channeled whelk samples from across the geographic range of channeled whelk (Table 7). Analysis using MICROCHECKER 2.2.3 revealed no evidence of scoring error or null alleles across the subset of channeled whelk samples. A total of 280 channeled whelk embryos went through DNA extraction, and 278 of those sample amplified across all four loci. The number of alleles for the four microsatellite loci ranged from 2 at Bcan65 to 7 at Bcan33 (Table 8). Mean expected heterozygosity (gene diversity) across loci ranged from 0.126 for BCAN65 and 0.377 for Bcan33 (Table 8).

Parentage Analysis

A total of 280 channeled whelk embryos sampled from 4 egg cases were used for sibship reconstruction, and a total of 4 microsatellite primers were used for analysis (Table 2, Table 7). Paternity estimates for Eggcase 1 and Eggcase 2

were based on all 4 microsatellite loci (Bcan65, Bcan62, Bcan27, Bcan33) (Table 7). Only 3 microsatellite loci were used for analysis of Eggcase 3 and Eggcase 4 because the Bcan65 microsatellite locus was monomorphic. Evidence of multiple paternity was found through the inferred paternal genotypes derived from sibship reconstruction. The average number of sires (inferred paternal genotypes) per eggcase was 9, with a range from 6 – 11 (Table 9).

DISCUSSION

Summary of Findings

This study was the first to assess the population genetic structure of channeled whelk, *Busycotypus canaliculatus*, throughout a significant portion of its geographic range from Massachusetts to South Carolina. A total of 227 channeled whelk samples from 10 resource areas along the US Atlantic coast was used to estimate levels of genetic diversity and identify the presence of population genetic structure based on analyses of 2570 SNP loci. Metadata including the R script used for analyses can be found in the Appendix (S. 5). There was significant population genetic structure among the 10 resource areas sampled in this study, with results from PCA, DAPC, and AMOVA analyses supporting the presence of seven genetically distinct populations: 1) MA, 2) RI, 3) mid-Atlantic 1 (OC, CT, ES, LT), 4) mid-Atlantic 2 (VB), 5) NCPS, 6) NCW, and 7) SC. This study also used channeled whelk egg cases to look for evidence of multiple paternity in this species. A total of 276 channeled whelk embryos samples from 4 eggcases was used to assess the incidence of multiple paternity through the use of sibship reconstruction analysis based on 4 microsatellite loci. The average number of inferred paternal genotypes was 9, supporting the presence of multiple paternity in channeled whelk.

Barriers to Dispersal

The complexity of the marine environment can create multiple types of barriers that limit dispersal including: geographic barriers, ocean currents, environmental limitations, and patchiness between suitable habitat (Treml et al. 2015). The dispersal potential of direct developing species are known to be impacted by many of these types of dispersal barriers including biogeographic features, physical features, and environmental factors (Weersing and Toonen 2009). Additionally, a negative correlation between dispersal ability and genetic divergence has been observed, where species with a life history strategy of reduced dispersal show a larger number of private alleles (alleles sequestered in a single population) and higher genetic distance among geographic samples (Castric and Bernatchez 2003; Bohonak 2004). The pairwise F_{ST} values among sampling locations in this study spanned several orders of magnitude (0.0001 – 0.465), indicating widely varying levels of divergence among sampling areas. Pairwise F_{ST} values were highest in comparisons involving SC to all other resource areas (0.280 – 0.465), which are elevated as compared to values commonly observed in studies of marine species (typically 0.01-0.02) (Waples 1987); however, similarly high levels of genetic divergence have been found in other direct developing marine invertebrates (F_{ST} = 0.244 – 0.479) (Cahill et al. 2017).

The highest levels of genetic divergence among channeled whelk sampled from different resource areas support the presence of significant barriers to gene flow (dispersal) that coincide with well-known biogeographic features. Within the

southern region (NCPS, NCW, SC), F_{ST} values increased as successive biogeographic barriers were crossed, indicating that Cape Hatteras (NCPS vs. 7 northern-most resource areas; average F_{ST} = 0.154) and Cape Fear (SC vs. all other resource area; average F_{ST} = 0.347) represent significant barriers to dispersal for channeled whelk and are driving the highest F_{ST} values observed in this study. Cape Hatteras is a known biogeographic barrier for many other marine species including black sea bass (*Centropristis striata)*, lined seahorse (*Hippocampus erectus)*, and many Molluscan species (Hale 2010; Briggs and Bowen 2012; Mccartney et al. 2013; Boehm et al. 2015; Pappalardo et al. 2015). Biogeographic barriers have also been identified as impediments to gene flow in other marine gastropods. A population genetics study using microsatellite markers for another marine snail with a geographic range along the US Atlantic coast, *Crepidula fornicata*, found evidence for strong genetic divergence toward the southern end of its range near the southern tip of the peninsula in Florida (F_{ST} = 0.067 – 0.099, $p < 0.001$).

While biogeographic features in the marine environment may explain hard barriers to dispersal, physical parameters like ocean currents and temperature may prove to be significant, yet less extreme barriers to gene flow (Treml et al. 2015). Levels of divergence in New England and the mid-Atlantic were an order of magnitude lower than comparisons with samples south of Cape Hatteras, suggesting more subtle barriers to gene flow in these areas. While there are no known biogeographic barriers limiting dispersal within the New England and mid-Atlantic sampling regions, other environmental factors may be reducing gene

flow for channeled whelk. These may include lack of suitable habitat, isolation by currents, or isolation by distance, however increased sampling between RI and OC would be needed to discriminate among these alternatives. Overall, analysis support the presence of four populations within these regions: 1) MA 2) RI 3) mid-Atlantic 1 (OC, CT, ES, LT) and 4) mid-Atlantic 2 (VB), although results differed slightly among the analyses. The DAPC analysis identified three distinct genetic populations; 1) MA and RI, 2) mid-Atlantic 1, and 3) mid-Atlantic 2 while the AMOVA analysis best supported three different genetic clusters; 1) MA, 2) RI, and 3) mid-Atlantic. However, the results of the AMOVA analysis separating mid-Atlantic 1 and mid-Atlantic 2 was also highly significant. Disagreement between clustering analyses were clarified by pairwise F_{ST} comparisons, which showed statistically significant levels of genetic divergence between RI and MI in the New England region, between the New England and mid-Atlantic region, and among samples from the mid-Atlantic region. Significant genetic differences between samples from the New England and mid-Atlantic regions were also found in the direct developing snail, *Crepidula convexa,* with samples from Massachusetts being significantly differentiated from samples from Virginia (F_{ST} = 0.057) (Cahill and Viard 2014).

A significant level of genetic differentiation was also observed between mid-Atlantic 1 and mid-Atlantic 2 (F_{ST} = 0.007 – 0.016), suggesting that environmental factors like temperature or currents may be limiting dispersal in this region. The mouth of the Chesapeake Bay separates these two genetically distinct populations in the mid-Atlantic region, with current patterns between the

Chesapeake Bay and the Atlantic Ocean potentially acting as a subtle barrier to dispersal for channeled whelk in that area; however, no significant pattern of IBD was present among mid-Atlantic samples, suggesting that simple straight line distance is not responsible for the observed pattern.

Latitudinal Differences in Diversity

Reduced dispersal ability resulting from direct development can decrease an effective population size, with estimates of effective population size (Ne) inversely related to the rate of loss of genetic diversity in a population with a finite number of individuals that are randomly mating (Eckert et al. 2008). The New England region had the lowest Ne estimates (32.4 – 65.6), with an increase in Ne estimates moving south towards the mid-Atlantic and southern regions (195.5 – 833.6). Inversely, the number of monomorphic loci decreased moving southward, with MA having the highest number of monomorphic loci (2373) and SC having the lowest number of monomorphic loci (439). Populations with higher numbers of monomorphic loci are less genetically diverse, indicating that levels of genetic diversity are higher in the mid-Atlantic and southern regions sampled, which align with what is expected in regions with higher levels of Ne estimates. Expected heterozygosity, another measure of genetic diversity, was also higher in the southern region (0.121 – 0.248) than in the mid-Atlantic (0.020 – 0.030) and New England (0.010 – 0.012) regions sampled, also indicating lower genetic diversity in the northern peripheral regions (MA and RI) of the geographic range of channeled whelk.

The lower levels of expected heterozygosity in northern peripheral populations of channeled whelk align with another direct developing snail *Crepidula convexa*, with a geographic range along the US Atlantic coast. Two northern populations of *Crepidula convexa* exhibited a slight decrease in measures of genetic diversity when compared to populations further south with the effect of reduced diversity in peripheral population being more pronounced than a pelagic larval dispersing congener *C. fornicata* (Cahill and Viard 2014). The central-peripheral population hypothesis predicts that peripheral populations will have reduced genetic variability when compared to central populations (Mayr 1963). Various studies have supported the central-peripheral hypothesis by revealing a decline in genetic diversity in populations at the end of species geographic range (Eckert et al. 2008; Ackiss et al. 2018). The pattern observed in this study conforms with what is expected by the central-peripheral population hypothesis, revealing that the northern peripheral populations (New England region) of channeled whelk have lower genetic diversity than the central and southern populations (Mid-Atlantic 1, Mid-Atlantic 2, NCPS, NCW, and SC).

Genetic Consequences of Direct Development

Species that exhibit direct development and reduced dispersal ability with localized populations tend to show a high level of genetic differentiation, reduced genetic diversity, and high levels of inbreeding (Wright 1943; Martel and Chia 1991; Palumbi 2003; Sanford and Kelly 2011; Mariani et al. 2012; Underwood and Darden 2019). The high levels of population differentiation found for

channeled whelk throughout the geographic range align with what has been observed in other direct-developing gastropods. Two marine snails, *Littorina sp*., with different dispersal modes showed varying levels of population genetic structure throughout their geographic range from northern California to Alaska with a moderate amount of population structure (F_{ST} = 0.063 – 0.320, p = < 0.05) for the direct developing species (*L. subrotundata*) as compared to a lack of population structure (F_{ST} = -0.003 – 0.006, p > 0.05) in the species with pelagic larvae (*L. scultulata*) (Kyle and Boulding 2000). In addition, many phyla (Enchinodermata, Cnidaria, Arthropoda, Bryozoa, Mollusca) exhibit higher levels of divergence in species that exhibit direct development (F_{ST} = 0.244 – 0.479) as compared to the lower levels of divergence seen in species with pelagic larvae (F_{ST} = 0.003 – 0.019) (Cahill et al. 2017). Divergence values seen in these phyla are similar to the divergence values recovered in this study and are consistent with reduced dispersal as a result of direct development.

A reduction in genetic diversity is characteristic of species that exhibit direct development. The sea slug *Costasiella ocellifera* is known to exhibit both pelagic larva and direct development. Populations of *C. ocellifera* that exhibit direct development have lower genetic diversity (He = 0.255 – 0.704) compared to populations of *C. ocellifera* that have pelagic larva (He = 0.661 – 0.762) (Ellingson and Krug 2016). Low levels of genetic diversity were found in channeled whelk populations (He = 0.010 – 0.248), which align with what is observed in other studies. While increased levels of inbreeding, caused by a reduction in genetic diversity, can be a consequence of direct development, high

levels of inbreeding were not observed in channeled whelk. The mean G_{IS} value (inbreeding value) for channeled whelk across all resource areas sampled was lower than estimates of G_{IS} values found in other direct developing molluscs (-0.005 vs. 0.184) (Addison and Hart 2005). Low levels of inbreeding have also been observed in *Crepidula convexa* (G_{IS} = - 0.220), a direct developing marine snail that exhibits multiple paternity (Le Cam et al. 2014; Cahill and Levinton 2016). The lower inbreeding levels seen in channeled whelk may be a result of the high number of males that are mating with an individual female, suggesting that polyandry may explain the maintenance of genetic diversity in channeled whelk populations.

A Relationship Between Polyandry and Genetic Diversity

The high levels of inbreeding and low levels of genetic diversity seen in populations of other direct developing gastropods were not observed in channeled whelk populations. This may be due to polyandry, a very common phenomenon in marine gastropods that can provide a buffer against high levels of inbreeding in small localized populations (Walker et al. 2007; Brante et al. 2011; Xue et al. 2014). Polyandry has been reported to increase genetic diversity within a population and reduce the likelihood of inbreeding through genetic bet-hedging (Mäkinen et al. 2007; Pearse and Anderson 2009). In the case of polyandry, genetic bet hedging involves females choosing multiple sires to provide a reservoir of genetic variability for their offspring (Watson 1991; Pearse and Anderson 2009). Multiple paternity has also been found in other direct

developing whelk in the family Buccinidae, including the knobbed whelk (*Busycon carica*) (Walker et al. 2007). However, higher levels of multiple paternity were found in channeled whelk when compared to the knobbed whelk (9 vs. 4 inferred fathers) (Walker et al. 2007). The high levels of multiple paternity observed in channeled whelk could be a biological mechanism that increases as fishing pressure and overexploitation of the resource increases, with the result of low inbreeding levels and overall maintenance of genetic diversity in localized populations of channeled whelk.

Pitfalls with Primers

Microsatellites were the original marker of choice for this study due to their utility both for assessing population genetic structure and multiple paternity; however, single nucleotide polymorphisms were ultimately chosen. Initially, there were challenges obtaining high quality DNA due to the polysaccharides and mucus found in the channeled whelk tissue samples. This issue is fairly common and has been reported in numerous published studies (Jaksch et al. 2016; Arseneau et al. 2017; Schultzhaus et al. 2019; Gomes-dos-Santos et al. 2020). Additionally, only 4 of 62 microsatellite markers amplified consistently across samples from several resource areas and had multiple alleles, while the remainder were either monomorphic or failed to amplify across resource areas. Consistent amplification failure is most likely due to the high levels of genetic divergence seen across the species range (Cahill and Viard 2014). However, because only a few markers (3 to 5) are needed for parentage analysis, the 4

microsatellite loci that successfully amplified were used for assessing the presence of multiple paternity in channeled whelk (Walker et al. 2002, 2007; Weetman et al. 2005; Brante et al. 2011; Ylitalo et al. 2018).

Fisheries Management Recommendations

Long-term sustainability of the channeled whelk commercial fishery may be threatened due to overharvest and removal of whelk before they are sexually mature. Management measures are inconsistent, with adjacent states having different harvest regulations (e.g. North Carolina). This study found 7 genetically distinct populations along the US Atlantic coast and inferred multiple drivers of population structure including: biogeographic barriers, current barriers, and possible lack of suitable habitat. The substantial stock structure present throughout the geographic range of channeled whelk underscores the need for new management strategies for the channeled whelk resource; however, a more comprehensive look at this resource is needed to identify population genetic structure on a finer scale.

Multiple genetic stocks were isolated within a state, and shared genetic stock was identified among states. In the mid-Atlantic region, Virginia appears to have two genetic stocks, one comprised of the Chincoteague, Eastern Shore, and Light Tower resources, and one surrounding Virginia Beach. The Chincoteague, Eastern Shore and Light Tower resources are a part of a larger genetic stock including Ocean City, MD. The finding of significant differences between sampling locations within Virginia (and North Carolina) and a lack of

genetic differentiation between samples from Virginia and Maryland suggests that state-by-state management might not be suitable for this resource. Regional management could be a possible change for this fishery; however, more information on migration patterns and what biophysical properties affect dispersal for channeled whelk will be needed before making a decision to better assess the connectivity of various genetic stock of channeled whelk throughout its range.

The addition of more sampling locations throughout the geographic range of channeled whelk and collection of environmental data for use in model development can be used to better understand the movement patterns and limits of dispersal for this species. Sampling effort should focus on the gaps between MA and RI and between RI and OC, which could provide a more fine-scale delineation of the population structure within the New England region and between New England and the mid-Atlantic regions. While this study was able to obtain samples from inshore assemblages, additional sampling should also include a genetic comparison among offshore assemblages because the channeled whelk resource is targeted there as well. Based on the results of this study, it is recommended that a channeled whelk working group including representatives of the channeled whelk resource (scientists, commercial fishers, and management agencies) from Massachusetts to north of Cape Hatteras be formed to further assess current regulations of the channeled whelk resource and how state-by-state management may or may not align with findings from this study.

TABLES

Table 1. Locations, population names, collection year, and number of channeled whelk sampled collected, as well as the number of samples selected for DArTseq, and the number of samples retained after quality filtering.

Location	Population Name	Collection Year	No. Samples Collected	No. Samples Selected for DArTseq	No. Samples Retained after Filtering
Buzzard Bay, MA	Massachusetts (MA)	2019	30	21	19
Rhode Island	Rhode Island (RI)	2018	12	12	12
Ocean City, MD	Ocean City (OC)	2018	50	34	34
Chincoteague, VA	Chincoteague (CT)	2018	41	25	17
Hog Island, VA	Eastern Shore (ES)	2018	50	16	11
Light Tower, VA	Light Tower (LT)	2018	25	23	21
Sand Bridge, VA	Virginia Beach (VB)	2018	50	34	34
Pamlico Sound, NC	North Carolina Pam (NCPS)	2018	30	30	28
Wilmington, NC	North Carolina Wilm (NCW)	2019	30	30	28
Charleston, SC	South Carolina (SC)	2015 - 2018	30	27	23

Table 2. Eggcases, collection date, number of unopened capsules for each eggcase, number of capsules selected, and number of channeled whelk embryos sampled from each eggcase.

Population	Collection Date	No. of Capsules	No. Of Capsules Selected	No. of Samples Selected from Capsules
Eggcase 1	March 2013	96	7	70
Eggcase 2	March 2013	76	7	70
Eggcase 3	February 2014	69	7	70
Eggcase 4	November 2014	93	7	70

Table 3a. Alternate hierarchical groupings of channeled whelk resource areas tested using Analysis of Molecular Variance (AMOVA). Resource area abbreviations are as follows: Buzzard Bay, Massachusetts (MA), Rhode Island (RI), Ocean City, Maryland (OC), Chincoteague, Virginia (CT), Eastern Shore, Virginia (ES), Light Tower, Virginia (LT), Virginia Beach, Virginia (VB), Pamlico Sound, North Carolina (NCPS), Wilmington, NC (NCW), Charleston, South Carolina (SC).

Resource Area	AMOVA 1	AMOVA 2	AMOVA 3	AMOVA 4	AMOVA 5
MA	1	1	1	1	1
RI	2	2	2	2	2
OC	3	3	3	3	3
CT	4	4	4	3	3
ES	4	5	4	3	3
LT	4	5	5	3	3
VB	5	6	6	4	3
NCPS	6	7	7	5	4
NCW	7	8	8	6	5
SC	8	9	9	7	6

Table 3b. Analysis of Molecular Variance (AMOVA) of channeled whelk resource areas including the variation explained (Est.Var.), the F-statistic, and corresponding p-values. Significance was assessed using 9999 permutations of the data. Statistically significant F-statistics are bolded.

Source of Variation	Est. Var.	F-statistic	p-value
AMOVA 1			
Among Resource Areas	47.279	**F_{RT} = 0.282**	0.001
Among Groups	0.000	F_{SR} = -0.016	0.999
Among Individuals	16.177	**F_{ST} = 0.270**	0.001
Within Individuals	106.207	**F_{IS} = 0.132**	0.001
Total	169.663	**F_{IT} = 0.367**	0.001
AMOVA 2			
Among Resource Areas	47.069	**F_{RT} = 0.283**	0.001
Among Groups	0.000	F_{SR} = -0.024	1.000
Among Individuals	16.177	**F_{ST} = 0.265**	0.001
Within Individuals	106.207	**F_{IS} = 0.132**	0.001
Total	169.453	**F_{IT} = 0.362**	0.001
AMOVA 3			
Among Resource Areas	46.390	**F_{RT} = 0.279**	0.001
Among Groups	0.000	F_{SR} = -0.019	0.991
Among Individuals	16.177	**F_{ST} = 0.265**	0.001
Within Individuals	106.207	**F_{IS} = 0.132**	0.001
Total	168.773	**F_{IT} = 0.362**	0.001
AMOVA 4			
Among Resource Areas	50.848	**F_{RT} = 0.296**	0.001
Among Groups	0.000	F_{SR} = -0.014	1.000
Among Individuals	16.177	**F_{ST} = 0.286**	0.001
Within Individuals	106.207	**F_{IS} = 0.132**	0.001
Total	173.232	**F_{IT} = 0.381**	0.001
AMOVA 5			
Among Resource Areas	58.535	**F_{RT} = 0.326**	0.001
Among Groups	0.000	F_{SR} = -0.011	1.000
Among Individuals	16.177	**F_{ST} = 0.318**	0.001
Within Individuals	106.207	**F_{IS} = 0.132**	0.001
Total	180.919	**F_{IT} = 0.409**	0.001

Table 4. Wrights F-statistics (F_{ST}) for 10 channeled whelk resource areas. F_{ST} values are on the lower diagonal of the table, with significant F_{ST} values denoted in bold. P-values are located on the upper diagonal of the table. Significance was assessed using 10,000 permutations of the data. Resource area abbreviations are as follows: Buzzard Bay, Massachusetts (MA), Rhode Island (RI), Ocean City, Maryland (OC), Chincoteague, Virginia (CT), Eastern Shore, Virginia (ES), Light Tower, Virginia (LT), Virginia Beach, Virginia (VB), Pamlico Sound, North Carolina (NCPS), Wilmington, NC (NCW), Charleston, South Carolina (SC).

	MA	RI	OC	CT	ES	LT	VB	NCPS	NCW	SC
MA	—	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
RI	**0.084**	—	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
OC	**0.107**	**0.041**	—	0.032	0.564	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
CT	**0.119**	**0.044**	0.003	—	0.492	0.005	< 0.001	< 0.001	< 0.001	< 0.001
ES	**0.128**	**0.047**	0.0003	0.0001	—	0.396	< 0.001	< 0.001	< 0.001	< 0.001
LT	**0.088**	**0.031**	**0.007**	**0.005**	0.001	—	< 0.001	< 0.001	< 0.001	< 0.001
VB	**0.081**	**0.035**	**0.016**	**0.011**	**0.011**	**0.007**	—	< 0.001	< 0.001	< 0.001
NCPS	**0.184**	**0.144**	**0.187**	**0.145**	**0.125**	**0.145**	**0.151**	—	< 0.001	< 0.001
NCW	**0.337**	**0.292**	**0.382**	**0.310**	**0.278**	**0.322**	**0.357**	**0.197**	—	< 0.001
SC	**0.411**	**0.360**	**0.465**	**0.381**	**0.344**	**0.397**	**0.438**	**0.280**	**0.047**	—

Pairwise FST
0.150 - 0.465
0.050 - 0.149
0.005 - 0.049
0.0001- 0.0049

Table 5a. Population summary statistics for the 7 genetically distinct populations of channeled whelk, including separate statistics for 111 females and 97 males. H_O is observed heterozygosity, H_e is the expected heterozygosity, and G_{IS} is the inbreeding coefficient. Abbreviations are as follows: Buzzard Bay, Massachusetts (MA), Rhode Island (RI), mid-Atlantic (Ocean City, Maryland, Chincoteague, Virginia, Eastern Shore, Virginia, Light Tower, Virginia), Virginia Beach, Virginia (VB), Pamlico Sound, North Carolina (NCPS), Wilmington, NC (NCW), Charleston, South Carolina (SC).

Groups		Ho	He	G_{IS}
Genetic Group:				
	MA	0.011	0.010	-0.033
	RI	0.013	0.012	-0.035
	mid-Atlantic	0.020	0.020	0.026
	VB	0.030	0.030	0.002
	NCPS	0.124	0.121	-0.012
	NCW	0.253	0.248	-0.007
	SC	0.232	0.238	0.026
Overall		**0.098**	**0.097**	**-0.005**
Females:				
	MA	0.009	0.008	-0.105
	mid-Atlantic	0.020	0.019	-0.043
	VB	0.031	0.030	-0.016
	NCPS	0.123	0.119	-0.017
	NCW	0.279	0.250	-0.088
	SC	0.232	0.223	-0.046
Overall		**0.116**	**0.108**	**-0.053**
Males:				
	MA	0.011	0.011	-0.046
	mid-Atlantic	0.020	0.020	-0.050
	VB	0.029	0.029	-0.013
	NCPS	0.125	0.118	-0.056
	NCW	0.227	0.236	0.029
	SC	0.231	0.231	-0.005
Overall		**0.107**	**0.108**	**-0.023**

Table 5b. Paired t-test comparing the mean G_{IS} for 111 female and 97 male channeled whelk including the standard error (Std. Err), 95% Confidence Intervals (95% CI), the t statistic (t), degrees of freedom (df), and p-value.

Mean	Std. Err	95% CI		t	df	p-value
		Lower	Upper			
-0.029	0.023	-0.087	0.029	-1.28	5	0.257

Table 6. Estimates of effective population size (N_e) for the 7 genetically distinct populations of channeled whelk based on the linkage disequilibrium model with random mating. Critical values were set at 0.05, 0.02, and 0.01. Abbreviations are as follows: Buzzard Bay, Massachusetts (MA), Rhode Island (RI), mid-Atlantic (Ocean City, Maryland, Chincoteague, Virginia, Eastern Shore, Virginia, Light Tower, Virginia), Virginia Beach, Virginia (VB), Pamlico Sound, North Carolina (NCPS), Wilmington, NC (NCW), Charleston, South Carolina (SC).

Genetic Group	N_e
MA	
Allele Freq. 0.01	32.4 (27.7 - 38.8)
Allele Freq. 0.02	32.4 (27.7 - 38.8)
Allele Freq. 0.05	137.7 (50.8 - infinite)
Monomorphic Loci	2373
RI	
Allele Freq. 0.01	65.6 (42.7 - 132.5)
Allele Freq. 0.02	65.6 (42.7 - 132.5)
Allele Freq. 0.05	19.9 (13.2 - 34.8)
Monomorphic Loci	2376
mid-Atlantic	
Allele Freq. 0.01	195.5 (184.4 - 207.9)
Allele Freq. 0.02	297.6 (256.2 - 353.6)
Allele Freq. 0.05	1612.5 (576.4 - infinite)
Monomorphic Loci	1592
VB	
Allele Freq. 0.01	712.6 (534.6 - 1063.3)
Allele Freq. 0.02	400.9 (307.7 - 571.6)
Allele Freq. 0.05	68382.3 (670.5 - infinite)
Monomorphic Loci	1736
NCPS	
Allele Freq. 0.01	508.0 (434.8 - 610.2)
Allele Freq. 0.02	270.0 (242.7 - 304.0)
Allele Freq. 0.05	328.5 (283.6 - 389.7)
Monomorphic Loci	981
NCW	
Allele Freq. 0.01	227.9 (216.1 - 241.0)
Allele Freq. 0.02	206.2 (195.8 - 217.7)
Allele Freq. 0.05	193.1 (183.3 - 203.9)
Monomorphic Loci	284
SC	
Allele Freq. 0.01	833.6 (664.2 - 1117.5)
Allele Freq. 0.02	833.6 (664.2 - 1117.5)
Allele Freq. 0.05	904.6 (685.3 - 1327.6)
Monomorphic Loci	439

Table 7. Microsatellite primers for channeled whelk, *Busycotypus canaliculatus*, including the forward and reverse 5' to 3' sequence, the T3 dye used, the annealing temperature (Ta °C), the repeat motif, and the allele size range.

Locus	Forward 5' - 3'	Reverse 5' to 3'	Dye	Ta ºC	Repeat Motif	Allele Size Range
Bcan65	TAACCCAACGCGCTAGTCAG	TTCTCGTCCACTGCCGATAC	FAM	57	AGAT	166 - 171
Bcan62	ATACATGTCCGTCGGTCTGC	TGGGCGTGCGTACATATTAC	NED	57	ACAT	224 - 248
Bcan27	CCAGCTCTCAAGAAATCCGTC	AGAGCTGAAGACCCTCCAAC	FAM	51	ACAT	225 - 233
Bcan33	CCCTGTGTGCATGTGGAAC	ACCTTGTTGTCACTCGATCAG	PET	51	AGAT	214 - 234

Table 8. Mean observed and expected heterozygosity of microsatellite primers across four populations.

Loci	No. of samples	No. of alleles	Ho	He
Bcan65	70	2	0.145	0.126
Bcan62	70	4	0.268	0.228
Bcan27	70	3	0.340	0.294
Bcan33	70	7	0.455	0.377

Table 9. Inferred paternal genotypes from sibship reconstruction. OFS is the number of offspring genotypes analyzed, IM is the number of inferred mothers, and IF is the number of inferred fathers.

Group	OFS	IM	IF
Eggcase 1	70	1	11
Eggcase 2	70	1	10
Eggcase 3	68	1	7
Eggcase 4	68	1	6

FIGURES

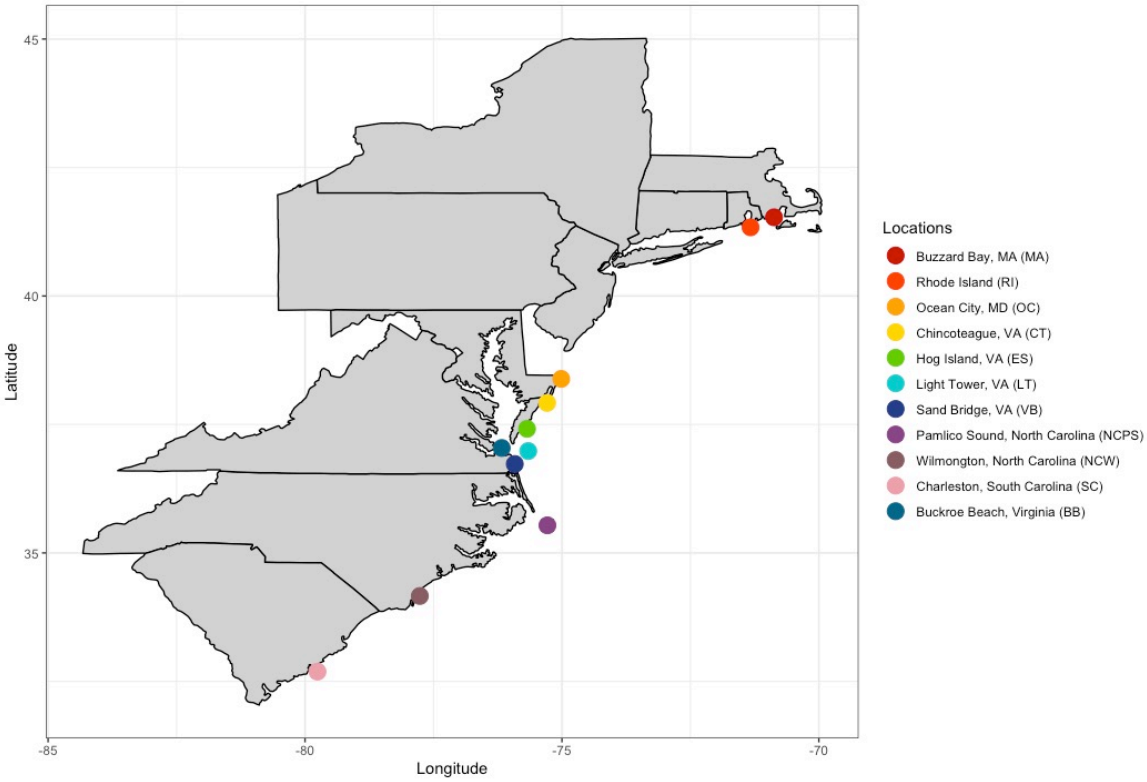


Figure 1. Geographic locations and abbreviations for channeled whelk resource areas sampled. A total of 282 individual samples were collected from 2015 – 2019.

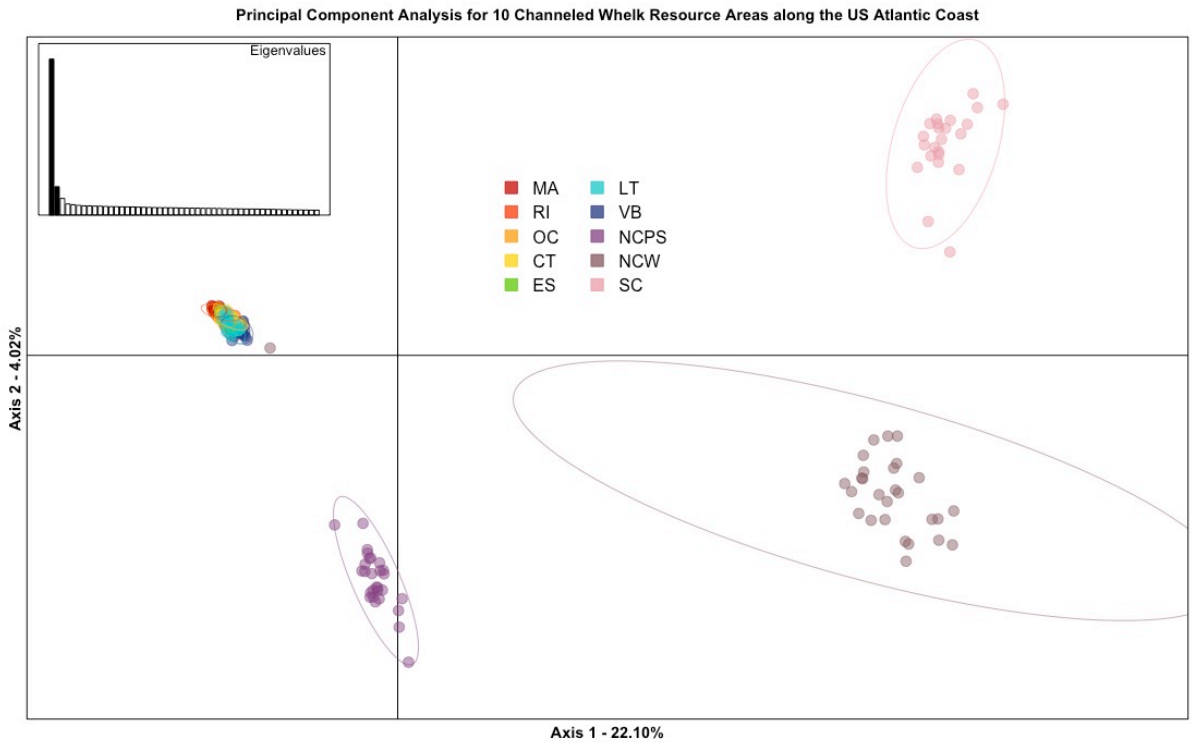


Figure 2. Principal Component Analysis (PCA) plotting the relative positioning for 10 channeled whelk resource areas along the US Atlantic coast using the first two principal components. Ellipses surround individuals grouped by resource areas. The eigenvalues calculated for this plot are highlighted in the top right corner. Axis 1 and 2 explained 22.10% and 4.02% of the variation in the data respectively. Resource area abbreviations are as follows: Buzzard Bay, Massachusetts (MA), Rhode Island (RI), Ocean City, Maryland (OC), Chincoteague, Virginia (CT), Eastern Shore, Virginia (ES), Light Tower, Virginia (LT), Virginia Beach, Virginia (VB), Pamlico Sound, North Carolina (NCPS), Wilmington, NC (NCW), Charleston, South Carolina (SC).

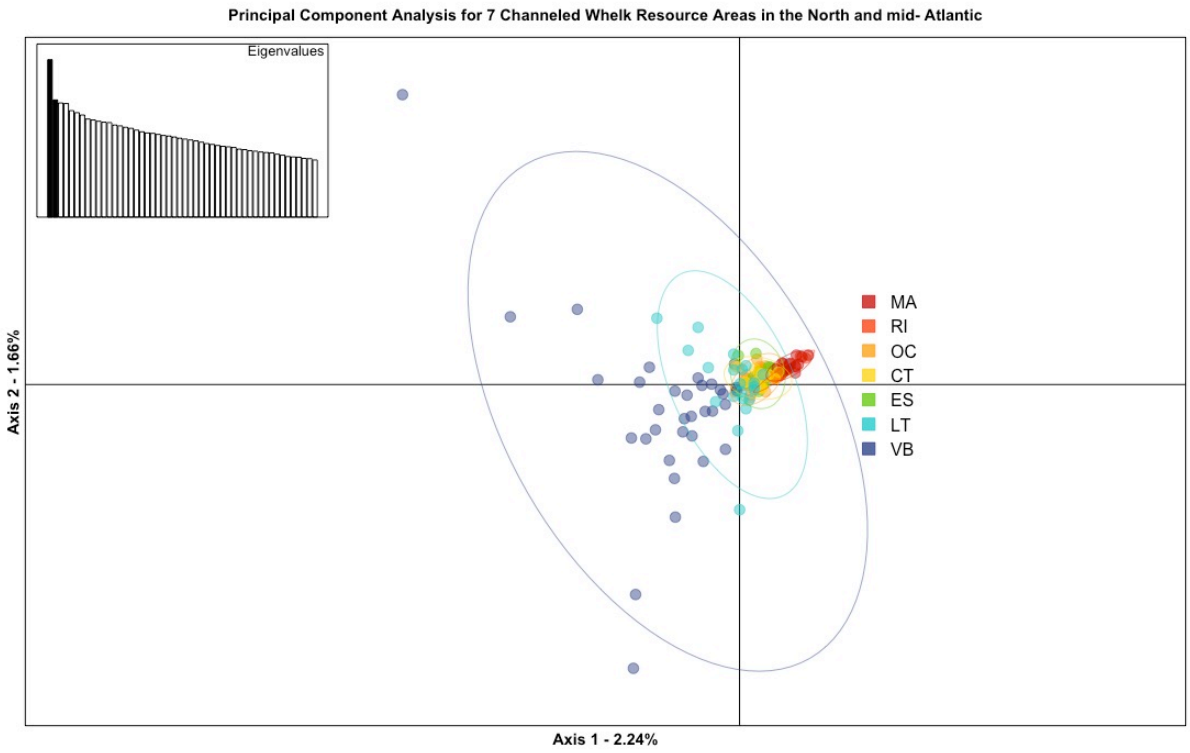


Figure 3. Principal Component Analysis (PCA) plotting the relative positioning for 7 channeled whelk resource areas along the US Atlantic coast using the first two principal component scores. Ellipses surround individuals grouped by resource areas. The eigenvalues calculated for this plot are highlighted in the top right corner. Axis 1 and 2 explained 2.24% and 1.66% of the variation in the data respectively. Resource area abbreviations are as follows: Buzzard Bay, Massachusetts (MA), Rhode Island (RI), Ocean City, Maryland (OC), Chincoteague, Virginia (CT), Eastern Shore, Virginia (ES), Light Tower, Virginia (LT), Virginia Beach, Virginia (VB).

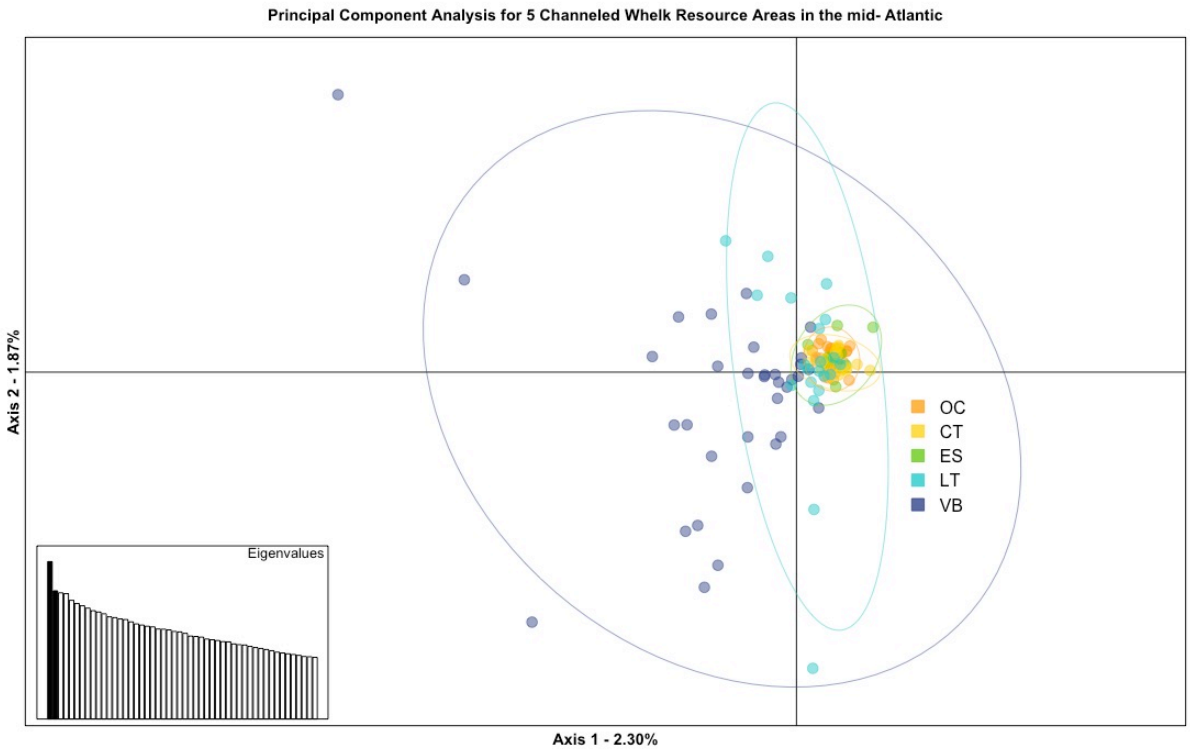


Figure 4. Principal Component Analysis (PCA) plotting the relative positioning for 5 channeled whelk resource areas along the US Atlantic coast using the first two principal component scores. Ellipses surround individuals grouped by resource areas. The eigenvalues calculated for this plot are highlighted in the top right corner. Axis 1 and 2 explained 2.30% and 1.87% of the variation in the data respectively. Resource area abbreviations are as follows: Ocean City, Maryland (OC), Chincoteague, Virginia (CT), Eastern Shore, Virginia (ES), Light Tower, Virginia (LT), Virginia Beach, Virginia (VB).

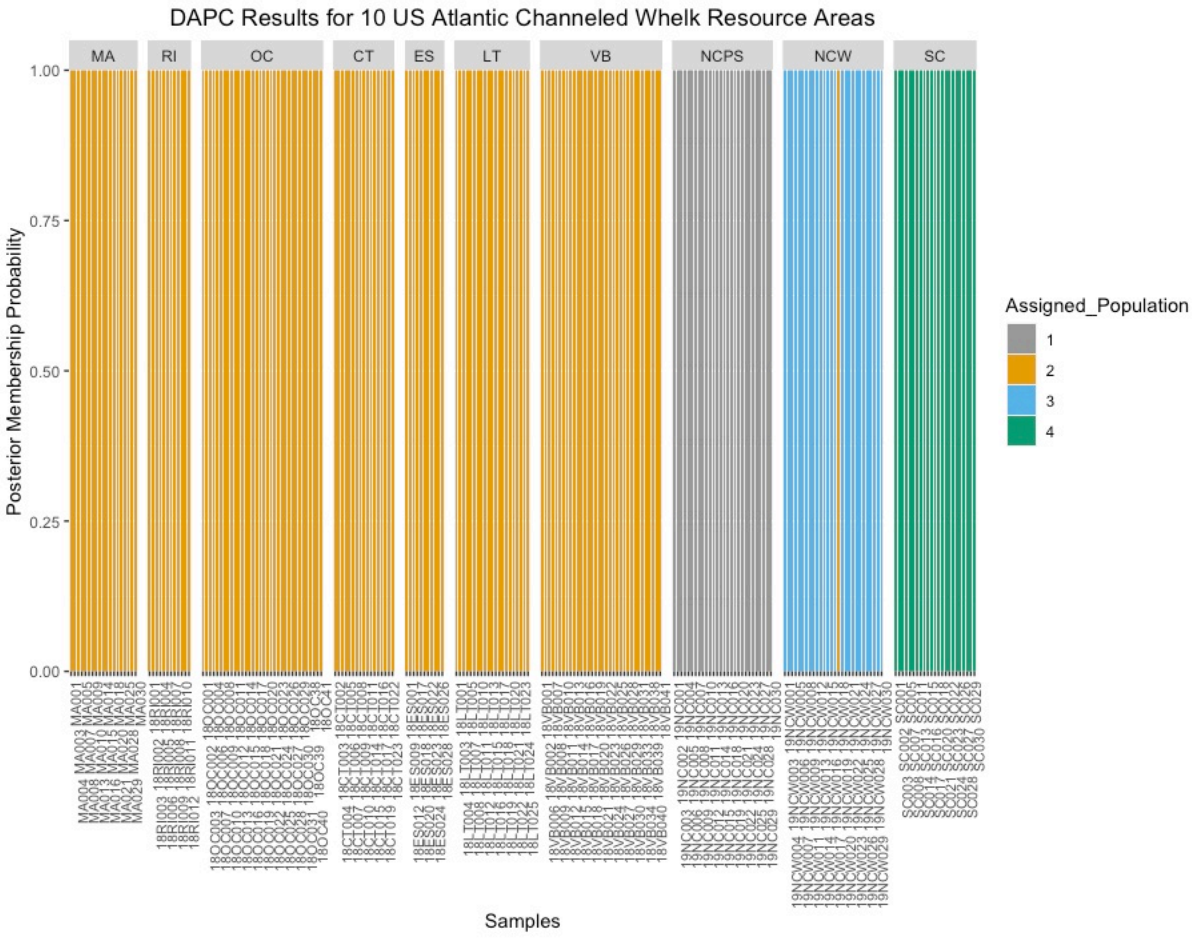


Figure 5. Discriminant Analysis of Principal Components (DAPC) plotting the posterior membership probability for 10 channeled whelk resource areas along the US Atlantic coast using 4 assigned populations. Resource area abbreviations are as follows: Buzzard Bay, Massachusetts (MA), Rhode Island (RI), Ocean City, Maryland (OC), Chincoteague, Virginia (CT), Eastern Shore, Virginia (ES), Light Tower, Virginia (LT), Virginia Beach, Virginia (VB), Pamlico Sound, North Carolina (NCPS), Wilmington, NC (NCW), Charleston, South Carolina (SC).

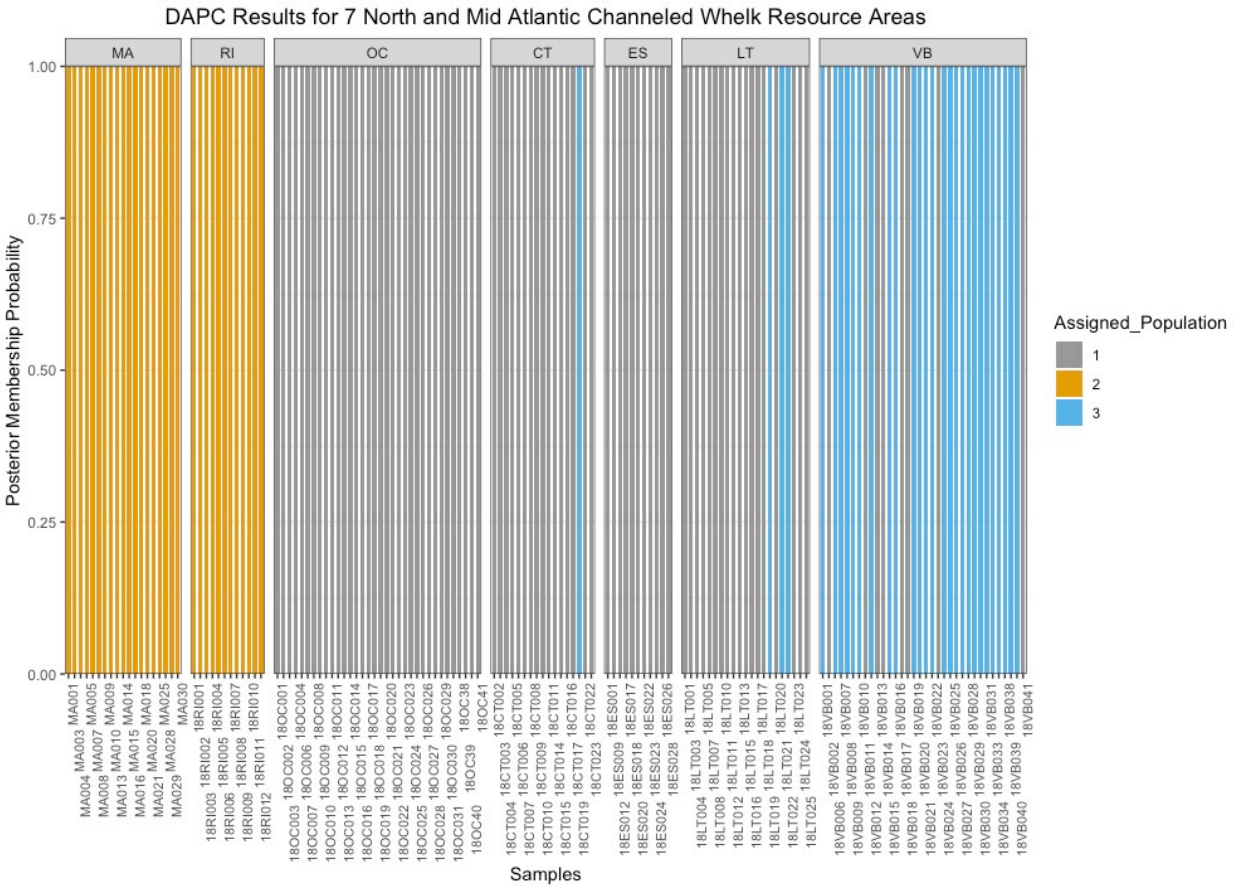


Figure 6. Discriminant Analysis of Principal Components (DAPC) plotting the posterior membership probability for 7 channeled whelk resource areas along the US Atlantic coast using 3 assigned populations. Resource area abbreviations are as follows: Buzzard Bay, Massachusetts (MA), Rhode Island (RI), Ocean City, Maryland (OC), Chincoteague, Virginia (CT), Eastern Shore, Virginia (ES), Light Tower, Virginia (LT), Virginia Beach, Virginia (VB).

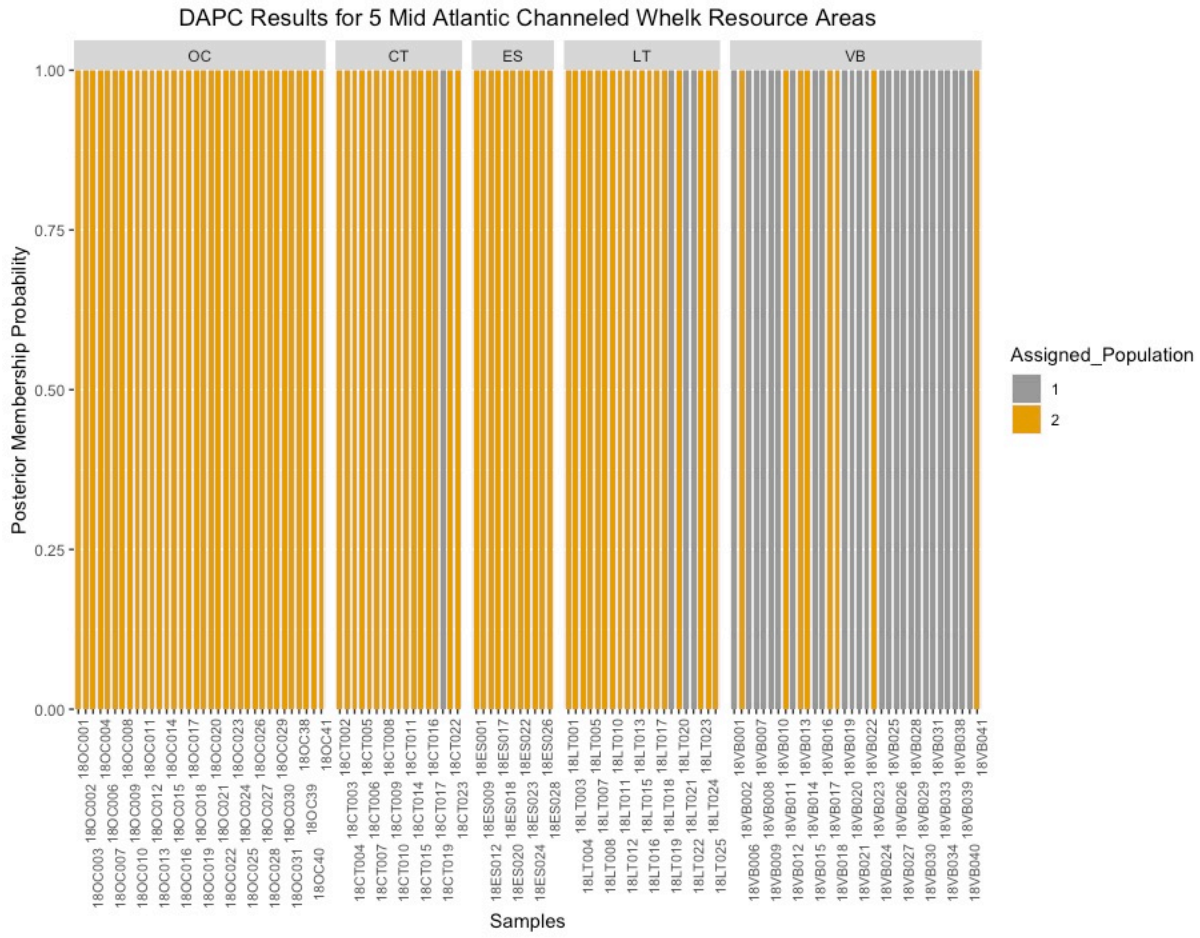


Figure 7. Discriminant Analysis of Principal Components (DAPC) plotting the posterior membership probability for 5 channeled whelk resource areas along the US Atlantic coast using 2 assigned populations. Resource area abbreviations are as follows: Ocean City, Maryland (OC), Chincoteague, Virginia (CT), Eastern Shore, Virginia (ES), Light Tower, Virginia (LT), Virginia Beach, Virginia (VB).

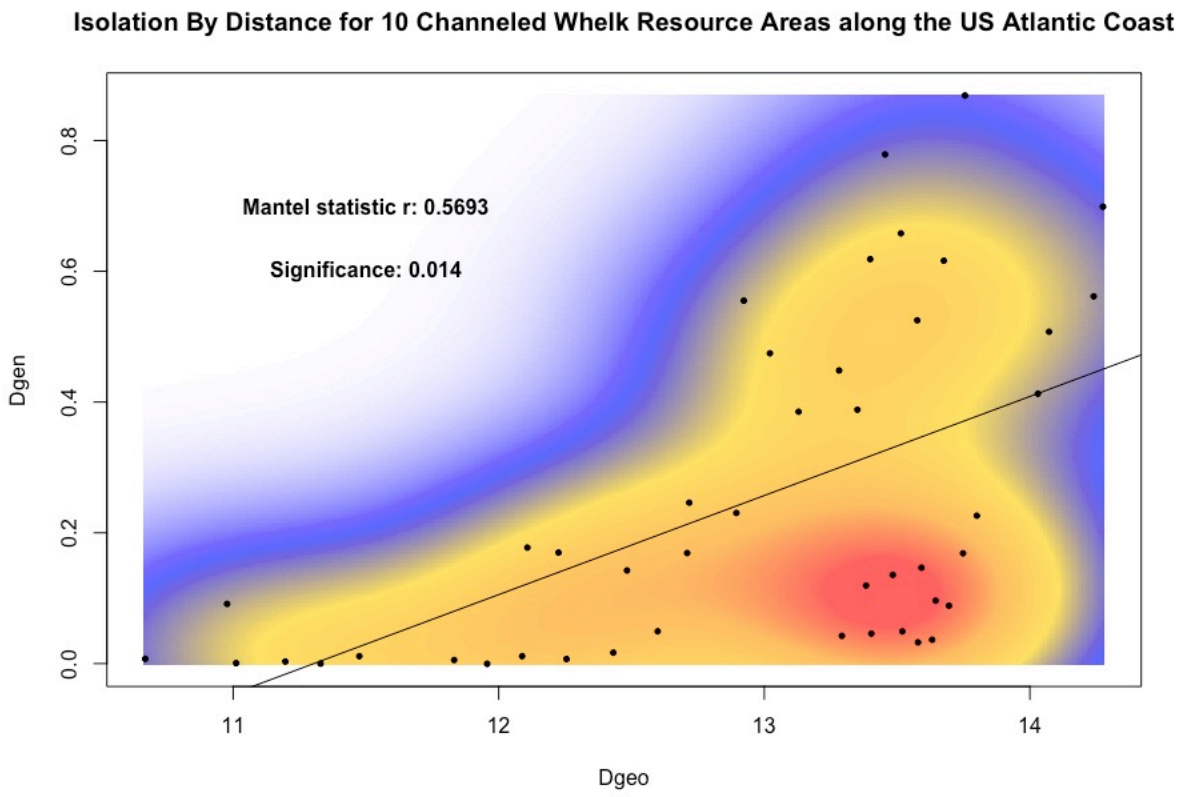


Figure 8. Mantel test using a matrix between Euclidean distance (Dgeo) and genetic distance (Dgen) to examine patterns of isolation by distance (IBD) for 10 channeled whelk resource areas along the US Atlantic coast. Significance was assessed using 999 permutations.

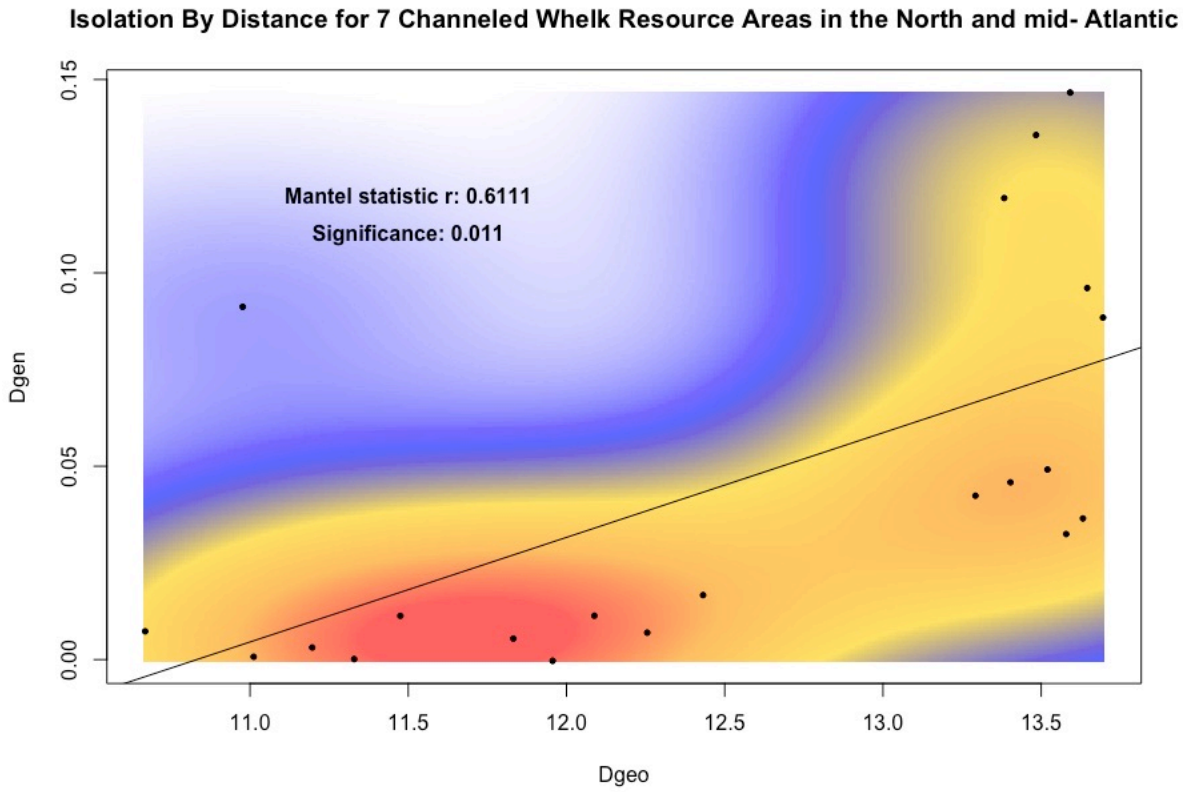


Figure 9. Mantel test using a matrix between Euclidean distance (Dgeo) and genetic distance (Dgen) to examine patterns of isolation by distance (IBD) for 7 channeled whelk resource areas along the US Atlantic coast. Significance was assessed using 999 permutations.

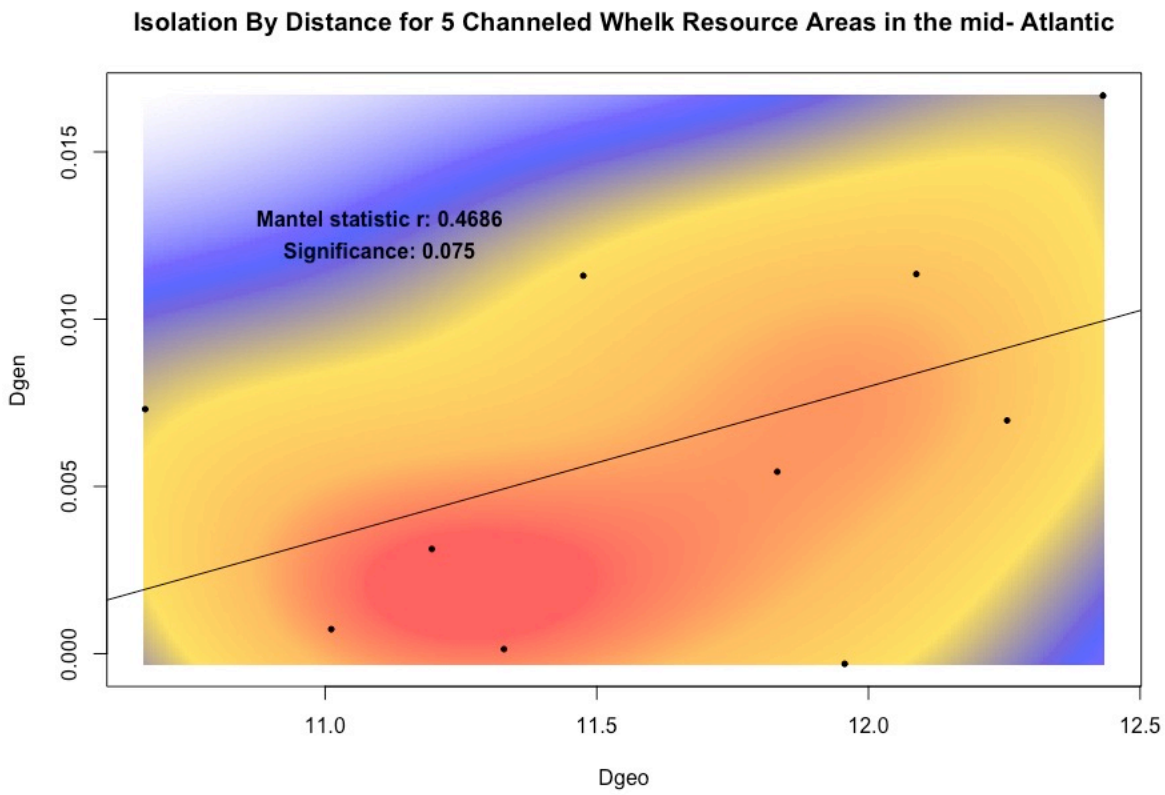


Figure 10. Mantel test using a matrix between Euclidean distance (Dgeo) and genetic distance (Dgen) to examine patterns of isolation by distance (IBD) for 5 channeled whelk resource areas along the US Atlantic coast. Significance was assessed using 999 permutations.

APPENDIX

S.1 Metadata for adult channeled whelk sampled from 10 resource areas along the US Atlantic coast.

Sample ID	Sample Location	Sampling Date	Shell Length (mm)	Shell Width (mm)	Sex	Lat	Lon
18MA001	Buzzard Bay, Massachusetts	7/9/19	155	81	F	41.531121	-70.873046
18MA002	Buzzard Bay, Massachusetts	7/9/19	157	86	F	41.531121	-70.873046
18MA003	Buzzard Bay, Massachusetts	7/9/19	165	89	F	41.531121	-70.873046
18MA004	Buzzard Bay, Massachusetts	7/9/19	143	72	M	41.531121	-70.873046
18MA005	Buzzard Bay, Massachusetts	7/9/19	139	82	F	41.531121	-70.873046
18MA006	Buzzard Bay, Massachusetts	7/9/19	142	74	M	41.531121	-70.873046
18MA007	Buzzard Bay, Massachusetts	7/9/19	129	67	M	41.531121	-70.873046
18MA008	Buzzard Bay, Massachusetts	7/9/19	134	72	M	41.531121	-70.873046
18MA009	Buzzard Bay, Massachusetts	7/9/19	137	73	F	41.531121	-70.873046
18MA010	Buzzard Bay, Massachusetts	7/9/19	138	74	M	41.531121	-70.873046
18MA011	Buzzard Bay, Massachusetts	7/9/19	154	86	F	41.531121	-70.873046
18MA012	Buzzard Bay, Massachusetts	7/9/19	147	81	F	41.531121	-70.873046
18MA013	Buzzard Bay, Massachusetts	7/9/19	N/A	92	F	41.531121	-70.873046
18MA014	Buzzard Bay, Massachusetts	7/9/19	126	68	M	41.531121	-70.873046
18MA015	Buzzard Bay, Massachusetts	7/9/19	94	58	F	41.531121	-70.873046
18MA016	Buzzard Bay, Massachusetts	7/9/19	N/A	67	M	41.531121	-70.873046
18MA017	Buzzard Bay, Massachusetts	7/9/19	131	71	M	41.531121	-70.873046
18MA018	Buzzard Bay, Massachusetts	7/9/19	137	70	M	41.531121	-70.873046
18MA019	Buzzard Bay, Massachusetts	7/9/19	142	75	M	41.531121	-70.873046
18MA020	Buzzard Bay, Massachusetts	7/9/19	139	77	M	41.531121	-70.873046
18MA021	Buzzard Bay, Massachusetts	7/9/19	144	78	M	41.531121	-70.873046
18MA022	Buzzard Bay, Massachusetts	7/9/19	98	48	M	41.531121	-70.873046
18MA023	Buzzard Bay, Massachusetts	7/9/19	132	65	M	41.531121	-70.873046
18MA024	Buzzard Bay, Massachusetts	7/9/19	123	64	M	41.531121	-70.873046
18MA025	Buzzard Bay, Massachusetts	7/9/19	129	68	M	41.531121	-70.873046
18MA026	Buzzard Bay, Massachusetts	7/9/19	133	72	M	41.531121	-70.873046
18MA027	Buzzard Bay, Massachusetts	7/9/19	126	70	M	41.531121	-70.873046
18MA028	Buzzard Bay, Massachusetts	7/9/19	132	71	M	41.531121	-70.873046
18MA029	Buzzard Bay, Massachusetts	7/9/19	137	73	M	41.531121	-70.873046
18MA030	Buzzard Bay, Massachusetts	7/9/19	140	73	M	41.531121	-70.873046
18RI001	Rhode Island	10/15/18	N/A	N/A	N/A	41.336159	-71.329855
18RI002	Rhode Island	10/15/18	N/A	N/A	N/A	41.336159	-71.329855
18RI003	Rhode Island	10/15/18	N/A	N/A	N/A	41.336159	-71.329855
18RI004	Rhode Island	10/15/18	N/A	N/A	N/A	41.336159	-71.329855
18RI005	Rhode Island	10/15/18	N/A	N/A	N/A	41.336159	-71.329855

18RI006	Rhode Island	10/15/18	N/A	N/A	N/A	41.336159	-71.329855
18RI007	Rhode Island	10/15/18	N/A	N/A	N/A	41.336159	-71.329855
18RI008	Rhode Island	10/15/18	N/A	N/A	N/A	41.336159	-71.329855
18RI009	Rhode Island	10/15/18	N/A	N/A	N/A	41.336159	-71.329855
18RI010	Rhode Island	10/15/18	N/A	N/A	N/A	41.336159	-71.329855
18RI011	Rhode Island	10/15/18	N/A	N/A	N/A	41.336159	-71.329855
18RI012	Rhode Island	10/15/18	N/A	N/A	N/A	41.336159	-71.329855
18OC001	Ocean City, Maryland	12/18/18	163	93	F	38.386951	-75.013776
18OC002	Ocean City, Maryland	12/18/18	137	68	M	38.386951	-75.013776
18OC003	Ocean City, Maryland	12/18/18	192	99	F	38.386951	-75.013776
18OC004	Ocean City, Maryland	12/18/18	208	118	F	38.386951	-75.013776
18OC005	Ocean City, Maryland	12/18/18	171	92	F	38.386951	-75.013776
18OC006	Ocean City, Maryland	12/18/18	135	69	F	38.386951	-75.013776
18OC007	Ocean City, Maryland	12/18/18	187	107	F	38.386951	-75.013776
18OC008	Ocean City, Maryland	12/18/18	122	68	M	38.386951	-75.013776
18OC009	Ocean City, Maryland	12/18/18	159	88	F	38.386951	-75.013776
18OC010	Ocean City, Maryland	12/18/18	130	72	F	38.386951	-75.013776
18OC011	Ocean City, Maryland	12/18/18	158	92	F	38.386951	-75.013776
18OC012	Ocean City, Maryland	12/18/18	184	103	F	38.386951	-75.013776
18OC013	Ocean City, Maryland	12/18/18	149	77	M	38.386951	-75.013776
18OC014	Ocean City, Maryland	12/18/18	80	38	F	38.386951	-75.013776
18OC015	Ocean City, Maryland	12/18/18	179	101	F	38.386951	-75.013776
18OC016	Ocean City, Maryland	12/18/18	171	93	F	38.386951	-75.013776
18OC017	Ocean City, Maryland	12/18/18	154	78	F	38.386951	-75.013776
18OC018	Ocean City, Maryland	12/18/18	134	69	F	38.386951	-75.013776
18OC019	Ocean City, Maryland	12/18/18	158	89	F	38.386951	-75.013776
18OC020	Ocean City, Maryland	12/18/18	152	80	F	38.386951	-75.013776
18OC021	Ocean City, Maryland	12/18/18	135	71	M	38.386951	-75.013776
18OC022	Ocean City, Maryland	12/18/18	168	88	F	38.386951	-75.013776
18OC023	Ocean City, Maryland	12/18/18	161	92	F	38.386951	-75.013776
18OC024	Ocean City, Maryland	12/18/18	107	54	F	38.386951	-75.013776
18OC025	Ocean City, Maryland	12/18/18	104	58	M	38.386951	-75.013776
18OC026	Ocean City, Maryland	12/18/18	168	98	F	38.386951	-75.013776
18OC027	Ocean City, Maryland	12/18/18	184	101	F	38.386951	-75.013776
18OC028	Ocean City, Maryland	12/18/18	177	102	F	38.386951	-75.013776
18OC029	Ocean City, Maryland	12/18/18	144	70	M	38.386951	-75.013776
18OC030	Ocean City, Maryland	12/18/18	128	64	F	38.386951	-75.013776
18OC031	Ocean City, Maryland	12/18/18	135	63	F	38.386951	-75.013776
18OC032	Ocean City, Maryland	12/18/18	134	65	F	38.386951	-75.013776
18OC033	Ocean City, Maryland	12/18/18	173	91	F	38.386951	-75.013776

18OC034	Ocean City, Maryland	12/18/18	157	79	M	38.386951	-75.013776
18OC035	Ocean City, Maryland	12/18/18	204	119	F	38.386951	-75.013776
18OC036	Ocean City, Maryland	12/18/18	111	53	F	38.386951	-75.013776
18OC037	Ocean City, Maryland	12/18/18	136	71	M	38.386951	-75.013776
18OC038	Ocean City, Maryland	12/18/18	84	42	F	38.386951	-75.013776
18OC039	Ocean City, Maryland	12/18/18	139	72	F	38.386951	-75.013776
18OC040	Ocean City, Maryland	12/18/18	170	93	F	38.386951	-75.013776
18OC041	Ocean City, Maryland	12/18/18	163	85	M	38.386951	-75.013776
18OC042	Ocean City, Maryland	12/18/18	174	91	F	38.386951	-75.013776
18OC043	Ocean City, Maryland	12/18/18	138	69	M	38.386951	-75.013776
18OC044	Ocean City, Maryland	12/18/18	172	92	F	38.386951	-75.013776
18OC045	Ocean City, Maryland	12/18/18	100	47	F	38.386951	-75.013776
18OC046	Ocean City, Maryland	12/18/18	171	90	F	38.386951	-75.013776
18OC047	Ocean City, Maryland	12/18/18	138	70	M	38.386951	-75.013776
18OC048	Ocean City, Maryland	12/18/18	165	88	F	38.386951	-75.013776
18OC049	Ocean City, Maryland	12/18/18	159	87	F	38.386951	-75.013776
18OC050	Ocean City, Maryland	12/18/18	157	84	F	38.386951	-75.013776
18CT001	Chincoteague, Virginia	11/20/18	152	N/A	F	37.919511	-75.287893
18CT002	Chincoteague, Virginia	11/20/18	138	80	M	37.919511	-75.287893
18CT003	Chincoteague, Virginia	11/20/18	164	95	M	37.919511	-75.287893
18CT004	Chincoteague, Virginia	11/20/18	152	82	M	37.919511	-75.287893
18CT005	Chincoteague, Virginia	11/20/18	163	94	F	37.919511	-75.287893
18CT006	Chincoteague, Virginia	11/20/18	141	70	M	37.919511	-75.287893
18CT007	Chincoteague, Virginia	11/20/18	175	102	M	37.919511	-75.287893
18CT008	Chincoteague, Virginia	11/20/18	195	106	F	37.919511	-75.287893
18CT009	Chincoteague, Virginia	11/20/18	148	74	M	37.919511	-75.287893
18CT010	Chincoteague, Virginia	11/20/18	167	94	M	37.919511	-75.287893
18CT011	Chincoteague, Virginia	11/20/18	190	103	F	37.919511	-75.287893
18CT012	Chincoteague, Virginia	11/20/18	138	72	M	37.919511	-75.287893
18CT013	Chincoteague, Virginia	11/20/18	166	90	M	37.919511	-75.287893
18CT014	Chincoteague, Virginia	11/20/18	150	79	M	37.919511	-75.287893
18CT015	Chincoteague, Virginia	11/20/18	143	77	M	37.919511	-75.287893
18CT016	Chincoteague, Virginia	11/20/18	203	118	F	37.919511	-75.287893
18CT017	Chincoteague, Virginia	11/20/18	190	104	F	37.919511	-75.287893
18CT018	Chincoteague, Virginia	11/20/18	172	95	M	37.919511	-75.287893
18CT019	Chincoteague, Virginia	11/20/18	211	118	F	37.919511	-75.287893
18CT020	Chincoteague, Virginia	11/20/18	222	125	F	37.919511	-75.287893
18CT021	Chincoteague, Virginia	11/20/18	200	110	F	37.919511	-75.287893
18CT022	Chincoteague, Virginia	11/20/18	172	87	M	37.919511	-75.287893
18CT023	Chincoteague, Virginia	11/20/18	184	105	F	37.919511	-75.287893

18CT024	Chincoteague, Virginia	11/20/18	203	115	F	37.919511	-75.287893
18CT025	Chincoteague, Virginia	11/20/18	183	95	F	37.919511	-75.287893
18CT026	Chincoteague, Virginia	11/27/18	N/A	101	F	37.919511	-75.287893
18CT027	Chincoteague, Virginia	11/27/18	162	86	M	37.919511	-75.287893
18CT028	Chincoteague, Virginia	11/27/18	N/A	86	F	37.919511	-75.287893
18CT029	Chincoteague, Virginia	11/27/18	146	74	F	37.919511	-75.287893
18CT030	Chincoteague, Virginia	11/27/18	147	73	F	37.919511	-75.287893
18CT031	Chincoteague, Virginia	11/27/18	171	90	F	37.919511	-75.287893
18CT032	Chincoteague, Virginia	11/27/18	N/A	82	M	37.919511	-75.287893
18CT033	Chincoteague, Virginia	11/27/18	155	80	M	37.919511	-75.287893
18CT034	Chincoteague, Virginia	11/27/18	196	103	F	37.919511	-75.287893
18CT035	Chincoteague, Virginia	11/27/18	128	66	M	37.919511	-75.287893
18CT036	Chincoteague, Virginia	11/27/18	211	107	F	37.919511	-75.287893
18CT037	Chincoteague, Virginia	11/27/18	170	88	F	37.919511	-75.287893
18CT038	Chincoteague, Virginia	11/27/18	213	117	F	37.919511	-75.287893
18CT039	Chincoteague, Virginia	11/27/18	166	83	F	37.919511	-75.287893
18CT040	Chincoteague, Virginia	11/27/18	141	71	F	37.919511	-75.287893
18CT041	Chincoteague, Virginia	11/27/18	138	72	M	37.919511	-75.287893
18ES001	Hog Island, Virginia	12/19/18	139	73	M	37.415635	-75.679773
18ES002	Hog Island, Virginia	12/19/18	132	68	F	37.415635	-75.679773
18ES003	Hog Island, Virginia	12/19/18	175	91	F	37.415635	-75.679773
18ES004	Hog Island, Virginia	12/19/18	153	79	F	37.415635	-75.679773
18ES005	Hog Island, Virginia	12/19/18	153	80	F	37.415635	-75.679773
18ES006	Hog Island, Virginia	12/19/18	181	100	F	37.415635	-75.679773
18ES007	Hog Island, Virginia	12/19/18	134	68	F	37.415635	-75.679773
18ES008	Hog Island, Virginia	12/19/18	141	75	M	37.415635	-75.679773
18ES009	Hog Island, Virginia	12/19/18	172	86	F	37.415635	-75.679773
18ES010	Hog Island, Virginia	12/19/18	135	73	F	37.415635	-75.679773
18ES011	Hog Island, Virginia	12/19/18	141	77	M	37.415635	-75.679773
18ES012	Hog Island, Virginia	12/19/18	138	68	M	37.415635	-75.679773
18ES013	Hog Island, Virginia	12/19/18	131	66	M	37.415635	-75.679773
18ES014	Hog Island, Virginia	12/19/18	129	67	M	37.415635	-75.679773
18ES015	Hog Island, Virginia	12/19/18	128	68	F	37.415635	-75.679773
18ES016	Hog Island, Virginia	12/19/18	178	90	F	37.415635	-75.679773
18ES017	Hog Island, Virginia	12/19/18	135	69	M	37.415635	-75.679773
18ES018	Hog Island, Virginia	12/19/18	191	96	F	37.415635	-75.679773
18ES019	Hog Island, Virginia	12/19/18	143	77	F	37.415635	-75.679773
18ES020	Hog Island, Virginia	12/19/18	177	93	F	37.415635	-75.679773
18ES021	Hog Island, Virginia	12/19/18	130	69	F	37.415635	-75.679773
18ES022	Hog Island, Virginia	12/19/18	139	71	F	37.415635	-75.679773

18ES023	Hog Island, Virginia	12/19/18	131	67	M	37.415635	-75.679773
18ES024	Hog Island, Virginia	12/19/18	127	63	M	37.415635	-75.679773
18ES025	Hog Island, Virginia	12/19/18	128	64	M	37.415635	-75.679773
18ES026	Hog Island, Virginia	12/19/18	135	65	M	37.415635	-75.679773
18ES027	Hog Island, Virginia	12/19/18	142	73	F	37.415635	-75.679773
18ES028	Hog Island, Virginia	12/19/18	129	62	F	37.415635	-75.679773
18ES029	Hog Island, Virginia	12/19/18	165	86	F	37.415635	-75.679773
18ES030	Hog Island, Virginia	12/19/18	143	73	F	37.415635	-75.679773
18ES031	Hog Island, Virginia	12/19/18	132	64	M	37.415635	-75.679773
18ES032	Hog Island, Virginia	12/19/18	144	69	M	37.415635	-75.679773
18ES033	Hog Island, Virginia	12/19/18	119	61	M	37.415635	-75.679773
18ES034	Hog Island, Virginia	12/19/18	141	70	F	37.415635	-75.679773
18ES035	Hog Island, Virginia	12/19/18	129	66	M	37.415635	-75.679773
18ES036	Hog Island, Virginia	12/19/18	146	75	F	37.415635	-75.679773
18ES037	Hog Island, Virginia	12/19/18	130	65	M	37.415635	-75.679773
18ES038	Hog Island, Virginia	12/19/18	128	60	M	37.415635	-75.679773
18ES039	Hog Island, Virginia	12/19/18	132	69	F	37.415635	-75.679773
18ES040	Hog Island, Virginia	12/19/18	130	66	M	37.415635	-75.679773
18ES041	Hog Island, Virginia	12/19/18	127	64	F	37.415635	-75.679773
18ES042	Hog Island, Virginia	12/19/18	134	69	M	37.415635	-75.679773
18ES043	Hog Island, Virginia	12/19/18	134	65	F	37.415635	-75.679773
18ES044	Hog Island, Virginia	12/19/18	125	62	M	37.415635	-75.679773
18ES045	Hog Island, Virginia	12/19/18	142	71	F	37.415635	-75.679773
18ES046	Hog Island, Virginia	12/19/18	165	84	F	37.415635	-75.679773
18ES047	Hog Island, Virginia	12/19/18	139	73	F	37.415635	-75.679773
18ES048	Hog Island, Virginia	12/19/18	127	65	M	37.415635	-75.679773
18ES049	Hog Island, Virginia	12/19/18	136	70	F	37.415635	-75.679773
18ES050	Hog Island, Virginia	12/19/18	144	73	F	37.415635	-75.679773
18LT001	Light Tower, Virginia	12/19/18	148	74	M	36.982738	-75.65605
18LT002	Light Tower, Virginia	12/19/18	153	77	F	36.982738	-75.65605
18LT003	Light Tower, Virginia	12/19/18	179	101	F	36.982738	-75.65605
18LT004	Light Tower, Virginia	12/19/18	142	73	F	36.982738	-75.65605
18LT005	Light Tower, Virginia	12/19/18	145	75	F	36.982738	-75.65605
18LT006	Light Tower, Virginia	12/19/18	132	68	F	36.982738	-75.65605
18LT007	Light Tower, Virginia	12/19/18	151	79	F	36.982738	-75.65605
18LT008	Light Tower, Virginia	12/19/18	134	66	F	36.982738	-75.65605
18LT009	Light Tower, Virginia	12/19/18	127	63	M	36.982738	-75.65605
18LT010	Light Tower, Virginia	12/19/18	135	71	M	36.982738	-75.65605
18LT011	Light Tower, Virginia	12/19/18	133	68	F	36.982738	-75.65605
18LT012	Light Tower, Virginia	12/19/18	208	113	F	36.982738	-75.65605

18LT013	Light Tower, Virginia	12/19/18	160	81	M	36.982738	-75.65605
18LT014	Light Tower, Virginia	12/19/18	154	76	M	36.982738	-75.65605
18LT015	Light Tower, Virginia	12/19/18	146	73	M	36.982738	-75.65605
18LT016	Light Tower, Virginia	12/19/18	157	77	F	36.982738	-75.65605
18LT017	Light Tower, Virginia	12/19/18	140	69	F	36.982738	-75.65605
18LT018	Light Tower, Virginia	12/19/18	147	72	M	36.982738	-75.65605
18LT019	Light Tower, Virginia	12/19/18	137	67	F	36.982738	-75.65605
18LT020	Light Tower, Virginia	12/19/18	155	78	F	36.982738	-75.65605
18LT021	Light Tower, Virginia	12/19/18	144	69	M	36.982738	-75.65605
18LT022	Light Tower, Virginia	12/19/18	N/A	71	M	36.982738	-75.65605
18LT023	Light Tower, Virginia	12/19/18	134	63	F	36.982738	-75.65605
18LT024	Light Tower, Virginia	12/19/18	N/A	68	M	36.982738	-75.65605
18LT025	Light Tower, Virginia	12/19/18	N/A	69	M	36.982738	-75.65605
18VB001	Sand Bridge, Virginia	12/17/18	186	100	F	36.730475	-75.918162
18VB002	Sand Bridge, Virginia	12/17/18	130	68	M	36.730475	-75.918162
18VB003	Sand Bridge, Virginia	12/17/18	154	77	M	36.730475	-75.918162
18VB004	Sand Bridge, Virginia	12/17/18	153	81	F	36.730475	-75.918162
18VB005	Sand Bridge, Virginia	12/17/18	137	71	M	36.730475	-75.918162
18VB006	Sand Bridge, Virginia	12/17/18	202	109	F	36.730475	-75.918162
18VB007	Sand Bridge, Virginia	12/17/18	163	86	F	36.730475	-75.918162
18VB008	Sand Bridge, Virginia	12/17/18	146	82	M	36.730475	-75.918162
18VB009	Sand Bridge, Virginia	12/17/18	187	107	F	36.730475	-75.918162
18VB010	Sand Bridge, Virginia	12/17/18	197	113	F	36.730475	-75.918162
18VB011	Sand Bridge, Virginia	12/17/18	173	98	M	36.730475	-75.918162
18VB012	Sand Bridge, Virginia	12/17/18	151	75	M	36.730475	-75.918162
18VB013	Sand Bridge, Virginia	12/17/18	201	110	F	36.730475	-75.918162
18VB014	Sand Bridge, Virginia	12/17/18	140	75	M	36.730475	-75.918162
18VB015	Sand Bridge, Virginia	12/17/18	185	101	F	36.730475	-75.918162
18VB016	Sand Bridge, Virginia	12/17/18	130	69	M	36.730475	-75.918162
18VB017	Sand Bridge, Virginia	12/17/18	158	84	M	36.730475	-75.918162
18VB018	Sand Bridge, Virginia	12/17/18	145	77	M	36.730475	-75.918162
18VB019	Sand Bridge, Virginia	12/17/18	167	89	F	36.730475	-75.918162
18VB020	Sand Bridge, Virginia	12/17/18	148	80	M	36.730475	-75.918162
18VB021	Sand Bridge, Virginia	12/17/18	176	93	F	36.730475	-75.918162
18VB022	Sand Bridge, Virginia	12/17/18	198	108	F	36.730475	-75.918162
18VB023	Sand Bridge, Virginia	12/17/18	182	101	F	36.730475	-75.918162
18VB024	Sand Bridge, Virginia	12/17/18	146	80	M	36.730475	-75.918162
18VB025	Sand Bridge, Virginia	12/17/18	154	80	M	36.730475	-75.918162
18VB026	Sand Bridge, Virginia	12/17/18	189	101	F	36.730475	-75.918162
18VB027	Sand Bridge, Virginia	12/17/18	186	100	F	36.730475	-75.918162

18VB028	Sand Bridge, Virginia	12/17/18	190	103	F	36.730475	-75.918162
18VB029	Sand Bridge, Virginia	12/17/18	146	75	M	36.730475	-75.918162
18VB030	Sand Bridge, Virginia	12/17/18	155	81	F	36.730475	-75.918162
18VB031	Sand Bridge, Virginia	12/17/18	125	64	M	36.730475	-75.918162
18VB032	Sand Bridge, Virginia	12/17/18	163	84	F	36.730475	-75.918162
18VB033	Sand Bridge, Virginia	12/17/18	136	70	M	36.730475	-75.918162
18VB034	Sand Bridge, Virginia	12/17/18	155	81	M	36.730475	-75.918162
18VB035	Sand Bridge, Virginia	12/17/18	190	102	F	36.730475	-75.918162
18VB036	Sand Bridge, Virginia	12/17/18	147	73	M	36.730475	-75.918162
18VB037	Sand Bridge, Virginia	12/17/18	169	89	F	36.730475	-75.918162
18VB038	Sand Bridge, Virginia	12/17/18	129	64	F	36.730475	-75.918162
18VB039	Sand Bridge, Virginia	12/17/18	131	65	M	36.730475	-75.918162
18VB040	Sand Bridge, Virginia	12/17/18	132	66	M	36.730475	-75.918162
18VB041	Sand Bridge, Virginia	12/17/18	187	103	F	36.730475	-75.918162
18VB042	Sand Bridge, Virginia	12/17/18	155	79	M	36.730475	-75.918162
18VB043	Sand Bridge, Virginia	12/17/18	214	114	F	36.730475	-75.918162
18VB044	Sand Bridge, Virginia	12/17/18	192	100	F	36.730475	-75.918162
18VB045	Sand Bridge, Virginia	12/17/18	156	85	M	36.730475	-75.918162
18VB046	Sand Bridge, Virginia	12/17/18	154	80	M	36.730475	-75.918162
18VB047	Sand Bridge, Virginia	12/17/18	147	77	M	36.730475	-75.918162
18VB048	Sand Bridge, Virginia	12/17/18	129	66	M	36.730475	-75.918162
18VB049	Sand Bridge, Virginia	12/17/18	135	70	F	36.730475	-75.918162
18VB050	Sand Bridge, Virginia	12/17/18	203	109	F	36.730475	-75.918162
19NCPS001	Pamlico Sound, North Carolina	3/26/19	111	55	F	35.537239	-75.282871
19NCPS002	Pamlico Sound, North Carolina	3/26/19	106	52	F	35.537239	-75.282871
19NCPS003	Pamlico Sound, North Carolina	3/26/19	94	50	M	35.537239	-75.282871
19NCPS004	Pamlico Sound, North Carolina	3/26/19	113	58	F	35.537239	-75.282871
19NCPS005	Pamlico Sound, North Carolina	3/26/19	115	57	F	35.537239	-75.282871
19NCPS006	Pamlico Sound, North Carolina	3/26/19	108	55	F	35.537239	-75.282871
19NCPS007	Pamlico Sound, North Carolina	3/26/19	119	60	F	35.537239	-75.282871
19NCPS008	Pamlico Sound, North Carolina	3/26/19	105	50	M	35.537239	-75.282871
19NCPS009	Pamlico Sound, North Carolina	3/26/19	104	48	M	35.537239	-75.282871
19NCPS010	Pamlico Sound, North Carolina	3/26/19	111	60	F	35.537239	-75.282871
19NCPS011	Pamlico Sound, North Carolina	3/26/19	102	53	M	35.537239	-75.282871
19NCPS012	Pamlico Sound, North Carolina	3/26/19	100	51	M	35.537239	-75.282871
19NCPS013	Pamlico Sound, North Carolina	3/26/19	123	62	F	35.537239	-75.282871
19NCPS014	Pamlico Sound, North Carolina	3/26/19	97	40	M	35.537239	-75.282871
19NCPS015	Pamlico Sound, North Carolina	3/26/19	102	52	M	35.537239	-75.282871
19NCPS016	Pamlico Sound, North Carolina	3/26/19	103	54	M	35.537239	-75.282871
19NCPS017	Pamlico Sound, North Carolina	3/26/19	108	52	F	35.537239	-75.282871

19NCPS018	Pamlico Sound, North Carolina	3/26/19	109	55	F	35.537239	-75.282871
19NCPS019	Pamlico Sound, North Carolina	3/26/19	105	51	F	35.537239	-75.282871
19NCPS020	Pamlico Sound, North Carolina	3/26/19	120	61	F	35.537239	-75.282871
19NCPS021	Pamlico Sound, North Carolina	3/26/19	N/A	50	F	35.537239	-75.282871
19NCPS022	Pamlico Sound, North Carolina	3/26/19	94	47	M	35.537239	-75.282871
19NCPS023	Pamlico Sound, North Carolina	3/26/19	115	59	F	35.537239	-75.282871
19NCPS024	Pamlico Sound, North Carolina	3/26/19	110	57	F	35.537239	-75.282871
19NCPS025	Pamlico Sound, North Carolina	3/26/19	112	56	F	35.537239	-75.282871
19NCPS026	Pamlico Sound, North Carolina	3/26/19	113	59	F	35.537239	-75.282871
19NCPS027	Pamlico Sound, North Carolina	3/26/19	91	42	F	35.537239	-75.282871
19NCPS028	Pamlico Sound, North Carolina	3/26/19	100	51	F	35.537239	-75.282871
19NCPS029	Pamlico Sound, North Carolina	3/26/19	103	52	M	35.537239	-75.282871
19NCPS030	Pamlico Sound, North Carolina	3/26/19	107	54	M	35.537239	-75.282871
19NCW001	Wilmington, North Carolina	5/5/19	164	86	F	34.162302	-77.765915
19NCW002	Wilmington, North Carolina	5/5/19	160	78	F	34.162302	-77.765915
19NCW003	Wilmington, North Carolina	5/5/19	152	79	F	34.162302	-77.765915
19NCW004	Wilmington, North Carolina	5/5/19	113	59	M	34.162302	-77.765915
19NCW005	Wilmington, North Carolina	5/5/19	136	69	M	34.162302	-77.765915
19NCW006	Wilmington, North Carolina	5/5/19	120	59	M	34.162302	-77.765915
19NCW007	Wilmington, North Carolina	5/5/19	119	56	F	34.162302	-77.765915
19NCW008	Wilmington, North Carolina	5/5/19	119	58	M	34.162302	-77.765915
19NCW009	Wilmington, North Carolina	5/5/19	125	64	M	34.162302	-77.765915
19NCW010	Wilmington, North Carolina	5/5/19	123	62	F	34.162302	-77.765915
19NCW011	Wilmington, North Carolina	5/5/19	140	71	F	34.162302	-77.765915
19NCW012	Wilmington, North Carolina	5/5/19	127	62	F	34.162302	-77.765915
19NCW013	Wilmington, North Carolina	5/5/19	113	56	M	34.162302	-77.765915
19NCW014	Wilmington, North Carolina	5/5/19	123	59	F	34.162302	-77.765915
19NCW015	Wilmington, North Carolina	5/5/19	109	52	M	34.162302	-77.765915
19NCW016	Wilmington, North Carolina	5/5/19	107	52	F	34.162302	-77.765915
19NCW017	Wilmington, North Carolina	5/5/19	119	60	M	34.162302	-77.765915
19NCW018	Wilmington, North Carolina	5/5/19	134	68	M	34.162302	-77.765915
19NCW019	Wilmington, North Carolina	5/5/19	112	52	M	34.162302	-77.765915
19NCW020	Wilmington, North Carolina	5/5/19	127	62	F	34.162302	-77.765915
19NCW021	Wilmington, North Carolina	5/5/19	108	53	M	34.162302	-77.765915
19NCW022	Wilmington, North Carolina	5/5/19	103	51	M	34.162302	-77.765915
19NCW023	Wilmington, North Carolina	5/5/19	158	81	F	34.162302	-77.765915
19NCW024	Wilmington, North Carolina	5/5/19	105	57	M	34.162302	-77.765915
19NCW025	Wilmington, North Carolina	5/5/19	142	72	F	34.162302	-77.765915
19NCW026	Wilmington, North Carolina	5/5/19	113	57	F	34.162302	-77.765915
19NCW027	Wilmington, North Carolina	5/5/19	114	57	F	34.162302	-77.765915

19NCW028	Wilmington, North Carolina	5/5/19	132	68	F	34.162302	-77.765915
19NCW029	Wilmington, North Carolina	5/5/19	153	76	F	34.162302	-77.765915
19NCW030	Wilmington, North Carolina	5/5/19	115	56	M	34.162302	-77.765915
SC001	Charleston, South Carolina	10/31/15	88	N/A	F	32.691005	-79.757874
SC002	Charleston, South Carolina	10/31/15	78	N/A	M	32.691005	-79.757874
SC003	Charleston, South Carolina	10/31/15	72	N/A	F	32.691005	-79.757874
SC004	Charleston, South Carolina	10/31/15	54	N/A	F	32.691005	-79.757874
SC005	Charleston, South Carolina	11/13/15	136	N/A	F	32.691005	-79.757874
SC006	Charleston, South Carolina	4/11/16	87	N/A	M	32.691005	-79.757874
SC007	Charleston, South Carolina	8/2/16	57	N/A	M	32.691005	-79.757874
SC008	Charleston, South Carolina	5/12/15	60	N/A	N/A	32.691005	-79.757874
SC009	Charleston, South Carolina	5/12/15	37	N/A	M	32.691005	-79.757874
SC010	Charleston, South Carolina	5/12/15	47	N/A	F	32.691005	-79.757874
SC011	Charleston, South Carolina	2/21/17	55	N/A	N/A	32.691005	-79.757874
SC012	Charleston, South Carolina	2/21/17	56	N/A	N/A	32.691005	-79.757874
SC013	Charleston, South Carolina	11/9/16	126	N/A	M	32.691005	-79.757874
SC014	Charleston, South Carolina	10/13/16	127.5	N/A	M	32.691005	-79.757874
SC015	Charleston, South Carolina	10/13/16	113	N/A	M	32.691005	-79.757874
SC016	Charleston, South Carolina	10/13/16	98	N/A	N/A	32.691005	-79.757874
SC017	Charleston, South Carolina	4/21/16	87	N/A	M	32.691005	-79.757874
SC018	Charleston, South Carolina	11/17/16	88	N/A	M	32.691005	-79.757874
SC019	Charleston, South Carolina	11/17/16	104	N/A	M	32.691005	-79.757874
SC020	Charleston, South Carolina	11/17/16	57	N/A	N/A	32.691005	-79.757874
SC021	Charleston, South Carolina	11/17/16	50	N/A	N/A	32.691005	-79.757874
SC022	Charleston, South Carolina	4/18/17	130	N/A	F	32.691005	-79.757874
SC023	Charleston, South Carolina	4/10/17	95	N/A	M	32.691005	-79.757874
SC024	Charleston, South Carolina	4/10/17	134	N/A	F	32.691005	-79.757874
SC025	Charleston, South Carolina	7/10/17	81	N/A	F	32.691005	-79.757874
SC026	Charleston, South Carolina	7/29/17	70	N/A	F	32.691005	-79.757874
SC027	Charleston, South Carolina	3/9/18	113	N/A	M	32.691005	-79.757874
SC028	Charleston, South Carolina	3/9/18	119	N/A	M	32.691005	-79.757874
SC029	Charleston, South Carolina	10/27/17	78	N/A	N/A	32.691005	-79.757874
SC030	Charleston, South Carolina	10/28/17	85	N/A	N/A	32.691005	-79.757874

S.2 The samples and amount of DNA used for DArTseq.

Sample ID	Nucleic Acid (ng/µl)	260/280	Amount of DNA sent for DArTseq
19MA022	71.6	1.44	20
19MA019	65.8	1.39	20
19MA024	56.8	1.4	20
19MA008	49.6	1.44	40
19MA010	39.6	1.52	40
19MA009	18	1.52	100
19MA002	17.9	2.16	100
19MA004	16.4	1.51	100
19MA005	13.1	1.71	100
19MA026	12.7	1.8	100
19MA029	11.8	1.46	100
19MA011	11.4	1.87	100
19MA023	10.9	1.96	100
19MA021	9.5	1.79	180
19MA012	9.4	1.77	180
19MA018	8.2	1.57	180
19MA006	7.8	2.06	180
19MA020	7.8	1.47	180
19MA028	7.8	1.55	180
19MA013	7.3	1.83	180
19MA017	7	1.88	180
19MA003	5.5	1.66	180
19MA030	5.3	1.55	180
19MA016	5.2	1.74	180
19MA001	5.1	2.05	180
19MA027	4.3	1.89	180
19MA015	3.3	1.94	180
19MA014	3	1.58	180
19MA025	3	2.54	180
19MA007	2.8	1.77	180
18RI011	291.9	1.72	7
18RI001	36.7	1.89	40
18RI005	21.9	1.97	60
18RI006	17.5	1.91	80
18RI008	15.7	2.06	80
18RI002	15.6	1.69	80

18RI007	15.1	1.48	80
18RI004	12.7	2.06	100
18RI010	12.3	2.06	100
18RI003	8.8	1.93	180
18RI012	7.3	1.65	180
18RI009	6.3	2.36	180
18OC004	95.3	2	20
18OC015	67.1	1.93	20
18OC023	52.9	1.93	20
18OC003	50.8	2	20
18OC031	46.7	1.94	40
18OC024	46.4	1.9	40
18OC014	45.6	2	40
18OC006	42.4	2.02	40
18OC002	39.9	1.95	40
18OC017	39.7	2	40
18OC009	38.4	1.95	40
18OC001	34.6	1.94	40
18OC016	34.3	2.02	40
18OC010	34	2	40
18OC029	33.9	2.05	40
18OC013	32.7	1.97	40
18OC022	32.6	1.95	40
18OC027	32.5	1.95	40
18OC008	32.2	1.98	40
18OC028	30	1.97	40
18OC021	27.3	2.03	60
18OC007	26.2	2.09	60
18OC011	23.2	1.9	60
18OC026	22.6	1.98	60
18OC018	20.9	2.05	60
18OC030	20	2.12	60
18OC012	16.4	2.07	80
18OC019	15.1	2.07	80
18OC020	10.9	2.27	120
18OC025	8.4	1.94	120
18CT004	19.5	2.13	80
18CT024	18.7	2.01	80
18CT014	18.4	2.38	80
18CT008	18.1	2.13	80

18CT012	17.3	2.24	80
18CT009	16.8	2.22	80
18CT001	12.6	2.08	120
18CT023	10.6	2.2	120
18CT020	10.1	1.96	120
18CT002	9.7	2.41	120
18CT006	9.7	2.16	120
18CT003	8.5	2.06	120
18CT025	6.7	2.3	180
18CT011	6.3	2.8	180
18CT017	5.5	2	180
18CT013	5.4	2.94	180
18CT016	5.3	2.68	180
18CT005	5.2	2.49	180
18CT010	5.2	3.15	180
18CT007	4.2	4.32	180
18CT021	4	2.31	180
18CT018	3.8	2.48	180
18CT019	3.6	2.77	180
18CT015	2.8	14.21	180
18CT022	-23.3	1.36	80
18ES024	93.6	2.04	20
18ES012	75	2	20
18ES014	74.1	2.03	20
18ES030	65.7	2.05	20
18ES019	64.6	2.04	20
18ES023	55.5	2	20
18ES020	53	2.01	20
18ES029	52	2.02	20
18ES026	51.7	2.03	20
18ES015	51.1	2.05	20
18ES005	49.3	2.04	40
18ES008	47.1	2.12	40
18ES001	46.8	2.14	40
18ES013	44.5	2.11	40
18ES016	44.1	2.04	40
18ES009	43	2.02	40
18ES028	39.2	1.99	40
18ES025	39	2.06	40
18ES022	37.3	2.03	40

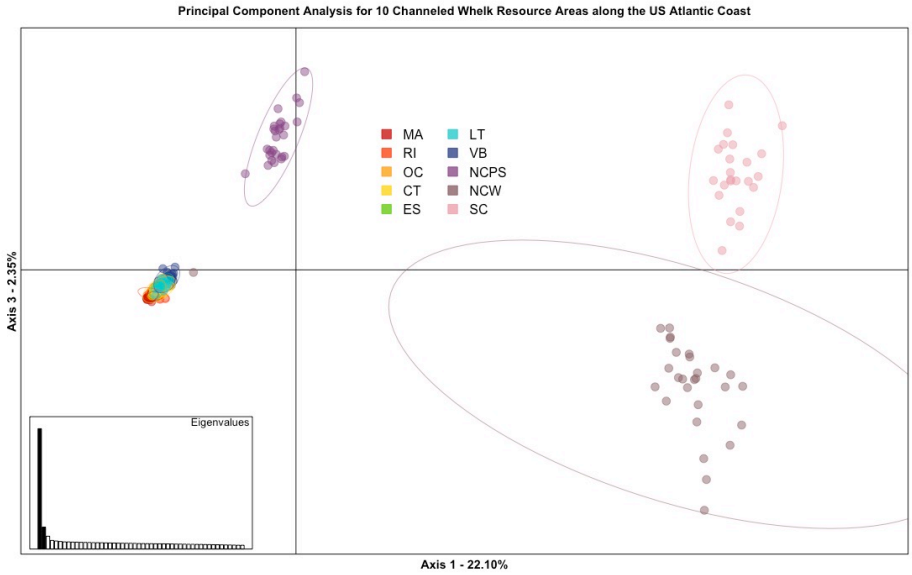
18ES017	34.7	2.08	40
18ES010	33.3	2.11	40
18ES004	33	2.09	40
18ES021	30.8	2.01	40
18ES027	29.8	2.03	60
18ES002	28.7	1.97	60
18ES018	24.4	1.99	60
18ES003	24.2	2.06	60
18ES011	18.8	2.24	60
18ES006	12.6	2.1	80
18ES007	5.7	2.45	80
18LT013	67.4	2.02	20
18LT012	61.6	2.01	20
18LT025	47	1.98	40
18LT011	42.1	2.05	40
18LT019	40.6	2.08	40
18LT005	40.3	2.14	40
18LT003	40.2	2.11	40
18LT023	39.7	2.05	40
18LT001	38.2	2.08	40
18LT004	37.5	2.04	40
18LT021	35.9	2.11	40
18LT017	35.8	2.01	40
18LT006	32.5	2.15	40
18LT007	32.5	2.19	40
18LT002	30.3	2.2	40
18LT022	29.3	2.15	40
18LT016	27.8	2.16	40
18LT010	27.5	2.1	40
18LT018	26.1	2.18	40
18LT024	25.7	2.14	40
18LT020	25.3	2.18	40
18LT009	20.3	2.21	80
18LT008	19.5	2.14	80
18LT015	18.9	2.26	80
18LT014	15.5	2.37	80
18VB002	47.4	1.93	40
18VB023	41.7	2	40
18VB013	41.3	2.07	40
18VB017	41.1	1.95	40

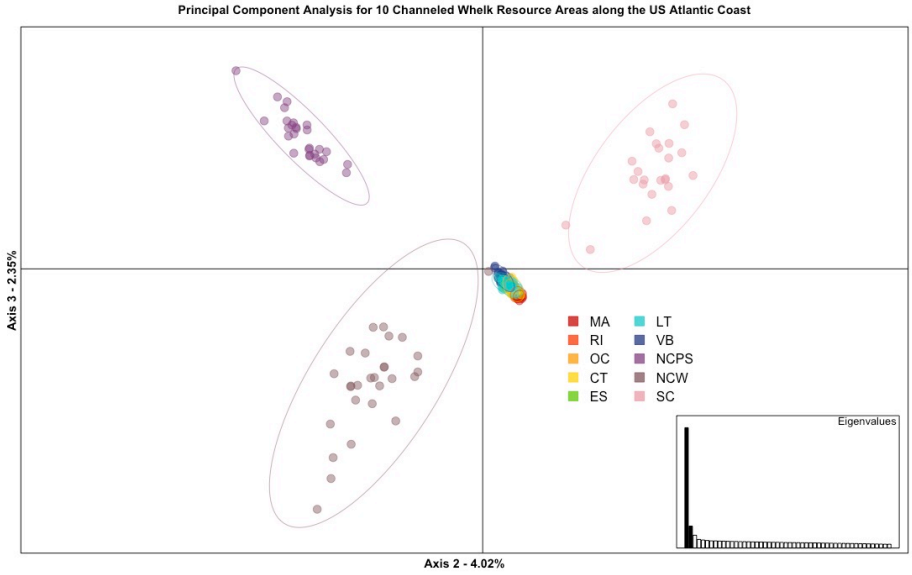
18VB026	39.7	2.02	40
18VB028	36.6	1.98	40
18VB025	36.3	2.07	40
18VB018	28.7	1.97	40
18VB007	27.8	1.99	40
18VB027	27.8	2.06	40
18VB021	26.4	1.97	40
18VB030	26	2.09	40
18VB019	25.5	2.04	40
18VB022	25.4	2.12	40
18VB001	22	2.19	60
18VB031	21.7	2.09	60
18VB016	20.8	2.04	60
18VB012	20.7	1.92	60
18VB020	20.7	2.14	60
18VB015	20.6	2	60
18VB024	20.5	2.08	60
18VB008	18.8	2.03	60
18VB014	18.2	2.13	60
18VB009	17.8	1.95	60
18VB010	17.8	1.98	60
18VB004	15	1.93	100
18VB003	14	2.04	100
18VB011	13.9	1.95	100
18VB006	10.8	2.05	100
18VB029	10.1	2.16	100
19NCPS017	1943.6	1.64	1
19NCPS018	1041.7	1.76	2
19NCPS021	917.6	1.79	2.2
19NCPS026	792	1.49	2.5
19NCPS028	649.8	1.75	2.8
19NCPS016	551.1	1.77	3.3
19NCPS019	272.2	1.87	6.6
19NCPS011	262.9	1.76	6.6
19NCPS009	262.8	1.66	6.6
19NCPS023	254.9	1.72	6.6
19NCPS002	186.3	1.86	10
19NCPS024	177.5	1.7	10
19NCPS013	136	1.96	10
19NCPS005	131.3	1.68	10

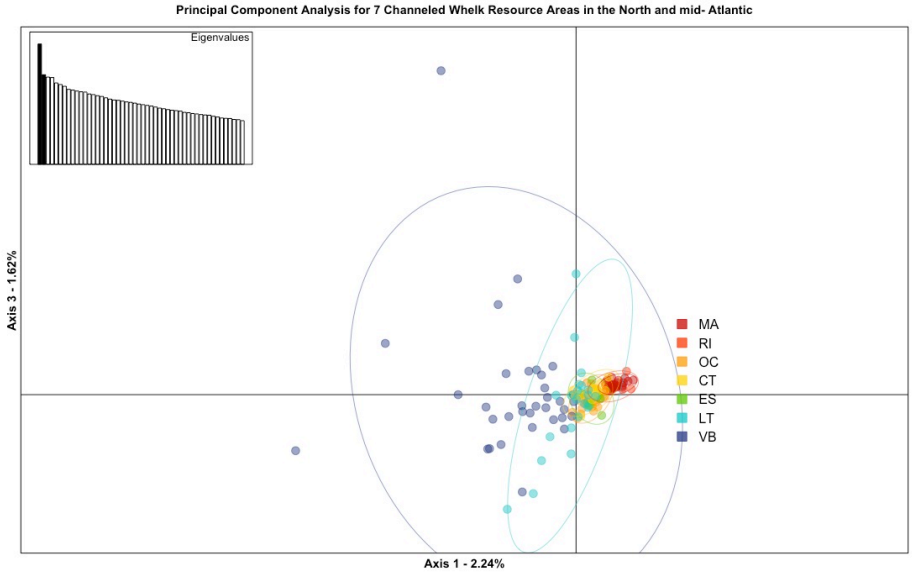
19NCPS015	130.8	2.08	10
19NCPS008	128.1	1.89	10
19NCPS027	99.4	1.95	20
19NCPS025	87.4	2.03	20
19NCPS014	86.3	2.17	20
19NCPS003	79.8	1.84	20
19NCPS010	75.6	1.8	20
19NCPS012	70.2	2.1	20
19NCPS007	65.5	2.15	20
19NCPS022	63.3	1.83	20
19NCPS029	61.3	2.02	20
19NCPS006	57.4	2.06	20
19NCPS030	54.7	2.09	20
19NCPS004	52.5	1.87	20
19NCPS020	47.7	2.13	40
19NCPS001	20.5	2.19	40
19NCW016	38	1.97	40
19NCW010	22.5	1.6	60
19NCW014	15.2	1.72	100
19NCW012	12.8	1.56	100
19NCW027	9.7	1.95	180
19NCW029	8.2	2.4	180
19NCW030	7.8	1.83	180
19NCW024	7.5	2.44	180
19NCW022	6.7	1.75	180
19NCW020	6.1	2.48	180
19NCW003	5	2.16	180
19NCW015	5	2.3	180
19NCW007	4.8	2.8	180
19NCW018	4.7	2.7	180
19NCW004	4.5	2.43	180
19NCW008	4.5	1.92	180
19NCW001	4.4	3.24	180
19NCW005	4.4	2.71	180
19NCW019	3.9	2.76	180
19NCW025	3.6	3.18	180
19NCW023	3.5	1.76	180
19NCW013	3.3	4.27	180
19NCW009	3.2	2.55	180
19NCW002	3.1	2.79	180

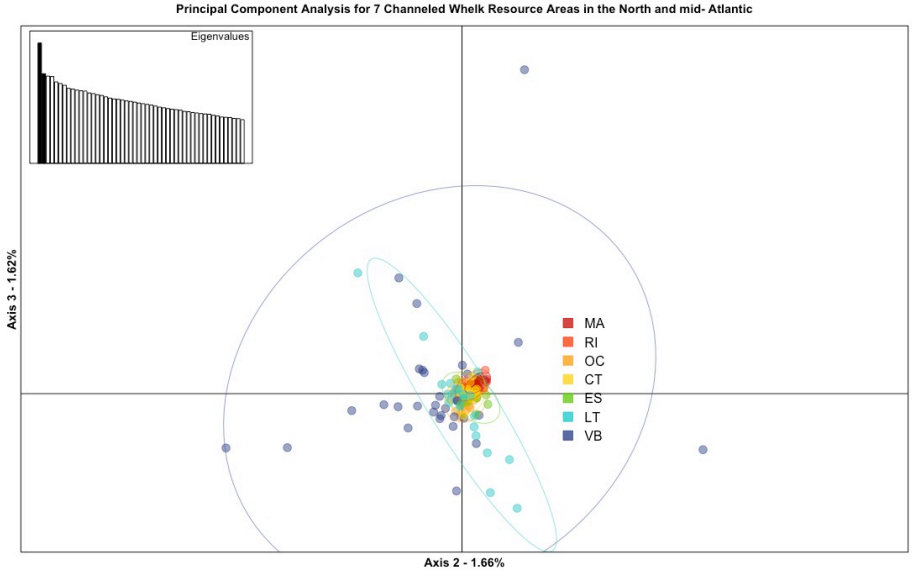
19NCW028	3	2.24	180
19NCW021	2.9	1.72	180
19NCW017	2.8	3.52	180
19NCW006	2.2	3.55	180
19NCW011	2.2	4.06	180
19NCW026	1.7	2.14	180
SC003	42.1	1.94	40
SC018	34.6	1.52	40
SC012	33	1.54	40
SC023	20.5	1.47	80
SC015	17.4	1.54	80
SC017	17.3	1.84	80
SC014	14.5	1.46	80
SC021	14.2	1.54	80
SC022	13.3	1.48	80
SC027	13.2	1.59	80
SC025	13	2.01	80
SC019	11.4	1.63	100
SC004	11.3	1.84	100
SC006	9.8	1.54	180
SC026	9.3	2.1	180
SC020	9.2	1.71	180
SC028	8.7	1.73	180
SC016	8.2	1.8	180
SC002	7.9	1.88	180
SC007	7.7	1.73	180
SC010	6.8	1.85	180
SC029	6.4	2.15	180
SC009	6.2	1.78	180
SC011	5.4	1.91	180
SC030	5	1.84	180
SC001	4.9	1.36	180
SC008	3.8	1.82	180
SC024	3.2	2.03	180
SC005	3.1	2.09	180
SC013	1	8.87	180

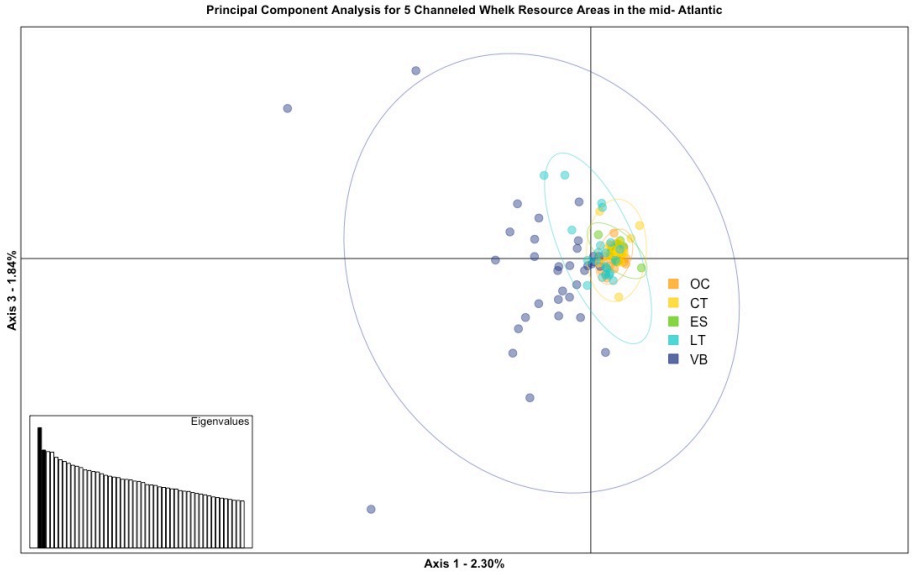
S.3 Supplemental PCAs from all, neutral, and outlier loci. This includes alternate axes for all PCAs used in this study.

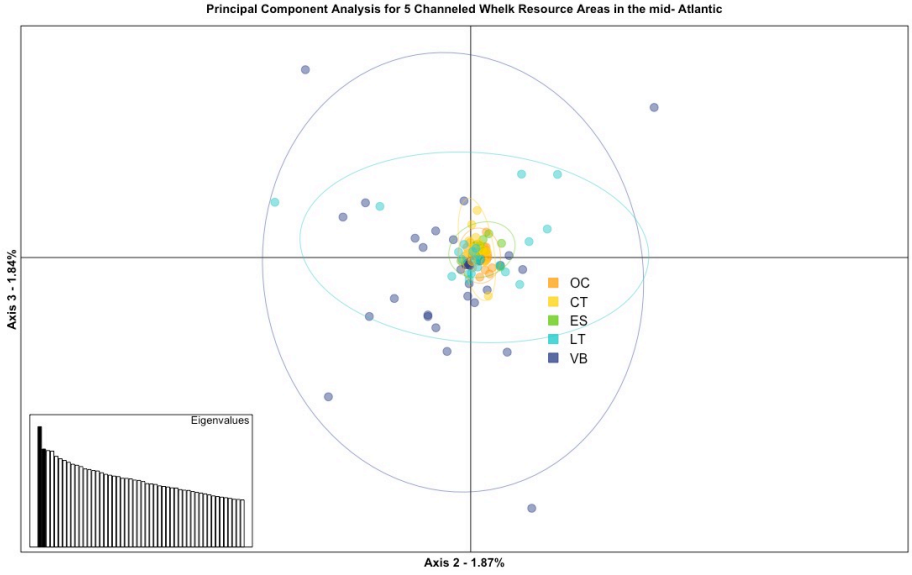




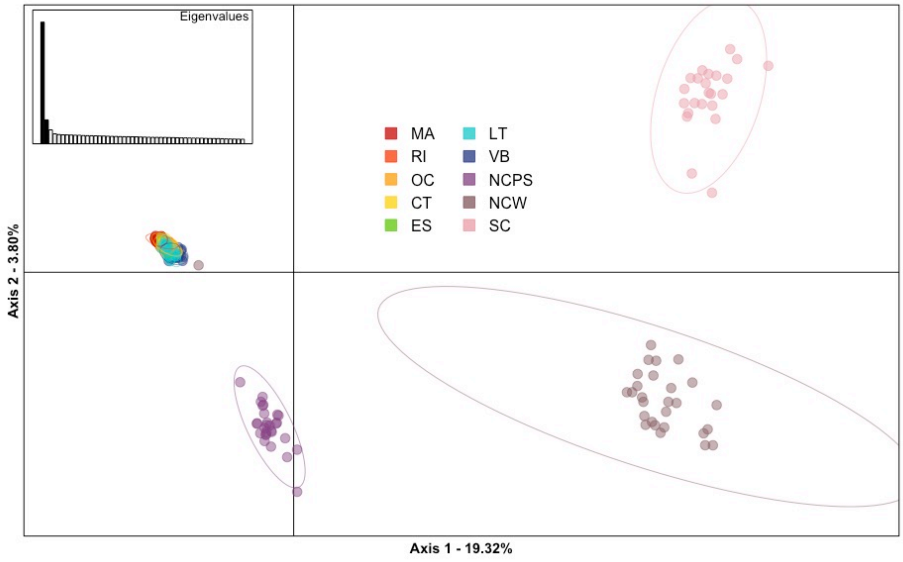

Principal Component Analysis for 7 Channeled Whelk Resource Areas in the North and mid- Atlantic
Eigenvalues
Axis 3 - 1.62%
Axis 1 - 2.24%
MA
RI
OC
CT
ES
LT
VB


Principal Component Analysis for 7 Channeled Whelk Resource Areas in the North and mid- Atlantic
Eigenvalues
Axis 3 - 1.62%
Axis 2 - 1.66%
MA
RI
OC
CT
ES
LT
VB

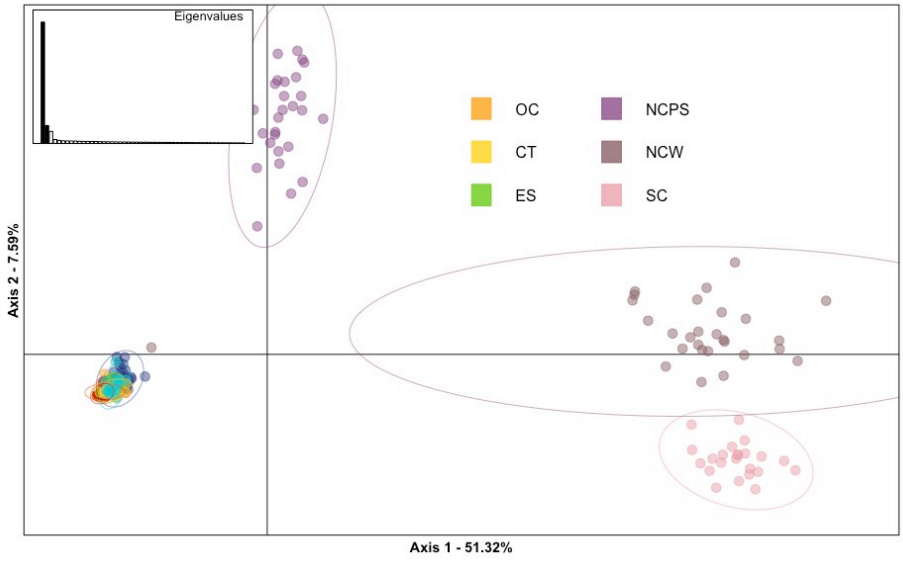

Principal Component Analysis for 5 Channeled Whelk Resource Areas in the mid- Atlantic
Axis 3 - 1.84%
Axis 1 - 2.30%
OC
CT
ES
LT
VB
Eigenvalues


Principal Component Analysis for 5 Channeled Whelk Resource Areas in the mid- Atlantic
Axis 3 - 1.84%
Axis 2 - 1.87%
OC
CT
ES
LT
VB
Eigenvalues

Principal Component Analysis for 10 Channeled Whelk Resource Areas along the US Atlantic Coast using neutral loci



Principal Component Analysis for 10 Channeled Whelk Resource Areas along the US Atlantic Coast using outlier loci



S.4 Metadata for microsatellite loci developed and used in this study. The primers that successfully amplified are bolded.

Primer	Forward 5' to 3'	Reverse 5' to 3'	Motif	Product Size
Bcan1	GGAAGAAAGGATGGACGGAC	TGCAATCCATCCATCCTCAC	ATCC	100
Bcan2	CTGTATGTCTGTCAGAACGCC	TCTGTCTCCCTATTCTTCGGC	ACAG	101
Bcan3	GAGGATAGGCGCCATGTAAG	CCGAGATTGATGCAGTATACCG	AGAT	100
Bcan4	**GGTTTACAAGTCGTCCCTCATG**	**AGGTAAAGCGCATTCTTGGTG**	**ACTC**	**100**
Bcan5	**GGATTGGAGGGAGACGGG**	**TCTCTGACTGAAGATAGGCCG**	**AATC**	**101**
Bcan6	CAAACAGACCGAACTACCTACC	GTAGGGAGGTGGGTATGTCG	ACCT	101
Bcan7	CTGCTCTGTGTGCTTGCG	TGATTTGGTGTAGACAGGACTG	ACGC	103
Bcan8	ACCTACCTACAGACGGACCG	AGTACGATGTGTTGAAGTGTGC	ACCT	100
Bcan9	CACTAGACCGGTCCTCAATG	ATTTGTACCCTTGTAGCACGTG	AATG	100
Bcan10	GTTTGCTGGTTGGTTGGCTC	AACCACCGAGTCAGCCAG	AACC	103
Bcan11	ACCTACCTACAGACGAACCG	GTCGTGTGGTAGGTAGGGC	ACCT	102
Bcan12	ACCTACCTACCTACAGACGG	GTGGGTCCGTTGTTTGAGAG	ACCT	102
Bcan13	GGTACCCAACTGACCGCTG	CGTGATTTGTGCAGGGATCG	AAAT	103
Bcan14	**GTTGAAAGGTGCTGTGATCAAC**	**CGCTCACACGCATGCATG**	**ACGC**	**101**
Bcan15	CAGACACAGAAACAGCAAGATG	TCTCTCTGGATCTTGCTGACAG	ACAG	101
Bcan16	**GGTTTACAAGTCGTCCCTCATG**	**AAGTATTCTCCAGGACTGCGC**	**ACTC**	**100**
Bcan17	AGGAAGAAACGCAGGCAGAC	ACGGATGGCTGGATGGAC	ACGG	100
Bcan18	ACTAAATTCTGGGCGCCC	GGCGCGGAGATTGGTATAAG	AACT	102
Bcan19	GTTGGTTGCTTGGACAAATGG	TACCAACCAACCGTACTCCC	AACC	102
Bcan20	GTGTACTCGTGGACAGACAG	CTGCCGGTCTGTCTGTCG	ACAG	100
Bcan21	TTTGCAATGCGGACACAGAC	AGACAGCCAGGACACAGTTG	ACAG	176
Bcan22	TCAGTTTGGCACGGATTTCTC	CTGGTGCGAGTGAGTGGTAG	ACAG	180
Bcan23	**GGCCTTTAACAACAGAGCCC**	**GTGTGTAACAGGCAGGAGTG**	**ACAG**	**187**
Bcan24	**TTTGTTGCTCTGCCTGTTTG**	**CAAGTAAGTGTGCAATTGAGGG**	**ACAG**	**193**
Bcan25	GCGGCGTCTCTCTCAGTATG	CTTGACTTCTTGACTCGGCG	ACAT	203
Bcan26	**ACAGCTGGAGGGATCTTGTC**	**TCGGACACTTCTGTAACTCCC**	**AAAC**	**204**
Bcan27	**CCAGCTCTCAAGAAATCCGTC**	**AGAGCTGAAGACCCTCCAAC**	**ACAT**	**211**
Bcan28	TGTTGCTCTGCCTGTTTGTC	CAAGTAAGTGTGCAATTGAGGG	ACAT	191
Bcan29	GCATGAAAGGCCGCAGTATATG	TTTAAGTCTTCGGCTGCCAC	AAAT	163
Bcan30	**TGCTGTACTTTGACAGGGAAC**	**TGCTGTACTTTGACAGGGAAC**	**AATC**	**170**
Bcan31	**ACAAGGACAGTGTGTGAGAC**	**CGTGCGAGTGTGTTACAAGG**	**AATC**	**180**
Bcan32	**TCCTGCGAGCATCCCTAATC**	**TTTGTCGGCAGCTTCAGTTG**	**ACCT**	**182**
Bcan33	**CCCTGTGTGCATGTGGAAC**	**ACCTTGTTGTCACTCGATCAG**	**AGAT**	**207**
Bcan34	**GTTGTGGCTGGATGTGAGATG**	**ACCTCTCTATTCGTCGGCAG**	**ACAT**	**129**
Bcan35	TCTCTGGTTATCTGTCGGTCTG	AGACAACTCAGTAGGCAGCG	ACAG	131
Bcan36	CGCGCGTAAATGTTAGTGTG	CGCGCGTAAATGTTAGTGTG	ACAT	136

Bcan37	**GCCAACATGTTACAGCTTGC**	**CCACTCAGGCCATATCAGGG**	**AATG**	**143**
Bcan38	**GATAAGGAGGAAGGCTTTGTGG**	**GGGTGTGTGGTGTATAAGAGAC**	**AAAG**	**152**
Bcan39	TCAATTGTGCAGAAACATGGTG	TGCCTAGCTTTATCCAGGAGTG	AG	130
Bcan40	GCCTGGTCGTATTCTAATGGC	TGTGTGTGTGAACTTTCGGC	AC	127
Bcan41	TATCCGAGCGTGCTTCTCTC	CGCGCGTGTGTGTCTTATAG	AC	161
Bcan42	ACTGAAACATGAGCAACTTTCC	AACATGAGCAACTTTCCACAC	AT	119
Bcan43	TGCGGGTAAGAAGCTAAGGAG	TGCGGGTAAGAAGCTAAGGAG	CG	200
Bcan44	ACCCAACGACACAATTCAGC	GCGGGTGAGGTGTGTAAGAG	AG	186
Bcan45	ACAGCGGGAGGAAAGGAAAC	GGACAGCGGGAGGAAAGG	AC	122
Bcan46	CCATCCTACACACGCGCG	CGCACCCACCCATCCTAC	CG	149
Bcan50	TCTTAACCCACATCTTAGCTGG	TCCTAACAGAACTAGCCGGC	ACCT	157
Bcan51	CAACAAACCAACCAACGTCC	CAACAAACCAACCAACGTCC	AGAT	144
Bcan52	**AAGTCGTCCCTCATGTCTCC**	**AAGTCGTCCCTCATGTCTCC**	**ACTC**	**140**
Bcan53	GGTAAGTAGGTCCGTCTGTAGG	GTCCGTCTGTAGGTAGGTCG	ACCT	137
Bcan54	GTGTCTCACCAAACACTGCG	ACTTGCCTGCAACTGAAACC	AAAC	154
Bcan55	GATCTGGATGGCGCTCTTTG	TGCAGACTCCTTAAAGAACCTC	ACTC	167
Bcan56	TCGCTCACTGTGTCCATGTG	TCGCTCACTGTGTCCATGTG	ACAG	174
Bcan57	CTTTCCTGGTTGAGTTTCCGG	TCCTGGTTTCTTTGTGTGCC	ACAG	178
Bcan58	**TGTCTCTCAACAGGGAATGC**	**AGAGTTACGGACAGCTGACG**	**ACAG**	**185**
Bcan59	**CTTCCGGCTGTCACTAAACC**	**TTAAACTGGCAGCGGTTGG**	**AAAC**	**191**
Bcan60	GAGAACTCTGTCGTGGCTAG	TGTTGTAACCGAAAGAAGCCC	AAAT	203
Bcan61	TTCTCTTCCAGACTCGTGCG	TCTCTTCCAGACTCGTGCG	AAAC	211
Bcan62	**ATACATGTCCGTCGGTCTGC**	**TGGGCGTGCGTACATATTAC**	**ACAT**	**203**
Bcan63	TGACAGACAATGAGTGATGGG	TGTTCTGACTTTAAACACACGC	ACGC	181
Bcan64	CCGATTGGTGCCTCGGTAC	TCTTCGTTTACCAGGTCCAGG	ACAG	169
Bcan65	**TAACCCAACGCGCTAGTCAG**	**TTCTCGTCCACTGCCGATAC**	**AGAT**	**150**

S.5 A text of the Rstudio code used for analyses.

```
################################
###   Sam Askin DART Data    ###
###     June 30, 2020        ###
################################

### Seting the Working Directory (Session -> Set Working Directory -> To Source File Location)
setwd("~/Desktop/ASKIN_THESIS_WHELK")

### Loading RData (for after filtering)
load("thesis.code.RData")


############################
### DArT Data Filtering ###
############################
install.packages("devtools")
library(devtools)
install.packages("BiocManager")
BiocManager::install(c("SNPRelate", "qvalue"))
install.packages("remotes")
remotes::install_github("green-striped-gecko/dartR", force = T)
library("dartR")

### Uploading DArT Files
gl <- gl.read.dart(filename = "Report_DBus19-4616_SNP_2.csv", ind.metafile = "metafile2.csv")
indNames(gl)
nPop(gl)
nLoc(gl)
popNames(gl)
levels(pop(gl))
table(pop(gl))
barplot(table(pop(gl)), las=2)

### Genlight Object Info
nInd(gl) # - returns the number of individuals in the genlight object
nLoc(gl) # - returns the number of loci
nPop(gl) # - returns the number of populations to which the individuals are assigned
indNames(gl) # - returns or sets labels for individuals
locNames(gl) # - returns or sets labels for loci
alleles(gl) # - returns or sets allelic states of each locus for each individual
ploidy(gl7hwe) # - returns or sets the ploidt of the individuals
pop(gl) # - returns or sets the population to which each individual belongs
levels(pop(gl)) # - unique list of population names
NA.posi(gl) # - returns loci with missing values
chr(gl) # - returs or sets chromosome for each locus

### Filtering Data (keep high quality loci, reduce noise), fill in the X with gl version
glX <- gl.filter.callrate(glX, method = "loc", threshold = 0.90) # - filter on call rate, at least 95% loci called
glX <- gl.filter.callrate(glX, method = "ind", threshold = 0.80) # - filter individuals on call rate, 90%
glX <- gl.filter.repavg(glX, threshold = 0.99, v=2) # - filter on reproducability, 100% reproducible
glX <- gl.filter.monomorphs(glX, v=1) # - filter out monomorphic loci and NAs
glX <- gl.filter.secondaries(glX,method= "best", v=2) # - filter out multiple SNP loci within a fragment
(secondaries) based on repeatability
glX <- gl.filter.maf(glX, threshold = 0.05) # - filter loci on minor allele freq less than 0.05
glX <- gl.filter.hwe(glX, alpha = 0.05, basis = "any", bon = T, v = 2) # - calculates the probabilities of
agreement with HWE
glX <- gl.filter.rdepth(glX, v=3) # - filters based on counts of sequence tags
glX <- gl.filter.heterozygosity(glX) # - filters based on observed heterozygosity
glX <- gl.filter.hamming() # - filters on number of base differences between two sequences
?gl.filter.X
```

```
### Genlight Filtering in dartR
gl.report.callrate(gl, method="ind", plot=TRUE) # - 90%
gl.report.callrate(gl, method="loc", plot=TRUE) # - 90%
gl2 <- gl.filter.callrate(gl, method = "loc", threshold = 0.90, v=3, plot=T)
gl3 <- gl.filter.callrate(gl2, method = "ind", threshold = 0.80, v=3, plot=T)
gl4 <- gl.filter.monomorphs(gl3, v=3) # - filter out monomorphic loci and NAs
gl.report.rdepth(gl4, plot=T) # - threshold between 5 - 30? could also reduce to 5 - 25
gl5 <- gl.filter.rdepth(gl4, lower = 5, upper = 25, v=3)
gl.report.repavg(gl5, plot=T) # - majority around 100%
gl6 <- gl.filter.repavg(gl5, threshold = 1, v=3)
gl7 <- gl.filter.hamming(gl6, v=3)
gl.report.callrate(gl7, method="ind", plot=TRUE) # -95%
gl8 <- gl.filter.callrate(gl7, method = "ind", threshold = 0.95, v=3, plot=T)
gl9 <- gl.filter.monomorphs(gl8, v=3) # - filter out monomorphic loci and NAsgl6 <-
gl.filter.secondaries(gl5,method= "random", v=3)
gl.report.callrate(gl9, "loc", plot=T) # - 96%
gl10 <- gl.filter.callrate(gl9, method = "loc", threshold = 0.96, v=3, plot=T)
gl.report.maf(gl10) # - threshold??
gl11 <- gl.filter.maf(gl10, threshold = 0.01, v=3)
gl12 <- gl.filter.secondaries(gl11,method= "random", v=3)

indNames(gl12)
nPop(gl12)
nLoc(gl12)
popNames(gl12)
levels(pop(gl12))
table(pop(gl12))
barplot(table(pop(gl12)), las=2)


### HWE filter in radiator
library(devtools)
devtools::install_github("thierrygosselin/radiator")
library(radiator)

### Converts genlight to tidy data format for radiator
tidygl12 <- tidy_genlight(gl12, gds=F, write=T, verbose=T)

### Selected filtering in radiator. Fill the X with tidy version
tidy.X <- filter_hwe(data=tidyglX) # - filters markers for hwe
tidy.12.2 <- filter_hwe(data=tidygl12) # - 2 threshold, 4 p-value, 2 blacklisted

### Go into folder radiator created from HWE filter. There will be a file of blacklisted markers for each p-value.
# - file name: blacklist.markers.hwd.0.0001.mid..value.2.hw.pop.threshold.tsv
# - Loci removed from filter:
#    CHROM1__100047245_12_C_T__12
#    CHROM1__100097038_35_G_A__35

### Now go to most recently filtered genlight object and make another genlight version with those loci removed
# - make sure to reformat Loci to correct form for dartR
library(dartR)
?gl.drop.loc
gl13 <- gl.drop.loc(gl12, loc.list=c("100047245-12-C/T", "100097038-35-G/A")) # - 2 blacklisted


### Outlier filtering in dartR
?gl.outflank()
gl14 <- gl.outflank(gl13)
gl.outlier <- write.csv(gl14$outflank$results$OutlierFlag, "gl.outlier.csv") # - no outlier loci
```

```
### Outlier filtering in pcadapt
install.packages("pcadapt")
library(pcadapt)
library(qvalue)
library(dartR)

df <- as.data.frame(gl13)
df[is.na(df)] <- 9 # replace NA with 9
write.table(df, file = "inputfile.outlier", row.names = F, col.names = F)
data <- read.pcadapt("inputfile.outlier", type = "lfmm") # LEA package uses same format
x <- pcadapt(input = data, K = 10)
plot(x, option = "screeplot")  # Here I choose 2.
plot(x , option = "manhattan")
x <- pcadapt(input = data, K = 4, min.maf = 0.01)
qval <- qvalue(x$pvalues)$qvalues
alpha <- 0.05
outliers <- which(qval < alpha)
outliers # 227 outliers for 0.01 maf
save(outliers, file = "outliers.0.01.Rdata")
gl.outlier <- gl.drop.loc(gl13, loc.list = locNames(gl13)[-outliers]) # drop neutral = keep outliers
gl.neutral <- gl.drop.loc(gl13, loc.list = locNames(gl13)[outliers]) # drop outlier = keep neutral
dim(gl.outlier$other$loc.metrics) # make sure number of loci has changed as well
dim(gl.neutral$other$loc.metrics)
dim(gl13$other$loc.metrics)

##########################################
### Final Filtered DArT Genlight Files ###
##########################################
gl13 # - strict HWE, 227 genotypes, 2,570 SNPs, 0.63% missing data
indNames(gl13)
nPop(gl13)
nLoc(gl13)
popNames(gl13)
levels(pop(gl13))
table(pop(gl13))
barplot(table(pop(gl13)), las=2)

### reports for genlight objects and changing population names
gl.report.callrate(gl13)
gl.report.heterozygosity(gl13)
gl.report.pa.pop(gl13)
gl.report.rdepth(gl13)
gl13$other$history
NewPop<- c("Chincoteague", "Eastern Shore", "Light Tower", "Massachusetts", "North Carolina Pam",
"North Carolina Wilm", "Ocean City", "Rhode Island", "South Carolina", "Virginia Beach")
popNames(gl13) <- NewPop
popNames(gl13)

gl.outlier # - 227 genotypes, 227 SNPs, 0.66% missing data
gl.outlier$other$history
NewPop<- c("Chincoteague", "Eastern Shore", "Light Tower", "Massachusetts", "North Carolina Pam",
"North Carolina Wilm", "Ocean City", "Rhode Island", "South Carolina", "Virginia Beach")
popNames(gl.outlier) <- NewPop
popNames(gl.outlier)

gl.neutral # - 227 genotypes, 2,343 SNPs, 0.63% missing data
gl.neutral$other$history
NewPop<- c("Chincoteague", "Eastern Shore", "Light Tower", "Massachusetts", "North Carolina Pam",
"North Carolina Wilm", "Ocean City", "Rhode Island", "South Carolina", "Virginia Beach")
popNames(gl.neutral) <- NewPop
popNames(gl.neutral)
```

```
##########################################
### Creating  Groupings of Data in dartR ###
##########################################
### All Populations
gl.all <- gl13
nPop(gl.all)
nLoc(gl.all)
indNames(gl.all)
popNames(gl.all)
gl.all
table(pop(gl.all))

### Mid-Atlantic Only Populations
gl.mid <- gl.keep.pop(gl.all, pop.list= c("Ocean City", "Chincoteague", "Eastern Shore", "Light Tower",
"Virginia Beach"))
nPop(gl.mid)
nLoc(gl.mid)
indNames(gl.mid)
popNames(gl.mid)
table(pop(gl.mid))

### Mid-Atlantic Minus VB
gl.mid.novb <- gl.keep.pop(gl.all, pop.list= c("Ocean City", "Chincoteague", "Eastern Shore", "Light Tower"))
nPop(gl.mid.novb)
nLoc(gl.mid.novb)
indNames(gl.mid.novb)
popNames(gl.mid.novb)
table(pop(gl.mid.novb))

### North and Mid- Atlantic
gl.midN <- gl.keep.pop(gl.all, pop.list=c("Massachusetts", "Rhode Island", "Ocean City","Chincoteague",
"Eastern Shore", "Light Tower", "Virginia Beach"))
nLoc(gl.midN)
indNames(gl.midN)
popNames(gl.midN)
table(pop(gl.midN))

### South and Mid-Atlantic
gl.midS <- gl.keep.pop(gl.all, pop.list=c("Ocean City","Chincoteague", "Eastern Shore", "Light Tower",
"Virginia Beach", "North Carolina Pam", "North Carolina Wilm", "South Carolina"))
nLoc(gl.midS)
indNames(gl.midS)
popNames(gl.midS)
table(pop(gl.midS))

### South Populations
gl.S <- gl.keep.pop(gl.all, pop.list=c("North Carolina Pam", "North Carolina Wilm", "South Carolina"))
nLoc(gl.S)
indNames(gl.S)
popNames(gl.S)
table(pop(gl.S))

### Population all male
gl.males <- gl.all[gl.all$other$ind.metrics$sex=="M", ]
gl.males
popNames(gl.males)
table(pop(gi.males))
gi.males <- gl2gi(gl.males)

### Population all female
gl.females <- gl.all[gl.all$other$ind.metrics$sex=="F", ]
```

```
gl.females
popNames(gl.females)
indNames(gi.females)
table(pop(gi.females))
gi.females <- gl2gi(gl.females)

### MA
gl.MA <- gl.keep.pop(gl.all, pop.list=c("Massachusetts"))
popNames(gl.MA)
indNames(gl.MA)
table(pop(gl.MA))
gi.MA <- gl2gi(gl.MA)

### RI
gl.RI <- gl.keep.pop(gl.all, pop.list=c("Rhode Island"))
popNames(gl.RI)
indNames(gl.RI)
table(pop(gl.RI))
gi.RI <- gl2gi(gl.RI)

### OC
gl.OC <- gl.keep.pop(gl.all, pop.list=c("Ocean City"))
popNames(gl.OC)
indNames(gl.OC)
table(pop(gl.OC))
gi.OC <- gl2gi(gl.OC)

### CT
gl.CT <- gl.keep.pop(gl.all, pop.list=c("Chincoteague"))
popNames(gl.CT)
indNames(gl.CT)
table(pop(gl.CT))
gi.CT <- gl2gi(gl.CT)

### ES
gl.ES <- gl.keep.pop(gl.all, pop.list=c("Eastern Shore"))
popNames(gl.ES)
indNames(gl.ES)
table(pop(gl.ES))
gi.ES <- gl2gi(gl.ES)

### LT
gl.LT <- gl.keep.pop(gl.all, pop.list=c("Light Tower"))
popNames(gl.LT)
indNames(gl.LT)
table(pop(gl.LT))
gi.LT <- gl2gi(gl.LT)

### VB
gl.VB <- gl.keep.pop(gl.all, pop.list=c("Virginia Beach"))
popNames(gl.VB)
indNames(gl.VB)
table(pop(gl.VB))
gi.VB <- gl2gi(gl.VB)

### NCPS
gl.NCPS <- gl.keep.pop(gl.all, pop.list=c("North Carolina Pam"))
popNames(gl.NCPS)
indNames(gl.NCPS)
table(pop(gl.NCPS))
gi.NCPS <- gl2gi(gl.NCPS)
```

```
### NCW
gl.NCW <- gl.keep.pop(gl.all, pop.list=c("North Carolina Wilm"))
popNames(gl.NCW)
indNames(gl.NCW)
table(pop(gl.NCW))
gi.NCW <- gl2gi(gl.NCW)

### SC
gl.SC <- gl.keep.pop(gl.all, pop.list=c("South Carolina"))
popNames(gl.SC)
indNames(gl.SC)
table(pop(gl.SC))
gi.SC <- gl2gi(gl.SC)

#######################################
### Isolation By Distance in dartR ###
#######################################
library(dartR)
dev.off()
?gl.ibd
par(col.main= "white")

### IBD for all 10 resource areas
glibd.all <- gl.ibd(gl.all, permutations= 999)
title(main="Isolation By Distance for 10 Channeled Whelk Resource Areas along the US Atlantic Coast",
col.main = "black")
par("usr")
text(11.5,.7, labels = "Mantel statistic r: 0.5693", col = "black", cex = 1, font = 2)
text(11.5,.6, labels = "Significance: 0.014", col = "black", cex = 1, font = 2)
glibd.all

### IBD for mid and North Atlantic resource areas
glibd.midN <- gl.ibd(gl.midN, permutations= 999)
title(main="Isolation By Distance for 7 Channeled Whelk Resource Areas in the North and mid- Atlantic",
col.main = "black")
par("usr")
text(11.5,.12, labels = "Mantel statistic r: 0.6111", col = "black", cex = 1, font = 2)
text(11.5,.11, labels = "Significance: 0.011", col = "black", cex = 1, font = 2)
glibd.midN

### IBD for mid Atlantic resource areas
glibd.mid <- gl.ibd(gl.mid, permutations= 999)
title(main="Isolation By Distance for 5 Channeled Whelk Resource Areas in the mid- Atlantic", col.main =
"black")
par("usr")
text(11.1,.013, labels = "Mantel statistic r: 0.4686", col = "black", cex = 1, font = 2)
text(11.1,.012, labels = "Significance: 0.075", col = "black", cex = 1, font = 2)
glibd.mid

### IBD for mid Atlantic minus VB resource areas
glibd.mid.novb <- gl.ibd(gl.mid.novb, permutations= 999)
title(main="Isolation By Distance for 4 Channeled Whelk Resource Areas in the mid- Atlantic", col.main =
"black")
par("usr")
text(11.3,.006, labels = "Mantel statistic r: .5498", col = "black", cex = 1, font = 2)
text(11.3,.005, labels = "Significance: 0.125", col = "black", cex = 1, font = 2)
glibd.mid.novb


#############################
### F Statistic in StAMPP ###
#############################
```

```
install.packages("StAMPP")
library(StAMPP)
?stamppConvert
?stamppFst

### Fst for all 10 Channeled Whelk resource areas, using 99% confidence intervals
gl.all.freq <- stamppConvert(gl.all, type = "genlight")
gl.all.fst.99 <- stamppFst(gl.all.freq, nboots = 10000, percent = 99, nclusters = 5)
gl.all.fst.99
Fsts.99 <- gl.all.fst.99$Fsts
P.vals.99 <- gl.all.fst.99$Pvalues
write.csv(Fsts.99,"~/Desktop/ASKIN_THESIS_WHELK//Fst.99.csv", row.names = FALSE)
write.csv(P.vals.99,"~/Desktop/ASKIN_THESIS_WHELK//P.vals.99.csv", row.names = FALSE)


############################
### PCA using Adegenet ###
############################
library(adegenet)
### Notes from Ellen Biesack
### Using Genind for Adegenet, converting genlight to genind
giX <- gl2gi(glX, v = 1)
### Reset all plot settings
dev.off()
### PCA
s.class(pca1$li, pop(gi8), col=transp(funky(15),.8), cpoint=1, cstar=0, axesell=F, label=NULL, grid=F,
cellipse = 0)
# col=transp(col,.6) uses your color object and makes the points transparent (change the number to make it
more or less transparent)
# cpoint=2 changes the size of your markers (you can also use pch=19 to change the marker shape - this is
a built-in R thing so you can look up the options)
# cstar=0 removes the lines from the center of the ellipses to the individual markers
# axesell=FALSE removes the "crosshairs" in your population ellipses
# label=NULL removes population labels (can also set label=popNames() or label=c("NC", "FL", "LA"))
# grid=F removes gridlines from plot background, although these may be helpful in determining the
dimensions of your plot if you want to adjust them
# cellipse = 0 removes the ellipses (just remove this part if you want to keep the ellipses)
### Adding legend to PCoA (first get the dimensions of your plot)
par("usr") # use these dimensions to figure out where your legend should go, and edit the first two options in
the below command for your x and y coordinates
legend(20,-1,c("CT","ES", "LT", "MA", "OC", "RI",
"VB"),fill=transp(funky(15),.8),border=transp(funky(15),.8),bty="n",pt.cex=.25,cex=.3) # pt.cex and cex
change the size of the points and font
### Adding eigenvalue barplot to the bottom left of the plot (if you want to change the position, just change it
to "bottomright" or "topleft", etc.)
add.scatter.eig(pca1$eig[1:20], 3,1,2, posi="bottomleft", sub="Eigenvalues", csub=1.0)
### Adjusting the dimensions of PCoA to zoom in and out using xlim and ylim
s.class(pca1$li, pop(gi5)) # use the par("usr") command to get the starting dimensions OR look at your plot
OR in the upper righthand corner you should see a d = # - that number is the width of the grids - count the
grids lengthwise and heightwise to get your dimensions
s.class(pca1$li, pop(gi7), xlim=c(-75,100), ylim=c(-100,50))
# remember that R will ignore some of your instructions if they would result in a plot that is a different
dimension from your window
### Getting Axis % Variation
max(pca1$eig)/sum(pca1$eig)*100
# for example, the first component (maximum eigenvalue) is 0.9823, so:
mtext("PC 1\n0.1429% of the variance",3,line=4)
# another way to do this is to sort the eigenvalues and pull out, for the example below, the second largest to
get your second component:
(sort(pca1$eig,T)[2])/sum(pca1$eig)*100
mtext("PC 2\n0.0266% of the variance",2,line=4)
### Make pretty PCoA
```

```
col <- c("aquamarine2","blue","red", "darkgoldenrod2")
s.class(pca1$li, pop(gi7), col=transp(funky(10),.5), cpoint=2, cstar=0, axesell=F, label=NULL, grid=F,
cellipse =3, xax=1, yax=3)
par("usr")
legend(-15,-.5, c("ES", "LT", "OC",
"VB"),fill=transp(funky(10),.8),border=transp(funky(10),.8),bty="n",pt.cex=.5,cex=.4)
add.scatter.eig(pca1$eig[1:17],col=heat.colors(50), posi="topleft", sub="Eigenvalues", csub=.3,xax=-20,
yax=-10)


### Notes for Channeled Whelk PCAs
### Color Assignment:
# - MA = red3
# - RI = orangered
# - OC = orange
# - CT = gold
# - ES = chartreuse3
# - LT = cyan3
# - VB = royalblue4
# - NCPS = orchid4
# - NCW = lightpink4
# - SC = lightpink2


### PCA for all 10 channeled whelk resource areas
dev.off()
par("usr")
gi.all <- gl2gi(gl.all, v = 1)
sum(is.na(gi.all$tab))
giXall <- scaleGen(gi.all, NA.method="mean")
class(giXall)
dim(giXall)
popNames(gi.all)
PopNames <- c("CT", "ES", "LT", "MA", "NCPS", "NCW", "OC", "RI", "SC", "VB") # - re-naming populations
for PCA
popNames(gi.all) <- PopNames
popNames(gi.all) # - this reorganized populations differently than assigned above. Make sure to check that #
of individuals per population match between gi and gl.
table(pop(gl.all))
table(pop(gi.all)) # new order: OC , VB, ES, RI, SC, NCPS, MA, NCW, CT, LT
dev.off()
pca1 <- dudi.pca(giXall,cent=F,scale=F,scannf=F,nf=3)
barplot(pca1$eig[1:50],main="PCA eigenvalues", col=heat.colors(50))
col2 <-c("orange", "royalblue4", "chartreuse3", "orangered", "lightpink2", "orchid4",  "red3", "lightpink4",
"gold", "cyan3")
par(oma=c(2,2,2,2))

# - Axis 1 and 2
s.class(pca1$li, pop(gi.all), col=transp(col2,.5), cpoint=2, cstar=0, axesell=F, label=NULL, grid=F, cellipse
=3, xax=1, yax=2)
par("usr")
col3 <-c("red3", "orangered", "orange", "gold", "chartreuse3", "cyan3", "royalblue4", "orchid4", "lightpink4",
"lightpink2")
legend(11,23, c("MA", "RI", "OC", "CT", "ES", "LT", "VB", "NCPS", "NCW", "SC"
),fill=transp(col3,.8),border=transp(col3,.8),bty="n",pt.cex=2,cex=1, ncol=2)
title(main="Principal Component Analysis for 10 Channeled Whelk Resource Areas along the US Atlantic
Coast", col.main = "black", outer = T, cex.main = .9)
max(pca1$eig)/sum(pca1$eig)*100 # Getting Axis % Variation
(sort(pca1$eig,T)[2])/sum(pca1$eig)*100 # % var explained by axis 2
mtext("Axis 1 - 22.10%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 1)
mtext("Axis 2 - 4.02%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 2)
add.scatter.eig(pca1$eig[1:50], posi="topleft", sub="Eigenvalues", csub=.3, xax=1, yax=2)
```

```
# - Axis 1 and 3
s.class(pca1$li, pop(gi.all), col=transp(col2,.5), cpoint=2, cstar=0, axesell=F, label=NULL, grid=F, cellipse
=3, xax=1, yax=3)
par("usr")
col3 <-c("red3", "orangered", "orange", "gold", "chartreuse3", "cyan3", "royalblue4", "orchid4", "lightpink4",
"lightpink2")
legend(11,23, c("MA", "RI", "OC", "CT", "ES", "LT", "VB", "NCPS", "NCW", "SC"
),fill=transp(col3,.8),border=transp(col3,.8),bty="n",pt.cex=2,cex=1, ncol=2)
title(main="Principal Component Analysis for 10 Channeled Whelk Resource Areas along the US Atlantic
Coast", col.main = "black", outer = T, cex.main = .9)
max(pca1$eig)/sum(pca1$eig)*100 # Getting Axis % Variation
(sort(pca1$eig,T)[3])/sum(pca1$eig)*100 # % var explained by axis 3
mtext("Axis 1 - 22.10%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 1)
mtext("Axis 3 - 2.35%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 2)
add.scatter.eig(pca1$eig[1:50], posi="bottomleft", sub="Eigenvalues", csub=.3, xax=1, yax=2)

# - Axis 2 and 3
s.class(pca1$li, pop(gi.all), col=transp(col2,.5), cpoint=2, cstar=0, axesell=F, label=NULL, grid=F, cellipse
=3, xax=2, yax=3)
par("usr")
col3 <-c("red3", "orangered", "orange", "gold", "chartreuse3", "cyan3", "royalblue4", "orchid4", "lightpink4",
"lightpink2")
legend(11,-5, c("MA", "RI", "OC", "CT", "ES", "LT", "VB", "NCPS", "NCW", "SC"
),fill=transp(col3,.8),border=transp(col3,.8),bty="n",pt.cex=2,cex=1, ncol=2)
title(main="Principal Component Analysis for 10 Channeled Whelk Resource Areas along the US Atlantic
Coast", col.main = "black", outer = T, cex.main = .9)
(sort(pca1$eig,T)[2])/sum(pca1$eig)*100 # % var explained by axis 2
(sort(pca1$eig,T)[3])/sum(pca1$eig)*100 # % var explained by axis 3
mtext("Axis 2 - 4.02%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 1)
mtext("Axis 3 - 2.35%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 2)
add.scatter.eig(pca1$eig[1:50], posi="bottomright", sub="Eigenvalues", csub=.3, xax=1, yax=2)

### PCA for mid Atlantic resource areas
gi.mid <- gl2gi(gl.mid, v = 1)
sum(is.na(gi.mid$tab))
giXmid <- scaleGen(gi.mid, NA.method="mean")
class(giXmid)
dim(giXmid)
popNames(gi.mid)
PopNames <- c("CT", "ES", "LT", "OC", "VB")
popNames(gi.mid) <- PopNames
popNames(gi.mid) # - this reorganized opulations differentlt than assigned above. Make sure to check that #
of individuals per population match between gi and gl.
table(pop(gl.mid))
table(pop(gi.mid)) # new order: OC , VB, ES, CT, LT
dev.off()
pca1 <- dudi.pca(giXmid,cent=F,scale=F,scannf=F,nf=3)
barplot(pca1$eig[1:50],main="PCA eigenvalues", col=heat.colors(50))
col2 <-c("orange", "royalblue4", "chartreuse3", "gold", "cyan3")
par(oma=c(2,2,2,2))

# - Axis 1 and 2
s.class(pca1$li, pop(gi.mid), col=transp(col2,.5), cpoint=2, cstar=0, axesell=F, label=NULL, grid=F, cellipse
=3, xax=1, yax=2)
par("usr")
col3 <- c("orange", "gold", "chartreuse3", "cyan3", "royalblue4")
legend(10,-1, c("OC", "CT", "ES", "LT", "VB"
),fill=transp(col3,.8),border=transp(col3,.8),bty="n",pt.cex=2,cex=1)
title(main="Principal Component Analysis for 5 Channeled Whelk Resource Areas in the mid- Atlantic",
col.main = "black", outer = T, cex.main = 0.9)
max(pca1$eig)/sum(pca1$eig)*100 # Getting Axis % Variation
```

```
(sort(pca1$eig,T)[2])/sum(pca1$eig)*100 # % var explained by axis 2
mtext("Axis 1 - 2.30%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 1)
mtext("Axis 2 - 1.87%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 2)
add.scatter.eig(pca1$eig[1:50], posi="bottomleft", sub="Eigenvalues", csub=.3, xax=1, yax=2)

# - Axis 1 and 3
s.class(pca1$li, pop(gi.mid), col=transp(col2,.5), cpoint=2, cstar=0, axesell=F, label=NULL, grid=F, cellipse
=3, xax=1, yax=3)
par("usr")
col3 <- c("orange", "gold", "chartreuse3", "cyan3", "royalblue4")
legend(10,-1, c("OC", "CT", "ES", "LT", "VB"
),fill=transp(col3,.8),border=transp(col3,.8),bty="n",pt.cex=2,cex=1)
title(main="Principal Component Analysis for 5 Channeled Whelk Resource Areas in the mid- Atlantic",
col.main = "black", outer = T, cex.main = 0.9)
max(pca1$eig)/sum(pca1$eig)*100 # Getting Axis % Variation
(sort(pca1$eig,T)[3])/sum(pca1$eig)*100 # % var explained by axis 3
mtext("Axis 1 - 2.30%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 1)
mtext("Axis 3 - 1.84%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 2)
add.scatter.eig(pca1$eig[1:50], posi="bottomleft", sub="Eigenvalues", csub=.3, xax=1, yax=2)

# - Axis 2 and 3
s.class(pca1$li, pop(gi.mid), col=transp(col2,.5), cpoint=2, cstar=0, axesell=F, label=NULL, grid=F, cellipse
=3, xax=2, yax=3)
par("usr")
col3 <- c("orange", "gold", "chartreuse3", "cyan3", "royalblue4")
legend(10,-1, c("OC", "CT", "ES", "LT", "VB"
),fill=transp(col3,.8),border=transp(col3,.8),bty="n",pt.cex=2,cex=1)
title(main="Principal Component Analysis for 5 Channeled Whelk Resource Areas in the mid- Atlantic",
col.main = "black", outer = T, cex.main = 0.9)
(sort(pca1$eig,T)[2])/sum(pca1$eig)*100 # % var explained by axis 2
(sort(pca1$eig,T)[3])/sum(pca1$eig)*100 # % var explained by axis 3
mtext("Axis 2 - 1.87%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 1)
mtext("Axis 3 - 1.84%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 2)
add.scatter.eig(pca1$eig[1:50], posi="bottomleft", sub="Eigenvalues", csub=.3, xax=1, yax=2)

### PCA for mid and North Atlantic resource areas
gi.midN <- gl2gi(gl.midN, v = 1)
sum(is.na(gi.midN$tab))
giXmidN <- scaleGen(gi.midN, NA.method="mean")
class(giXmidN)
dim(giXmidN)
popNames(gi.midN)
PopNames <- c("CT", "ES", "LT", "MA", "OC", "RI", "VB")
popNames(gi.midN) <- PopNames
popNames(gi.midN) # - this reorganized opulations differentlt than assigned above. Make sure to check that
# of individuals per population match between gi and gl.
table(pop(gl.midN))
table(pop(gi.midN)) # new order: OC , VB, ES, RI, MA, CT, LT
dev.off()
pca1 <- dudi.pca(giXmidN,cent=F,scale=F,scannf=F,nf=3)
barplot(pca1$eig[1:50],main="PCA eigenvalues", col=heat.colors(50))
col2 <-c("orange", "royalblue4", "chartreuse3", "orangered", "red3", "gold", "cyan3")
par(oma=c(2,2,2,2))

# - Axis 1 and 2
s.class(pca1$li, pop(gi.midN), col=transp(col2,.5), cpoint=2, cstar=0, axesell=F, label=NULL, grid=F, cellipse
=3, xax=1, yax=2)
par("usr")
col3 <-c("red3", "orangered", "orange", "gold", "chartreuse3", "cyan3", "royalblue4")
legend(15,15, c("MA", "RI", "OC", "CT", "ES", "LT",
"VB"),fill=transp(col3,.8),border=transp(col3,.8),bty="n",pt.cex=2,cex=1)
```

```
title(main="Principal Component Analysis for 7 Channeled Whelk Resource Areas in the North and mid-Atlantic", col.main = "black", outer = T, cex.main = .9)
max(pca1$eig)/sum(pca1$eig)*100 # Getting Axis % Variation
(sort(pca1$eig,T)[2])/sum(pca1$eig)*100 # % var explained by axis 2
mtext("Axis 1 - 2.24%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 1)
mtext("Axis 2 - 1.66%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 2)
add.scatter.eig(pca1$eig[1:50], posi="topleft", sub="Eigenvalues", csub=.3, xax=1, yax=2)

# - Axis 1 and 3
s.class(pca1$li, pop(gi.midN), col=transp(col2,.5), cpoint=2, cstar=0, axesell=F, label=NULL, grid=F, cellipse=3, xax=1, yax=3)
par("usr")
col3 <-c("red3", "orangered", "orange", "gold", "chartreuse3", "cyan3", "royalblue4")
legend(15,15, c("MA", "RI", "OC", "CT", "ES", "LT",
"VB"),fill=transp(col3,.8),border=transp(col3,.8),bty="n",pt.cex=2,cex=1)
title(main="Principal Component Analysis for 7 Channeled Whelk Resource Areas in the North and mid-Atlantic", col.main = "black", outer = T, cex.main = .9)
max(pca1$eig)/sum(pca1$eig)*100 # Getting Axis % Variation
(sort(pca1$eig,T)[3])/sum(pca1$eig)*100 # % var explained by axis 3
mtext("Axis 1 - 2.24%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 1)
mtext("Axis 3 - 1.62%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 2)
add.scatter.eig(pca1$eig[1:50], posi="topleft", sub="Eigenvalues", csub=.3, xax=1, yax=2)

# - Axis 2 and 3
s.class(pca1$li, pop(gi.midN), col=transp(col2,.5), cpoint=2, cstar=0, axesell=F, label=NULL, grid=F, cellipse=3, xax=2, yax=3)
par("usr")
col3 <-c("red3", "orangered", "orange", "gold", "chartreuse3", "cyan3", "royalblue4")
legend(15,15, c("MA", "RI", "OC", "CT", "ES", "LT",
"VB"),fill=transp(col3,.8),border=transp(col3,.8),bty="n",pt.cex=2,cex=1)
title(main="Principal Component Analysis for 7 Channeled Whelk Resource Areas in the North and mid-Atlantic", col.main = "black", outer = T, cex.main = .9)
(sort(pca1$eig,T)[2])/sum(pca1$eig)*100 # % var explained by axis 2
(sort(pca1$eig,T)[3])/sum(pca1$eig)*100 # % var explained by axis 3
mtext("Axis 2 - 1.66%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 1)
mtext("Axis 3 - 1.62%", col = "black", line = 0, cex = .9, font = 2, outer = T, side = 2)
add.scatter.eig(pca1$eig[1:50], posi="topleft", sub="Eigenvalues", csub=.3, xax=1, yax=2)

#########################
### DAPC in adegenet ###
#########################
library(adegenet)
?find.clusters
### first example is with all 10 channeled whelk resource areas
# -  we have 10 sampling locations, so max.n.clust = 10. Keep as many PCs as possible
# - choose K based off the "knee"
par(mfrow=c(1,1))

### DAPC for all 10 channeled whelk resource area
grp.all <- find.clusters(gi.all, max.n.clust = 10, n.iter = 1000000, stat = "BIC")
grp.all$Kstat
grp.all$grp # - Kstat = 4, compare original pop assignment to inferred assignment in table and plot, check for known biogeographic barriers and Fst values
table(pop(gi.all), grp.all$grp)
table.value(table(pop(gi.all), grp.all$grp), col.lab=paste("inferred", 1:6),
          row.lab=paste(levels(pop(gi.all))))
# - DAPC, applying group assignment to DAPC analyses
# - dont want too many or too few PCs because discrimination too high or low, 80% variance good, common to keep 3 discriminate functions, PC1, PC2, PC3
dapc.grp.all <- dapc(gi.all, grp.all$grp)

### DAPC for mid Atlantic channeled whelk resource areas
```

```
grp.mid <- find.clusters(gi.mid, max.n.clust = 5, n.iter = 1000000, stat = "BIC")
grp.mid$Kstat #2
table(pop(gi.mid), grp.mid$grp)
table.value(table(pop(gi.mid), grp.mid$grp), col.lab=paste("inferred", 1:2),
          row.lab=paste(levels(pop(gi.mid))))
dapc.grp.mid <- dapc(gi.mid, grp.mid$grp)

### DAPC for mid and North Atlantic channeled whelk resource areas
grp.midN <- find.clusters(gi.midN, max.n.clust = 7, n.iter = 1000000, stat = "BIC")
grp.midN$Kstat #3
table(pop(gi.midN), grp.midN$grp)
table.value(table(pop(gi.midN), grp.midN$grp), col.lab=paste("inferred", 1:4),
          row.lab=paste(levels(pop(gi.midN))))
dapc.grp.midN <- dapc(gi.midN, grp.midN$grp)


################################
### Plotting DAPC in GGplot ###
################################
library(ggplot2)
library(reshape2) # - there was some weirdness with this being discontinued but it still seemed to work?
# - first example is with all 10 channeled whelk resource areas

### DAPC Plot forall 10 channeled whelk resource areas
# - create dataframe
dapc.results.all <- as.data.frame(dapc.grp.all$posterior)
dapc.results.all$pop <- pop(gi.all)
dapc.results.all$pop
dapc.results.all$indNames <- indNames(gi.all)
head(dapc.results.all)
# - get dataframe in ggplot format
dapc.results.all <- reshape2::melt(dapc.results.all)
head(dapc.results.all)
colnames(dapc.results.all) <- c("gi_Pop", "Sample", "Assigned_Population",
"Posterior_Membership_Probability")
head(dapc.results.all)
dapc.results.all$gl_Pop2 <- factor(dapc.results.all$gi_Pop, levels = c("MA", "RI", "OC", "CT", "ES", "LT",
"VB", "NCPS", "NCW", "SC"))
cbPalette <- c("#999999", "#E69F00", "#56B4E9", "#009E73", "#F0E442","#CC79A7", "#0072B2",
"#D55E00", "black")
dapc.results.all$Assigned_Population
dev.off()
# - structure plot
structure <- ggplot(dapc.results.all, aes(x=Sample, y=Posterior_Membership_Probability,
fill=Assigned_Population)) +
  geom_bar(stat = "identity", width = 0.7) +
  xlab("Samples") +
  ylab("Posterior Membership Probability") +
  ggtitle("DAPC Results for 10 US Atlantic Channeled Whelk Resource Areas") +
  theme(plot.title = element_text(hjust = 0.5)) +
  scale_fill_manual(values = cbPalette) +
  scale_y_continuous(expand = c(0, 0)) +
  facet_grid(~gl_Pop2, scales = "free", space = "free") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, size = 8)) +
  scale_x_discrete(guide = guide_axis(n.dodge = 3)) +
  guides(fill=guide_legend(title="Assigned_Population"))
structure

### DAPC Plot for mid Atlantic channeled whelk resource areas
dapc.results.mid <- as.data.frame(dapc.grp.mid$posterior)
dapc.results.mid$pop <- pop(gi.mid)
dapc.results.mid$pop
```

```
dapc.results.mid$indNames <- indNames(gi.mid)
head(dapc.results.mid)
dapc.results.mid <- reshape2::melt(dapc.results.mid)
head(dapc.results.mid)
colnames(dapc.results.mid) <- c("gi_Pop", "Sample", "Assigned_Population",
"Posterior_Membership_Probability")
head(dapc.results.mid)
dapc.results.mid$gl_Pop2 <- factor(dapc.results.mid$gi_Pop, levels = c("OC", "CT", "ES", "LT", "VB"))
structure.mid <- ggplot(dapc.results.mid, aes(x=Sample, y=Posterior_Membership_Probability,
fill=Assigned_Population)) +
  geom_bar(stat = "identity", width = 0.7) +
  xlab("Samples") +
  ylab("Posterior Membership Probability") +
  ggtitle("DAPC Results for 5 Mid Atlantic Channeled Whelk Resource Areas") +
  theme(plot.title = element_text(hjust = 0.5)) +
  scale_fill_manual(values = cbPalette) +
  scale_y_continuous(expand = c(0, 0)) +
  facet_grid(~gl_Pop2, scales = "free", space = "free") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, size = 8)) +
  scale_x_discrete(guide = guide_axis(n.dodge = 3)) +
  guides(fill=guide_legend(title="Assigned_Population"))
structure.mid

### DAPC Plot for mid and North Atlantic channeled whelk resource areas
dapc.results.midN <- as.data.frame(dapc.grp.midN$posterior)
dapc.results.midN$pop <- pop(gi.midN)
dapc.results.midN$pop
dapc.results.midN$indNames <- indNames(gi.midN)
head(dapc.results.midN)
dapc.results.midN <- reshape2::melt(dapc.results.midN)
head(dapc.results.midN)
colnames(dapc.results.midN) <- c("gi_Pop", "Sample", "Assigned_Population",
"Posterior_Membership_Probability")
head(dapc.results.midN)
dapc.results.midN$gl_Pop2 <- factor(dapc.results.midN$gi_Pop, levels = c("MA", "RI", "OC", "CT", "ES",
"LT", "VB"))
structure.midN <- ggplot(dapc.results.midN, aes(x=Sample, y=Posterior_Membership_Probability,
fill=Assigned_Population)) +
  geom_bar(stat = "identity", width = 0.7) +
  xlab("Samples") +
  ylab("Posterior Membership Probability") +
  ggtitle("DAPC Results for 7 Mid Atlantic Channeled Whelk Resource Areas") +
  theme(plot.title = element_text(hjust = 0.5)) +
  scale_fill_manual(values = cbPalette) +
  scale_y_continuous(expand = c(0, 0)) +
  facet_grid(~gl_Pop2, scales = "free", space = "free") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, size = 8)) +
  scale_x_discrete(guide = guide_axis(n.dodge = 3)) +
  guides(fill=guide_legend(title="Assigned_Population"))
structure.midN


###################################
### Genepop Files for all data ###
###################################
library(radiator)
?genomic_converter

### all
INDIVIDUALS <- gl.all$ind.names
STRATA <- gl.all$other$ind.metrics$pop
TARGET_ID <- gl.all$ind.names
```

```
df.all <- data.frame(INDIVIDUALS, STRATA, TARGET_ID)
write.csv(df.all,"~/Desktop/ASKIN_THESIS_WHELK//all.strata.csv", row.names = FALSE)
# - convert csv to tab delimited file.
genomic_converter(gi.all, strata = "all.strata.txt", output = "genepop", filename = "genepop.all")

### MA
INDIVIDUALS <- gl.MA$ind.names
STRATA <- gl.MA$other$ind.metrics$pop
TARGET_ID <- gl.MA$ind.names
df.MA <- data.frame(INDIVIDUALS, STRATA, TARGET_ID)
write.csv(df.MA,"~/Desktop/ASKIN_THESIS_WHELK//MA.strata.csv", row.names = FALSE)
# - convert csv to tab delimited file.
genomic_converter(gi.MA, strata = "MA.strata.txt", output = "genepop", filename = "genepop.MA")

### RI
INDIVIDUALS <- gl.RI$ind.names
STRATA <- gl.RI$other$ind.metrics$pop
TARGET_ID <- gl.RI$ind.names
df.RI <- data.frame(INDIVIDUALS, STRATA, TARGET_ID)
write.csv(df.RI,"~/Desktop/ASKIN_THESIS_WHELK//RI.strata.csv", row.names = FALSE)
# - convert csv to tab delimited file.
genomic_converter(gi.RI, strata = "RI.strata.txt", output = "genepop", filename = "genepop.RI")

### OC
INDIVIDUALS <- gl.OC$ind.names
STRATA <- gl.OC$other$ind.metrics$pop
TARGET_ID <- gl.OC$ind.names
df.OC <- data.frame(INDIVIDUALS, STRATA, TARGET_ID)
write.csv(df.OC,"~/Desktop/ASKIN_THESIS_WHELK//OC.strata.csv", row.names = FALSE)
# - convert csv to tab delimited file.
genomic_converter(gi.OC, strata = "OC.strata.txt", output = "genepop", filename = "genepop.OC")

### CT
INDIVIDUALS <- gl.CT$ind.names
STRATA <- gl.CT$other$ind.metrics$pop
TARGET_ID <- gl.CT$ind.names
df.CT <- data.frame(INDIVIDUALS, STRATA, TARGET_ID)
write.csv(df.CT,"~/Desktop/ASKIN_THESIS_WHELK//CT.strata.csv", row.names = FALSE)
# - convert csv to tab delimited file.
genomic_converter(gi.CT, strata = "CT.strata.txt", output = "genepop", filename = "genepop.CT")

### ES
INDIVIDUALS <- gl.ES$ind.names
STRATA <- gl.ES$other$ind.metrics$pop
TARGET_ID <- gl.ES$ind.names
df.ES <- data.frame(INDIVIDUALS, STRATA, TARGET_ID)
write.csv(df.ES,"~/Desktop/ASKIN_THESIS_WHELK//ES.strata.csv", row.names = FALSE)
# - convert csv to tab delimited file.
genomic_converter(gi.ES, strata = "ES.strata.txt", output = "genepop", filename = "genepop.ES")

### LT
INDIVIDUALS <- gl.LT$ind.names
STRATA <- gl.LT$other$ind.metrics$pop
TARGET_ID <- gl.LT$ind.names
df.LT <- data.frame(INDIVIDUALS, STRATA, TARGET_ID)
write.csv(df.LT,"~/Desktop/ASKIN_THESIS_WHELK//LT.strata.csv", row.names = FALSE)
# - convert csv to tab delimited file.
genomic_converter(gi.LT, strata = "LT.strata.txt", output = "genepop", filename = "genepop.LT")

### VB
INDIVIDUALS <- gl.VB$ind.names
STRATA <- gl.VB$other$ind.metrics$pop
```

```
TARGET_ID <- gl.VB$ind.names
df.VB <- data.frame(INDIVIDUALS, STRATA, TARGET_ID)
write.csv(df.VB,"~/Desktop/ASKIN_THESIS_WHELK//VB.strata.csv", row.names = FALSE)
# - convert csv to tab delimited file.
genomic_converter(gi.VB, strata = "VB.strata.txt", output = "genepop", filename = "genepop.VB")

### NCPS
INDIVIDUALS <- gl.NCPS$ind.names
STRATA <- gl.NCPS$other$ind.metrics$pop
TARGET_ID <- gl.NCPS$ind.names
df.NCPS <- data.frame(INDIVIDUALS, STRATA, TARGET_ID)
write.csv(df.NCPS,"~/Desktop/ASKIN_THESIS_WHELK//NCPS.strata.csv", row.names = FALSE)
# - convert csv to tab delimited file.
genomic_converter(gi.NCPS, strata = "NCPS.strata.txt", output = "genepop", filename = "genepop.NCPS")

### NCW
INDIVIDUALS <- gl.NCW$ind.names
STRATA <- gl.NCW$other$ind.metrics$pop
TARGET_ID <- gl.NCW$ind.names
df.NCW <- data.frame(INDIVIDUALS, STRATA, TARGET_ID)
write.csv(df.NCW,"~/Desktop/ASKIN_THESIS_WHELK//NCW.strata.csv", row.names = FALSE)
# - convert csv to tab delimited file.
genomic_converter(gi.NCW, strata = "NCW.strata.txt", output = "genepop", filename = "genepop.NCW")

### SC
INDIVIDUALS <- gl.SC$ind.names
STRATA <- gl.SC$other$ind.metrics$pop
TARGET_ID <- gl.SC$ind.names
df.SC <- data.frame(INDIVIDUALS, STRATA, TARGET_ID)
write.csv(df.SC,"~/Desktop/ASKIN_THESIS_WHELK//SC.strata.csv", row.names = FALSE)
# - convert csv to tab delimited file.
genomic_converter(gi.SC, strata = "SC.strata.txt", output = "genepop", filename = "genepop.SC")

### MALES
INDIVIDUALS <- gl.males$ind.names
STRATA <- gl.males$other$ind.metrics$pop
TARGET_ID <- gl.males$ind.names
df.males <- data.frame(INDIVIDUALS, STRATA, TARGET_ID)
write.csv(df.males,"~/Desktop/ASKIN_THESIS_WHELK//males.strata.csv", row.names = FALSE)
# - convert csv to tab delimited file.
genomic_converter(gi.males, strata = "males.strata.txt", output = "genepop", filename = "genepop.males")

### FEMALES
INDIVIDUALS <- gl.females$ind.names
STRATA <- gl.females$other$ind.metrics$pop
TARGET_ID <- gl.females$ind.names
df.females <- data.frame(INDIVIDUALS, STRATA, TARGET_ID)
write.csv(df.females,"~/Desktop/ASKIN_THESIS_WHELK//females.strata.csv", row.names = FALSE)
# - convert csv to tab delimited file.
genomic_converter(gi.females, strata = "females.strata.txt", output = "genepop", filename =
"genepop.females")


#####################################################
### Ne Estimator Tutorial for Mac Explanation ###
#####################################################
library(radiator)
?genomic_converter
# - edit all strata so that mid-Atlantic minus VB are one population. Call this strata Ne. Should be 7 strata.
genomic_converter(gi.all, strata = "Ne.strata.txt", output = "genepop", filename = "genepop.all.Ne")

# - Ne is the estimation of how much of the population is reproductive/ cntributing alleles to next generation
```

```
# - genetic drift, more individuals contributing but lesser diversity amongst individuals
# - assumptions of Ne to keep in mind equal male and female, no overlapping generations, etc. (read up on this)
# - Ne Estimator window -> linkage with random, heterozygote excess, coancestry, temporal (0,1.5) in methods
# - What is best given my available data.
# - uncheck temporal (not relevant to samples)
# - Linkage disequilibrium -> can be linked physically or linked because non-random mating or inbreeding
# - Heterozygote excess -> how population if small wil have a def of homo and excess of het because of drift (can be transient, look for evidence of bottlenecks -> easy to lose rare homozygotes )
# - Molecular coancestry -> looking for alleles that are identical by descent (inherited from same ancestor) and if more, then probably small generation
# - file format is genepop


###################################
### divMigrate-Online Tutorial ###
###################################
# - https://popgen.shinyapps.io/divMigrate-online/
# - read in a genepop file, chose the statistic you want to use to calculate relative migration
# - download and save results matrix and network plot
library(diveRsity)
?divMigrate


###################################
### Genalex Files for all data ###
###################################
library("dartR")
?gl2genalex()

genalex.females <- gl2genalex(gl.females, outfile ="genalex.females.csv")
genalex.males <- gl2genalex(gl.males, outfile = "genalex.males.csv")


################################################
### Calculating Ho and He to verify Genalex ###
################################################
library(adegenet)
library(hierfstat)
library(pegas)
div <- summary(gi.MA)
mean(div$Hobs)
mean(div$Hexp)
div <- summary(gi.RI)
mean(div$Hobs)
mean(div$Hexp)
div <- summary(gi.OC)
mean(div$Hobs)
mean(div$Hexp)
div <- summary(gi.CT)
mean(div$Hobs)
mean(div$Hexp)
div <- summary(gi.ES)
mean(div$Hobs)
mean(div$Hexp)
div <- summary(gi.LT)
mean(div$Hobs)
mean(div$Hexp)
div <- summary(gi.VB)
mean(div$Hobs)
mean(div$Hexp)
```

```
div <- summary(gi.NCPS)
mean(div$Hobs)
mean(div$Hexp)
div <- summary(gi.NCW)
mean(div$Hobs)
mean(div$Hexp)
div <- summary(gi.SC)
mean(div$Hobs)
mean(div$Hexp)



######################
### Map in GGplot ###
######################
setwd("~/Desktop/ASKIN_THESIS_WHELK")
lat.lon <- read.csv(file = 'lat.lon.csv')
is.data.frame(lat.lon)

dev.off()
states <- map_data("state")
east_coast <- subset(states, region %in% c("massachusetts", "new jersey", "rhode island",
"maryland","pennsylvania", "connecticut", "new york", "virginia", "north carolina", "south carolina"))

map <- ggplot() +
  geom_polygon(data = east_coast,aes(x = long, y = lat, group = group), fill = "lightgray", color = "black")+
  coord_fixed(1)+
  geom_point(data=lat.lon, aes(x=lat.lon$lon, y=lat.lon$lat, color = as.factor(lat.lon$Resource.Area)), size =
5)+
  labs(x = "Longitude", y = "Latitude")+
  scale_color_manual(values=c("red3", "orangered", "orange", "gold", "chartreuse3", "cyan3", "royalblue4",
"orchid4", "lightpink4", "lightpink2", "deepskyblue4"),
                     name="Locations",
                     breaks= lat.lon$Resource.Area,
                     labels=c("Buzzard Bay, MA (MA)", "Rhode Island (RI)", "Ocean City, MD (OC)", "Chincoteague,
VA (CT)", "Hog Island, VA (ES)", "Light Tower, VA (LT)", "Sand Bridge, VA (VB)", "Pamlico Sound, North
Carolina (NCPS)", "Wilmongton, North Carolina (NCW)", "Charleston, South Carolina (SC)", "Buckroe
Beach, Virginia (BB)"))
map
```

LITERATURE CITED

Ackiss, A. S., C. E. Bird, Y. Akita, M. D. Santos, K. Tachihara, and K. E. Carpenter. 2018. Genetic patterns in peripheral marine populations of the fusilier fish Caesio cuning within the Kuroshio Current. Ecology and Evolution.

Addison, J. A., and M. W. Hart. 2005. Spawning, copulation and inbreeding coefficients in marine invertebrates. Biology Letters.

Arseneau, J. R., R. Steeves, and M. Laflamme. 2017. Modified low-salt CTAB extraction of high-quality DNA from contaminant-rich tissues. Molecular Ecology Resources.

Avise, J. C., A. Tatarenkov, and J.-X. Liu. 2011. Multiple mating and clutch size in invertebrate brooders versus pregnant vertebrates. Proceedings of the National Academy of Sciences 108(28):11512–11517.

Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, E. U. Selker, W. A. Cresko, and E. A. Johnson. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE.

Barnes, R. D. 1987. Invertebrate Zoology, 5th edition. WB Saunders company.

Barroso, C. M., M. Nunes, C. A. Richardson, and M. H. Moreira. 2005. The gastropod statolith: A tool for determining the age of Nassarius reticulatus. Marine Biology.

Begg, G. A., K. D. Friedland, and J. B. Pearce. 1999. Stock identification and its role in stock assessment and fisheries management: An overview. Page Fisheries Research.

Blankenberg, D. et al. 2010. Mnipulation of FASTQ data with Galaxy. Bioinformatics 26(14):1783–1785.

Boehm, J. T., J. Waldman, J. D. Robinson, and M. J. Hickerson. 2015. Population genomics reveals seahorses (Hippocampus erectus) of the western Mid-Atlantic coast to be residents rather than vagrants. PLoS ONE.

Bohonak, A. J. 2004. Dispersal, Gene Flow, and Population Structure. The Quarterly Review of Biology.

Bourret, V., M. P. Kent, C. R. Primmer, A. Vasemägi, S. Karlsson, K. Hindar, P. McGinnity, E. Verspoor, L. Bernatchez, and S. Lien. 2013. SNP-array reveals genome-wide patterns of geographical and potential adaptive divergence across the natural range of Atlantic salmon (Salmo salar). Page Molecular Ecology.

Brante, A., M. Fernández, and F. Viard. 2011. Microsatellite evidence for sperm storage and multiple paternity in the marine gastropod Crepidula coquimbensis. Journal of Experimental Marine Biology and Ecology 396:83–88.

Briggs, J. C., and B. W. Bowen. 2012. A realignment of marine biogeographic provinces with particular reference to fish distributions. Journal of Biogeography.

Bunje, P. 2003a. The Mollusca: Sea slugs, squid, snails, and scallops.

Bunje, P. 2003b. The Gastropoda: Snails and slugs, limpets, and sea hares.

Cahill, A. E., A. De Jode, S. Dubois, Z. Bouzaza, D. Aurelle, E. Boissin, O. Chabrol, R. David, E. Egea, J. B. Ledoux, B. Mérigot, A. A. T. Weber, and A. Chenuil. 2017. A multispecies approach reveals hot spots and cold spots of diversity and connectivity in invertebrate species with contrasting dispersal modes. Molecular Ecology.

Cahill, A. E., and J. S. Levinton. 2016. Genetic differentiation and reduced genetic diversity at the northern range edge of two species with different dispersal modes. Molecular Ecology.

Cahill, A. E., and F. Viard. 2014. Genetic structure in native and non-native populations of the direct-developing gastropod Crepidula convexa. Marine Biology.

Le Cam, S., F. Riquet, J. A. Pechenik, and F. Viard. 2014. Paternity and gregariousness in the sex-changing sessile marine gastropod crepidula convexa: Comparison with other protandrous crepidula species. Journal of Heredity.

Castric, V., and L. Bernatchez. 2003. The rise and fall of isolation by distance in the anadromous brook charr (Salvelinus fontinalis Mitchill). Genetics.

Cunha, R. L., C. Grande, and R. Zardoya. 2009. Neogastropod phylogenetic relationships based on entire mitochondrial genomes. BMC Evolutionary Biology.

Dakin, E. E., and J. C. Avise. 2004. Microsatellite null alleles in parentage analysis.

Davis, J. P., and R. T. Sisson. 1988. Aspects of the Biology Relating to the Fisheries Management of New England Populations of the Whelks Busycotypus-Canaliculatus and Busycon-Carica. Journal of Shellfish Research.

Deleware Division of Fish and Wildlife. 2020. Commercial Shellfish Regulations.

Do, C., R. S. Waples, D. Peel, G. M. Macbeth, B. J. Tillett, and J. R. Ovenden. 2014. NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Molecular Ecology Resources.

Dohner, M., N. E. Phillips, and P. A. Ritchie. 2018. Fine-scale genetic structure across a New Zealand disjunction for the direct-developing intertidal whelk Cominella maculosa (Gastropoda: Buccinidae). Biological Journal of the Linnean Society 123:593–602.

Eckert, C. G., K. E. Samis, and S. C. Lougheed. 2008. Genetic variation across species' geographical ranges: The central-marginal hypothesis and beyond.

Edmundson, S. A. 2016. Channeled Whelk (Busycotypus canaliculatus) Ecology in Relation to the Fishery in Vineyard and Nantucket Sounds, Massachusetts.

Edwards, A. L., and M. G. Harasewych. 1988. Biology of the recent species of the subfamily Busyconinae. Journal of Shellfish Research 7:467–472.

Ellingson, R. A., and P. J. Krug. 2016. Reduced genetic diversity and increased reproductive isolation follow population-level loss of larval dispersal in a marine gastropod. Evolution.

Emery, A. M., I. J. Wilson, S. Craig, P. R. Boyle, and L. R. Noble. 2001. Assignment of paternity groups without access to parental genotypes: Multiple mating and developmental plasticity in squid. Molecular Ecology 10(5):1265–1278.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics.

Faircloth, B. C. 2008. MSATCOMMANDER: Detection of microsatellite repeat arrays and automated, locus-specific primer design. Molecular Ecology Resources.

Feng, X. J., G. F. Jiang, and Z. Fan. 2015. Identification of outliers in a genomic scan for selection along environmental gradients in the bamboo locust, Ceracris kiangsu. Scientific Reports.

Fisher, R. A. 2015. Channeled whelk Busycotypus canaliculatus in the mid-Atlantic region. Coherent approach to Busycon/Busycotypus fishery management along the US Altantic seaboard. Page 16th International Conference on Shellfish Restoration. Charleston, South Carolina.

Fisher, R. A., and D. B. Rudders. 2017. Population and Reproductive Biology of the Channeled Whelk, Busycotypus canaliculatus , in the US Mid-Atlantic . Journal of Shellfish Research.

Fleming, A. M., M. M. Dohner, N. E. Phillips, and P. A. Ritchie. 2018. Genetic connectivity among populations of two congeneric direct-developing whelks varies across spatial scales. New Zealand Journal of Marine and Freshwater

Research.

Galindo, J., J. W. Grahame, and R. K. Butlin. 2010. An EST-based genome scan using 454 sequencing in the marine snail Littorina saxatilis. Journal of Evolutionary Biology.

Gomes-dos-Santos, A., M. Lopes-Lima, L. F. C. Castro, and E. Froufe. 2020. Molluscan genomics: the road so far and the way forward.

Gosselin, T., M. Lamothe, F. Devloo-Delva, and P. Grewe. 2020. radiator: RADseq Data Exploration, Manipulation and Visualization using R.

Gruber, B., P. J. Unmack, O. F. Berry, and A. Georges. 2018. dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. Molecular Ecology Resources.

Guo, S. W., and E. A. Thompson. 2006. Performing the Exact Test of Hardy-Weinberg Proportion for Multiple Alleles. Biometrics.

HABE, T. 1972. The Japan's Smallest Gastropod, Ammonicera japonica, sp. nov. Venus.

Hale, S. S. 2010. Biogeographical Patterns of Marine Benthic Macroinvertebrates Along the Atlantic Coast of the Northeastern USA. Estuaries and Coasts.

Hamming, R. W. 1950. Error Detecting and Error Correcting Codes. Bell System Technical Journal.

Harding, J. M. 2012. Observations on the Early Life History and Growth Rates of Juvenile Channel Whelks Busycotypus canaliculatus (Linnaeus, 1758) . Journal of Shellfish Research.

Harrang, E., S. Lapègue, B. Morga, and N. Bierne. 2013. A high load of non-

neutral amino-acid polymorphisms explains high protein diversity despite moderate effective population size in a marine bivalve with sweepstakes reproduction. G3: Genes, Genomes, Genetics.

Haymer, D. S. 2015. Random Amplified Polymorphic Dnas and Microsatellites: What Are They, and Can They Tell Us Anything We Don't Already Know? Annals of the Entomological Society of America 87(6):717–722.

Helyar, S. J., J. Hemmer-Hansen, D. Bekkevold, M. I. Taylor, R. Ogden, M. T. Limborg, A. Cariani, G. E. Maes, E. Diopere, G. R. Carvalho, and E. E. Nielsen. 2011. Application of SNPs for population genetics of nonmodel organisms: New opportunities and challenges. Molecular Ecology Resources.

Huey, R. B. 2002. Physiological Consequences of Habitat Selection. The American Naturalist.

J., A., J. F. Crow, and M. Kimura. 1971. An Introduction to Population Genetics Theory. Population (French Edition).

Jaksch, K., A. Eschner, T. V. Rintelen, and E. Haring. 2016. DNA analysis of molluscs from a museum wet collection: A comparison of different extraction methods. BMC Research Notes.

Jeffery, N. W., B. F. Wringe, M. C. McBride, L. C. Hamilton, R. R. E. Stanley, L. Bernatchez, M. Kent, M. Clément, J. Gilbey, T. F. Sheehan, P. Bentzen, and I. R. Bradbury. 2018. Range-wide regional assignment of Atlantic salmon (Salmo salar) using genome wide single-nucleotide polymorphisms. Fisheries Research.

Jennions, M. D., and M. Petrie. 2000. Why do females mate multiply? A review of the genetic benefits. Biological Reviews 75:21–64.

Jombart, T. 2008. Adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics.

Jones, A. G., C. M. Small, K. A. Paczolt, and N. L. Ratterman. 2010. A practical guide to methods of parentage analysis. Molecular Ecology Resources 10:6–30.

Kamel, S. J., R. K. Grosberg, and J. A. Addison. 2014. Multiscale patterns of genetic structure in a marine snail (Solenosteira macrospira) without pelagic dispersal. Marine Biology 161:1603–1614.

Kilian, A., P. Wenzl, E. Huttner, J. Carling, L. Xia, H. Blois, V. Caig, K. Heller-Uszynska, D. Jaccoud, C. Hopper, M. Aschenbrenner-Kilian, M. Evers, K. Peng, C. Cayla, P. Hok, and G. Uszynski. 2012. Diversity arrays technology: A generic genome profiling technology on open platforms. Methods in Molecular Biology.

Krebs, J. E., E. S. Goldstein, and S. T. Kilpatrick. 2011. Lewin's Genes XInternatio. Jones and Bartlett Publishers, Salisbury, Massachusetts.

Kyle, C. J., and E. G. Boulding. 2000. Comparative population genetic structure of marine gastropods (Littorina spp.) with and without pelagic larval dispersal. Marine Biology.

Leaché, A. D., and J. R. Oaks. 2017. The Utility of Single Nucleotide Polymorphism (SNP) Data in Phylogenetics. Annual Review of Ecology, Evolution, and Systematics.

Lotterhos, K. E., and M. C. Whitlock. 2014. Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. Molecular Ecology.

Luu, K., E. Bazin, and M. G. B. Blum. 2017. pcadapt: an R package to perform genome scans for selection based on principal component analysis. Page Molecular Ecology Resources.

Magalhaes, H. 1948. An Ecological Study of Snails of the Genus Busycon at Beaufort, North Carolina. Ecological Monographs 18(3):377–409.

Mäkinen, T., M. Panova, and C. André. 2007. High levels of multiple paternity in Littorina saxatilis: Hedging the bets? Journal of Heredity 98(7):705–711.

Mamoozadeh, N. R., J. E. Graves, and J. R. McDowell. 2019. Genome-wide SNPs resolve spatiotemporal patterns of connectivity within striped marlin (Kajikia audax), a broadly distributed and highly migratory pelagic species. Evolutionary Applications.

Mariani, S., K. T. C. A. Peijnenburg, and D. Weetman. 2012. Independence of neutral and adaptive divergence in a low dispersal marine mollusc. Marine Ecology Progress Series 446:173–187.

Martel, A., and F. S. Chia. 1991. Drifting and dispersal of small bivalves and gastropods with direct development. Journal of Experimental Marine Biology and Ecology.

Maryland Department of Natural Resources. 2020. Commercial Shellfish Regulations.

Mayr, E. 1963. Animal Species and Evolution. Harvard University Press,

Cambridge.

Mccartney, M. A., M. L. Burton, and T. G. Lima. 2013. Mitochondrial DNA differentiation between populations of black sea bass (Centropristis striata) across Cape Hatteras, North Carolina (USA). Journal of Biogeography.

McDowell, J. R., and H. L. Brightman. 2018. High level of genetic connectivity in a deep-water reef fish, Caulolatilus microps. Journal of Fish Biology.

McInerney, C. E., A. L. Allcock, M. P. Johnson, D. A. Bailie, and P. A. Prodöhl. 2011. Comparative genomic analysis reveals species-dependent complexities that explain difficulties with microsatellite marker development in molluscs. Heredity.

Meirmans, P. G. 2020. genodive version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. Molecular Ecology Resources.

Misener, S., S. A. Krawetz, S. Rozen, and H. Skaletsky. 2003. Primer3 on the WWW for General Users and for Biologist Programmers. Page Bioinformatics Methods and Protocols.

Narum, S. R., and J. E. Hess. 2011. Comparison of FST outlier tests for SNP loci under selection. Molecular Ecology Resources.

Nelson, G. A., S. H. Wilcox, R. Glen, and T. L. Pugh. 2018. No Title. Page Massachusetts Division of Marine Fisheries Technical Report TR-66. Gloucester, Massachusetts.

New Jersey Division of Fish and Wildlife. 2020. Commercial Shellfish Regulations.

Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills, and P. Shipley. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes.

Palumbi, S. R. 2003. Population genetics, demographic connectivity, and the design of marine reserves. Ecological Applications.

Pappalardo, P., J. M. Pringle, J. P. Wares, and J. E. Byers. 2015. The location, strength, and mechanisms behind marine biogeographic boundaries of the east coast of North America. Ecography.

Peakall, R., and P. E. Smouse. 2012. GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics.

Pearse, D. E., and E. C. Anderson. 2009. Multiple paternity increases effective population size.

Peemoeller, B. J., and B. G. Stevens. 2013. Age, size, and sexual maturity of channeled whelk (Busycotypus canaliculatus) in buzzards bay, Massachusetts. Fishery Bulletin.

Pembleton, L. W., N. O. I. Cogan, and J. W. Forster. 2013. StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. Molecular Ecology Resources.

Power, A. J., C. J. Selers, and R. L. Walker. 2009. Grrowth and Sexual Maturity of the Knobbed Whelk, Busycon carica (Gmelin, 1791), From a Commercially HArvested Population in Costal Georgia. Occasional Papers of the University of Georgia Marine Extension Service.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155(2):945–959.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

RAWLINGS, T. A. 2007. Adaptations to Physical Stresses in the Intertidal Zone: The Egg Capsules of Neogastropod Molluscs. American Zoologist.

Robins, R. 2018. Personal Communication.

Rodhouse, P. G., and E. M. C. Hatfield. 1990. Age determination in squid using statolith growth increments. Fisheries Research.

Roper, C. F. E., and K. J. Boss. 1982. The Giant Squid. Scientific American.

Rousset, F. 2008. GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources.

Russello, M. A., S. L. Kirk, K. K. Frazer, and P. J. Askey. 2012. Detection of outlier loci and their utility for fisheries management. Evolutionary Applications.

Sanford, E., and M. W. Kelly. 2011. Local Adaptation in Marine Invertebrates. Annual Review of Marine Science 3:509–535.

Sauvage, C., N. Bierne, S. Lapègue, and P. Boudry. 2007. Single Nucleotide polymorphisms and their relationship to codon usage bias in the Pacific oyster Crassostrea gigas. Gene.

Schultzhaus, J. N., C. R. Taitt, B. Orihuela, M. Smerchansky, Z. S. Schultzhaus, D. Rittschof, K. J. Wahl, and C. M. Spillmann. 2019. Comparison of seven methods for DNA extraction from prosomata of the acorn barnacle,

Amphibalanus amphitrite. Analytical Biochemistry.

Seeb, J. E., G. Carvalho, L. Hauser, K. Naish, S. Roberts, and L. W. Seeb. 2011. Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms.

Sefc, K. M., and S. Koblmüller. 2009. Assessing parent numbers from offspring genotypes: The importance of marker polymorphism. Journal of Heredity 100(2):197–205.

Selkoe, K. A., and R. J. Toonen. 2006. Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers.

Shalack, J. D. 2007. Movement and Behavior of Whelks (Family Melongenidae) in Georgia Coastal Waters. University of Georgia.

Slatkin, M. 2006. Isolation by Distance in Equilibrium and Non-Equilibrium Populations. Evolution.

Stanley, S. 1970. Shell Form And Life Habits Of The Bivalvia (Mollusca). Memoir of the Geological Society of America 125.

Strathmann, R. R. 1990. Why life histories evolve differently in the sea. American Zoologist.

Teske, P. R., I. Papadopoulos, G. I. Zardi, C. D. McQuaid, M. T. Edkins, C. L. Griffiths, and N. P. Barker. 2007. Implications of life history for genetic structure and migration rates of southern African coastal invertebrates: Planktonic, abbreviated and direct development. Marine Biology.

Teske, P. R., J. Sandoval-Castillo, M. Sasaki, and L. B. Beheregaray. 2015. Invasion success of a habitat-forming marine invertebrate is limited by lower-

than-expected dispersal ability. Marine Ecology Progress Series 536:221–227.

Treml, E. A., J. Roberts, P. N. Halpin, H. P. Possingham, and C. Riginos. 2015. The emergent geography of biophysical dispersal barriers across the Indo-West Pacific. Diversity and Distributions.

Underwood, E. B., and T. L. Darden. 2019. Population Genetic Structure and Diversity of the Invasive Island Apple Snail Pomacea maculata (Perry, 1810) in South Carolina and Georgia. Journal of Shellfish Research.

Vaux, F., S. F. K. Hills, B. A. Marshall, S. A. Trewick, and M. Morgan-Richards. 2017. A phylogeny of Southern Hemisphere whelks (Gastropoda: Buccinulidae) and concordance with the fossil record. Molecular Phylogenetics and Evolution.

Vieira, M. L. C., L. Santini, A. L. Diniz, and C. de F. Munhoz. 2016. Microsatellite markers: What they mean and why they are so useful. Genetics and Molecular Biology 39(3):1415–4757.

Virginia Marine Resource Comminsion. 2020. Pertaining to channeled whelk. Regulation 4 VAC 20-890-10 ET SEQ.

Walker, B. R. L., A. J. Power, M. Sweeney-reeves, E. Covington, M. Mitchell, and T. Recicar. 2008. Growth, Migration, Population Structure and Sex Ratio of Four Whelk Species (Family Melongenidae) Within Wassaw Sound, Georgia. Occasional Papers of the University of Georgia Marine Extension Service.

Walker, D. E., A. J. Power, M. Sweeney-Reeves, and J. C. Avise. 2007. Multiple

paternity and female sperm usage along egg-case strings of the knobbed whelk, Busycon carica (Mollusca; Melongenidae). Marine Biology 151:53–61.

Walker, D., B. A. Porter, and J. C. Avise. 2002. Genetic parentage assessment in the crayfish Orconectes placidus, a high-fecundity invertebrate with extended maternal brood care. Molecular Ecology.

Wang, J. 2004. Sibship Reconstruction from Genetic Data with Typing Errors. Genetics 166(4):1963–1979.

Waples, R. S. 1987. A Multispecies Approach to the Analysis of Gene Flow in Marine Shore Fishes. Evolution.

Waples, R. S., and O. Gaggiotti. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity.

Watson, P. J. 1991. Multiple paternity as genetic bet-hedging in female sierra dome spiders, Linyphia litigiosa (Linyphiidae). Animal Behaviour.

Weersing, K., and R. J. Toonen. 2009. Population genetics, larval dispersal, and connectivity in marine systems. Marine Ecology Progress Series.

Weetman, D., L. Hauser, P. W. Shaw, and M. K. Bayes. 2005. Microsatellite markers for the whelk Buccinum undatum. Molecular Ecology Notes.

WoRMS Editorial Board. 2020. World Register of Marine Species.

Wray, G. A., and R. A. Raff. 1991. The evolution of developmental strategy in marine invertebrates. Trends in Ecology and Evolution.

Wright, S. 1938. Size of population and breeding structure in relation to evolution.

Science.

Wright, S. 1943. Isolation by Distance. Genetics 28(2):114–138.

Xu, T., J. Sun, J. Lv, H. Kayama Watanabe, T. Li, W. Zou, G. W. Rouse, S. Wang, P. Y. Qian, Z. Bao, and J. W. Qiu. 2017. Genome-wide discovery of single nucleotide polymorphisms (SNPs) and single nucleotide variants (SNVs) in deep-sea mussels: Potential use in population genomics and cross-species application. Deep-Sea Research Part II: Topical Studies in Oceanography.

Xue, D., T. Zhang, and J. X. Liu. 2014. Microsatellite evidence for high frequency of multiple paternity in the marine gastropod Rapana venosa. PLoS ONE 9(1):e86508.

Ylitalo, H., T. Oliver, I. Fernandez-Silva, J. Wood, and R. Toonen. 2018. Who's your daddy? A behavioral and genetic study of multiple paternity in a polygamous marine invertebrate, Octopus oliveri.