ABSTRACT

Title of Dissertation:	POPULATION GENETICS OF EASTERN OYSTER *Crassostrea virginica* RESTORATION IN THE CHESAPEAKE BAY
	Katherine M. Hornick, Doctor of Philosophy, and 2020
Dissertation directed by:	Associate Professor, Louis V. Plough, Marine, Estuarine, Environmental Science

The strategic release of captive-bred organisms is one of the most popular methods to restore species, but concerns exist regarding genetic impacts on natural populations over the long-term. Slow recovery of depleted eastern oyster *C. virginica* populations in the Chesapeake Bay prompted a large-scale hatchery-based restoration program consisting of the mass-release of hatchery-produced juveniles from local, wild broodstock. This dissertation characterized the genetic impact of this program, with the overall goal of understanding how characteristics of species life-history interact with hatchery practices to shape genetic variation in populations over short and long-time scales. In Chapter 2, analysis of genetic diversity changes resulting from hatchery production under two spawning designs (mass- and controlled-spawns) revealed substantial reductions in diversity and the number of breeders from parents to offspring, due primarily to high variance in reproductive success among adults in hatchery culture. In Chapter 3, high-resolution genomic data was used in a population genetic analysis comparing diversity of

restored reefs in Harris Creek with variable planting histories and husbandry practices to 'wild' Chesapeake Bay oyster reefs. While restored reefs showed similar levels of diversity as wild reefs, strong positive relationships between planting frequency or broodstock numbers and genetic diversity were found, suggesting that hatchery practices can significantly impact diversity in natural populations. These genomic data also permitted the investigation of local adaptation and genotype by environment associations which revealed that salinity was correlated with loci putatively under selection, suggesting potential fitness tradeoffs for sourcing non-local broodstock. In Chapter 4, an individual-based model was created using biological and demographic data from Chesapeake Bay oysters to simultaneously evaluate the impact of multiple hatchery practices on natural population genetic diversity over time scales not possible with empirical methods. Overall, hatchery practices had a large effect on genetic diversity in most scenarios, but spawning practices (mass or controlled) and broodstock rotation were more important than broodstock number, suggesting that broodstock-limited programs may have other options to maintain diversity. In summary, these studies advance our understanding of how marine supplementation impacts both neutral and adaptive variation and will provide critical information for future oyster restoration efforts.

POPULATION GENETICS OF EASTERN OYSTER *Crassostrea virginica* RESTORATION IN THE CHESPEAKE BAY

by

Katherine Marie Hornick

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2020

Advisory Committee:
Professor Louis V. Plough, Chair
Dr. James J. Pierson
Dr. Matthew Gray
Dr. Jeffrey Cornwell
Dr. Standish K. Allen, Jr.

Foreword

This dissertation contains three research chapters that were conducted in collaborative efforts with my academic advisor. Chapter 2 is published in an academic journal with myself as the primary author and major contributor with my advisor as a co-author also contributing. At the end of each chapter, I acknowledge my collaborators.

Dedication

To my parents, who have worked tirelessly to provide me with the best educational opportunities, have been an endless source of support, and have always encouraged me to follow my dreams.

Acknowledgements

This dissertation is the culmination of five years of experiments, laboratory work, bioinformatic analyses, model building, and interpretations. It not only represents my efforts, but those that helped me throughout this process. This research would not have been possible without them, nor would I be the person and scientist that I am today without their guidance and support.

I would first like to thank my advisor, Dr. Louis Plough, whose strong support, guidance, and expert mentoring allowed me to succeed. Louis has taught me to think critically about every analysis, to interpret my results in a biological context, and countless other things. Without fail, Louis always made time for me whenever I had a question, no matter how busy he was with other responsibilities. I am grateful for him giving me space to learn, grow, and develop as a scientist, and for always believing in me. During times when I was in the weeds or feeling helpless, he always had the ability to pull me back onto the right course. The opportunities he has provided me with have contributed to my growth as a scientist and will be useful for my future career. Louis is truly an outstanding scientist, teacher, and mentor, and I cannot thank him enough.

Next, I would like to thank my committee members, Dr. Jamie Pierson, Dr. Matthew Gray, Dr. Ken Paynter, Dr. Stan Allen, and Dr. Jeff Cornwell. My committee members were extremely supportive of my research and provided insightful comments on my projects. I would like to thank Jamie, who has supported me both scientifically and personally throughout this process and who has always been so generous with his time whenever I had questions. Discussions with Jamie helped fuel interesting research directions, and his insights and advice outside of graduate research have been helpful

throughout this process. I would like to thank Ken for always providing useful and constructive directions and insight during committee meetings. I would like to thank Matt Gray for always providing a practical perspective during committee meetings. I would like to thank Stan Allen, for motivating me to become a better scientist. Finally, I would like to thank Jeff Cornwell for serving on my committee the day of my dissertation defense.

Many funding institutions, people, and funding sources supported the research within this dissertation. First, thank you to the Horn Point Laboratory for providing a Graduate Assistant Fellowship to support the first year of my PhD. Thank you to the Horn Point Laboratory Education Committee and the National Shellfisheries Association for providing support for travel to academic meetings that have greatly enriched my educational and research experiences. Thank you to the University of Maryland Department of Marine Estuarine Environmental Science and the University of Maryland Graduate School for Dean's Fellowship awards that helped support my research and allowed me to participate in academic meetings. I would like to acknowledge and thank the Horn Point Laboratory Oyster Hatchery and the Paynter Lab for sharing their data and providing insight and assistance throughout the duration of my graduate work. I would like to thank Ralph Kimes, Gordy Dawson, and the whole maintenance crew in AREL for always coming to the rescue when I was struggling with a maintenance issue. I would like to thank my fellow Plough Lab members especially Lexy McCarty, Ben Lee, Shannon Hood, and Katie McFarland for their encouragement and support. Finally, I would also like to thank Anna Priester for teaching me oyster hatchery and nursery

culture 101, and for kindly watching my oysters while I was out of town and Samantha Gleich for being so supportive and encouraging.

I would like to thank my family for their constant love and support, who have been instrumental in helping me navigate the challenges of balancing work and life, and handle hardships throughout this process with grace. My mother, Debby Hornick, father, David Hornick, and sister, Elizabeth Sayre have supported my passion and curiosity for the natural world since the days of my seashell and shark tooth collection, and the turtle talk at Kiawah Island. I would not have made it where I am today without them. A special thanks to my sister Liz, and brother in law Justin Sayre, for always allowing me to invite myself over for a weekend away from eastern shore and for providing me with endless snacks and laughs. I would not have been able to finish this degree without your support.

Table of Contents

List of Tables

Chapter 2

Chapter 3

eastern oysters. Hatchery plantings denotes the number of seasons a restored site was planted with hatchery-produced oysters.

Table 2. Sampling site of eastern oysters, station ID (buoy), depth of buoy in meters, the sampling frequency (annual), and latitudinal range of sampling buoys. Distance represents the distance from the buoy to the oyster sampling site. Continuous sampling frequency represents buoys that sample in 15-minute increments.

Table 3. Summary of data filtering procedures: rows refer to filtering steps; columns refer to statistics for each step. For columns, 'sites' refers to individual polymorphisms (SNPs, indels, or complex polymorphisms), and 'Inds' refers to individuals. 'Start', 'End', and 'Removed' refer, respectively, to the number of each unit before the filtering step, the number after the filtering step, and the number removed with the filter.

Table 4. Descriptive statistics for each *C. virginica* sampling site, including observed heterozygosity (H_o), expected heterozygosity (H_e), inbreeding coefficient (F_{IS}), allelic richness (A_r), effective population size (N_e), and relatedness Ritland (1996) and Wang (2016).

Table 5. Redundancy analysis (RDA) results for the full dataset for all *C. virginica* populations (7,710 SNPs) and the inner Bay populations (excluding Wachapreague; 7,710 SNPs). Correlated represents SNPs identified as outliers in RDA and correlated with environmental parameters including mean salinity (salinity), maximum water temperature (maxWT), minimum water temperature (minWT), minimum pH (minpH), mean pH (mpH), and minimum dissolved oxygen (DO; minDO). Significant SNPs represent those were considered to be significant outliers if the q-value was <0.05 (see Methods for details).

Table 6. Redundancy analysis (RDA) results for potentially neutral and adaptive *C. vrignica* SNP datasets including all sites and only the inner Bay sites (excluding Wachapreague). The environmental parameters include mean salinity (salinity), maximum water temperature (maxWT), mean water temperature (mWT), minimum pH (minpH), mean pH (mpH), and minimum dissolved oxygen (DO; minDO).

Table 7. BLAST matches from sequences identified as being putatively under divergent selection from oyster *Crassostrea virginica* populations in the studied region. SNPs are located in the eastern oyster genome and chromosomes. Protein IDs and names are derived from the eastern oyster protein sequences of the genome. Gene ontologies related to the identified protein (GO ID) were retrieved using protein names and the environmental variable correlated with proteins is listed.

Table 8. Effective population size (N_e) for *C. virginica* sampling site from this study (Chesapeake Bay) and from Bernatchez et al., (2019) (Canada)

Chapter 4

Table 1. Demographic and life-history parameters applied to wild and hatchery populations of *C. virginica*

Table 2. Twenty-seven model scenarios of hatchery-based supplementation including those varying spawning type (controlled or mass-spawn), broodstock management (recycled or rotated), and broodstock numbers (N=10, 25, and 500). Mass-spawn scenarios also tests the genetic impact of varying migration rates (closed, medium, and open) for rotated and recycled broodstock.

Table 3. Parameters tested in sensitivity analysis of hatchery supplementation IBM of *C. virignica*

Table 4. Summary table of changes in genetic diversity metrics including effective population size (N_e), allelic richness (A_r), and observed heterozygosity (H_o) throughout the course of model simulations for mass and controlled-spawn scenarios. Spawning types are further divided by broodstock practice (rotated and recycled), and periods during the course of simulations including supplementation (Supp. generation 3), post-supplementation early (Post-supp. early, generation 9), and post-supplementation late (Post-supp. Late, generation 27). Genetic diversity metrics are represented as percent difference relative to control scenarios.

List of Figures

Chapter 1

Chapter 2

Figure S3. Mean global relatedness in *C. virginica* a supplemented reef (HC), wild broodstock source parents, and hatchery-produced offspring. Relatedness: Lynch and Ritland (Lynch and Ritland, 1999) relatedness estimator calculated by COANCESTRY (Wang, 2011).

Figure S4. Discriminant analysis of principle component (DAPC) among mass-spawned *C. virginica* hatchery cohorts (offspring). Number of PCs to retain was determined via the a-spline optimization approach in Adegenet (28 PCs retained).

Chapter 3

Figure 1. Map of sampling locations of *C. virginica* eastern oysters within the Chesapeake Bay. Abbreviations of sampling sites are presented in Table 1.

Figure 2. Linkage disequilibrium decay (R^2) and distance across the 10 *C. virginica* chromosomes with loess best fit. The critical R^2 was calculated to be 0.09 and the green line represents the point at which the loess curve intersects the critical R^2.

Figure 3. Venn diagram with number of SNPs identified as outliers by three methods: pcadapt, OutFLANK, and Bayescan prior to SNP thinning (full dataset).

Figure 4. Effect of hatchery practices on metrics of genetic diversity in Harris Creek restored reefs of *C. virginica*. The effect of the number of broodstock on genetic diversity metrics of restored oyster reefs including, A. effective population size (N_e) and B. relatedness of restored reefs. The effect of the number of hatchery planting seasons on genetic diversity metrics of restored reefs including C. N_e, D. relatedness, and E. observed heterozygosity. The effect of the F. broodstock male-to-female ratio in the hatchery on observed heterozygosity of restored reefs.

Figure 5. Heatmap of pairwise F_{ST} for *C. virginica* populations using the putatively neutral SNPs. Populations are ordered from north to south (from HCR1, Harris Creek 1 to W, Wachapreague). Abbreviations of sampling sites are presented in Table 1.

Figure 6. Discriminant analysis of principal components (DAPC) among wild and restored *C. virginica* populations based on 2,810 neutral unlinked SNPs. Abbreviations of sampling sites are presented in Table 1.

Figure 7. Plot of STRUCTURE results for 11 *C. virginica* populations using the neutral SNPs. Plots of individual admixture determined using the program STRUCTURE at the K recommended by the dK method (K=4).

Figure 8. Isolation-by-distance (IBD) relationship between A. all *C. virginica* site pairs and B. only wild *C. virginica* pairs where linearized pairwise F_{ST} values ($F_{ST}/(1- F_{ST})$) are regressed over marine distance. Circles represent pairwise comparisons and the regression line (blue) is fitted with a 95% confidence limits (grey).

oysters (blue; ages 2–11) in the virtual Harris Creek wild population during a 50-generation model run.

Figure 5. Comparison of genetic diversity metrics from simulated populations vs empirical estimates from the Chesapeake Bay. Panel A shows the effective population size (N_e), panel B shows allelic richness (A_r), panel C shows observed heterozygosity (H_o), and Panels D–F show pairwise FST of Harris Creek-Chesapeake Bay (D), Harris Creek-Broodstock (E), and Chesapeake Bay-Broodstock (F). The dotted line represents the empirical estimates from Chapter 3 across 2,210 RAD loci from mixed-age cohorts of oysters collected from Harris Creek (Harris Creek), Little Choptank (Chesapeake Bay), and States Bank (Broodstock). Note that the scales differ across panels.

Figure 6. Raw modeling results for a control scenario and a mass spawned, recycled broodstock, 10 broodstock restoration scenario. Estimates of effective population size (N_e) are shown in panel A for the control scenario, and panel B for the supplementation scenario. Change in population sizes are depicted for the control scenario (panel C), and the restored scenario (panel D). Data for each population is presented, including the supplemented population (HC- blue lines), a local population that receives migrants from the supplemented population (CB -green lines), and the wild broodstock population (BR-red lines). Estimates were based on a subsample of 10,000 *C. virginica* mixed-aged individuals (ages 2–11) and the line represents the average across 10 simulation replicates. The vertical dotted lines represent the start and the end of the supplementation period.

Figure 7. Model simulation results for genetic diversity metrics across varying spawning and broodstock management scenarios. The varying scenarios include controlled-spawns (controlled) and mass-spawns of recycled (M-recycled) and rotated (M-rotated) broodstock and results are compared to control scenarios (percent difference). Genetic diversity metrics are shown over time during the simulations (shown as generations in the grey bars above plots). Panels A shows effective population size (N_e) estimates. Panel B shows allelic richness (A_r) and Panel C shows observed heterozygosity (H_o). Estimates of genetic diversity metrics were based on a subsample of 10,000 *C. virginica* mixed-aged individuals (ages 2–11) and are averages across 10 simulation replicates.

Figure 8. Model simulation results for genetic diversity metrics across varying migration and broodstock management scenarios. Panels A and B show effective population size (Ne) estimates for mass-spawned recycled broodstock and rotated broodstock, respectively. Panels C and D show allelic richness (Ar) estimates for mass-spawned recycled broodstock and rotated broodstock, respectively. Panels E and F show observed heterozygosity (Ho) estimates for mass spawned recycled broodstock and rotated broodstock, respectively. Different colored bars represent different number of broodstock (N=10, 25, and 500). The migration rates tested are closed (no migration), medium (5% migration), and open (15% migration). Genetic diversity metrics are represented as percent difference relative to control scenarios. Estimates were based on a subsample of

10,000 C. virginica mixed-aged individuals (ages 2–11) and are averages across 10 simulation replicates.

Figure 9. Sensitivity simulation results (shown in generations, grey boxes above plots) for genetic diversity metrics. Genetic diversity metrics are represented as percent difference relative to control scenarios using the standard modeling parameters for all scenarios. Panels A and B show effective population size (N_e) estimates for sensitivity analyses of the geometric reproductive distribution (varying *P*) and poly number of males and females in a reproductive event, respectively. Panels C and D show allelic richness (A_r) estimates for sensitivity analyses of the geometric reproductive distribution (varying *P*) and poly number of males and females in a reproductive event, respectively. Panels E and F show observed heterozygosity (H_o) estimates for sensitivity analyses of the geometric reproductive distribution (varying *P*) and poly number of males and females in a reproductive event, respectively. Values corresponding to the numbers of the x-axis for the geometric distribution (Panels A, C, and E) are 0.00002, 0.0002, 0.002, 0.02, and 0.2 for 1, 2, 3, 4, and 5, respectively. Values corresponding to the numbers of the x-axis for the poly number (Panels B, D, and F) are 0.3X, 0.03X, 0.003X, 0.0003X, and 0.00003X for 1, 2, 3, 4, and 5 (X=population size), respectively. Tukey's boxplots were calculated from 5 replicates.

Chapter 1: Introduction to Dissertation

Genetic diversity, population persistence, and conservation

The maintenance of biodiversity is one of the most important conservation concerns, as populations of organisms are being reduced at an alarming rate, and an increasing number of species require human intervention to prevent extinction (Frankham et al. 2002). Genetic diversity is a critical measure of biodiversity that impacts population viability (Frankham 1996, 2005) and can be defined as the variety of genetic variants (alleles) and genotypes present in a population (Frankham et al. 2002). International conservation policy recognizes biodiversity at three levels: ecosystem, species, and genetic, and that management should aim to retain all three (Convention on Biological Diversity 2007). Therefore, characterization and management of genetic diversity is important and has been facilitated by advancements in molecular techniques.

The extent of genetic diversity is determined by the interaction among a few key evolutionary forces. From a theoretical viewpoint, genetic diversity reflects the balance between the appearance and disappearance of genetic variants (alleles) and is governed by the rate of allele loss and fixation. For example, loci with neutral alleles are largely influenced by the stochastic force of genetic drift, or the random fluctuation of allele frequency across generations. The deterministic force of natural selection also changes allele frequencies. The interplay of these forces drives evolutionary changes in populations over time. Thus, by studying the genetic diversity of populations, it is

possible to understand the interaction of evolutionary and demographic forces shaping those populations.

The rate of loss of genetic diversity can be predicted from the effective population size (N_e; Wright 1931; Crow and Kimura 1970), a fundamental parameter in evolutionary biology and conservation genetics. Effective population size is the size of an idealized population that loses genetic diversity at the same rate as the focal population (Wright 1931) and integrates genetic effects with the life history of the species. The effective population size (N_e) (larger is better) is closely linked to core genetic parameters of the population that determine long term viability, such as inbreeding, genetic drift, or maintenance of genetic diversity across generations (Frankham 1996; Frankham et al. 2014). Based on N_e theory, smaller populations lose diversity faster than larger populations, but the loss of genetic diversity can be slowed by maintaining a constant population size (Vilà et al. 2003), ensuring equal sex ratios (Melampy and Howe 1977), maintaining non-overlapping generations (Crow and Denniston 1988), and equalizing family sizes (Frankham et al. 2000). Effective population size takes into account the average level of inbreeding, or the rate of change in average homozygosity, which increases as population size decreases (Crow, 1954). While the N_e is often much smaller than abundances or census size (number of reproductively mature individuals), a population risks extinction when N_e becomes too small (Hedrick and Kalinowski 2000). Therefore, N_e is an important metric of diversity used to monitor population health in conservation management (e.g. Hare et al. 2011).

In addition to N_e, commonly used measures to compare genetic diversity within natural populations at the molecular level include heterozygosity and allelic richness. At

a single locus, heterozygosity is the proportion of individuals that are heterozygous (have two different alleles at that locus); heterozygosity ranges between zero and one. Inbreeding, for example, directly reduces heterozygosity by increasing the proportion of homozygotes relative to random expectations. Another valuable measure of genetic diversity is the total number of alleles at a locus, a metric that is more sensitive to the loss of genetic diversity than heterozygosity. However, this metric, unlike heterozygosity, is highly dependent on sample size. Allelic richness is the measure of allelic diversity that accounts for differences in sample size by using a rarefaction method. Together, these metrics allow for standing genetic diversity in natural populations to be quantified and compared.

While conservation often requires a focus on immediate demographic concerns, retention of genetic diversity in threatened species is critical if extinction is to be avoided. The extent of genetic diversity within a population collectively influences the ability of that population to persist over the long-term or during environmental stress and plays a critical role in the ecological dynamics of a community, potentially supporting sustained levels of harvest (Reusch et al. 2005; Vellend and Geber 2005; Agashe 2009). Declines in genetic diversity can reduce fitness and initiate a negative feedback loop that leads to smaller population sizes, drift, and additional inbreeding (Gilpin and Soule 1986). Declines in genetic diversity have been associated with an increased risk of extinction (Saccheri et al. 1998) reduced population growth rate (Hanski and Saccheri 2006), and reduced potential for response to environmental change (Willi et al. 2006; Allendorf et al. 2013).

The use of genetic markers to measure genetic diversity for conservation purposes has only been readily available over the last few decades (Allendorf et al. 2010). Allozyme analysis (Hunter and Markert 1957) was the first technique to reveal the extensive genetic variation found in natural populations. With the discovery of polymerase chain reaction (PCR), the use of DNA markers became available and the first studies of DNA variation examined mitochondrial DNA (Avise 1986). Microsatellite DNA markers (from nuclear DNA) were first discovered in the 1980s and became widely used to quantify genetic diversity within a conservation context in the 1990s (Schlotterer et al. 1998). Today, there are numerous genetic marker types and approaches that allow quantification of genetic diversity to address a variety of questions in evolution and ecology, including the most ubiquitous marker, single nucleotide polymorphisms (SNP), which is often the unit of measurement in genome-wide datasets (Syvänen 2005). With advances in genomic technologies, researchers can now survey entire genomes from multiple individuals across populations, providing highly accurate estimates of key diversity parameters. Furthermore, these large marker data sets can provide insight into population size (Beaumont et al. 2002; Waples and Do 2008) and population history (Luikart et al. 1998; Kuhner 2006) to provide critical information for species conservation.

The genetics of captive breeding and supplementation

The efficacy of captive breeding programs to enhance natural populations remains controversial among scientists, managers, and the public (Snyder et al. 1996;

Bowkett 2009). Captive breeding can produce large numbers of individuals for stock supplementation or restoration. However, it can cause rapid genetic changes that can undermine long-term population persistence and reduce the fitness of captive bred organisms when released in the natural environment (Frankham 2008). While the goal of captive breeding for re-introduction or supplementation is to increase the demographic representation while maintaining genetic diversity (Allendorf et al. 2010), avoiding genetic changes such as adaptation to captivity are often difficult. These changes have been documented in a variety of taxa, from invertebrates to apex predators (Frankham 2008). However, captive breeding is often a ubiquitous element of species management in order to avoid extinction. High profile and successful captive breeding programs include those involving the conservation of waterfowl (Hayes 2002), the California condor (*Gymnogyps californianus*) (Ralls and Ballou 2004; Walters et al. 2010), black footed ferret (*Mustela nigripes*) (Wisely et al. 2003), and steelhead trout (*Oncorhynchus mykiss*) (Berejikian and Doornik 2018).

Adaptation to captivity can be a significant issue for organisms that live in artificial environments such as hatcheries (Frankham and Loebel 1992). Typically, captive breeding programs require hatchery-based laboratory (or nursery) propagation of the early life stages of species, which may impose an altered environment when the potential for genetic change (natural selection) is substantial (Christie et al. 2016; Plough et al. 2016). Recent studies have shown significantly altered gene expression and methylation patterns in hatchery-produced fish (Christie et al. 2016; Le Luyer et al. 2017). Selection within a generation in the hatchery, or adaptation to the hatchery environment over successive generations, can lead to adaptive mismatches with local

environmental conditions (Waal et al. 2013), decreased fitness of hatchery-produced stocks in the natural environment (Araki et al. 2008; Lorenzen et al. 2012; Christie et al. 2014), and depressed wild population fitness via hatchery-reared/wild hybridizations (Hedgecock and Coykendall 2007). Selecting the source of broodstock in supplementation programs is therefore important and should be informed by the scale and degree of local adaptation in wild populations (McKay et al. 2005), which is often unknown. Whether intentional or unintentional, minimizing selection in captivity is important to the long-term success of captive breeding programs, particularly if population supplementation is desired (McPhee 2004).

Captive breeding and supplementation can also have profound and rapid effects on the genetic diversity of the wild recipient populations, a subject that has received considerable attention both theoretically and empirically (Grant et al. 2017; Kitada 2018). Reductions in genetic diversity and effective population size (N_e) have been documented in wild receiving populations when large numbers of hatchery-produced individuals from small broodstock number are released (e.g. Ryman and Laikre 1991; Ryman et al. 1995; Christie et al. 2012). This effect (reduction of N_e in wild receiving populations produced by small broodstock size) has been dubbed the Ryman-Laikre effect (Ryman and Laikre 1991; Ryman et al. 1995), and remains a concern for supplementation programs of salmonids and other fish (e.g. Ryman and Laikre 1991; Hedgecock and Coykendall 2007; Laikre et al. 2010; Christie et al. 2012). To reduce the potential for the Ryman-Laikre effect, large numbers of wild rotated broodstock and maintaining natural spawning sex ratios is recommended. Unfortunately, following such recommendations can be hampered by the status of the wild population or limited hatchery resources/capacity. For

example, the number of broodstock a hatchery facility spawns may be limited by space and/or labor, or it may be difficult to obtain broodstock harboring sufficient genetic diversity, especially in heavily exploited populations. Hatcheries capable of using large numbers of native broodstock that are renewed regularly may achieve the goal of high-diversity, self-sustaining populations (Adkison 1995; Caughley and Gunn 1996; Heggenes et al. 2006), with little to no genetic impact (e.g. Heggenes et al. 2006; Gow et al. 2011; Katalinas et al. 2017; Berejikian and Doornik 2018). However, major gaps still exist in understanding genetic diversity changes resulting from hatchery-based intervention for marine species with complex life-history features.

The genetics of shellfish supplementation

Among the most emblematic taxa representative of stock depletion, habitat destruction, and overfishing are marine shellfish, due to their crucial roles in coastal ecosystem function, their economic importance, and their precipitous population declines in recent times due to anthropogenic impacts. Restoration or supplementation of marine shellfish, primarily through juvenile seeding (Gaffney 2006; Laing et al. 2006), has increased significantly over the last few decades (e.g. oyster and mussels in Port Phillip Bay, Australia, The Nature Conservancy; eastern oysters in New York Harbor, Billion Oyster Project, scallops in Martha's Vineyard, Massachusetts, Martha's Vineyard Shellfish Group Inc., and Olympia oysters in Puget Sound, Puget Sound Restoration Fund). Shellfish restoration faces many of the same challenges that fish supplementation programs face, but also have a number of unique challenges related to restoring reef

structures and maintaining the genetic diversity of both hatchery-propagated and wild populations.

A striking difference between fish and oyster supplementation is the requirement of physical habitat (shell or other hard surface) for the early life-stage to settle and attach. Many oyster species form reefs, and larvae generally attach to other oysters or oyster shell. Dense reefs can be formed by the settlement and growth of successive generations of oysters on the shells of their predecessors. Large-scale reef development is a relatively long and complex process that involves interactions among a variety of physical and biotic factors (Bahr and Lanier 1981; Kennedy and Sanford 1999; Coen and Luckenbach 2000). In addition, areas in need of restoration are often either "recruitment limited", "substrate limited," or both (Brumbaugh and Coen 2009). While recruitment-limited environments lack sufficient local broodstock (reproductively mature individuals) to naturally populate reef structures, substrate-limited environments lack reef structures to which shellfish larvae can attach. Therefore, in areas that are substrate-limited, it is common for projects to involve some form of reef construction from an appropriate substrate type.

Many marine shellfish species exhibit the 'periodic' life-history strategy (e.g. Winemiller and Rose 1992) that is typified by high fecundity (one female can produce millions of offspring) and high mortality (type-III) early in the life cycle (e.g. Gaffney 2006; Plough 2016; Plough et al. 2016). These life history features can result in extremely high variance in reproductive success among parents (e.g. Hedgecock 1994; Lallias et al. 2010) leading to small N_e in hatchery populations (Hedgecock and Sly 1990; Boudry et al. 2002; Appleyard and Ward 2006; Lind et al. 2009). These life-history

features also render shellfish populations more vulnerable to loss of variation than might be expected from their great abundance. The high potential for unequal reproductive success among spawning adults is called the 'sweepstakes reproductive success' (SRS; Hedgecock 1994), in which a small proportion of the spawning population produces the bulk of annual recruitment. The implication of the SRS hypothesis for conservation is that these seemingly large marine populations may have effective population sizes that are orders of magnitudes smaller than census sizes. In addition, the reduction of diversity due to SRS likely increases in programs using only a small number of broodstock and/or mass spawning methods (Hedgecock 1994; Hedgecock and Coykendall 2007; Hedgecock and Pudovkin 2011). Typically, mass spawning protocols hold numbers of broodstock in a single tank, and after gametes are released, cross-fertilization among individuals occurs randomly. In contrast, single pair-matings, in which sperm from one male and eggs from one female are crossed in isolated vessels, are a more controlled way of producing hatchery offspring. Ideally, offspring from crosses should be reared separately until settlement, and numbers should be equalized prior to planting in order to maximize genetic diversity (Camara and Vadopalas 2009; Cooper et al. 2010; Gruenthal et al. 2014). Thus, the degree of diversity loss in hatchery-produced stocks is often related to hatchery protocols (e.g. Lind et al. 2009), which can vary substantially among programs.

In shellfish culture, inadvertent culling of (i.e. selection against) slow growing larvae (e.g. Taris et al. 2006) has been shown to alter the genetic composition of hatchery-produced populations which resulted in directional changes to growth rate and physiology. Moreover, these impacts may carry over to the fitness of the adult stage or may be inherited by the next generation in fish (Christie et al. 2016). However, despite

these studies, the fitness trade-offs resulting from domestication selection are understudied, especially in species with a complex life-cycle that includes pelagic larvae such as shellfish (Emlet and Sadro 2006; Crean et al. 2011). While domestication selection is increasingly under study (Waples 1991; Araki et al. 2007, 2008; Christie et al. 2012b; Skaala et al. 2019), very little data exists on even the best studied species, and essentially no data exists on domestication selection in oysters (but see McFarland et al. 2020). Few efforts to explore the potential for domestication selection during the larval stage of bivalve exist (e.g. McFarland et al. 2020), but future work is necessary to understand the potential impacts of domestication selection in the hatchery and the impact on natural populations.

Limited studies have been conducted to examine the genetic impact of shellfish restoration, though few such programs of any scale have been initiated (Gaffney 2006; Camara and Vadopalas 2009). One example involves the long-term genetic monitoring of a great scallop seeding program in the Bay of Brest, France (Morvezen et al. 2016). In this study, genetic analysis of samples from enhanced populations demonstrated little to no effect on wild population genetic diversity and it was suggested that gene flow from surrounding populations and/or the reproductive input of undetected sub-populations may have buffered the Ryman-Laikre effect (Morvezen et al. 2016). Hatchery-based supplementation programs are increasing in frequency and scope, and there should be additional opportunities to study the genetics of shellfish restoration within the US and elsewhere. Within the US, the Chesapeake Bay represents a fantastic test case to study the genetics of shellfish restoration on an unprecedented scale.

Oyster fishery decline in the Chesapeake Bay and the rise of oyster restoration

Oysters are the most economically important group of bivalve shellfish (FAO 2018), and provide critical ecosystem services (Coen et al. 2007; Grabowski and Peterson 2007). The eastern oyster (*Crassostrea virginica*) is a commercially and ecologically important species that inhabits coastal waters and estuaries along the Atlantic and Gulf Coasts of North America. Oysters provide critical ecosystem services such as improving water quality and clarity, providing complex hard-bottom habitat, and promoting biodiversity (Newell 1988; Coen et al. 2007). Oyster reefs promote biodiversity in several of ways: they provide refuge for fishes and invertebrates, act as coupling between benthic and pelagic systems, and bolster fishery production (Lenihan and Peterson 1998).

Eighty-five percent of oyster reefs have been lost worldwide (Beck et al. 2011), and oyster populations in the Chesapeake Bay have declined dramatically due to overharvest, habitat destruction, and disease outbreaks (Rothschild et al. 1994; Wilberg et al. 2011). In 1884, the Chesapeake Bay fishery reached peak harvests of 615,000 metric tons (Rothschild et al. 1994) but annual harvests have since dropped to ~3% of the fishery's peak (Tarnowski, 2016). In the Maryland portion of Chesapeake Bay, abundance of eastern oysters has been estimated at ~ 0.3% of the abundance before the onset of commercial fishing (Wilberg et al. 2011).

During this period of decline, a variety of management and restoration efforts have been undertaken to counter the effects of overharvest and habitat loss with limited success (Brooks 1891; Kennedy and Breisch, 1983; Kennedy, 1989). These initial efforts were implemented to maintain the fishery and successful restoration was measured by

increased harvests, which was a perspective that influenced fishery management policies for decades (Haven et al. 1981; Kennedy and Breisch 1983; Rothschild et al. 1994; Hargis and Haven, 1999). In addition, poor legislation, management decisions, and failure to react to available scientific information have contributed to resource mismanagement (Kennedy and Breisch 1983). Private oystering on leased grounds was promoted as a solution to the decline in wild populations (Brooks 1891), but the industry failed to shift from harvesting public beds to private culture (Kennedy and Breisch 1983; Keiner 2011). Maryland failed to establish an oyster aquaculture industry and wild populations were further harvested with little restrictions until recent political efforts. In the late 1880s, the Maryland Oyster Commission and the Maryland Oyster Police formed to improve and protect the fishery. In the early 1900s, a minimum size limit of three inches was enforced on the fishery as well as a shell tax to provide shell for replenishment activities (Kennedy and Breisch 1983). In 1960, the State of Maryland began the repletion program, which consisted of planting oyster shell dredged from the upper bay in different areas to restore habitat (Kennedy and Breisch 1983; Rothschild et al. 1994). Another part of this program involved moving seed oysters from state seed bars to areas of low recruitment to increase abundances in those areas.

Oyster restoration efforts have also focused on the placement of alternate substrate to serve as a site for oyster recruitment and growth (Nestlerode et al. 2007). Over time, these alternate substrates become covered with a layer of shells to produce a 3-D structure (a reef) that rises off the bottom, and thus provides a suitable habitat and substrate for settlement of oyster larvae. Oyster shell from the upper Bay deposits have been dredged and moved to selected restoration sites around the Bay to allow for natural

recruitment or to receive hatchery-produced oysters (Kennedy et al. 1996). In addition, shell from the oyster shucking and packing industry have been planted like dredged shell, with a portion of it used in hatcheries to produce of spat-on-shell (juveniles). Alternate substrates for reef construction have been used in the absence of significant shell deposits. These materials include granite, concrete, limestone marl, pelletized coal ash, and steel slag, as well as shells from other benthic organisms or fossil shells (Schuhmacher and Schillak 1994; Nestlerode et al. 2007; U.S. Army Corps of Engineers 2012).

More recently, the complex, ecological communities associated with oyster reefs have gained significant attention, and there has been a shift in management objectives toward rehabilitation of impaired resources and habitat in order to restore ecological function (Kennedy and Breisch 1983; Rodney and Paynter 2006; Grabowski and Peterson 2007). Instead of being maximized for harvest, restoration sites have been designed for maximum habitat and ecological improvement. In the late 1980's and early 1990's, a number of oyster restoration sites were established throughout Chesapeake Bay, for the purpose of habitat improvement (increasing bottom quality for growing oysters) or population enhancement (increasing oyster abundances). Some of these sites were designated sanctuaries (off-limits to fishing), while others allowed fishing after five years (reserves; MD DNR 2016). During 1961-2016, Maryland established 29 oyster sanctuaries, with the three largest sanctuaries established in 2010 in the Choptank River (Harris Creek, Broad Creek, and the Little Choptank). The Choptank River sanctuaries were designed with the goal of sustaining or improving 25% of the remaining oyster

resource in Maryland, allowing oysters to live longer, spawn without harvest pressure, and develop disease resistance over time (MD DNR 2016).

Since 2010, additional restoration efforts have involved deploying spat-on-shell in depleted oyster habitat (Mann and Powell 2007), although artificial reefs have been constructed in the three largest oyster sanctuaries since 2010 (Westby et al. 2016). The University of Maryland Center for Environmental Sciences Horn Point Laboratory (HPL) Oyster Hatchery has been responsible for the hatchery culture and deployment of juvenile oysters into these sanctuaries. The HPL Hatchery produces cultured spat on shell (juveniles) through mass spawning of local, wild, rotated broodstock. Since 2011, billions of spat have been planted in the Harris Creek Sanctuary (completed in 2016; Figure 2) with routine monitoring of restoration progress (primarily abundance) indicating success in achieving many of the initial goals of the program (Lane et al. 2011, Paynter et al. 2013, 2014; Westby et al. 2016). Monitoring results demonstrate that 98% of reefs in Harris Creek met minimum thresholds in oyster densities (number of oysters per square meter) and positive trends in water clarity compared to areas with fewer oysters (Westby et al. 2016). In addition, the first eastern oyster stock assessment for Maryland's three largest oyster sanctuaries was completed to estimate how abundance have changed over time and as a result of restoration (Damiano and Wilberg 2019).

Despite the large investment to restore oyster populations in Maryland over the past two decades, there has been little attempt to evaluate success or failure of these efforts (Kennedy et al. 2011). According to the 2016 Oyster Management Review conducted by the Maryland Department of Natural Resources on the sanctuaries established in Maryland in 2009 or 2010, survival of oysters remained the same or

increased since the creation of sanctuaries (MD DNR 2016). In addition, biomass in sanctuaries also increased, achieving the highest recorded biomass levels within the past 26 yeas of the Chesapeake Bay (MD DNR 2016). In general, pre- and post-restoration data at sites have not been recorded, which makes estimating success of restoration activities difficult (Kennedy et al. 2011). For example, in Maryland, 53 of 108 reefs receiving substrate and 59 of 92 receiving wild seed were monitored pre- and post-restoration.

Oyster restoration and supplementation programs are largely new, and rigorous evaluation of their efficacy is still in a trial-and-error stage. While many restoration programs (including Chesapeake Bay) use the total number of oysters planted and subsequent survival or planted individuals as a measure of success, such metrics are inadequate for assessing the long-term resilience of restored populations. The return of ecosystem services has also been an important metric to evaluating the success of restoration projects. Previous studies have shown that biodiversity, including genetic diversity is positively associated with the provision of ecosystem services (Hooper et al. 2005). Seagrass restoration projects have demonstrated the importance of genetic diversity in transplants (Hughes and Stachowicz 2004) and increased genetic diversity has been linked to enhanced ecosystem resistance to disturbance (Hughes and Stachowicz 2004; van Katwijk et al. 2009). However, despite such advances in restoration practices in other fields, many oyster restoration and supplementation efforts do not incorporate genetic monitoring into restoration programs or as a means of measuring success.

Recent research on the genetics of wild and restored oysters

The mechanisms that define the scale and pattern of population connectivity, genetic diversity, and effective population size remain poorly understood for eastern oysters. In marine species with high fecundity, high early mortality, and broad larval dispersal, the dynamics of even a relatively pristine population can be difficult to predict. While recent particle tracking models have resolved oyster larval dispersal in great detail (North et al. 2008; Narváez et al. 2012), tracking larval dispersal under natural conditions is extremely difficult, and thus much is still unknown about the fate of larvae during this period. Previous analyses using genetic markers support the view that eastern oysters have the potential for long distance dispersal among its populations (Reeb and Avise 1990; Rose et al. 2006), but much is needed to understand the processes affecting genetic diversity in this species, as well as understanding the genetic impacts that hatchery propagation and supplementation on natural populations, as large-scale efforts are currently underway within the Chesapeake Bay.

To date, only limited research has been conducted to characterize the potential genetic impacts of current or previous oyster supplementation efforts in the Chesapeake Bay, due in part to the lack of wild reefs exist in the region (reviewed in Beck et al. 2011). Rose (2008) investigated the genetic impact of out-planting disease resistant oyster strains in the Great Wicomico River. Utilizing Ryman-Laikre models (Ryman and Laikre 1991; Ryman et al. 1995) to investigate the genetic impact of a single year and sustained out-planting efforts, Rose (2008) detected substantial declines in wild population N_e except when the hatchery N_e was high (>100) and reproductive contribution was low (5-10%). These data provide one of the first glimpses of how oyster

supplementation can reduce genetic diversity of natural populations, though aquaculture lines used were already reduced in diversity compared to the recommended wild broodstock sources.

While the Ryman-Laikre model (Ryman and Laikre 1991) and its extensions (Waples et al. 2016b) are useful for predicting changes in N_e of wild receiving populations, the underlying assumptions of these models are often violated by oyster life-history features. First, the underlying model assumes a single population with wild and hatchery-produced components, which is often not the case with oysters (and rarely in nature). Next, Ryman Laikre models assume discrete generations, which is violated for many marine species including oysters that are relatively long-lived and have overlapping generations. While efforts to account for age-structure have been made for semelparous species (e.g. Waples and Do 1994), little is known how the Ryman-Laikre effect operates in iteroparous species. Therefore, alternative approaches will be required to better assess the genetic changes associated with hatchery-based restoration of oysters.

Fortunately, the genomic revolution of the early 2000s has given rise to a variety of more precise and powerful molecular methods that may make restoration planning more effective in the future. Recent technological advancements in next generation sequencing (NGS) have dramatically increased genotyping resolution, reduced associated costs, and together with the development of bioinformatic tools, have opened the door for population genetic studies in any species (Davey et al. 2011). Historically, conservation genetics has relied on the use of allozymes and microsatellites markers at low density throughout the genome (10s of markers), which only provided information on neutral genetic processes and variation. In transitioning to the genomics era, the use of genome-

wide single nucleotide polymorphisms allows profiling of both neutral and adaptive variation, and has provided insight into the genetic basis of local adaptation across an ever-growing number of non-model marine organisms with complex life-history features (e.g. Hauser and Carvalho 2008; Sanford and Kelly 2011; Silliman 2019; Bernatchez et al. 2019; Vendrami et al. 2019). These patterns revealed from adaptive variation, in many cases, can provide unique insights into a species' evolutionary potential and resilience (Stapley et al. 2010; Guo et al. 2015; Funk et al. 2016). The expansion of studies is due to the fact that technologies are improving and provide higher quality data at a decreasing cost. Importantly, many key biological questions can be answered with data from only a fraction of the genome (assuming equal spacing and genome-wide coverage), and genome reduction sequencing methods have become increasingly popular (e.g. restriction digest and capture enrichment sequencing (Peterson et al. 2012; Kozarewa et al. 2015). These methods allow for the analysis of high-resolution markers in many individuals, but also result in higher coverage and increased accuracy. In addition, sophisticated experimental, statistical, and modeling frameworks are being developed to differentiate candidate loci governing local adaptation from genomic variation driven by neutral processes (e.g. Günther and Coop 2013; Whitlock and Lotterhos 2015), and seascape genomics allow for contextualization of adaptive and neutral variation in marine environments.

Recently, sequencing, assembly, and annotation of the eastern oyster genome was completed (NCBI Projects: PRJNA379157 and PRJNA36014) which has paved the way for a variety of detailed studies of oysters including population genetics, adaptation, and physiological responses to the environment. This resource can improve our understanding

of evolution, facilitate selective breeding, enhance oyster cultivation practices, and restoration monitoring. This genome, and the genome of many other non-model species will be necessary tools to meet the challenges of conservation biology in a rapidly changing environment.

Research focus

This dissertation aims to characterize the genetic impact of a large-scale oyster restoration program, with the overall goal of understanding how characteristics of species life-history interact with hatchery practices (and logistical constrains of these practices) to shape genetic variation in populations that impact both short- and long-term persistence. To accomplish this, I used a combination of experimental, field, and modeling approaches (Figure 3) to examine short-term effects of oyster restoration (e.g. hatchery experiments) and longer-term effects (e.g. field analysis of diversity), including the creation of an individual-based model (IBM) to examine the genetic legacy of species life-history and hatchery practices under a large range of scenarios that could not be examined via field or experimental work. This dissertation contributes to a growing body of knowledge in understanding the genetic impacts of hatchery practices as well as for understanding the mechanisms, scale, and patterns of population structure and local adaptation in marine species with complex life-history features. While this work is based within the Chesapeake Bay, the implications of this work are transferable to other oyster restoration programs. Importantly, this is one of the first opportunities to examine the genetic impacts of large-scale hatchery-based shellfish supplementation.

In Chapter 2, I tracked genetic diversity changes resulting from hatchery production with two spawning designs (mass- and controlled-spawns) using microsatellite markers and parentage methods in an experimental hatchery context. Microsatellite markers were used to characterize basic descriptions of genetic diversity in restored and wild oyster populations as well as diversity of a restored reef in Harris Creek. Building off of the results from Chapter 2, Chapter 3 utilized high-resolution genomic data to compare genetic diversity of restored reefs in Harris Creek with variable planting histories and broodstock numbers to 'wild' (not intensively restored) Chesapeake Bay oyster reefs. These high-resolution genomic data permitted the investigation of population structure, local adaptation, and the extent at which environmental gradients influence genetic variation among these populations in a heterogeneous estuarine environment, which has not been extensively characterized for this species within the Chesapeake Bay. As a continuation of my exploration of the impacts of hatchery-based supplementation of oysters, in Chapter 4 I utilized the empirical data obtained from Chapters 2 and 3, to parametrize and create an individual-based genetic model to examine hatchery and wild oyster population dynamics within a large-scale supplementation program. The model was used to evaluate the impact of varying hatchery practices on wild population genetic diversity. Results of this work provide important baseline data for monitoring genetic diversity in Harris Creek and provide a reference point or benchmark for improving or informing current and future oyster restoration programs and their management. The findings of my doctoral research will hopefully inform management of hatchery populations and aid in the restoration of declining populations across an array of species with complex life-history features.

Figures Chapter 1

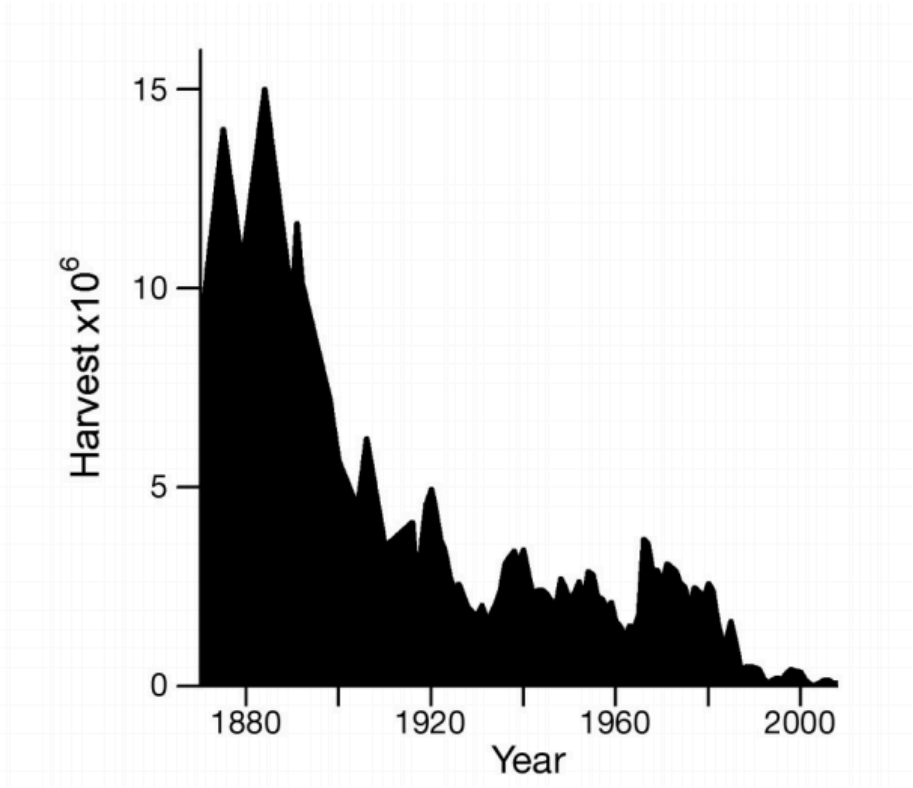


Figure 1. Oyster harvests in the Chesapeake Bay have substantially declined over the past 150 years (Wilberg et al. 2011).

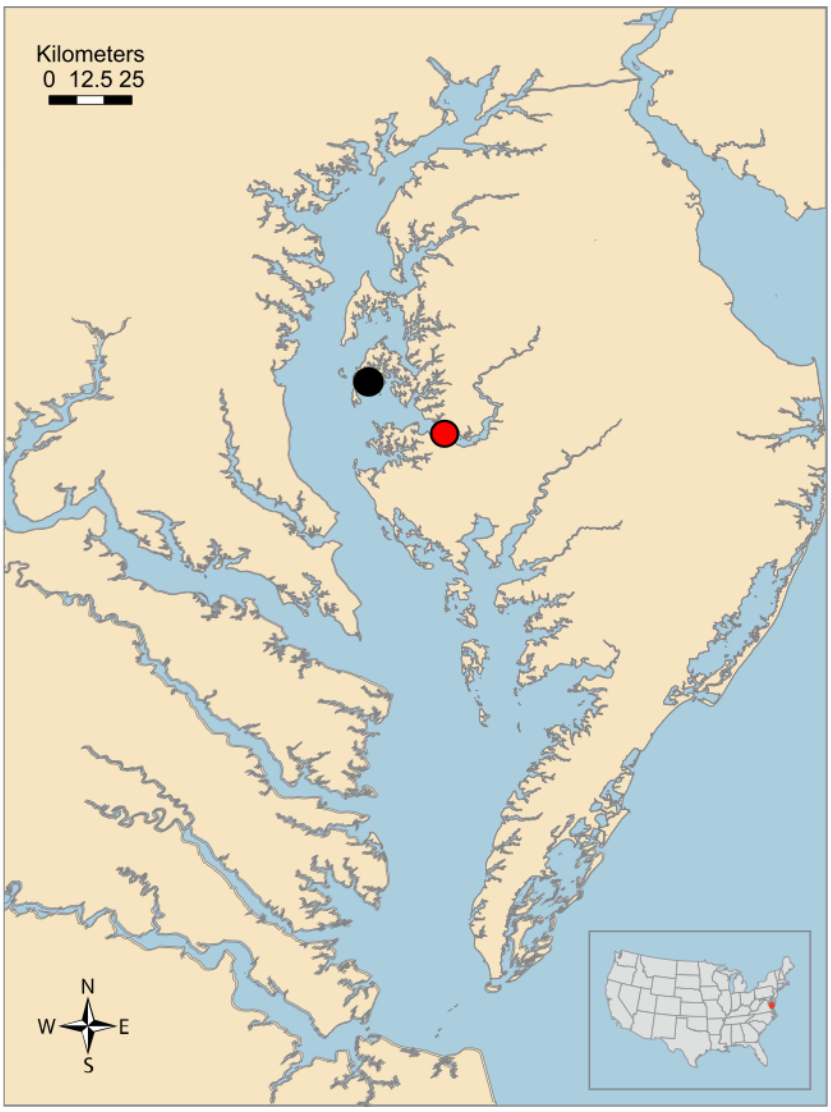


Figure 2. Location of the Harris Creek Sanctuary (black) and the Horn Point Laboratory Oyster Hatchery (red) in the Chesapeake Bay, Maryland.

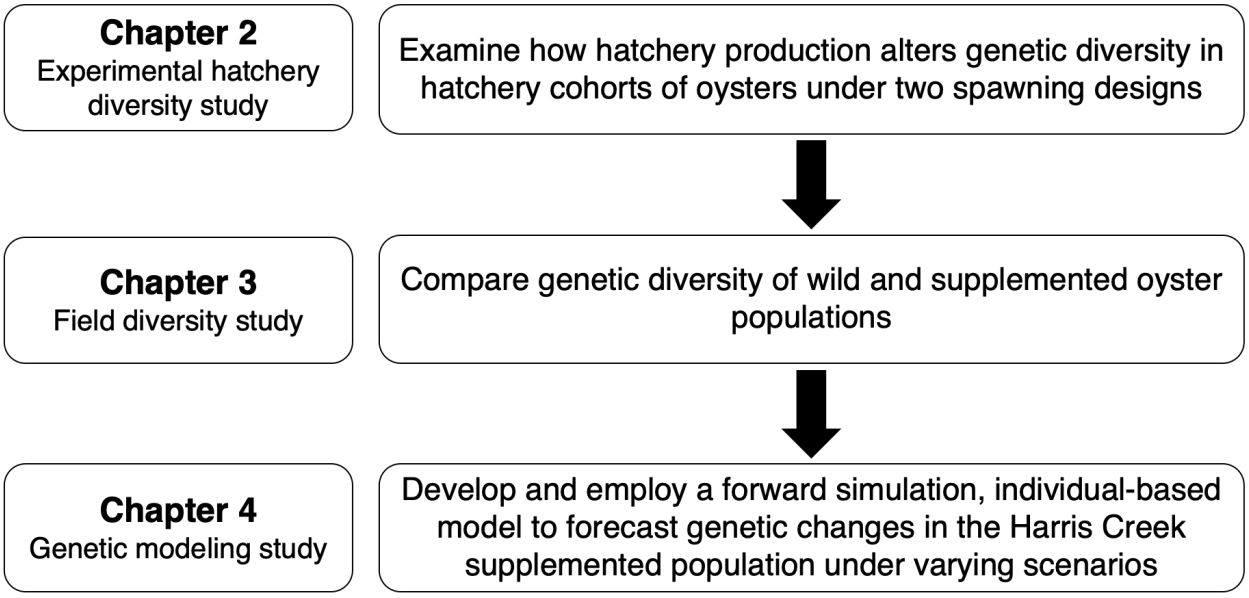


Figure 3. Conceptual diagram of the three research chapters that comprise this dissertation on the genetic impacts of hatchery-based oyster restoration. This topic is investigated within the Harris Creek sanctuary in Maryland using experimental approaches in Chapter 2, field approaches and high-resolution genomic data in Chapter 3, and a modeling approach in Chapter 4, which incorporates empirical genetic data obtained from previous chapters.

Chapter 2: Tracking genetic diversity in a large-scale eastern oyster restoration program: effects of hatchery propagation and initial characterization of diversity on restored vs. wild reefs

© Heredity
https://doi.org/10.1038/s41437-019-0202-6

Abstract

The release of hatchery-propagated fish and shellfish is occurring on a global scale, but the genetic impacts of these practices are often not fully understood and rarely monitored. Slow recovery of depleted eastern oyster populations in the Chesapeake Bay, USA has prompted a hatchery-based restoration program focused in the Choptank River, Maryland consisting of the mass release of hatchery-produced juveniles from local, wild broodstock. To evaluate potential genetic effects of this program, we (1) examined changes in genetic diversity (allelic richness, heterozygosity) and the effective number of breeders (N_b) over the hatchery production cycle with microsatellite-based parentage of natural, mass- and controlled-spawned cohorts, and (2) compared genetic diversity and effective population size (N_e) of a restored reef to wild source populations. Mass-spawned cohorts showed high variance in reproductive contribution, particularly among males, leading to a 45% average reduction in N_b from spawning adult numbers and higher relatedness—lower magnitude reductions in heterozygosity and significant reductions in allelic richness were also observed. While controlled-spawns (single-male fertilizations of pooled eggs) reduced male variance, overall reproductive variance (V_k) remained high. Finally, oysters sampled from a restored reef displayed comparable N_e,

genetic diversity, and relatedness to samples from wild populations, with no significant genetic differentiation among them. Overall, the hatchery-based results and initial field-based population genetic analyses suggest that despite reductions in diversity from parents to offspring owing to high V_k, enhancement with rotated, wild broodstock appears to have maintained genetic diversity in a restored reef population compared to proximal wild populations.

Introduction

Large-scale releases of plant and animal populations has increased worldwide to augment overexploited fisheries, forests, and wildlife (e.g. Laikre et al. 2010). For coastal fisheries, which have seen sharp declines over the last half century (Jackson et al. 2001; Myers and Worm 2003), the release of hatchery-propagated stocks has become an important component of fisheries management strategies, used for restocking, stock enhancement, or ranching in 100 s of species around the world (see Bell et al. 2008 for definitions; Laikre et al. 2010; Lorenzen et al. 2012). Restocking aims to re-establish a locally extinct commercial species and/or to restore depleted spawning biomass to a level where it can provide regular substantial yields, whereas stock enhancement aims at sustaining or improving fisheries in the face of decline (Bell et al. 2008; Lorenzen et al. 2012). Although the genetic risks associated with restocking and stock enhancement have been well described (e.g. Blankenship and Leber 1995; Lorenzen et al. 2012), these programs are rarely monitored (Laikre et al. 2010).

Restocking and stock enhancement are effective at increasing the abundances of depleted or declining populations in the short-term, but these programs risk dramatically

altering the genetic composition and diversity of populations, which can negatively impact long-term population resilience (reviewed in Frankham et al. 2010). For example, unintentional domestication selection (Frankham 2008), can result in the release of individuals with adaptive mismatches to local environmental conditions (e.g. Waal et al. 2013), and reduced fitness compared with their wild-born counterparts (Araki et al. 2007, 2008). Programs using a reduced number of broodstock can cause reductions in wild population genetic diversity and effective population size (N_e), leading to increased rates of genetic drift and increased potential for inbreeding, thereby limiting evolutionary potential (Allendorf and Ryman 1987; Christie et al. 2012a). When hatchery N_e is lower than wild (recipient) N_e and reproductive contribution to wild populations is high, reductions in diversity of wild populations can be acute, producing the so-called Ryman-Laikre effect, which is a concern for the enhancement of salmonids and other fish (Ryman and Laikre 1991; Gold et al. 2008; Christie et al. 2012a). Genetically aware, "conservation-based" strategies are increasingly used in enhancement and/or restocking programs to maintain high-diversity, self-sustaining populations (Caughley and Gunn 1996; Heggenes et al. 2006). However, less is known about the long-term genetic impacts of these programs and genetic monitoring is often limited or absent (e.g. Laikre et al. 2010).

Restoration of marine shellfish populations, primarily through juvenile seeding (Gaffney 2006; Laing et al. 2006), is increasing as coastal populations succumb to habitat degradation and overfishing (e.g. Beck et al. 2011). Although the potential for diversity loss is widely appreciated, monitoring is rare and few studies of the genetic impacts of these programs have been conducted (e.g. Gaffney 2006; Camara and Vadopalas 2009).

Characteristics of marine shellfish life-history, including high fecundity, high early mortality, and high variance in reproductive success (e.g. Hedgecock and Pudovkin 2011; Plough 2016; Plough et al. 2016), may exacerbate many of the underlying genetic risks associated with restocking or restoration, and can limit Ne to a small fraction of the census population size (He et al. 2012). Few studies have attempted to connect genetic changes in the hatchery (larval) phase to the genetic composition and diversity of wild or restored populations (but see (Hanley et al. 2016), and less is known about the Ryman-Laikre effect in iteroparous species like shellfish (Waples et al. 2016b). A number of laboratory studies have examined diversity during the larval stages of shellfish species (Boudry et al. 2002; Lind et al. 2009; Lallias et al. 2010), but less work has been done to connect these changes in the hatchery phase to the genetic composition and diversity in wild or restored populations. Except for work by Morvezen et al. (2016), there is a dearth of data on the genetic changes associated with shellfish hatchery-based enhancement and/or restocking programs from field samples or experiments, and no studies have monitored the genetic impacts of an intensive restoration program.

In this study, we examined the genetic impacts of a largescale restoration program for the eastern oyster Crassostrea virginica (Gmelin, 1791) in the Chesapeake Bay, with two primary objectives: (1) to characterize how genetic diversity changes during the hatchery production phase under two natural spawning protocols, and (2) to examine how genetic diversity of a restored reef population compares to surrounding wild reefs, to provide an initial assessment of how genetic diversity has been maintained or lost as a result of the restoration program. Oysters once supported major fisheries along the US East Coast but have declined to < 1 % of their historic abundances due to overfishing,

habitat destruction, and disease (Newell 1988; Rothschild et al. 1994). In the Chesapeake Bay, historically one of the most productive oyster fisheries on the US East Coast, oyster restoration activities of various sizes have been undertaken over the last half century to restore the fishery and ecosystem services provided by oyster reefs (e.g. Kennedy et al. 2011). Recently, a federal mandate to restore 20 Chesapeake Bay tributaries by 2025 has provided support for more comprehensive restoration in the Choptank River region (Maryland, USA), with the first of three sub-tributaries, Harris Creek (MD), completed in 2016 (Westby et al. 2017). The University of Maryland Center for Environmental Science's (UMCES) Horn Point Laboratory (HPL) Oyster Hatchery has been producing spat (juvenile oysters) on shell for the Harris Creek Sanctuary (and other locations), through natural (temperature-stimulated) group or "mass" spawning of wild Choptank River broodstock (e.g. Wallace et al. 2008). However, no genetic monitoring of this program has been conducted, and the potential impacts of spawning protocol, hatchery propagation, and planting on genetic diversity are unknown.

To quantify genetic changes during hatchery production (objective 1), we conducted parentage analyses on six mass-spawned cohorts and three controlled-spawned cohorts (isolated natural spawns with single-male fertilizations) using up to nine microsatellite markers. Parental contribution (variance in reproductive success) and genetic diversity metrics, including effective number of breeders, were calculated in the offspring at the spat (planting) stage and compared with their parents. To quantify the broader population genetic impact of this restoration program (i.e., genetic diversity in restored vs wild populations; objective (2), we estimated and compared N_e, genetic diversity metrics, and relatedness among five wild broodstock sources, an aquaculture

line, and a restored Harris Creek reef sample planted in 2012, to provide an initial assessment of how diversity is maintained following enhancement with hatchery-produced cohorts. This work provides important baseline data for monitoring genetic diversity of oyster restoration efforts in Chesapeake Bay, and more broadly, for examining genetic impacts of other large-scale shellfish restoration efforts ongoing or planned in the US (e.g., Brumbaugh and Coen 2009; Dinnel et al. 2009; Holley et al. 2018).

Materials and Methods

Spawning methods, collection of broodstock or hatchery sources, and sampling of spat

Oyster cohorts (offspring from mass- and controlled-spawns) were produced at the UMCES HPL Oyster Hatchery in Cambridge, MD, USA over the summers of 2014–2016.Wild Choptank River broodstock was used for the production of all cohorts (Table 1), except for cohort M6, which was produced from the Louisiana-derived "LoLA" aquaculture line, obtained from the Aquaculture Genetics and Breeding Technology Center (ABC) at the Virginia Institute of Marine Sciences. All broodstock were conditioned in the HPL Oyster Hatchery at ambient salinity and 20 °C for 6–8 weeks prior to spawning. The HPL Oyster Hatchery produces all spat on shell for restoration planting using a mass-spawn protocol

For the mass-spawned cohorts (M1–M6; Table 1), groups of ~ 50 ripe broodstock were placed in an open aquarium (spawning table) with heated (~ 28–30 °C) flowthrough filtered Choptank River water (~10–12 ppt; hereafter seawater) to initiate spawning. As

oysters spawned, females and males were removed from the spawning table and placed into separate containers (one for all males and one for all females) and allowed to finish spawning. When spawning ceased, pooled eggs were counted volumetrically, and an appropriate amount of pooled sperm was used to fertilize the pool of eggs (Galtsoff 1964). For all mass-spawned cohorts, spawning order was recorded for males and females and fertilization occurred no more than 45 min after the first oyster was observed to spawn. For the controlled-spawned cohorts (C1–C3; Table 1), broodstock oysters were placed in individual 1.8 L aquaria (Aquaneering, San Diego, CA) arranged on a multi-tiered rack with heated, flowing seawater (independent in- and out-flow for each vessel). When evidence of spawning was observed, water to the particular aquaria was turned off to allow gamete accumulation. The time of spawning for each oyster was noted. Once a sufficient number of spawning males and females were identified, eggs were pooled, counted volumetrically, and divided evenly among eight containers (the number of males that spawned) for individual male fertilizations (no sperm competition). Cohorts M1–M3 were spawned in June 2014, M4–M6 in June 2015, and C1–C3 in July 2016. Tissue samples (adductor muscle or mantle) were collected from all adult broodstock and preserved in 70–95% ethanol. All broodstock adults (LoLA and wild) were also utilized in the population genetic analyses of diversity, population structure, and N_e (see below). In addition, 48 adult oysters (mixed-age classes) were sampled from a restored reef in the Harris Creek Sanctuary (Seed 2; 38.71298 N, −76.31985 W) in 2015, which had been planted with hatchery-produced oysters in 2012 (Fig. 1). Whole individual spat were randomly sampled from each cohort and stored in tissue lysis buffer at − 80 °C prior to extraction. One hundred twenty spat were sampled for cohorts M1–M3 and C1–C3, and

115 spat were sampled for cohorts M4–M6 (Table 1). All details concerning the larval culturing and juvenile nursing are presented in the Supplementary Materials and Methods.

DNA extraction and microsatellite genotyping

DNA was extracted from adult tissue and spat using the E. Z.N.A.® Tissue DNA Kit (Omega-Biotek, Norcross, GA) or the Qiagen DNeasy Tissue kit (Qiagen Inc, Valencia, CA), following the protocol for animal tissues. DNA concentrations were estimated using a Qubit Fluorometer (Thermo Fisher Scientific Inc., Waltham, MA, USA) and all samples were diluted to a concentration of 10 ng μL^{-1}. Genetic analysis was conducted on nine microsatellite loci previously developed by Brown et al. (2000), Reece et al. (2004), Carlsson and Reece (2007), Wang and Guo (2007), and Wang et al. (2009) and were named as in the source publication (see Supplementary Table S1 for loci ID). Samples from the restored Harris Creek reef and cohorts M1–M6 (broodstock parents and offspring) were genotyped at all nine microsatellite loci while cohorts C1–C3 were genotyped at six of these loci (Supplementary Table S1). Polymerase chain reactions were carried out in volumes of 12.5 µl with 1Å~ GoTaq® flexi PCR buffer (Promega, Madison, WI, USA), 200 µM dNTPs, 0.04 µM of the M13- tailed forward primer, 0.16 µM of both the reverse primer and the M13 fluorescent dye-labeled primer, and 1.25 units of GoTaq® Flexi DNA Polymerase (Promega, Madison, WI, USA). The PCR cycling program was run in two phases; the first to amplify the target microsatellite marker, and the second to incorporate a fluorescently labeled dye attached to the M13 tag sequence (see (Schuelke 2000) for details).

Microsatellite PCR products were pooled, precipitated, and shipped dry to the Arizona State University DNA Laboratory, where fragment analysis was performed on an ABI 3730 capillary sequencer. Genotypes were scored by eye using the LIZ500 (Applied Biosystems) as an internal size standard on the Peak Scanner v1.0 software (Applied Biosystems) and verified using Genemapper (Applied Biosystems). For each cohort, 25% of the offspring were genotyped in duplicate to calculate genotype error rate at each locus.

Parentage assignment

Parentage assignment was performed in CERVUS v.3.0.7 (Marshall et al. 1998; Kalinowski et al. 2007) using the parent–pair analysis option and the Delta method. For each cohort, simulations in CERVUS were run to estimate the critical values of Delta for each cohort with strict (95%) and relaxed (80%) levels of statistical confidence in assignment (Kalinowski et al. 2007). Each set of simulations was specific to the number of loci typed, the number of candidate parents, and the locus error rates for that particular cohort. Putative candidate parent pairs were considered accurate if the offspring-dam-sire trio statistical confidence was ≥ 95% (strict confidence). We excluded individuals assigned at relaxed confidence (80%), or those that were not assigned at all, from downstream analysis.

Effective number of breeders (N_b) and related statistical analyses

After assigning offspring to broodstock parent pairs, the mean number of offspring (k) per parent and variance in reproductive success (V_k) was calculated for male

and female broodstock from each cohort. V_k was standardized by the mean number of offspring in a family (σ^2/μ^2) and is reported as such from here on. From these parameters, the inbreeding number of breeders (N_{bI}) was estimated for each sex following equations 1 and 2 from Christie et al. (2012). Equation 6 from Waples and Waples (2011) was used to estimate N_b (hereafter N_{bS}) using information from sibship reconstruction analysis in COLONY v2.0.6.1 (Jones and Wang 2010). We then compared the estimated N_{bI} to the number of successful broodstock spawned (N) for all cohorts. Differences in V_k between males and females were tested within each cohort using a two-tailed parametric F-test in R (R Core Team 2017). Skew in reproductive contribution of males and females was assessed via Chi-square goodness-of-fit tests in R, with the null hypothesis of equal reproductive contribution.

Population genetic analyses and statistical comparisons

Allelic richness (A_r) for each locus and cohort was calculated in FSTAT v2.9.3.2 (Goudet 2002), which uses a sample-size independent rarefaction analysis of allelic richness. Observed heterozygosity (H_o) was calculated in the R package "hierfstat" (Goudet 2005). Tests for Hardy–Weinberg equilibrium (HWE) were performed in GenoDive v2.0b27 (Meirmans and Van Tienderen 2004) using 9,999 one-sided permutation tests, with P-values corrected for multiple tests via sequential Bonferroni correction (Rice 1989). For markers that differed significantly from HWE, null allele frequencies were estimated using Microchecker v.2.2.3 (Van Oosterhout et al. 2004) and FreeNA (Chapuis and Estoup 2007). To test whether parents and offspring from a given cohort differed in genetic diversity metrics (A_r, H_o), paired Wilcoxon signed-rank tests

between parents and offspring were performed (see Nei 1987, p. 183), with P-values adjusted for multiple tests using the false discovery rate correction (Benjamini and Hochberg 1995). Relatedness was estimated for parents and offspring (grouped by broodstock source) from hatchery cohorts and for the Harris Creek sample using the Lynch and Ritland method (Lynch and Ritland 1999) in COANCESTRY v1.0 (Wang 2011), because it performs well in panmictic populations (e.g., Oliehoek et al. 2006). Contemporary effective population sizes (N_e) were calculated for adult broodstock source populations, hatchery-produced oysters, and a restored population from Harris Creek reef using the linkage disequilibrium (LD) method (Waples and Do 2008) implemented in NeEstimator v2.01 (Do et al. 2014; minor allele frequencies < 0.05 excluded). For this analysis, if there were multiple adult broodstock groups from the same reef/population, they were pooled within population (i.e., Sandy Hill 1, 2, and 3 were combined into a single Sandy Hill population; see Table 3). Finally, all wild broodstock source populations were also combined into a single population for an overall wild Choptank River N_e estimate. Although the mixed-age adult samples can be used to estimate N_e (Waples et al. 2014), the offspring cohorts, which represent a single age-class, were used to estimate N_b (Waples 2005). To obtain offspring N_e estimates, raw offspring N_b estimates (LD method) were adjusted according to Waples et al. (2014) using three life-history traits: adult life span = 15 (10–20 years in undisturbed populations, Powell and Cummins 1985), age at maturity (α) = 2 (averaged values from Galtsoff 1964; Rothschild et al. 1994; Powell et al. 2013), and variation in age-specific fecundity CVf = 0.65 (from Mroch et al. 2012; Mann et al. 2014).

To examine the extent of genetic differentiation among the broodstock sources, hatchery-produced offspring, and restored Harris Creek sample, F_{ST} (Nei's G_{ST}; Nei 1973) was calculated in GenoDive. Heterogeneity in allelic frequencies between pairs of samples was tested in GenoDive (G-test; 9,999 permutations). Significance criteria were adjusted for the number of simultaneous tests using sequential Bonferroni corrections (Rice 1989). F_{ST} values (Weir 1996) were recalculated in FreeNA using the 'excluding null alleles' method (Chapuis and Estoup 2007) to correct the positive bias induced by the presence of null alleles on F_{ST} estimation. Finally, population structure was assessed among broodstock population sources and their offspring, with discriminant analysis of principle components (DAPC) in the R package 'Adegenet' (Jombart 2008).

Results

Genotyping results and null alleles

Across the nine cohorts examined, a total of 1,299 individuals were genotyped, 846 (141 parents, 705 offspring) of which comprised the mass-spawns and 405 (45 parents, 360 offspring) of which comprised the controlled-spawns (Table 1). A total of 12,547 genotypes were scored with an average error rate of 1.9% across all samples. One locus (Cvi6) was excluded from cohort M6 because of inconsistencies with allele scoring. Null alleles were detected in five (RUCV3, RUCV114, Cvi2i23, Cvi2i4, and Cvi6) out of the nine loci (Table S1), which is consistent with previous studies using the same markers (Rose et al. 2006; Arnaldi et al. 2018). Five out of the nine microsatellite loci showed Hardy–Weinberg disequilibrium with an excess of homozygotes in most of the studied

populations (Supplementary Table S1). All population genetic analyses were performed with and without these loci and the two data sets provided similar results.

Parentage results, variance in reproductive success, and changes in diversity

Parentage assignment of offspring to parental pairs in CERVUS ranged from 74 to 89% (mean 83.2%) at the 95% confidence cutoff level (Supplementary Table S2). We observed no difference in assignment success between mass- and controlled-spawns (Wilcoxon signed-rank test; $P > 0.05$). Parentage assignment results were then used to assess variance in reproductive success and skew among parents within mass-spawned and controlled-spawn cohorts.

For the mass-spawned cohorts, reproductive success varied widely among individuals—individual males produced 0–58.1% of assigned offspring (mean = 8.7%, mode = 5%) and individual females produced 0–32.6% of assigned offspring (mean = 8.3%, mode = 0%). The standardized variance in reproductive success was higher in males than females in all mass-spawned cohorts (64% higher mean V_k compared with females; F-test, $F_{0.05,(5,5)} = 0.143$, $P = 0.03$; Fig. 2). Among individuals that were observed to have spawned, 5.8% of males and 5.6% of females effectively produced no offspring (averaged over all six cohorts). Male reproductive contribution differed significantly from the expectation of equal contribution, indicating evidence of significant reproductive skew in all six mass-spawns (Supplementary Figure S1). Female reproductive contribution was also highly variable and differed significantly from the expectation of equality in five of the six mass-spawned cohorts; however, the skew was less than that of males (Supplementary Figure S1, i.e., less significant Chi-square

deviations). Male reproductive contribution appeared to be affected by male spawning order, and males that spawned later contributed more offspring (P = 0.03, Supplementary Figure S2). Overall, high variance in reproductive success among parents resulted in a substantial reduction in N_{bI} (average 45%) and relatively low N_{bI}/N ratios (Table 2). N_b estimates varied depending on the method used, but all N_{bS} estimates were lower than N_{bI} (N_{bI} ranged from 7.62 to 17.58, whereas N_{bS} ranged from 4.13 to 9.85; Table 2). The magnitude of reduction in N_{bI} appeared to be affected by the sex ratios of each cohort, with a marginally significant trend (P = 0.09, compared with alpha = 0.05) of decreasing variance for sex ratios closer to one (Table 2). For example, the highest (0.84) N_{bI}/N ratio was observed in cohort M3, where the sex ratios were closest to one, whereas the lowest (0.36) N_{bI}/N ratio was observed in cohort M2, where the sex ratios were farthest from one (0.4). Finally, statistically significant reductions in allelic richness (A_r) from parents to offspring were found in all the mass-spawned cohorts (Wilcoxon signed-rank test $P<0.03$, Table 2), while reductions in H_o between parents and offspring were not significant ($P> 0.05$, Table 2). Results for the controlled-spawns were generally similar to the mass-spawns with high variance in reproductive success among individuals. Individual males produced 3.3–42.5% of offspring (average = 12.5%, mode = 3.26%), whereas individual females produced 0–57.5% of offspring (average = 14.3%, mode = 1.1%). Patterns of V_k between males and females followed the opposite trend of the mass-spawns, with higher variance for females compared with males (65% higher than males; F-test, $F_{0.05,(2,2)} = 0.375$, $P>0.05$; Fig. 2). Among individuals that were observed to have spawned, 100% of males had offspring assigned to them vs. 95.2% of females, both of which are higher than in the mass-spawned cohorts. Again, the high V_k among parents led to a reduction in N_{bI}

(average = 47%) and relatively low N_{bI}/N ratios (Table 2). N_{bS} was lower than N_{bI} in all controlled-spawned cohorts (Table 2). We focus on N_{bI} for the remainder of the study because it can directly be compared with the initial number of spawning adults. Statistically significant reductions in A_r between parents and offspring were found in all the controlled-spawn cohorts (Wilcoxon signed-rank test $P< 0.03$, Table 2), whereas H_o was maintained in the offspring of cohorts C1 and C3 and increased in the offspring of cohort C3 ($P>0.05$, Table 2).

Global relatedness was higher in all hatchery-produced offspring (Supplementary Figure S3; average increase = 0.065). Relatedness in wild broodstock sources ranged from 0.002 to 0.011, whereas the mean relatedness of wild hatchery-produced offspring ranged from 0.03 to 0.129. As expected, relatedness in LoLA broodstock parents (0.117) was higher than wild broodstock sources and also increased in the hatchery-produced offspring (0.208). Offspring from the controlled-spawned cohort broodstock source (TB) had the smallest relative increase in relatedness (0.028), whereas offspring within Chlora Point (CLP; mass-spawned cohort) exhibited the greatest relative increase in relatedness (0.118). Relatedness in the Harris Creek restored sample was comparable to that in wild broodstock sources and was relatively low overall (0.012).

Genetic diversity and differentiation among wild broodstock and restored oysters

Genetic diversity (Ar and Ho) of wild adult broodstock ranged from 9.33 to 14.38 and 0.51 to 0.68, respectively (Table 2). The LoLA (aquaculture) adults displayed the lowest allelic richness overall (8.56; Table 2) but had comparable levels of heterozygosity (0.68). Overall, the restored Harris Creek sample had the second highest

allelic richness (14.19) and comparable levels of heterozygosity (0.59) to wild broodstock populations. Pairwise F_{ST} analyses (Nei's G_{ST}) revealed low genetic differentiation among most broodstock sources. All pairwise F_{ST} estimates between wild broodstock source populations were small, ranging from − 0.001 to 0.022 (Table 3)—none were significant after Bonferroni correction. Recalculating F_{ST} excluding null alleles in FreeNA produced very similar estimates to un-corrected values (Supplementary Table S4). Pairwise F_{ST} estimates between the Harris Creek and wild broodstock source populations were small and non-significant, ranging from − 0.001 to 0.019 (Table 3). Pairwise F_{ST} estimates between the LoLA broodstock adults and Chesapeake Bay wild populations were higher, ranging from 0.054 to 0.079, and all comparisons (adult broodstock: BBY, CLP, HC, TB, SH1, SH2, and SH3) were statistically significant (G-test; $P<0.001$). Interestingly, pairwise FST estimates between hatchery-produced offspring (i.e. the cohorts produced from wild broodstock) and wild broodstock parent sources were typically higher than comparisons among only adult broodstock ($0.011 < F_{ST} < 0.148$; Supplementary Table S5). Similar to the FST results, analyses of population structure via DAPC revealed two or three major clusters, with LoLA broodstock adults (and offspring) grouping distinctly from Choptank River wild broodstock and the restored Harris Creek sample (Figs. 3a, b). More structure was evident among the offspring cohorts (e.g., SH2 and SH1 vs. SH3; Supplementary Figure S4), and broodstock sources and offspring tended to group together, as expected (Fig. 3a). More subtle genetic differences were observed between Sandy Hill (SH) broodstock samples (e.g., SH2 adults vs. others) despite being sampled from a common population (Fig. 3b).

The Harris Creek sample clustered in the center of the five other wild Choptank River broodstock sources (Fig. 3b).

Effective population sizes of wild broodstock sources and the Harris Creek population

Contemporary effective population size (N_e) estimates varied among the different wild (broodstock source) populations but were fairly consistent overall, ranging from 75.3 to 129.6 with a mean of 102. Most of the estimates were bounded at the 95% confidence limits (95% confidence interval range 52.1–119.2; Table 4). Two Ne estimates were unbounded—Black Buoy and Chlora Point, but these had relatively small sample sizes (N = 34 and N = 22, respectively). The N_e estimate for the pooled wild broodstock sources was highest at 366.8 and had the largest 95% confidence interval range (200.1–1424.3). The N_e estimate for the Harris Creek sample was similar to the range of values estimated for wild populations at 68.3—confidence limits for the wild populations and the restored sample overlapped substantially (Table 4). As expected, N_e estimates were systematically higher in wild populations compared to the representative aquaculture line (LoLA), and effective sizes were systematically higher in wild broodstock parents compared to their hatchery-reared offspring (Table 4).

Discussion

Changes in genetic diversity resulting from hatchery propagation and spawning protocol

Across all hatchery-produced cohorts of oysters (mass- and controlled-spawns), offspring displayed lower genetic diversity, higher global relatedness, and reduced N_b

and N_e compared with their wild adult progenitors. The reduction in N_b was the most substantial among the diversity metrics examined, reflecting the high variance in reproductive success observed among parents. As expected, patterns of diversity loss differed between the two spawning protocols, but the controlled spawns did not reduce overall reproductive variance (see below). Indeed, controlled fertilizations of oysters may not necessarily reduce diversity loss during hatchery cultivation. For example, Boudry et al. (2002) found high V_k of cohorts of Pacific oysters despite controlled fertilizations. Post-spawning genotype-dependent larval mortality may also contribute to high V_k among parents (e.g. Plough 2016; Plough 2018).

In the mass-spawned cohorts, where we expected V_k to be greatest, we observed higher skew in males compared with females, which is consistent with results from previous studies of fish and shellfish (e.g. Bekkevold et al. 2002; Boudry et al. 2002). The high male variance can possibly be explained by differences in sperm quality or quantity among males in the competitive, pooled fertilization environment of a mass-spawn. Sperm competition can influence the proportion of eggs fertilized and differences in sperm traits (quality), such as sperm motility and velocity, can lead to variable contributions among males (Gaffney et al. 1993; Withler and Beacham 1994; Wedekind et al. 2007). However, male reproductive variance may not be attributed to sperm traits alone (Linhart et al. 2005; Kaspar et al. 2007). Fertilization rates may also depend on gametic compatibility (e.g., Gaffney et al. 1993), which was not tested in this study. Other characteristics of the mass-spawn protocol, such as the timing of gamete release, may also lead to imbalanced reproductive contributions. For example, we found that males that released gametes later in the mass-spawns contributed more offspring

(Supplementary Figure S2). Fertilizing with pooled sperm from natural spawns is a common practice in shellfish restoration hatcheries (e.g., in Maryland, HPL Oyster Hatchery; in Martha's Vineyard, Massachusetts, Emma Green-Beach, Martha's Vineyard Shellfish Group Inc., personal communication; and in New York Harbor, Rebecca Resner, Billion Oyster Project, personal communication), but the inability to control the timing of gamete release or gamete output may increase variance in reproductive success among parents.

In contrast to the mass-spawn results, the controlled-spawns effectively reduced male reproductive variance, but increased female reproductive variance, resulting in a similar reduction in N_b (low ratios of N_b/N), and effectively negating any benefits from reducing the male variance component. The inadvertent increase in female variance likely resulted from several factors associated with the controlled-spawning design, which relied on natural spawns within self-contained vessels. First, the number of eggs per female was not normalized across individuals before fertilization (they were pooled and then divided evenly); thus, differences in fecundity or effective fecundity among females could have contributed to female reproductive variance. Second, egg quality or viability may have varied among females, with reduced egg viability at the time of fertilization for females that spawned earlier compared with females that spawned closer to the time of time of fertilization. Eggs from some females were not fertilized for 2.5 h after the first observation of spawning (see Supplementary Table S7 for detailed information on timing of gamete release), and thus, they may have been less viable at the time of fertilization. Though we lack detailed information on temporal changes to gamete viability in *C. virginica*, an experimental study of gamete viability in the broadcast spawning bivalve

Cerastoderma edule showed that percent fertilization decreased 50% after 2 h (André and Lindegarth 1995), indicating that a 2.5-hour window could reduce fertilization success during spawning. Natural spawning protocols are clearly less ideal than directed fertilizations via strip-spawning of adults. However, the goal of these experiments was to control fertilizations (and sperm competition) within the framework of typical HPL Oyster Hatchery protocols that use natural spawning for oyster restoration. Employing strip-spawning would make large numbers of pairwise fertilizations logistically easier, but this is not always preferred or possible, especially if programs are broodstock limited, and differential survival of larvae can still result in skewed contributions and high V_k among parents (Boudry et al. 2002; Lallias et al. 2010). Overall, natural, individual spawns are challenging to execute on a small-scale, and thus, this practice is not likely to be viable for large-scale restoration hatcheries. Overcoming large variance in family sizes remains a major obstacle towards maximizing genetic diversity and N_e in hatchery-propagated individuals for restoration. The findings from this study, that there is an average loss of ~ 45% of the N_b between parents and offspring in the mass-spawns, could help to inform decisions about spawning practices and the number of broodstock used for eastern oyster or other shellfish restoration programs.

Genetic differentiation between wild, hatchery-produced, and restored oysters

Wild broodstock source populations within the Choptank River showed no genetic differentiation, which is consistent with previous studies reporting high gene flow among oyster populations in Chesapeake Bay (e.g. Rose et al. 2006) and Delaware Bay (He et al. 2012). Despite low F_{ST} values overall, Rose et al. (2006) observed a weak

pattern of isolation by distance among oyster populations in Chesapeake Bay, indicating some barriers to gene flow over 100 s of km. High gene flow among mid-Bay populations may also be driven by a long history of human-mediated adult and juvenile oyster movement, planting, and transplantation throughout the Chesapeake Bay (Kennedy and Breisch 1983; Mann and Powell 2007). Wild Choptank River and restored Harris Creek reef oysters differed genetically (statistically significant pairwise F_{ST}) from the LoLA aquaculture line, which is expected for a selectively bred, "closed" hatchery line that was developed with genetically divergent wild progenitors from Louisiana (Appleyard and Ward 2006; Champagnon et al. 2012). Finally, the restored Harris Creek oyster population showed no genetic differentiation from the local broodstock populations (e.g., Table 3 and Fig. 3), which is consistent with the fact that these local populations were used to produce oysters that are planted in Harris Creek. However, the power provided by eight microsatellite markers is unlikely to be sufficient to resolve fine-scale population structure among reefs or restored populations over such a small geographic scale.

Comparisons of contemporary effective population sizes between wild and restored reefs

Contemporary N_e estimates in this study agree generally with the magnitude of values reported for eastern oyster populations in the Delaware Bay (37–437) by He et al. (2012), but are lower than those reported for the James River (535–1 516) by Rose et al. (2006), and higher than that reported for the Delaware Bay (33.8) by Hedgecock et al. (1992). Although there is a rather wide range of N_e estimates reported across locations, most estimates are fairly low ($N_e < 1000$). N_e estimates for two of the broodstock sources

in this study were unbounded (BBY and CLP; sample sizes < 34, which is low for N_e estimation via the LD method; (Waples and Do 2010), but estimates for the remaining populations were bounded, and thus provide insight into the range of possible N_e for the wild broodstock populations in the Choptank River (75.3–129.6). Compared with estimates for wild oysters in the Choptank River, the restored oyster sample from Harris Creek had similar levels of genetic diversity and effective population size (N_e = 68.3), which suggests that the Harris Creek restoration program has not caused significant declines in genetic diversity, at least based on the single reef sampled and metrics examined.

Though we did not detect major declines in diversity for this restoration program, genetic analyses of similar restocking or enhancement efforts in marine fish have yielded mixed results in terms of the severity of associated genetic changes. Although some studies show that stock enhancement produced effectively no change to population genetic diversity (e.g., Heggenes et al. 2006; Gow et al. 2011; Katalinas et al. 2017) other studies have shown rapid declines in diversity and severe Ryman-Laikre effects after only a few generations (Gold et al. 2008; Karlsson et al. 2008; Christie et al. 2012a). Disparate results among studies may be driven by specific characteristics of the program (i.e., husbandry practices), the initial status of wild populations, or possibly the design of the monitoring study itself (e.g., number of samples, type and number of markers used). In the current study, a lack of salient genetic decline may reflect specific HPL hatchery broodstock management and spawning practices that are implemented to minimize declines in genetic diversity. For example, the HPL hatchery uses 100–1000 s of wild broodstock each year while past broodstock are rotated out and retired after 3–4 years of

use (Stephanie Alexander, HPL oyster hatchery, personal communication). Moreover, deliberate out-planting of many distinct cohorts of spat (i.e., produced from spawns with different sets of parents), ensures that a given restored reef will comprise multiple spawns, from multiple years, from hundreds of parents. Of course, caution must be exercised when inferring the genetic impact of oyster restoration in Harris Creek based on the sampling of a single restored reef. The analysis of additional restored populations with variable planting histories, and the use of larger marker data sets (e.g., 1000s of SNPs) is ongoing and should provide a more conclusive picture of the genetic impacts of hatchery propagation and planting on in these populations. Effective population size (N_e) estimates in marine animals tend to be much smaller than the census population size (N) (e.g. Hedrick 2005; Palstra and Ruzzante 2008), and extremely low N_e/N ratios (10^{-3}–10^{-6}) have been reported in many marine invertebrates and fishes (Hedgecock 1994; Hauser et al. 2002; Hedrick 2005; Hauser and Carvalho 2008), perhaps suggesting high variation in reproductive contribution among adults. So, how should the N_e estimates of oysters in this and previous studies be considered in a conservation or management context? Applying relatively cautious conservation thresholds (e.g. the 50/500 rule; Franklin 1980; Franklin and Frankham 1998), N_e estimates of oysters are on the order of magnitude of what is required for inbreeding avoidance, but are less than the prescribed target for preserving long-term evolutionary viability or quantitative genetic variation ($N_e > 500$). However, concerns about the genetic risks for populations with small N_e may be slightly overblown if there is frequent gene flow (i.e. replenishing genetic variation despite low N_e), which is typical for populations of marine animals like oysters that have

a highly dispersive larval stage (e.g. Gaffney 2006; Hauser and Carvahlo 2008; Palstra and Ruzzante 2008).

Although low N_e estimates may overstate the level of genetic risk to some marine populations, it has also been argued that estimates of N_e/N ratios in marine species may be artificially low (downwardly biased) when true N_e is actually quite high (Palstra and Ruzzante 2008; Waples 2016). Using simulations, Waples (2016) examined the range of life-history characteristics (e.g., longevity, age at maturity, fecundity, and variance in reproductive success that increase with age) that would be required to generate tiny N_e/N ratios (e.g., < 0.001) observed in empirical studies of fish and shellfish, and determined that unless very large sample sizes were used (e.g., 10s of thousands), N_e estimates could be seriously biased. This does not mean that all low N_e estimates are wrong, but that we should be very cautious about the interpretation of the N_e estimates in this and previous studies using relatively few markers and small sample sizes. If estimates across samples or markers are bimodal in distribution (very low and high) and/or are typically unbounded, a very large N_e cannot be rejected out of hand. Given that the N_e estimates in this study are generally consistent with previous Ne estimates of oyster populations in the Chesapeake Bay (and most estimates were bounded), it seems likely that they are reflective of truly low N_e for populations in the Choptank River.

Conclusions

This study provides an initial examination of the genetic impacts of a large-scale eastern oyster restoration program, and the overarching findings are relevant for future eastern oyster restoration programs as well as for the restoration of other marine shellfish

with similar life-history features. Overall, hatchery-based propagation of oysters led to substantial reductions in diversity and the N_b from parents to offspring (a decline of ~ 45% in N_b on average), which can be explained by the high variance in reproductive success (V_k) among adults in mass-spawns and hatchery-based larval culture. Experiments employing controlled, natural spawn protocols with independent fertilizations of pooled eggs failed to reduce V_k compared with mass-spawns. Despite high V_k and diversity loss in the offspring of individuals cohorts, estimates of genetic diversity metrics and Ne from field-based samples did not indicate major losses of genetic diversity in hatchery-planted oysters in Harris Creek, MD, at least based on the single restored population examined. Additional restored populations will need to be analyzed before any definitive conclusions can be made about the genetic impact of this restoration program. The use of a large number of rotated, wild broodstock, from which mixed larval batches are planted and re-planted over time may have helped to minimize diversity loss within restored reef sites. Based on the results of this study, a number of standard recommendations could be made for oyster (or other shellfish) restoration programs, including using large numbers of rotated, wild broodstock, and implementing pair-cross matings (Camara and Vadopalas 2009). However, pair-cross matings and/or strip-spawning may not be feasible for some restoration programs, especially those that employ mass-spawns or are broodstock limited.

Data archiving Genotype data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.h3kv180.

Acknowledgements

We thank the HPL Oyster Hatchery for their help with spawns, broodstock conditioning, and larval/juvenile culturing, Anna Priester for assistance with hatchery work, Scott Bingham and ASU DNA Laboratory for microsatellite sequencing, and Kennedy Paynter for Harris Creek sample collection. We also thank Katherine McFarland for helpful comments on previous versions of the manuscript. Support for this work was provided by the Deerbrook Charitable Trust (DCT 15-30 to LVP), Maryland Sea Grant (5-230360 to KMH and LVP), and the University of Maryland Center for Environmental Science.

Tables Chapter 2

Table 1. Detailed cohort information of *C. virginica* populations used in this study.

Cohort	Source	GPS	N	Females	Males	Offspring sampled
M1	Sandy Hill (SH1)	38.60N, -76.13W	23	16	7	120
M2	Sandy Hill (SH2)	38.59N, -76.10W	21	15	6	120
M3	Sandy Hill (SH3)	38.60N, -76.12W	21	11	10	120
M4	Chlora Point (CLP)	38.63N, -76.14W; 38.63N, -76.15W	22	10	12	115
M5	Black Buoy (BBY)	38.58N, -76.04W	34	11	23	115
M6	Louisiana (LoLA)	NA	20	9	11	115
C1	States Bank (TB)	-76.04, 38.57	15	7	8	120
C2	States Bank (TB)	-76.04, 38.57	15	7	8	120
C3	States Bank (TB)	-76.04, 38.57	15	7	8	120

M1–M6 represent the mass-spawned cohorts; C1–C3 represent the controlled-spawned cohorts.

Table 2. Summary of genetic diversity results for parents and offspring of each *C. virginica* cohort

Cohort	Type	A_r ± s.e.	P-value[1]	H_o ± s.e.	P-value[2]	N_{bI}	N_{bI}/N	N_{bS}
M1	Adult	12.56 ± 2.04	0.007*	0.67 ± 0.06	0.36	9.99	0.43	9.85
	Offspring	8.46 ± 1.16		0.55 ± 0.06				
M2	Adult	11.71 ± 1.64	0.007*	0.58 ± 0.03	1	7.62	0.36	4.37
	Offspring	6.69 ± 0.81		0.55 ± 0.03				
M3	Adult	11.57 ± 1.76	0.007*	0.68 ± 0.06	0.09	17.58	0.84	4.13
	Offspring	8.36 ± 1.22		0.54 ± 0.07				
M4	Adult	12.11 ± 2.26	0.007*	0.65 ± 0.09	0.46	13.63	0.62	9.32
	Offspring	7.26 ± 1.23		0.58 ± 0.10				
M5	Adult	14.38 ± 2.84	0.012*	0.63 ± 0.09	0.09	17.8	0.52	4.57
	Offspring	10.11 ± 1.63		0.52 ± 0.10				
M6	Adult	8.56 ± 1.31	0.007*	0.68 ± 0.08	0.08	11.29	0.56	6.03
	Offspring	6.01 ± 0.75		0.44 ± 0.10				
C1	Adult	11 ± 2.62	0.031*	0.58 ± 0.10	1	8.19	0.55	4.04
	Offspring	7.33 ± 1.57		0.58 ± 0.11				
C2	Adult	9.33 ± 1.94	0.031*	0.51 ± 0.09	1	6.11	0.41	4.05
	Offspring	5.95 ± 1.01		0.51 ± 0.11				
C3	Adult	11.16 ± 2.70	0.031*	0.56 ± 0.10	0.47	9.38	0.63	4.74
	Offspring	7.41 ± 1.64		0.60 ± 0.11				
HC	---	14.19 ± 2.45	---	0.59 ± 0.08	---	---	---	

P-values from Wilcoxon signed-rank tests after Benjamini and Hochberg correction: Ar, allelic richness[1] or H_o, observed heterozygosity[2] between parents and offspring. N_{bI}/N represents the ratio of number of breeders from Christie et al. (2012) to the number of broodstock spawned and

N_{bS} represents N_b based on sibship reconstruction from Waples and Waples (2011). HC represents the restored Harris Creek sample. M1–M6 represent the mass-spawned cohorts; C1–C3 represent the control-spawn cohorts.

Table 3. F_{ST} estimates between all *C. virginica* wild and broodstock source populations (below diagonal) and G-test *P*-value (above diagonal).

	BBY	*CLP*	*HC*	*LoLa*	*SH1*	*SH2*	*SH3*
BBY	_ _	0.556	0.193	**< 0.001**	0.012	0.013	0.18
CLP	0.009	_ _	0.283	**< 0.001**	0.036	0.005	0.061
HC	-0.001	0.004	_ _	**< 0.001**	0.353	**< 0.001**	0.016
LoLa	**0.056**	**0.054**	**0.058**	_ _	**< 0.001**	**< 0.001**	**<0.001**
SH1	0.022	0.017	0.007	**0.079**	_ _	0.05	0.306
SH2	0.021	0.02	0.019	**0.077**	0.011	_ _	**0.001**
SH3	0.016	0.012	0.009	**0.06**	-0.001	0.018	_ _

Bonferroni correction is given in bold ($P < 0.002$). Population codes are explained in Table 1 or Figure 1; HC represents the supplemented Harris Creek sample.

Table 4. Effective population sizes (and confidence intervals (CIs)) of wild, hatchery-produced and restored *C. virginica* populations based on the linkage disequilibrium method (Waples and Do 2008).

	Parents			Offspring		
	N	N_e	CI	N	N_e^a	CI^a
Restored population						
Harris Creek	48	68.3	(41.6, 148.6)	--	--	--
Aquaculture line						
LoLA	20	31.1	(18.6, 69.9)	92	10.5	(8.4, 13)
Wild populations						
Black Buoy	34	129.6	(52.1, ∞)	102	21.9	(18, 26.6)
Chlora Point	22	∞	(86, ∞)	98	4.8	(3.8, 6.6)
Sandy Hill	65	75.3	(54.7, 119.2)	298	22.3	(19.7, 25)
Pooled	166	178.2	(126.2, 281)	781	36.7	(33.3, 40.6)

N_e^a CI^a represent adjusted N_e and CIs of hatchery-produced offspring according to Waples et al. (2014).

Figures Chapter 2

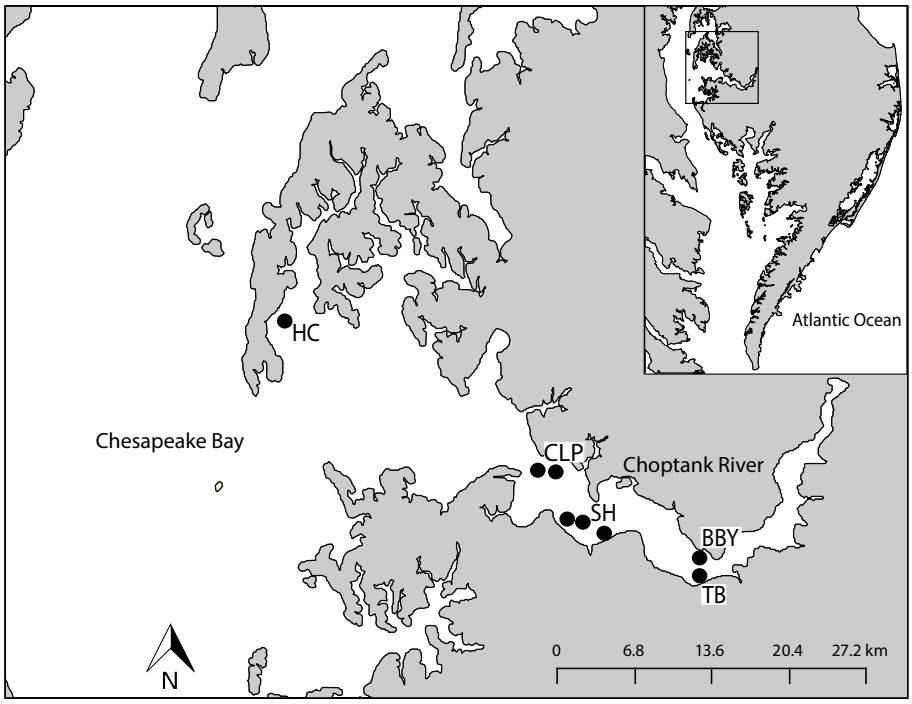


Figure 1. Map of Chesapeake Bay showing the approximate locations of wild *C. virginica* broodstock sources and the Harris Creek Sanctuary. 'CLP' is Chlora Point., 'SH' is Sandy Hill, 'BBY' is Black Buoy, 'TB' is States Bank and 'HC' is the Harris Creek Sanctuary site.

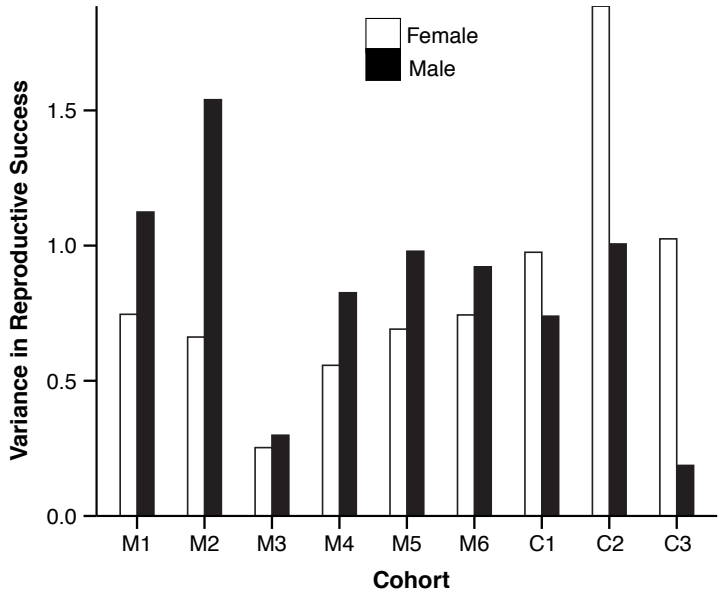


Figure 2. Standardized variance in reproductive success for males and females across *C. virginica* cohorts. M1–M6 correspond to the mass-spawned cohorts; M6 corresponds to the LoLA aquaculture line, and C1–C3 correspond to the controlled-spawn cohorts.

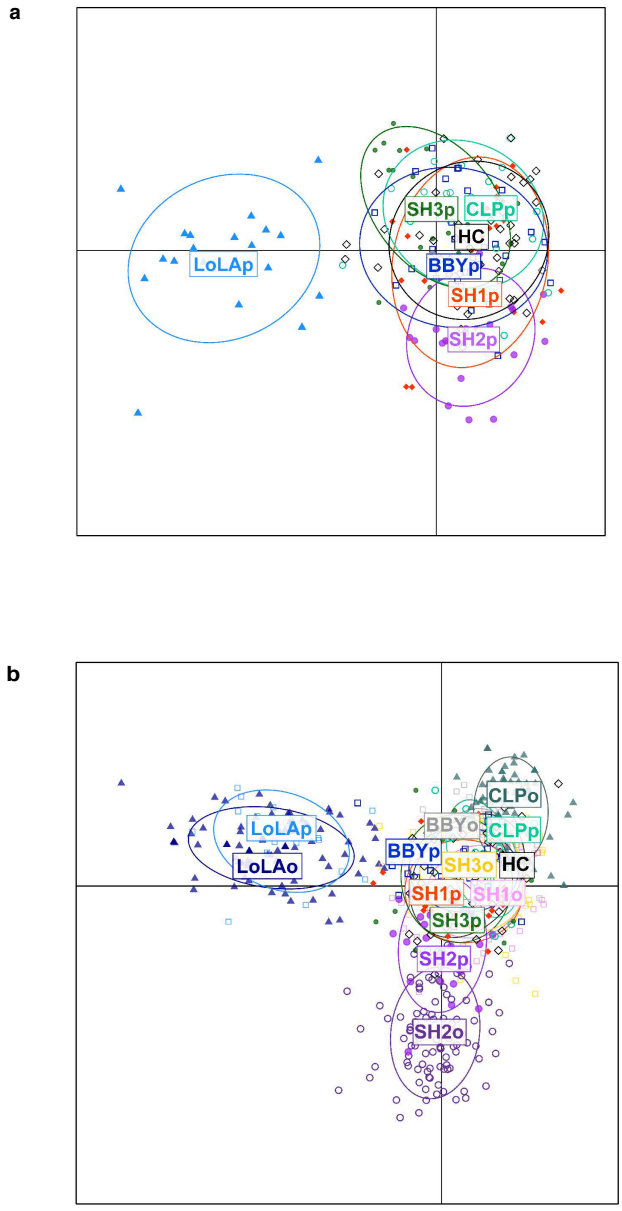


Figure 3. Discriminant analysis of principle components (DAPC) among wild and restored *C. virginica* populations. Results are shown for adult samples (broodstock and restored populations) only (panel a) and for all individuals (adults and offspring; panel b). Number of PCs to retain for each set of analyses was determined via the a-spline optimization approach in Adegenet (29, and 23 PCs retained for all, and parents only, respectively). Letters at the end of population names/codes represent parent (p) or offspring (o) groups.

Supplementary Information

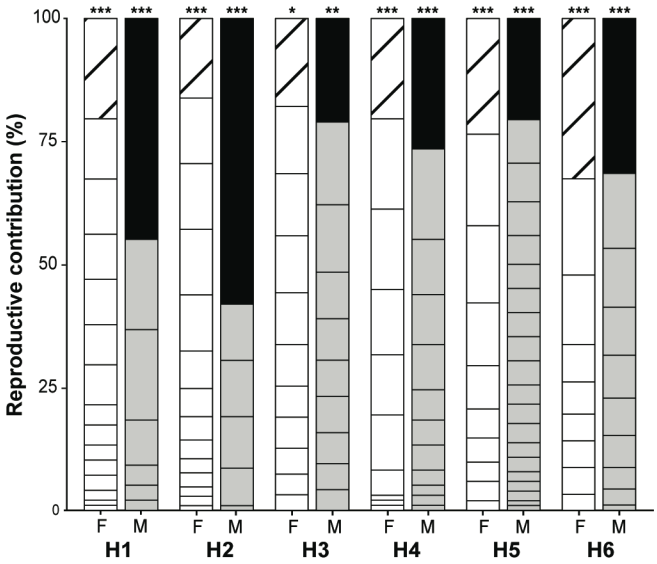


Figure S1: Reproductive contribution (% of total offspring assigned) for males (M) and females (F) in each *C. virginica* cohort. Hatched (female) and black (male) bars represent the contribution of the most dominant female or male in each cohort. Asterisks indicate significance of chi-square goodness-of-fit tests for equal reproductive contributions of males and females (* $P < 0.05$, ** $P < 0.001$, *** $P < 0.00001$)

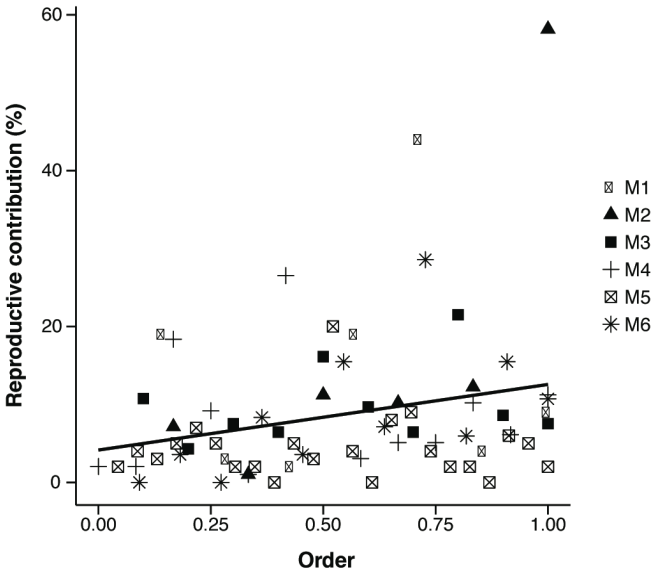


Figure S2. Male reproductive order and reproductive contribution (%) for all mass-spawned *C. virginica* cohorts.

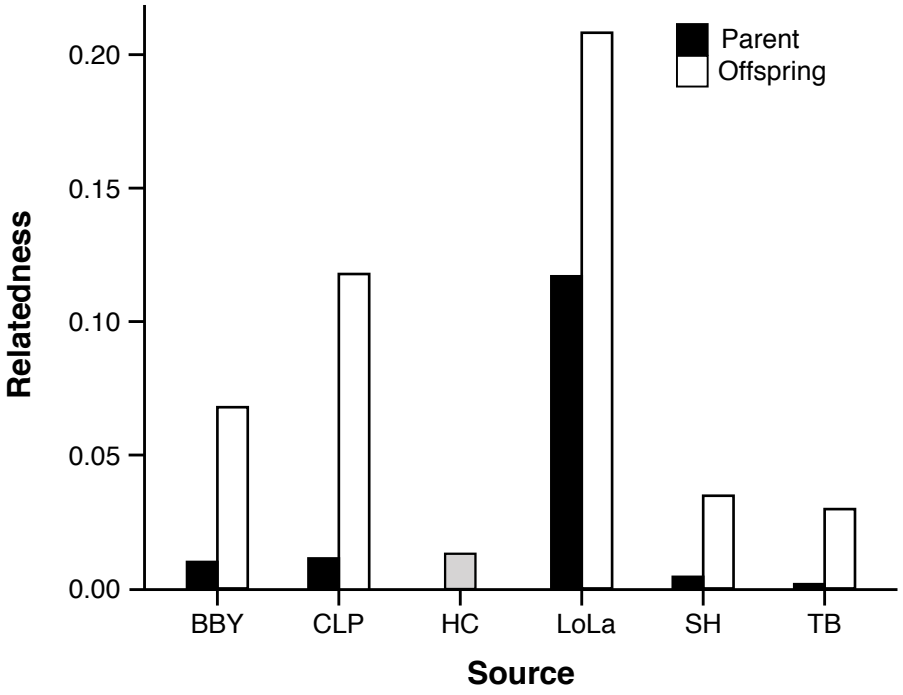


Figure S3. Mean global relatedness in *C. virginica* a supplemented reef (HC), wild broodstock source parents, and hatchery-produced offspring. Relatedness: Lynch and Ritland (Lynch and Ritland, 1999) relatedness estimator calculated by COANCESTRY (Wang, 2011).

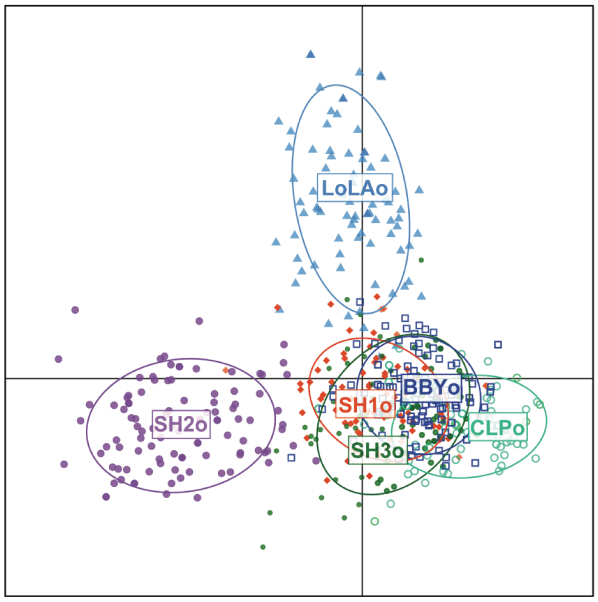


Figure S4. Discriminant analysis of principle component (DAPC) among mass-spawned *C. virginica* hatchery cohorts (offspring). Number of PCs to retain was determined via the a-spline optimization approach in Adegenet (28 PCs retained).

Table S1. Genetic diversity summary statistics for each microsatellite locus and *C. virginica* cohort including the total number of alleles (A), expected (H_e) and observed heterozygosity (H_o), F_{IS}, HWE deviation test (HWE), null allele frequency (Null), and polymorphic information content (PIC).

		Parents									Offspring									
Locus		**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**C1**	**C2**	**C3**	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**C1**	**C2**	**C3**	**HC**
RUCV3																				
	A	7	7	8	6	7	4	--	--	--	7	6	7	5	7	4	--	--	--	8
	H_e	0.69	0.77	0.82	0.78	0.71	0.67	--	--	--	0.65	0.75	0.76	0.80	0.73	0.68	--	--	--	0.77
	H_o	0.48	0.33	0.67	0.46	0.50	0.50	--	--	--	0.50	0.51	0.60	0.40	0.54	0.42	--	--	--	0.6
	F_{IS}	0.31	0.57	0.19	0.42	0.29	0.26	--	--	--	0.23	0.32	0.22	0.50	0.26	0.24	--	--	--	0.22
	HWE	+	**	+	**	+	+	--	--	--	+	***	**	**	**	***	--	--	--	**
	Null	0.17	0.39	0.08	0.25	0.18	0.13	--	--	--	0.13	0.19	0.13	0.33	0.16	0.25	--	--	--	0.12
	PIC	0.65	0.71	0.78	0.72	0.65	0.60	--	--	--	0.60	0.70	0.72	0.76	0.69	0.62	--	--	--	0.72
RUCV45																				
	A	20	17	19	20	24	15	21	16	15	14	13	16	19	18	10	18	13	15	26
	He	0.95	0.92	0.96	0.95	0.95	0.91	0.98	0.94	0.95	0.81	0.86	0.90	0.87	0.88	0.85	0.93	0.87	0.91	0.96
	Ho	1.00	0.76	0.67	0.96	0.93	0.95	0.93	0.80	0.87	0.80	0.83	0.90	0.95	0.79	0.85	0.93	0.86	0.92	0.83
	FIS	-0.06	0.18	0.31	-0.01	0.13	-0.04	0.05	0.15	0.09	0.01	0.04	0.00	-0.09	0.10	-0.03	-0.01	0.01	-0.02	0.13
	HWE	+	**	**	+	+	+	+	+	+	+	+	+	+	**	+	+	+	+	**
	Null	-0.04	0.09	0.17	-0.01	0.06	-0.03	0.01	0.06	0.03	0.00	0.02	-0.01	-0.05	0.05	-0.01	-0.01	0.00	-0.01	0.06
	PIC	0.92	0.89	0.93	0.92	0.93	0.88	0.94	0.90	0.91	0.79	0.84	0.88	0.86	0.87	0.83	0.92	0.85	0.90	0.94
RUCV61																				
	A	18	18	14	18	19	11	13	11	17	15	14	13	12	16	10	11	11	15	19
	He	0.95	0.93	0.92	0.94	0.95	0.86	0.94	0.90	0.94	0.89	0.82	0.87	0.86	0.91	0.84	0.85	0.88	0.88	0.93
	Ho	0.73	0.62	0.71	0.64	0.71	0.65	0.60	0.60	0.67	0.91	0.70	0.73	0.97	0.79	0.51	0.66	0.79	0.75	0.76

	FIS	0.24	0.34	0.23	0.33	0.26	0.25	0.37	0.38	0.30	-0.02	0.15	0.17	-0.13	0.13	0.32	0.22	0.10	0.16	0.19
	HWE	*	**	*	**	**	*	**	**	**	+	***	**	**	**	**	**	**	**	**
	Null	0.12	0.20	0.12	0.19	0.14	0.13	0.20	0.19	0.16	-0.01	0.10	0.09	-0.06	0.07	0.25	0.12	0.05	0.08	0.10
	PIC	0.92	0.90	0.89	0.92	0.93	0.82	0.90	0.86	0.90	0.87	0.81	0.85	0.84	0.89	0.82	0.83	0.86	0.87	0.92
RUCV114																				
	A	5	6	5	3	5	5	7	5	5	5	6	5	3	5	4	5	5	4	5
	He	0.78	0.70	0.73	0.35	0.62	0.50	0.63	0.67	0.66	0.71	0.64	0.64	0.31	0.74	0.38	0.62	0.42	0.55	0.57
	Ho	0.48	0.38	0.57	0.23	0.41	0.55	0.47	0.40	0.27	0.46	0.44	0.45	0.30	0.50	0.27	0.28	0.23	0.46	0.31
	FIS	0.39	0.46	0.22	0.36	0.34	-0.10	0.26	0.41	0.61	0.37	0.31	0.29	0.05	0.34	0.20	0.55	0.46	0.16	0.45
	HWE	**	**	+	+	+	+	+	*	**	**	**	**	*	**	**	**	**	**	**
	Null	0.23	0.27	0.11	0.19	0.20	-0.09	0.12	0.21	0.40	0.23	0.19	0.16	0.01	0.20	0.16	0.38	0.30	0.09	0.29
	PIC	0.72	0.64	0.66	0.32	0.56	0.44	0.58	0.61	0.59	0.66	0.59	0.57	0.29	0.70	0.35	0.57	0.40	0.50	0.52
RUCV148																				
	A	3	4	3	3	3	3	2	4	2	2	3	3	2	3	3	2	4	2	4
	He	0.40	0.57	0.33	0.35	0.19	0.27	0.37	0.41	0.33	0.36	0.56	0.26	0.19	0.11	0.20	0.24	0.22	0.24	0.24
	Ho	0.44	0.48	0.43	0.23	0.21	0.3	0.2	0.27	0.4	0.23	0.49	0.20	0.17	0.07	0.14	0.28	0.21	0.25	0.25
	F_{IS}	-0.10	0.17	-0.08	0.28	-0.07	-0.11	0.47	0.35	-0.22	0.35	0.13	0.23	0.11	0.34	0.37	-0.15	0.02	-0.07	-0.03
	HWE	+	+	+	+	+	+	+	+	+	**	+	**	*	**	**	+	*	+	+
	Null	-0.05	0.11	-0.04	0.19	-0.05	-0.07	0.28	0.22	-0.11	0.21	0.05	0.12	0.06	0.19	0.19	-0.07	0.04	-0.04	0.00
	PIC	0.35	0.49	0.35	0.32	0.11	0.25	0.29	0.37	0.27	0.29	0.49	0.24	0.17	0.10	0.19	0.21	0.20	0.21	0.22
Cvi2i23																				
	A	15	16	15	18	20	8	--	--	--	11	9	12	10	15	7	--	--	--	18
	H_e	0.87	0.92	0.86	0.92	0.92	0.85	--	--	--	0.80	0.75	0.68	0.77	0.86	0.84	--	--	--	0.89
	H_o	0.87	0.85	0.9	0.909	0.912	1	--	--	--	0.71	0.79	0.64	0.69	0.64	0.62	--	--	--	0.87

	FIS	0.00	0.08	-0.05	0.01	0.01	-0.19	--	--	--	0.12	-0.06	0.07	0.11	0.26	0.29	--	--	--	0.02
	HWE	+	+	+	+	+	*	--	--	--	**	+	**	**	**	+	--	--	--	+
	Null	-0.01	0.03	-0.05	-0.01	-0.01	-0.10	--	--	--	0.08	-0.04	0.04	0.07	0.16	0.15	--	--	--	0.01
	PIC	0.84	0.89	0.82	0.89	0.90	0.81	--	--	--	0.78	0.71	0.66	0.75	0.84	0.82	--	--	--	0.87
Cvi2i4																				
	A	15	17	14	14	18	9	--	--	--	11	11	13	10	13	9	--	--	--	18
	He	0.93	0.93	0.93	0.92	0.92	0.85	--	--	--	0.89	0.72	0.88	0.85	0.90	0.81	--	--	--	0.93
	Ho	0.77	0.84	0.95	0.86	0.79	0.55	--	--	--	0.43	0.42	0.63	0.64	0.64	0.30	--	--	--	0.70
	FIS	0.18	0.10	-0.02	0.06	0.14	0.36	--	--	--	0.52	0.42	0.28	0.25	0.29	0.60	--	--	--	0.25
	HWE	**	*	+	+	*	**	--	--	--	**	***	**	**	**	***	--	--	--	**
	Null	0.08	0.04	-0.02	0.02	0.07	0.20	--	--	--	0.34	0.27	0.16	0.14	0.17	0.46	--	--	--	0.13
	PIC	0.91	0.90	0.90	0.89	0.90	0.81	--	--	--	0.88	0.68	0.86	0.83	0.89	0.78	--	--	--	0.91
Cvi4767E-VIMS																				
	A	18	16	16	17	19	11	13	13	18	11	13	15	12	12	10	12	6	14	18
	He	0.95	0.92	0.94	0.94	0.94	0.87	0.94	0.95	0.96	0.91	0.80	0.90	0.89	0.88	0.79	0.87	0.74	0.89	0.93
	Ho	0.78	0.57	0.67	0.68	0.71	0.90	0.73	0.73	0.67	0.91	0.79	0.83	0.87	0.73	0.72	0.73	0.59	0.81	0.84
	GIS	0.18	0.38	0.30	0.28	0.25	-0.03	0.23	0.25	0.31	-0.01	0.01	0.08	0.01	0.18	0.02	0.159	0.21	0.10	0.10
	HWE	*	**	**	**	**	+	*	**	**	+	+	**	+	**	+	**	**	**	*
	Null	0.09	0.23	0.16	0.15	0.13	-0.03	0.10	0.11	0.17	-0.01	0.00	0.04	0.00	0.10	0.04	0.09	0.11	0.05	0.05
	PIC	0.92	0.89	0.92	0.92	0.92	0.84	0.90	0.91	0.92	0.88	0.78	0.89	0.87	0.87	0.76	0.85	0.69	0.88	0.91
Cvi6																				
	A	13	12	11	10	--	11	10	7	10	15	11	8	6	--	8	9	6	8	13
	He	0.89	0.90	0.84	0.85	--	0.87	0.84	0.78	0.80	0.88	0.72	0.80	0.67	--	0.77	0.81	0.66	0.74	0.84
	Ho	0.46	0.59	0.65	0.82	--	0.7	0.53	0.40	0.47	0.60	0.42	0.63	0.69	--	0.52	0.75	0.54	0.53	0.52

FIS	0.50	0.35	0.23	0.04	--	0.20	0.37	0.54	0.42	0.32	0.18	0.22	-0.03	--	0.08	0.08	0.19	0.29	0.38
HWE	**	**	*	+	--	*	**	**	**	+	+	**	+	--	***	*	**	**	**
Null	0.32	0.20	0.12	0.01	--	0.10	0.22	0.31	0.26	0.19	0.27	0.12	-0.03	--	0.19	0.04	0.08	0.16	0.23
PIC	0.86	0.86	0.80	0.82	--	0.83	0.79	0.73	0.75	0.89	0.68	0.76	0.63	--	0.73	0.78	0.61	0.70	0.81

Population codes are explained in Table 1 and Figure 1
Levels of significance of HWE deviation test are indicated using astericks (*** $P \leq 0.001$; ** $P \leq 0.01$; * $P \leq 0.05$ and + conformance to HWE

Table S2. Parentage assignment results for nine hatchery *C. virginica* cohorts using CERVUS 3.0 software.

Cohort	N_{total}	Confidence level of assignment	Critical Delta	Observed assignments	Expected assignments
M1	***120***	95%	***0.37***	(98) 82%	(116) 97%
		80%	***0***	(100) 83%	120 (100%)
		Unassigned		(20) 17%	(0) 0%
M2	***119***	95%	***0.24***	(98) 88%	(116) 98%
		80%	***0***	(107) 90%	(118) 99%
		Unassigned		(12) 10%	(1) 0%
M3	***114***	95%	***0.56***	(93) 82%	(109) 95%
		80%	***0***	(95) 83%	(113) 100%
		Unassigned		(19) 17%	(1) 0%
M4	***115***	95%	***0.21***	(98) 85%	(113) 98%
		80%	***0***	(98) 85%	(115) 100%
		Unassigned		(17) 15%	(0) 0%
M5	***115***	95%	***0.67***	(102) 89%	109 (94%)
		80%	***0***	(108) 94%	(115) 100%
		Unassigned		(7) 6%	(0) 0%
M6	***114***	95%	***1.24***	(92) 81%	(95) 84%
		80%	***0***	(99) 77%	(113) 99%
		Unassigned		(15) 13%	(1) 0%
C1	***114***	95%	***1.09***	(92) 81%	(101) 88%
		80%	***0***	(105) 92%	(113) 99%
		Unassigned		(9) 8%	(1) 1%
C2	***117***	95%	***1.25***	(87) 74%	96 (82%)
		80%	***0***	(90) 77%	(116) 99%
		Unassigned		(27) 23%	(1) 1%
C3	***118***	95%	***0.67***	(103) 87%	(110) 93%
		80%	***0***	(105) 89%	(117) 99%
		Unassigned		(13) 11%	(1) 1%

N_{total} is the number of juveniles included in the analysis. The critical Delta scores and expected number of parentage assignments were determined by simulation of parentage analysis (see methods).

Table S3. Mean global relatedness in *C. virginica* all mass-spawned and controlled-spawned cohorts (parents and offspring) and a supplemented reef (HC). Relatedness: Lynch and Ritland (Lynch and Ritland 1999) relatedness estimator calculated by COANCESTRY (Wang 2011).

Cohort	M1	M2	M3	M4	M5	M6	C1	C2	C3	HC
Parent	0.004	0.014	0.011	0.011	0.010	0.117	-0.017	-0.005	-0.003	0.013
Offspring	0.098	0.193	0.077	0.129	0.068	0.208	0.102	0.167	0.135	

Table S4. Matrix of pairwise F_{ST} (Weir, 1996) between all *C. virginica* populations calculated with FreeNA using the excluding null alleles (ENA) correction (Chapuis and Estoup 2007).

	BBY	*CLP*	*HC*	*LoLa*	*SH1*	*SH2*	*SH3*
BBY	_ _						
CLP	0.009	_ _					
HC	0.004	0.002	_ _				
LoLA	0.051	0.051	0.055	_ _			
SH1	0.017	0.02	0.005	0.071	_ _		
SH2	0.023	0.019	0.019	0.069	0.009	_ _	
SH3	0.011	0.012	0.008	0.056	-0.000	0.015	_ _

Population codes are explained in Table 1 or Figure 1

Table S5. F_{ST} estimates between all *C. virginica* populations including offspring.

	BBYp	*BBYo*	*CLPp*	*CLPo*	*HC*	*LoLAp*	*LoLAo*	*SH1p*	*SH1o*	*SH2p*	*SH2o*	*SH3p*	*SH3o*
BBYp	_ _												
BBYo	0.012	_ _											
CLPp	0.009	**0.028**	_ _										
CLPo	0.004	**0.054**	**0.011**	_ _									
HC	0.004	**0.025**	0.001	**0.043**	_ _								
LoLAp	**0.054**	**0.084**	**0.056**	**0.092**	**0.058**	_ _							
LoLAo	**0.095**	**0.135**	**0.093**	**0.13**	**0.098**	0.011	_ _						
SH1p	0.017	**0.025**	0.022	**0.073**	0.007	**0.079**	**0.129**	_ _					
SH1o	**0.049**	**0.058**	**0.024**	**0.09**	**0.036**	**0.199**	**0.167**	0.005	_ _				
SH2p	0.02	**0.044**	0.021	**0.073**	0.019	**0.077**	**0.12**	0.011	**0.038**	_ _			
SH2o	**0.113**	**0.148**	**0.105**	**0.168**	**0.108**	**0.151**	**0.197**	**0.098**	**0.13**	**0.04**	_ _		
SH3p	0.012	**0.024**	0.016	**0.059**	0.009	**0.06**	**0.113**	0.001	**0.037**	0.018	**0.1**	_ _	
SH3o	**0.034**	**0.045**	**0.035**	**0.077**	**0.027**	**0.092**	**0.142**	**0.023**	**0.053**	**0.05**	**0.134**	0.001	_ _

Significance of each pairwise F_{ST} value calculated by 9,999 permutations. Significance after Bonferroni correction is given in bold ($P < 0.0006$). Population codes are explained in Table 1 or Figure 1. Letters at the end of population names/codes represent parent (p) or offspring (o) groups.

Table S6. Female (F1-F7) and male (M1-M8) spawning times and fertilizations times for *C. virginica* controlled-spawn cohorts (C1-C3)

Cohort	**C1**				**C2**				**C3**			
	C1F1	9:54	*C1M1*	9:50	*C2F1*	11:18	*C2M1*	11:05	*C3F1*	12:25	*C3M1*	12:19
	C1F2	11:13	*C1M2*	9:54	*C2F2*	11:18	*C2M2*	11:05	*C3F2*	13:03	*C3M2*	12:37
	C1F3	11:13	*C1M3*	9:55	*C2F3*	11:20	*C2M3*	11:10	*C3F3*	13:04	*C3M3*	12:37
	C1F4	11:14	*C1M4*	9:58	*C2F4*	11:21	*C2M4*	11:11	*C3F4*	13:11	*C3M4*	12:46
	C1F5	11:14	*C1M5*	10:04	*C2F5*	11:37	*C2M5*	11:11	*C3F5*	13:16	*C3M5*	12:55
	C1F6	11:14	*C1M6*	10:28	*C2F6*	11:41	*C2M6*	11:11	*C3F6*	13:18	*C3M6*	13:08
	C1F7	11:18	*C1M7*	11:03	*C2F7*	12:26	*C2M7*	11:11	*C3F7*	13:18	*C3M7*	13:11
			C1M8	11:09			*C2M8*	11:12			*C3M8*	13:18
Fertilization	**12:46**				**13:34**				**14:31**			
Max time	**2:52**		**2:56**		**2:16**		**2:29**		**2:06**		**2:12**	

Max time (hh:mm) represents length of time sperm and oocytes were held from first male and female in each cohort prior to fertilization.

Chapter 3: Genome-wide analysis of restored and natural eastern oyster populations reveal local adaptation and positive impacts of planting effort and broodstock number

Abstract

The release of captive-bred plants and animals has increased worldwide to augment declining species. However, insufficient attention has been given to understanding how neutral and adaptive genetic variation are partitioned within and among proximal natural populations, and the patterns and drivers of gene flow over small spatial scales, which can be important for restoration success. A seascape genomics approach was used to investigate population structure, local adaptation, and the extent to which environmental gradients influence genetic variation among wild and restored populations of Chesapeake Bay eastern oysters *Crassostrea virginica*. I also investigated the impact of hatchery practices on neutral genetic diversity of restored reefs and quantified the broader genetic impacts of large-scale hatchery-based shellfish restoration. Restored reefs showed similar levels of diversity as wild reefs, and striking relationships were found between planting frequency and broodstock numbers and genetic diversity metrics (effective population size and relatedness) suggesting that hatchery practices can have a major impact on diversity. Despite long-term restoration activities, haphazard historical translocations, and high dispersal potential of larvae that could homogenize allele frequencies among populations, moderate neutral population genetic structure was uncovered. Moreover, environmental factors, namely salinity, temperature, and dissolved

oxygen, play a major role in the distribution of neutral and adaptive genetic variation. For marine invertebrates in heterogeneous seascapes, collecting broodstock from large populations experiencing similar environments to candidate sites may provide the most appropriate sources for restoration and ensure population resilience in the face of rapid environmental change. This is one of a few studies to demonstrate empirically that hatchery practices have a major impact on the retention of genetic diversity. Overall, these results contribute to the growing body of evidence for fine-scale genetic structure and local adaptation in broadcast spawning marine species and provide novel information for the management of an important fisheries resource.

Introduction

Anthropogenic impacts to aquatic environments including habitat loss, species introductions, overharvesting, and climate change have severely degraded ecosystems and reduced populations of species worldwide, with coastal marine environments among the most severely affected (Lotze et al. 2006). To counteract these impacts, reestablish ecosystem function, and build resiliency, restoration activities, including population supplementation with translocated stock from wild populations or captive-reared offspring, have become important fisheries management strategies (see Bell et al. 2008 for definitions and objectives, Lorenzen et al. 2012). While these activities have increased population abundances (e.g. Berejikian and Doornik 2018), they may also have profound genetic impacts that can reduce long-term population resilience (reviewed in Frankham et al. 2010). Therefore, understanding patterns of neutral and adaptive genetic variation is critical to establishing restoration programs that aim to preserve genetic

diversity, maintain historic gene flow and local adaptation, and promote resilience in the face of rapid environmental change (Laikre et al. 2010; Flanagan et al. 2018). While recent advances in genomics allow more precise quantification of neutral variation and the identification of adaptive loci affected by the environment (Baird et al. 2008; Allendorf et al. 2010), more work is needed to link these approaches with practical aspects of species restoration (e.g. Breed et al. 2018).

A key issue for many restoration programs is the degree to which genetic diversity is maintained in hatchery-produced individuals compared to wild populations. Reductions in genetic diversity and effective population size (N_e, the evolutionary analog to census population size) have been documented in wild populations when large numbers of hatchery-produced individuals from a small number of broodstock are released (Ryman and Laikre 1991; Ryman et al. 1995; Christie et al. 2012a). In the short-term, reductions in genetic diversity impact population viability due to the negative impacts associated with inbreeding depression. In the long-term, reductions in genetic diversity impact population viability because populations with insufficient genetic diversity harbor less adaptive potential and are thus expected to be more vulnerable to environmental variability (Lande 1995; Willi et al. 2006). While "genetically aware" restoration programs exist (i.e. broodstock are selected from local populations and carefully planned breeding protocols are utilized), the severity of associated genetic changes remains variable (e.g. Heggenes et al. 2006; Gow et al. 2011; Christie et al. 2012). Relatively little work has been done to connect captive breeding and restoration practices to the genetic diversity of restored populations.

Another central and often controversial issue for population restoration is the choice of appropriate broodstock material (Broadhurst et al. 2008).Transplanting foreign genotypes with lower fitness than local genotypes can have important implications for restoration success and the long-term viability of restored populations (Helenurm 1998; Galloway and Fenster 2000; Montalvo and Ellstrand 2001; Hufford and Mazer 2003). Restoration guidelines advocate the use of local, wild broodstock (e.g. Brumbaugh et al. 2006), but these guidelines often assume high connectivity and minimal population structure among marine species with planktonic dispersal. However, recent studies of marine species indicate both limited effective dispersal and local adaptation over small-scales may be more common than previously hypothesized (Hauser and Carvalho 2008; Sanford and Kelly 2011; Silliman 2019; Bernatchez et al. 2019). Therefore, the choice of appropriate genetic material for increased success in population restoration programs requires an understanding of population structure and patterns of adaptation across a broad range of environments scales.

Restoration of native oysters has increased due to the worldwide decline of ecologically, economically, and culturally significant species (Kirby 2004; Beck et al. 2011). For bivalves, this practice often includes seeding hatchery-propagated juveniles into wild populations (Gaffney 2006; Laing et al. 2006). While the genetic impacts associated with shellfish restoration have been documented (Boudry et al. 2002; Camara and Vadopalas 2009; Lind et al. 2009), few studies have examined genetic changes associated with shellfish restoration programs (Lallias et al. 2010; Morvezen et al. 2016; Arnaldi et al. 2018; Hughes et al. 2019; Hornick and Plough 2019; Jaris et al. 2019). Only recently have patterns of neutral and adaptive genetic variation in wild populations of

shellfish been uncovered using high-resolution genomic methods (Silliman 2019; Lehnert et al. 2019; Bernatchez et al. 2019; Vendrami et al. 2019; Miller et al. 2019). Marine shellfish exhibit complex life cycles which include both planktonic larval stages and benthic juvenile and adult stages. Life-history features such as high-fecundity, type III survivorship, and high variance in reproductive success (Hedgecock and Pudovkin 2011; Plough et al. 2016) can reduce N_e and genetic diversity in hatchery-produced stocks and exacerbate the negative genetic impacts associated with restoration. While genetic information is frequently integrated in terrestrial ecosystem restoration planning (Rice and Emery 2003; McKay et al. 2005; Leimu and Fischer 2008), it is considered but rarely integrated into marine restoration planning (Baums 2008; for exceptions see Hämmerli and Reusch 2002; Camara and Vadopalas 2009; Fraser et al. 2011).

Oysters are the most economically important group of bivalve shellfish (FAO 2018), and provide critical ecosystem services (Coen et al. 2007; Grabowski et al. 2012), but 85% of oyster reefs have been lost worldwide (Beck et al. 2011). Therefore, interest in restoring oyster populations has grown, with the greatest amount of restoration efforts focused on the eastern oyster (*Crassostrea virgnica* Gmelin 1791). The eastern oyster historically has been the most important native shellfish species in North America ranging from the Gulf Coast to the maritime provinces in Canada (Wilberg et al. 2011), but contemporary oyster populations have declined to ~1% of historic abundances (Mackenzie, 2007; Wilberg et al., 2011). In the Chesapeake Bay, a variety of management and restoration efforts have been undertaken, including seed translocations within and between Bay tributaries, the construction of reef habitat using fresh and dredged shell, designation of oyster sanctuaries or reserves, and supplementing reefs with

hatchery-produced juveniles or large adults (Kennedy and Breisch 1983). Recently, a federal mandate to restore 20 Chesapeake Bay tributaries by 2025 provides support for large-scale restoration in the Choptank River (Maryland, USA), with the first sanctuary, Harris Creek, completed in 2016 (Westby et al. 2017). The University of Maryland Center for Environmental Science's (UMCES) Horn Point Laboratory (HPL) Oyster Hatchery produces spat (juvenile oysters) for Harris Creek (and other tributaries), through mass-spawning of local, wild broodstock. Hornick and Plough (2019) conducted initial characterization of the neutral genetic impacts of this program. However, the analysis of additional wild and restored populations using high-resolution genome-wide markers is necessary to infer patterns of neutral and adaptive genetic variation of Chesapeake Bay oyster populations. This information will permit a more complete understanding of the genetic impacts of large-scale hatchery-based oyster restoration. Furthermore, verifying whether restored reefs have maintained similar levels of genetic variation to proximal wild oyster populations is important because it provides insight into their future resilience. In addition, accounting for local adaptation in the restoration of oysters may increase long-term effectiveness of restoration programs. Great potential for local adaptation exists for oyster populations in the highly variable, human-impacted coastal and estuarine environments that they reside, so understanding the extent of adaptive variation in these populations is critical (Funk et al. 2012).

In this study, next-generation sequencing and a more expansive sampling of restored and wild reefs than previous studies were used to examine the genetic impact of a large-scale hatchery-based restoration program for eastern oysters in the Chesapeake Bay. While previous studies have investigated genetic diversity of hatchery-produced

eastern oysters (Hornick and Plough 2019; Hughes et al. 2019; Jaris et al. 2019), genetic structure (Rose et al. 2006) and local adaptation of wild oyster populations (Bernatchez et al. 2019), none have used genome-wide marker data to characterize impacts of restoration on adaptive and neutral genetic variation. By characterizing patterns of genetic variation within and among restored and wild eastern oyster populations, I quantified the broader population genetic impacts of large-scale hatchery-based shellfish restoration, investigated population structure, local adaptation, and the extent at which environmental gradients influence genetic variation among these populations. This is the first study to include fine-scale sampling of restored shellfish populations with variable hatchery-planting efforts as well as utilizing thousands of high-resolution single nucleotide polymorphisms (SNPs) to characterize neutral and adaptive genetic variation and structure of restored and wild oyster populations in the Chesapeake Bay. Understanding the extent of genetic variability in wild and restored oyster populations and how the variation is structured across broad environmental gradients will be important information for planning future oyster restoration programs and their management.

Materials and Methods

Sample collection

Oysters were collected between 2015 and 2018 from nine sites throughout the Chesapeake Bay (Table 1 and Figure 1). For the Harris Creek sites, divers sampled putative wild oysters (based on sampling location and reef characteristics), recently recruited juveniles (spat), and adult oysters from sites with variable hatchery planting

efforts (Table 1 and Figure 1). Samples collected from restored reefs in Harris Creek included sites planted with hatchery oysters during one season, two seasons, and four seasons (a season occurs during the summer/fall and may involve more than one hatchery planting event) to assess genetic changes associated with planting frequency. For the wild Maryland populations, oysters were obtained from the Choptank River hatchery broodstock source population States Bank (Figure 1). Wild Virginia populations included oysters from sites with no previous hatchery-produced restoration plantings at the scale of the program in Harris Creek (tens of millions of seed planted each year). All samples represent mixed-age cohorts (see Table 1 for average length of oysters from each site), except the recently recruited spat sample from Harris Creek (HCS). Tissues were sampled from adductor muscle or mantle and preserved in 70-95% ethanol until DNA extraction (N = 556 individuals).

Library preparation and sequencing

Genomic DNA was extracted using the E.Z.N.A.® Tissue DNA Kit (Omega-Biotek, Norcross, GA), following the protocol for animal tissues with RNAse-A treatment following manufacturer instructions. A modified double digest restriction-site associated DNA (ddRAD) protocol (Peterson et al. 2012) was used to simultaneously discover and genotype individuals at thousands of single nucleotide polymorphisms (SNPs). Briefly, genomic DNA was digested using the enzymes *EcoRI* and *SphI* (New England Biolabs). Barcoded adapters were ligated onto the digested fragments and fragments were pooled and size selected. Flow cell adapters with one of five index-identifiers were ligated to each set of pooled individuals (48) using 12 PCR cycles prior

to sequencing. Paired-end 250-bp sequencing was conducted on the Illumina HiSeq 2500 platform at Genewiz, Inc. (South Plainfield, NJ) spread across two and a half lanes. Two sampling sites were spread across libraries to minimize batch effects from library preparation and sequencing of 14 replicates were included to estimate sequencing and genotyping errors.

Bioinformatics and genotyping

Read quality was evaluated using FastQC v.0.11.5 (Andrews 2010). Raw sequences were demultiplexed using the *process_radtags* component of Stacks v.2.0 (Catchen et al. 2013). Read mapping and SNP calling were performed using the dDocent pipeline v. 2.7.7 (Puritz et al. 2014) with default settings unless otherwise noted. Trimmed reads were directly mapped to the latest release of the *C. virginica* genome (NCBI Bioprojects: PRJNA379157 and PRJNA376014, accession numbers: NC_035780.1 – NC_035789.1) using the MEM algorithm of Burrows-Wheeler Aligner (BWA; Li and Durbin, 2009) with parameters A (match score), B (mismatch score), and O (gap penalty) set to 1,2, and 5, respectively, which are appropriate for genomic data of marine species (Puritz, unpublished data). Freebayes v1.2.0-dirty (Garrison and Marth 2012) was used to obtain raw variant calls and SNP genotypes. Complete details about methods used to create the final SNP dataset including data filtering can be found in the Table 2.

Outlier detection and defining datasets

Three outlier detection methods with different underlying models were used to partition SNPs into groups of putatively neutral versus SNPs putatively under directional selection: Bayescan (v.2.1) (Foll and Gaggiotti 2008) with prior model of 10,000 following recommendations of Lotterhos and Whitlock (2015), 10,000 iterations, a burn-in of 200,000 steps and a false discovery rate (FDR) of 0.05; OutFLANK (v.0.2) (Whitlock and Lotterhos 2015) with default options (LeftTrimFraction=0.05, RightTrimFraction=0.05, Hmin=0.1, 11) and a q-threshold of 0.05, and pcadapt (v.4.1.0) with a q-value threshold of 0.05 (Luu et al. 2017). For Bayescan and OutFLANK analyses, individuals were grouped into populations by sampling site. For pcadapt, multiple values of K (principal components) were tested, and the final number of PC axes retained was determined by visual inspection of the scree plot. Multiple methods were utilized to minimize the occurrence of false positives as recommended by Hoban et al. (2016).

The dataset was then split into "neutral" and "outlier" components with the final outlier dataset consisting of all SNPs identified as outliers under directional selection by at least one of the approaches, and all outliers detected in the redundancy analysis (RDA; details below), with the neutral dataset consisting of all remaining SNPs. SNPs that were detected as outliers using at least two detection methods were classified as SNPs putatively under divergent selection. All outlier analyses were repeated using the same criteria excluding the Virginia coastal Bay Wachapreague sample to test for selection within Chesapeake Bay samples.

Inclusion of loci that are strongly linked (high linkage disequilibrium) can lead to biases in downstream analyses if independence of loci is assumed (Willis et al. 2017). For inferences of genetic diversity and population structure, a dataset was created that excluded SNPs in close proximity in the genome. Thinning of the neutral dataset to retain one SNP per 2137 bp was performed in VCFTOOLS using the thin function (Danecek et al. 2011). The appropriate thinning distance was determined by calculating R^2 separately for SNPs on the same chromosome (intrachromosomal pairs) and for unlinked SNPs (interchromosomal pairs). The critical R^2 was estimated from the unlinked loci by root transforming the R^2 values and taking the 95th percentile of the distribution as the threshold beyond which the LD is caused by physical linkage (Breseghello and Sorrells 2006). The relationship between LD decay and genetic distance was summarized by fitting a second-degree smoothed locally-weighted linear regression (LOESS) curve (Cleveland 1979) to intrachromosomal R^2 data in R. The distance the loess curve intercepted the critical R^2 was identified as the threshold for LD decay (Figure 2).

Genetic diversity and effective size of wild and restored oysters

The thinned SNP dataset (2,810 SNPs) was used to calculate observed (H_o) and expected (H_e) heterozygosity, and the F_{IS} inbreeding coefficient in hierfstat v0.04-22 (Goudet 2005; R Core Team 2019). Confidence intervals for population-specific F_{IS} were determined using the *boot.ppfis* function in hierfstat with 1,000 bootstrap replicates. Relatedness was estimated for wild and restored oysters using the R package related v.0.8 (Pew et al. 2015). The Ritland estimator (Ritland 1996) was used because it has the least bias with small sample sizes (Wang 2017)

Contemporary genetic effective population size (N_e) was estimated using the single-sample linkage disequilibrium method (Hill 1981; Waples 2006; Waples and Do 2010) as implemented in NeEstimator v2.1 (Do et al. 2014) under a random mating model. The Harris Creek spat sample (which represents a single cohort), provides information most relevant to estimating N_b (Waples 2005), while the mixed-age cohorts (all remaining sites) provide information relevant to estimating N_e (Waples et al. 2014). The neutral dataset was used for N_e estimation (i.e. excluding loci putatively under selection) as suggested by Waples (2006) and P_{Crit} was set to 0.02 (alleles with frequencies <0.02 are excluded), which balances effects of precision and bias (Waples and Do 2010). Confidence intervals were based on the jackknife method (Jones et al. 2016). While the spat sample provides information relevant to estimating N_b, there is some influence from background N_e per generation (Waples et al. 2014). To overcome bias due to overlapping generations using the LD N_e method (Waples et al. 2014), the raw N_b estimate from the Harris Creek spat sample was adjusted using three life-history traits as in Hornick and Plough (2019) (Waples et al. 2014).

The contribution of reef size, number of broodstock, male-to-female ratio of broodstock, and planting frequency to genetic diversity metrics of restored reefs was investigated (mixed-cohort samples) using generalized linear models. The relationship between N_e and H_o at restored reefs to planting frequency, number of broodstock used each planting season, male-to-female ratio of broodstock used each planting season, and reef size (acres; Table 3) were examined. For this analysis, samples from a restored reef from Harris Creek planted with hatchery-produced oysters in 2012 that was genotyped with nine microsatellite markers was included (Hornick and Plough 2019). To ensure that

heterozygosity of all individuals was measured on the same scale despite differences in marker information, the standardized multilocus heterozygosity, the sum of observed average heterozygosity in a population (Coltman et al. 1999) was calculated using the R package inbreedR v.0.3.2 (Stoffel et al. 2016). Significant correlations between the predictors and genetic diversity metrics of restored reefs were calculated in R.

Genetic differentiation, population structure, and population assignment

All analyses related to neutral population genetic structure were performed using the thinned neutral dataset. The extent of genetic differentiation between the sampling sites was evaluated using pairwise estimates of F_{ST} (Weir and Cockerham 1984) with the *genet.dist* function in hierfstat. Isolation by distance (IBD, Sokal 1979) was evaluated using a Mantel test of pairwise F_{ST} values coded as $F_{ST}/(1–F_{ST})$ as a function of water distance between sampling sites (calculated by drawing routes between all sites on Google Earth) as implemented in adegenet v.2.1.1 (Jombart 2008; Jombart and Ahmed 2011).

Two approaches were used to investigate neutral spatial genetic structure: the multivariate discriminant analysis of principal components (DAPC) and the Bayesian clustering algorithm implemented in STRUCTURE v.2.3.4 (Pritchard et al. 2000). Clustering identification was performed by cross-validated DAPC implemented in the r package adegenet (Jombart 2008; Jombart and Ahmed 2011). Individuals were grouped based on sampling site. Cross-validation was performed over a range of 1-478 PCs with 500 replicates to determine the number of principal components to retain and to avoid overfitting during discrimination. After the number of optimal PCs was identified, a

second cross-validation was performed for a narrower range of principal components (±10 of the previously identified optimum). Membership of individuals to clusters were defined by independent k-means, using the Bayesian Information Criterion (BIC). Next, the Bayesian clustering method STRUCTURE v.2.3.4 (Pritchard et al. 2000) was used to identify the number of distinct genetic clusters (K) with a burn-in of 50,000 iterations followed by an additional 200,000 Markov chain Monte Carlo (MCMC) steps, using prior sampling location information and the no-admixture model, which is preferred when levels of divergence between populations are low (Hubisz et al. 2009). Fifteen replicates of K from 1 to 11 were performed, where K is the number of population clusters. Replicates were summarized and visualized using the CLUMPAK server (Kopelman et al. 2015). The K method in STRUCTURE HARVESTER was used to determine the optimal K (Earl and vonHoldt 2012).

The ability to assign individuals to their sample of origin was tested with subsets of SNPs using R package Assigner v.0.5.6 (Gosselin et al. 2016). Assigner uses a training data set to identify highly discriminatory loci (based on F_{ST}), followed by a leave-one-out method on an independent test data set to test the assignment (Anderson 2010). Half of the individuals from each sampling site were used as a training data set, and the remaining individuals were used to assign individuals to populations. Data sets with 50, 100, 200, 500, 1000, 2000, 5000, with the highest F_{ST} were used, as well as all SNPs (7710), with ten iterations per data set.

Associations between environmental variables and genetic variation

A redundancy analysis (RDA) was performed as a genotype-environment association method to detect SNPs putatively under selection based on correlations with environmental variables (Forester et al. 2018) using the r package vegan v.2.5-5 (Oksanen et al. 2019). When compared to differentiation-based outlier methods, RDA can detect even weak multi-locus signatures of selection for multiple environmental variables (Rellstab et al. 2015; Forester et al. 2018). Genotypes for all SNPs and sampling sites were used with environmental data for each locality obtained from the Maryland Department of Natural Resources Eyes on the Bay program (http://eyesonthebay.dnr.maryland.gov/) and the Chesapeake Bay Program (http://data.chesapeakebay.net/) from buoys located closest to each of the eleven sampling sites (Table 2). Environmental variables considered to be important for oysters were downloaded (three related to salinity and temperature, and two related to dissolved oxygen (DO) and pH from 2014-2018). The environmental parameters used were annual minimum, mean, and maximum salinity and water temperature, and annual mean and minimum DO and pH, at a given site. Some buoys had continuous monitoring (data every 15 minutes; Table 2) while others sampled only two or four times per month, so the minimum resolution available for all sites, sampling twice per month over four years, was used (N=96 for each variable). Missing values for environmental data were replaced with the median using the R package RANDOMFOREST v.4.6-14 (Liaw and Wiener 2001) so that (96 data points for each environmental parameter).

For pre-analysis data filtering, correlations between environmental variables were evaluated using the Pearson correlation coefficient and when two variables were highly

correlated ($|r| \geq 0.7$), only one variable was retained. The resulting set of variables included two variables related to temperature (annual maximum and minimum water temperature), one related to salinity (annual mean salinity), two related to pH (annual mean and minimum pH), and one related to DO (annual minimum DO). RDA requires complete data frames, so missing genotype data was imputed by using the most common genotype across individuals (Forester et al. 2018). Significance (alpha ≤ 0.05) of the global RDA and significance of each RDA axis were assessed using an ANOVA with 999 permutations. Candidate outlier SNPs were identified using the distribution of their loadings on each significant RDA axis, so SNPs with loadings located at tail of the distribution were more likely to be under selection. The tail cut-off of ± 3 standard deviations (SDs) from the mean loading of each axis was used to identify candidate SNPs and to minimize false positives and false negatives (Forester et al. 2018). The covariate most strongly correlated (highest correlation coefficient) between each candidate SNP and environmental variable was identified to group candidates by potential driving environmental parameters. In order for a SNP to be considered a significant outlier by RDA, we followed the methodology implemented in pcadapt (Luu et al. 2017) and using the script from Capblancq et al. (2018). Briefly, SNP loadings were recovered from the RDA, and only loadings of the most informative ordination axis were retained by determining the amount of information retained on the different axes of RDA. A Mahalanobis distance was calculated using the covRob function in the r package robust v.0.4-18.2 (Wang et al. 2014). The false discovery rate (FDR) was then adjusted by computing q-values and considered a SNP to be a significant outlier if q-value was <0.05.

RDA was conducted to disentangle the relative contribution of environmental and spatial components driving neutral and adaptive genetic variation (Borcard et al. 1992; Liu 1997; Legendre and Fortin 2010; Bie et al. 2012). RDA is a useful multivariate regression technique when running regression analyses with multivariate predictors (space and environment) and multivariate responses (here, allele frequencies of SNPs). For this analysis, the thinned neutral dataset (2,810 SNPs) and the putatively adaptive dataset (see Results) was used with environmental variables and spatial variables (X and Y coordinates). Spatial variables based on the x-y coordinates were defined using the principal coordinates of neighborhood matrices (PCNMs), also known as Moran's eigenvector maps (MEM) using the *pcnm* function in vegan. Half of the PCNM variables with positive eigenvalues were retained, which has been suggested in similar contexts (Manel et al. 2012; Fitzpatrick and Keller 2015).

Genetic data and environmental variables were compared in a partial RDA specifying geography (PCNMs) as a third conditioned matrix so that the analysis conditions on geographic location. To assess the correlation between genotype and each spatial/environmental variable, an analysis of variance (ANOVA) was performed with 1,000 permutations one variable at a time and variables with $p \leq 0.1$ were retained. Then, the variance inflation factor (VIF; vif.cca function implemented in vegan) was calculated to evaluate multicollinearity of all retained variables (Hair et al. 1995; Zuur et al. 2010; James et al. 2013) variables with VIF>=10 were excluded (Hair et al. 1995). The ordistep function from the R package vegan was used to select the most important explanatory variables among those retained. The final RDA was assessed using an ANOVA and marginal ANOVAs (1,000 permutations) to assess the contribution of each environmental

variable. Next, to explain how much of the genetic variation in *C. virginica* is uniquely explained by environmental variables, how much is uniquely explained by geography, and how much is due to the combined effect of the two, variance components of the RDA were partitioned by running 3 models: a full model with environmental and geographic variables; a partial model in which geography explains genetic data conditioned on important environmental variables; a partial model in which important environmental variables explain genetic data conditioned on geography. This analysis allowed for distinguishing between how much of the total explainable neutral and adaptive variance was due to the environment (after removing geographical effects), how much was due to geography (after removing environmental effects), and how much was due to the joint effect of both factors.

Results

Genotyping results and outlier detections

After filtering, the final dataset consisted of 7,710 SNPs from 478 individuals with a genotyping call rate of 97.6% and genotyping error rate of 0.022%. Detailed information about the number of sites retained at each filtering step can be found in Table 3.

Three outlier detection methods were used to identify SNPs putatively under divergent selection. The number of outliers identified by each method and analysis, and the overlap between methods is shown in Figure 3. Using the filtered dataset with all sites, pcadapt was the least conservative (442 SNPs), OutFLANK was intermediate (117

SNPs), and BAYESCAN was the most conservative (20 SNPs) with the lowest number of outliers detected. Finally, excluding the coastal Bay site, (Wachapreague) pcadapt was the least conservative (813 SNPs), OutFLANK was intermediate (18 SNPs), and BAYESCAN was the most conservative (9 SNPs) with the lowest number of outliers detected (data not shown). Any SNP identified as an outlier in at least one method was removed from the neutral dataset. The outlier dataset consisted of SNPs identified as an outlier in at least two methods (BAYESCAN, OutFLANK, pcadapt, details below) and SNPs identified to be significant outliers in RDA.

Linkage disequilibrium and genetic diversity of wild and restored oysters

The critical R^2 calculated from the intra-chromosomal LD analysis was 0.0999 (root transformed 95th percentile of intra-chromosal LD; Breseghello and Sorrells 2006). The point at which the loess curve (fit to the intra-chromosomal LD) intercepted the critical R^2 was determined as the average LD decay. Based on these criteria, SNPs were thinned at a distance of 2137 bp (Figure 2, averaged across 10 chromosomes). After removing linked SNPs, the thinned neutral dataset consisted of 2,810 SNPs.

The critical R^2 calculated from the intra-chromosomal LD analysis excluding W sample was 0.0998 (Breseghello and Sorrells 2006). SNPs were thinned at a distance of 1603 bp. After removing linked SNPs, the thinned neutral dataset consisted of 2,842 SNPs.

To explore patterns of genetic diversity, mean expected heterozygosity (H_e), observed heterozygosity (H_o), allelic richness (A_r), inbreeding coefficients (F_{IS}), relatedness, and effective population size (N_e) were calculated for each sampling site

using the thinned neutral dataset. H_e was similar between sampling sites, ranging from 0.212–0.237, while H_o differed more substantially among sites ranging from 0.185–0.243 (Table 4). All samples displayed higher levels of H_o than H_e except HCS (restored), HCW (MD wild), BR, and TS (VA wild). The Harris Creek restored sample HCS displayed the lowest H_o overall (0.180), while the restored Harris Creek site HCR1 displayed the highest H_o overall (0.243). Excluding the single cohort HCS sample, all Harris Creek restored samples displayed slightly higher levels of H_o than both wild Maryland and Virginia populations (Table 4). Allelic richness was similar between sampling sites ranging from 1.944–1.99 (Table 4), but showed an interesting trend for restored oyster samples, and increased as planting frequency increased (Figure 4a). The coastal Bay W sample displayed the lowest allelic richness (Table 4). F_{IS} values ranged from -0.066 (HCR1) to 0.165 (HCS) and about half of all F_{IS} coefficients were negative. The restored mixed cohort sites had the lowest F_{IS} overall (HCR1, HCR2, HCR4) as well as the coastal Bay (W) site. Global relatedness trends ranged from 0.0022–0.066 (lowest in TS and highest in W; Table 4). Relatedness of mixed cohort restored reefs decreased as planting frequency increased and the HCS sample had the lowest relatedness of all restored samples using both methods. For the Maryland wild samples, LC had the lowest relatedness using (0.0057). For the wild Virginia samples, TS has the lowest relatedness the inner Bay wild Virginia samples had lower relatedness than wild Maryland samples overall (Table 4).

Effective population size effects of hatchery planting frequency and broodstock number

Estimates of N_e were variable across sampling sites, ranging from 71.5 (HCR1) to 584.8 in (LC) (Table 4). The wild Maryland samples ranged from 75.9–584.8, the wild Virginia samples ranged from 124.0–408.0, and the restored Harris Creek samples ranged from 71.5–335.1. All but one (TS) of the estimates were bounded at the jackknife confidence limits (jackknife confidence interval range 43.2–26771.9 Table 4). While the adjusted point estimate of N_b from a single cohort of juveniles was the lowest (66.3), the upper confidence limit was higher than all MD samples except LC and HCR4. The N_e estimates for the restored Harris Creek samples increased as hatchery planting frequency increased and were higher than N_e estimates from the wild MD populations HCW and TB (Table 4). Overall, the N_e estimates for the Harris Creek restored samples were similar to the range of values estimated for wild populations in Maryland and Virginia, and confidence limits for the wild and restored populations overlapped substantially.

The number of broodstock used for hatchery plantings was significantly positively correlated with N_e ($P = 0.027$, $R^2 = 0.912$, Figure 4a) and relatedness ($P = 0.021$, $R^2 = 0.997$, Figure 4b) of restored reefs. The number of hatchery planting seasons was significantly positively correlated with N_e ($P = < 0.001$, $R^2 = 0.999$, Figure 4c) and relatedness ($P = 0.020$, $R^2 = 0.995$, Figure 4d) of restored reefs. The number of hatchery planting seasons at each site was also positively correlated with H_o ($R^2 = 0.75$, Figure 4e), and the average broodstock sex ratio was positively correlated with H_o ($R^2 = 0.926$, Figure 4f), but neither were statistically significant ($p > 0.08$). There was a non-significant negative relationship associated with reef size (acres) and N_e ($P = 0.61$, $R^2 = -0.33$) and H_o ($P = 0.72$, $R^2 = -0.65$). Overall, the relationships between genetic diversity

metrics and hatchery practices (planting effort and broodstock size) were positive, strong, and highly predictive.

Genetic differentiation, population structure, and population assignment

Pairwise F_{ST} estimates between wild and restored populations were small, ranging from 0.001–0.032 (Figure 5). All pairwise F_{ST} estimates were highest between the coastal Bay Wachapreague (W) site and all other sites ($0.021 < F_{ST} < 0.032$). Pairwise F_{ST} estimates between HCR1 and the inner Bay sites were higher than comparisons among other inner Bay populations ($0.008 < F_{ST} < 0.012$). Similar to F_{ST} results, analyses of population structure via DAPC revealed four major population clusters, with the coastal Bay (W) sample grouping distinctly from all wild and restored inner Bay sites (Figure 6). In addition, subtle genetic differences were observed between the HCR1 site and the rest of the sites from Harris Creek (Figure 6). Analysis in STURCTURE (Figure 7) also suggested four clusters based on both the mean likelihood values (L(K)) and the Evanno method (deltaK). Finally, Mantel tests showed a significant correlation between pairwise F_{ST} and water distance for the neutral dataset, indicating a moderate trend of isolation by distance (adjusted $R^2 = 0.139$, $P = 0.001$; Figure 8), even when restored samples were removed ($R^2 = 0.056$, $P = 0.012$; Figure 8).

Assignment success using Assigner depended both on the number of markers used (more was generally better) and the sample site (Figure 9). Median overall assignment accuracy was 27, 38, 38, 46, 54, 58, 62, and 64% when using the top (based on F_{ST}) 50, 100, 200, 500, 1000, 2000, 5000, and all SNPs, respectively. One hundred percent of individuals from Wachapreague were correctly assigned to their sampling site

using only the top 50 markers. In contrast, only 11% of individuals from Tangier Sound (TS) could be assigned to their sampling site using all SNPs (Figure 9). In addition, sites TS and HCS both had little to no improvement in assignment accuracy by increasing the number of markers. In contrast, JR had a nearly four-fold increase in assignment accuracy (14% to 52%) by increasing from 50 to 500 SNPs.

Genotype by environment association results

The global model of the genotype-environment RDA conducted using the full filtered dataset (7,710 SNPs) to detect candidate loci under selection was highly significant (P = 0.001) (Figure 10). There were six significant ($p < 0.05$) RDA axes which returned 269 unique SNPs (121 were significant) that loaded ± 3 SD from the mean loading on each axis: 152 SNPs detected on RDA axis 1, 34 SNPs detected on RDA axis 2, 24 SNPs detected on RDA axis 3, 22 SNPs detected on RDA axis 4, 22 SNPs detected on RDA axis 5, and 15 SNPs detected on RDA axis 6. The majority (32.3%) of candidate SNPs (87) were most strongly correlated with mean salinity (43 SNPs were significant at the alpha 0.05 level). The SNPs correlated with the remaining predictors, and their significance is shown in Table 4. Ninety-nine of these significant SNPs were also detected as outliers in at least two of the differentiation-based methods (54 overlapped when thinned). Interestingly, most of the SNPs correlated with salinity were located on chromosomes five and six (Figure 11).

For the first RDA using the neutral, unlinked dataset (2,810 SNPs), geography (PCNM2) and environmental variables (mean salinity, minimum DO, minimum pH, and mean pH) were selected for RDA. The global RDA was highly significant and explained

1.1% of the genetic variation (P = 0.001; Table 5). Partitioning of total variance analysis (comparing the full model with a partial model conditioned on environmental variables and a partial model conditioned on geography) indicated that the environment explains 84.9% of the total explainable genetic variance after removing variance explained by geography (P = 0.001); geography explains 14.6% of the genetic variance after removing variance explained by environment (P = 0.001); and environment and geography together have a joint effect of 5.3% on genetic variance (also the proportion of variance in which climate and geography cannot be separated due to collinearity). RDA analysis indicated mean salinity and mean pH to be the most important predictors of neutral genetic variation among all environmental and spatial variables considered, respectively.

For the RDA based on the SNPs identified as being putatively adaptive (9 SNPs), a single geographic variable (PCNM2) and six environmental variables (mean salinity, maximum water temperature, minimum DO, minimum pH, and mean pH) were selected for RDA. The global RDA was highly significant and explained 23.1% of the genetic variation (P = 0.001). Partitioning of total variation indicated that the environment explained 94.4% of the genetic variation (P = 0.001); geography explained 3.5.% of the genetic variance (P = 0.001); the environment and geography together have a joint effect of 2.1% on genetic variance. RDA analysis indicated that mean salinity and minimum DO to be the most important predictors of adaptive genetic variation among all environmental and spatial variables considered.

The global model of the genotype-environment RDA conducted for the inner Bay samples to detect candidate loci under selection was highly significant (P = 0.001) (Figure 10). There were six significant (P = 0.001) RDA axes which returned 145 unique

candidates (3 SNPs were significant) that loaded +-3 SD from the mean loading on each axis: 37 SNPs detected on RDA axis 1, 41 SNPs detected on RDA axis 2, 19 SNPs detected on RDA axis 3, 16 SNPs detected on RDA axis 4, 19 SNPs detected on RDA axis 5, and 12 SNPs detected on RDA axis 6. The SNPs correlated with the environmental predictors is shown in Table 5. All of these SNPs were also detected as outliers in at least two of the differentiation-based methods.

For the RDA based on the unlinked neutral SNPS for the inner Bay samples, a single geographic variable (PCNM2) and five environmental variables (mean salinity, mean water temperature, minimum DO, and minimum pH) were selected for RDA. The global RDA was highly significant and explained 0.59% of the genetic variation (P = 0.001; Table 6). Partitioning of total variation indicated that the environment explained 81.7% of the genetic variation (P = 0.001); geography explained 18% of the genetic variance (P = 0.001); the environment and geography together have a joint effect of 3.6% on genetic variance. RDA analysis indicated mean salinity and mean water temperature to be the most important predictors of neutral genetic variation among all environmental and spatial variables considered, respectively.

For the RDA based on the SNPs identified as being putatively adaptive for the inner Bay samples (excluding coastal Bay W sample) (6 SNPs), a single geographic variable (PCNM2) and five environmental variables (mean salinity, maximum water temperature, minimum DO, and mean pH) were selected for RDA. The global RDA was highly significant and explained 7.5% of the genetic variation (P = 0.001). Partitioning of total variation indicated that the environment explained 91.2% of the genetic variation (P = 0.01), geography explained 8.2.% of the genetic variance (P = 0.012), and the

environment and geography together had a joint effect of 0.6% on genetic variance. RDA analysis indicated that mean salinity and maximum water temperature were the most important predictors of adaptive genetic variation among all environmental and spatial variables considered.

Functional annotation of outlier loci

The SNPs identified as outliers in at least two genome-scan methods (Bayescan, pcadapt, OutFLANK) and RDA were distributed across all 10 chromosomes. Of the SNPs that were identified on the eastern oyster genome (211 SNPs), 84 were located within genes, 23 of which were uncharacterized proteins, and 62 of which had gene ontology (GO) annotations. Several genes were involved in ion binding and osmoconformation including sodium/hydrogen exchanger beta-like, propionyl-CoA carboxylase alpha chain, mitochondrial, calcium uptake protein 3 mitochondrial-like isoform X4, and extracellular tyrosine-protein kinase PKDCC-like. Several genes were involved in cellular nitrogen metabolic processes including ATP-dependent DNA helicase DDX11-like and nicotinamide/nicotinic acid mononucleotide adenylyltransferase 1-like isoform X2, and carbohydrate metabolic processes including sucrase-isomaltase 2 intestinal-like and maltase-glucoamylase 2 intestinal-like (Table 7).

The SNPs identified in at least two methods and RDA excluding the W sample, were distributed across all 10 chromosomes. Of the SNPs that were identified on the eastern oyster genome, 74 were located within genes, 17 of which were uncharacterized proteins, and 48 of which had gene ontology (GO) annotations. Most of the genes were involved in ion binding including Kv channel-interacting protein 4-like isoform X16,

lipoprotein receptor-related protein 2-like, and oxidoreductase activity including xanthine dehydrogenase/oxidase-like isoform X2 and ecto-NOX disulfide-thiol exchanger 2-like.

Discussion

Understanding patterns of neutral and adaptive genetic variation is critical to establishing population restoration programs that aim to preserve genetic diversity, maintain genetic structure, and promote resilience in the face of rapid environmental change. However, for marine shellfish species with complex life-history features, this information is often unavailable or is not integrated into management. A RAD-Seq approach was used to characterize patterns of genetic variation within and among wild and restored eastern oyster populations in the Chesapeake Bay and the high-resolution data was used to investigate population structure, local adaptation, and the extent at which environmental gradients influence genetic variation among these populations.

There are four major findings of this study that provide critical information for management of eastern oysters, which typify the periodic, broadcast spawning life history of many other marine animal species (Winemiller and Rose 1992). First, restored oyster reefs in Harris Creek, MD had similar levels of genetic diversity compared with proximal wild populations. Second, the number of broodstock used for spat production and the frequency of restoration planting had strong (highly predictive) and positive associations with metrics of genetic diversity including N_e and relatedness. Third, despite previous restoration efforts, frequent historical translocations, and high dispersal potential of oyster larvae, we uncovered a moderate degree of neutral population genetic structure in wild and restored Chesapeake Bay oyster populations which suggests that fine scale

population structure can exist over small scales for marine shellfish. Finally, strong correlations between environmental variables and outlier loci were found suggesting that local adaption or genotype by environment interactions are driving the adaptive differentiation of oysters over relatively small scales. This adds to the growing evidence of fine-scale genetic structure and local adaptation in marine species. These results suggest that sourcing wild broodstock from large, local populations experiencing similar environments to candidate sites is likely to provide the most appropriate sources for hatchery-based restoration of oysters.

Comparison of genetic diversity and N_e between restored vs. wild oysters

In general, estimates of genetic diversity in these Chesapeake Bay oyster populations were comparable to other published datasets. Notably, restored oysters from Harris Creek had comparable levels of genetic diversity to wild oysters from Maryland and Virginia. More than half (6/11) of the estimated inbreeding values (F_{IS}) across sampling locations were negative indicating heterozygosity excess, and those that were positive were lower than those observed using SNP datasets in Canadian eastern oyster populations (F_{IS}=0.191–0.211; Bernatchez et al. 2019), and other oyster species, such as the black lip pearl oyster ($F_{IS}\geq0.5$; Lal et al. 2018). Inbreeding levels were also lower than what was observed in a recent study of oyster populations in the lower Chesapeake Bay using 48 SNPs (F_{IS}=0.02–0.156, Turley et al. 2019), and in Rhode Island using microsatellites (F_{IS}=0.00–0.47; Jaris et al. 2019), and comparable to those observed in a recent study of the Olympia oyster using genome-wide SNPs (F_{IS} = –0.09–0.133; Silliman 2019). Mean heterozygosity (H_o and H_e) was within the range of or slightly

lower than what has been observed in studies using similar markers (SNPs) in oysters. For example, observed heterozygosity levels were similar to those observed in Canadian eastern oyster populations (Bernatchez et al., 2019), but lower than what was observed in Delaware Bay oysters (0.329–0.343; Thongda et al. 2018). Relatedness of restored and wild populations was similar to values previously reported in wild (0.002–0.011) and restored (0.012) populations in Chesapeake Bay (Hornick and Plough 2019) and substantially lower than that of hatchery-produced offspring (0.03–0.129) except for HCR1 (0.030), which was at the lower end of this range. Overall, these results suggest that genetic diversity of restored and wild oyster populations in the upper Chesapeake Bay are comparable and that large-scale hatchery-based restoration has not caused significant declines in diversity, at least based on the reefs sampled and metrics examined. This was also found previously (Hornick and Plough 2019), albeit with limited sampling and marker resolution.

Another important metric of comparison between wild and restored oysters is N_e, which allows for prediction of a population's current and future viability. In general, N_e estimates in this chapter were similar in the magnitude of values reported for eastern oyster populations in other regions of the US east coast. For example, estimates of N_e were similar to previous estimates for oysters in the Delaware Bay (37–437; He et al. 2012), in the James River (535–1 516; Rose et al. 2006), and in the Choptank River (68.3–178.2; Hornick and Plough 2019), but are higher than those reported in the Delaware Bay (33.8) by Hedgecock et al. (1992) (see Chapter 2 for a discussion of N_e estimates in oysters and their conservation context). However, Chesapeake Bay N_e estimates, these and others, are much lower than those estimated by Bernatchez et al.

(2019) for Canadian oyster populations, which utilized a similar RAD-seq genotyping approach (examined genome-wide SNPs; Table 8). In fact, point estimates of N_e from this chapter (and associated confidence intervals) are consistently an order of magnitude lower than the Canadian sites (N_e range 236.8–7071.7, Table 8), except for one Canadian population (COC), which was of a similar order of magnitude to our estimates (Table 8). The difference between estimates from US vs. Canadian population estimates may be due to any number of environmental, exploitative, and demographic differences between the regions and we acknowledge the caveats associated with comparing these two datasets (e.g. different restriction enzymes used and different numbers of SNPs examined, etc.), as well as the numerous caveats associated with N_e estimation in general (Waples et al. 2013, 2014, 2016a). Still, the differences are substantial, and it is possible that the N_e of Canadian populations is much larger than populations along the US east coast, which have experienced much more intensive harvest pressure and human impacts led to population declines (Beck et al. 2011). For example, Canadian populations of oysters are characterized as "fair" compared to the descriptor for Chesapeake Bay populations, which is "poor" (based on Beck et al. 2011). The finding that restored reefs in Chesapeake Bay exhibit similar genetic diversity to wild populations in the region is important, but perhaps less impressive if one considers that substantial population declines of oysters have occurred in the Chesapeake Bay over the last century (Beck et al. 2011). Thus, comparisons between restored and contemporary wild reefs overlook the potentially large differences between present and historical diversity (i.e. shifting baselines). If these estimates of N_e in Canadian populations are accurate, and if they are broadly reflective of reduced anthropogenic

impacts over time (e.g. lower fishing pressure; Beck et al. 2011), N_e of Chesapeake Bay oyster populations (wild or restored) are still much reduced compared to what they likely were in the past. Therefore, maintaining diversity of extant Chesapeake Bay wild populations should only be a minimum target.

Effect of planting history and broodstock size on restored reef diversity

Another major finding from this study is that the number of broodstock used for hatchery plantings and the number of hatchery planting seasons significantly impact diversity at restored reefs in Harris Creek. Highly significant and strongly predictive positive relationships between planting effort and broodstock size and genetic diversity metrics (N_e and relatedness) as well as a positive correlation between broodstock male-to-female ratio and observed heterozygosity were found. These are fascinating results because to date, few studies have been able to assess how hatchery production techniques can directly (and positively) impact genetic diversity of cultured and supplemented populations of shellfish. A similar result was found in a recent study of eastern oysters, in which the ratio of males-to-females in broodstock was positively correlated with metrics of genetic diversity of hatchery-produced eastern oyster cohorts (Hughes et al. 2019). However, Hughes et al. (2019) was not focused on restoration specifically and the experiment was on a much smaller scale. Using an individual-based model, Katalinas et al. (2019) investigated how stock enhancement practices such as the number of breeders and relative contribution of stocked fish impact levels of genetic diversity on the wild spawning population of red drum in South Carolina (Katalinas et al. 2019). Model results indicated that in order to maintain genetic diversity of the wild population, the stock

enhancement program should use at least 10 effective breeders in the hatchery (replaced annually), with mean contributions of stocked fish at less than 30% (Katalinas et al. 2019). Future simulation-based work incorporating shellfish life-history features and empirical genetic data would be useful for quantifying genetic diversity changes associated with varying hatchery practices. It is clear that the use of large broodstock numbers from multiple local sites, and the planting of multiple cohorts over many planting seasons will increase diversity of restored sites, especially when initial broodstock numbers are limited. More empirical work is needed to understand how hatchery practices directly influence genetic diversity of supplemented populations, especially in species with complex life-history features that may make maintaining genetic diversity in the hatchery less manageable (e.g. Hornick and Plough 2019). Nevertheless, these relatively simple and modifiable hatchery or husbandry practices (broodstock number, male-to-female ratio of broodstock, and planting frequency) may offer a straightforward way to achieve short-term goals of abundance increase while also approaching long-term goals of maintaining diversity and promoting self-sustaining wild populations.

Population structure and adaptive divergence across environmental gradients

Contemporary population structure of Chesapeake Bay eastern oysters is a product of diverse factors including larval dispersal and behavior, natural selection over environmental gradients, genetic drift, and demographic history. Though weak or negligible genetic structure is often assumed for marine broadcast spawning species over short spatial scales (e.g. 10s of km; Bradbury et al. 2008), subtle but significant

population structure among oyster populations in the Chesapeake region was uncovered as evidenced by genetic clustering of proximal sites and significant isolation by distance (IBD) over the length of the estuary. This contrasts with the expectation that decades of replenishment and restoration activities in Maryland and Virginia, which have led to substantial movement of oysters, would homogenize allele frequencies and limit any signatures of environmental and geographic population structure. Given the fact that larval periods of 2-3 weeks would allow for dispersal distances well beyond the scale of structure found, it is likely that the heterogeneous estuarine environment is driving this structure. Redundancy analysis (RDA) indicated that environmental gradients have a stronger effect on genetic variation than distance-based isolating factors such as genetic drift. In regions like maritime provinces in Canada, which have experienced less fishing pressure and human-assisted migration, studies have shown that environmental factors play a critical role in the distribution of neutral and putatively adaptive genetic variation in oysters (Bernatchez et al. 2019). Alternatively, it is possible that the observed patterns of structure may, in part, reflect ancestral population structure since the last movements of oyster occurred during recent times (c.a. 2006), but more work would be required to test this hypothesis. Future studies incorporating coalescent modeling approaches could provide important information regarding the historical relationships of Chesapeake Bay oysters (e.g. Chen et al. 2017; Díaz et al. 2018).

The genetic by environment analysis (redundancy analysis, RDA) indicates that salinity was the most important predictor of both neutral and adaptive variation. The observed neutral population structure uncovered in this chapter may be related to the influence of salinity on larval dispersal. Salinity influences larval duration, growth, and

survival (Davis 1958; Hidu and Haskin 1978; Kennedy 1996) during the period when oysters disperse. Salinity is also a critical factor that cues vertical swimming behavior and transport of oyster larvae in laboratory (Hidu and Haskin 1978). Eastern oysters have great capacity to tolerate a range of salinities (e.g. Shumway 1996) and a number of studies suggest adaptation related to salinity tolerance in oysters (Newkirk 1978; Buroker 1983; King et al. 1994; Bible and Sanford 2016; She et al. 2018; Bernatchez et al. 2019). While oysters lack the ability to adjust extracellular fluids, they have a compensatory machinery for transporting osmotically active solutes including free amino acids (FAAs) (Pierce and Amende 1981; Zhao et al. 2012). Genes correlated with salinity in our RDA analyses were, as expected, involved in osmoconformation (sodium/hydrogen exchanger beta-like), hydrolase activity (sucrase-isomaltase intestinal-like and metabolism and maltase-glucoamylase intestinal-like) and G-protein coupled receptor activity (adhesion G protein-coupled receptor L3-like; Table 7). Eierman and Hare (2014) found phosphorylation and hydrolysis of peptides to be part of the physiological response to osmotic stress. Moreover, hydrolase activity was found to be down regulated in oysters in response to low salinity (Jones et al. 2019) indicating a direct link between response to salinity and this pathway. Interestingly, most of the genes associated with salinity were located on chromosome five and six (Figure 11); however, significant SNPs were located across all chromosomes which is suggestive of locally adapted variation being pervasive throughout multiple genomic regions (i.e. is polygenic). These results provide insight into the mechanisms of salinity adaptation in oysters and the SNPs identified here provide potential targets for genetic monitoring programs and information regarding the selection of broodstock sources in the Chesapeake Bay.

Other variables such as temperature and dissolved oxygen (DO) were also correlated with genetic variation, but the effect size and significance were lower. Temperature can considerably affect oxygen consumption, energy metabolism, mitochondrial efficiency, and thus ATP-related functions (Abele et al. 2002; Sokolova 2004; Chamberlin 2004; Cherkasov et al. 2006; Ivanina et al. 2012). Many of the SNPs associated with temperature were located within genes known to have ATP-related functions (extracellular tyrosine-protein kinase PKDCC-like, ATP-dependent DNA helicase DDX11-like; Table 7) suggesting that adaptive divergence could be linked with thermal adaptation and energetic metabolism, as revealed in other *Crassostrea* species (Li et al. 2017). Similarly, this result suggests that in oysters, temperature may determine the amount of energy spent on maintenance and growth, as most biological processes, i.e. protein synthesis and degradation, are temperature dependent. When exposed to hypoxia and anoxia, oysters reduce metabolic rate up to 90% and decrease oxygen consumption (Shumway and Koehn 1982). Low DO levels are common in marine and estuarine systems such as the Chesapeake Bay (Taft et al. 1980; Breitburg 1992). where the temporal and spatial severity of hypoxic episodes has been exacerbated by anthropogenic sources (Diaz 2001; Hagy et al. 2004; Kemp et al. 2005). Many of the SNPs associated with DO were located within genes known to function in catalytic activity and metabolism (uncharacterized protein LOC111121258 isoform X2, ecto-NOX disulfide-thiol exchanger 2-like, and sucrase-isomaltase intestinal-like). Despite some interesting and sensible results for DO, there is a possibility of type II error for temperature and DO associations because the range of values for these variables were not as dynamic as those of salinity (Figure 12B), and the sampling resolution of the environmental data was

limited (twice each month). Future studies including finer resolution within-bay population and environmental sampling may reveal further patterns of selection and differentiation and may potentially impact the broad-scale correlations observed here.

Restoration implications

Results from this chapter provide evidence that oyster populations in the Chesapeake Bay may be locally adapted to prominent environmental features, particularly salinity, which has direct management implications. First, the finding of local adaptation over small spatial scales suggests limiting introgression from divergent populations (Conover 1998; do Prado et al. 2018) by favoring the use of local wild broodstock for restoration. Collecting local broodstock could be beneficial because nearby populations are likely to be more connected by gene flow and experience similar environments. However, because geographic distance did not significantly predict neutral or adaptive variation, matching environmental conditions of collection and restoration sites may be more important for broodstock and/or seed selection than geographic distance alone (McKay et al. 2005; Bischoff and Hurault 2013). This result is particularly important for estuarine species restoration, as stark environmental change over small geographic scales is common (Elliott and McLusky 2002; McLusky and Elliott 2007). Whether the use of local broodstock results in increased survival of planted individuals and better restoration outcomes merits future investigation, as results of previous studies have been mixed. For example, Bible and Sanford (2016) performed reciprocal transplants of Olympia oyster offspring from three sites in San Francisco Bay and found that oysters of local origin tended to survive better than locally non-adapted sources.

These results suggested that local adaptation may occur even within a single estuary (Bible and Sanford 2016). A recent study of low salinity survival, showed that larval survival at a given salinity seems to be matched to the salinity of the parental population (or conditioning salinity) (Scharping et al. 2019). Whether this survival is a result of adult acclimation vs. local adaptation merits future work. However, local sources do not always perform better than all other sources (Leimu and Fischer 2008; Hereford 2009). A recent study of eastern oysters documented significant genetic by environment (GxE) variation in survival and growth, but no evidence for local adaptation (Hughes et al. 2017). Thus, the benefits of using local broodstock may depend on idiosyncrasies of the specific set of populations under study, including the amount of standing diversity in the system and the connectivity among populations. Restoration plans aimed at conserving multiple interconnected reefs will likely capture an important axis of adaptive variation and maintain genetic diversity of restored populations. Future work should incorporate a larger number of populations to quantify the spatial scale of local adaptation (Hice et al. 2012) and reciprocal transplant experiments to determine if local populations perform better than non-local counterparts.

Despite the potential benefits of using locally adapted broodstock, broodstock collection from local populations may not be ideal or feasible for restoration in some cases. In areas such as Australia and Europe, native populations of shellfish have been driven to local extinction, so sourcing broodstock locally is not an option (Beck et al. 2011). In addition, some local oyster populations may lack a sufficient amount of genetic variability to adapt to rapidly changing climatic conditions (Montalvo et al. 1997; Rice and Emery 2003; McKay et al. 2005; Harris et al. 2006; Jones 2013), or they may be

small and/or inbred (Leimu et al. 2006). In such cases, sourcing broodstock from a number of local and/or regional populations may be the only solution. As shown in this chapter, the intensity of selection gradients and rate of gene flow can vary widely, so it remains difficult to prescribe a standard geographic distance as a scale for local adaptation. In the Chesapeake Bay, the availability of fine-scale environmental data can potentially aid in identifying the drivers of adaptive differences between reefs rather than just geographic distance, in order to delineate zones by environmental distance to be used as guides for broodstock selection. Therefore, the idea of using a more widely available "coarsely adapted" mixture of broodstock sources that contain genetic variation for further adaptive fine-tuning may increase restoration success by aiding in the preservation of a species' adaptive potential (Lesica and Allendorf 1999; Rice and Emery 2003). Nevertheless, these results indicate that when considering broodstock sources based on adaptive differentiation in heterogeneous environments, collecting broodstock from large populations from similar environments to candidate sites should form the basis of broodstock sourcing guidelines.

Conclusions

This study provides comprehensive characterization of neutral and adaptive population structure of restored and wild oysters in the Chesapeake Bay and is the first study to investigate genetic changes of restored reefs associated with variable hatchery-planting frequencies. The results obtained here suggest that using large numbers of local, wild broodstock in hatchery-based restoration programs and planting of reefs multiple times (especially if broodstock numbers are low), can increase diversity. Furthermore,

results from this chapter contribute to the growing body of evidence that adaptive differentiation can occur over very fine geographic scales in marine species and suggest that this structuring is at least partly driven by spatial heterogeneity in environmental parameters like salinity, temperature, and dissolved oxygen. The eastern oyster is a commercially exploited species with large-scale restoration efforts underway in the Chesapeake Bay and in other regions (e.g. Brumbaugh and Coen 2009; Dinnel et al. 2009; Holley et al. 2018), and an understanding of spatial patterns of neutral and adaptive genetic differentiation can inform management to ensure the efficiency of restoration and sustainability of oyster populations in the future. More broadly, this chapter demonstrates the utility of genomic-based approaches and provides an overall framework for other studies that aim to integrate genomics into conservation management for enhanced restoration success.

Acknowledgements

I thank the HPL Oyster Hatchery, Kennedy Paynter, Melissa Southworth, and PG Ross for providing oyster samples, and Genewiz for DNA sequencing. I also thank Katherine McFarland for helpful comments on previous versions of the manuscript and Katherine Silliman for online workbooks that assisted with genomic analyses. Support for this work was provided by the by the Deerbrook Charitable Trust (DCT 15-30 to LVP), Maryland Sea Grant (5-230360 to KMH and LVP), and Maryland Sea Grant (5-230890 to KMH and LVP).

Tables Chapter 3

Table 1. Location, latitudinal range, type, sample size, number of samples successfully genotyped (Ngen) and size ranges for each Chesapeake Bay sampling site of eastern oysters. Hatchery plantings denotes the number of seasons a restored site was planted with hatchery-produced oysters.

Site	Abbreviation	Type	hatchery plantings	Latitude	Longitude	Location	N	Ngen	Size (mm)
Harris Creek	HCR1	restored	2014	38.735323	-76.30243	MD	50	33	103.7 ± 17.7
Harris Creek	HCR2	restored	2015, 2016	38.711485	-76.316936	MD	50	37	57.3± 8.8
Harris Creek	HCR4	restored	2011, 2013, 2014, 2017	38.731909	-76.302688	MD	50	43	67.4± 1.9
Harris Creek	HCS		2015, 2016	38.715637	-76.320025	MD	60	49	
Harris Creek	HCW	wild		38.710212	-76.318738	MD	60	53	
Little Choptank	LC	wild		38.5368	-76.254303	MD	48	47	87.4± 10.9
States Bank	TB	wild		38.57	-76.04	MD	48	48	127.3 ± 23.5
Beverly's Rock	BR	wild		37.5322	-76.253	VA	50	48	45-95
James River	JR	wild		37.012	-76.468	VA	48	47	75.2 ± 16.3
Tangier Sound	TS	wild		37.78303	-75.94814	VA	50	38	62.6 ± 13.3
Wachapreague	W	wild		37.612233	-75.66795	VA	50	35	102.7 ± 24.2

Table 2. Sampling site of eastern oysters, station ID (buoy), depth of buoy in meters, the sampling frequency (annual), and latitudinal range of sampling buoys. Distance represents the distance from the buoy to the oyster sampling site. Continuous sampling frequency represents buoys that sample in 15-minute increments.

Site	Station ID	Depth (m)	Sampling Freq.	Latitude	Longitude	Distance (km)
HCR1	XFG4618	0.3 m above bottom	continuous	38.74323	-76.30338	0.88
HCR2	XFG2810	0.3 m above bottom	continuous	38.712517	-76.316803	0.12
HCR4	XFG4618	0.3 m above bottom	continuous	38.74323	-76.30338	1.26
HCS	XFG2810	0.3 m above bottom	continuous	38.712517	-76.316803	0.45
HCW	XFG2810	0.3 m above bottom	continuous	38.712517	-76.316803	0.31
LC	EE2.2	1m depth intervals	2x/month	38.52609	-76.30408	4.49
TB	ET5.2	1m depth intervals	4x/month	38.5807	-76.0587	2.01
BR	CB5.3	1m depth intervals	15x4	37.91011	-76.17137	42.63
JR	LE5.3	1m depth intervals	2x/month	36.99044	-76.47544	2.49
TS	EE3.2	1m depth intervals	4x/month	37.98139	-75.92423	22.16
W	XBM8828	0.3 m above bottom	continuous	38.14825	-75.28622	68.37

Abbreviations of sampling sites are presented in Table 1.

Table 3. Summary of data filtering procedures: rows refer to filtering steps; columns refer to statistics for each step. For columns,'sites' refers to individual polymorphisms (SNPs, indels, or complex polymorphisms), and 'Inds' refers to individuals. 'Start', 'End', and 'Removed' refer, respectively, to the number of each unit before the filtering step, the number after the filtering step, and the number removed with the filter.

Filter	Start sites	End sites	Start Inds	End Inds	Removed sites	Removed Inds
Minor allele count <3, Mean site quality < 20, & Mean site call rate < 0.5	4544551	538478	570	570	4006073	0
Minimum depth =5	538478	538478	570	570	0	0
Filter_missing_ind script; Ind call rate < 0.5	538478	538478	570	491	0	79
Minor allele frequency < 0.05, mean site call rate < 0.9	538478	15731	491	491	522747	0
pop_missing_filter script; call rate 0.75 in one population	15731	14725	491	491	1006	0
dDocent_filters script	14725	7040	491	491	7685	0
Decomposed to allelic primitives	8163	7832	491	491	331	0
Hardy-Weinberg equilibrium	7832	7796	491	491	36	0
Max alleles=2	7796	7710	491	491	86	0
Remove reps	7710	7710	491	478	0	13

Table 4. Descriptive statistics for each *C. virginica* sampling site, including observed heterozygosity (H_o), expected heterozygosity (H_e), inbreeding coefficient (F_{IS}), allelic richness (A_r), effective population size (N_e), and relatedness Ritland (1996) and Wang (2016).

Sites	H_o	H_e	F_{IS}	A_r	N_e (CI) 0.20	N_e (CI) 0.10	Ritland	Wang
HCR1	0.243	0.231	-0.066 (-0.067, -0.048)	1.978	71.5 (38.2, 259.5)	76.9 (41.8, 266.6)	0.0293	0.0495
HCR2	0.242	0.230	-0.060 (-0.061, -0.043)	1.983	156.8 (116.1, 235.9)	161.3 (120.8, 238.0)	0.0213	0.0485
HCR4	0.234	0.227	-0.037 (-0.038, -0.019)	1.987	335.1 (222.4, 657.3)	349.1 (231.8, 684.6)	0.0080	0.0381
HCS	0.180	0.218	0.165 (0.163, 0.186)	1.982	66.3 (32.2, 312.8)	68.9 (33.6, 324.3)	0.0075	-0.1195
HCW	0.227	0.237	0.034 (0.033, 0.052)	1.989	75.9 (43.2, 196.8)	77.3 (44.0, 201.2)	0.0081	-0.0736
LC	0.223	0.223	0.0002 (-0.002, 0.017)	1.986	543.3 (387.7, 897.2)	584.8 (415.3, 976.4)	0.0057	0.0170
TB	0.232	0.228	-0.024 (-0.026, -0.009)	1.983	139.1 (91.4, 267.1)	143.5 (94.7, 273.8)	0.0116	0.0199
BR	0.216	0.226	0.038 (0.037, 0.056)	1.990	408.0 (199.3, 26771.9)	417.7 (203.6, 39441.8)	0.0032	-0.0331
TS	0.185	0.219	0.146 (0.144, 0.167)	1.987	124.0 (44.5, ∞)	130.8 (47.3, ∞)	0.0022	-0.1029
JR	0.224	0.223	-0.01 (-0.01, 0.008)	1.988	370.3 (227.3, 935.0)	390.4 (236.4, 1043.6)	0.0064	0.0246
W	0.220	0.212	-0.046 (-0.048, - 0.027)	1.944	203.0 (122.2, 538.8)	231.6 (138.6, 636.6)	0.0662	0.1097

HCS N_e and CIs represent adjusted N_e and CIs according to Waples et al. (2014). Abbreviations of sampling sites are presented in Table 1.

Table 5. Redundancy analysis (RDA) results for the full dataset for all populations (7,710 SNPs) and the inner Bay populations (excluding Wachapreague; 7,710 SNPs). Correlated represents SNPs identified as outliers in RDA and correlated with environmental parameters including mean salinity, maximum water temperature, minimum water temperature, minimum pH, mean pH, and minimum dissolved oxygen (DO). Significant SNPs represent those were considered to be significant outliers if the q-value was <0.05 (see Methods for details).

	All populations		**Inner Bay populations**	
Parameter	*correlated*	*significant*	*correlated*	*significant*
Salinity	87	43	36	1
Maximum water temp	38	18	27	0
Minimum water temp	41	17	22	1
Minimum pH	33	10	19	0
Mean pH	42	18	19	0
Minimum DO	28	15	22	1

Table 6. Redundancy analysis (RDA) results for potentially neutral and adaptive SNP datasets including all sites and only the inner Bay sites (excluding Wachapreague). The environmental parameters include mean salinity (salinity), maximum water temperature (maxWT), mean water temperature (mWT), minimum pH (minpH), mean pH (mpH), and minimum dissolved oxygen (DO; minDO).

Dataset	Variable types	Significant variables	Adjusted R^2	*p*-value
All sites				
	Global	—	0.011	0.001
2,810 neutral SNPS	Spatial	PCNM2	0.146	0.001
	Environmental	salinity	0.849	0.001
		mpH		
		minpH		
		minDO		
nine SNPS potentially under selection	Global	—	0.231	0.001
	Spatial	PCNM2	0.035	0.001
	Environmental	salinity	0.944	0.001
		mpH		
		maxWT		
		minDO		
		minpH		
Inner Bay sites				
	Global	—	0.0059	0.001
2,842 neutral SNPS	Spatial	PCNM2	0.18	0.001
	Environmental	salinity	0.817	0.001
		mWT		
		minDO		
		minpH		
six SNPS potentially under selection	Global	—	0.075	0.001
	Spatial	PCNM2	0.082	0.012
	Environmental	salinity	0.912	0.01
		maxWT		
		minDO		
		mpH		

Significance of the global model and significance of each variable in the partial RDA were evaluated using an ANOVA (10,000 permutations).

Table 7. BLAST matches from sequences identified as being putatively under divergent selection from oyster *Crassostrea virginica* populations in the studied region. SNPs are located in the eastern oyster genome and chromosomes. Protein IDs and names are derived from the eastern oyster protein sequences of the genome. Gene ontologies related to the identified protein (GO ID) were retrieved using protein names and the environmental variable correlated with proteins is listed.

Correlated variable	Chromosome	SNP position	Protein ID	Protein name	GO ID
minDO	chr1	42499514	XP_022318158.1	uncharacterized protein LOC111121258 isoform X2	GO:0006464;GO:0016301;GO:0043167
minDO	chr3	30266951	XP_022321732.1	protein sel-1 homolog 3-like	GO:0003674
minDO	chr4	20204343	XP_022300581.1	scm-like with four MBT domains protein 1	GO:0003674;GO:0005634;GO:0009058;GO:0034641
minDO	chr4	59161068	XP_022330446.1	neuropeptide F receptor-like	GO:0003674;GO:0005575;GO:0007165;GO:0007267
minDO	chr5	68843212	XP_022341851.1	ecto-NOX disulfide-thiol exchanger 2-like	GO:0005886;GO:0016491
minDO	chr5	68843434	XP_022341851.1	ecto-NOX disulfide-thiol exchanger 2-like	GO:0005886;GO:0016491
minDO	chr6	936212	XP_022288317.1	uncharacterized protein LOC111100577 isoform X12	GO:0043167
minDO	chr8	61940713	XP_022296589.1	uncharacterized protein LOC111106271	GO:0003674
minDO	chr8	53732997	XP_022300965.1	uncharacterized protein LOC111109176	
minDO	chr8	29659627	XP_022340780.1	centrosomal protein of 55 kDa-like	GO:0000278;GO:0003674;GO:0007165;GO:0051301

MxWT	chr1	53871422	XP_022286506.1	zinc finger protein Aiolos-like	GO:0003674	
MxWT	chr1	37752842	XP_022305935.1	nuclear receptor subfamily 1 group D member 2-like	GO:0003677;GO:0003700;GO:0005634;GO:0007165;GO:0009058;GO:0032991;GO:0034641;GO:0043167	
MxWT	chr2	7864021	XP_022313407.1	prostaglandin E2 receptor EP4 subtype-like	GO:0003674;GO:0005575;GO:0007165	
MxWT	chr4	1441592	XP_022333886.1	uncharacterized protein LOC111130906		
MxWT	chr4	23923663	XP_022333953.1	arrestin domain-containing protein 3-like		
MxWT	chr5	3083788	XP_022335658.1	1-phosphatidylinositol 4-5-bisphosphate phosphodiesterase gamma-1-like isoform X7	GO:0003674;GO:0006629;GO:0007165;GO:0009056	EC:3.1.4.11;EC:3.1.4.3;Phosphoinositide phospholipase C;Phospholipase C
MxWT	chr5	25069837	XP_022337704.1	centromere protein N-like isoform X2	GO:0007059;GO:0051276;GO:0065003	
MxWT	chr6	32082204	XP_022288785.1	probable ATP-dependent RNA helicase DDX58	GO:0043167	
MxWT	chr7	57219780	XP_022291424.1	uncharacterized protein LOC111102827	GO:0003674;GO:0005575;GO:0007165;GO:0048646	
MS	chr1	54789010	XP_022291696.1	rac guanine nucleotide exchange factor B-like isoform X3	GO:0003674	
MS	chr1	19401881	XP_022292455.1	angiopoietin-related protein 7-like		

MS	chr1	3055672	XP_022296641.1	E3 ubiquitin-protein ligase TRIM71-like	GO:0043167
MS	chr1	9430078	XP_022313050.1	G patch domain-containing protein 11-like	GO:0003674
MS	chr1	54741363	XP_022318158.1	uncharacterized protein LOC111121258 isoform X2	GO:0006464;GO:0016301;GO:0043167
MS	chr1	40736489	XP_022328667.1	patatin-like phospholipase domain-containing protein 2	GO:0003674;GO:0006629;GO:0009056
MS	chr2	22206407	XP_022315237.1	adhesion G protein-coupled receptor L3-like	GO:0003674;GO:0005575;GO:0007165
MS	chr2	43516534	XP_022317650.1		
MS	chr2	43516819	XP_022317651.1	myosin heavy chain striated muscle-like isoform X8	GO:0005856;GO:0008092;GO:0032991;GO:0043167
MS	chr3	61872567	XP_022320666.1	LOW QUALITY PROTEIN: low-density lipoprotein receptor-related protein 2-like	GO:0043167
MS	chr3	13041257	XP_022320827.1	polygalacturonase 1 beta-like protein 1	
MS	chr4	59301805	XP_022328845.1	uncharacterized protein LOC111127854	
MS	chr4	14408011	XP_022332702.1	coiled-coil domain-containing protein 171-like isoform X4	
MS	chr5	72463351	XP_022291782.1	uncharacterized protein LOC111103071 isoform X2	
MS	chr5	71889468	XP_022336578.1	zinc finger protein ZIC 4-like	GO:0003674
MS	chr5	73028262	XP_022337078.1	ATP-dependent RNA helicase DDX3X-like isoform X3	GO:0043167

MS	chr5	75161604	XP_022338224.1	ATP-dependent DNA helicase DDX11-like	GO:0003677;GO:0004386;GO:0005694;GO:0016887;GO:0034641;GO:0043167	EC:3.6.1.3;EC:3.6.1.15;Adenosinetriphosphatase;Nucleoside-triphosphate phosphatase
MS	chr5	72787167	XP_022338241.1	centrosomal protein of 19 kDa-like		
MS	chr5	73332073	XP_022338483.1	probable G-protein coupled receptor L3-like	GO:0003674;GO:0005575;GO:0007165	
MS	chr5	67489624	XP_022339347.1	lysosome-associated membrane glycoprotein 1-like	GO:0005575	
MS	chr5	62335338	XP_022340859.1	Na(+)/H(+) exchanger beta-like	GO:0005575;GO:0022857;GO:0042592	
MS	chr5	62335542	XP_022340859.1	Na(+)/H(+) exchanger beta-like	GO:0005575;GO:0022857;GO:0042592	
MS	chr5	73057556	XP_022341265.1	protein FAM124A-like isoform X2		
MS	chr5	66359736	XP_022342097.1	myoferlin-like isoform X19	GO:0005575	
MS	chr5	66381543	XP_022343779.1	uncharacterized protein LOC111136903 isoform X4		
MS	chr5	71392872	XP_022343889.1	uncharacterized protein LOC111136978		
MS	chr5	61804560	XP_022346086.1	uncharacterized protein LOC111138422		
MS	chr6	33055474	XP_022286676.1	uncharacterized protein LOC111099612 isoform X2	GO:0003674;GO:0007165	
MS	chr6	33055712	XP_022286676.1	uncharacterized protein LOC111099612 isoform X2	GO:0003674;GO:0007165	
MS	chr6	32397315	XP_022286836.1	sucrase-isomaltase intestinal-like	GO:0005975;GO:0016798	

MS	chr6	32397627	XP_022286836.1	sucrase-isomaltase intestinal-like	GO:0005975;GO:0016798
MS	chr6	32371906	XP_022286837.1	maltase-glucoamylase intestinal-like	GO:0005975;GO:0016798
MS	chr6	37347819	XP_022286861.1	cell division cycle protein 20 homolog	GO:0006464;GO:0009056;GO:0030234
MS	chr6	37113676	XP_022288271.1	propionyl-CoA carboxylase alpha chain mitochondrial-like isoform X2	GO:0043167
MS	chr6	37113888	XP_022288271.1	propionyl-CoA carboxylase alpha chain mitochondrial-like isoform X2	GO:0043167
MS	chr6	31381121	XP_022288369.1	putative transferase CAF17 homolog mitochondrial	GO:0003674
MS	chr6	43149879	XP_022289320.1	Kv channel-interacting protein 4-like isoform X16	GO:0043167
MS	chr8	46972312	XP_022298284.1	xanthine dehydrogenase/oxidase-like isoform X2	GO:0016491;GO:0043167
MS	chr8	41054116	XP_022310198.1	GTPase IMAP family member 4-like	GO:0043167
MS	chr10	18623025	XP_022311665.1	centromere protein F-like	
minWT	chr1	15360452	XP_022286854.1	MAM and LDL-receptor class A domain-containing protein 1-like isoform X2	GO:0003674;GO:0005575
minWT	chr3	48562353	XP_022320940.1	follistatin-related protein 1-like isoform X3	GO:0043167
minWT	chr5	36513326	XP_022298711.1	spermatogenesis-associated protein 20-like isoform X3	GO:0003674
minWT	chr5	62781849	XP_022305585.1	extracellular tyrosine-protein kinase PKDCC-like	GO:0006464;GO:0016301;GO:0043167

minWT	chr5	68753201	XP_022341843.1	centrosomal protein of 63 kDa-like isoform X9		
minWT	chr5	69858610	XP_022344414.1	ankyrin repeat domain-containing protein 34B-like isoform X2	GO:0003674	
minWT	chr9	103219837	XP_022324961.1	tetratricopeptide repeat protein 28-like isoform X4	GO:0003674	
mpH	chr3	69869130	XP_022322065.1	transient receptor potential cation channel subfamily V member 5-like isoform X7	GO:0005575;GO:0022857	
mpH	chr3	68487745	XP_022331683.1	endoribonuclease Dicer-like	GO:0004518	
mpH	chr5	40272733	XP_022340021.1	2C4-dienoyl-CoA reductase mitochondrial-like		
mpH	chr6	37335763	XP_022289242.1	E3 ubiquitin-protein ligase TRIM33-like isoform X3	GO:0005622;GO:0043167	
minpH	chr3	50442087	XP_022320066.1	glycine receptor subunit alpha-3-like	GO:0005575;GO:0007165;GO:0022857	
minpH	chr3	31340804	XP_022322746.1	uncharacterized protein LOC111124176		
minpH	chr3	68487745	XP_022331683.1	endoribonuclease Dicer-like	GO:0004518	EC:3.1.30;EC:3.1.26;EC:3.1.26.3; Acting on ester bonds;Acting on ester bonds;Ribonuclease III
minpH	chr4	2910965	XP_022332925.1	uncharacterized protein LOC111130315	GO:0003674	
minpH	chr4	9343106	XP_022333019.1	stress-induced-phosphoprotein 1-like	GO:0003674	
minpH	chr5	83486745	XP_022340236.1	uncharacterized protein LOC111134936 isoform X2	GO:0006464;GO:0007165	

minpH	chr5	2625449	XP_022342384.1	A-kinase anchor protein 10, mitochondrial-like	
minpH	chr5	25778951	XP_022346035.1	LOW QUALITY PROTEIN: inositol 1,4,5-trisphosphate receptor type 2-like	GO:0005783;GO:0022857;GO:0043167
minpH	chr9	70716765	XP_022304136.1	glycoprotein 3-alpha-L-fucosyltransferase A-like	GO:0005575;GO:0006464;GO:0009058;GO:0016757
minpH	chr9	11888150	XP_022307492.1	LOW QUALITY PROTEIN: 77 kDa echinoderm microtubule-associated protein-like	GO:0043167
minpH	chr9	53281019	XP_022316072.1	uncharacterized protein LOC111119843 isoform X3	GO:0003674;GO:0007165

Table 8. Effective population size (N_e) for *C. virginica* sampling site from this study (Chesapeake Bay) and from Bernatchez et al. (2019) (Canada)

Sites (Canada)	N	N_e (CI) 0.20	Sites (Chesapeake Bay)	N_e (CI) 0.20
BOU	40	3983.6 (2541.7, 9167.7)	HCR1	71.5 (38.2, 259.5)
COC	38	236.8 (69.1, ∞)	HCR2	156.8 (116.1, 235.9)
CRB	33	2240.7 (1772.3, 3043.2)	HCR4	335.1 (222.4, 657.3)
CRQ	40	5574.3 (3745.6, 10872.1)	HCS	66.3 (32.2, 312.8)
INK	38	7071.7 (4321.2, 19386.6)	HCW	75.9 (43.2, 196.8)
MAL	35	3017.8 (2249.9, 4575.8)	LC	543.3 (387.7, 897.2)
MIR	37	6297.4 (4078.9, 13777.6)	TB	139.1 (91.4, 267.1)
MIS	39	6254.2 (3959.3, 14834.4)	BR	408.0 (199.3, 26771.9)
RIC	39	1180.5 (832.4, 2017.6)	TS	124.0 (44.5, ∞)
SHD	39	2952.9 (2196.1, 4499.7)	JR	370.3 (227.3, 935.0)
SHM	38	2917.6 (1743, 8855.4)	W	203.0 (122.2, 538.8)
SSI	30	1917.6 (1503.5, 2644)		
TAB	40	5966.3 (4119.9, 10796)		

Abbreviations of Chesapeake Bay sampling sites are presented in Table 1.
Abbreviations of Canada sampling sites are presented in Bernatchez et al., (2018)

Figures Chapter 3

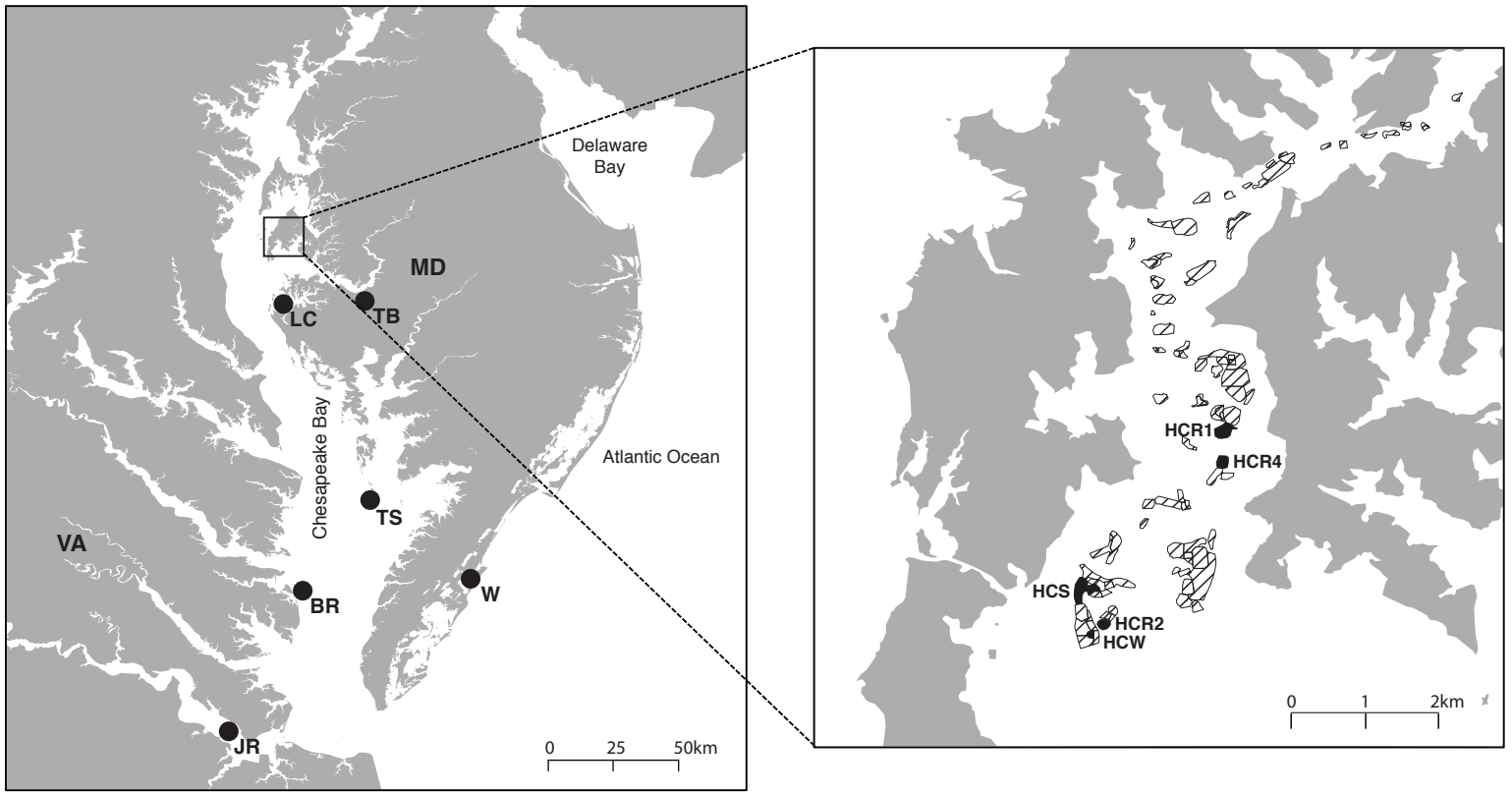
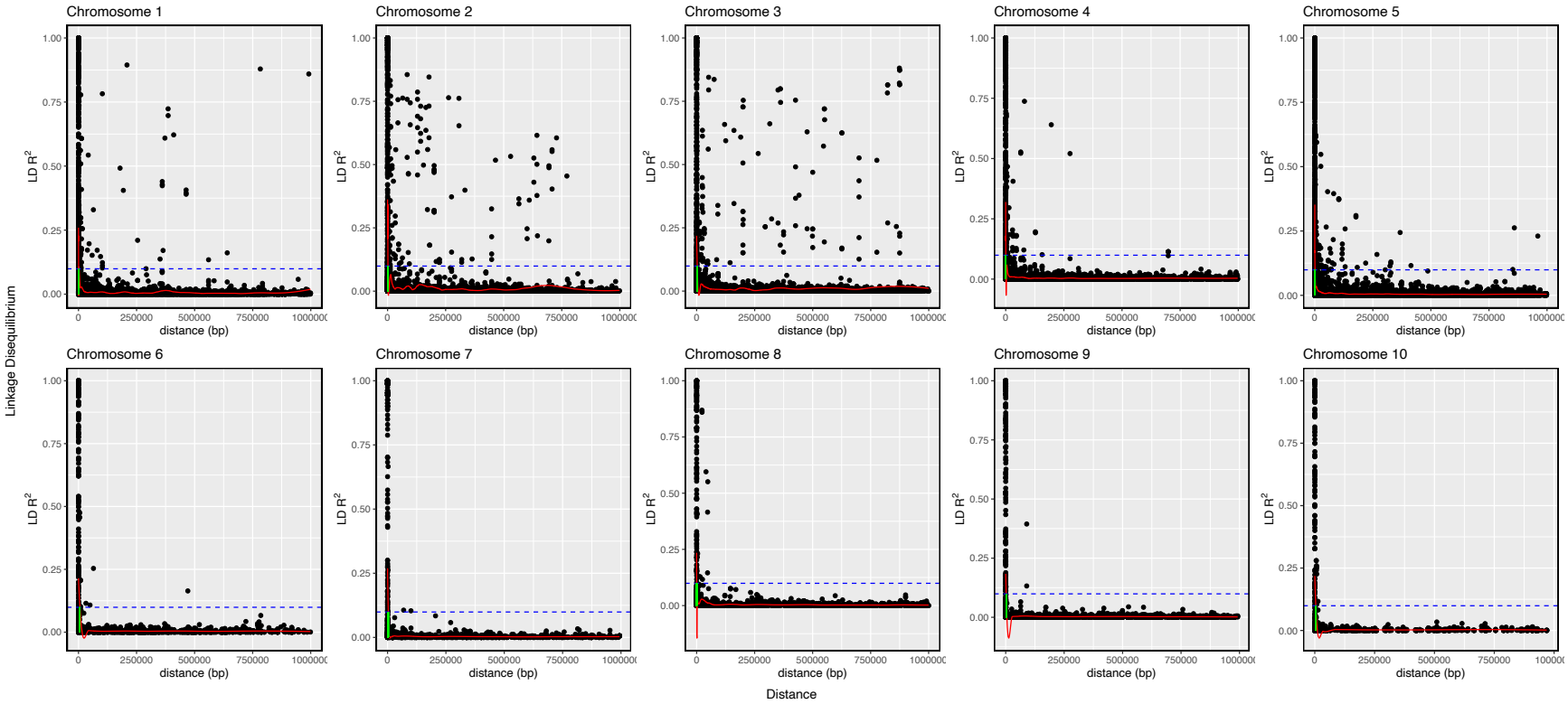


Figure 1 Map of sampling locations of eastern oysters within the Chesapeake Bay. Abbreviations of sampling sites are presented in Table 1.


Chromosome 1
Chromosome 2
Chromosome 3
Chromosome 4
Chromosome 5
Chromosome 6
Chromosome 7
Chromosome 8
Chromosome 9
Chromosome 10
Linkage Disequilibrium
LD R2
distance (bp)
Distance

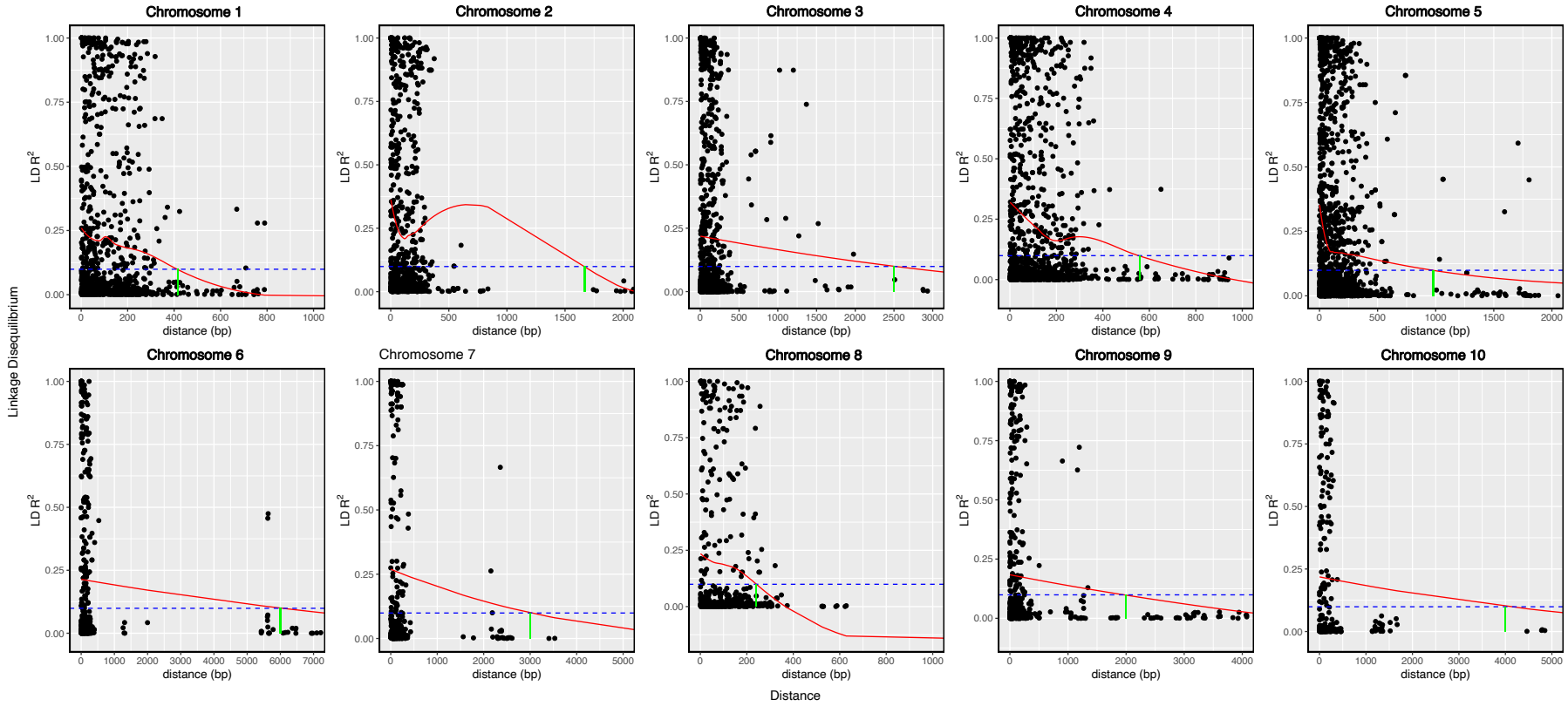


Figure 2. Linkage disequilibrium decay (R^2) and distance across the 10 *C. virginica* chromosomes with loess best fit. The critical R^2 was calculated to be 0.09 and the green line represents the point at which the loess curve intersects the critical R^2.

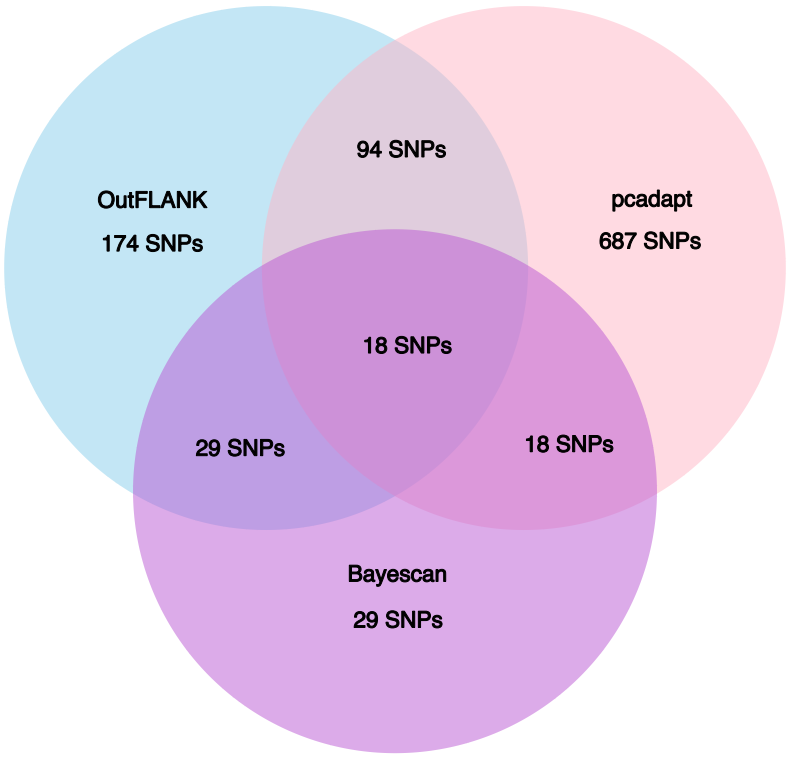


Figure 3. Venn diagram with number of SNPs identified as outliers by three methods: pcadapt, OutFLANK, and Bayescan prior to SNP thinning (full dataset).

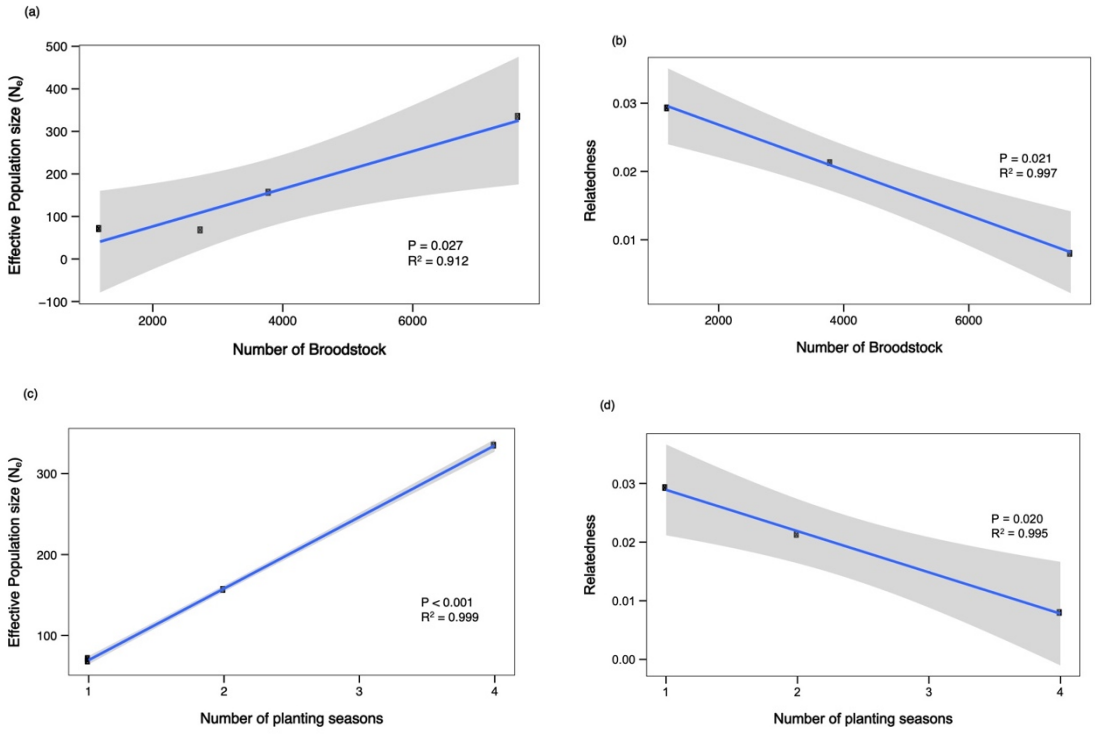


Figure 4. Effect of hatchery practices on metrics of genetic diversity in Harris Creek restored reefs. The effect of the number of broodstock on genetic diversity metrics of restored reefs including, a. effective population size (N_e) and b. relatedness of restored reefs. The effect of the number of hatchery planting seasons on genetic diversity metrics of restored reefs including c. N_e, d. relatedness, and e. observed heterozygosity. The effect of the f. broodstock male-to-female ratio in the hatchery on observed heterozygosity of restored
reefs.

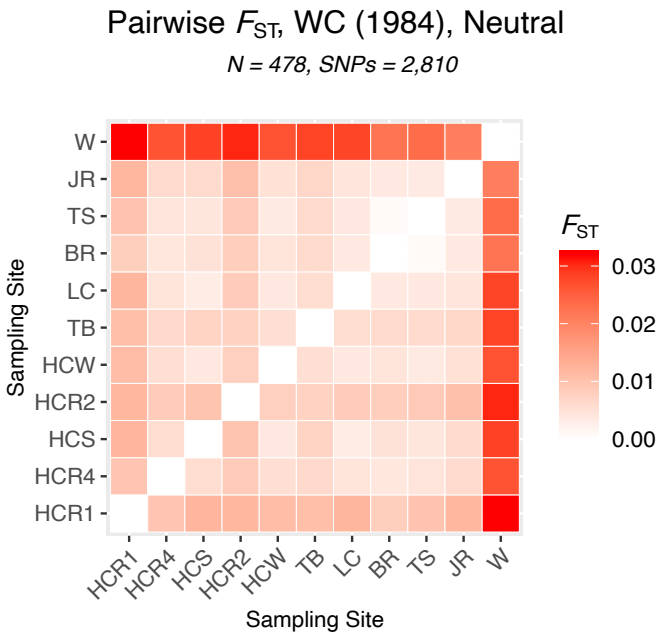


Figure 5. Heatmap of pairwise F_{ST} for *C. virginica* populations using the putatively neutral SNPs. Populations are ordered from north to south (from HCR1, Harris Creek 1 to W, Wachapreague). Abbreviations of sampling sites are presented in Table 1.

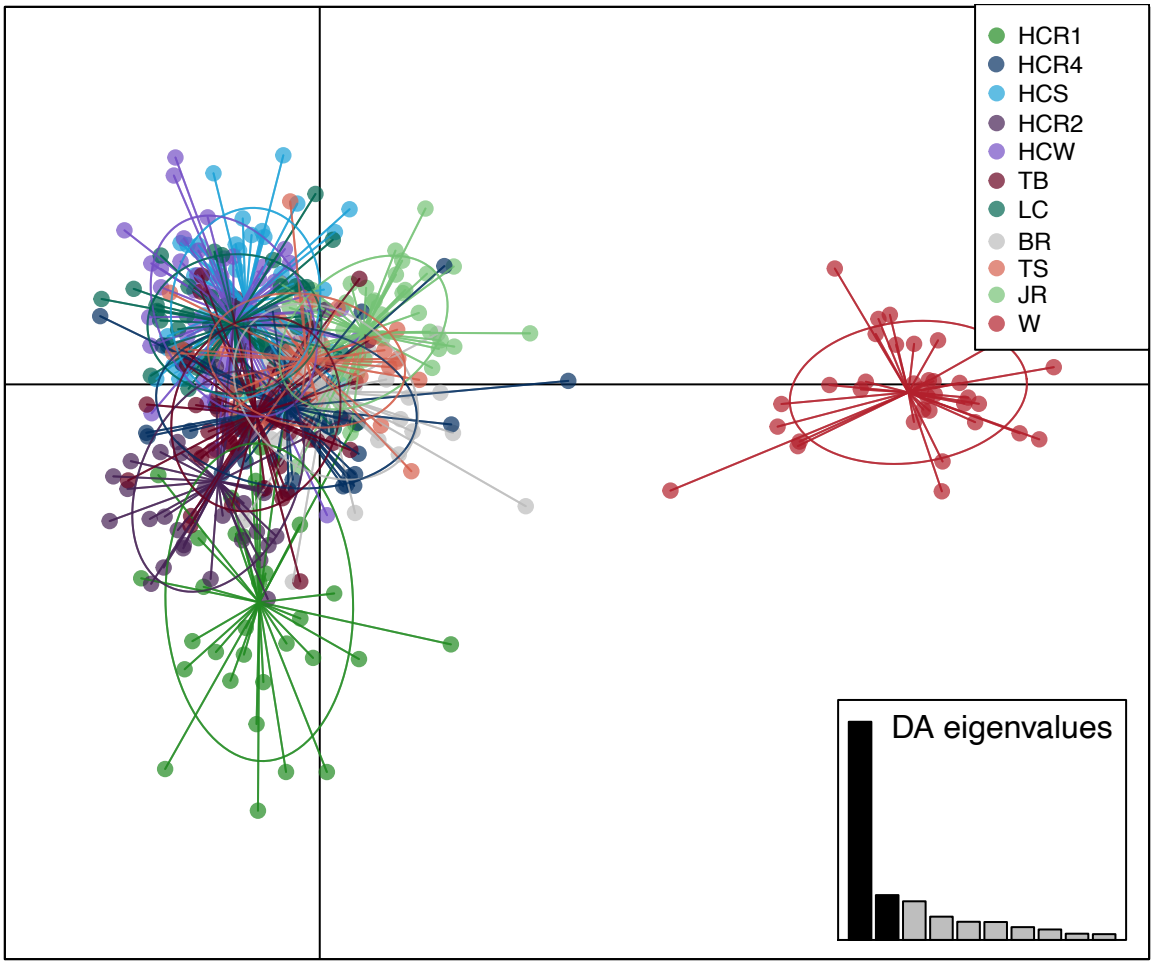


Figure 6. Discriminant analysis of principal components (DAPC) among wild and restored *C. virginica* populations based on 2,810 neutral unlinked SNPs. Abbreviations of sampling sites are presented in Table 1

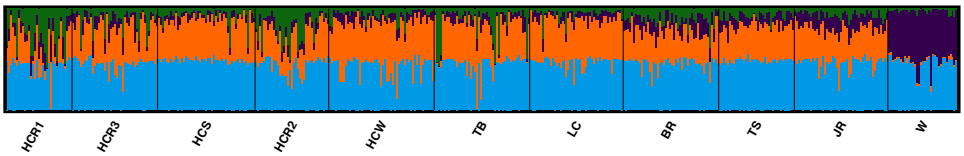


Figure 7. Plot of STRUCTURE results for 11 *C. virginica* populations using the neutral SNPs. Plots of individual admixture determined using the program STRUCTURE at the K recommended by the dK method (K=4).

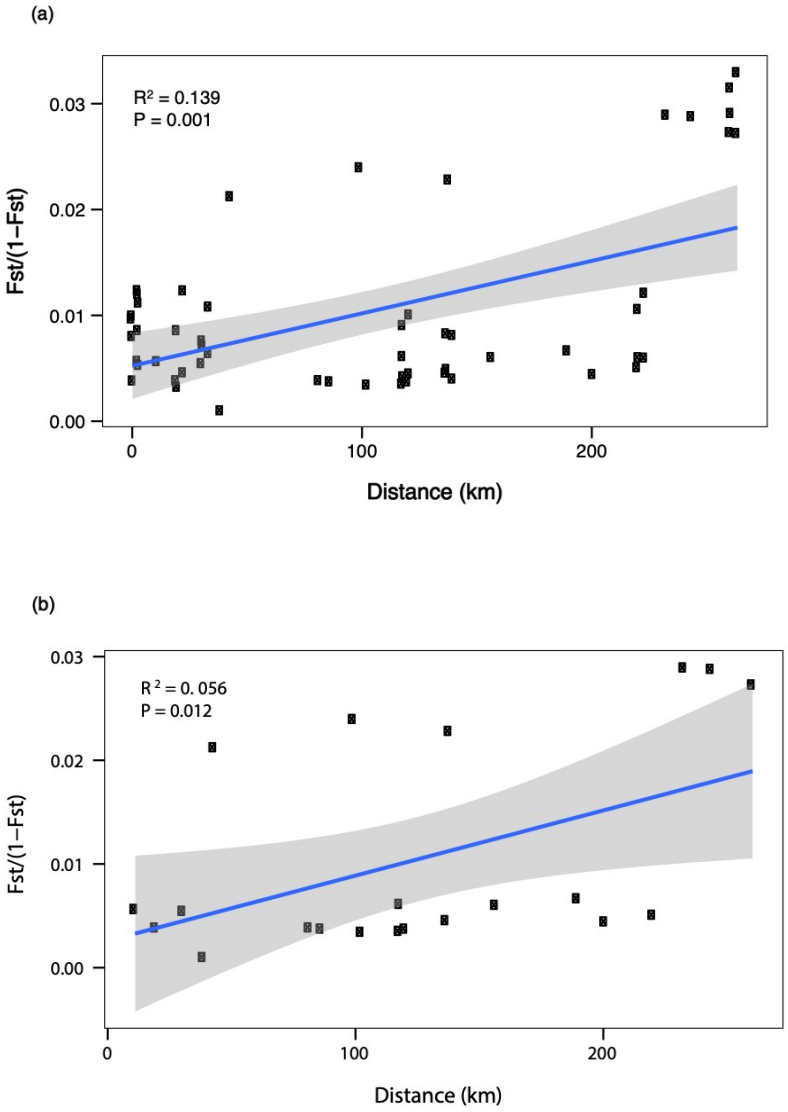


Figure 8. Isolation-by-distance (IBD) relationship between a. all *C. virginica* site pairs and b. only wild *C. virginica* pairs where linearized pairwise F_{ST} values (F_{ST}/(1- F_{ST})) are regressed over marine distance. Circles represent pairwise comparisons and the regression line (blue) is fitted with a 95% confidence limits (grey).

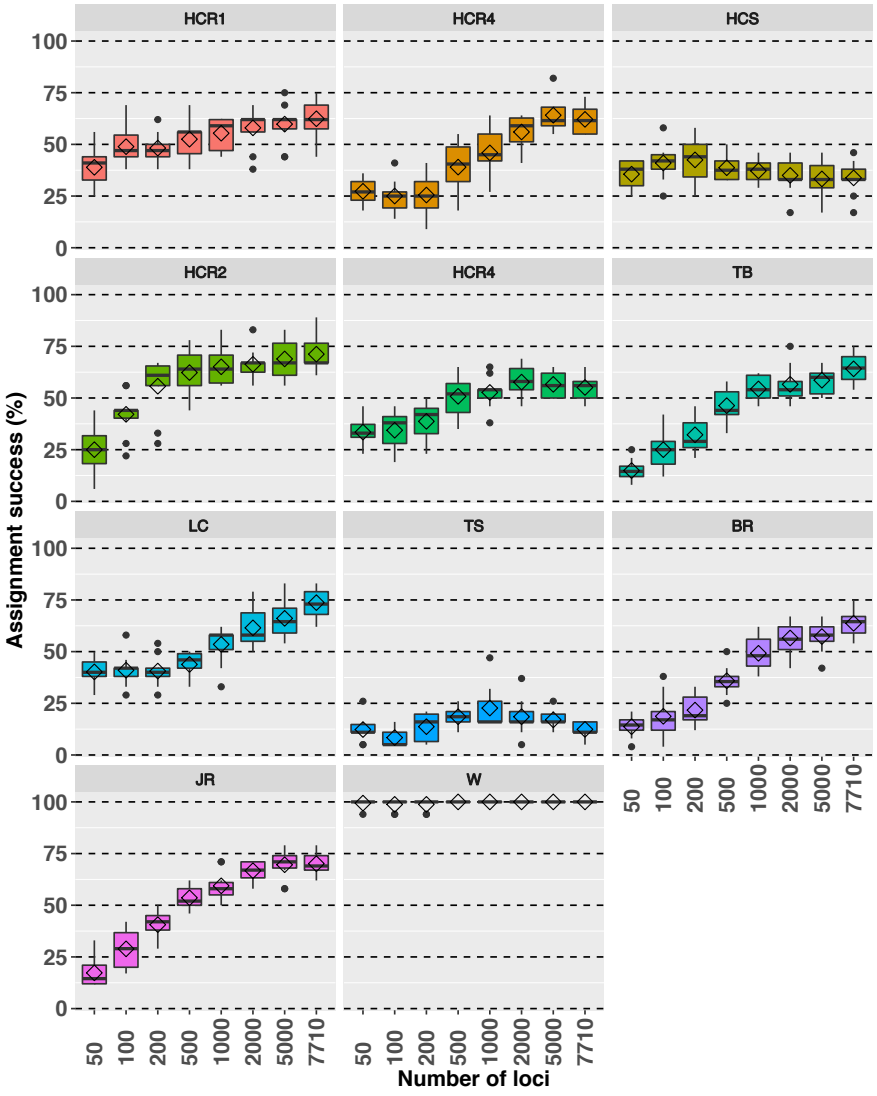


Figure 9. Assignment success of individuals to their sample of origin using the training, holdout, leave-one-out technique implemented by Assigner and GSI Sim. Number of loci used represents the highest F_{ST} markers identified in the training data set. Abbreviations of sampling sites are presented in Table 1.

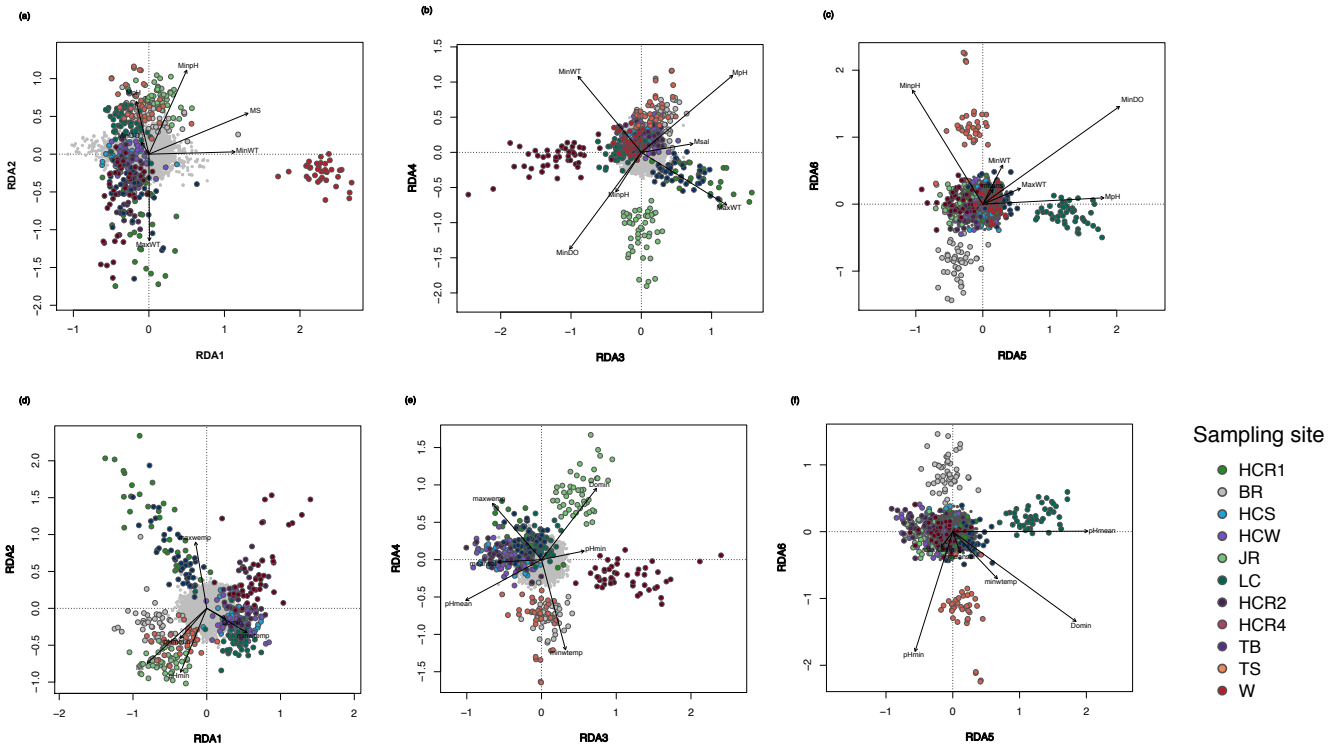


Figure 10. Redundancy analysis (RDA) for polygenic adaptation analyses performed using the 7,710 SNPs on significant axes. These include: A. axes 1 and 2, B. axes 3 and 4, and C. axes 5 and 6. Redundancy analysis (RDA) for polygenic adaptation analyses performed using the 7,710 SNPs excluding W sample on significant axes. These include: D. axes 1 and 2, E. axes 3 and 4, and F. axes 5 and 6. Arrows represent environmental variables (pHmean: mean pH, pHmin: minimum pH, meansal: mean salinity, minwtemp: minimum water temperature, maxwemp: maximum water temperature, and Domin: minimum dissolved oxygen). Large colored circles and small gray circles represent sampling sites and SNPs, respectively. Abbreviations of sampling sites are presented in Table 1.

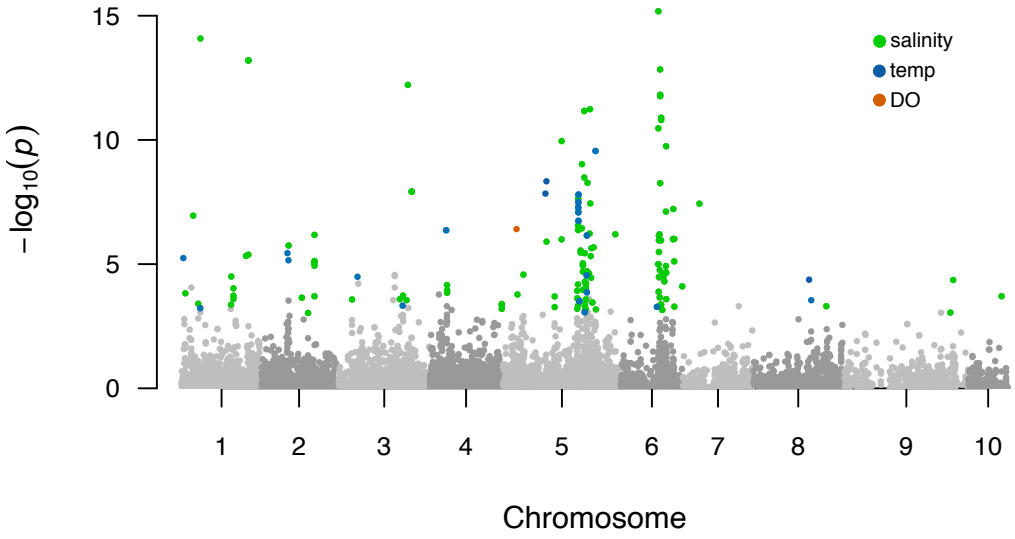
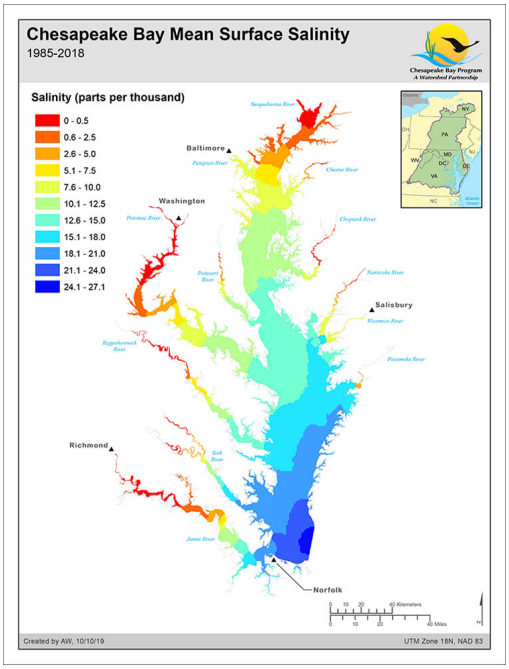


Figure 11. Manhattan plot showing p-values from RDA (all 6 significant axes) for 7,710 SNPs aligned by position on chromosomes 1-10. Colored dots correspond to 145 significant SNPs identified as outliers by RDA that were correlated with environmental parameters (salinity, temperature, and DO) prior to SNP thinning. Note significant clustering of significant RDA outliers on chromosomes 5 and 6.

A.



B.

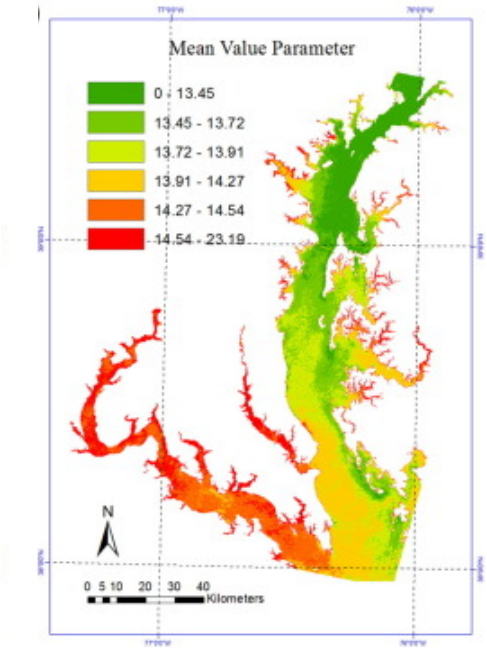


Figure 12. Chesapeake Bay mean A. surface salinity from 1985-2018 from the Chesapeake Bay Program and B. water temperature from 1984-2011 from Ding and Elmore (2015).

Chapter 4: Examining the genetic impact of hatchery-based oyster restoration using an individual-based model

Abstract

The strategic release of captive-bred organisms into the environment is one of the most popular methods to restore populations in forestry, fisheries, and wildlife management. However, concerns exist regarding genetic impacts on wild receiving populations over the long-term, especially for species with complex life-history features. We developed a forward-simulating, individual-based model (IBM) of oyster population genetics in the Chesapeake Bay using the simuPOP framework to examine the impacts of various restoration scenarios and hatchery practices on genetic diversity of simulated natural populations of eastern oysters. Simulations incorporated complex life-history features and demographic empirical genetic data from oysters in the Chesapeake Bay, including age structure, recruitment variability, variance in reproductive success, and polygamous mating, which, has not been incorporated into previous models. In addition, diagnostic model results indicated that high variance in reproductive success produced data consistent with previous experiments and expectations. We quantified changes in key metrics of genetic diversity (effective population size, N_e, heterozygosity, H_o, and allelic richness, A_r) in recipient populations resulting from a range of simulated restoration scenarios and hatchery practices including spawning practices (controlled or mass), broodstock numbers (N=10, 25, and 500), and broodstock practices (recycled or rotated). The impact of varying migration rates of the receiving population with two local wild populations on genetic diversity were examined (closed, 5%, and 10%; two-way

migration). Across simulations, N_e was the most sensitive metric to varying hatchery practices. Large broodstock numbers (N=500) resulted in a net increase (relative to control scenarios) in N_e and H_o of a receiving population compared to small and medium broodstock numbers across all scenarios. Spawning practices had a large effect on N_e, with controlled spawns resulting in the smallest changes in genetic diversity metrics throughout simulations. Overall, spawning practices and broodstock rotation has larger effects on diversity, suggesting that programs that are broodstock limited may still be able to maintain diversity in their programs by altering other practices. The IBM from this chapter is the first to incorporate complex life-history features of shellfish species in a fully flexible modeling framework and thus it may be useful for evaluating future supplementation strategies for oysters and for investigating restoration impacts of other marine species with similar life-history features.

Introduction

The strategic release of captive-bred organisms is one of the most popular methods to restore populations in forestry, fisheries, and wildlife management (Laikre et al. 2010) and recent predictions suggest thousands of additional species may soon require such measures to prevent extinction (Seddon et al. 2005; Mendelson et al. 2006; Attard et al. 2016). In the marine environment, the release of hatchery-produced stocks (hereafter "supplementation") has become an important fisheries management strategy used for restocking, stock enhancement, and ranching (e.g. see Bell et al. 2008 for definitions; Laikre et al., 2010). Programs exist for hundreds of species around the world (Bartley

and Leber 2004; Lorenzen et al. 2013). While supplementation programs contribute to the maintenance of economically, socially, and ecologically important fish populations (Araki et al. 2008), numerous studies have documented potentially adverse impacts on the genetic integrity and evolutionary potential of wild populations (Ryman and Laikre 1991; Araki et al. 2008; Fraser 2008; Christie et al. 2012a, b). These negative genetic impacts are often at odds with conservation, restoration, and fisheries management goals, and there is great interest in increasing monitoring and improving the ability to predict short- and long-term genetic and evolutionary impacts of supplementation programs.

There are two key areas of potential genetic risks associated with supplementation that affect the short- and long-term viability of wild populations. First, adaptation to captivity (i.e. domestication selection) can lead to reduced fitness in hatchery produced individuals and adaptive mismatches with local environmental conditions (Araki and Schmid 2010; Waal et al. 2013; Christie et al. 2014). Second, releasing large numbers of hatchery-produced individuals from small numbers of broodstock can lead to reductions in genetic diversity and effective population size (N_e, the evolutionary analog to census population size; (Ryman and Laikre 1991; Ryman et al. 1995; Laikre et al. 2010). The retention of genetic diversity is important to the long-term evolutionary potential of a population as reduced genetic diversity has been associated with an increased risk of population extinction (Saccheri et al. 1998), reduced population growth (Hanski and Saccheri 2006), and reduced potential for response to environmental change (Waples 1991). In some cases, systematic genetic monitoring has been undertaken to better characterize the impacts of captive breeding on genetic diversity and to optimize conservation actions (e.g. La Haye et al. 2017). While advancements in genomic methods

have allowed for monitoring the genetic diversity of supplemented populations, diagnostic tools assessing how changes in hatchery practices impact genetic diversity of supplemented populations are scarce.

Restoration of marine shellfish populations primarily through juvenile seeding (Gaffney 2006; Laing et al. 2006) has increased due to the worldwide decline of ecologically, economically, and culturally significant shellfish species (Beck et al. 2011). The high ecological and economic value of eastern oysters *Crassostrea virginica* has prompted wide-ranging restoration efforts across its native range (Damiano and Wilberg 2019),with restoration approaches often including hatchery production and planting of spat on shell (juvenile oysters). In the Chesapeake Bay, Maryland, a federal mandate to restore 20 tributaries by 2025 has provided support for large-scale hatchery-based restoration in the Choptank River tributary region. The first of four sub-tributary sanctuaries, Harris Creek, was completed in 2016 (Westby et al. 2017). The University of Maryland Center for Environmental Science's (UMCES) Horn Point Laboratory (HPL) Oyster Hatchery produces spat on shell for oyster sanctuaries through mass-spawning of local, wild, rotated broodstock (following recommendations by the Nature Conservancy, Brumbaugh et al. 2006). Initial characterization of the genetic impacts of this program uncovered expected decreases in genetic diversity of hatchery-produced cohorts (i.e. spat were less diverse than parents), but genotyping of the supplemented reefs revealed that genetic diversity was maintained relative to surrounding natural reefs (Hornick and Plough 2019, Chapter 3). In addition, the level of genetic diversity at supplemented reefs was significantly and positively associated with specific hatchery practices, including the number of broodstock used for spawning and the frequency a reef was supplemented

(Chapter 3), which may provide relatively simple guidelines for increasing or maintaining genetic diversity in these restoration programs. While this initial work investigated, retrospectively, how different spawning and planting practices affected genetic diversity of hatchery-produced cohorts of oysters over 5 years, it is impossible to vary different hatchery practices in concert (e.g. broodstock management, spawning strategies) to evaluate long-term impacts on genetic diversity of natural populations. Long-term ecological and genetic impacts of supplementation are often difficult to monitor as changes in genetic diversity often manifest slowly (reviewed in Frankham et al. 2010), and some programs have only existed for a short period of time. Moreover, genetic monitoring can be expensive and time consuming (Fussi et al. 2016) and resources are rarely available for genetic analysis or monitoring after the restoration has been completed. As hatchery-based supplementation of oysters is increasingly utilized as a component of restoration and fisheries management, novel approaches integrating species biology, genetics, and management decisions are needed to evaluate the long-term genetic impacts and to refine hatchery programs to maximize retention of diversity.

Computer simulations are increasingly used to model the complex evolutionary dynamics of populations over space and time, incorporating known demographic and genetic information from real populations to determine how genetic composition, diversity, and even adaptive traits may change under various scenarios including hatchery-based supplementation (e.g. Strand 2002; Hoban et al. 2012). Individual based models (IBMs) are one such approach and consider each individual of a population as an independent entity and track the events (e.g. birth, death, growth, and development) that occur within the simulation for each individual. IBMs thus provide tremendous flexibility

in their application to a variety of ecological and evolutionary questions. For example, IBMs have long been applied to human populations, including studying the evolution of complex human diseases (Peng et al. 2007) and the influence of human movements on local-scale malaria transmission (Pizzitutti et al. 2018). Recently, studies have applied IBMs to population viability analysis for a large variety of taxa including birds (Letcher et al. 1998), mammals (Yiming et al. 2003), and insects (Griebeler and Seitz 2002) as well as for the reintroduction of extinct species (Kramer-Schadt et al. 2004). Recent advances in IBM approaches provide approaches to evaluate demographic and genetic responses of marine species to multiple interacting factors (population genetic IBMs). For example, population genetic IBMs have been used to examine the impacts of cultured fish on genetic diversity and fitness in wild populations (e.g. Waples and Do 1994; Oota and Matsuishi 2005; Darden et al. 2017; Mims et al. 2019; Katalinas et al. 2019). Because these genetic IBMs provide an opportunity to simultaneously consider demographic and genetic effects, they provide a powerful tool to assist management of captive breeding and supplementation programs in the face of epistemic uncertainty and complex conservation management decisions (Balkenhol and Landguth 2011). However, despite previous attempts to model the genetics of hatchery-based supplementation (e.g. Oota and Matsuishi 2005; Katalinas et al. 2019), none have adequately incorporated complex life-history features of marine species such as age structure, recruitment variability, variance in reproductive success, and polygamous mating.

In this chapter, an individual-based model (IBM) was developed in simuPOP (Peng and Kimmel 2005) to examine the genetic impacts of various hatchery-based restoration strategies, with the goal of informing future restoration planning of shellfish

and species with complex life-history features. simuPOP uses a flexible scripting language in Python to allow operators to control complex demographic features (e.g. polygamous mating, variance in reproductive success, and recruitment variability) from the species under study, and is one of the few simulators to allow such options (Hoban et al. 2012). We utilized life-history, demographic, and empirical genetic data from eastern oysters in the Chesapeake Bay (Hornick and Plough 2019; Ch. 3; Damiano and Wilberg 2019) to simulate natural conditions and examine changes in the genetic diversity of a wild receiving population under several restoration scenarios and validate against empirical data. The IBM allowed us to track genetic diversity changes in a supplemented population resulting from various hatchery management strategies and afforded the opportunity to do large-scale experiments with simulated populations, which could not be achieved empirically. The goal of this work is to determine how major, tangible (i.e. changeable) hatchery activities and husbandry practices can affect the maintenance of diversity in wild populations. More broadly, we evaluate the utility of this IBM in providing unique and complementary insights (relative to results from empirical studies) on how hatchery-based supplementation of marine species impacts genetic diversity and population resiliency over time.

Methods

Individual-based model in simuPOP

Forward-time genetic simulations in simuPOP v.1.1.9 (Peng and Kimmel 2005; Peng and Amos 2008) were conducted. First diagnostic models were developed that

tested and incorporated key demographic and life-history features of eastern oysters (described below). Once the model was developed and validated against empirical data and theoretical expectations for oysters, life-history features were incorporated into full simulations. Then simulations were run to examine the genetic impacts of hatchery-based supplementation on natural populations under a variety of scenarios.

Basic framework of the model and demographic parameters

Input values representative of Chesapeake Bay eastern oyster populations were used to initialize simulation models (Table 1), including observed allele frequencies, age-structure, age-specific survival rates, migration rates, and variance in reproductive success. To simulate the life history and historical conditions in Harris Creek, the model contained three open subpopulations: Harris Creek (HC) (the wild recipient population), Chesapeake Bay (CB) (representing a local population that could interact via dispersal with the restored population), and the broodstock population (BR) (wild population from which individuals were brought into the hatchery to spawn). Individuals in each of the three subpopulations were randomly assigned sex and age (1-11) at initialization. Migration from two other subpopulations was modeled because supplemented oyster populations may be part of a metapopulation where migration occurs among subpopulations. Model parameters, notations, and parameter values and sources or equations used to estimate parameter values can be found in Table 1. The basic framework of the model is detailed in Figure 1.

For model simulations, one time-step equals one year and one reproductive cycle (individuals aged 2–11 reproduce, see below) and generation time is two years (i.e.

individuals are reproductively mature at age-2; conservative generation time based on Burkenroad 1931; Coe 1936; Hayes and Menzel 1981; O'Beirn et al. 1996). Age-0 individuals were assumed to be spat to remove computation intensity required to model/simulate billions of larvae, of which only a small percentage make it to the spat stage (Hedgecock 1994; Hedgecock and Pudovkin 2011). During each time-step (reproductive cycle), two-way migration between each subpopulation occurred (15% rate; randomly choose individuals from the existing), and individuals transitioned to the next age-class according to age-specific survival probabilities (Table 1), while individuals greater than 11 years old were removed from the population (i.e. died). While eastern oysters have been hypothesized to live up to 20 years (Galtsoff 1964) the onset of two diseases, MSX and Dermo have reduced the adult life span significantly in a salinity-dependent manner (Powell et al. 1996; Paraso et al. 1999; Calvo et al. 2001; Harding et al. 2008). In addition, *C. virginica* from plantings in Maryland have been reported to survive at least 9 years (Paynter et al. 2010), so the maximum age for simulations was assumed to be 11 years. Increasing the maximum age to 15 did not lead to changes in core genetic parameters (data not shown). Survival rates for spat (age 0) matched values from Harris Creek, estimated in Damiano and Wilberg (2019). Age-specific survival rates are not available for eastern oysters (only size based estimates available; Doering 2019; Damiano and Wilberg 2019), so small oysters were assumed to represent age 1, and market-sized oysters (3"– 4 ½") represented age 3+ (Paynter et al. 2010). Variability in mortality for a given age and each year exists in this region (Damiano and Wilberg 2019), so natural mortality rates were averaged for market-sized oysters from 1989-2015 and

this value was applied as a static survival rate for age 3+ oysters. The mortality rate for age 2 oysters was the average of the spat and market survival rates.

The simulated populations were smaller than the actual population sizes in Harris Creek and in other Chesapeake Bay oyster populations due to computational constraints. Starting abundances of 100,000 were used, which allowed for the examination of restoration combinations without compromising the number of simulation replicates that were carried out in order to decrease variation around the mean. This large population size is reasonably reflective of a moderate sized tributary population that is not overwhelmingly influenced by genetic drift. Increasing the population size to one million did not lead to changes in the core parameters or basic model framework (data not shown).

Model diagnostics and comparisons with field data

Once constructed, the model was then tuned to match empirical (observed) distributions of oyster reproductive success and recruitment, as well as current levels of genetic diversity and gene flow, measured as F_{ST} (described in detail below). Specifically, parameters were incorporated in the model to match data on the abundance of individuals in Harris Creek, recruitment dynamics within Harris Creek, and natural mortality estimates for different age-classes of oysters within the region (Damiano and Wilberg 2019). In addition, the model incorporated high variance in reproductive success among parents, and the resulting N_e was compared with empirical estimates from Chapter 3. Finally, model parameters such as migration rates were tuned to reflect empirical F_{ST}

values, and diversity levels were tuned to reflect empirical estimates of observed heterozygosity and allelic richness in these populations.

Mating in the model

Mating occurred each time-step (year) by generating offspring via polygamous mating of parents from reproductively mature age-classes (2–11) using modules available within simuPOP. Offspring were created from a geometric distribution to reflect high variance in reproductive success among parents (high variance in offspring number produced per parent, V_k). This distribution allowed for a small adult proportion to successfully spawn and contribute to offspring within each subpopulation to account for 'sweepstake reproductive success' (SRS) that is hypothesized to occur in marine species with high fecundities and type-III survivorship (Hedgecock 1994; Hedgecock and Pudovkin 2011; Hoban et al. 2013). Mating was polygamous for both males and females, and the poly number (number of mates per sex) for each sex varied each generation.

Since the expected distribution of reproductive success for wild oysters is unknown, data on distribution of reproductive success was taken from other marine species and variables were adjusted to match empirically measured N_e values for oysters. Empirical data on the distribution of reproductive success in brown trout Serbezov et al. (2012) followed an approximate gamma distribution, which reflects high variance in reproductive success (V_k). In addition, a recent simulation study used the geometric distribution to model species with high V_k (Hoban et al. 2013). Therefore, reproductive events followed a geometric distribution which produced high V_k in simulated populations for which observed N_e values were approximately similar to estimates from

natural populations in Harris Creek using high-resolution SNP panels (Chapter 3). During the model development and validation process, Poisson and binomial distributions were also explored to model offspring distribution but were ultimately ruled out because simulated V_k and resultant N_e estimates did not closely match empirical estimates from field and lab experiments (Chapters 2 and 3).

Recruitment and migration of individuals

The number of offspring produced (i.e. recruits) followed a gamma distribution to reflect variation in recruitment in this region (Damiano and Wilberg 2019, Figure 2). Recruitment was compared with empirical data from Damiano and Wilberg (2019) (using only data from 1989–2006 as this represents data unimpacted by supplementation activities) (Figure 2). Recruitment in model simulations followed a gamma distribution that depended on the size of the reproductively mature population (i.e. mass of the distribution shifted with the size of the population using the NumPy package in python v.3.7.3):

Recruits = gamma (shape=1, scale=size of reproductively mature population).

After the recruitment step, migration of individuals among subpopulations occurred to mimic recruitment during the larval stage only (i.e. dispersal). Mixing between the populations occurred through the two-way migration of individuals at a rate that matched empirical genetic differentiation estimates (F_{ST}) between populations.

Genetic parameters of the model

Two hundred-fifty neutral, independent bi-allelic loci evolving under a strict infinite sites model with a mutation rate $u=1x10^{-8}$ were modeled to track changes in diversity and effective size during simulations. Preliminary simulations were conducted with 100 up to 1,000 markers and the precision and accuracy of N_e estimates appeared to plateau at around 250 markers, so this was chosen as the number of markers to include to reduce computational time.

Pre-scenario simulation burn-in

Simulations incorporated a 100-year (50 generations) burn-in period to establish pre-restoration conditions. Mutation-drift equilibrium was obtained by running the three wild subpopulations for 100 generations (i.e. the burn-in), starting from individuals with a starting allele frequency corresponding to a starting H_e of ~0.223, which approximates the observed H_e in polymorphic RADSeq SNPs (Ch. 3). After verifying that the three subpopulations had reached stable equilibrium, confirmed by the convergence of the effective population size (N_e), observed heterozygosity (H_o) and allelic richness (A_r), hatchery simulations began.

Framework for hatchery-based restoration scenarios

For scenarios including hatchery-based supplementation, reproductively mature individuals were captured from the broodstock population and brought into the hatchery to spawn. Hatchery offspring were produced by either mass- (polygamous mating) or controlled- (i.e. pair matings) spawns of reproductively mature (ages 2-11) hatchery

broodstock, depending on the scenario. Following spawning, hatchery broodstock were either returned to the wild broodstock population (rotated) or were used again in the hatchery (recycled), depending on the scenario. Each year, a specific number of individuals produced in the hatchery were introduced to the wild Harris Creek population based on numbers collected by the HPL Oyster Hatchery. The movement of broodstock into the hatchery and recently settled spat produced by the hatchery population to the Harris Creek population (i.e. the planting of spat) was simulated as migration events. Hatchery-based supplementation plantings occurred every year for 8 years (the length of time Harris Creek has been supplemented) and were followed by a 50-year recovery period to examine any residual effects. Scenarios were compared to control runs where a hatchery restoration program was never initiated.

Genetic diversity metrics measured during simulations

Genetic diversity metrics were estimated for the 250 markers based on sampling 10,000 individuals (>1% of the population Marandel et al., 2019) excluding the spat age class, to mimic typical field sampling of juvenile or adult oysters. In the model, genetic parameter estimates were made at the conclusion of the model burn-in period to define initial conditions and then at each generation (two years) throughout the simulation period in both control scenarios (no hatchery supplementation) and scenarios including hatchery supplementation. Genetic diversity metrics commonly estimated and determined to have conservation and evolutionary significance were tracked (Crow and Kimura 1970; Hare et al. 2011; Allendorf et al. 2013). These metrics included allelic richness (A_r), observed heterozygosity (H_o), and effective population size (N_e), and were

calculated using the r package hierfstst v.0.04-22 (Goudet 2005). Effective population size (N_e) was estimated using NeEstimator V2 (Do et al. 2014) and corrected for overlapping generations as in Hornick and Plough (2019). The jackknife method was used to estimate confidence intervals (CIs) across loci (Jones et al. 2016), and the lowest allele frequency was set at 0.02 to further minimize bias due to rare alleles (Waples and Do 2010).

Model scenarios of oyster hatchery supplementation

A total of 27 simulation scenarios were run, with 10 replicates each (see Table 2). Simulation scenarios examined the genetic impact of spawning practices (controlled vs. mass) using different numbers of broodstock (N=10, 25, and 500). The impact of mass-spawning on genetic diversity of a supplemented population was chosen because it is an extremely common practice used to produce spat in shellfish restoration hatcheries (e.g., in Maryland, HPL Oyster Hatchery; in Martha's Vineyard, Massachusetts, Emma Green-Beach Martha's Vineyard Shellfish Group Inc, personal communication; and in New York Harbor, Rebecca Resner, Billion Oyster Project, personal communication). For mass-spawn scenarios, the impact of rotating and recycling broodstock was investigated. For control-spawn scenarios, broodstock had to be rotated because paired matings are produced by strip-spawning, which is lethal to oysters (i.e. can't be recycled in the subsequent year). A second set of scenarios were run that examined the impact of varying migration scenarios (closed, medium, open). These scenarios only manipulated migration rates within the context of mass-spawns with recycled and rotated broodstock (N=10, 25, and 500 broodstock for each). The migration rates (two-way migration among the three

subpopulations) that were varied for these scenarios included closed (no migration), medium (5% migration), and open (15% migration). All hatchery-based restoration scenarios are shown in Table 2. For all model scenarios, genetic diversity metrics were calculated each generation and compared as percent change relative to control runs where hatchery-based supplementation was never initiated. Additional simulations were run: one long-term supplementation period (25 years, 12.5 generations) and one simulation was run for a 200-year (100 generations) post-supplementation period based on the most extreme supplementation scenario (recycled broodstock 10). There were no changes throughout this extended recovery period in any genetic diversity metric, indicating that any residual effects, if present, were seen within the 50-year recovery period.

Sensitivity analyses

Sensitivity analyses were conducted to understand the extent to which the model results might vary given uncertainties in certain inputs and parameters. Field data was lacking for certain reproductive and early life-history features, including the poly number (number of mates each oyster has in a given spawn), the number of males and females for each mating event, and the variance in reproductive success and the relative contribution of parents. A simple sensitivity analysis was performed in which the poly number (the number of females each male mates with during each mating event and vice versa) and the geometric distribution (to vary the number of offspring assigned to each parent for a reproductive event, Table 3) were varied. The number of offspring for each mating event using a geometric distribution follows mean $1/p$ and variance $(1-p)/p^2$, so p was varied in sensitivity analyses. (Table 3). The poly number was changed by $\pm$10% of a

proportion of the population size (Table 3). The relative contribution of offspring was modeled with five different values of *p* of the geometric distribution, varied by ±10% (Table 3, 5 times for each, 5 replicates) and the sensitivity of population genetic diversity estimates was examined. Sensitivity was measured as the percent change from the base case in the model (Geometric $p = 0.002$, Poly number = 2 to 0.003*popsize; Table 3) in which the parameters assumed their nominal values (similar approach described by Saltelli et al. 2000; Cacuci 2003). In addition, extreme values were run to investigate how they affected genetic metrics in the model.

Results

Model diagnostics and comparisons with field data

Modeling multiple distributions of offspring numbers and reproductive contribution of breeding-age individuals showed that the geometric distribution produced data most similar to empirical N_e estimates for oysters in the Chesapeake Bay, empirical distributions of reproductive success from Serbezov et al. (2012), and simulated distributions of reproductive success with high V_k from Hoban et al. (2012; Figure 3). Secondly, manipulating the scale and shape of the recruitment function within simulations produced population trends similar to what was observed for recruitment and population size by Damiano and Wilberg (2019; Figure 2 and Figure 4).

Simulated levels of genetic diversity matched empirical estimates based on pre-existing genomic data (Figure 5). Empirical estimates of N_e based on mixed-age adult samples fell within 95th percentiles of simulated values in all generations (Figure 5a).

Empirical estimates of A_r fell within the 95th percentiles of simulated values for the first half of simulations, but not for the second half (Figure 5b). However, by the end of the model run, the simulated value of A_r (1.922) was within 3.4% of the empirical estimate (1.989). Estimates of H_o from simulations were slightly higher than empirical estimates throughout the validation period but fell within the 95th percentiles of the empirical estimate by the end of the validation period (Figure 5c). Pairwise F_{ST} among populations throughout the validation period fell within the 95th percentiles of empirical F_{ST} estimates (Figure 5d–f). Despite variable marker and sample size in empirical studies, close agreement between simulated and empirical estimates for genetic diversity metrics and life-history features indicate that the IBM can accurately simulate realistic values of genetic diversity observed in wild oyster populations.

Model scenarios

Twenty-seven different scenarios of hatchery-based restoration were simulated to investigate changes in genetic diversity of a receiving population of oysters. These scenarios can be grouped into three general categories: mass spawned simulations, controlled spawn simulations, and migration rate simulations. And within those categories, the number of broodstock were varied (N=10, 25, 500) and the broodstock management was varied (recycled or rotated broodstock; see Table 2). In Figure 6, The raw model output for an example scenario run is shown (mass spawned, recycled broodstock, N=10 broodstock), which is indicative of the general trends for population size and diversity metrics across many restoration scenarios relative to a control scenario (no supplementation). In general, decreases in N_e occurred during the beginning and

middle of the supplementation period, with the most dramatic decreases occurring in the Harris Creek population (supplemented; Figure 6b). However, during the post-supplementation period, N_e generally increased and then stabilized, but was still higher relative to control scenarios towards the end of the simulation period for all populations (Figure 6b and see Figures 7a, 8a and b and text below for more complete results for each specific scenario). Increases in N_e were coupled with increases in population size in all populations, especially within Harris Creek during the supplementation period (Figure 6). Abundances of all subpopulations stabilized during the end of the supplementation period (Figure 6d). This pattern in N_e was consistent across all scenarios except for those with large broodstock number (N=500) where the trend was different, and N_e increased relative to the control during the supplementation period.

Controlled-spawn scenarios

For the controlled-spawn scenarios, changes in genetic diversity metrics were influenced by the number of broodstock (Table 4, Figure 7). Small numbers of broodstock (N=10) was the only scenario that resulted in an initial decrease in N_e relative to the control (31.3%). All other broodstock sizes (N=25, and 500) resulted in net increases in N_e throughout the simulation period with greater increases with larger broodstock number (Table 4). Following supplementation (post-supplementation early), N_e increased then leveled off by the end of the simulation period (Figure 7a). Allelic richness decreased slightly throughout the supplementation period, with greater decreases occurring in smaller broodstock numbers (N=10 and 25). By the end of the simulation period, A_r increased in all controlled-spawn scenarios relative to control scenarios, with a

stepwise increase associated with broodstock number (Figure 7b). In all controlled spawn scenarios, A_r decreased slightly during the supplementation period, with the relative decrease following similar trends as N_e and H_o related to the number of broodstock (Figure 7b). Slight decreases in H_o were observed during the supplementation period but eventually H_o increased relative to the control simulations across all controlled-spawn scenarios by the end of the simulation period (Figure 7c).

Mass-spawn scenarios

For the mass-spawn scenarios, changes in genetic diversity metrics were influenced by the number of broodstock and broodstock practice (recycled vs. rotated) (Table 4, Figure 7). For recycled broodstock scenarios, decreases in N_e were observed throughout the supplementation period for small and medium broodstock (N=10 and 25). Eventually, N_e increased relative to control scenarios by the end of the simulation period, with greater increases with larger broodstock number (Table 4). Similar trends were found for N_e in rotated broodstock scenarios. Small broodstock (N=10) was the only scenario that resulted in an initial decrease in N_e relative to the control (19.2%). All other broodstock sizes (N=25 and 500) resulted in net increases in N_e throughout the simulation period with greater increases with larger broodstock number (Table 4). Following supplementation (early post-supplementation, generation 9, Table 4), N_e increased then leveled off by the end of the simulation period. For recycled broodstock scenarios, A_r decreased slightly throughout the supplementation and post-supplementation early period, with greater decreases occurring in smaller broodstock numbers (N=10 and 25; Figure 7b). By the end of the simulation period, A_r increased in all recycled mass-spawn

scenarios relative to control scenarios, except in those with small broodstock (N=10, -2.36% decrease, Table 4). For rotated broodstock scenarios, A_r decreased slightly throughout the supplementation period, with greater in scenarios with smaller broodstock numbers (N=10 and 25). By the end of the simulation period, A_r increased in all rotated mass-spawn scenarios relative to control scenarios (Table 4, Figure 7b). For both recycled and rotated broodstock scenarios, there were slight decreases in H_o during the supplementation period but eventually H_o increased relative to the control by the end of the simulation period (Table 4; Figure 7b).

Varying migration scenarios

For scenarios varying migration rates, changes in diversity were influenced by the number of broodstock and migration rate, but the changes associated with varying migration rates were rather small (Figure 8). For recycled broodstock scenarios, increases in N_e were observed throughout the supplementation period only for large broodstock (N=500) across all migration scenarios, with the largest increases in the open scenarios (152.8% increase, med=145.9% increase, open=114.3% increase; Figure 8a). N_e increased for all broodstock numbers relative to control scenarios by the end of the simulation period (Figure 8a). For rotated broodstock scenarios, there were decreases in N_e throughout the supplementation period only for small broodstock (N=10) across all migration scenarios with the largest decreases in the closed scenarios (28.5% decrease, med=23.9% decrease, open=17.4% decrease; Figure 8b). Eventually N_e increased relative to control scenarios by the end of the simulation period (Figure 8b). For recycled broodstock scenarios, A_r decreased slightly throughout the supplementation and post-

supplementation period, with greater decreases occurring in smaller broodstock numbers (N=10 and 25) and less migration (closed and medium Figure 8c). By the end of the simulation period, A_r increased in only recycled scenarios with large broodstock numbers (N=500) and the open scenarios with medium broodstock (N=25) relative to control scenarios. For rotated broodstock scenarios, A_r decreased slightly throughout the supplementation and post-supplementation period, with greater decreases in scenarios with smaller broodstock numbers (N=10 and 25) than those with larger broodstock numbers (N=500) (Figure 8d). By the end of the simulation period, A_r increased in all rotated scenarios relative to control scenarios (Figure 8d). For both recycled and rotated broodstock scenarios, there were slight decreases in H_o during the supplementation period, but H_o increased relative to the control by the end of the simulation period across all scenarios except recycled broodstock closed (N=10) (Figures 8e and f). Overall, the pace of the recovery from reduced genetic diversity was faster in scenarios that included migration (Figure 8).

Sensitivity analyses

Sensitivity analyses were conducted to understand how model output and results might vary given the uncertainties in certain inputs and parameters. Analyses focused specifically on poly number and variations in the distribution of reproductive success among parents modeled as a geometric distribution. Within the model framework, the geometric distribution is used to create the distribution of offspring from each parent, and the poly number determines the number of males that mate with females and vice versa for each mating event.

In general, we found that genetic diversity metrics were most sensitive to variation in the geometric distribution used to model reproductive success (see Table 3). Across sensitivity analyses, the most sensitive genetic diversity metric was N_e. As the *p* parameter increased in the geometric distribution, the sensitivity of N_e increased with the largest increases occurring at the highest value ($p = 0.2$). For the poly number sensitivity analyses, as the poly number increased, N_e decreased, but the changes were not as dramatic as those for the geometric distribution. In general, genetic diversity metrics were less sensitive to changes in the poly number except when extreme values were implemented (Figure 9). The largest relative change in N_e from varying the poly number (Figure 9b) occurred at the beginning of simulations (year 10) in the smallest treatment (68% decrease relative to control). Across all sensitivity analyses, sensitivity of H_o generally increased throughout the sensitivity simulation period (Figure 9e and f).

Discussion

In this chapter, an individual-based genetic model was developed that incorporated the complex demographic and life-history features of oysters in the Chesapeake Bay to examine changes in genetic diversity associated with varying hatchery supplementation strategies. Overall, effective population size (N_e) was the only metric to exceed a 10% change compared to the control or base scenario, and was much more sensitive to varying broodstock practices (recycled vs. rotated) and number (N=10, 25, and 500) as well as the spawning type (mass vs. controlled) compared to observed heterozygosity (H_o) and allelic richness (A_r). In addition, while simulations with large

broodstock numbers (N=500) resulted in relative increases in diversity metrics, simulations with medium broodstock numbers (N=25) that were rotated also resulted in increases across most diversity metrics. This result is encouraging and may suggest that the use of large broodstock numbers is not the only way to increase or maintain diversity in a restoration program (at least based on this model). Interestingly, the amount of migration or connectivity did not have a major impact on the genetic diversity of a proximal wild population simulated in the model. However, the rate of migration from local populations mitigated genetic diversity losses in the restored population and resulted in faster replenishment of lost diversity after supplementation. The sensitivity analyses demonstrated that large changes in the mean and variance of the modeled distribution for variance in reproductive success (geometric) impacted N_e significantly. Overall, this study demonstrates the utility of a genetic simulation approach for examining how complex hatchery restoration scenarios can impact genetic diversity over the short- and long-term. This chapter also highlights the importance of understanding, to the extent possible, the reproductive and life history features of the species in question. The results from this study may be useful for planning future oyster supplementation programs.

Model assumptions and caveats

In the creation of the IBM, a few simplifying assumptions were made, which may have affected model outcomes and inferences. First, no sex change was assumed in individuals over time. Oysters are protandric hermaphrodites, and the ratio of female to male oysters increases with the mean size of oysters (Burkenroad 1931; Harding et al.

2013). Second, equal fecundity across age classes was assumed as this was very complicated to code within the simuPOP framework. The number of eggs per female increases non-linearly with oyster size (Mann and Evans 1998) and it is unclear how including this in the model may have impacted results. Third, equal growth, survival, and fitness of hatchery-produced and wild oysters was assumed. This assumption is contradictory to a number of finfish studies which demonstrate lower fitness of hatchery-produced fish compared to wild counterparts (e.g. Araki et al. 2008, 2009). It remains unknown if hatchery culture results in fitness decreases in oysters subsequently planted in the field. Fourth, migration rate was fixed (15% each year) within the model to reflect empirical F_{ST} values. Only two other populations were included, while in reality the Harris Creek metapopulation may contain more than two populations, with migration rates that likely vary substantially each year. Including the two local populations within the model framework allowed for insight into how migration may impact changes in genetic diversity resulting from hatchery-based supplementation.

Another important caveat is that the IBM models neutral genetic diversity and structure only. Adaptive genetic diversity was not included, which would directly influence demography and may be associated with higher probabilities of supplementation success and persistence when individuals are sourced from nearby populations (Neraas and Spruell 2001; Weeks et al. 2011; DeHaan and Bernall 2013) or those with similar environments (Wang and Bradburd 2014). Future restoration modeling efforts with the IBM could incorporate fitness parameters and examine the impacts of domestication selection or relative fitness differences between wild and hatchery produced individuals.

Impact of spawning practices and broodstock number on genetic diversity

Effective population size was the genetic diversity metric most sensitive to hatchery-based supplementation and to varying hatchery management strategies in the model simulations. In contrast, allelic richness was only mildly impacted by supplementation, and heterozygosity was even less sensitive, which is congruent with a number of empirical studies examining diversity changes in hatchery-reared populations. For example, reductions in allelic richness were more common than reductions in heterozygosity in early generations of hatchery-reared populations (reviewed by Araki and Schmid 2010). Segovia-Viadero et al. (2016) found reductions in allelic richness but not in expected heterozygosity in early generation hatchery-reared populations of urchins compared with wild populations. Furthermore, the authors found a significantly reduced N_e in hatchery-reared populations (Segovia-Viadero et al. 2016). Similarly, using an IBM for red drum stocking strategies, Katalinas et al. (2019) found N_e and A_r to be more sensitive than heterozygosity to the impact of fish stocking. Overall, results from this simulation study, as well as other modeling and empirical studies, indicate that monitoring of hatchery-based oyster supplementation programs should focus on how N_e changes in receiving populations.

Model results indicated that that either using large numbers of rotated broodstock from the wild or utilizing controlled spawns when broodstock numbers are small, can both positively impact the genetic diversity of a receiving natural population (Blankenship and Leber 1995; Munro and Bell 1997; Fisch et al. 2015). A similar finding of the positive influence of broodstock on restored population N_e was shown in Chapter 3

with field sampling and population genetic approaches. For example, N_e of supplemented reefs was significantly correlated to the number of broodstock used for supplementation. This is also consistent with a previous simulation study by Waples and Do (1994), who found that the number of breeders was the single most important factor determining the genetic impact of hatchery-based supplementation of Pacific salmonids. Recycled (mass-spawn) broodstock scenarios produced the smallest increases in genetic diversity metrics across simulations. This result is similar to a previous study that found relationships between inbreeding coefficient and the rate of broodstock rotation as well as the number of broodstock used in the hatchery (Duchesne and Bernatchez 2002). Overall, the results of this Chapter join those of an increasing number of investigations demonstrating that hatchery-based supplementation may, in some circumstances, result in positive or negligible change in genetic diversity of receiving populations compared to wild populations (e.g. Waples and Do 1994; Cuenco 1994; Ryman et al. 1995; Wang and Ryman 2001; Duchesne and Bernatchez 2002; Katalinas et al. 2019). Nevertheless, these simulations show how specific hatchery practices can alter diversity of restored populations and may help to guide decisions in the hatchery to maximize retention of diversity.

Impact of migration on genetic diversity metrics

The degree of connectivity among restored and non-restored populations had a clear impact on genetic diversity in model scenarios. During the supplementation period, the supplemented population consistently had greater diversity in scenarios with migration compared to scenarios without migration or with lower migration rates. In the

post-supplementation period, initial decreases in genetic diversity due to supplementation were eventually reversed relative to the control in populations that were experienced migration. In addition, the pace of this rebound in diversity was greater with higher levels of migration. This finding is similar to results from Duchesne and Bernatchez (2002) who showed that supplemented populations that are connected to a larger metapopulation through gene flow recover from the genetic risks of inbreeding more quickly than isolated populations. This idea is also supported (indirectly) by a study of scallop seeding in the Bay of Brest, France by Morvezen et al (2016), who found very little decrease in diversity due to stocking, which was possibly mitigated by high gene flow from undetected subpopulations. Given the concordance between initial modeling efforts and empirical observations, it is clear that restoration programs should try to take into account the level of gene flow between supplemented and neighboring populations when designing restoration plans. These results also suggest that assuming complete genetic isolation of supplemented populations may lead to overly conservative prediction regarding the rates of diversity loss in these populations. Knowledge of connectivity among populations proximal to the supplemented population is critical for predicting the long-term impacts of supplementation on diversity.

Implications for hatchery practices or restoration strategies in shellfish

The IBM created in this study can be used to inform hatchery practices and guide supplementation of eastern oysters in the Chesapeake Bay, and possibly other shellfish species with similar life history features. The IBM created in this study is a fully flexible tool that accommodates constraints on production facilities, including the number of

hatchery broodstock, spawning type (controlled or mass) and/or broodstock practices (rotated or recycled). Furthermore, uncertainty in system or population specific parameters (e.g. abundance of wild population, variance in reproductive success, and poly number of species, and recruitment variability) can also be modeled. Therefore, this model can be used as a research tool for a variety of hatchery-based supplementation programs with varying goals. Even with a wide range of uncertainty surrounding certain parameters in the simulation (i.e. poly number and reproductive distribution of parents), clear evidence for the importance of spawning practice (controlled vs. mass), broodstock number (N=10, 25, and 500) and connectivity/migration between populations driving genetic outcomes of simulations was found. Continued efforts to understand these parameters, their variability across programs, and how they interact with the genetics of a specific system will inform the potential success of future hatchery-based restoration efforts, especially for marine shellfish with complex life-history features.

The results from this chapter revealed several management-relevant outcomes and a number of standard recommendations can be made for hatchery-based supplementation of oysters. First, if available, large numbers of broodstock (0.5% of wild population abundance) should be utilized, with the best practices involving controlled-spawns, then mass-spawns with rotated broodstock. In addition, if large broodstock numbers are not available, rotating smaller broodstock numbers and utilizing mass- or controlled-spawns can lead to the maintenance and even increase in genetic diversity of receiving populations. However, pair-cross matings and strip spawning may not be feasible in some programs. Therefore, the most effective strategy will depend on the hatchery facility, labor, and broodstock numbers available. Overall, this work highlights the importance of

understanding the life-history of the target species, specifically the variance of reproductive success, and supports the use of IBMs as a tool for developing responsible species management strategies.

Acknowledgements

I thank Matt Damiano, Kathryn Doering, and Mike Wilberg for providing empirical data used to parameterize the model. I also thank Kenny Rose and Allan Strand for early help on model conceptualization. I thank Bo Peng, for his conistsent help while building the model. I finally thank Tsetso Bachvaroff. and Institute of Marine and Environmental Technology for providing computational resources. Support for this work was provided by Maryland Sea Grant (5-230360 to KMH and LVP), and Maryland Sea Grant (5-230890 to KMH and LVP).

Chapter 4 Tables

Table 1. Demographic and life-history parameters applied to wild and hatchery populations of *C. virginica*

Parameter	Value	Source/Equation
Starting abundance	100,000	downscaled Damiano & Wilberg (2018)
Adult survival rate ages 3+	0.74	Damiano & Wilberg (2018); Paynter et al. (2011)
Age 0-1 survival rate	0.5	Damiano & Wilberg (2018)
Ages 1-2 survival rate	0.6	Damiano & Wilberg (2018)
Mean generation time (years)	2	Burkenroad (1931); Coe (1936); Hayres & Menzel (1981); O'Beirn et al. (1996)
Maximum age	11	Paynter et al. (2010); Powell et al. (1996); Paraso et al. (1999); Ragone Calvo et al. (2001); Harding et al. (2008)
Migration rate	15%	FST from Hornick & Plough (*unpubl.*)
Number supplemented	1–50X wild population size	Horn Point Laboratory Oyster Hatchery data
Poly number (wild)	2–0.003X wild population size	NA
Poly number (hatchery)	1–number of broodstock	Levinton (2005)
Recruitment distribution	gamma	Damiano & Wilberg (2018); gamma (shape=1, scale=size of reproductively mature population).
Reproductive distribution (hatchery)	gamma	(shape=0.01, scale =10000) Hornick & Plough (2019)
Reproductive distribution (wild)	geometric	Hornick & Plough (2019)

Table 2. Twenty-seven model scenarios of hatchery-based supplementation including those varying spawning type (controlled or mass-spawn), broodstock management (recycled or rotated), and broodstock numbers (N=10, 25, and 500). Mass-spawn scenarios also tests the genetic impact of varying migration rates (closed, medium, and open) for rotated and recycled broodstock.

Spawning and broodstock practices			**Varying migration rates**					
Controlled-spawn	***Mass-spawn***		***Mass-spawn***					
			recycled			*rotated*		
rotated	*recycled*	*rotated*	*closed*	*medium*	*open*	*closed*	*medium*	*open*
10	10	10	10	10	10	10	10	10
25	25	25	25	25	25	25	25	25
500	500	500	500	500	500	500	500	500

Table 3. Parameters tested in sensitivity analysis of hatchery supplementation IBM of *C. virignica*

	Geometric (p)[1]	**Poly number**
1	0.00002	0.00003*population size
2	0.0002	0.0003*population size
3	0.002	0.003*population size
4	0.02	0.03*population size
5	0.2	0.3*population size

[1]Number of offspring for each mating event follows a geometric distribution with mean $1/p$ and variance $(1-p)/p^2$

Table 4. Summary table of changes in genetic diversity metrics including effective population size (N_e), allelic richness (A_r), and observed heterozygosity (H_o) throughout the course of model simulations for mass and controlled-spawn scenarios. Spawning types are further divided by broodstock practice (rotated and recycled), and periods during the course of simulations including supplementation (Supp. generation 3), post-supplementation early (Post-supp. early, generation 9), and post-supplementation late (Post-supp. Late, generation 27). Genetic diversity metrics are represented as percent difference relative to control scenarios.

	Controlled-spawn			**Mass-spawn**					
	Rotated			**Recycled**			**Rotated**		
	Supp.	***Post-supp. (early)***	***Post-supp. (late)***	***Supp.***	***Post-supp. (early)***	***Post-supp. (late)***	***Supp.***	***Post-supp. (early)***	***Post-supp. (late)***
N_e									
10	-31.3	319.6	71.5	-76.4	119.8	98.3	-19.2	514.6	130.3
25	75.9	477.4	206	-47.3	222.3	183.8	52.6	597.3	71.9
500	368.9	769.1	235.9	117.2	436.8	238	132.5	701.5	173.7
A_r									
10	-1.75	-1.31	0.32	-2.38	-2.2	-2.36	-1.76	-1.46	0.32
25	-1.66	-0.89	1.07	-2.17	-1.46	0.11	-1.66	-0.84	1.07
500	-0.62	0.16	1.82	-0.83	0	1.18	-0.93	-0.21	1.5
H_o									
10	-0.81	0.11	4.24	-0.8	-3.08	0.21	-0.83	-0.62	1.15
25	-1.08	0.25	5.36	-0.81	0.07	3.28	-1.18	0.73	3.91
500	-0.03	1.63	4.56	-0.09	0.84	5.13	-0.25	0.35	3.45

Chapter 4 Figures

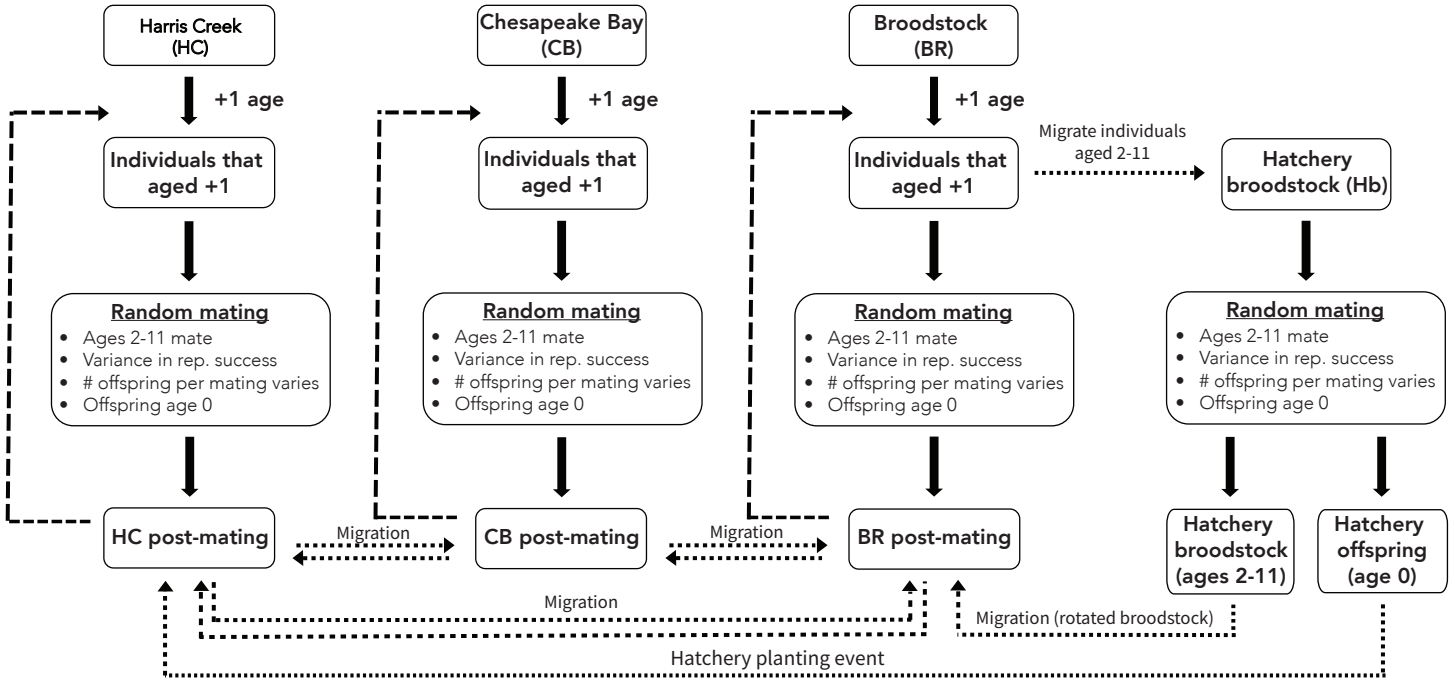


Figure 1. Flow chart depicting the stages of the individual-based model in simuPOP. Dotted arrows represent migration events and dashed arrows represent one time-step (year) in the model

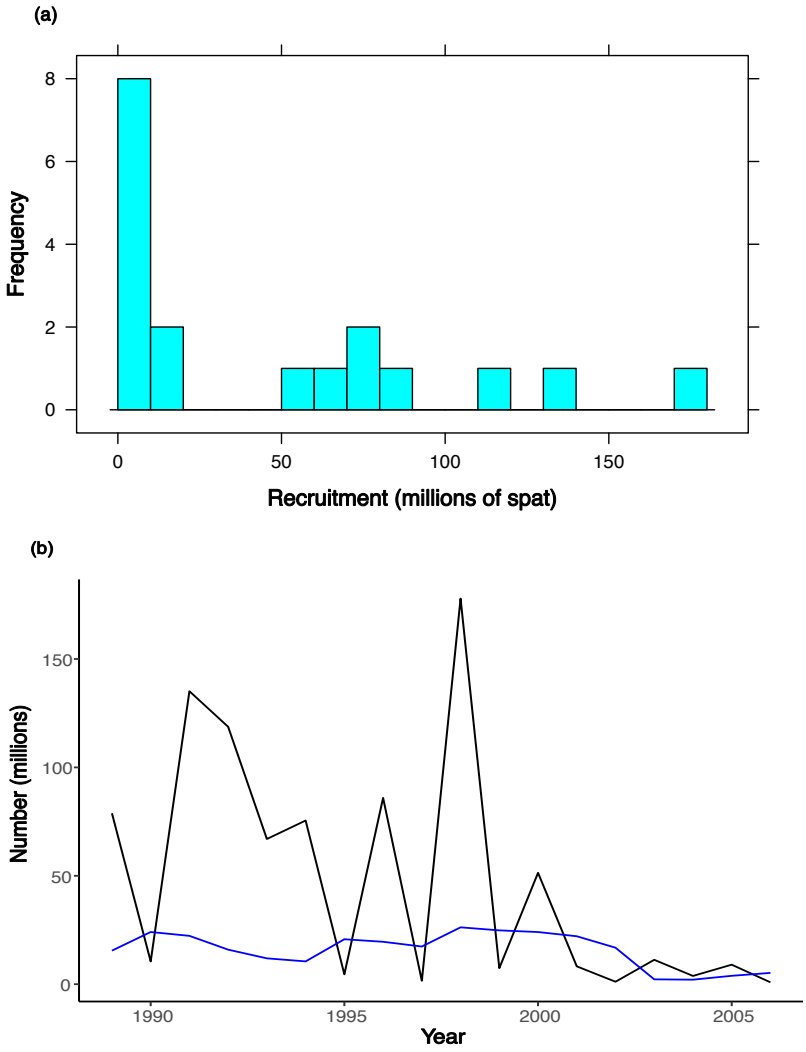


Figure 2. Empirical data on *C. virginica* spat recruitment. Panel A. shows the distribution of spat recruitment and Panel B. shows the abundance of spat (black) and market sized oysters (blue) in Harris Creek from 1989–2006 in Harris Creek, MD from Damiano and Wilberg (2019).

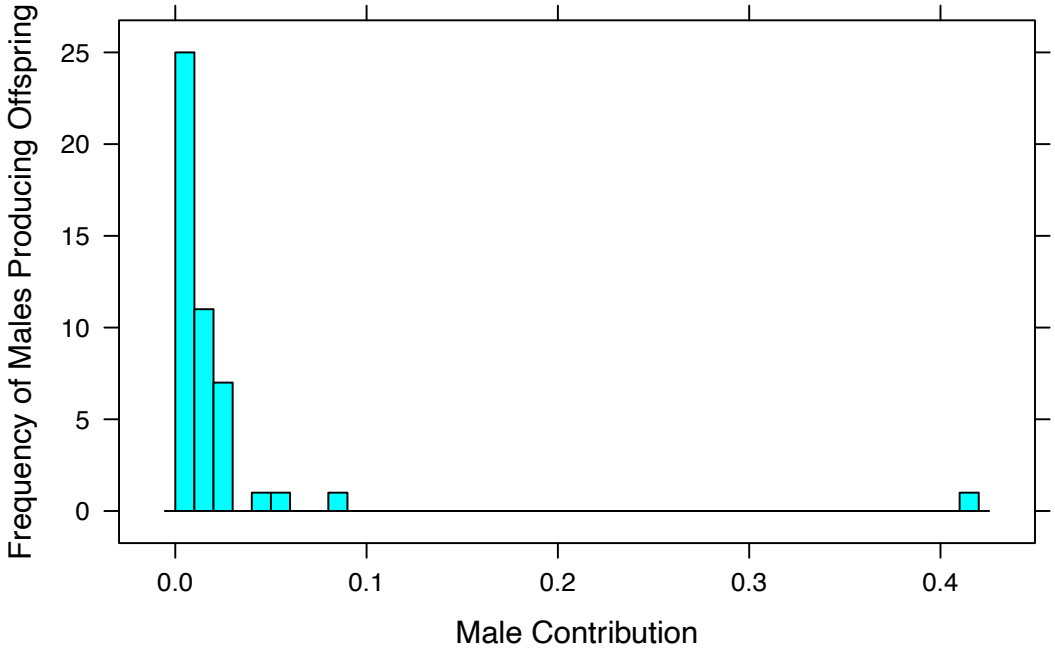


Figure 3. Simulated distribution of *C. virginica* male reproductive success for one reproductive event. The histogram represents only males that contributed to offspring (110). Males not contributing any offspring represented 99% of the breeding population (123,168 individuals).

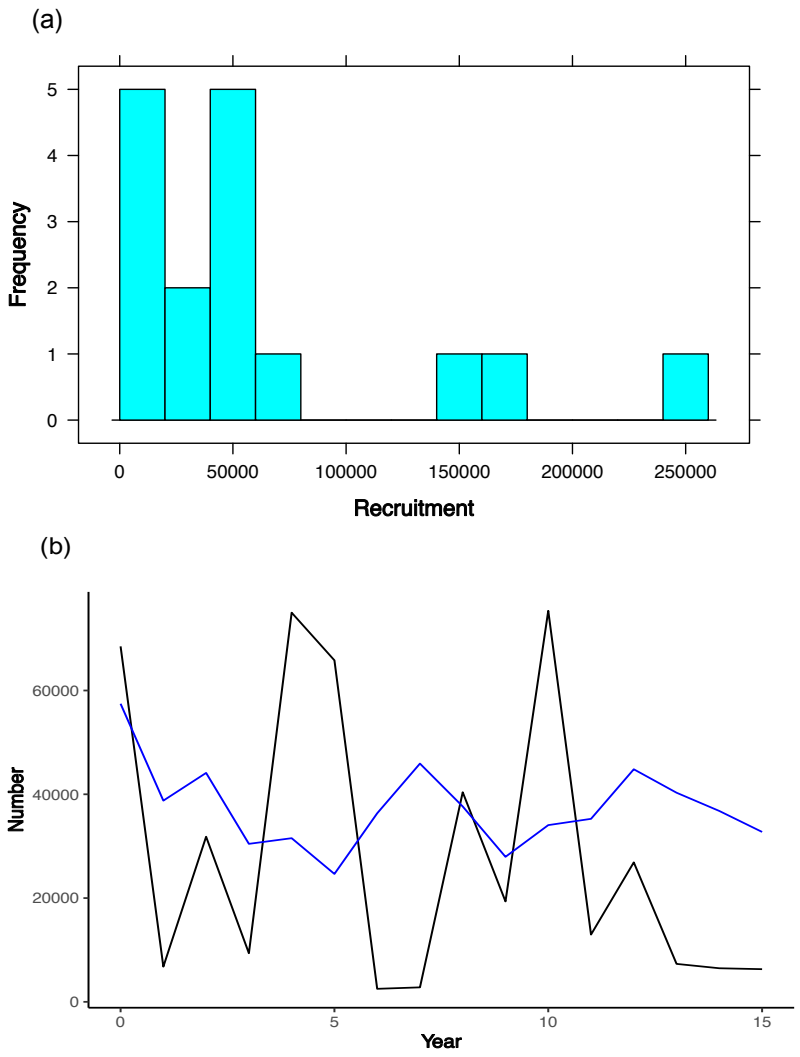


Figure 4. Modeling results for recruitment of *C. virginica* spat. Panel A shows the distribution of spat recruitment and panel B shows spat (black) and reproductively mature oysters (blue; ages 2–11) in the virtual Harris Creek wild population during a 50-generation model run.

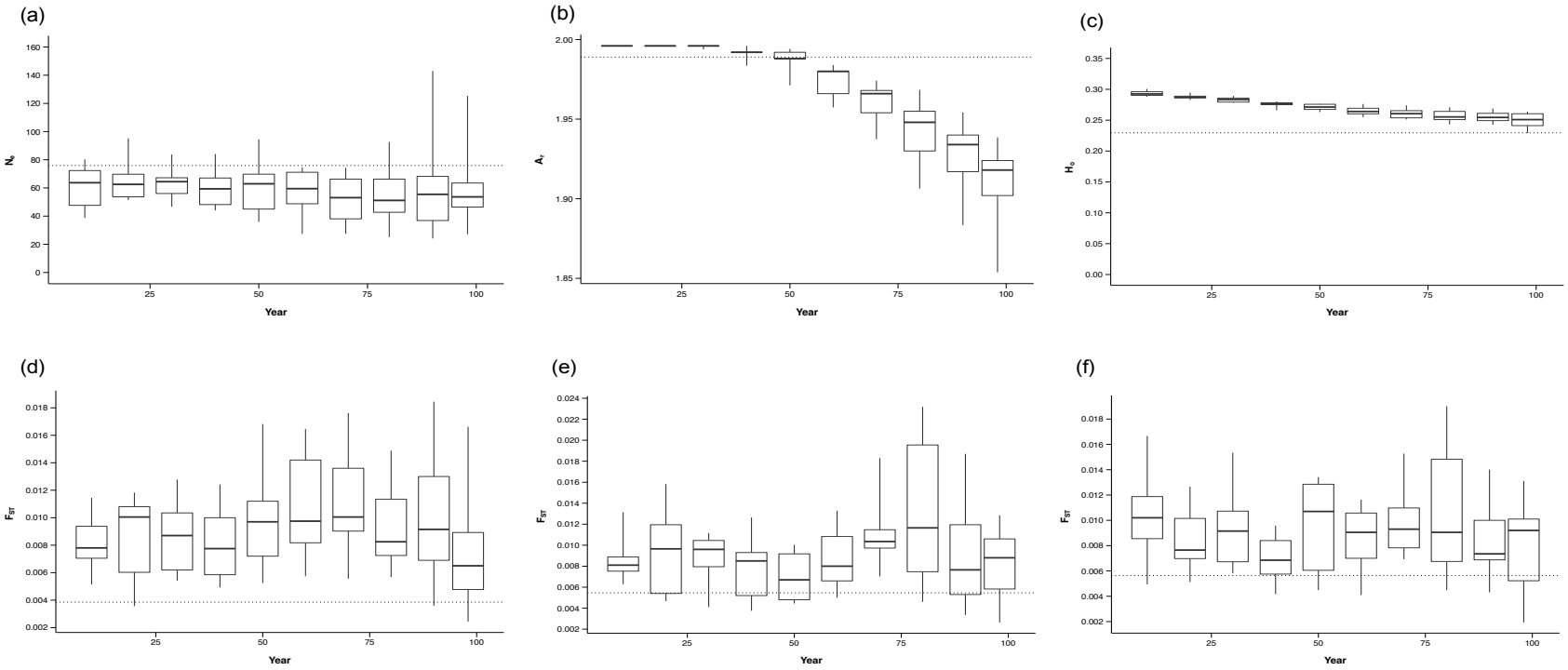


Figure 5. Comparison of genetic diversity metrics from simulated populations vs empirical estimates from the Chesapeake Bay. Panel a shows the effective population size (N_e), panel b shows allelic richness (A_r), panel c shows observed heterozygosity (H_o), and Panels d–f show pairwise F_{ST} of Harris Creek-Chesapeake Bay (d), Harris Creek-Broodstock (e), and Chesapeake Bay-Broodstock (f). The dotted line represents the empirical estimates from Chapter 3 across 2,210 RAD loci from mixed-age cohorts of oysters collected from Harris Creek (Harris Creek), Little Choptank (Chesapeake Bay), and States Bank (Broodstock). Note that the scales differ across panels.

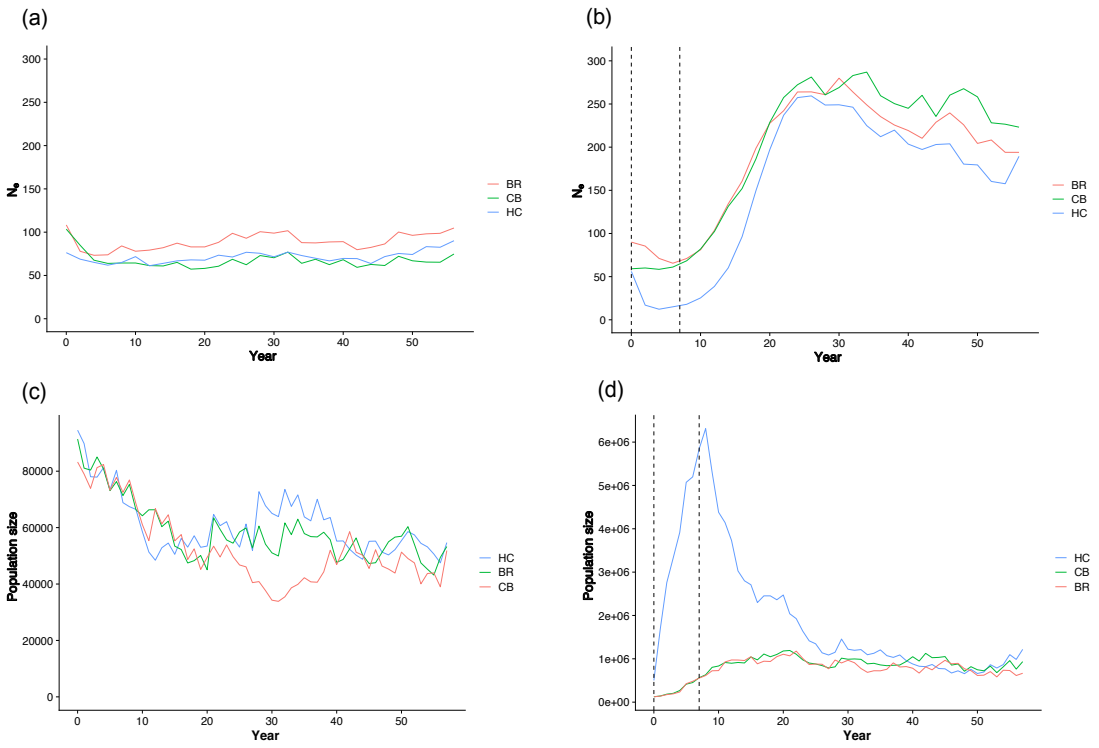


Figure 6. Raw modeling results for a control scenario and a mass-spawned, recycled broodstock, 10 broodstock restoration scenario for 50 years. Estimates of effective population size (N_e) are shown in panel A for the control scenario, and panel B for the supplementation scenario. Change in population sizes are depicted for the control scenario (panel C), and the restored scenario (panel D). Data for each population is presented, including the supplemented population (HC- blue lines), a local population that receives migrants from the supplemented population (CB -green lines), and the wild broodstock population (BR- red lines). Estimates were based on a subsample of 10,000 *C. virginica* mixed-aged individuals (ages 2–11) and the line represents the average across 10 simulation replicates. The vertical dotted lines represent the start and the end of the supplementation period.

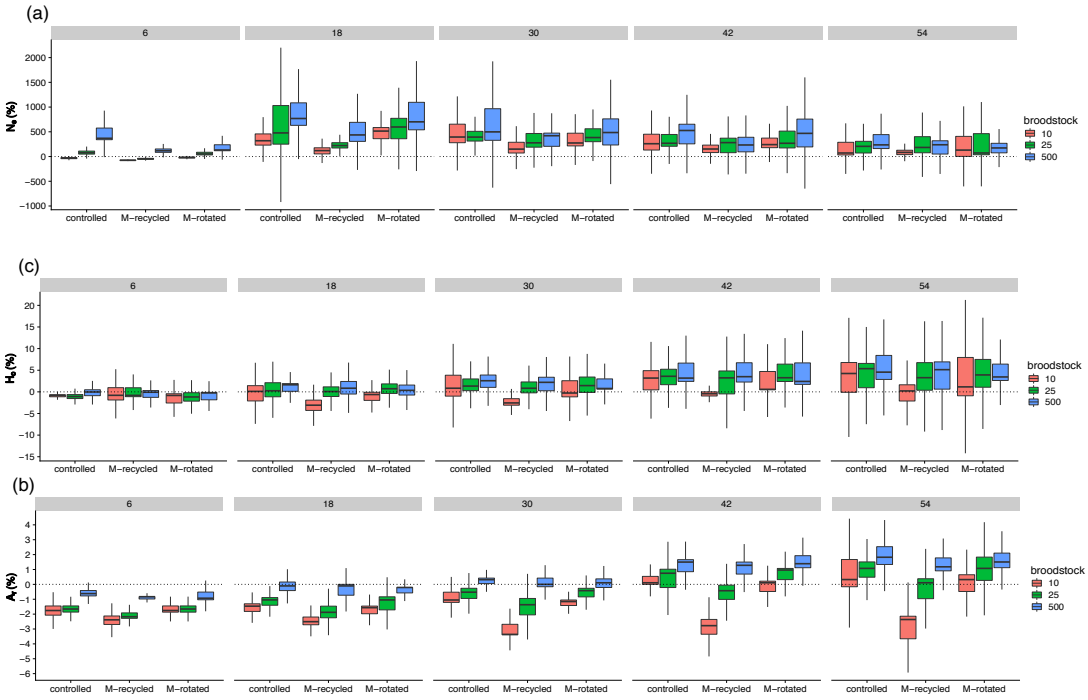


Figure 7. Model simulation results (shown in years, grey boxes above plots) for genetic diversity metrics across varying spawning and broodstock management scenarios. The varying scenarios include controlled-spawns (controlled) and mass-spawns of recycled (M-recycled) and rotated (M-rotated) broodstock and results are compared to control scenarios (percent difference). Genetic diversity metrics are shown over time during the simulations (shown as generations in the grey bars above plots). Panel a shows effective population size (N_e) estimates, Panel b shows allelic richness (A_r) and Panel c shows observed heterozygosity (H_o). Estimates of genetic diversity metrics were based on a subsample of 10,000 *C. virginica* mixed-aged individuals (ages 2–11) and are averages across 10 simulation replicates

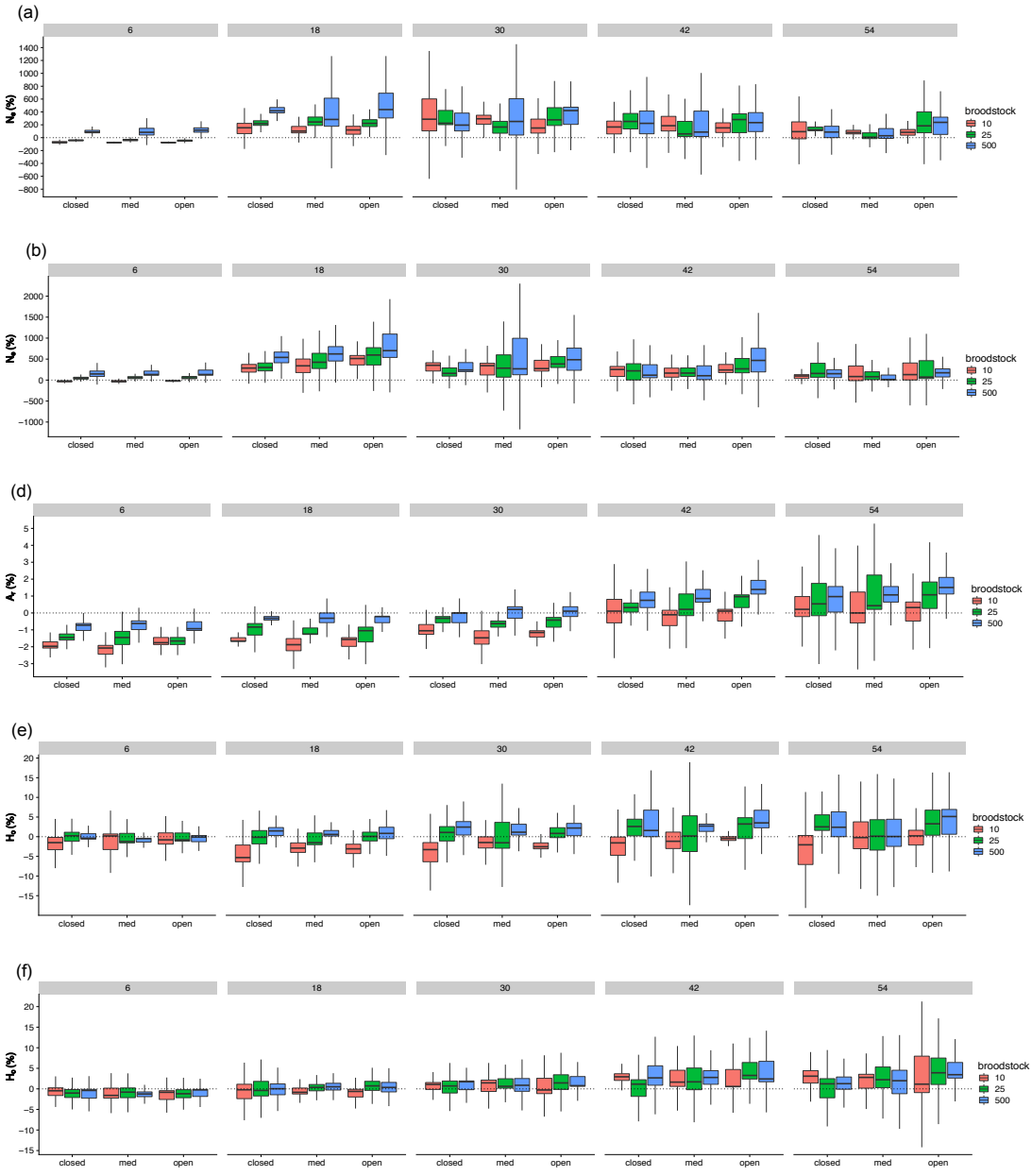


Figure 8. Model simulation results (shown in years, grey boxes above plots) for genetic diversity metrics across varying migration and broodstock management scenarios. Panels a and b show effective population size (N_e) estimates for mass-spawned recycled broodstock and rotated broodstock, respectively. Panels c and d show allelic richness (A_r) estimates for mass-spawned recycled broodstock and rotated broodstock, respectively. Panels e and f show observed heterozygosity (H_o) estimates for mass spawned recycled broodstock and rotated broodstock, respectively. Different colored bars represent different number of broodstock (N=10, 25, and 500). The migration rates tested are closed (no migration), medium (5% migration), and open (15% migration). Genetic

diversity metrics are represented as percent difference relative to control scenarios. Estimates were based on a subsample of 10,000 *C. virginica* mixed-aged individuals (ages 2–11) and are averages across 10

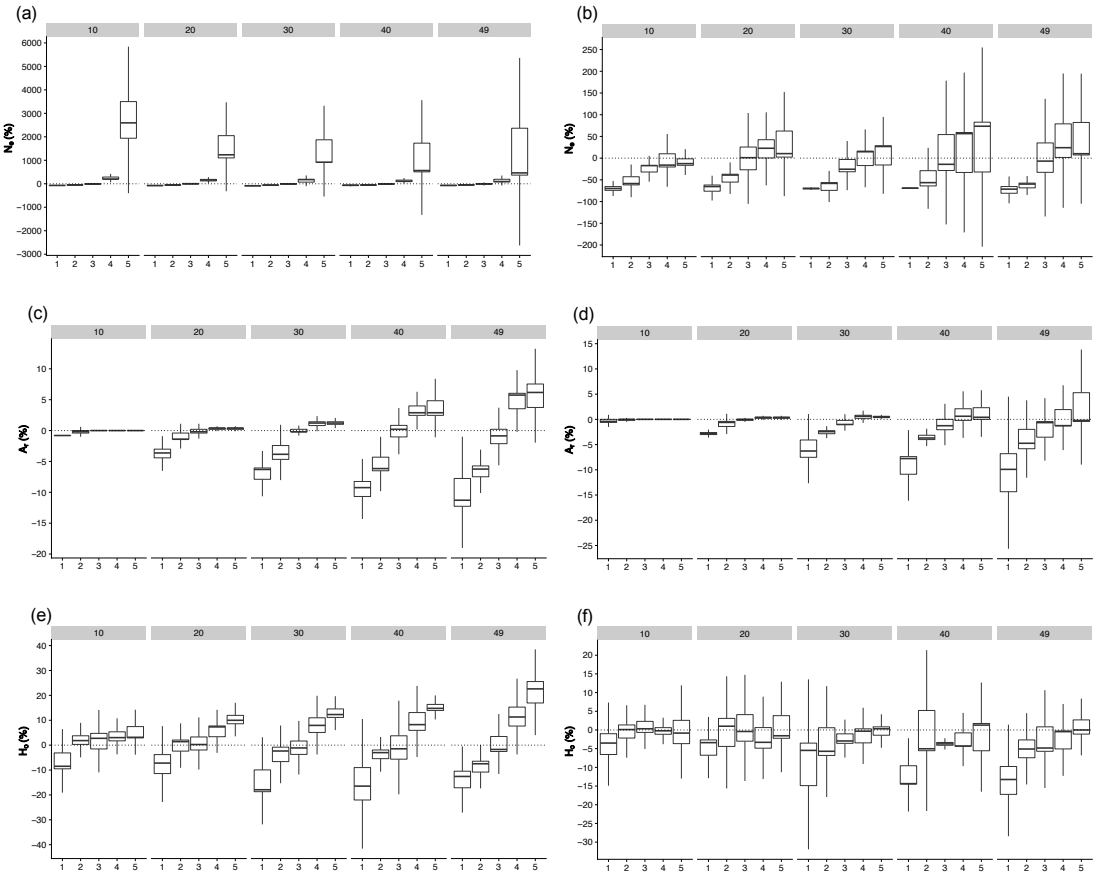


Figure 9. Sensitivity simulation results (shown in years, grey boxes above plots) for genetic diversity metrics. Genetic diversity metrics are represented as percent difference relative to control scenarios using the standard modeling parameters for all scenarios. Panels a and b show effective population size (N_e) estimates for sensitivity analyses of the geometric reproductive distribution (varying *P*) and poly number of males and females in a reproductive event, respectively. Panels c and d show allelic richness (A_r) estimates for sensitivity analyses of the geometric reproductive distribution (varying *P*) and poly number of males and females in a reproductive event, respectively. Panels E and F show observed heterozygosity (H_o) estimates for sensitivity analyses of the geometric reproductive distribution (varying *P*) and poly number of males and females in a reproductive event, respectively. Values corresponding to the numbers of the x-axis for the geometric distribution (Panels a, c, and e) are 0.00002, 0.0002, 0.002, 0.02, and 0.2 for 1, 2, 3, 4, and 5, respectively. Values corresponding to the numbers of the x-axis for the poly number (Panels b, d, and f) are 0.3X, 0.03X, 0.003X, 0.0003X, and 0.00003X for 1, 2, 3, 4, and 5 (X=population size), respectively. Tukey's boxplots were calculated from 5 replicates.

Chapter 5: Synthesis, conclusions and future work

General summary of dissertation scope and findings

My dissertation focused on the genetic impacts of hatchery-based supplementation of the eastern oyster (*Crassostrea virginica*), a foundation species in North America that has been severely depleted by overfishing, habitat destruction, and disease. This work represents one of only a few efforts to explore the genetic impacts of oyster hatchery-based supplementation from multiple perspectives in a species with complex life-history features including broadcast spawning, high fecundity, type III survivorship, iteroparity, and overlapping generations. While hatchery practices can be implemented to maintain diversity in hatchery-produced individuals, characteristics of species life-history and biology may pose challenges to maintaining this diversity, and these reproductive and biological factors are not always straightforward to control. Particularly for oysters, high-fecundity and Type-III survivorship can lead to reductions in genetic diversity even under controlled-spawning designs as shown in Chapter 2 and elsewhere. Key metrics of population genetic diversity, such as effective population size and population structure, are not well characterized in this species, further challenging restoration efforts. With the predicted expansion of captive breeding, the work in this dissertation is of importance to understand how genetic diversity and structure is generated in natural oysters populations and how the genetic diversity of hatchery-produced and of supplemented populations can be maximized, as well as the potential

impacts large-scale hatchery-based supplementation can have on natural populations. Below, I discuss the relevance of my dissertation research for marine animal conservation generally and recommend next steps for the effective design of hatchery-based supplementation programs for shellfish that may increase resiliency in the face of anthropogenic climate change.

Chapter 2 recap and future studies

Chapter 2 investigated genetic diversity changes associated with hatchery production of oysters via mass- and controlled- spawns. The results indicated that high variance in reproductive success that is commonly observed in oysters is not easily controlled during hatchery production, even in a controlled-spawning framework. However, only three cohorts per spawning type were analyzed, and this study would benefit from a greater sample size, and a more organized controlled-spawning framework. A future study could repeat this experiment with an increased number of cohorts in addition to comparing diversity changes in hatchery cohorts produced via mass-, controlled-, and strip-spawning methods. Strip-spawning could minimize the 'timing' issues that occurred in the lab during the controlled-spawns (e.g. the lag between spawning and fertilization). In addition, sampling all individuals present during spawning (even those that did not spawn) would be valuable for more complete understanding of diversity changes resulting from hatchery production, and because some of the potential parents in the previous study may have been missed.

Chapter 3 recap and future studies

In Chapter 3, the impacts of varying hatchery practices on observed genetic diversity were examined in the field at restored sites in Harris Creek. The striking and predictable relationships between broodstock number and planting frequency and genetic diversity at restored reefs in Harris Creek merit future investigation. Continued monitoring within Harris Creek, as well as other oyster sanctuaries in the Chesapeake Bay, and perhaps across captive breeding and supplementation programs of shellfish would be beneficial. Within Harris Creek, increased sampling resolution of restored reefs spanning the entire sanctuary would be particularly valuable. Furthermore, preserving tissue samples from broodstock and restored populations within the sanctuary will allow the tracking of genetic changes over time. Monitoring genetic diversity over time is important to evaluate the long-term success of this program and to better understand diversity changes associated with captive breeding and supplementation of shellfish. Overall, results from this chapter highlight the importance in utilizing molecular tools to better monitor genetic change in hatchery and wild populations of oysters to inform "best practices" in hatchery management to support declining wild populations.

The use of high-resolution genomic data from Chapter 3 also provided insight into neutral and adaptive variation of Chesapeake Bay oyster populations and the environmental factors driving this variation. Using genotype data derived from RAD sequencing of oysters in the Chesapeake Bay and a population in Virginia Coastal Bays, neutral and putatively adaptive markers of genetic variation were identified to infer the potential drivers of both gene flow and adaptive differentiation. Evidence for population genetic structure was found despite the potential for high levels of gene flow, with

populations in the Chesapeake Bay exhibiting significant isolation by distance. Using multivariate environmental association analyses, strong associations between salinity, temperature, pH, and dissolved oxygen and putative adaptive SNPs at fine spatial scales were found. Salinity was found to be the major predictor of both neutral and adaptive genetic variation in the Chesapeake Bay. Despite some interesting and sensible results for dissolved oxygen, there is a possibility of type II error for temperature and dissolved oxygen associations because the range of values for these variables were not as dynamic as those of salinity, and our sampling resolution of our environmental data was limited (twice per month). Future studies including finer resolution within-bay population and environmental sampling may reveal further patterns of selection and differentiation and may potentially impact the broad-scale correlations observed. This work contributes to the field of seascape genomics by providing support for potential adaptive differentiation driven by coastal environmental gradients and how this information can potentially be utilized for broodstock sourcing decisions. While our findings suggest that oyster populations within the Chesapeake Bay are locally adapted, future work should incorporate reciprocal transplant experiments to examine more directly the effects of variable salinity on eastern oyster fitness.

Chapter 4 recap and future studies

The individual based model (IBM) created in this Chapter is one of the first of its kind for marine shellfish populations and can be used to explicitly test and compare the effects of certain hatchery practices on genetic diversity. This model also has potential to actually guide more genetically aware supplementation plans for eastern oysters in the

Chesapeake Bay, as well as shellfish species with similar life history features. The model is a flexible population genetics simulation tool that incorporates realistic constraints on production facilities. These include the number of hatchery broodstock, spawning practices (controlled or mass), broodstock use and maintenance (rotated from the wild or recycled within the hatchery), and uncertainty in system parameters (abundance of wild population, variance in reproductive success, and poly number of species–number of males that mate with females and vice versa, and recruitment variability). Therefore, this model can be used as a research tool for a variety of hatchery-based supplementation programs with varying goals. Furthermore, the utility of an IBM genetic modeling approach in evaluating the impact of varying hatchery practices on observed genetic diversity metrics of a supplemented wild population is demonstrated. Even with a wide range of uncertainty surrounding key parameters in the simulation (i.e. poly number and the relative reproductive contribution among parents), we found evidence of the importance of spawning practice (controlled vs. mass), broodstock number (N=10, 25, and 500) and connectivity/migration between populations driving genetic outcomes of simulations. Continued efforts to understand these parameters, how they vary across programs, and how they interact with the gene flow and N_e of a specific system will inform the potential success of future hatchery-based restoration efforts.

Hatchery-based supplementation systems are complex, and the results from Chapter 4 should not be used in isolation, but in consideration of other issues that would be useful to test in future modeling efforts. We did not include adaptive genetic diversity in the model (i.e. natural selection on traits), which would directly influence demography and may be linked to higher probabilities of supplementation success and persistence when

individuals are sourced from nearby populations (Neraas and Spruell 2001; Weeks et al. 2011; DeHaan and Bernall 2013) or those with similar environments (Wang and Bradburd 2014). Future supplementation efforts will likely incorporate more powerful genomic approaches in selecting and evaluating source populations for supplementation (He et al. 2016). In addition, future scenarios testing long-term supplementation with varying wild population size, adaptation to culture conditions, or domestication selection, as well as the impact of fishing pressure on the observed genetic diversity of wild populations would be valuable next steps. Spatial demographic processes that may influence population genetics of oyster populations present a challenge when attempting to generalize the change of persistence across different systems. A simulation-based approach, particularly one that is spatially explicit, can help evaluate dynamic patterns of occupancy and emulate spatial processes in way that cannot be described by simple metapopulation models or conceptual frameworks. Simulation-based models cannot provide empirical information regarding connectivity of oysters within the Chesapeake Bay and within other systems, but they can provide crucial demographic insights that are impossible to achieve by other means. The utility of population genetic simulation approaches will likely only grow as the need for hatchery-based restoration and management strategies continues (Scribner et al. 2016), and as the fields of landscape and metapopulation ecology continue to converge (Howell et al. 2018).

Overall research findings, limitations, and future directions

It is important to note that this dissertation did not explore the potential for domestication selection during captive breeding of bivalve shellfish, which is understudied. Understanding the potential for domestication selection, and if it poses demonstrable fitness risks is imperative to the future success of captive breeding and supplementation programs of shellfish. For example, in programs that are broodstock-limited, the potential use of domesticated aquaculture strains for restoration has been considered. Future work investigating the potential impacts domesticated oysters and shellfish can have on natural populations is necessary and new and emerging genomic tools will facilitate this research.

The results of my dissertation research highlight that hatchery practices (both intentional and not) have important genetic consequences for restored populations and can have long-term implications for species conservation. However, there is a clear need to establish a stronger link between restoration research and restoration practices. While management decisions may need to be made with incomplete information, monitoring genetic diversity within restored populations provides additional data/tools for decision making and could increase conservation and restoration success by giving managers more information to aid in restoration planning and implementation.

Bibliography

Abele D, Heise K, Pörtner HO, Puntarulo S (2002) Temperature-dependence of mitochondrial function and production of reactive oxygen species in the intertidal mud clam *Mya arenaria*. J Exp Biol 205:1831

Adkison M (1995) Population differentiation in Pacific salmon: local adaptation, genetic drift, or the environment? Canadian Journal of Fish and Aquatic Sciences 52:2762–2777

Agashe D (2009) The Stabilizing Effect of Intraspecific Genetic Variation on Population Dynamics in Novel and Ancestral Habitats. The American Naturalist 174:255–267.

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. Nature Reviews Genetics 11:697–709.

Allendorf FW, Luikart G, Aitken SN (2013) Conservation and the genetics of populations, 2nd ed. John Wiley & Sons, Hoboken

Allendorf FW, Ryman N (1987) Genetic management of hatchery stocks. In: Population genetics and fishery management eds. 1987 eds. 1987. pp 141–159

Anderson EC (2010) Assessing the power of informative subsets of loci for population assignment: standard methods are upwardly biased. Molecular Ecology Resources 10:701–710.

André C, Lindegarth M (1995) Fertilization efficiency and gamete viability of a sessile, free-spawning bivalve, *Cerastoderma edule*. Ophelia 43:215–277

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Appleyard SA, Ward RD (2006) Genetic diversity and effective population size in mass selection lines of Pacific oyster (*Crassostrea gigas*). Aquaculture 254:148–159.

Aquaculture Genetics and Breeding Technology Center (ABC) (2009) The aquaculture genetics and breeding technology's oyster breeding programme. Virginia Institute of Marine Science, Gloucester Point, VA, p 28

Araki H, Berejikian BA, Ford MJ, Blouin MS (2008) Synthesis: Fitness of hatchery-reared salmonids in the wild. Evolutionary Applications 1:342–355.

Araki H, Cooper B, Blouin MS (2007) Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. Science 318:100–103.

Araki H, Cooper B, Blouin MS (2009) Carry-over effect of captive breeding reduces reproductive fitness of wild-born descendants in the wild. Biology Letters 5:621–624.

Araki H, Schmid C (2010) Is hatchery stocking a help or harm? Evidence, limitations and future directions in ecological and genetic surveys. Aquaculture 308:S2–S11.

Arnaldi KG, Walters LJ, Hoffman EA (2018) Effects of time and harvest on genetic diversity of natural and restored oyster reefs. Restoration Ecology 26:943–951.

Attard CRM, Möller LM, Sasaki M, et al (2016) A novel holistic framework for genetic-based captive-breeding and reintroduction programs: Holistic Framework for Reintroductions. Conservation Biology 30:1060–1069.

Avise J (1986) Mitochondrial DNA and the evolutionary genetics of higher animals. Philosophical Transactions of the Royal Society of London B, Biological Sciences 312:325–342.

Bahr LM, Lanier WP (1981) The ecology of intertidal oyster reefs of the south Atlantic coast: a community profile. FWS/OBS/81.15. U.S. Fish and Wildlife Service, Washington, D.C., USA.

Baird NA, Etter PD, Atwood TS, et al (2008) Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. PLoS ONE 3:e3376.

Balkenhol N, Landguth EL (2011) Simulation modelling in landscape genetics: on the need to go further. Molecular Ecology 20:667–670.

Bartley DM, Leber KM, FAO (eds) (2004) Marine ranching. Food and Agriculture Organization of the United Nations, Rome.

Baums IB (2008) A restoration genetics guide for coral reef conservation. Molecular Ecology 17:2796–2811.

Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian Computation in Population Genetics. Genetics 162:2025

Beck MW, Brumbaugh RD, Airoldi L, et al (2011) Oyster Reefs at Risk and Recommendations for Conservation, Restoration, and Management. BioScience 61:107–116.

Bekkevold D, Hansen MM, Loeschcke V (2002) Male reproductive competition in spawning aggregations of cod (*Gadus morhua,* L.). Molecular Ecology 11:91–102.

Bell JD, Leber KM, Blankenship HL, et al (2008) A new era for restocking, stock enhancement and sea ranching of coastal fisheries resources. Reviews in Fisheries Science 16:1–9.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological) 57:289–300.

Berejikian BA, Doornik DMV (2018) Increased natural reproduction and genetic diversity one generation after cessation of a steelhead trout (Oncorhynchus mykiss) conservation hatchery program. PLOS ONE 13:e0190799.

Bernatchez S, Xuereb A, Laporte M, et al (2019) Seascape genomics of eastern oyster (*Crassostrea virginica*) along the Atlantic coast of Canada. Evolutionary Applications 12:587–609.

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

Bie T, Meester L, Brendonck L, et al (2012) Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. Ecology Letters 15:740–747.

Bischoff A, Hurault B (2013) Scales and drivers of local adaptation in *Brassica nigra* (Brassicaceae) populations. American Journal of Botany 100:1162–1170.

Blankenship HLEE, Leber KM (1995) A Responsible Approach to Marine Stock Enhancement. American Fisheries Society Symposium 15:167–175

Borcard D, Legendre P, Drapeau P (1992) Partialling out the Spatial Component of Ecological Variation. Ecology 73:1045–1055.

Boudry P, Collet B, Cornette F, et al (2002) High variance in reproductive success of the Pacific oyster (C*rassostrea gigas*, Tunberg) revealed by microsatellite-based patentage analysis of multifactorial crosses. Aquaculture 204:283–296

Bowkett AE (2009) Recent Captive-Breeding Proposals and the Return of the Ark Concept to Global Species Conservation. Conservation Biology 23:773–776.

Bradbury IR, Laurel B, Snelgrove PVR, et al (2008) Global patterns in marine dispersal estimates: the influence of geography, taxonomic category and life history. Proc R Soc B 275:1803–1809.

Breed MF, Harrison PA, Bischoff A, et al (2018) Priority Actions to Improve Provenance Decision-Making. BioScience 68:510–516.

Breitburg DL (1992) Episodic Hypoxia in Chesapeake Bay: Interacting Effects of Recruitment, Behavior, and Physical Disturbance. Ecological Monographs 62:525–546.

Breseghello F, Sorrells ME (2006) Association Mapping of Kernel Size and Milling Quality in Wheat (*Triticum aestivum* L.) Cultivars. Genetics 172:1165–1177.

Broadhurst LM, Lowe A, Coates DJ, et al (2008) Seed supply for broadscale restoration: maximizing evolutionary potential. Evolutionary Applications 1:587–597.

Brooks WK (1891) The oyster, re-issued, 1996 Edition with a foreword by K.T. Paynter, Jr. Johns Hopkins University Press, Baltimore, MD, 230 pp.

Brown BL, Franklin DE, Gaffney PM, Hong M, Dendanto D, Kornfield I (2000) Characterization of microsatellite loci in the Eastern oyster, Crassostrea virginica. Mol Ecol 9:2217–2219

Brumbaugh RD, Beck MW, Coen LD, Craig L, Hicks P. (2006) A Practitioners' Guide to the design and monitoring of shellfish restoration projects: an ecosystem services approach. The Nature Conservancy, Arlington, VA., MRD Educational Report No. 22. 28 pp.

Brumbaugh RD, Coen LD (2009) Contemporary Approaches for Small-Scale Oyster Reef Restoration to Address Substrate Versus Recruitment Limitation: A Review and Comments Relevant for the Olympia Oyster, *Ostrea lurida* Carpenter 1864. Journal of Shellfish Research 28:147–161.

Burkenroad MD (1931) Sex in the Louisiana Oyster, *Ostrea virginica*. Science 74:71–72.

Buroker NE (1983) Population genetics of the American oyster *Crassostrea virginica* along the Atlantic coast and the Gulf of Mexico. Mar Biol 75:99–112.

Cacuci D (2003) Sensitivity and Uncertainty Analysis, Volume I: Theory

Calvo L, Wetzel RL, Burreson EM (2001) Development and verification of a model for the population dynamics of the protistan parasite, *Perkinsus marinus*, within its host, the eastern oyster, Crassostrea virginica, in Chesapeake Bay. Journal of Shellfish Research 20:231–241.

Camara MD, Vadopalas B (2009) Genetic Aspects of Restoring Olympia Oysters and Other Native Bivalves: Balancing the Need for Action, Good Intentions, and the Risks of Making Things Worse. Journal of Shellfish Research 28:121–145.

Capblancq T, Luu K, Blum MGB, Bazin E (2018) Evaluation of redundancy analysis to identify signatures of local adaptation. Mol Ecol Resour 18:1223–1233.

Carlsson J, Reece KS (2007) Eight PCR primers to amplify ESTlinked microsatellites in the Eastern oyster, Crassostrea virginica genome: Primer note. Mol Ecol Notes 7:257–259

Catchen J, Hohenlohe PA, Bassham S, et al (2013) Stacks: an analysis tool set for population genomics. Molecular Ecology 22:3124–3140.

Caughley G, Gunn A (1996) Conservation biology in theory and practice. Blackwell Science, Cambridge, Mass., USA

Chamberlin ME (2004) Top-down control analysis of the effect of temperature on ectotherm oxidative phosphorylation. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology 287:R794–R800.

Champagnon J, Elmberg J, Guillemain M, et al (2012) Conspecifics can be aliens too: A review of effects of restocking practices in vertebrates. Journal for Nature Conservation 20:231–241.

Chapuis M-P, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. Molecular Biology and Evolution 24:621–631.

Chen C, Lu RS, Zhu SS, et al (2017) Population structure and historical demography of *Dipteronia dyeriana* (Sapindaceae), an extremely narrow palaeoendemic plant from China: implications for conservation in a biodiversity hot spot. Heredity 119:95–106.

Cherkasov AS, Biswas PK, Ridings DM, et al (2006) Effects of acclimation temperature and cadmium exposure on cellular energy budgets in the marine mollusk Crassostrea virginica: linking cellular and mitochondrial responses. Journal of Experimental Biology 209:1274–1284.

Christie MR, Ford MJ, Blouin MS (2014) On the reproductive success of early-generation hatchery fish in the wild. Evolutionary Applications 7:883–896.

Christie MR, Marine ML, Fox SE, et al (2016) A single generation of domestication heritably alters the expression of hundreds of genes. Nature Communications 7:1–6. =

Christie MR, Marine ML, French R a, et al (2012a) Effective size of a wild salmonid population is greatly reduced by hatchery supplementation. Heredity 109:254–260.

Christie MR, Marine ML, French RA, Blouin MS (2012b) Genetic adaptation to captivity can occur in a single generation. Proceedings of the National Academy of Sciences 109:238–242.

Cleveland WS (1979) Robust Locally Weighted Regression and Smoothing Scatterplots. Journal of the American Statistical Association 74:829–836.

Clyde L. Mackenzie (2007) Causes underlying the historical decline in eastern oyster (*Crassostrea Virginica* GMELIN, 1791) landings. Journal of Shellfish Research 26:927–938

Coe WR (1936) Environment and sex in the oviparous oyster Ostrea virginica. The Biological Bulletin 71:353–359.

Coen LD, Brumbaugh RD, Bushek D, et al (2007) Ecosystem services related to oyster restoration. Marine Ecology Progress Series 341:303–307.

Coen LD, Luckenbach MW (2000) Developing success criteria and goals for evaluating oyster reef restoration: Ecological function or resource exploitation? Ecological Engineering 15:323–343.

Coltman DW, Pilkington JG, Smith JA, Pemberton, JM (1999) Parasite-mediated selection against inbred soay sheet in a free-living island population. Evolution 53:1259–1267.

Convention on Biological Diversity (2007). Available online at: https://www.cbd.int/doc/meetings/cop-bureau/cop-bur-2007/cop-bur-2007-10-14-en.pdf

Conover DO (1998) Local Adaptation in Marine Fishes: Evidence and Implications for Stock Enhancement. Bulletin of Marine Science 62:17

Cooper AM, Miller LM, Kapuscinski AR (2010) Conservation of population structure and genetic diversity under captive breeding of remnant coaster brook trout (*Salvelinus fontinalis*) populations. Conservation Genetics 11:1087–1093.

Crean AJ, Monro K, Marshall DJ (2011) Fitness consequences of larval traits persist across the metamorphic boundary. Evolution 65:3079–3089.

Crow JF (1954) Breeding structure of populations. II. Effective population number. Statistics and mathematics in biology. Iowa State College Press, 543-556.

Crow JF, Denniston C (1988) Inbreeding and variance effective population size numbers. Evolution 42:482–495.

Crow JF, Kimura M (1970) An introduction to population genetics theory. Burgess Pub. Co., Minneapolis

Cuenco ML (1994) A Model of an Internally Supplemented Population. Transactions of the American Fisheries Society 123:277–288.

Damiano MD, Wilberg MJ (2019) Population dynamics of eastern oysters in the Choptank River Complex, Maryland during 1989–2015. Fisheries Research 212:196–207. https://doi.org/10.1016/j.fishres.2018.12.023

Danecek P, Auton A, Abecasis G, et al (2011) The variant call format and VCFtools. Bioinformatics 27:2156–2158.

Darden TL, Robinson JD, Strand AE, Denson MR (2017) Forecasting the genetic impacts of net pen failures on Gulf of Mexico cobia populations using individual-based model simulations. Journal of the World Aquaculture Society 48:20–34.

Davey JW, Hohenlohe PA, Etter PD, et al (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Reviews Genetics 12:499–510. https://doi.org/10.1038/nrg3012

Davis HC (1958) Survival and growth of clam and oyster larvae at different salinities. The Biological Bulletin 114:296–307.

DeHaan PW, Bernall SR (2013) Spawning Success of Bull Trout Transported above Main-Stem Clark Fork River Dams in Idaho and Montana. North American Journal of Fisheries Management 33:1269–1282.

Díaz F, Luís A. Lima A, Nakamura AM, et al (2018) Evidence for Introgression Among Three Species of the Anastrepha fraterculus Group, a Radiating Species Complex of Fruit Flies. Frontiers in Genetics 9:359.

Diaz RJ (2001) Overview of Hypoxia around the World. J Environ Qual 30:275–281.

Dinnel PA, Peabody B, Peter-Contesse T (2009) Rebuilding Olympia Oysters, Ostrea lurida Carpenter 1864, in Fidalgo Bay, Washington. Journal of Shellfish Research 28:79–85.

Do C, Waples RS, Peel D, et al (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Molecular Ecology Resources 14:209–214.

do Prado FD, Vera M, Hermida M, et al (2018) Parallel evolution and adaptation to environmental factors in a marine flatfish: Implications for fisheries and aquaculture management of the turbot (*Scophthalmus maximus*). Evolutionary Applications 11:1322–1341.

Doering KL (2019) Patterns of oyster natural mortality in Chesapeake Bay, Maryland during 1991-2017 and its relationships with environmental factors and disease. University of Maryland

Duchesne P, Bernatchez L (2002) An analytical investigation of inbreeding in multi-generation supportive breeding. Conservation Genetics 3:45–58.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4:359–361.

Elliott M, McLusky DS (2002) The need for definitions in understanding estuaries. Estuarine, Coastal and Shelf Science 55:815–827.

Emlet RB, Sadro SS (2006) Linking stages of life history: How larval quality translates into juvenile performance for an intertidal barnacle (*Balanus glandula*). Integrative and Comparative Biology 46:334–346.

FAO (Food and Agriculture Organization of the United Nations) (2018). Fishery and aquaculture statistics. Aquaculture production. Retrieved from http://www.fao.org/fishery/static/Yearbook/YB2016_USBcard/navigation/index_content_aquaculture_e.htm#B

Fisch KM, Kozfkay CC, Ivy JA, et al (2015) Fish Hatchery Genetic Management Techniques: Integrating Theory with Implementation. North American Journal of Aquaculture 77:343–357.

Fitzpatrick MC, Keller SR (2015) Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. Ecology Letters 18:1–16.

Flanagan SP, Forester BR, Latch EK, et al (2018) Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation. Evolutionary Applications 11:1035–1052.

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

Forester BR, Lasky JR, Wagner HH, Urban DL (2018) Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. Molecular Ecology 27:2215–2233.

Frankham R (2005) Genetics and extinction. Biological Conservation 126:131–140.

Frankham R (1996) Relationship of genetic variation to population size in wildlife. Conservation Biology 10:1500–1508.

Frankham R (2008) Genetic adaptation to captivity in species conservation programs. Molecular Ecology 17:325–333.

Frankham R, Ballou J, Briscoe D (2010) Introduction to conservation genetics., 2nd edn. Cambridge University Press., Cambridge, UK ; New York

Frankham R, Ballou JD, Briscoe DA, McInnes KH (2002) Introduction to Conservation Genetics, 1st edn. Cambridge University Press

Frankham R, Bradshaw CJA, Brook BW (2014) Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. Biological Conservation 170:56–63.

Frankham R, Loebel DA (1992) Modeling problems in conservation genetics using captive Drosophila populations: Rapid genetic adaptation to captivity. Zoo Biol 11:333–342.

Frankham R, Manning H, Margan SH, Briscoe DA (2000) Does equalization of family sizes reduce genetic adaptation to captivity? Animal Conservation 3:357–363.

Franklin IR (1980) Evolutionary change in small populations. In: Soule ME, Wilcox BA (eds). Conservation Biology: An Evolutionary Ecological Perspective. Sinauer Associates, Massachusetts, p 135–140

Franklin IR, Frankham R (1998) How large must populations be to retain evolutionary potential? Animal Conservation 1:69–70.

Fraser DJ (2008) How well can captive breeding programs conserve biodiversity? A review of salmonids. Evolutionary applications 1:535–86.

Fraser DJ, Weir LK, Bernatchez L, et al (2011) Extent and scale of local adaptation in salmonid fishes: review and meta-analysis. Heredity 106:404–404

Funk WC, Lovich RE, Hohenlohe PA, et al (2016) Adaptive divergence despite strong genetic drift: genomic analysis of the evolutionary mechanisms causing genetic differentiation in the island fox (*Urocyon littoralis*). Molecular Ecology 25:2176–2194.

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012) Harnessing genomics for delineating conservation units. Trends Ecol Evol 27:489–496.

Fussi B, Westergren M, Aravanopoulos F, et al (2016) Forest genetic monitoring: an overview of concepts and definitions. Environmental Monitoring and Assessment 188:493.

Gaffney PM (2006) The role of genetics in shellfish restoration. Aquatic Living Resources 19:277–282.

Gaffney PM, Bernat CM, Allen SK (1993) Gametic incompatibility in wild and cultured populations of the eastern oyster, *Crassostrea virginica* (Gmelin). Aquaculture 115:273–284.

Galloway LF, Fenster CB (2000) Population differentiation in an annual legume: local adaptation. Evolution 54:1173–1181.

Galtsoff PS (1964) The American oyster, *Crassostrea virginica* Gmelin. Fish Bull 64:1–480

Garrison E, Marth G (2012) Haplotype-based variant detection from short-read sequencing. arXiv:12073907 [q-bio.GN]

Gilpin ME, Soule ME (1986) Minimum viable populations: processes of species extinction. In: Conservation Biology: The Science of Scarcity and Diversity (ed. Soule´ ME). Sinauer, Sunderland, MA, pp. 19–34.

Gold JR, Ma L, Saillant E, et al (2008) Genetic effective size in populations of hatchery-raised red drum released for stock enhancement. Transactions of the American Fisheries Society 137:1327–1334.

Gosselin T, Anderson EC, Bradbury IR (2016). assigner: Assignment analysis with GBS/RAD data using R.

Goudet J (2002). FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www2.unil.ch/popgen/softwares/fstat.htm Updated from Goudet (1995)

Goudet J (2005) hierfstat, a package for r to compute and test hierarchical F-statistics. Molecular Ecology Notes 5:184–186.

Gow JL, Tamkee P, Heggenes J, et al (2011) Little impact of hatchery supplementation that uses native broodstock on the genetic structure and diversity of steelhead trout revealed by a large-scale spatio-temporal microsatellite survey. Evolutionary Applications 4:763–782.

Grabowski JH, Brumbaugh RD, Conrad RF, et al (2012) Economic Valuation of Ecosystem Services Provided by Oyster Reefs. BioScience 62:900–909.

Grabowski JH, Peterson CH (2007) Restoring oyster reefs to recover ecosystem services. In: Theoretical Ecology Series. Elsevier, pp 281–298

Grant WS, Jasper J, Bekkevold D, Adkison M (2017) Responsible genetic approach to stock restoration, sea ranching and stock enhancement of marine fishes and invertebrates. Reviews in Fish Biology and Fisheries 27:615–649.

Griebeler EM, Seitz A (2002) An individual based model for the conservation of the endangered Large Blue Butterfly, *Maculinea arion* (Lepidoptera: Lycaenidae). Ecological Modelling 156:43–60.

Gruenthal KM, Witting DA, Ford T, et al (2014) Development and application of genomic tools to the restoration of green abalone in southern California. Conservation Genetics 15:109–121.

Günther T, Coop G (2013) Robust identification of local adaptation from allele frequencies. Genetics 195:205–220.

Guo B, DeFaveri J, Sotelo G, et al (2015) Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. BMC Biology 13:19.

Hagy JD, Boynton WR, Keefe CW, Wood KV (2004) Hypoxia in Chesapeake Bay, 1950–2001: Long-term change in relation to nutrient loading and river flow. Estuaries 27:634–658.

Hämmerli A, Reusch T (2002) Local adaptation and transplant dominance in genets of the marine clonal plant Zostera marina. Marine Ecology Progress Series 242:111–118.

Hanley TC, Hughes AR, Williams B, et al (2016) Effects of intraspecific diversity on survivorship, growth, and recruitment of the eastern oyster across sites. Ecology 97:1518–1529.

Hanski I, Saccheri I (2006) Molecular-level variation affects population growth in a butterfly metapopulation. PLoS Biology 4:e129.

Harding J, Mann R, Southworth M (2008) Shell Length-at-age Relationships in James River, Virginia, Oysters (*Crassostrea virginica*) Collected Four Centuries Apart. Journal of Shellfish Research 27:1109–1115.

Harding JM, Powell EN, Mann R, Southworth MJ (2013) Variations in eastern oyster (*Crassostrea virginica*) sex-ratios from three Virginia estuaries: protandry, growth and demographics. Journal of the Marine Biological Association of the United Kingdom 93:519–531.

Hare MP, Nunney L, Schwartz MK, et al (2011) Understanding and estimating effective population size for practical application in marine species management. Conservation Biology 25:438–449.

Hair JF Jr., Anderson RE, Tatham RL, Black WC (1995). Multivariate Data Analysis (3rd ed). New York: Macmillan.

Hargis WJ Jr., Haven DS (1999) Chesapeake oyster reefs, their importance, destruction and guidelines for restoring them, in: MW Luckenbach, R Mann and JA Wesson, editors. Oyster reef habitat restoration: a synopsis and synthesis of approaches. VIMS Press, Gloucester Point, VA.

Harris JA, Hobbs RJ, Higgs E, Aronson J (2006) Ecological restoration and global climate change. Restoration Ecology 14:170–176.

Hauser L, Adcock GJ, Smith PJ, et al (2002) Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). Proceedings of the National Academy of Sciences 99:11742–11747.

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

Haven DS, Hargis, WJ Jr., Kendall P (1981) The oyster industry of Virginia: Its status, problems and promise. 1981 (revised). Va. Inst. Mar. Sci., Spec. Pap. Mar. Sci. No.4: 1024 pp.

Hayes FN (2002) Natural History, Captive Management and Survival of the New Zealand Brown Teal. Brown Teal Conservation Trust publication 2002 146 pp.

Hayes PF, Menzel RW (1981) The reproductive system of early setting *Crassostrea virginica* (Gmelin) in the northern Gulf of Mexico, and its implications for population recruitment. The Biological Bulletin 160:80–88.

He X, Johansson M, Heath D (2016) Role of genomics and transcriptomics in selection of reintroduction source populations. Conservation Biology 30:1010–1018.

He Y, Ford SE, Bushek D, et al (2012) Effective population sizes of eastern oyster *Crassostrea virginica* (Gmelin) populations in Delaware Bay, USA. Journal of Marine Research 70:357–379.

Hedgecock D (1994) Does variance in reproductive success limit effective population sizes of marine organisms? Genetics and Evolution of Aquatic Organisms 122–134.

Hedgecock D, Chow V, Waples RS (1992) Effective population numbers of shellfish broodstocks estimated from temporal variance in allelic frequencies. Aquaculture 108:215–232.

Hedgecock D, Coykendall K (2007) Genetic Risks of Marine Hatchery Enhancement: The Good, the Bad, and the Unknown. In: Bert TM (ed) Ecological and Genetic Implications of Aquaculture Activities. Springer Netherlands, Dordrecht, pp 85–101

Hedgecock D, Pudovkin AI (2011) Sweepstakes reproductive success in highly fecund marine fish and shellfish: A review and commentary. Bulletin of Marine Science 87:971–1002.

Hedgecock D, Sly F (1990) Genetic drift and effective population sizes of hatchery-propagated stocks of the Pacific oyster, *Crassostrea gigas*. Aquaculture 88:21–38.

Hedrick PW (2005) Large variance in reproductive success and the Ne/N ratio. Evolution 438:753–754.

Hedrick PW, Kalinowski ST (2000) Inbreeding depression in conservation biology. Annual Review of Ecology and Systematics 31:139–162.

Heggenes J, Beere M, Tamkee P, Taylor EB (2006) Genetic diversity in steelhead before and after conservation hatchery operation in a coastal, boreal river. Transactions of the American Fisheries Society 135:251–267.

Helenurm K (1998) Outplanting and Differential Source Population Success in *Lupinus guadalupensis*. Conservation Biology 12:10

Hereford J (2009) A quantitative survey of local adaptation and fitness trade-offs. The American Naturalist 173:579–588.

Hice LA, Duffy TA, Munch SB, Conover DO (2012) Spatial scale and divergent patterns of variation in adapted traits in the ocean: Adaptive genetic variation in the ocean. Ecology Letters 15:568–575.

Hidu H, Haskin HH (1978) Swimming speeds of oyster larvae *Crassostrea virginica* in different salinities and temperatures. Estuaries 1:252.

Hill WG (1981) Estimation of effective population size from data on linkage disequilibrium. Genetical Research 38:209–216.

Hoban S, Bertorelle G, Gaggiotti OE (2012) Computer simulations: tools for population and evolutionary genetics. Nature Reviews Genetics 13:110–122.

Hoban S, Kelley JL, Lotterhos KE, et al (2016) Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. The American Naturalist 188:379–397.

Hoban SM, Mezzavilla M, Gaggiotti OE, et al (2013) High variance in reproductive success generates a false signature of a genetic bottleneck in populations of constant size: a simulation study. BMC Bioinformatics 14:309.

Holley JR, McComas KA, Hare MP (2018) Troubled waters: Risk perception and the case of oyster restoration in the closed waters of the Hudson-Raritan Estuary. Marine Policy 91:104–112.

Hooper DU, Chapin FS, Ewel JJ, et al (2005) Effects of Biodiversity on Ecosystem Functioning: A Consensus of Current Knowledge. Ecological Monographs 75:3–35.

Hornick KM, Plough LV (2019) Tracking genetic diversity in a large-scale oyster restoration program: effects of hatchery propagation and initial characterization of diversity on restored vs. wild reefs. Heredity 123:92–105.

Howell PE, Muths E, Hossack BR, et al (2018) Increasing connectivity between metapopulation ecology and landscape ecology. Ecology 99:1119–1128.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour 9:1322–1332.

Hufford KM, Mazer SJ (2003) Plant ecotypes: genetic differentiation in the age of ecological restoration. Trends in Ecology & Evolution 18:147–155.

Hughes AR, Hanley TC, Byers JE, et al (2019) Genetic diversity and phenotypic variation within hatchery-produced oyster cohorts predict size and success in the field. Ecological Applications 0:e01940.

Hughes AR, Hanley TC, Byers JE, et al (2017) Genetic by environmental variation but no local adaptation in oysters (*Crassostrea virginica*). Ecology and Evolution 7:697–709.

Hughes AR, Stachowicz JJ (2004) Genetic diversity enhances the resistance of a seagrass ecosystem to disturbance. Proceedings of the National Academy of Sciences 101:8998–9002.

Hunter RL, Markert CL (1957) Histochemical demonstration of enzymes separated by zone electrophoresis in starch gels. Science 125:1294–1295.

Ivanina AV, Kurochkin IO, Leamy L, Sokolova IM (2012) Effects of temperature and cadmium exposure on the mitochondria of oysters (*Crassostrea virginica*) exposed to hypoxia and subsequent reoxygenation. Journal of Experimental Biology 215:3142–3154.

Jackson JBC, Kirby MX, Berger WH, et al (2001) Historical collapse overfishing of and the recent coastal ecosystems. Science 293:629–638.

James G, Witten D, Hastie T, Tibshirani R (eds) (2013) An introduction to statistical learning: with applications in R. Springer, New York

Jaris H, Brown DS, Proestou DA (2019) Assessing the contribution of aquaculture and restoration to wild oyster populations in a Rhode Island coastal lagoon. Conservation Genetics 20:503–516.

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405.

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

Jones AT, Ovenden JR, Wang Y-G (2016) Improved confidence intervals for the linkage disequilibrium method for estimating effective population size. Heredity (Edinb) 117:217–223.

Jones HR, Johnson KM, Kelly MW (2019) Synergistic Effects of Temperature and Salinity on the Gene Expression and Physiology of Crassostrea virginica. Integrative and Comparative Biology 59:306–319.

Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. Mol Ecol Res 10:551–555

Jones TA (2013) Ecologically Appropriate Plant Materials for Restoration Applications. BioScience 63:211–219.

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology 16:1099–1106.

Karlsson S, Saillant E, Bumguardner BW, et al (2008) Genetic identification of hatchery-released red drum in Texas bays and estuaries. North American Journal of Fisheries Management 28:1294–1304.

Kaspar V, Kohlmann K, Vandeputte M, et al (2007) Equalizing sperm concentrations in a common carp (*Cyprinus carpio*) sperm pool does not affect variance in proportions of larvae sired in competition. Aquaculture 272:S204–S209.

Katalinas CJ, Brenkert K, Darden T, Denson MR (2017) A genetic assessment of a red drum, *Sciaenops ocellatus*, stock enhancement program. Journal of the World Aquaculture Society 49:523–539.

Katalinas CJ, Robinson JD, Strand AE, et al (2019) Forecasting the genetic influences of red drum stock enhancement in South Carolina. Fisheries Management and Ecology 26:224–235.

Keiner C (2011) The oyster question: Scientists, watermen, and the Maryland Chesapeake Bay since 1880. Isis 102:147–148.

Kemp W, Boynton W, Adolf J, et al (2005) Eutrophication of Chesapeake Bay: historical trends and ecological interactions. Marine Ecology Progress Series 303:1–29.

Kennedy VS (1989) The Chesapeake Bay oyster fishery: traditional management practices. In: J. F. Caddy, editor. Marine invertebrate fisheries: their assessment and management. New York, NY: Wiley. pp. 455–477.

Kennedy VS, Newell RIE, Eble AF (eds) (1996) The eastern oyster Crassostrea virginica. Maryland Sea Grant, College Park, MD

Kennedy VS, Breisch LL (1983) Sixteen Decades of Political Management of the Oyster Fishery in Maryland's Chesapeake Bay. Journal of Experimental Management 16:153–171.

Kennedy VS, Breitburg DL, Christman MC, et al (2011) Lessons Learned from Efforts to Restore Oyster Populations in Maryland and Virginia, 1990 to 2007. Journal of Shellfish Research 30:719–731.

Kennedy VS, Newell RIE, Eble AF, eds., (1996) The Eastern Oyster: *Crassostrea virginica*. Maryland Sea Grant, College Park, MD, 734 pp.

Kennedy VS, Sanford LP (1999) Characteristics of relatively unexploited beds of the eastern oyster, *Crassostrea virginica*, and early restoration programs. Pages 25-46 in Oyster reef habitat restoration: A synopsis and synthesis of approaches. MW Luckenbach, R Mann, JAWesson, eds. Gloucester Point, VA: Virginia Institute of Marine Science Press.

King TL, Ward R, Zimmerman EG (1994) Population Structure of Eastern Oysters (*Crassostrea virginica*) Inhabiting the Laguna Madre, Texas, and Adjacent Bay Systems. Can J Fish Aquat Sci 51:215–222.

Kirby MX (2004) Fishing down the coast: Historical expansion and collapse of oyster fisheries along continental margins. Proceedings of the National Academy of Sciences 101:13096–13099.

Kitada S (2018) Economic, ecological and genetic impacts of marine stock enhancement and sea ranching: A systematic review. Fish and Fisheries 19:511–532. https://doi.org/10.1111/faf.12271

Kopelman NM, Mayzel J, Jakobsson M, et al (2015) Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Molecular Ecology Resources 15:1179–1191.

Kozarewa I, Armisen J, Gardner AF, et al (2015) Overview of target enrichment strategies. Current Protocols in Molecular Biology 112:7.21.1-7.21.23.

Kramer-Schadt S, Revilla E, Wiegand T, Breitenmoser U (2004) Fragmented landscapes, road mortality and patch connectivity: modelling influences on the dispersal of Eurasian lynx: Lynx dispersal in fragmented landscapes. Journal of Applied Ecology 41:711–723.

Kuhner MK (2006) LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. Bioinformatics 22:768–770.

La Haye MJJ, Reiners TE, Raedts R, et al (2017) Genetic monitoring to evaluate reintroduction attempts of a highly endangered rodent. Conserv Genet 18:877–892.

Laikre L, Schwartz MK, Waples RS, Ryman N (2010) Compromising genetic diversity in the wild: Unmonitored large-scale release of plants and animals. Trends in Ecology and Evolution 25:520–529.

Laing I, Walker P, Areal F (2006) Return of the native – is European oyster (Ostrea edulis) stock restoration in the UK feasible? Aquatic Living Resources 19:283–287.

Lal MM, Southgate PC, Jerry DR, Zenger KR (2018) Genome-wide comparisons reveal evidence for a species complex in the black-lip pearl oyster *Pinctada margaritifera* (Bivalvia: Pteriidae). Sci Rep 8:191.

Lallias D, Taris N, Boudry P, et al (2010) Variance in the reproductive success of flat oyster Ostrea edulis L. assessed by parentage analysis in natural and experimental conditions. Genetics Research 92:175–187

Lande, R., 1995. Mutation and conservation. Conservation Biology 9:782–791.

Lane H, Michaelis A, Paynter K. (2011) Oyster restoration monitoring 2010. Submitted to the Oyster Recovery Partnership.

Le Luyer J, Laporte M, Beacham TD, et al (2017) Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon. Proceedings of the National Academy of Sciences 114:12964–12969.

Legendre P, Fortin M-J (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. Molecular Ecology Resources 10:831–844.

Lehnert SJ, DiBacco C, Van Wyngaarden M, et al (2019) Fine-scale temperature-associated genetic structure between inshore and offshore populations of sea scallop (*Placopecten magellanicus*). Heredity 122:69–80.

Leimu R, Fischer M (2008) A Meta-Analysis of Local Adaptation in Plants. PLoS ONE 3:e4010.

Leimu R, Mutikainen P, Koricheva J, Fischer M (2006) How general are positive relationships between plant population size, fitness and genetic variation? Journal of Ecology 94:942–952.

Lenihan HS, Peterson CH (1998) How habitat degradation through fishery disturbance through fishery disturbance enhances impacts hypoxia on oyster reefs. Ecological Applications 8:128–140.

Lesica P, Allendorf FW (1999) Ecological genetics and the restoration of plant communities: Mix or match? Restoration Ecology 7:42–50.

Letcher BH, Priddy JA, Walters JR, Crowder LB (1998) An individual-based, spatially-explicit simulation model of the population dynamics of the endangered red-cockaded woodpecker, *Picoides borealis*. Biological Conservation 86:1–14

Li A, Li L, Song K, et al (2017) Temperature, energy metabolism, and adaptive divergence in two oyster subspecies. Ecol Evol 7:6151–6162.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760.

Liaw A, Wiener M (2001) classification and regression by randomForest. R News 2:18–22.

Lind CE, Evans BS, Knauer J, et al (2009) Decreased genetic diversity and a reduced effective population size in cultured silver-lipped pearl oysters (*Pinctada maxima*). Aquaculture 286:12–19.

Linhart O, Rodina M, Gela D, et al (2005) Spermatozoal competition in common carp *(Cyprinus carpio*): What is the primary determinant of competition success? Reproduction 130:705–711.

Liu Q (1997) Variation partitioning by partial redundancy analysis (RDA). Environmetrics 8:75–85.

Lorenzen K, Agnalt A-L, Blankenship HL, et al (2013) Evolving context and maturing ccience: Aquaculture-based enhancement and restoration enter the marine fisheries management toolbox. Reviews in Fisheries Science 21:213–221.

Lorenzen K, Beveridge MCM, Mangel M (2012) Cultured fish: Integrative biology and management of domestication and interactions with wild fish. Biological Reviews 87:639–660.

Lotterhos KE, Whitlock MC (2015) The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. Molecular Ecology 24:1031–1046.

Lotze HK, Lenihan HS, Bourque BJ, et al (2006) Depletion, Degradation, and Recovery Potential of Estuaries and Coastal Seas. Science 312:1806.

Luikart G, Sherwin WB, Steele BM, Allendorf FW (1998) Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. Molecular Ecology 7:963–974.

Luu K, Bazin E, Blum MGB (2017) pcadapt: an R package to perform genome scans for selection based on principal component analysis. Molecular Ecology Resources 17:67–77.

Lynch M, Ritland K (1999) Estimation of pairwise relatedness with molecular markers. Genetics 152:1753–1766

Manel S, Gugerli F, Thuiller W, et al (2012) Broad-scale adaptive genetic variation in alpine plants is driven by temperature and precipitation. Molecular Ecology 21:3729–3738.

Mann R, Evans DA (1998) Estimation of oyster Crassostrea virginica standing stock, larval production and advective loss in relation to observed recruitment in the James River, Virginia. J Shellfish Res 17:239–253

Mann R, Powell EN (2007) Why oyster restoration goals in the Chesapeake Bay are not and probably cannot be achieved. Journal of Shellfish Research 26:905–917.

Mann R, Southworth M, Carnegie RB, Crockett RK (2014) Temporal variation in fecundity and spawning in the eastern oyster, *Crassostrea virginica*, in the Piankatank River, Virginia. Journal of Shellfish Research 33:167–176.

Marandel F, Lorance P, Berthelé O, et al (2019) Estimating effective population size of large marine populations, is it feasible? Fish and Fisheries 20:189–198.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 5:639–55.

Maryland Department of Natural Resources (MD DNR) (2016). Oyster management review: 2010-2015. Prepared by Maryland Department of Natural Resources, Draft Report July 2016. 128 pp. Available at: https://dnr.maryland.gov/fisheries/Documents/FiveYearOysterReport.pdf

McFarland K, Plough LV, Nguyen M, Hare MP (2020) Are bivalves susceptible to domestication selection? Using starvation tolerance to test for potential trait changes in eastern oyster larvae. bioRxiv. 02.26.966176

McKay JK, Christian CE, Harrison S, Rice KJ (2005) "How local is local?"-A review of practical and conceptual issues in the genetics of restoration. Restoration Ecology 13:432–440.

McLusky DS, Elliott M (2007) Transitional waters: A new approach, semantics or just muddying the waters? Estuarine, Coastal and Shelf Science 71:359–363.

McPhee ME (2004) Generations in captivity increases behavioral variance: considerations for captive breeding and reintroduction programs. Biological Conservation 115:71–77.

Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: Two programs for the analysis of genetic diversity of asexual organisms. Molecular Ecology Notes 4:792–794.

Melampy MN, Howe HF (1977) Sex ratio in the tropical tree *Triplaris americana* (Polygonaceae). Evolution 31:867–872.

Mendelson JR, Lips KR, Gagliardo RW, et al (2006) Confronting amphibian declines and extinctions. Science 313:48–48.

Miller AD, Hoffmann AA, Tan MH, et al (2019) Local and regional scale habitat heterogeneity contribute to genetic adaptation in a commercially important marine mollusc (*Haliotis rubra*) from southeastern Australia. Molecular Ecology 28:3053–3072.

Mims MC, Day CC, Burkhart JJ, et al (2019) Simulating demography, genetics, and spatially explicit processes to inform reintroduction of a threatened char. Ecosphere 10:e02589.

Montalvo AM, Ellstrand NC (2001) Nonlocal transplantation and outbreeding depression in the subshrub *Lotus scoparius* (Fabaceae). Am J Bot 88:258–269.

Montalvo AM, Rice SLW, Buchmann SL, et al (1997) Restoration biology: A population biology perspective. Restoration Ecology 5:277–290.

Morvezen R, Boudry P, Laroche J, Charrier G (2016) Stock enhancement or sea ranching? Insights from monitoring the genetic diversity, relatedness and effective population size in a seeded great scallop population (*Pecten maximus*). Heredity 117:142–148.

Mroch RM, Eggleston DB, Puckett BJ (2012) Spatiotemporal variation in oyster fecundity and reproductive output in a network of no-take reserves. Journal of Shellfish Research 31:1091–1101.

Munro JL, Bell JD (1997) Enhancement of marine fisheries resources. Reviews in Fisheries Science 5:185–222.

Myers RA, Worm B (2003) Rapid worldwide depletion of predatory fish communities precipitated marked changes in coastal ecosystem structure and function. Nature 423:280–283

Narváez DA, Klinck JM, Powell EN, et al (2012) Modeling the dispersal of eastern oyster (*Crassostrea virginica*) larvae in Delaware Bay. Journal of Marine Research 70:381–409.

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Nat Acad Sci USA 70:3321–3323

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, NY

Neraas LP, Spruell P (2001) Fragmentation of riverine systems: the genetic effects of dams on bull trout (*Salvelinus confluentus*) in the Clark Fork River system. Molecular Ecology 10:1153–1164.

Nestlerode JA, Luckenbach MW, O?Beirn FX (2007) Settlement and survival of the oyster *Crassostrea virginica* on created oyster reef habitats in Chesapeake Bay. Restoration Ecology 15:273–283.

Newell RIE (1988) Ecological changes in Chesapeake Bay: are they the result of overharvesting the American oyster, Crassostrea virginica? In: Understanding the estuary. Advances in Chesapeake Bay research. US EPA CBP/TRS 24/88. CRC Publication 129. Chesapeake Bay Consortium, Solomons, MD, p 536–546

Newell RIE, Kennedy VS, Manuel JL Meritt D (2005) Behavioral responses of *Crassostrea ariakensis* and *Crassostrea virginica* larvae to environmental change under spatially realistic conditions. Final report to Maryland Department of Natural Resources, May 2005.

Newkirk G (1978) Interaction of genotype and salinity in larvae of the oyster *Crassostrea virginica*. Mar Biol 48:227–234.

North E, Schlag Z, Hood R, et al (2008) Vertical swimming behavior influences the dispersal of simulated oyster larvae in a coupled particle-tracking and hydrodynamic model of Chesapeake Bay. Marine Ecology Progress Series 359:99–115.

O'Beirn FX, Heffernan PB, Walker RL, Jansen ML (1996) Young-of-the-Year Oyster, Crassostrea virginica, Reproduction in Coastal Georgia. Estuaries 19:651.

Oliehoek PA, Windig JJ, van Arendonk JAM, Bijma P (2006) Estimating relatedness between individuals in general populations with a focus on their use in conservation programs. Genetics 173:483–496

Oksanen J (2015) Vegan: an introduction to ordination. http://cran.r-project.org/web/packages/vegan/vignettes/intro-vegan.pdf.

Oota T, Matsuishi T (2005) Increase of inbreeding by stocking on wild population assessed by using in dividual-based life history model. Fisheries Science 71:73–78.

Palstra FP, Ruzzante DE (2008) Genetic estimates of contemporary effective population size: What can they tell us about the importance of genetic stochasticity for wild population persistence? Molecular Ecology 17:3428–3447.

Paraso MC, Ford SE, Powell EN, et al (1999) Modeling the MSX parasite in eastern oyster (*Crassostrea virginica*) populations. II. Salinity effects. Journal of Shellfish Research 18:501–516

Paynter KT, Lane H, Michaelis, A. (2013) Paynter lab annual monitoring and research summary 2012. Submitted to the Oyster Recovery Partnership.

Paynter KT, Michaelis A, Handschy A. (2014) Paynter lab annual monitoring and research summary 2013. Submitted to the Oyster Recovery Partnership

Paynter KT, Politano V, Lane HA, et al (2010) Growth rates and prevalence of *Perkinsus marinus* in restored oyster populations in Maryland. Journal of Shellfish Research 29:309–317.

Peng B, Amos CI (2008) Forward-time simulations of non-random mating populations using simuPOP. Bioinformatics 24:1408–1409.

Peng B, Amos CI, Kimmel M (2007) Forward-Time Simulations of Human Populations with Complex Diseases. PLoS Genetics 3:e47.

Peng B, Kimmel M (2005) simuPOP: a forward-time population genetics simulation environment. Bioinformatics 21:3686–3687.

Peterson BK, Weber JN, Kay EH, et al (2012) Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS ONE 7:e37135.

Pew J, Muir PH, Wang J, Frasier TR (2015) related: an R package for analysing pairwise relatedness from codominant molecular markers. Molecular Ecology Resources 15:557–561.

Pierce SK, Amende LM (1981) Control mechanisms of amino acid-mediated cell volume regulation in salinity-stressed molluscs. J Exp Zool 215:247–257.

Pizzitutti F, Pan W, Feingold B, et al (2018) Out of the net: An agent-based model to study human movements influence on local-scale malaria transmission. PLOS ONE 13:e0193493.

Plough LV (2016) Genetic load in marine animals: A review. Current Zoology 62:567–579.

Plough LV (2018) Fine-scale temporal analysis of genotype-dependent mortality at settlement in the Pacific oyster Crassostrea gigas. J Exp Mar Biol Ecol 501:90–98

Plough LV, Shin G, Hedgecock D (2016) Genetic inviability is a major driver of type III survivorship in experimental families of a highly fecund marine bivalve. Molecular Ecology 25:895–910.

Powell EN, Cummins H (1985) Are molluscan maximum life spans determined by long-term cycles in benthic communities? Oecologia 67:177–182

Powell EN, Klinck J, Hofmann EE (1996) Modeling diseased oyster populations. II. Triggering mechanisms for Perkinsus marinus epizootics. Journal of Shellfish Research 15:141–165

Powell EN, Morson JM, Ashton-Alcox KA, Kim Y (2013) Accommodation of the sex ratio in eastern oysters Crassostrea virginica to variation in growth and mortality across the estuarine salinity gradient in Delaware Bay. J Mar Biol Assoc UK 93:533–555

Pritchard JK, Stephens M, Donnelly P (2000) Inference of Population Structure Using Multilocus Genotype Data. Genetics 155:945

Puritz JB, Hollenbeck CM, Gold JR (2014) dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. PeerJ 2:e431

R Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/

R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/

Ralls K, Ballou JD (2004) Genetic status and management of California condors. The Condor 106:215–228.

Reeb CA, Avise JC (1990) A genetic discontinuity in a continuously distributed species: mitochondrial DNA in the American oyster, Crassostrea virginica. Genetics 124:397–406

Reece KS, Ribeiro WL, Gaffney PM, Carnegie RB, Allen SK (2004) Microsatellite marker development and analysis in the eastern oyster (*Crassostrea virginica*):

Confirmation of null alleles and non-mendelian segregation ratios. J Hered 95:346–352

Rellstab C, Gugerli F, Eckert AJ, et al (2015) A practical guide to environmental association analysis in landscape genomics. Molecular Ecology 24:4348–4370.

Reusch TBH, Ehlers A, Hammerli A, Worm B (2005) Ecosystem recovery after climatic extremes enhanced by genotypic diversity. Proceedings of the National Academy of Sciences 102:2826–2831.

Rice KJ, Emery NC (2003) Managing microevolution: restoration in the face of global change. Frontiers in Ecology and the Environment 1:469–478.

Rice WR (1989) Analyzing Tables of Statistical Tests. Evolution 43: 223–225.

Ritland K (1996) Estimators for pairwise relatedness and individual inbreeding coefficients. Genetical Research 67:175–185.

Rodney WS, Paynter KT (2006) Comparisons of macrofaunal assemblages on restored and non-restored oyster reefs in mesohaline regions of Chesapeake Bay in Maryland. Journal of Experimental Marine Biology and Ecology 335:39–51.

Rose CG (2008) Population genetics of the eastern oyster in Chesapeake Bay. Doctor in Philosophy Thesis University of Maryland, College Park (http://drum.lib.umd.edu/ bitstream/1903/8994/1/Rose_umd_0117E_10044.pdf).

Rose CG, Paynter KT, Hare MP (2006) Isolation by distance in the eastern oyster, Crassostrea virginica, in Chesapeake Bay. Journal of Heredity 97:158–170.

Rothschild BJ, Ault JS, Goulletquer P, Heral M (1994) Decline of the Chesapeake Bay oyster population: A century of habitat destruction and overfishing. Marine Ecology Progress Series 111:29–40.

Ryman N, Jorde PE, Laikre L (1995) Supportive breeding and variance effective population size. Conservation Biology 9:1619–1628.

Ryman N, Laikre L (1991) Effects of supportive breeding on the genetically effective population size. Conservation Biology 5:325–329

Saccheri I, Kuussaari M, Kankare M, et al (1998) Inbreeding and extinction in a butterfly metapopulation. Nature 392:491–494.

Saltelli A, Chan K, Scott EM (eds) (2000) Sensitivity analysis. Wiley, Chichester ; New York

Sanford E, Kelly MW (2011) Local Adaptation in Marine Invertebrates. Annual Review of Marine Science 3:509–535.

Scharping R, Plough L, Meritt D, North E (2019) Low-salinity tolerance of early-stage oyster larvae from a mesohaline estuary. Marine Ecology Progress Series 613:97–106.

Schlotterer C, Ritter R, Harr B, Brem G (1998) High mutation rate of a long microsatellite allele in Drosophila melanogaster provides evidence for allele-specific mutation rates. Molecular Biology and Evolution 15:1269–1274.

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology 18:233–234.

Schuhmacher H, Schillak L (1994) Integrated electrochemical and biogenic deposition of hard material - a natural-like colonization substrate. Bulletin of Marine science 55:672–679.

Scribner KT, Lowe WH, Landguth E, et al (2016) Applications of genetic data to improve management and conservation of river fishes and their habitats. Fisheries 41:174–188.

Seddon PJ, Soorae PS, Launay F (2005) Taxonomic bias in reintroduction projects. Animal Conservation 8:51–58.

Segovia-Viadero M, Serrão EA, Canteras-Jordana JC, Gonzalez-Wangüemert M (2016) Do hatchery-reared sea urchins pose a threat to genetic diversity in wild populations? Heredity.

Serbezov D, Jorde PE, Bernatchez L, et al (2012) Life history and demographic determinants of effective/census size ratios as exemplified by brown trout (*Salmo trutta*): Life history and demographic determinants of effective/census size ratios. Evolutionary Applications 5:607–618.

She Z, Li L, Meng J, et al (2018) Population resequencing reveals candidate genes associated with salinity adaptation of the Pacific oyster *Crassostrea gigas.* Scientific Reports 8:8683.

Shumway SE (1996) Natural environmental factors. In: Kennedy VS, Newell RIE, and Eble AE, eds. The Eastern Oyster: Crassostrea virginica. Maryland Sea Grant College, College Park, Maryland. 734 pp.

Shumway S, Koehn R (1982) Oxygen consumption in the American oyster *Crassostrea virginica*. Mar Ecol Prog Ser 9:59–68.

Silliman K (2019) Population structure, genetic connectivity, and adaptation in the Olympia oyster (*Ostrea lurida*) along the west coast of North America. Evolutionary Applications 12:923–939.

Skaala Ø, Besnier F, Borgstrøm R, et al (2019) An extensive common-garden study with domesticated and wild Atlantic salmon in nature reveals impact on smolt production and shifts in fitness traits. Evolutionary Applications 1–16.

Snyder NFR, Derrickson SR, Beissinger SR, et al (1996) Limitations of Captive Breeding in Endangered Species Recovery. Conservation Biology 10:338–348.

Sokal RR (1979) Testing statistical significance of geographic variation patterns. Systematic Zoology 28:227.

Sokolova IM (2004) Cadmium effects on mitochondrial function are enhanced by elevated temperatures in a marine poikilotherm, *Crassostrea virginica* Gmelin (Bivalvia: Ostreidae). Journal of Experimental Biology 207:2639–2648.

Stapley J, Reger J, Feulner PGD, et al (2010) Adaptation genomics: the next generation. Trends in Ecology & Evolution 25:705–712.

Stoffel MA, Esser M, Kardos M, et al (2016) inbreedR: an R package for the analysis of inbreeding based on genetic markers. Methods Ecol Evol 7:1331–1339.

Strand AE (2002) metasim 1.0: an individual-based environment for simulating population genetics of complex population dynamics. Mol Ecol Notes 2:373–376.

Syvänen A-C (2005) Toward genome-wide SNP genotyping. Nature Genetics 37:S5–S10.

Taft J, Taylor W, Hartwig E, Loftus R (1980) Seasonal oxygen depletion in Chesapeake Bay. Estuaries and Coasts 3:242–247.

Taris N, Ernande B, Mccombie H, Boudry P (2006) Phenotypic and genetic consequences of size selection at the larval stage in the Pacific oyster (*Crassostrea gigas*). Journal of Experimental Marine Biology and Ecology 333:147–158

Tarnowski, M. T. (2016). Maryland oyster population status report. Annapolis, MD: Maryland Department of Natural Resources.

Thongda W, Zhao H, Zhang D, et al (2018) Development of SNP panels as a new tool to assess the genetic diversity, population structure, and parentage analysis of the eastern oyster (*Crassostrea virginic*a). Marine Biotechnology 20:385–395.

Turley B, Reece K, Shen J, et al (2019) Multiple drivers of interannual oyster settlement and recruitment in the lower Chesapeake Bay. Conservation Genetics 20:1057–1071.

U.S. Army Corps of Engineers (USACE) (2012). Chesapeake Bay oyster recovery: native oyster restoration master plan, Maryland and Virginia. Prepared by U.S. Army

Corps of Engineers, Baltimore and Norfolk Districts, 290 pp. Available at: http://www.chesapeakebay.net/channel_files/18195/cb_oystermasterplan_march2012_low-res.pdf

van Katwijk MM, Bos AR, de Jonge VN, et al (2009) Guidelines for seagrass restoration: Importance of habitat selection and donor population, spreading of risks, and ecosystem engineering effects. Marine Pollution Bulletin 58:179–188.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4:535–538.

Vellend M, Geber MA (2005) Connections between species diversity and genetic diversity: Species diversity and genetic diversity. Ecology Letters 8:767–781.

Vendrami DLJ, De Noia M, Telesca L, et al (2019) RAD sequencing sheds new light on the genetic structure and local adaptation of European scallops and resolves their demographic histories. Scientific Reports 9:7455.

Vilà C, Sundqvist A, Flagstad Ø, et al (2003) Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant. Proceedings of the Royal Society of London Series B: Biological Sciences 270:91–97.

Waal SD, Balkhair M, Al-Mashikhi A, Khoom S (2013) Investigating the translocation and seeding of wild *Haliotis mariae* wood, 1828, in the sultanate of Oman. Journal of Shellfish Research 32:315–323.

Wallace RK, Waters P, Rikard FS (2008) Oyster Hatchery Techniques. Publication 4302, Southern Region Aquaculture Center, Stoneville, MS

Walters JR, Derrickson SR, Michael Fry D, et al (2010) Status of the California condor (*Gymnogyps californianus*) and efforts to achieve its recovery. The Auk 127:969–1001.

Wang IJ, Bradburd GS (2014) Isolation by environment. Molecular Ecology 23:5649–5662.

Wang J (2011) COANCESTRY: a program for simulating, estimating and analyzing relatedness and inbreeding coefficients. Mol Ecol Res 11:141–145

Wang J (2017) Estimating pairwise relatedness in a small sample of individuals. Heredity 119:302–313. https://doi.org/10.1038/hdy.2017.52

Wang J, Ryman N (2001) Genetic effects of multiple generations of supportive breeding. Conservation Biology 15:1619–1631

Wang YP, Shi YH, Guo XM (2009). Identification and characterization of 66 EST-SR Markers in the eastern oyster Crassostrea virginica (Gmelin). J Shellfish Res 28:227–234

Wang Y, Guo X (2007) Development and characterization of EST-SSR markers in the eastern oyster Crassostrea virginica. Marine Biotechnology 9:500–511.

Wang J, Zamar R, Marazzi A, Yohai V, Salibian-Barrera M, Maronna R, Zivot E, Rocke D, Martin D, Maechler M, Konis K (2014). robust: Robust Library. R package version 0.4-16, URL http://CRAN.R-project.org/package=robust.

Waples R, Luikart G, Faulkner JR, Tallmon DA (2013) Simple life-history traits explain key effective population size ratios across diverse taxa. Proceedings of the Royal Society B: Biological Sciences 280:20131339.

Waples RK, Larson WA, Waples RS (2016a) Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. Heredity 117:233–240.

Waples RS (2016) Tiny estimates of the Ne/N ratio in marine fishes: Are they real? Journal of Fish Biology 89:2479–2504.

Waples RS (1991) Genetic interactions between hatchery and wild salmonids: lessons from the Pacific Northwest. Canadian Journal of Fisheries and Aquatic Sciences 48:124–133. https://doi.org/10.1139/f91-311

Waples RS (2005) Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? Molecular Ecology 14:3335–3352.

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conservation Genetics 7:167–184.

Waples RS, Antao T, Luikart G (2014) Effects of overlapping generations on linkage disequilibrium estimates of effective population size. Genetics 197:769–780.

Waples RS, Do C (2008) LDNE: A program for estimating effective population size from data on linkage disequilibrium. Molecular Ecology Resources 8:753–756.

Waples RS, Do C (1994) Genetic risk associated with supplementation of Pacific Salmonids: captive broodstock programs. Canadian Journal of Fisheries and Aquatic Sciences 51:310–329.

Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. Evol Appl 3:244–262.

Waples RS, Hindar K, Karlsson S, Hard JJ (2016b) Evaluating the Ryman–Laikre effect for marine stock enhancement and aquaculture. Current Zoology 62:617–627.

Wedekind C, Rudolfsen G, Jacob A, et al (2007) The genetic consequences of hatchery-induced sperm competition in a salmonid. Biological Conservation 137:180–188.

Waples RS, Waples RK (2011) Inbreeding effective population size and parentage analysis without parents. Mol Ecol Res 11:162–171

Weeks AR, Sgro CM, Young AG, et al (2011) Assessing the benefits and risks of translocations in changing environments: a genetic perspective: Translocations in changing environments. Evolutionary Applications 4:709–725.

Weir BS (1996) Genetic data analysis II: Methods for discrete population genetic data. Sinauer Associates, Sunderland, MA.

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. Evolution 38:1358–1370.

Westby SR, Bruce DB, Slacum W, French E, Sowers A, Weissberger E. (2016) 2015 oyster restoration implementation update. Silver Spring, MD: NOAA.

Westby SR, Sowers A, Slacum W, Paynter K (2017) 2016 Oyster reef monitoring report. NOAA, Silver Spring, MD

Whitlock MC, Lotterhos KE (2015) Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of FST. The American Naturalist 186:S24–S36.

Wilberg MJ, Livings ME, Barkman JS, et al (2011) Overfishing, disease, habitat loss, and potential extirpation of oysters in upper Chesapeake Bay. Marine Ecology Progress Series 436:131–144.

Willi Y, Van Buskirk J, Hoffmann AA (2006) Limits to the Adaptive Potential of Small Populations. Annual Review of Ecology, Evolution, and Systematics 37:433–458.

Willis SC, Hollenbeck CM, Puritz JB, et al (2017) Haplotyping RAD loci: an efficient method to filter paralogs and account for physical linkage. Molecular Ecology Resources 17:955–965.

Winemiller KO, Rose KA (1992) Patterns of Life-History Diversification in North American Fishes: implications for Population Regulation. Canadian Journal of Fisheries and Aquatic Sciences 49:2196–2218.

Wisely SM, McDonald DB, Buskirk SW (2003) Evaluation of the genetic management of the endangered black-footed ferret (*Mustela nigripes*). Zoo Biology 22:287–298.

Withler RE, Beacham TD (1994) Genetic consequences of the simultaneous or sequential addition of semen from multiple males during hatchery spawning of chinook salmon (*Oncorhynchus tshawytscha*). Aquaculture 126:11–23.

Wright S (1931) Evolution in Mendelian Populations. Genetics 16:97–159

Yiming L, Zhongwei G, Qisen Y, et al (2003) The implications of poaching for giant panda conservation. Biological Conservation 111:125–136.

Zhao X, Yu H, Kong L, Li Q (2012) Transcriptomic Responses to Salinity Stress in the Pacific Oyster Crassostrea gigas. PLoS ONE 7:e46244.

Zuur AF, Ieno EN, Elphick CS (2010) A protocol for data exploration to avoid common statistical problems: Data exploration. Methods in Ecology and Evolution 1:3–14.