2021

A Quantitative Genetic Analysis Of Commercial Traits In Polyploid Crassostrea Virginica, With An Evaluation Of Strategies For Genetic Improvement Of Triploids

Joseph L. Matt
William & Mary - Virginia Institute of Marine Science, josephmatt13@gmail.com

A Quantitative Genetic Analysis of Commercial Traits in Polyploid *Crassostrea virginica*, with an Evaluation of Strategies for Genetic Improvement of Triploids

A Dissertation
Presented to

The Faculty of the School of Marine Science

William & Mary

In Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

by

Joseph L Matt

August 2021

APPROVAL PAGE

This dissertation is submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

Joseph L Matt

Approved by Committee, August 2021

Stan Allen, Ph.D.
Committee Chair / Advisor

Jan McDowell, Ph.D.

Jessica Moss Small, Ph.D.

Emily Rivest, Ph.D.

Louis Plough, Ph.D.
University of Maryland Center for Environmental Science
Cambridge, MD

Dedicated to Dr. Dennis Matt. I got my Ph.D., Dad.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Thank you, Stan. You guided me through many great opportunities, including my PhD program.

Thank you, Jan McDowell, Jessica Moss Small, Emily Rivest, and Louis Plough for your guidance and teachings during my PhD program.

Thank you, Jessica Moss Small, for your guidance and support in all aspects of this project, including field logistics, data management, and data analysis.

Thank you, Peter Kube, for your time and patience in teaching me much about quantitative genetics and statistical analysis.

Thank you, Jen Hay, Cathy Cake, and Linda Schaffner, for your support during my time at VIMS.

Thank you, Eric Guévélou. You were always there to support, teach, and collaborate.

Thank you, ABC: Amanda Chesler-Poole, Lauren Gregg, Karen Sisler, Kate Ritter Sage, Shelley Katsuki, Joana Sousa, Nate Geyerhahn, and Kemarin Kim. Much of the hatchery and field work in this dissertation were done by this staff of world-class aquaculturists.

Thank you, Nate Geyerhahn. Our comradery and discussions were a special support for me. You cared and taught me, and you guided me through some major challenges.

Thank you, Debbrah Pelata, for helping me manage the funds for the project and my travel.

Thank you, Julie Krask, for your regular support and care for me during this program.

Thank you, Pasha Maher, Josh Shenker, Pierce Few, and Will Brennan. You listened, cared, and helped me when things were tough.

Thank you, Stephanie Wilson. You were there for me. You were always compassionate, understanding, and helpful. You kept me afloat.

Thank you, Mom, for your warmth and love.

Thank you, Dad. You have been the ultimate advocate. You regularly encouraged and inspired me and put time in to help me with my writing.

ABSTRACT PAGE

Triploids are a popular product in commercial oyster aquaculture and make up most of the hatchery-produced *Crassostrea virginica* farmed in the Chesapeake Bay. Despite their importance to commercial aquaculture, the potential of genetically improving triploid *C. virginica* from selective breeding and breeding strategies for their improvement had not been evaluated. In this dissertation, the prospect of improving triploid *C. virginica* through selective breeding was assessed with a quantitative genetic analysis from a field test, and breeding strategies for genetically improving triploids were compared by computer simulation.

Heritability and genetic correlations involving commercial traits in triploids, including mass mortality associated with late spring conditions, or "triploid mortality," were estimated from twenty paternal half-sib triploid families and forty full-sib tetraploid families reared at three sites in the Chesapeake Bay. A triploid mortality event only occurred at a site on the bayside of the Eastern Shore of Virginia (Nandua Creek), with three triploid families having survival less than 0.70 between May 7 and July 9. The heritability of survival in triploid families during the triploid mortality event was high (1.06 ± 0.32), suggesting selective breeding can reduce the risk of these mortalities in the future. Genetic correlations between survival in triploids at Nandua Creek and the other two sites, York River and Choptank River, were low (0.46 ± 0.22, 0.46 ± 0.24), indicating a weak relationship between genes causing "triploid mortality" and genes causing mortality at York River and Choptank River. Heritability for total weight, meat weight, and shape traits in triploids was often high (> 0.30) and higher than that previously reported for diploid *C. virginica*. Genetic correlations between traits in triploids and tetraploids were always positive and ranged from 0.30 to 1. Although the positive genetic correlations indicate that selecting for genetic improvement of tetraploids will also lead to genetic improvement in triploids, the estimates had high standard errors, leaving the strength of the relationship unclear.

Breeding strategies for genetically improving triploids were compared by simulation with a focus on the effect of genetic correlations between ploidies. The strategies were 1) separate diploid and tetraploid family breeding programs, without phenotyping triploids and 2) a single family breeding program phenotyping diploids, triploids, and tetraploids. The strategy of phenotyping all ploidies resulted in more genetic improvement of triploids when between-ploidy genetic correlations were low (0.33 – 0.66), and the two strategies had similar results at higher genetic correlations (0.75 – 0.90). The higher or similar improvement of triploids across moderate genetic correlations suggests the single breeding program is the better approach, especially if robust estimates of between-ploidy genetic correlations are unavailable.

Potential exists to genetically improve triploid *C. virginica* in the Chesapeake Bay through selective breeding, including reducing the risk of "triploid mortality." Phenotyping diploids, tetraploids, and triploids in a single breeding program is likely to yield the highest possible improvement in triploids if using family selection. Future studies should assess the benefit of applying genomic selection to polyploid oyster breeding. Genomic selection may be highly advantageous for improving triploids because it could enable identification of individual diploids and tetraploids that have the highest genetic value for triploid production.

A Quantitative Genetic Analysis of Commercial Traits in Polyploid *Crassostrea virginica*, with an Evaluation of Strategies for Genetic Improvement of Triploids

Chapter 1: Introduction

Oyster aquaculture is an environmentally sustainable form of food production that can provide ecosystem services and economic growth to coastal communities. Farming oysters requires no feed or clearing of land, both major sources of green-house gas emissions from other sources of animal protein, and oyster farming produces less than 0.5% of the greenhouse gas emissions of terrestrial livestock farming (Ray et al., 2019). Additionally, farming bivalves like oysters provides ecosystem services such as nutrient removal, habitat provisioning, and water filtration (Alleway et al., 2019; Shumway, 2011). Many coastal communities have lost the ecosystem and economic benefit once provided by wild oyster populations. In many regions, including regions in North America, Australia, and Europe, wild oysters are at less than 1% of their historical abundance (Beck et al., 2011). Farming oysters in regions with depleted wild populations can help restore the ecosystem services once provided by oysters and revive industries that have dwindled with the collapse of the wild fisheries.

A promising form of aquaculture for growth and economic viability is hatchery-based aquaculture. In hatchery-based aquaculture, animals are propagated in a hatchery rather than captured in the wild, meaning the number of animals available for farming is not constrained by the abundance of wild populations. Additionally, hatchery-based aquaculture allows control over the breeding and thereby influence over the genetic constitution of a controlled population. By affecting the genetics of a controlled population, desirable characteristics in the animals can be expressed, which increase the economic viability of farming by reducing the cost of production or increasing the value of the product.

A widely applied method of affecting the genetics of aquaculture species is polyploid induction. Polyploidy is the condition of having three or more chromosome sets and it occurs

naturally in many species of plants and some species of animals (Otto and Whitton, 2000). Many fish and shellfish species have been induced to be polyploid for research and commercial evaluation (reviews by Beaumont and Fairbrother, 1991; Thorgaard, 1986). A common goal of polyploid induction for aquaculture is the production of sterile animals. In theory, sterility could bypass poor performance associated with sexual development, such as slower growth, reduced flesh quality, and higher mortality (Lincoln et al., 1974; Refstie et al., 1977; Stanely et al., 1981). Sterility in animals was expected to manifest from triploidy, both because of the sterility found in triploid plants (e.g., Kihara, 1951), and the expectation that an odd number of chromosome sets would cause meiotic dysfunction. Triploidy has been induced in many species of fish and shellfish with varying results among species, from partial to total sterility, and has only become commercially useful in a few species, such as some species of trout (*Onchorhynchus spp.*) and oysters (*Crassostrea* spp.) (Piferrer et al., 2009).

Triploid oysters were originally induced by chemical treatment of fertilized eggs. Stanley et al. (1981) induced triploidy in *C. virginica* by treating fertilized eggs with the fungal antibiotic cytochalasin B (CB), which blocks the extrusion of a polar body during meiosis. Induction of triploidy by CB was later applied to other oyster species, including *Crassostrea gigas* (Downing and Allen, 1987) and *Saccostrea glomerata* (Nell et al., 1994). Chemically induced triploids were found to have lower fecundity than diploids (Allen and Downing, 1990; Barber and Mann, 1991; Cox et al., 1996) and to maintain their meat weight when meat weight in diploid oysters decreased as a consequence of spawning (Allen, 1988; Allen and Downing, 1986). The high meat quality of triploid oysters in the spawning season suggested they could be marketable when diploids were not, making triploid oysters a commercially valuable option (Allen, 1988). Although chemically induced triploids became a substantial portion of commercial

production in the Pacific Northwest of the United States in the late 1980s (Allen, 1988), there were some limitations with the commercialization of chemically induced triploids. Most notably, the treatment rarely produced populations that were 100% triploid, instead often containing a mix of triploids and diploids (e.g., Stanley et al., 1981). Triploid populations with any percentage of diploid "contamination" could reduce the commercial value of the population (assuming triploids are the more valuable product), as well as warrant costly re-attempts of chemical induction to produce populations with a higher percentage of triploids. A method that would consistently produce populations that were 100% triploid was in demand to eliminate the inefficiencies of chemical induction.

More reliable production of triploid oysters became possible with the induction of tetraploid oysters. Tetraploids were of interest to commercial aquaculture because of the possibility of crossing diploids to tetraploids to create triploid progeny, which had been demonstrated in plants (Kihara, 1951) and invertebrates (Astaurov, 1969). Guo and Allen (1994) first demonstrated production of tetraploid oysters by chemical induction, then Guo et al. (1996) demonstrated 100% triploid progeny from crossing tetraploids to diploids. Like chemically induced triploids, triploids made from crossing diploids to tetraploids, or "mated triploids," have been found to have reduced fecundity (Guévélou et al., 2019; Jouaux et al., 2010; Matt and Allen, 2021) and maintain meat weight during the spawning season (Matt et al., 2020). Crossing tetraploids and diploids is now a widely applied method of commercial production of triploid oysters in many parts of the world, including the West Coast, Gulf Coast, and East Coast of the United States.

Mated triploid *C. virginica* are the preferred product in the lower Chesapeake Bay of the United States, a region where hatchery-based oyster aquaculture has grown dramatically over

the last 15 years. From 2005 to 2018, the number of market sized oysters sold by hatchery-based production has increased from less than 5 million to over 30 million (Hudson, 2019). Triploids have comprised nearly all of the production, consisting of 80 to 97% of hatchery-produced oysters planted on Virginia farms since surveys were initiated in 2009 (Hudson, 2019; Murray and Oesterling, 2010).

Commercial aquaculture of *C. virginica* in the lower Chesapeake Bay has benefitted from both polyploidy and selective breeding. The Aquaculture Genetics and Breeding Technology Center (ABC) at the Virginia Institute of Marine Science (VIMS) has been supplying commercial hatcheries with selectively bred diploid and tetraploid brood stock since 2004. Selective breeding at ABC was originally mass selection for faster growth and resistance to diseases caused by the protozoans *Haplosporidium nelsoni* and *Perkinsus marinus* (Frank-Lawale et al., 2014; Ragone Calvo et al., 2003). Mass selection involved selecting individuals for mating based on the phenotypes of the selection candidates (e.g., the largest oysters). ABC has since adopted family breeding, which entails producing and phenotyping sets of siblings, referred to as families. In contrast to mass selection, family breeding is conducive to tracking genetic relatedness of individuals in the population, which is important for controlling inbreeding and applying quantitative genetic analyses.

Quantitative genetics is the study of the inheritance of traits that vary continuously based on the effects of many genes. Quantitative genetic theory can be used with statistical analysis to draw conclusions about inheritance without needing to understand the underlying genetic mechanisms, making quantitative genetic analysis a fitting tool for selective breeding. With phenotypes and genetic relatedness, genetic parameters that can guide breeding strategy – heritability and genetic correlation – can be estimated from linear models. Heritability is the

proportion of phenotypic variation in a trait that is explained by genetic variation, and genetic correlation indicates how two traits are genetically related. Heritability and genetic correlations are fundamental to understanding the effects selective breeding may have on future generations and are therefore core genetic parameters for developing a data-driven strategy for selective breeding.

Quantitative genetic analysis has been applied to several species of oysters for the purpose of selective breeding (Allen et al., 2021; de Melo et al., 2019; Kube et al., 2018, 2011; Vu et al., 2020). In all instances, analyses have suggested selective breeding can increase the genetic value in future generations, thereby improving trait values (e.g., growth rate). In the selectively bred populations of diploid *C. virginica* managed by ABC, heritability has been sufficient to increase the genetic value of the population for the primary traits of interest, survival and total weight (Allen et al., 2021). In the tetraploid populations selected by ABC, non-zero heritabilities for traits suggest increases in genetic value are upcoming (Allen, Kube, Small, in prep). Interestingly, quantitative genetic analysis has never been applied to triploid oysters, leaving the potential to improve traits in triploids unclear.

Unusual mortality events in recent years were an impetus to extend quantitative genetic analysis to triploid *C. virginica* in the lower Chesapeake Bay. "Triploid mortality," defined as the mass mortality of near-market sized triploids in late spring, has been reported from oyster farms in the Chesapeake Bay since 2012 and has been the subject of several empirical studies (Guévélou et al., 2019; Matt et al., 2020; Ritter, 2019). A typical triploid mortality event exceeds 20% mortality of the stock but can approach much higher levels (50-85%) in some years. The mortality events are unusual because they are not associated with typical stressors of *C. virginica*, such as regional pathogens, poor husbandry, or extreme environmental conditions

(Guévélou et al., 2019; Matt et al., 2020). By applying quantitative genetic analysis to a triploid mortality event, selective breeding can be evaluated as an option to reduce the risk of these mortalities in future generations.

Applying quantitative genetic analysis to many commercial traits in triploid *C. virginica* would best inform a breeding strategy for triploid improvement. Collectively analyzing survival, growth traits, and shape traits would indicate the degree of potential improvement of these traits through estimates of heritability and inform inter-trait genetic correlations that can guide breeding strategy. Genetic correlations between traits guide breeding strategy by informing how selecting for one trait affects another trait. For example, a positive genetic correlation between size and survival may indicate that selecting for faster growing triploids will result in triploids with greater resistance to "triploid mortality." Another valuable application of genetic correlations is estimation of genotype by environment interactions (GxE). A GxE is when the effects of genes change differently with a change in the environment, and the strength of a GxE can be estimated by genetic correlation of the same trait measured in two environments (Falconer and Mackay, 1996). GxEs are important to breeding when the goal is to improve animals in multiple environments, such as for breeding for triploids that perform well across the variable environments in the Chesapeake Bay.

Breeding to improve triploids is indirect and involves alternative strategies that have not been evaluated. Due to their low fecundity (Allen and Downing, 1990; Jouaux et al., 2010; Matt and Allen, 2021) and the inviability of triploid x triploid crosses (Guo and Allen, 1994b), genetic improvement in mated triploids must derive from selective breeding of diploid and tetraploid oysters used for triploid production. Two major breeding strategies exist for improving mated triploids: 1) select diploids and tetraploids based on the performance of diploids and tetraploids

or 2) select diploids and tetraploids based on the performance of diploids, tetraploids, and triploids. Selecting based on diploid and tetraploid performance is the current strategy at ABC, where diploids and tetraploids are selected in parallel family breeding programs. The strategy resulting in greater triploid improvement is likely determined by genetic correlations between traits in diploids and tetraploids and traits in triploids. The higher the correlation between ploidies, the greater the parallel improvement in triploids from selecting based on diploid and tetraploid performance. Genetic correlations between traits in different ploidies have never been measured in any animals (but have in plants [e.g., Nyine et al., 2018]) likely because breeding tetraploid animals for triploid improvement is unique to oysters.

Estimates of genetic correlations between ploidies may not clearly indicate the better breeding strategy for triploid improvement. A negative correlation or a correlation near 1 would lead to clear guidance for a breeding strategy. However, the relative outcome of each strategy would not be apparent at moderate, positive genetic correlations (e.g., 0.6-0.8). With quantitative genetic theory alone, it is conceivable that either strategy is superior or the strategies are equivalent at such moderate correlations. A better understanding of the values for genetic correlation at which each strategy is advantageous may be gained from simulations. Simulations are regularly used in evaluating breeding strategies (e.g., Hickey and Gorjanc, 2012; Rutten et al., 2002; Wang et al., 2003) and are suitable to account for the many interacting factors involved in estimating the outcome of breeding strategies. Evaluating the strategies by simulation could guide polyploid breeding of *C. virginica* and be applicable to polyploid breeding in other species.

The goal of this dissertation is to advance the science of polyploid breeding by applying quantitative genetic principles to the aim of genetically improving triploid *C. virginica*. Two

major topics were addressed: 1) the outlook for genetic improvement of triploid *C. virginica* in the Chesapeake Bay through selective breeding, including the potential to breed for resistance to "triploid mortality" and 2) the better strategy to breed for genetic improvement in triploids. Chapters 2 and 3 describe analyses of a field test. Twenty triploid and forty tetraploid families related by tetraploid sire were reared at three sites in the Chesapeake Bay from fall of 2018 to fall of 2019. Chapter 2 focuses on quantitative genetic analysis of survival in the triploid and tetraploid families, including an analysis of survival associated with "triploid mortality." Chapter 3 examines the heritabilities and genetic correlations for other commercial traits — total weight, meat weight, meat yield, and shape traits. Chapter 4 is a comparison of breeding strategies for triploid improvement via simulation. The dissertation concludes with a synthesis that highlights how breeding for triploid *C. virginica* could best proceed and offers suggestions for future research into polyploid oyster breeding.

Chapter 2: Genetic basis of survival and “triploid mortality” from a field trial with families of triploid and tetraploid *Crassostrea virginica*

1. Introduction

Hatchery-based aquaculture has increased the commercial yield of the Eastern oyster, *Crassostrea virginica*, in the Chesapeake Bay. Approximately 35 million hatchery-produced oysters were sold each year from commercial farms in Virginia from 2014 to 2018 (Hudson, 2019). The production has partially restored an oyster industry that dwindled because of the collapse of the wild fishery (Schulte, 2017). The general underutilization of leased ground in the Chesapeake Bay suggests the industry has substantial room to grow (Beckensteiner et al., 2020; Bosch et al., 2010).

Currently, the hatchery-based aquaculture industry in the Virginia portion of the Chesapeake Bay consists of mostly triploids. Triploids have comprised 80 to 97% of the hatchery-produced oysters planted on Virginia farms between 2009 and 2018 (Hudson, 2019; Murray and Oesterling, 2010). In most industries, including in Virginia, triploids are commercially produced by mating tetraploid oysters to diploid oysters (Guo et al., 1996), which reliably produces 100% triploid progeny (e.g., Guo et al., 1996; Matt and Allen, 2014). Triploid oysters are a popular commercial product because of their added value in terms of faster growth (Dégremont et al., 2012; Harding, 2007) and higher meat yield during the warmer months (Allen, 1988). Growth and meat yield advantages are at least partially due to reduced fecundity in triploids, which is associated with a consistent meat weight during the spawning season (Allen and Downing, 1986; Matt et al. 2020).

A pressing issue in farming triploid oysters in the Chesapeake is unusual mortality often associated with late spring conditions, or “triploid mortality” (Guévélou et al., 2019; Matt et al., 2020). Reports from oyster farms in the Chesapeake Bay since 2012, as well as empirical studies (Guévélou et al., 2019; Matt et al., 2020), have defined triploid mortality as the mass

mortality of near-market sized (76mm) triploids in late spring. A typical triploid mortality event exceeds 20% mortality of the stock but can approach much higher levels (50-85%) in some years. The mortality events are unusual because they are not associated with typical stressors of *C. virginica*, such as regional pathogens, poor husbandry, extreme temperatures or salinities, or sudden changes in temperatures or salinities (Guévélou et al., 2019; Matt et al., 2020). The name, "triploid mortality," represents an early hypothesis that the mortality events were tied to the triploid condition. There is some evidence that triploids are especially susceptible to the late spring mortality events (Matt et al., 2020), however some studies have found diploids can die in similar fashion (Guévélou et al., 2019; Ritter, 2019).

In several instances, selective breeding has led to improved survival in oysters. Selective breeding increased survival in *C. virginica* in response to the epizootics caused by *Haplosporidium nelsoni* and *Perkinsus marinus* (Guo et al., 2003; Ragone Calvo et al., 2003) that caused mass mortalities of oysters in the Chesapeake Bay (Andrews, 1988), and increased survival of *Crassostrea gigas* in exposure to the Ostreid herpesvirus type I (OsHV-1) (Camara et al., 2017; Dégremont et al., 2015; Kube et al., 2018) that has caused mass mortalities of oysters in in western Europe (Peeler et al., 2012; Segarra et al., 2010), New Zealand (Keeling et al., 2014), and Australia (Paul-Pont et al., 2014). Breeding has also been effective in increasing survival in response to mortality events that occur without an attributable pathogen. Dégremont et al. (2010) measured a response to selection in resistance to "summer mortality" in diploid *C. gigas*. Summer mortality has been long reported in *C. gigas* across its distribution, and like triploid mortality, would occur in the warm months without evidence of a responsible pathogen or environmental condition (Glude, 1975; Koganezawa, 1975; Samain and McCombie, 2008).

A recent study by Ritter (2019) suggests selective breeding may also have an influence on triploid mortality. In their study, triploid half-sib families that varied based on tetraploid sire were deployed to a site where triploid mortality events regularly occur. A mortality event matching the criteria for triploid mortality occurred and there was substantial variance in survival among triploid families during the late spring window (Ritter, 2019). Variance in mortality among the families suggests transmittable genetic effects from the tetraploid sires, or additive genetic effects, may have been contributed to variation in survival and indicates selective breeding for resistance to triploid mortality may be effective.

Selective breeding for resistance to triploid mortality is complicated by the reproductive sterility of triploids. Fecundity is low in triploid oysters (e.g., Allen and Downing, 1990; Jouaux et al., 2010; Matt and Allen, 2021), and triploid x triploid crosses result in low survival and almost exclusively aneuploid progeny (Guo and Allen, 1994b). Due to the sterility of triploids, genetic improvement must derive from selective breeding of the diploid and tetraploid oysters used for their production.

The Aquaculture Genetics and Breeding Technology Center (ABC) at the Virginia Institute of Marine Science (VIMS) has been selecting for improvement in commercial traits of diploid oysters since it was created in 1997, and more recently has been programmatically selecting for improvement in tetraploids primarily for improvement in commercial triploids. Both diploids and tetraploids are selected in parallel family breeding programs. The diploid program has demonstrated regular genetic gains in survival in mesohaline and oligohaline sites since targeted selection began in 2015 (Allen et al., 2021), and positive heritability for survival in tetraploids suggests genetic gains are upcoming (Allen, Kube, Small, in prep).

It has been unclear if parallel diploid and tetraploid breeding programs are effective in improving triploid *C. virginica* and if it can be an effective strategy for breeding for resistance to triploid mortality. The principal question is, will improvement in tetraploid and diploid traits correlate with improvements in triploids? In other words, what is the correlation between genetic effects on traits in diploids and traits in triploids, and between traits in tetraploids and triploids. Genes improving diploid and tetraploid traits may also improve traits in triploids, meaning there is a high, positive genetic correlation, or the traits could be largely controlled by different genes, signifying a lower genetic correlation. These types of genetic correlations among different ploidies have never been measured in any animals, in part because breeding tetraploid and diploid animals for triploid improvement is unique to oysters (but not plants [e.g., Nyine et al., 2018]). It was necessary for ABC to establish these genetic correlations to determine the prognosis for the parallel breeding programs, including if parallel breeding programs may be an effective way to select for resistance to triploid mortality. Previous studies in *C. gigas* evaluating the effect of ploidy on resistance to summer mortality suggest for some traits, the genetic correlation between triploids and diploids and tetraploids are positive – selected diploid lines were found to transmit resistance to triploids (Dégremont et al., 2010), and selected tetraploid lines were found to transmit OsHV-1 resistance to triploids (Benabdelmouna, 2014; Dégremont et al., 2015).

Evaluating genetic parameters in polyploid populations, such as genetic correlations and heritability, require different methods than used for diploid populations. Additive genetic variance and breeding values for a trait are estimated based on the phenotypes and the genetic covariance among individuals. Genetic covariances among individuals can be derived from a pedigree using probabilities related to the extent of genetic similarity by descent (Wright, 1922;

Henderson, 1976). The standard methods of estimating genetic covariances from a pedigree assume all individuals are diploid (Henderson, 1976). When estimating genetic parameters in pedigrees with multiple ploidy levels, such as triploid and tetraploid, the standard methods provide inaccurate estimations of genetic covariance because the number of sets of chromosomes inherited from sire and dam can be disproportionate. Hamilton and Kerr (2018) developed rules for estimating genetic covariances from a mixed-ploidy pedigree and wrote an R package, polyAinv, to calculate the genetic covariance matrix from a user-provided pedigree. Although verified by simulation, the package has never been used to calculate genetic parameters in extant polyploid populations.

The primary objective in this study was to evaluate the additive genetic basis of late spring mortality events in triploid *C. virginica*, which would include the first measures of heritability in triploid oysters. Additionally, a goal was to estimate the genetic correlations between survival in triploid and tetraploid families, the first estimates of genetic correlations between mated polyploid animals. Together, these parameters can inform the possibility that selective breeding can reduce triploid mortality events and could also suggest strategies for optimal improvement of triploids. As part of a department-wide effort by ABC, twenty triploid and forty tetraploid families related by tetraploid sire were deployed to an oligohaline site and mesohaline site for a field test from fall of 2018 to fall of 2019. The same 20 triploid families were also deployed to a site with a history of triploid mortality events, and diploid and tetraploid lines were deployed to all sites for reference. Survival was monitored in families and lines, and in polyploid families, additive genetic variation and genetic correlations were estimated using univariate and multivariate animal models.

2. Methods

2.1 Brood stock and crosses

2.1.1 Families

Triploid and tetraploid families were produced by crossing individuals from diploid and tetraploid lines. Diploids consisted of oysters from the ABC line DEBY LEW (Ragone Calvo et al., 2003) that has been selected in a low salinity environment (8-15 ppt) or from DEBY LYN selected in a higher salinity environment (15-23 ppt) (Table 1). Tetraploids were from several VIMS lines: 4GEN, 4VBOY, 4OBOY, 4LGT, and 4GNL. The 4GEN line originated from ABC diploid lines and has been propagated by ABC since 2003, reared in the York River near Gloucester Point, Virginia and in the Rappahannock River near Topping, Virginia. 4VBOY, 4OBOY, and 4LGT all have germplasm from Louisiana material, originating as hybrids between diploid lines from Louisiana and the 4GEN line. Tetraploid ABC lines were created from the process of Guo and Allen (1994) and therefore there was an intermediate triploid stage between diploid and tetraploid. To make Louisiana derivative tetraploid lines, tetraploid males were crossed to a Louisiana diploid line to make triploid hybrids; the triploid hybrids were crossed again to the same Louisiana line to make tetraploids (Peachy and Allen, 2016). 4VBOY and 4OBOY were produced using the OBOY line (developed by Jerome La Peyre, Leonhardt et al., 2017) and the 4GEN line. 4LGT was produced using the ABC LGT line originating from wild oysters from Louisiana Grant Terre (Frank-Lawale et al., 2014). Finally, the 4GNL are F_1 hybrids produced by mating 4GEN and 4VBOY.

A total of 46 tetraploid families and 23 triploid half-sib families were spawned for this study on two dates: June 14, 2017 (24 tetraploid, 12 triploid) and July 10, 2017 (22 tetraploid, 11 triploid) at VIMS' Kauffman Aquaculture Center (Topping, Virginia). All crosses were conducted

via strip spawning (Allen and Bushek, 1992). Tetraploid families were made by crossing one male and one female. Triploid half-sib families were produced by crossing one tetraploid male with pooled eggs from five diploid females. The pool of diploid eggs was different between the two different spawning dates.

Only sperm from tetraploid males was used to make crosses. Sperm from each male was split into three aliquots to fertilize eggs from two tetraploid females and one aliquot of diploid eggs, producing two full-sib tetraploid families and one paternal half-sib triploid family. Eggs from five females from diploid line DEBY LEW were pooled and fertilized with sperm from 12 tetraploid males unique to the first spawn, and eggs from five females from diploid line DEBY LYN were pooled and fertilized with sperm from 11 tetraploid males unique to the second spawn.

2.1.2 Reference lines

Diploids and tetraploids from other spawns were included in the experiment (Table 2) as reference groups. For diploids, individuals from lines annually propagated by VIMS ABC (e.g. Frank-Lawale et al., 2014) were used as diploid reference groups: DEBY LYN (herein, DEBY), XB, HNRY, LOLA, and LILY. In addition, a synthetic diploid group was comprised of a mix of oysters from several VIMS families made in 2017 and labelled LFAMS for this study. DEBY, XB, and HNRY have been selected in higher salinity environments (18-22 ppt), whereas LOLA, LILY, and the population sampled to make LFAMS have been selected in low salinity (8-12 ppt). For tetraploid reference lines, individuals from lines annually propagated were used – 4GEN, 4GNL, and 4VBOY. A fourth tetraploid reference group was made, comprised from various VIMS tetraploid families made in 2017. Only individuals from families made from 4OBOY and 4LGT

were selected for the reference group, which was labelled 4OBLT. All reference lines were spawned in 2017 at the ABC research hatchery in Gloucester Point, VA, or at the VIMS Kauffman Aquaculture Center in Topping, VA.

2.2 Larval Rearing, Initial Field Deployment, and Ploidy Verification

Families (crosses in 2.1.1) were reared separately in individual tanks at the VIMS Kauffman Aquaculture Center. Larvae were reared at ~25°C, water was changed every two days, and larvae were fed *Pavlova* sp., *Chaetocerous neogracile*, and *Tetraselmis* sp. based on their development stage. Larvae were harvested once they reached the pediveliger stage. Harvesting consisted of sieving the mature pediveligers and storing them in a moist coffee filter at 4°C. Immature larvae were returned to tanks. Thereafter, pediveligers were harvested every two days for a total of three harvests (day 1, 3, and 5 of harvest). All three harvests for each family were combined and a maximum of 60,000 pediveligers were transferred to individual downweller systems containing oyster shell ground to ~400µm in diameter. When oysters could be retained on a 500µm screen, families were transferred to individual silos in a land-based upweller system at the ABC research hatchery where they were reared on unfiltered water from the York River.

Once most individuals within a family were larger than 5mm in diameter, each family was deployed into 3mm mesh bags contained in double-tier bottom cages in the York River (Gloucester Point, VA) at a maximum density of 4000 oysters per bag. The number of bags deployed per family ranged from one to three, and the time of deployment for each family ranged from August 28, 2017, to November 16, 2017.

Reference lines (2.1.2) were grown in a similar manner. Diploid reference lines were reared at the ABC Gloucester Point hatchery, and tetraploid reference lines were reared at the Kauffman Aquaculture Center.

All polyploid families and reference lines were sampled to verify ploidy via flow cytometry (FCM) (Allen, 1983). All FCM measurements were made with a Sysmex-Partex Cyflow Space flow cytometer (Partec GmbH, Münster, Germany) using DAPI as a stain. Fifteen to twenty-five individuals (1-2 mm shell length) were haphazardly selected from each family or cross while in an upwelling system and were analyzed individually by disaggregating cells from whole individuals. Samples were taken from 4GEN, 4GNL, and 4VBOY for ploidy verification. For 4OBLT, the families that made up 4OBLT were verified.

2.3 Field Test

Families and reference lines were deployed in a controlled field test from summer of 2018 to the fall of 2019 at three sites in the Chesapeake Bay: York River (Gloucester Point, VA), Choptank River (Cambridge, MD), and Nandua Creek (Pungoteague, VA) (Figure 1). At the York River and Choptank River, oysters were deployed in baskets on an adjustable longline system. At Nandua Creek, oysters were deployed in vexar bags in single-tier bottom cages.

Individuals from the seed deployed in 2017 were haphazardly selected from each family or reference line on June 4th and 5th of 2018 and deployed to sites for the field test. For families, seed from all deployed bags were combined prior to stocking individuals for the experiment. Each triploid family was deployed to each site. Tetraploid families were only deployed to York River and Choptank River due to logistical constraints. Diploid high salinity reference lines (DEBY, XB, HNRY) were only deployed to the high salinity sites, York River and

Nandua Creek, whereas diploid low salinity reference lines (LOLA, LILY, LFAMS) were deployed to the low salinity site, Choptank River. Each triploid family, diploid reference line, and tetraploid reference line was deployed in three units of 120 oysters per site, except for 4VBOY which was deployed in two units of 60. Each tetraploid family was deployed in three units of 150 oysters to York River and Choptank River, except for six families that were only deployed to York River because of low numbers. Five families deployed to York River and five families deployed to Choptank River were deployed in two units of 150. One family was deployed in two units of 120 per site. For all stocking densities (60,120, 150), oysters occupied less than half their container for the duration of the field test.

2.4 Environmental Conditions

Temperature was measured at each site during the field trial by a HOBO® TidbiT temperature logger (Onset Computer Corporation, Bourne, MA, USA) by placing a logger within a unit (bag or basket). Salinity was measured at Nandua Creek during the field trial with a HOBO ® conductivity logger. Calibrations for conductivity were conducted during each site visit with a portable conductivity meter, and conductivity data were converted to salinity in parts per thousand (ppt) using HOBOware Conductivity Assistant (Onset Computer Corporation, Bourne, MA, USA). Some mechanical failures produced gaps in the salinity data at Nandua Creek. Salinity data for the other sites were accessed from existing monitoring programs. Salinity data for the York River site were accessed via the Virginia Estuarine and Coastal Observing System's (VECOS, conducted by the Chesapeake Bay National Estuarine Research Reserve - Virginia) continuous monitoring station in Gloucester Point, Virginia (station ID: York RiverK005.40). The

continuous monitoring station in Gloucester Point collected data with a YSI 6600 data sonde and was approximately 1000 meters from the York River site. For the Choptank River site, salinity data collected using a YSI handheld sonde were provided by the Horn Point Laboratory at the University of Maryland in Cambridge, Maryland. Water assessed for salinity was from within 100 meters of the Choptank River site.

2.5 Survival

Survival was assessed within each unit (bag or basket) in the fall of 2018 (York River: Nov. 26, Choptank River: Nov. 12, Nandua Creek: Dec. 12), spring of 2019 (York River: April 29, Choptank River: May 6, Nandua Creek: May 7), summer of 2019 (York River: July 1, Choptank River: July 8, Nandua Creek: July 9), and fall of 2019 (York River: Oct. 21, Choptank River: Oct. 28, Nandua Creek: Oct. 28). Assessment involved counting live oysters and removing empty shells. While units were assessed for survival, oysters were removed without replacement for measurements. For the first survival assessment in the fall of 2018, survival within each unit was calculated as live oysters divided by the number of oysters deployed. For all subsequent assessments, survival within each unit was calculated with the following equation:

$$\text{Survival}_t = \frac{\text{Live}_t}{\text{Live}_{t-1} - \text{Samples Removed}_{t-1}} \text{ x } \text{Survival}_{t-1}$$

where t indicates sampling time and Samples Removed are the number of live oysters sampled without replacement.

2.6 Late Spring Survival

A particular span of time was used to define survival in late spring, so-called 'late spring survival' – the span between the April/May sampling and the early July sampling. Survival for

this period was calculated as the number of live oysters in early July divided by the number of live oysters in the units in April/May (after sampling).

2.7 Genetic analyses

Survival data from triploid families and tetraploid families were analyzed using linear mixed models in ASReml (Gilmour et al. 2015). Survival data within units were converted to binary data from individual animals (dead=0, alive=1) for the linear mixed models. For triploid families, the following model was used:

$$y = \mu + Spawn + Unit_{Family} + Animal + \varepsilon \quad (1)$$

where y is a vector of measurements, μ is the mean of the measurements, Spawn is a fixed effect indicating whether the individual was spawned in June or July of 2017, $Unit_{Family}$ is a random effect indicating unit nested within family, Animal is the random genetic effect linked to the pedigree, and ε is the residual variation. Data from tetraploid families were analyzed with the following mixed model:

$$y = \mu + Spawn + Unit_{Family} + Family + Animal + \varepsilon \quad (2)$$

where Family represents the random effect of the full-sib family, and Animal represents the random additive genetic effect linked to the pedigree.

Narrow-sense heritabilities (h^2) were estimated for cumulative survival (summer 2018 to fall 2019) and late spring survival (spring 2019 to summer 2019) in each population on both the observed and underlying scale. Herein, population refers to a combination of ploidy and site (e.g. triploids at York River). Narrow-sense heritabilities on the observed scale (h_o^2) were calculated with the following equation in triploid families:

$$h_o^2 = \frac{\sigma_a^2}{\sigma_u^2 + \sigma_a^2 + \sigma_\varepsilon^2}$$

and the following equation in tetraploid families:

$$h_o^2 = \frac{\sigma_a^2}{\sigma_u^2 + \sigma_f^2 + \sigma_a^2 + \sigma_\varepsilon^2}$$

where σ_a^2 , σ_u^2 , σ_ε^2 , and σ_f^2 represent the estimated variance of the Animal, Unit, Residual, and Family effects, respectively, from univariate models 1 and 2. The denominators of the heritability equations represent the phenotypic variation (σ_p^2). Standard errors were estimated using post-analysis in ASReml (.pin file). Heritability was adjusted to the underlying scale (h_u^2) with the following equation from Dempster and Lerner (1950):

$$h_u^2 = h_o^2 \frac{\left(p(1-p)\right)}{z^2}$$

where p is the proportion survival, and z is the height of the standard normal curve corresponding to p. The value for p was the average survival for the data set being analyzed. The z value was calculated with the following commands in R (R Core Team, 2019):

a = qnorm (p)

z = dnorm (a)

where a is the z score of the normal distribution corresponding to p, qnorm is a quantile function of the normal distribution, and dnorm is a density function of the normal distribution. Standard errors were calculated for heritabilities on the underlying scale with the following equation:

$$\left(\frac{h_u^2}{h_o^2}\right) h_{o_{se}}^2$$

where $h^2_{o_{se}}$ is the standard error of the heritability on the observed scale.

Genetic correlations were estimated as:

$$r_{g\,(1,2)} = \frac{\sigma_{a(1,\ 2)}}{\sqrt{\sigma^2_{a1}}\sqrt{\sigma^2_{a2}}}$$

where $\sigma_{a\,(1,\ 2)}$ represents the covariance in the Animal term for trait 1 and 2, and σ^2_{a1} and σ^2_{a2} represent the estimated variance of the Animal term for trait 1 and 2, respectively. Standard errors were estimated using post-analysis in ASReml (.pin file).

Genetic correlations were estimated for survival over the entire field trial (herein, survival) and late spring survival among 1) triploid families at different sites and 2) between triploid and tetraploid families. A bivariate model was used for all genetic correlation estimates. Between triploid families at different sites, a bivariate version of model (1) was used with covariances for residual and unit effects set to 0. Genetic correlations between triploid and tetraploid families were estimated with a bivariate version of model (2), with variance for the Family effect in triploid families set to 0 and covariances set to 0 for Residual, Unit, and Family effects.

2.8 Mixed-Ploidy Pedigrees

All genetic analyses require linking Animal term in the mixed model to a specified pedigree using an inverse additive relationship matrix. Inverse additive relationship matrices were produced using the polyAinv package in R (Hamilton and Kerr, 2018), which applies rules that are appropriate for pedigrees with different ploidies including odd-numbered ploidies (Hamilton and Kerr, 2018). The matrices were created in a step-wise process. First, a standard three-column pedigree (individual, sire, dam) was obtained from the ABC Oyster Breeding database operated by the Commonwealth Scientific and Industrial Research Organization

(CSIRO). Due to the spawning design (i.e., pooled eggs to produce the triploid families rather than using individuals), dams were not assigned to any triploid individual in the pedigree. The pedigree was then edited to a seven-column pedigree for polyAinv, as instructed in the help page, by adding four columns to each row: "sire gamete ploidy level", "sire lambda," "dam gamete ploidy level," and "dam lambda." A "2" was assigned for "sire gamete ploidy level" and "dam gamete ploidy level" for each tetraploid, while each triploid was assigned a "2" for "sire gamete ploidy level" and "1" for "dam gamete ploidy level." "Sire lambda" and "dam lambda" were set to 0 because double recombination was assumed absent in this analysis. ASReml "ginverse" objects (.giv files) were produced from polyAinv, which were used in ASReml as inverse additive relationship matrices.

3. Results

3.1 Families

Forty of the 46 tetraploid x tetraploid families and 20 of the 23 tetraploid x diploid families had sufficient numbers of oysters to be deployed in the field trial: 20 tetraploid and 10 triploid families came from the first spawn (June 2017) and 20 tetraploid families and 10 triploid families came from the second spawn (July 2017). All tetraploid lines were represented in the families that were deployed (Table 3). The tetraploid sires were mostly from 4LGT (7), 4GEN (6), and 4OBOY (5), with only a few coming from 4GNL (3) and 4VBOY (2). Similarly, the tetraploid dams were mostly from 4OBOY (12), 4GEN (10), and 4LGT (9), with lesser representation from 4GNL (6) and 4VBOY (3).

3.2 Ploidy

All samples from triploid families were verified triploid. In one tetraploid family, 2 of 15 sampled were triploid, and in another tetraploid family, 1 of 15 was diploid. All other tetraploid families were certified as 100% tetraploid, including the tetraploid families that made up 4OBLT, and all samples (25/25) from 4GEN, 4GNL, and 4VBOY.

3.3 Survival

By the end of the field trial, average survival was higher in diploids and triploids than in tetraploids (Figure 2 and Table 4). For York River, mean survival was 0.78 in diploid reference lines, 0.76 in triploid families, 0.53 in tetraploid families, and 0.46 in tetraploid reference lines. At Choptank River, mean survival was 0.68 in diploid reference lines, 0.49 in triploid families, and 0.08 in tetraploid reference lines. Tetraploid families at Choptank River had low survival between fall and winter of 2019 with many families having no remaining oysters. The field test for tetraploid families at Choptank River was therefore terminated in winter of 2019, and no data from tetraploid families at Choptank River were analyzed in this study. For Nandua Creek, mean survival was 0.84 in diploid reference lines, 0.77 in triploid families, and 0.63 in tetraploid reference lines. (Tetraploid families were not deployed to Nandua Creek.)

Survival by the end of the trial varied among diploid reference lines selected for low salinity and little among diploid reference lines selected for high salinity (Table 5). At Choptank River (low salinity), mean survival was 0.79 for LILY, 0.74 for LFAMS, and 0.50 for LOLA. At York River (high salinity), mean survival was 0.82 for XB, 0.80 for HNRY, and 0.72 for DBY. At Nandua Creek (also high salinity), mean survival was 0.90 for HNRY, 0.86 for DBY, and 0.77 for XB.

Tetraploid reference lines had higher variation in survival than diploid reference lines (Table 5). At York River, mean survival was 0.73 for 4OBLT (a mix of 2017 tetraploid families made from

4OBOY and 4LGT), 0.39 for 4GNL, and 0.27 for 4GEN. (Survival of 4VBOY at York River was not calculated because of a counting error.) At Choptank River, 4OBLT had the highest mean survival Choptank River (0.24), followed by 0.09 for 4GNL and 0 for 4GEN and 4VBOY. 4OBLT also had the highest mean survival at Nandua Creek (0.75), followed by 0.60 for 4GNL, 0.48 for 4GEN, and 0.69 for 4VBOY.

There was substantial variation for survival among triploid families and among tetraploid families by the end of the field trial (Figure 2). The range in survival in tetraploid families at York River was 0.10 – 0.77. For triploid families, variation was highest at Choptank River (0.32 – 0.89), followed by Nandua Creek (0.50 – 0.99) and York River (0.60 – 0.92). Mortality in the families occurred at different times at different sites. Tetraploid families and triploid families at York River had the highest mortality in late summer/early fall of 2019, resulting in a decrease of 0.15 and 0.08, respectively. At Choptank River, the greatest mortality in the triploid families was between fall of 2018 and spring of 2019 (decrease of 0.21). For Nandua Creek, highest mortality in triploid families occurred during the 'late spring survival' period of 2019, an average decrease of 0.11.

3.4 Late Spring Survival

In diploid reference lines, late spring survival across all sites ranged from 0.89 to 1.0 (Table 6). Late spring survival among tetraploid reference lines across all sites ranged from 0.81 to 0.96. In triploid families, late spring survival ranged from 0.60 to 1.0 across all sites, and late spring survival among tetraploid families at York River ranged from 0.64 to 1.0.

3.5 Genetic Analysis of Overall Survival

Substantial additive genetic variation for survival existed for triploid and tetraploid families. Narrow-sense heritability on the underlying scale was high for triploid families at Nandua Creek (0.81), moderate for triploid families at Choptank River (0.44) and York River (0.20), and moderate for tetraploid families at York River (0.24) (Table 7).

Genetic correlations for survival among populations were all positive and varied from moderate (0.46 – 0.62) to high (0.85 – 0.97) (Table 8). Moderate genetic correlations existed between triploid families at York River and Nandua Creek (0.46), and Choptank River and Nandua Creek (0.46). A very high genetic correlation was found between triploid families at York River and triploid families at Choptank River (0.97). Genetic correlations between tetraploid and triploid families were moderate for York River and Choptank River sites, 0.55 and 0.62, respectively. At the York River, the genetic correlation between tetraploid and triploid families was high (0.85). Genetic correlations were moderate between tetraploid families at York River and 1) triploid families at Choptank River (0.55) and 2) triploid families at Nandua Creek (0.62).

3.6 Genetic Analysis of Late Spring Survival

Narrow-sense heritability on the underlying scale for late spring survival varied substantially among populations (Table 9). For triploids, heritability was low at York River ($h_u^2 =$ 0.09) and moderate at Choptank River ($h_u^2 = 0.42$). For triploid families at Nandua Creek, heritability for late spring survival was very high ($h_u^2 = 1.06$). For the only group of tetraploid families (at York River), heritability for late spring survival was moderate ($h_u^2 = 0.36$).

Genetic correlations between populations varied from 0.21 to 0.89 for late spring survival (Table 10). Genetic correlations were low between triploid families at Nandua Creek

and triploid families at York River (0.32) and Choptank River (0.21), and higher between York River and Choptank River (0.81). Genetic correlations between tetraploid and triploid families were moderate between York River and Choptank River (0.32) and higher between York River and York River (0.89) and York River and Nandua Creek (0.76). Standard errors for estimates of genetic correlation in late spring survival were high, ranging from 0.22 to 0.54.

3.7 Environmental conditions

Temperature loggers were deployed in 2019 and collected data from April 3 to October 21 at York River, from May 6 to October 28 at Choptank River, and from March 21 to October 28 at Nandua Creek (Figure 4). For the time when data were available at all sites (May 6 to October 21), mean average daily temperature was 24.9°C at York River, 24.7°C at Choptank River, and 26.5°C at Nandua Creek. Between the survival assessment in spring 2019 and summer 2019 (May 7 to July 9), mean average daily temperature was 24.1°C at York River, 23.8°C at Choptank River, and 26.5°C at Nandua Creek.

A salinity logger was deployed to Nandua Creek on December 11, 2018. Salinity data from York River and Choptank River were accessed from VECOS and provided by Horn Point Laboratory from June 1, 2018 to October 28th, 2019. From the time data were available at all sites (December 11, 2018 – October 28, 2019), mean average daily salinity was 15.9 at York River, 7.7 at Choptank River, and 13.7 at Nandua Creek (Figure 5). Salinity was lowest at each site between November 2018 and April 2019, when the mean average daily salinity was 13.0 at York River, 10.9 at Nandua Creek, and 5.5 at Choptank River.

4. Discussion

4.1 Quantitative Genetic Analysis of Late Spring Survival at Nandua Creek

Survival in late spring was treated as a distinct trait for quantitative genetic analysis for the primary objective of evaluating whether selective breeding would be an option to reduce severe mortality of triploid oysters during late spring, often referred to as "triploid mortality"(Guévélou et al., 2019; Matt et al., 2020). The mortality events occur in May and June and result in the loss of many (> 0.20) near market sized triploid oysters (Guévélou et al., 2019; Matt et al., 2020; Ritter, 2019). Over 20 oyster farms in Virginia have reported mortality events matching such criteria since 2012 (Karen Hudson, personal communication), including a farm in Nandua Creek, where the mortality events occur regularly (Matt et al., 2020; Ritter, 2019).

A large proportion of the variation in late spring survival among triploid families at Nandua Creek was attributed to additive genetic effects. In late spring, defined in this study as between May 7 and June 9 at Nandua Creek, one triploid family had survival of 0.79 and three had survival less than 0.70. Meanwhile, most of the 20 triploid families had survival greater than 0.90. The narrow-sense heritability estimate for late spring survival in triploid families at Nandua Creek was 1.06 (0.32) (standard error), implying that all or a large proportion the variation in survival was due to additive genetic effects.

High variation in survival of triploid families reared at Nandua Creek was previously observed by Ritter (2019). Like this study, Ritter (2019) tested half-sib triploid families that varied based on their tetraploid sire. Ritter (2019) deployed twelve triploid half-sib families, made from crossing 12 tetraploid sires to pooled diploid females, to Nandua Creek from March 2017 to June 2018. By the end of experiment, cumulative survival among triploid families

ranged from 0.22 to 0.88, with most of the mortality occurring in the late spring — between March and June of 2018. The results in Ritter (2019) corroborate the results in this study, and although not demonstrated, the high variation in survival among triploid families in Ritter (2019) suggests that the heritability of late spring survival in triploids at Nandua Creek was high in that study.

The very high heritability estimate for late spring survival at Nandua Creek in this study (1.06) is not without precedent. Heritability estimates have also been very high in field trials targeting "summer mortality" in diploid *C. gigas* (Dégremont et al., 2010, 2007). Dégremont (2007) measured summer survival in juveniles (4-6 months old) of 43 full-sib families across three sites along the French coast. Over only a few months, survival varied from 0.02 to 0.82, resulting in narrow-sense heritability estimates of 0.47 (0.20), 0.89 (0.40), and 1.08 (0.46). In a follow-up study, Dégremont et al. (2010) measured realized heritability of summer survival through divergent selection. "High" (high survival) and "low" (low survival) lines were produced from oysters that were retained in safe conditions, separate from the field trials, so they were full sibs of families that had high and low survival in the field in Dégremont (2007). Large differences in survival between juveniles in the "high" and "low" lines over two generations resulted in realized heritability estimates ranging from 0.55 (0.18) to 1.02 (0.20), supporting the finding in Dégremont (2007) that heritability for this trait was very high.

The findings by Dégremont (2007, 2010) are pertinent because summer mortality in *C. gigas* has been discussed as analogous to triploid mortality events in *C. virginica* (Guévélou et al., 2019; Matt et al., 2020). Both occur during the reproductive season and have a hypothesized etiology of a physiological - environmental interaction involving gametogenesis (Koganezawa, 1975; Matt et al., 2020; Samain and McCombie, 2008). The "high" (high survival) and "low"

(low survival) lines originating from Dégremont (2007, 2010) were renamed “Resistant” = high survival; “Susceptible” = low survival and subject of several follow-up studies in attempts to understand the mechanism behind the genetically-based susceptibility (Huvet et al., 2010, 2008; Samain et al., 2007). Gonad development was a major factor distinguishing the lines – oysters in the Susceptible lines often showed a higher percent of gonad area in histological sections and spawning was often “partial” rather than “massive” (Huvet et al., 2010, 2008; Samain et al., 2007). However, a mechanism explaining the difference in survival between the two contrasting lines was unresolved (Samain, 2011). Investigations of the physiological mechanisms underlying variation in late spring survival in triploid *C. virginica* could benefit from an approach similar to that of Dégremont et al. (2010), Huvet et al. (2010, 2008), and Samain et al. (2007) – triploids made from diploid and tetraploid families with the highest and lowest estimated breeding values could be produced and tested to better understand the cause of the mortalities.

Although a strong genetic basis was observed for late spring survival of triploids at Nandua Creek, it is unclear if the selection pressures at Nandua Creek are representative for the more widely observed syndrome of late spring mortalities in triploid *C. virginica*. Many farms in Virginia have reported late spring mortalities in triploids, and although most reports have come from the bayside of the Eastern Shore of Virginia (K. Hudson, personal communication) in presumably similar environments as Nandua Creek, reports have also come from the western section of the Chesapeake Bay and the seaside of the Eastern Shore of Virginia (Atlantic Ocean). Because Nandua Creek was the only site where a late spring mortality event was observed, there is no way to know if the genetic basis of survival at Nandua Creek is expressed in other sites with late spring mortality. Genetic correlations between a late spring mortality event at Nandua and other site(s) are critical to this understanding, and high correlation would indicate

that selecting for higher survival at Nandua Creek would be an effective strategy for improving survival at other commercial farms. The biggest challenge in estimating these correlations is that these mortality events are hard to capture because they do not consistently occur at the same sites every year (Guévélou et al., 2019; Matt et al., 2020). Logistically, deploying enough triploid families across numerous sites is onerous. As for late spring survival at the other sites, we can assume that the low genetic correlation between Nandua Creek and Choptank River or Nandua Creek and York River indicate that the causes of mortality may not be the same, i.e., are not specific to late spring mortality. Genetic correlations with other sites experiencing a late spring mortality event, as well as more robust estimates from additional field trials, are important in shaping a breeding plan to reduce late spring mortalities of triploid *C. virginica* on commercial farms.

4.2 Examination of the "Triploid Mortality" Event at Nandua Creek

This study primarily focused on evaluating the additive genetic basis of "triploid mortality" for the first time, but also shared methods from previous studies of examining environmental and genetic factors that may be associated with "triploid mortality." The high mortality in several triploid families at Nandua Creek fit the definition of "triploid mortality" because many (> 20%) near market sized triploid oysters died in late spring (May 7 to July 9), with little mortality (< 10%) before or after the event. Several studies, including this study, have identified a genetic by environment (G x E) interaction associated with triploid mortality events (Guévélou et al., 2019; Matt et al., 2020; Ritter, 2019). The same studies have examined several

possible causes or correlating factors, including regional pathogens, size, brood stock origin, gonad morphology, and female fecundity, all of which have been found to have no clear causative association (Guévélou et al., 2019; Matt et al., 2020; Ritter, 2019). In this study, salinity and temperature were monitored as possible factors, and ploidy was examined as possibly having an association with suscepibility.

Physical conditions, such as temperature, have not been identified as a distinguishing factor for sites affected by triploid mortality events. Matt et al. (2020) found similar annually daily temperatures from mid-April to August for a site with a triploid mortality event and nearby test sites without an event. In this dissertation, however, mean daily temperature during late spring (May 7 to July 9) was more than 2 degrees higher at Nandua Creek than at the unaffected sites, and during late May, mean daily temperature was approximately 4 degrees higher at Nandua Creek than the other sites (Figure 4). Although inherent site differences in temperature, such as those related to tidal flushing or residence time, cannot be ruled out, the difference in temperature measured among sites may have been due to gear type. Temperature loggers were placed into a bag or basket containing oysters, and baskets were intertidal at York River and Choptank River (long-line system) and subtidal at Nandua Creek (bottom cages). The intertidal sites tracked together closely in temperature, whereas the subtidal site was regularly warmer. It is possible that episodes of desiccation, particularly at night, may have caused lower average daily temperatures at the intertidal sites. Considering findings from Matt et al. (2020), high temperature remains an unlikely distinction of high-risk sites, however, culture method, such as an intertidal vs. subtidal, may be worth investigating at affected sites like Nandua Creek as a possible factor that affects the severity of the mortality events.

Several empirical studies, including this one, have evaluated the question whether mortality episodes in late spring in the lower Chesapeake Bay are specific to triploids or not. Matt et al. (2020), like this study, found diploids of Virginia genetic origin had low mortality during a triploid mortality event at Nandua Creek. The three diploid reference lines tested in this study had mortality of 0.04, 0.06, and 0.11 during the triploid mortality event, much lower than the high mortality observed in several triploid families (0.21, 0.35, 0.40, 0.40). In contrast, Ritter (2019) and Guévélou et al. (2019) found instances of high levels of mortality in diploids during a triploid mortality event at Nandua Creek and a nearby creek (Nassawadox Creek), respectively. Results from Ritter (2019) and Guévélou et al. (2019) showed diploids can die at high levels during late spring like triploids, however, the relative susceptibility based on ploidy remains unresolved and thus warrants that diploids continue to be tested alongside triploids.

4.3 Analysis of Overall Survival

4.3.1 Survival Among Diploid Lines

Survival among diploid reference lines involved a comparison of diploids made from mass selection and family breeding. Three of the diploid reference lines tested in this study were ABC mass-selected lines (Frank-Lawale et al., 2014): LOLA, DBY, and XB, and three were products of family breeding at ABC (Allen et al., 2021): LILY, LFAMS, and HNRY. The oysters from the family breeding program had much higher survival than the mass-selected lines in low salinity. LILY and LFAMS were top survivors at Choptank River, having much higher survival (> 0.20) than LOLA. At the mesohaline sites, little difference was found among the diploid lines. Higher survival was expected in the lines made from families because oysters in the family breeding program are a product of continued selection of the mass-selected lines. ABC

switched to family breeding in diploids in 2014, starting with a base population of families founded in large part by the mass-selected lines (Allen et al., 2021). At the same time, ABC ceased selecting the mass-selected lines, instead focusing on maintaining their current state. Additional trials that benchmark the performance of lines derived from the family breeding program (LILY, HNRY) to mass-selected lines (LOLA, DBY, XB) would be valuable in demonstrating the improvements made from family breeding and may help guide decision making for commercial growers on what seed to purchase for best performance.

4.3.2 Survival Among Triploid Families

The differing timing of mortality in triploid families among the sites suggested that survival manifested as a different trait (i.e., had a different etiology) at each site. Based on the results of the field trial, the causes of high survival were hypothesized to be low salinity tolerance at Choptank River, resistance to "triploid mortality" at Nandua Creek, and disease resistance at York River. Most of the mortality in triploids at Choptank River occurred in the winter and early spring of 2019, presumably as a result of the low salinity, which was an average of 5.5 during this time. A salinity of 5 is the minimum of the range at which *C. virginica* naturally occurs (5-40) (Galstoff, 1964) and is an uncommonly low salinity for oysters bred by ABC. Mortality in triploids at Nandua Creek primarily occurred in late spring, matching criteria for "triploid mortality." In the York River, most of the mortality occurred during the summer and fall of 2019; cumulative mortality in several triploid families at York River decreased by over 0.15 from July 1 to October 21 of 2019. Disease from *P. marinus* was hypothesized as an important factor in these mortalities because prevalence and intensity of infection from *P. marinus* increases with age (Andrews, 1988) and are at their peak in the late summer in the lower York River (Burreson and Ragone Calvo, 1996).

Based on quantitative genetic analysis, the "three" causes of mortality may only be only two: resistance to "triploid mortality," and general "robustness." Between York River and Nandua Creek and between Choptank River and Nandua Creek, the genetic correlation for survival during the field trial was 0.46 (standard error = 0.24 or 0.22). Albeit an imprecise estimate, the values suggest a weak relationship existed between the genes determining survival at Nandua Creek and the other two sites. In contrast, the genetic correlation between York River and Choptank River was 0.97 (0.14). The very high genetic correlation between York River and Choptank River indicates that despite the different timing and conditions in which mortality at these sites occurred, survival at each site was likely influenced by many of the same genes.

The high genetic correlation between triploid families in the York River and triploid families in the Choptank River in this study suggests that for at least survival, salinity may not be a major agent of G x E interaction for triploids. The same does not seem true for diploids. The genetic correlation between survival at York River and Choptank River in diploid *C. virginica* families was much lower (0.57 ± 0.14) (Allen et al., 2021). A major caveat of comparing the results from this study with that in Allen et al. (2021) is that the data for diploids are based on many year classes and hundreds of families, while the genetic parameters for triploid *C. virginica* from this study represent only one year class of 20 families. Data from additional year classes of triploids are needed to verify this difference between diploids and triploids. However, there is reason to expect that triploid oysters could be generalists across salinity zones. Polyploidy has often been attributed to an increased tolerance to stresses in plants (Dubcovsky and Dvorak, 2007; Levin, 2002), possibly because of an increase in extent of heterozygosity (Otto and Whitton, 2000).

4.4 Methods of Data Analysis

Polyploid populations were analyzed for genetic parameters in this study, which required different methods than that standard for diploid populations. The analyses in this study relied on adopting methods developed by Gallais (2003), Kerr (2012), and Hamilton and Kerr (2018), which incorporate ploidy in the calculation of the coefficient of coancestry, coefficient of inbreeding, and coefficient of relationship. Typically, these interconnected elements are calculated for and among individuals in a pedigree with methods assuming each individual is diploid (Henderson, 1976). The elements, which summarize genetic relationships, can then be used in estimation of genetic parameters and breeding values in the form of a covariance matrix, referred to as the additive relationship matrix (the inverse of the matrix is required to solve the mixed model equation). Incorporating ploidy into the calculations of the coefficients can lead to more accurate estimates of genetic relationships when pedigrees contain polyploids because the probability of inbreeding and dynamics of gene flow differ based on ploidy (Gallais, 2003; Hamilton and Kerr, 2018; Kerr et al., 2012).

Calculating genetic relationships based on ploidy is particularly important for pedigrees that contain multiple ploidy levels, as in the current study, because tetraploids and diploids can have unequal genetic contributions to offspring and thus disproportionate genetic relationships with progeny. For example, a tetraploid sire and triploid offspring share a higher proportion of genes than a tetraploid sire and tetraploid offspring, simply because in the former, 2/3 of the genome of the triploid is contributed by the tetraploid, and in the latter case – 1/2. Thus, if standard methods for diploids (Henderson, 1976) were used in this study, the estimated genetic

relationship among triploid half-sibs would have been 0.25. With the rules by Hamilton and Kerr (2018), the relationship was 0.33. The difference has implications for the estimations of breeding values, heritability, and genetic correlations.

The crossing design in this study prioritized examining the additive genetic effects from tetraploid sires. Additive genetic effects (or breeding value) represent the expected performance of offspring if an individual is mated at random in the population (Falconer and Mackay, 1996). By mating each tetraploid in this study to five diploid dams, survival of triploid offspring measured the contribution from various dams, resulting in a robust estimate of the additive genetic effect of each tetraploid sire on the survival in triploids. Additionally, crossing the same diploid dams to tetraploid sires (same five diploid dams for first spawn, same five diploid dams for second spawn) partly controlled for diploid effects among half-sib families. The design was efficient at assessing variance in additive genetic effects among tetraploid sires, however it was at the expense of the accuracy of genetic relationships. Each triploid family was assumed to contain an equal proportion of individuals descended from each diploid dam, yet the actual proportion of triploids from each dam was unknown. Deviations from the assumption of equal representation across dams would increase the impact diploid effects had on the results and would decrease the accuracy of the estimate of genetic parameters. More reliable estimates of genetic parameters are achievable from a design with only single pair matings.

Population structure in the base population of tetraploids may have affected estimates of the genetic parameters. When estimating genetic parameters from a pedigree, the default assumption is that the individuals of the base population (founders) derive from a single randomly mating population (Isik and Holland, 2017; Muff et al., 2019). In the genetic analysis

in this study, tetraploids were assumed to have derived from a genetically homogenous population, however sires were from several tetraploid lines. The tetraploid lines all share ancestry with 4GEN, which descended from crosses between oysters from Grand Terre, Louisiana (LA) and the XB line, however each line also derived from other genetic sources. The 4LGT line derives from 4GEN crossed with diploids from LA Grand Terre, 4OBOY and 4VBOY derive from 4GEN crossed with diploids from LA Oyster Bayou, and 4GNL is a hybrid of 4GEN and 4VBOY. Thus, unaccounted genetic structure may have led to biased estimates of genetic parameters and overestimations of heritability (Isik and Holland, 2017; Wolak and Reid, 2017). In practice, the influence that subpopulations exert on the estimation of genetic parameters, referred to as genetic group effects, can be assessed by including this information in statistical analysis. However, for this work there were too few families to do a genetic groups analysis. Genetic group effects from the various tetraploid lines can be evaluated in future studies once larger numbers of families descended from each line are attained in the pedigree.

4.5 Selective Breeding for Higher Survival in Triploids

The prospect of improving survival in triploids from selective breeding is promising because substantial additive genetic variation for survival was estimated in this study, however the process of improving triploids is complicated by the inability to select them directly. Triploid oysters are effectively reproductively sterile. Fecundity is low (e.g., Allen and Downing, 1990; Jouaux et al., 2010; Matt and Allen, 2021), but more problematic is that triploid x triploid crosses result in low survival and almost exclusively aneuploid progeny (Guo and Allen, 1994b).

Therefore, triploids cannot be bred directly and triploid improvement must come from selection in diploids, tetraploids, or both.

With family breeding, the fundamental question with targeting triploid improvement is what is the nature of the field tests for estimating breeding values. There are three primary choices (Table 11). First, test diploid and tetraploid families, selectively breed both in parallel, and expect that the triploids improve from the improvement in the diploids and tetraploids. Second, test only triploids in the field and estimate breeding values in diploid and tetraploid families based on their genetic relationship with the triploid families. Third, test all three ploidies – diploid, triploid, and tetraploid – in the field and estimate breeding values based on performance of all three sets of families. Each approach is discussed below. In chapter 4, approaches are examined through simulations.

The first approach, producing and testing only diploid and tetraploid families for triploid improvement, is the most simple but also the most uncertain approach. The approach is simple because a pre-existing program for diploid improvement need not be altered, and a parallel tetraploid program can proceed independently, which can simplify logistics. The approach is uncertain because improvement in triploids is dependent on genetic correlations between traits in triploids and traits in diploids and tetraploids, and these genetic correlations have not been established for commercial traits in *C. virginica*. The genetic correlations determine the sign (+/-) and size of the response to selection in the triploid traits (rg_{xy} – Table 11, column a).

A primary objective in this study was to estimate the genetic correlation between survival in triploids and tetraploids because genetic correlations involving triploid traits had never been measured. All correlations were positive, which suggests selectively breeding

tetraploids alone can have a positive effect on the survival of triploids. Additionally, several genetic correlations were high, such as overall survival between tetraploids and triploids at York River (0.85), and late spring survival between tetraploids at York River and triploids at Nandua Creek (0.76), suggesting a substantial correlated effect between tetraploids and triploids could occur for survival. The genetic correlation was lower for other traits (overall survival: 0.55, 0.62). More importantly, because the standard errors were always high (range: 0.23 to 0.29), the genetic correlations between survival in tetraploids and triploids, and thus the outcome of selectively breeding tetraploids and expecting improvement in triploids, remains largely uncertain.

In the second approach, testing only triploid families eliminates the dependency on the genetic correlation with traits in diploids and tetraploids. For this approach, triploid families would be produced that are half-sibs with diploid and tetraploid families, but only triploids would be tested in the field (Table 11, column b). Diploids and tetraploid families would be selected based on the performance of the triploids and their relationships in the pedigree. A drawback with this strategy is selection of diploids and tetraploids occurs at the half-sib level. Triploid families and tetraploid families will share a tetraploid parent, while triploid families and diploid families will share a diploid parent. Compared to selecting at the full-sib family level, which is possible in approach one and three, selecting at the half-sib level has a lower possible maximum accuracy of selection (r_{ux} or r_{uy} – Table 11), defined as the correlation between the estimated breeding value and the true breeding value. Thus, testing only triploid families is likely to have a lower rate of improvement in triploids compared to the other strategies.

The third approach, testing related diploid, triploid, and tetraploid families, is the most complete because the genetic correlation between traits in triploids and traits in diploids and tetraploids can be estimated and incorporated into the selection decision process. By involving the genetic correlation in the estimation of breeding values, any non-zero genetic correlation could increase the accuracy of selection by helping to distinguish diploid or tetraploid full-sib families. Testing related diploid, triploid, and tetraploid families is the best default strategy because it allows genetic correlations between ploidies, the primary determinant of the most advantageous strategy in terms of between-family breeding (Table 11), to be estimated while selectively breeding for triploid improvement.

To reduce triploid mortality events, arguably the most pressing issue for triploid improvement in the Chesapeake Bay, selective breeding can proceed most quickly by testing related diploid, triploid, and tetraploid families. Because the events are defined by high mortality in triploids, triploids need to be measured to establish the additive genetic basis of the trait. Testing related diploids and tetraploids can identify traits with a substantial genetic correlation with triploid mortality that could increase selection accuracy. Traits routinely measured in the ABC diploid and tetraploid family breeding program, such as 18 month survival, total weight, and meat weight at York River and Horn Point (Allen et al., 2021) warrant evaluation as genetically correlated traits, and given the high genetic correlation between late spring survival in tetraploids in the York River and late spring survival in triploids at Nandua Creek (0.76), late spring survival in diploids and tetraploids may be worth measuring on an annual basis. Extending field tests of diploids and tetraploids to sites expected to experience triploid mortality may also be valuable. Although the low variance in late spring survival in diploid and tetraploid lines at Nandua Creek in this study suggests testing diploids and

tetraploids there may not be worthwhile, families may exhibit higher variance than lines, and other studies suggest diploids may have substantial late spring mortality in some years (Guévélou et al., 2019; Ritter, 2019).

5. Tables

Table 1. Spawning design for tetraploid families ($4n_x$) and triploid families ($3n_x$) produced from spawns on June 14, 2017 and July 10, 2017. Sperm from tetraploid sires (numbered 1 to 23) were split three ways and crossed to two tetraploid dams (numbered 1 to 46) and an aliquot of eggs from diploids. Eggs from diploids were pooled from five DEBY LEW diploids in the first spawn and five DEBY LYN diploids in the second spawn.

	Spawn 1: June 14, 2017				Spawn 2: July 10, 2017			
Tetraploid Families	$4n_1$ $4n_2$	$4n_3$ $4n_4$	…	$4n_{23}$ $4n_{24}$	$4n_{25}$ $4n_{26}$	$4n_{27}$ $4n_{28}$	…	$4n_{45}$ $4n_{46}$
	↑	↑	…	↑	↑	↑	…	↑
Tetraploid Dams	1, 2	3, 4	…	23, 24	25, 26	27, 28	…	45, 46
	x	x	…	x	x	x	…	x
Tetraploid Sires	1	2	…	12	13	14	…	23
	x	x	…	x	x	x	…	x
Diploid Dams	DEBY LEW	DEBY LEW	…	DEBY LEW	DEBY LYN	DEBY LYN	…	DEBY LYN
	↓	↓	…	↓	↓	↓	…	↓
Triploid Families	$3n_1$	$3n_2$	…	$3n_{12}$	$3n_{13}$	$3n_{14}$	…	$3n_{23}$

Table 2: Diploid (2N) and tetraploid (4N) lines of *Crassostrea virginica* included in the experiment. Lines are distinguished by salinity environment in which they were selected: high (18-22 ppt) or low (8-12 ppt). Tetraploid lines have been reared in the York River (Gloucester Point, Virginia) and Rappahannock River (Topping, Virginia)

Cross	Ploidy	Salinity
LOLA	2N	low
LILY	2N	low
LFAMS	2N	low
DEBY	2N	high
XB	2N	high
HNRY	2N	high
4GEN	4N	–
4GNL	4N	–
4VBOY	4N	–
4OBLT	4N	–

Table 3. Tetraploid brood stock represented in the 40 tetraploid families and 20 triploid families that survived to be deployed in the field trial. Descriptions of tetraploid lines (4GEN, 4GNL, 4LGT, 4OBOY, 4VBOY) are in the Methods.

	Sires	Dams	Totals
4GEN	6	10	16
4GNL	3	6	9
4LGT	7	9	16
4OBOY	5	12	17
4VBOY	2	3	5
Totals	23	40	

Table 4: Minimum, mean, and maximum overall survival (survival between summer of 2018 and fall of 2019) in diploid lines, tetraploid lines, triploid families, and tetraploid families of *Crassostrea virginica* measured at three sites in the Chesapeake Bay (York River, Choptank River, Nandua Creek). 2N=diploids, 3N=triploids, 4N=tetraploids.

	York River				Choptank River				Nandua Creek			
	n	min	mean	max	n	min	mean	max	n	min	mean	max
2N Lines	3	0.72	0.78	0.82	3	0.50	0.68	0.79	3	0.77	0.84	0.90
4N Lines	3	0.27	0.46	0.73	4	0	0.08	0.24	4	0.48	0.63	0.75
3N Families	20	0.60	0.76	0.92	20	0.32	0.49	0.89	20	0.50	0.77	0.99
4N Families	40	0.10	0.53	0.77	–	–	–	–	–	–	–	–

Table 5: Mean survival of diploid lines (top) and tetraploid lines (bottom) of *Crassostrea virginica* at York River (YR), Choptank River (CR), and Nandua Creek (ND). Standard errors are in parentheses. "–" represents lines not deployed. *Survival of 4VBOY was not calculated at YR because of a counting error.

Cross	YR	CR	ND
LOLA	–	0.50 (0.05)	–
LILY	–	0.79 (0.04)	–
LFAMS	–	0.74 (0.04)	–
DEBY	0.72 (0.04)	–	0.86 (0.05)
HNRY	0.80 (0.05)	–	0.90 (0.03)
XB	0.82 (0.02)	–	0.77 (0.08)
4GEN	0.27 (0.05)	0	0.48 (0.06)
4GNL	0.39 (0.05)	0.09 (0.04)	0.60 (0.07)
4VBOY	*	0	0.69 (0.02)
4OBLT	0.73 (0.02)	0.24 (0.05)	0.75 (0.03)

Table 6: Minimum, mean, and maximum late spring survival (survival between spring of 2019 and summer of 2019) in diploid lines, tetraploid lines, triploid families, and tetraploid families of *Crassostrea virginica* measured at three sites in the Chesapeake Bay (York River, Choptank River, Nandua Creek). Lines or families with less than 50 animals per replicate (on average) were excluded from calculation of minimum, mean, and maximum. Interval survival was not calculated in tetraploid lines at CR because all tetraploid lines had less than 50 individuals per rep at CR by spring of 2019. 2N=diploids, 3N=triploids, 4N=tetraploids.

	York River				Choptank River				Nandua Creek			
	n	min	mean	max	n	min	mean	max	n	min	mean	max
2N Lines	3	0.94	0.97	1.0	3	0.93	0.94	0.94	3	0.89	0.93	0.96
4N Lines	3	0.81	0.88	0.96	4	–	–	–	4	0.87	0.88	0.90
3N Families	20	0.91	0.97	1.0	20	0.79	0.94	1.0	20	0.60	0.88	1.0
4N Families	40	0.64	0.91	1.0	–	–	–	–	–	–	–	–

Table 7: Overall survival: Additive genetic variation (σ_a^2), phenotypic variation (σ_p^2), and narrow-sense heritability (h^2) on the observed and underlying scale for survival between summer of 2018 and fall of 2019 in triploid (3N) and tetraploid (4N) families of *Crassostrea virginica* measured at three sites in the Chesapeake Bay (York River, Choptank River, Nandua Creek). Standard errors are in parentheses.

	4N York River	3N York River	3N Choptank River	3N Nandua Creek
σ_a^2	0.04	0.02	0.07	0.08
σ_p^2	0.25	0.18	0.25	0.18
observed h^2	0.15 (0.06)	0.11 (0.05)	0.28 (0.11)	0.42 (0.13)
underlying h^2	0.24 (0.09)	0.20 (0.10)	0.44 (0.17)	0.81 (0.25)

Table 8: Overall survival: Genetic correlations for survival from summer of 2018 to fall of 2019 between triploid (3N) families of *Crassostrea virginica* (top) and between triploid and tetraploid (4N) families (bottom) measured at three sites in the Chesapeake Bay (York River, Choptank River, Nandua Creek).

3N York River	3N Choptank River	0.97 (0.14)
3N York River	3N Nandua Creek	0.46 (0.24)
3N Choptank River	3N Nandua Creek	0.46 (0.22)
4N York River	3N York River	0.85 (0.23)
4N York River	3N Choptank River	0.55 (0.29)
4N York River	3N Nandua Creek	0.62 (0.24)

Table 9: Late spring survival: Additive genetic variation (σ_a^2), phenotypic variation (σ_p^2), and narrow-sense heritability (h^2) on the observed and underlying scale for late spring survival, measured between spring of 2019 and summer of 2019, in triploid (3N) and tetraploid (4N) families of *Crassostrea virginica* measured at three sites in the Chesapeake Bay (York River, Choptank River, Nandua Creek). Standard errors are in parentheses.

	4N_York River	3N_York River	3N_Choptank River	3N_Nandua Creek
σ_a^2	0.01	0.0006	0.008	0.04
σ_p^2	0.09	0.03	0.07	0.11
observed h^2	0.13 (0.05)	0.02 (0.02)	0.12 (0.06)	0.40 (0.12)
underlying h^2	0.36 (0.14)	0.09 (0.10)	0.42 (0.23)	1.06 (0.32)

Table 10: Late spring survival: Genetic correlations for late spring survival, measured between spring and summer of 2019, between triploid (3N) families of *Crassostrea virginica* at Nandua Creek and triploid and tetraploid (4N) families at the York River and Choptank River. Standard errors are in parentheses.

3N York River	3N Choptank River	0.81 (0.45)
3N York River	3N Nandua Creek	0.32 (0.39)
3N Choptank River	3N Nandua Creek	0.21 (0.30)
4N York River	3N York River	0.89 (0.54)
4N York River	3N Choptank River	0.32 (0.39)
4N York River	3N Nandua Creek	0.76 (0.22)

Table 11: Breeding strategies (columns a, b, c) for improving triploid oysters using family breeding. Each strategy is described by the families produced, the families tested (e.g., phenotyped from a field trial), the trait that is used to estimate breeding values, the trait targeted for improvement, a simplified version of the elements dictating the response of the target trait to selection, a scenario in which each strategy is advantageous, and the maximum possible accuracy (correlation of estimated breeding value and true breeding value) from the breeding scheme. In the example, only two traits exist, a trait in triploids (y) and a trait in diploids or tetraploids (x). *i :* intensity of selection, r_{ux} : accuracy of selection on trait x based on index u for trait x, r_{uy} : accuracy of selection on trait y based on index u for trait y, rg_{xy} : genetic correlation between trait x and trait y, σ_{ay} : additive genetic standard deviation for trait y. Index u refers to estimating breeding values using univariate or multivariate animal models described in methods.

	a	b	c
Produce	2N + 4N	2N + 3N + 4N	2N + 3N + 4N
Test	2N + 4N	3N	2N + 3N + 4N
Selected Trait	2N or 4N *(x)*	3N *(y)*	3N *(y)*
Target Trait	3N *(y)*	3N *(y)*	3N *(y)*
Simplified Response Equation	$i \cdot r_{ux} \cdot rg_{xy} \cdot \sigma_{ay}$	$i \cdot r_{uy} \cdot \sigma_{ay}$	$i \cdot r_{uy} \cdot \sigma_{ay}$
Max. Accuracy (r_{ux} or r_{uy})	0.71	0.50	0.71

6. Figures

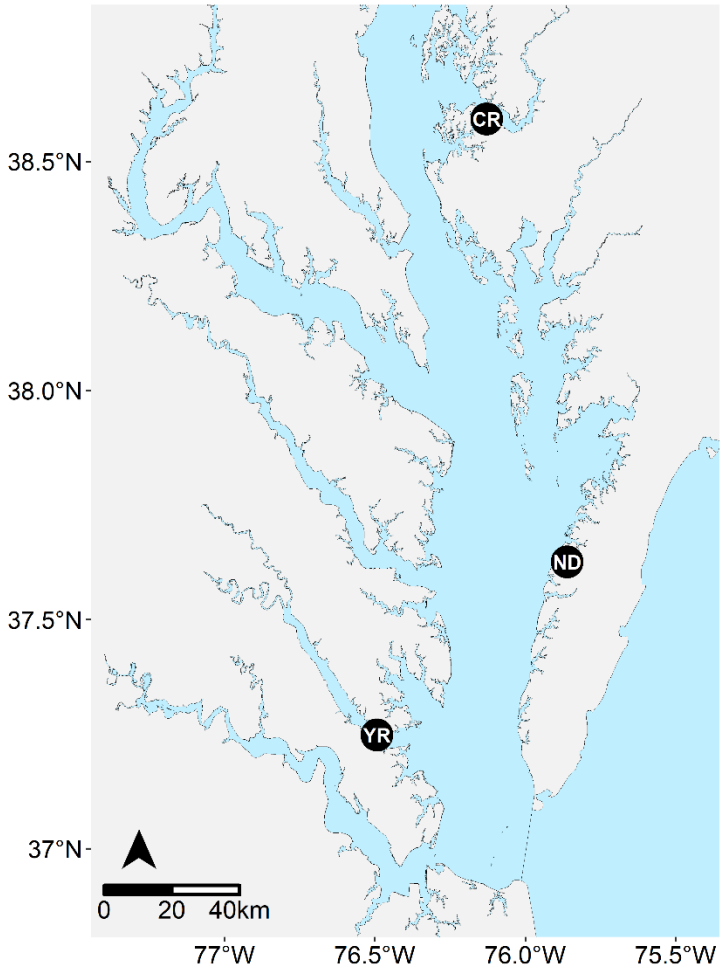


Figure 1: Map of sites in the Chesapeake Bay where triploid and tetraploid families of *Crassostrea virginica* were deployed for field testing. YR=York River, CR=Choptank River, ND=Nandua Creek.

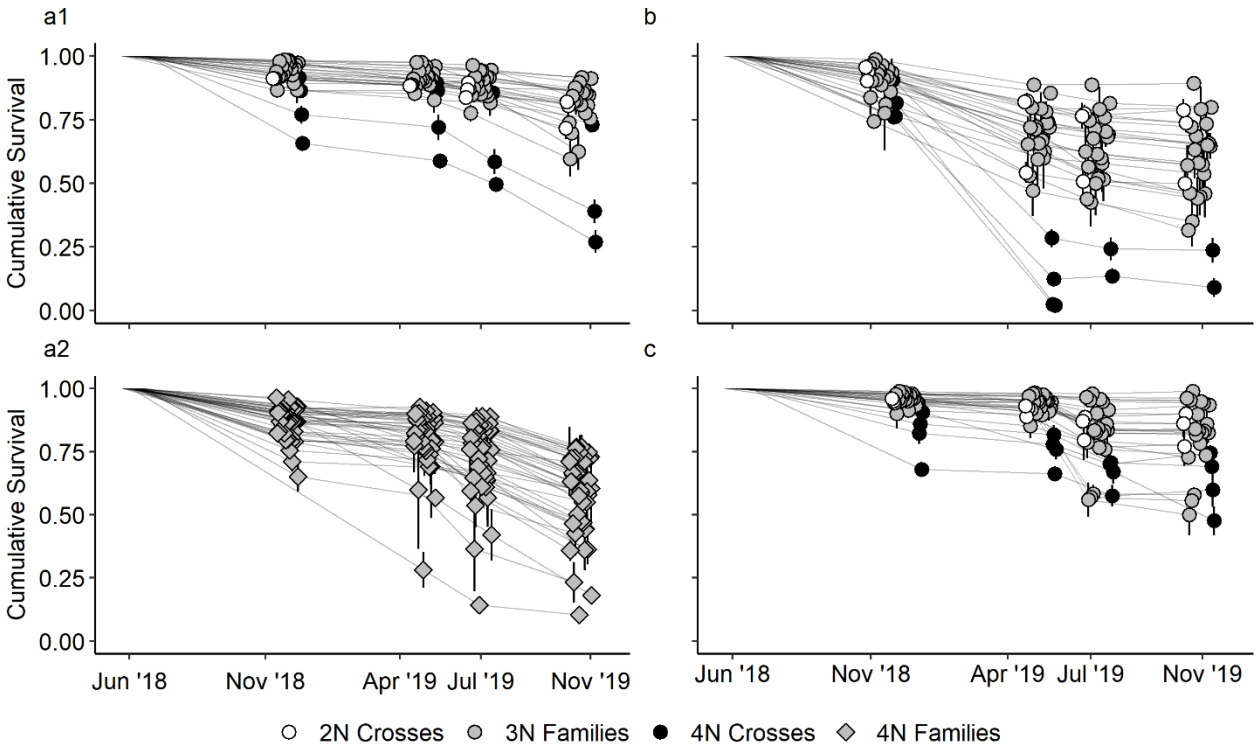


Figure 2: Survival in diploid lines, triploid families, tetraploid lines, and tetraploid families from June 2018 to November 2019 at a) York River, b) Choptank River, and c) Nandua Creek. Diploid lines, triploid families, and tetraploid lines are plotted separately (a1) from tetraploid families (a2) at the York River. Error bars represent standard error. 2N=diploid, 3N=triploid, 4N=tetraploid.

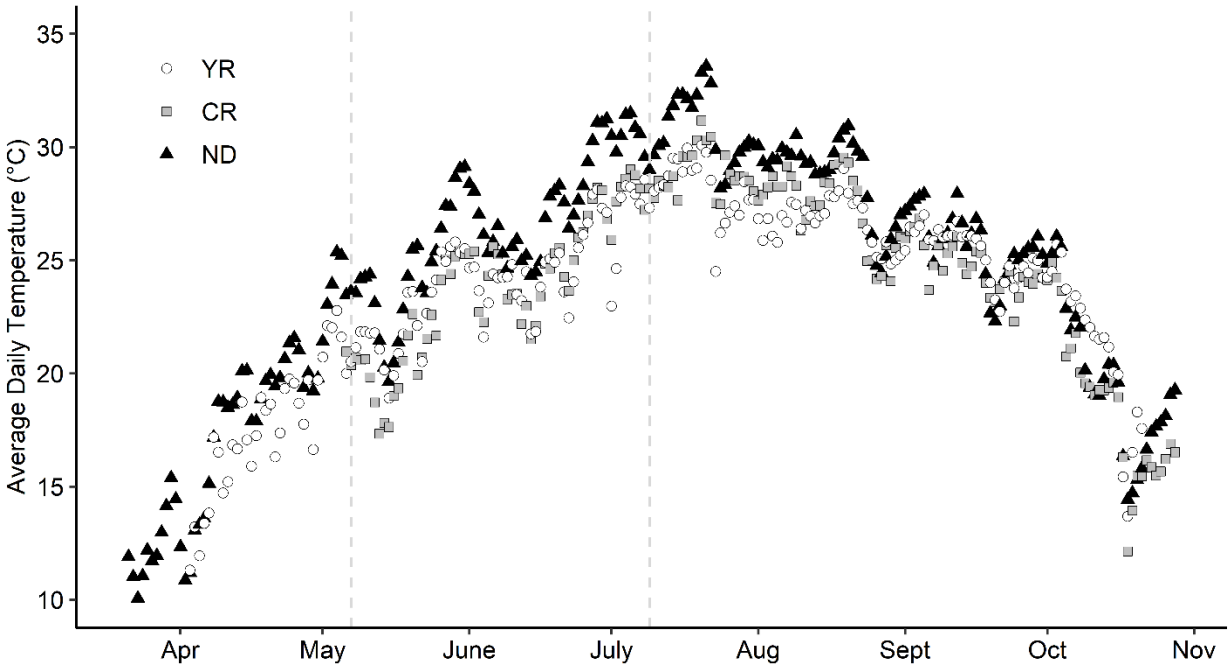


Figure 3: Average daily temperature (degrees Celsius) from March 21^{s}, 2019, to October 28th, 2019, at York River (YR), Choptank River (CR), and Nandua Creek (ND). Temperature data were collected at each site by HOBO® temperature loggers. Dotted lines represent the survival assessments in spring of 2019 (May 7) and summer of 2

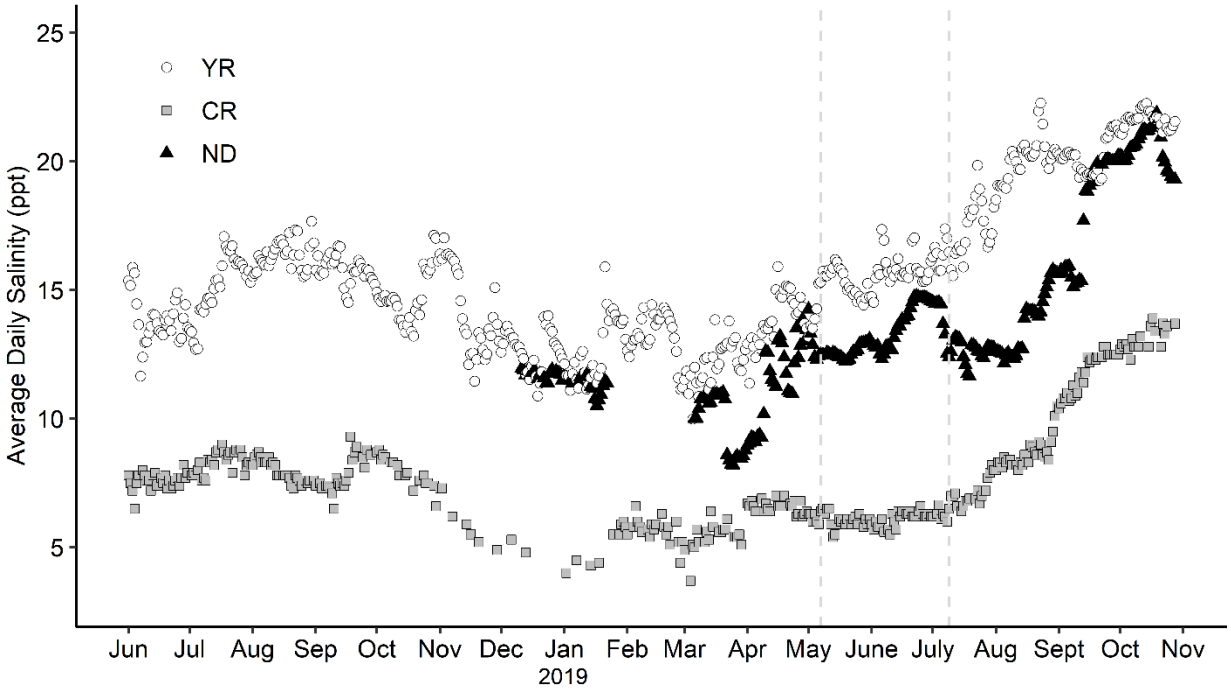


Figure 4: Average daily salinity (ppt) from June 1st, 2018 to October 28th, 2019 at York River (YR), Choptank River (CR), and Nandua Creek (ND). Salinity for YR is from Virginia Estuarine and Coastal Observing System monitoring station in Gloucester Point, Virginia (YRK005.40). Salinity for CR is from Horn Point Laboratory, University of Maryland, in Cambridge, Maryland. Salinity for ND collected by HOBO® conductivity logger.

Chapter 3: Genetic parameters of growth and quality traits in triploid and tetraploid *Crassostrea virginica*

1. Introduction

Quantitative genetic analysis has provided a data-driven strategy for selective breeding of *Crassostrea virginica* in the Chesapeake Bay. After years of effective mass selection of diploid oysters for higher survival and growth rate (Frank-Lawale et al., 2014; Ragone Calvo et al., 2003), the Aquaculture Genetics and Breeding Technology Center (ABC) at the Virginia Institute of Marine Science (VIMS) began applying quantitative genetic principles in the form of a family breeding program (Allen et al., 2021). Analyzing variance among diploid families for commercial traits has yielded estimates of heritability and genetic correlation that have shaped the strategy for selecting diploid oysters suited for high and low salinity zones (Allen et al., 2021).

Quantitative genetic methods can be used to develop a breeding strategy to improve the popular commercial product, triploid *C. virginica*. Triploids are the major crop of hatchery-based oyster aquaculture in Virginia, as nearly all (85%) of the farmed oysters in Virginia are triploid (Hudson, 2018). Thus far, the only trait in triploids that has been examined using quantitative genetic methods is survival (Chapter 2). Survival in triploids had substantial heritability (often greater than 0.30), indicating a high potential for improvement from breeding. Applying quantitative genetic methods to other traits of commercial importance, such as total weight, meat yield, and shape, can further inform the potential to improve triploids and guide an optimal breeding strategy.

Analysis of all commercial traits in triploids can inform inter-trait genetic relationships that affect breeding strategy. Traits are often genetically related because they are influenced by many of the same genes (Falconer and Mackay, 1996). Some genes may increase the values of

both traits they affect, while others may affect traits antagonistically. The net effect of shared genes on two traits can be estimated from quantitative genetic analysis as a genetic correlation, and genetic correlations between traits are important to breeding strategy because they indicate how selecting for one trait affects another trait. Genetic correlations can be considered favorable for the breeding program if selecting for a desired change in one trait results in a desired change in the other. Genetic correlations can also be unfavorable, in that selecting for a desired change in a trait results in an undesirable change in another trait. Typically, unfavorable correlations between commercially important traits are managed with a trade-off — traits are selected in a way so that some improvement in one trait is sacrificed to reduce adverse changes in the other trait. Such unfavorable genetic correlations have been identified between growth rate and shell shape characteristics in *C. virginica* and *Crassostrea gigas*, leading to breeding strategies of sacrificing some improvement in growth rate to avoid undesirable changes in shape (Allen et al., 2021; Kube et al., 2011; Ward et al., 2005).

Important for triploids are genetic correlations with episodic mortality in late spring, or "triploid mortality" (Guévélou et al., 2019; Matt et al., 2020). Triploid mortality is the mass mortality of near market size (76mm) triploid *C. virginica* in late spring that occurs without evidence of typical stressors, such as regional pathogens, poor husbandry, or extreme environmental conditions (Guévélou et al., 2019; Matt et al., 2020). A quantitative genetic analysis of 20 triploid families found variance in survival during a triploid mortality event had high heritability, suggesting resistance can be gained from selective breeding (Chapter 2). Genetic correlations can indicate if selection in other commercial traits can improve resistance to "triploid mortality" via favorable correlations, or if selecting for resistance will have undesirable effects on other commercial traits from unfavorable correlations.

Genetic correlations can also be used to evaluate how to improve triploids for multiple environments most effectively. The breeding goal at ABC is to genetically improve oysters cultured throughout the Chesapeake Bay, which is a spatially heterogenous estuary. Selecting for oysters for one environment in the Bay may not effectively improve them for other environments because of genotype by environment interactions (GxE). A GxE indicates the effects of genes change differently with a change in the environment. For example, genotypes with the highest value for size at one site may have the average or lowest value for size at another site. The strength of a GxE can be estimated with quantitative genetic analysis as a genetic correlation of the same trait measured in each environment (Falconer and Mackay, 1996). High, positive genetic correlations indicate weak GxE, whereas low or negative correlations indicate strong GxE. To improve animals most effectively for multiple environments, a GxE may justify treating the same trait measured in multiple environments as distinct traits, or even managing and selecting separate populations suited for each environment (e.g., James, 1961). In Chapter 2, a moderate GxE was found for survival in triploids between a site with a triploid mortality event and two other sites without an event. Analysis of more commercial traits in triploids across sites can determine the extent of GxE in other important traits and guide how triploids can be best selected for improvement across the Chesapeake Bay.

Genetic correlations between traits in different ploidies are important to breeding strategy for triploid oysters. The indirect nature of genetically improving triploids means genetic correlations with traits in other ploidies are essential to breeding strategy. Triploid *C. virginica* are reproductively sterile and commercially produced by mating diploids to tetraploids (e.g., Guo et al., 1996), and therefore genetic improvement of triploids derives from selectively

breeding diploids and tetraploids. The inability to propagate triploids yields two major options for improving triploids: 1) select diploids and tetraploids based solely on their performance or 2) select diploids and tetraploids based on the performance of diploids, tetraploids, and triploids. Traits in triploids can be genetically improved by selecting diploid and tetraploid families in parallel programs provided there is a favorable genetic correlation between the traits in diploids and tetraploids and the traits in triploids. The higher the correlation, the more the triploids will improve in parallel with the other ploidies. So far, the only genetic correlations estimated between different ploidies in oysters has been for survival between tetraploids and triploids (Chapter 2), which was strong and favorable between tetraploids and triploids tested at the same site (0.85). Genetic correlations between ploidies for other commercial traits are valuable to further evaluate which of the two breeding approaches is better for genetically improving triploid oysters.

The primary objective in this study was to estimate heritabilities and genetic correlations for commercial traits in triploid oysters to inform breeding strategy for triploid improvement. Total weight, meat weight, meat yield, and shape traits were measured in oysters from the same 20 triploid and 40 tetraploid families from Chapter 2 and were analyzed using univariate and multivariate animal models. Genetic correlations were estimated between traits measured at the same site, between the same traits measured in different environments, and between traits in triploids and tetraploids.

2. Methods

Tetraploid and triploid families in this study were the same as those described in Chapter 2.

2.1 Brood stock and crosses

Triploid and tetraploid families were produced by the staff at ABC by crossing individuals from diploid and tetraploid lines. In brief, diploid brood stock were from the Virginia Institute of Marine Science (VIMS) DEBY line (Ragone Calvo et al., 2003) that had been selected in either a low salinity environment (8-15 ppt) (DEBY LEW) or higher salinity environment (15-23 ppt) (DEBY LYN). Tetraploid brood stock were from several VIMS tetraploid lines: 4GEN, 4VBOY, 4OBOY, 4LGT, or 4GNL. Tetraploid and triploid families were produced from these brood stock in two spawns in June and July of 2017 at the VIMS Kauffman Aquaculture Center (Topping, Virginia). All crosses were conducted via strip spawning (Allen and Bushek, 1992). For each spawn, five diploid females were stripped, and eggs were pooled and split into aliquots. Individual tetraploid males were stripped, and sperm from each male was split into three aliquots. Two of the aliquots were used to fertilize eggs from individual tetraploid females, and one was used to fertilize an aliquot of diploid eggs. In result, two full-sib tetraploid families and one paternal half-sib triploid family were produced from each tetraploid sire. Diploid DEBY LEW were used in the first spawn and diploid DEBY LYN were used in the second spawn. Twenty triploid families and forty tetraploid families were produced in total, half from each spawn (Chapter 2).

2.2 Larval Rearing and Initial Field Deployment

Larvae from each family were raised separately in individual larval tanks by the staff at ABC, and seed from each family was kept separate in downweller and upweller systems at the hatchery in Gloucester Point, VA. Once most individuals within a family were larger than 5mm in diameter, families were deployed into bottom cages in the York River (Gloucester Point, VA). The time of deployment for each family ranged from August 28, 2017, to November 16, 2017.

2.3 Ploidy Verification

Fifteen or twenty-five individuals were sampled from each tetraploid and triploid family for ploidy verification and were analyzed individually. DNA content was verified via flow cytometry (FCM) (Allen, 1983). All FCM measurements were made with a Sysmex-Partec Cyflow Space flow cytometer (Partec GmbH, Münster, Germany) using DAPI as a stain.

2.4 Field Test

Tetraploid and triploid families were deployed in a controlled field test from the summer of 2018 to the fall of 2019 at three sites in the Chesapeake Bay: York River (Gloucester Point, VA), Choptank River (Cambridge, MD), and Nandua Creek (Pungoteague, VA). At the York River and Choptank River, oysters were deployed in baskets on an adjustable longline system. At Nandua Creek , oysters were deployed in vexar bags in single-tier bottom cages.

Individual oysters were haphazardly selected from each family on June 4th and 5th of 2018 and deployed to sites for the field test. Each triploid family was deployed to each site, while tetraploid families were only deployed to York River and Choptank River due to logistical constraints. Each triploid family was deployed in three baskets of 120 oysters for York River and Choptank River and three bags of 120 oysters for Nandua Creek. Each tetraploid family was deployed in three baskets of 150 oysters to York River and Choptank River, except for six families that were only deployed to York River because of low numbers. Five tetraploid families deployed to York River and five tetraploid families deployed to Choptank River were deployed in two baskets of 150. One tetraploid family was deployed in two baskets of 120 per site. All field work was done with the assistance of the ABC staff.

Salinity was recorded during the field test at Nandua Creek using a conductivity logger and reported in Chapter 2 (Figure 2.4). Salinity data for York River and Choptank River during

the field test were obtained from the Chesapeake Bay National Estuarine Research Reserve – Virginia, which maintains a continuous monitoring station in Gloucester Point, Virginia (station ID: YRK005.40), and from the Horn Point Laboratory at the University of Maryland in Cambridge, Maryland, respectively (Table 1).

2.5 Measurements

Families were sub-sampled and measured four times during the field test: fall 2018, spring 2019, summer 2019, and fall 2019. For triploid families, 15 individuals were sampled per unit (bag or basket) on each sampling date for all metrics (45 per family). For tetraploid families, what measures were taken varied according to season. In the fall of 2018 and 2019, 10 individuals per basket (30 per family) were sampled for all metrics. Shells were cleaned of sediment and fouling organisms before measuring. Measuring involved the following: total weight, wet meat weight, shell length (greatest distance from umbo to bill), shell width (greatest distance perpendicular to length), and shell height ("greatest distance between the outside of the closed valves, measured at a right angle to the shell commissure" [Galstoff, 1964]). To measure meat weight, shucked meats were drained on a mesh net and then weighed. In spring 2019, only total weight was measured on 15 individuals per basket (45 per family). In summer of 2019, 15 individuals per basket (45 per family) were measured for all metrics except wet meat weight. All measurements were completed with assistance of the ABC staff. All data were stored in the ABC Oyster Breeding database (i.e., Allen et al., 2021).

2.6 Traits for Quantitative Genetic Analysis

Five traits were selected for quantitative genetic analyses: total weight, meat weight, meat yield, width index, and height index. Meat yield was calculated as the ratio of meat weight to total weight. The width index is shell width divided by shell length, and the height index is shell height divided by shell length.

2.7 Genetic Analyses

Data were analyzed with univariate and multivariate mixed models in ASReml (Gilmour et al. 2015). For analyzing data from triploid families, the following mixed model was used:

$$y = \mu + Spawn + Unit_{Family} + Animal + \varepsilon \quad (1)$$

where y is a vector of measurements, μ is the mean of the measurements, Spawn is a fixed effect indicating whether the individual was spawned in June or July of 2017, $Unit_{Family}$ is a random effect indicating unit nested within family, Animal is the random genetic effect linked to the pedigree, and ε is the residual variation. For analyzing data from tetraploid families, the following mixed model was used:

$$y = \mu + Spawn + Unit_{Family} + Family + Animal + \varepsilon \quad (2)$$

where Family represents the random effect of the full-sib family, and Animal represents the random additive genetic effect linked to the pedigree.

Data were analyzed in stages. First, models were iteratively fit, from univariate to multivariate, to estimate genetic correlations among sampling periods for each trait (e.g., total weight in fall of 2018 vs. total weight in spring of 2019). For data from triploid families, a four trait multivariate version of model (1) was ultimately fit to estimate genetic correlations among sampling periods in each population. Herein, a population is defined as the combination of

ploidy and site (e.g., triploid families at York River). A bivariate version of model (2) was ultimately fit for data from tetraploid families to estimate genetic correlations among sampling periods. Inter-trait estimations of covariance were included for Unit, Family, and Animal terms, while covariances for Residual terms were set to 0.

Genetic correlations were estimated as:

$$r_{g\,(1,2)} = \frac{\sigma_{a(1,\ 2)}}{\sqrt{\sigma^2_{a1}}\sqrt{\sigma^2_{a2}}}$$

where $\sigma_{a\,(1,\ 2)}$ represents the covariance in the Animal term for trait 1 and 2, and σ^2_{a1} and σ^2_{a2} represent the estimated variance of the Animal term for trait 1 and 2, respectively. Standard errors were estimated using post-analysis in ASReml (.pin file).

For many traits, data from different sampling periods were pooled to represent a single trait for genetic analyses. Data from different sampling periods were only pooled if genetic correlations between sampling periods were high in all populations. Prior to pooling, data were standardized by dividing by the phenotypic standard deviation. After data were pooled, all mixed models involved an additional fixed effect, Measure Date, to account for the multiple sampling periods. All subsequent genetic analyses were performed on standardized, pooled data.

Narrow-sense heritability (h^2) was estimated for each trait in each population. Univariate models were used to estimate heritabilities in triploid and tetraploid families. For traits in triploid families, h^2 was estimated as:

$$h^2 = \frac{\sigma^2_a}{\sigma^2_u + \sigma^2_a + \sigma^2_\varepsilon}$$

and in tetraploid families was estimated as:

$$h^2 = \frac{\sigma_a^2}{\sigma_u^2 + \sigma_f^2 + \sigma_a^2 + \sigma_\varepsilon^2}$$

where σ_u^2 , σ_a^2 , σ_ε^2 , and σ_f^2 represent the estimated variance of the Unit, Animal, Residual, and Family effects, respectively. Common full-sib family effects (c^2) were estimated for tetraploid families as:

$$c^2 = \frac{\sigma_f^2}{\sigma_u^2 + \sigma_a^2 + \sigma_a^2 + \sigma_\varepsilon^2}$$

Common full-sib family effects could not be estimated in triploids because triploid families were paternal half-sib families, not full-sib families. Standard errors were estimated using post-analysis in ASReml (.pin file).

From combined data, genetic correlations were estimated among various sets of data: 1) between traits of the same population, 2) between triploid families at different sites, and 3) between triploid and tetraploid families.

(1) Within-population genetic correlations were estimated for each trait in each population. For triploid families, a bivariate model was used and inter-trait covariances were estimated for Unit, Animal, and Residual terms. Bivariate models that included estimations of inter-trait covariance for Unit, Family, Animal, and Residual terms were used for tetraploid families.

(2) Between-site genetic correlations were estimated for each trait in triploid populations. A bivariate model was used for between-site genetic correlations with covariances for Residual and Unit effects set to 0.

(3) Genetic correlations were estimated between-ploidy using a two trait version of model (2). Variance for the Family effect in data from triploid families was set to 0, and covariances were set to 0 for Residual, Unit, and Family effects.

2.8 Genetic Correlations with Survival

Genetic correlations were estimated between the five traits measured in this chapter (total weight, meat weight, meat yield, width index, height index) and survival. Data for survival is reported in Chapter 2 and comprised survival for two periods of time: 1) the entire field trial (summer 2018 to fall 2019), herein "overall survival," and 2) survival between spring 2019 and summer 2019, herein "late spring survival." Late spring survival was of interest because it is the defined period for triploid mortality events in the Chesapeake Bay (e.g., Guévélou et al., 2019; Matt et al., 2020). Genetic correlations between overall survival and the traits measured in this chapter were estimated in triploid and tetraploid families at each site. Genetic correlations between late spring survival and the traits measured in this chapter were estimated for Nandua Creek only, where a triploid mortality event occurred (Chapter 2). For genetic correlations involving survival or late spring survival, a two trait version of model (1) was used for triploids and a two trait version of model (2) was used for tetraploids. Covariances were set to 0 for Residual and Family effects.

2.9 Mixed-Ploidy Pedigrees

All genetic analyses included an inverse additive relationship matrix as the pedigree information. Inverse additive relationship matrices were produced from pedigrees using the polyAinv package in R (Hamilton and Kerr, 2018), which applies rules that are appropriate for pedigrees with different ploidies (Hamilton and Kerr, 2018). The pedigree assumed no double

reduction, and all triploids were assigned a '0' for dam (as in Chapter 2). Inverse additive relationship matrices were produced by polyAinv as "ginverse" objects, which were used in ASReml as the pedigree information.

2.10 Statistics

Phenotypic data was described qualitatively, with mean and standard deviation, to summarize the distribution of the data as in Allen et al. (2021). Estimates of heritability and genetic correlation were considered statistically different from zero if the 95% confidence interval of the estimate did not contain zero.

3. Results

3.1 Ploidy

All samples from triploid families were verified triploid. In one tetraploid family, 2 of 15 sampled individuals were triploid, and in another tetraploid family, 1 of 15 sampled individuals was diploid. All other samples from tetraploid families were tetraploid.

3.2 Phenotypic Data

Oysters were measured in fall of 2018 (York River: Nov. 26, Choptank River: Nov. 12, Nandua Creek: Dec. 12), spring of 2019 (York River: April 29, Choptank River: May 6, Nandua Creek: May 7), summer of 2019 (York River: July 1, Choptank River: July 8, Nandua Creek: July 9), and fall of 2019 (York River: Oct. 21, Choptank River: Oct. 28, Nandua Creek: Oct. 28). Tetraploid families at Choptank River were measured in fall of 2018, but because of high mortality likely caused by low salinity, the field test for tetraploid families at Choptank River was terminated in the winter of 2018.

By fall of 2019, mean total weight in triploid families was 118.3 g ± 24.4 (standard deviation) at York River, 87.3 g ± 20.9 in triploid families at Nandua Creek, 74.5 g ± 19.9 in tetraploid families at York River, and 66.5 ± 18.5 in triploid families at Choptank River (Figure 1a). Mean meat weight by fall of 2019 was 21.0 g ± 5.8 in triploid families at York River, 14.6 g ± 4.7 in triploid families at Choptank River, 13.1 ± 3.5 in triploid families at Nandua Creek, and 12.2 g ± 4.2 in tetraploid families at York River (Figure 1b) (Table 2).

Mean meat yield decreased over time in all populations (Figure 1c). Between fall of 2018 and fall of 2019, mean meat yield decreased in tetraploid families at York River (0.22 ± 0.03 to 0.16 ± 0.04), triploid families at York River (0.22 ± 0.03 to 0.18 ± 0.03), triploid families at Nandua Creek (0.19 ± 0.03 to 0.15 ± 0.03), and triploid families at Choptank River (0.24 ± 0.04 to 0.22 ± 0.04). The mean meat yield in tetraploid families at Choptank River in fall of 2018 was 0.23 ± 0.04.

Mean width index ranged from 0.60 ± 0.06 to 0.72 ± 0.08, and mean height index ranged from 0.28 ± 0.04 to 0.35 ± 0.05 (Figure 1d, 1e). Averaged over all sampling periods, mean width index was 0.69 ± 0.08 in triploid families at Choptank River and tetraploid families at Choptank River, 0.67 ± 0.07 in triploid families at York River, 0.65 ± 0.07 in tetraploid families at York River, and 0.63 ± 0.07 in triploid families at Nandua Creek. Mean height index averaged over all sampling periods was 0.34 ± 0.05 in tetraploid families at York River, 0.31 ± 0.04 in triploid families at York River and Nandua Creek, 0.30 ± 0.05 in tetraploid families at Choptank River (fall of 2018 only), and 0.29 ± 0.04 in triploid families at Choptank River (Table 2).

3.3 Genetic Correlations Among Sampling Periods

Genetic correlations among sampling periods varied from 0.34 to > 1. For triploid families in the York River, genetic correlations between fall of 2018 and subsequent sampling periods tended to decrease over time, such as for total weight, meat yield, meat weight, and width index (Appendix 1). For triploid families in the Choptank River, there was a nearly opposite pattern; for total weight, meat weight, and height index, genetic correlations between measurements in fall of 2018 and subsequent sampling periods tended to increase over time. Genetic correlations among sampling periods for Nandua Creek were high (0.83 to > 1) for total weight and meat weight except for between fall of 2018 (the start of the field trial) and fall of 2019 (the end of the field trial), when genetic correlations were especially low (total weight: 0.62; meat weight: 0.34). Additionally, many of the genetic correlation estimates at Nandua Creek had large standard errors (> 0.50). Genetic correlations among sampling periods in tetraploid families at York River were high (≥ 0.90) for all traits. Genetic correlations for tetraploid families at Choptank River were not available because of the early mortality that terminated the field test.

3.4 Pooling data

Data from summer of 2019 and fall of 2019 were pooled and used for estimates of heritability and genetic correlation. Data were pooled for these sampling periods because genetic correlations between these periods were high in all populations. Data were pooled so estimates of heritability and genetic correlation would be more robust.

3.5 Heritability

Narrow-sense heritability in triploid families ranged from 0.50 to 0.64 for total weight, 0.10 to 0.34 for meat weight, 0.43 to 0.76 for meat yield, 0.10 to 0.33 for width index, and 0.47 to 0.73 for height index (Table 3). Heritability estimates were statistically different from 0 (95%

confidence intervals did not include 0) for all traits in triploids except meat weight and width index in triploids at Nandua Creek. Narrow-sense heritability in tetraploid families at York River was 0.40 for total weight, 0.14 for meat weight, 0.22 for weight index, and 0.34 for height index. Heritability estimates were only statistically different from 0 for total weight and height index in tetraploid families.

3.6 Between-trait Correlations

Genetic correlations between traits for weight (total weight and meat weight) and between traits for shape (width index and height index) were positive and statistically significant from 0. Genetic correlations between total weight and meat weight ranged from 0.67 to 0.78, and genetic correlations between width index and height index ranged from 0.55 to 0.76 (Table 4).

Genetic correlations between size and shape traits in triploid families at York River and between total weight and meat yield in triploid families at Choptank River and Nandua Creek were negative and statistically significant from 0. For triploid families at York River, genetic correlations between total weight and width index (-0.61), total weight and height index (-0.66), and meat weight and height index (-0.69) were negative. Genetic correlations between total weight and meat yield were -0.55 for triploids at Choptank River and -0.93 for triploids at Nandua Creek (Table 4).

3.7 Between-site Correlations in Triploids

Genetic correlations between sites were positive for all traits in triploids (Table 5). Estimates ranged from 0.64 to 0.92 between triploids at York River and Choptank River, 0.15 to 0.87 between triploids at York River and Nandua Creek, and 0.52 to 0.92 between triploids at Choptank River and Nandua Creek (Table 5).

3.8 Between-Ploidy Correlations

All genetic correlations between tetraploid families at York River and triploid families at York River, Choptank River, and Nandua Creek were positive (Table 6). Genetic correlations between tetraploids and triploids were high for total weight, ranging from 0.75 (vs. triploids at York River) to 1.15 (vs. triploids at Nandua Creek) with standard errors ranging from 0.19 to 0.30. For meat weight, genetic correlations varied from 0.32 to 1.50 and had especially high standard errors (ranging from 0.48 to 0.75). Genetic correlations for meat yield ranged from 0.30 to 1.04 (standard errors: 0.21 to 0.30). For width index, between-ploidy genetic correlations varied from 0.14 to 0.61 (standard errors: 0.29 to 0.43) and for height index, 0.51 to 0.93 (standard errors 0.21 to 0.32) (Table 6).

3.9 Correlations with Survival

Genetic correlations between survival and the traits measured in this chapter varied from -0.65 to 0.94 (Table 7). For triploid families at York River and Choptank River, genetic correlations between survival and meat yield were negative (-0.65 and -0.64) and statistically significant from 0. The genetic correlation between survival and meat weight for triploid families at Nandua Creek was positive (0.94) and statistically significant from 0.

3.10 Correlations with Late Spring Survival at Nandua Creek

Late spring survival was measured in triploid families as the interval survival from May of 2019 to July of 2019 (Chapter 2). Genetic correlations were estimated between late spring survival in triploid families at Nandua Creek and the other traits (total weight, meat weight, meat yield, width index, and height index) measured at each sampling time (fall 2018, spring 2019, summer 2019, and fall 2019) at Nandua Creek. All genetic correlations were positive except for a correlation with width index in fall of 2018 (-0.05, standard error: 0.44). The only correlation estimates significantly different 0 were with meat weight in fall of 2018 (0.75, standard error: 0.31) and meat weight in spring of 2019 (0.56, standard error: 0.27) (Table 8).

4. Discussion

4.1 Heritability and Genetic Correlation in Triploids

Total weight, meat weight, meat yield, and shape traits in triploids have a substantial additive genetic basis. Estimates of heritability were high for total weight, meat yield, and height index (0.50 to 0.73) and moderate for meat weight and width index (0.10 to 0.34). The heritabilities in triploids are higher than that reported for diploid *C. virginica* (Allen et al., 2021), but are less precise, principally because hundreds of more families were analyzed in Allen et al. (2021). Additional trials can provide more precise estimates of heritability; however, these initial results are promising that size and quality traits in triploids are under substantial additive genetic control and can be significantly improved through selective breeding.

Unfavorable genetic correlations were identified between size and shape traits in triploids. Genetic correlations between total weight and width index and between meat weight and width index were moderately high and negative (-0.66, -0.69) for triploids at the York River. Similarly, Ward et al. (2005) and Kube et al. (2011) found a negative genetic correlation between width index and size traits in diploid *C. gigas*, and Allen et al. (2021) found a negative genetic correlation between total weight and shape traits in diploid *C. virginica*. The unfavorable correlations indicate that selecting for only size can result in selecting for more oysters of unacceptable shape, which could have negative economic consequences for farmers (e.g., Kube et al. 2011). A breeding strategy to maintain shape while improving size (e.g., Allen et al., 2021; Ward et al., 2005) may be optimal for triploid *C. virginica*.

Meat yield had an unfavorable genetic correlation with several traits in triploids. Genetic correlations between meat yield and total weight were negative for triploids at Choptank River and Nandua Creek, as were genetic correlations between meat yield and survival for triploids at York River and Choptank River. Most of the estimates were moderately high (-0.55, -0.64, -0.65), however the correlation between meat yield and total weight at Nandua Creek was very high (-0.93). The highly unfavorable genetic correlations indicate selecting for meat yield in triploid *C. virginica* would result in a substantial sacrifice to size and survival. Thus, selecting for meat yield in triploids may only be advantageous if it is highly influential to the economic value of the oysters. Meat yield may be an imprecise measure of the quality and economic value of farmed oysters in the Chesapeake Bay and thus not worth selecting in triploids. Meat yield is the ratio of meat weight to total weight, and it is an imprecise measure of quality because it is affected by both changes in meat weight and shell weight. Thus, selecting for higher meat yield may result in lower shell weight, and decreasing shell weight may

be undesirable for the high end 'boxed' and 'half-shell' market to which most hatchery-based oysters in Virginia are sold (Hudson, 2018). Lighter and likely thinner shells may lower the value of oysters for these markets because it may impair the ability to shuck oysters without damaging the shell. In fact, the ability to shuck oysters has been identified as a quality attribute to food retailers and restaurant chefs from the supply chain of oysters farmed in the Chesapeake Bay (Love et al., 2020). Increasing the meat quality without decreasing shell quality may be possible by selecting for a trait that more accurately measures the ratio of meat relative to the volume of the shell cavity.

Quality related to the meat to shell ratio may be measured accurately and with high throughput with a condition index. A representative measurement for quality related to meat may be the extent the volume of the oyster cavity is occupied by soft tissue, or the 'condition' of the oyster. Many methods of measuring condition, such as a water displacement method (e.g., Galtsoff, 1964), are not practical for a large-scale oyster breeding program because estimating the volume of the cavity is time consuming. However, Lawrence and Scott (1982) demonstrated a time efficient method of estimating cavity volume that could be suitable for the high throughput of animals in oyster breeding programs. Vu et al. (2020) recently applied the principle from Lawrence and Scott (1982) in a condition index for *Crassostrea angulata* as the ratio of wet meat weight to cavity volume and found the index was variable (CV = 0.50) and heritable (0.10 ± 0.03). Provided it is variable and heritable in triploid *C. virginica*, selecting for a higher condition index may be a good option to improve the meat quality of farmed oysters in the Chesapeake Bay.

Selecting for higher condition in triploids may improve resistance to "triploid mortality." A strong and favorable genetic correlation existed between overall survival and meat weight at

Nandua Creek (0.93, standard error = 0.23). The correlation implies association between meat weight and "triploid mortality" because most of the variance in survival at Nandua Creek was due to a triploid mortality event (Chapter 2). Genetic correlations involving a more specific estimate of survival during the triploid mortality event, late spring survival at Nandua Creek, also revealed a positive association with meat weight. Genetic correlations between late spring survival and meat weight in fall of 2018 and meat weight in spring of 2019 were positive and statistically significant from 0, indicating triploid families with higher meat weight tended to have higher survival in late spring. Meat weight is not a measurement of condition *per se* because it is affected by the extent the volume of the oyster cavity is occupied by soft tissue (condition) and size. However, genetic correlations with late spring survival at Nandua Creek were stronger for meat weight than total weight, suggesting condition, and not size, may be genetically associated with resistance to "triploid mortality." Future studies can further examine a genetic relationship between condition and "triploid mortality" with a phenotype that more exactly represents condition, such as the ratio of wet meat weight to cavity volume from Vu et al. (2020).

All genetic correlations between sites were lower than unity (1), indicating some level of GxE for all traits in triploids. Of the 15 between-site genetic correlations in triploids, 0 were above 0.95, 2 were above 0.90, and 5 were above 0.80. The traits measured in this chapter did not have the same GxEs as survival in Chapter 2, where a weak GxE existed between York River and Choptank River (0.97) and a moderate GxE existed between Nandua Creek and the other sites (0.46) (Chapter 2). In the size and quality traits in this chapter, GxEs were found across all sites. A moderate GxE (< 0.80) existed between each combination of sites for at least one trait.

The GxEs in triploid traits means selective breeding strategy can largely influence the genetic improvement made for each environment. The major selective breeding strategies of improving traits in two environments were originally outlined by James (1961) and are relevant to developing strategy for triploid *C. virginica*. One strategy is to only select animals for one environment. Maximum genetic improvement will be made for the selected environment because all breeding effort will be allocated to it, and provided the genetic correlation between sites is positive, improvements will also be made in the other environment. This single environment strategy may be optimal if genetic correlations between sites are high (weak GxE) or if improvement in one environment is prioritized, neither of which fit well with breeding for triploid *C. virginica* in the Chesapeake Bay. Other strategies involve directly selecting for improvement at each site — a single population could be selected for improvement in both sites or two separate populations could be selected for improvement at each site. The strategy that results in greater improvement at each site is largely based on the between-site genetic correlation (James, 1961; Mulder et al., 2006; Mulder and Bijma, 2005). Mulder et al. (2006) simulated both strategies and found that if the between-site genetic correlation was higher than 0.61, the single population strategy led to higher improvement in both environments. The findings of Mulder et al. (2006) are conceptually useful; however, the correlation values distinguishing superiority of each strategy are unlikely to be appropriate guides for breeding of *C. virginica*. Mulder et al. (2006) simulated a progeny-testing scheme, which has little in common with the current family-based breeding of *C. virginica* in the Chesapeake Bay (Allen et al., 2021). Sae-Lim et al. (2016) suggested the genetic correlation at which the single population strategy is superior may be higher for family-based breeding; however, this has not been simulated. Simulating the single and multiple population approach for a family-based scheme

would be valuable to make data-driven decisions on how to select triploid *C. virginica* across the Chesapeake Bay.

4.2 Heritability and Genetic Correlation Involving Tetraploids

Estimates of heritability for traits in tetraploids were moderate and similar to that in diploid *C. virginica* (Allen et al., 2021). The estimates suggest improvement in tetraploid traits is possible through selective breeding, which could ultimately result in improvement in triploids. The extent improvements in tetraploids will result in improvement of triploids depends largely on the genetic correlation between tetraploid and triploid traits.

Genetic correlations between traits in tetraploids and triploids were always positive, but imprecise. Standard errors ranged from 0.19 to 0.75, and thus most estimates could not be statistically distinguished as positive or negative (95% confidence intervals included positive and negative numbers). The between-ploidy correlations in this chapter inform how selecting for tetraploid traits would affect traits in triploid offspring. For example, if genetic correlations were high, tetraploids selected for faster growth would be expected to yield triploid offspring with faster growth. More precise estimates of the genetic correlations between tetraploids and triploids would be valuable in determining the value in directly improving tetraploid traits to improve triploids.

Precise genetic correlations between ploidies (e.g., tetraploid and triploid) may require a vast amount of data. Bijma and Bastiaansen (2014) examined factors influencing precision of purebred-crossbred genetic correlations, which are similar to between-ploidy correlations. Sires

in crossbreeding schemes (i.e., tetraploids) are crossed to dams of their own line (i.e., tetraploids) and dams of another line (i.e., diploids) and genetic correlations between pure lines and hybrids (i.e., tetraploids and triploids) are estimated. An important variable affecting precision in Bijma and Bastiaansen (2014) was the reliability of the estimated breeding value of the sires, which was shown to be largely dependent on the number of dams mated per sire. The higher the number of dams per sire, the higher the reliability of the estimated breeding value of the sires and the higher the precision. For breeding operations at ABC, it is typical to mate two dams per sire (Allen et al., 2021), which may mean precision of between-ploidy correlations will remain low even with a large amount of data. Using the equation for calculating the standard error of the genetic correlation from Bijma and Bastiaansen (2014) with values appropriate for polyploid breeding at ABC regarding dams per sire (2), offspring measured per full-sib family (45), heritability (0.3), common family effect (0.02), and value of true correlation (0.70), the standard error did not get below 0.05 until 750 half-sib families were tested (i.e. 1500 full sib families of tetraploids and 1500 full sib families of triploids).

Even with precise estimates of between-ploidy genetic correlations, the better approach to improve triploids may not be obvious. The two major types of approaches are 1) select diploids and tetraploids based solely on their performance or 2) incorporate performance of triploids into the breeding program. At extreme values for the genetic correlations, the superior breeding approach is apparent. For example, if the genetic correlations were very high (> 0.95), there would be little value in measuring triploids, and selecting diploids and tetraploids based solely on their performance would be optimal. If the genetic correlations were very low or negative, diploid and tetraploid performance would not correlate well with triploid performance, making measurements of triploids an important part of breeding for triploid

improvement. Less clear are the relative outcomes of each approach if the genetic correlations are moderate (0.50 – 0.90). This question is the subject of the next chapter, wherein approaches for triploid improvement are evaluated through simulation with genetic correlation as a variable.

5. Tables

Table 1. Salinity, latitude, longitude, and gear used at each site triploid and tetraploid families of *Crassostrea virginica* were tested.

Site	Salinity*	Latitude	Longitude	Gear
York River (YR)	13-17	37.248	-76.499	longline
Choptank River (CR)	6-8	38.593	-76.128	longline
Nandua Creek (ND)	11-15	37.626	-75.861	bottom cages

*Represents the first and third quartile of data from June 1, 2017 to October 28, 2018 (YR, CR) or December 11, 2017 to October 28, 2018 (ND). Data for ND collected by a conductivity logger (Chapter 2). Data for YR from Virginia Estuarine and Coastal Observing System measurement station 37.247, -76.499. Data for CR Horn Point Laboratory at the University of Maryland in Cambridge, Maryland.

Table 2: Sample mean ($\bar{x}$), standard deviation (SD), and coefficient of variation (CV) for tetraploids (4N) and triploids (3N) of *Crassostrea virginica* measured at various sites in summer and fall of 2019 when oysters were 27 to 30 months old. YR= York River, CR= Choptank River, ND= Nandua Creek.

		Summer 2019				Fall 2019			
		4N YR	3N YR	3N CR	3N ND	4N YR	3N YR	3N CR	3N ND
Total Weight (g)	$\bar{x}$	49.1	66.2	32.3	64.4	74.5	118.3	66.5	87.3
	SD	13.0	14.9	9.3	15.9	19.9	24.4	18.5	20.9
	CV	0.26	0.22	0.29	0.25	0.27	0.21	0.28	0.24
Meat Weight (g)	$\bar{x}$	--	15.9	7.5	10.4	12.2	21.0	14.6	13.1
	SD	--	4.5	2.8	2.7	4.2	5.8	4.7	3.5
	CV	--	0.28	0.37	0.26	0.34	0.27	0.32	0.27
Meat Yield	$\bar{x}$	--	0.24	0.23	0.16	0.16	0.18	0.22	0.15
	SD	--	0.04	0.04	0.03	0.04	0.03	0.04	0.03
	CV	--	0.17	0.17	0.19	0.25	0.17	0.18	0.20
Width Index	$\bar{x}$	0.66	0.69	0.68	0.66	0.64	0.65	0.72	0.60
	SD	0.07	0.07	0.08	0.07	0.06	0.07	0.08	0.06
	CV	0.10	0.09	0.11	0.10	0.10	0.10	0.11	0.10
Height Index	$\bar{x}$	0.34	0.31	0.29	0.32	0.34	0.29	0.28	0.31
	SD	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04
	CV	0.13	0.14	0.15	0.13	0.13	0.13	0.14	0.13

Table 3. Narrow-sense heritability (h^2) with standard errors in parentheses for traits in tetraploid families (4N) and triploid families (3N) of *Crassostrea virginica* measured at various sites in summer and fall of 2019 when oysters were 27 to 30 months old. YR= York River, CR= Choptank River, ND= Nandua Creek.

	4N YR	3N YR	3N CR	3N ND
Total Weight	0.40 (0.19)	0.50 (0.15)	0.64 (0.20)	0.53 (0.17)
	0.08 (0.07)	--	--	--
Meat Weight	0.14 (0.13)	0.34 (0.12)	0.33 (0.12)	0.10 (0.07)
	0.10 (0.06)	--	--	--
Meat Yield	0.30 (0.15)	0.54 (0.16)	0.76 (0.20)	0.43 (0.15)
	0.01 (0.06)	--	--	--
Width Index	0.22 (0.12)	0.31 (0.12)	0.26 (0.10)	0.10 (0.06)
	0.04 (0.04)	--	--	--
Height Index	0.34 (0.15)	0.73 (0.21)	0.54 (0.17)	0.47 (0.15)
	0.04 (0.06)	--	--	--

Table 4. Within-population genetic correlations with standard errors in parentheses for traits in tetraploid families (4N) and triploid families (3N) of *Crassostrea virginica* measured at various sites in summer and fall of 2019 when oysters were 27 to 30 months old. YR= York River, CR= Choptank River, ND= Nandua Creek.

Trait 1	Trait 2	4N YR	3N YR	3N CR	3N ND
Total Weight	Meat Weight	0.78 (0.16)	0.67 (0.14)	0.74 (0.12)	0.77 (0.16)
Total Weight	Meat Yield	-0.35 (0.36)	-0.39 (0.23)	-0.55 (0.19)	-0.93 (0.08)
Total Weight	Width Index	-0.26 (0.34)	-0.61 (0.18)	-0.26 (0.27)	0.36 (0.29)
Total Weight	Height Index	-0.14 (0.35)	-0.66 (0.15)	0.23 (0.26)	0.18 (0.26)
Meat Weight	Meat Yield	0.37 (0.45)	0.42 (0.22)	0.14 (0.27)	-0.45 (0.38)
Meat Weight	Width Index	-0.23 (0.51)	-0.25 (0.27)	-0.18 (0.28)	0.54 (0.36)
Meat Weight	Height Index	-0.36 (0.52)	-0.69 (0.16)	-0.08 (0.28)	0.23 (0.34)
Meat Yield	Width Index	0.07 (0.39)	0.43 (0.23)	0.19 (0.26)	-0.14 (0.33)
Meat Yield	Height Index	-0.11 (0.37)	-0.05 (0.26)	-0.44 (0.21)	-0.10 (0.27)
Width Index	Height Index	0.76 (0.17)	0.64 (0.16)	0.72 (0.14)	0.55 (0.24)

Table 5. Among-site genetic correlations with standard errors in parentheses for traits in triploid families of *Crassostrea virginica* measured at various sites in summer and fall of 2019 when oysters were 27 to 30 months old.

Trait	Site 1	Site 2	
Total Weight	York River	Choptank River	0.76 (0.14)
	York River	Nandua Creek	0.58 (0.19)
	Choptank River	Nandua Creek	0.88 (0.10)
Meat Weight	York River	Choptank River	0.64 (0.19)
	York River	Nandua Creek	0.15 (0.35)
	Choptank River	Nandua Creek	0.92 (0.25)
Meat Yield	York River	Choptank River	0.88 (0.07)
	York River	Nandua Creek	0.87 (0.09)
	Choptank River	Nandua Creek	0.64 (0.17)
Width Index	York River	Choptank River	0.85 (0.13)
	York River	Nandua Creek	0.46 (0.29)
	Choptank River	Nandua Creek	0.52 (0.28)
Height Index	York River	Choptank River	0.92 (0.07)
	York River	Nandua Creek	0.76 (0.14)
	Choptank River	Nandua Creek	0.88 (0.09)

Table 6. Between-ploidy genetic correlations with standard errors in parentheses for traits in triploid (3N) and tetraploid (4N) families of *Crassostrea virginica* measured at various sites in summer and fall of 2019 when oysters were 27 to 30 months old.

Trait	Site of 4N	Site of 3N	
Total Weight	York River	York River	0.75 (0.30)
		Choptank River	0.92 (0.22)
		Nandua Creek	1.15 (0.19)
Meat Weight	York River	York River	0.32 (0.52)
		Choptank River	0.36 (0.48)
		Nandua Creek	1.50 (0.75)
Meat Yield	York River	York River	0.67 (0.25)
		Choptank River	0.30 (0.30)
		Nandua Creek	1.04 (0.21)
Width Index	York River	York River	0.61 (0.29)
		Choptank River	0.39 (0.37)
		Nandua Creek	0.14 (0.43)
Height Index	York River	York River	0.93 (0.21)
		Choptank River	0.73 (0.26)
		Nandua Creek	0.51 (0.32)

Table 7: Genetic correlations between overall survival and other traits within triploid (3N) and tetraploid (4N) families of *Crassostrea virginica* measured at three sites in the Chesapeake Bay (York River, Choptank River, Nandua Creek). Survival was measured between summer of 2018 and fall of 2019, while all other traits were measured in summer and fall of 2019.

Trait 1	Trait 2	4N York River	3N York River	3N Choptank River	3N Nandua Creek
Survival	Total Weight	0.25 (0.27)	0.33 (0.26)	0.49 (0.22)	0.32 (0.24)
Survival	Meat Weight	0.51 (0.40)	-0.13 (0.31)	0.05 (0.30)	0.94 (0.23)
Survival	Meat Yield	0.33 (0.31)	-0.65 (0.20)	-0.64 (0.19)	0.15 (0.26)
Survival	Width Index	0.04 (0.32)	-0.16 (0.30)	0.10 (0.29)	0.59 (0.25)
Survival	Height Index	0.18 (0.24)	0.32 (0.30)	0.42 (0.25)	0.23 (0.25)

Table 8: Genetic correlations between late spring survival and other traits measured in fall 2018, spring 2019, summer of 2019, and fall 2019 in triploid families of *Crassostrea virginica* at Nandua Creek. Standard errors are in parentheses.

Fall 2018	Total Weight	0.32 (0.26)
	Meat Weight	0.75 (0.31)
	Meat Yield	0.16 (0.28)
	Width Index	-0.05 (0.44)
	Height Index	0.15 (0.28)
Spring 2019	Total Weight	0.30 (0.29)
	Meat Weight	0.56 (0.27)
	Meat Yield	0.14 (0.30)
	Width Index	0.44 (0.23)
	Height Index	0.30 (0.33)
Summer 2019	Total Weight	0.28 (0.26)
	Meat Weight	1.07 (0.59)
	Meat Yield	0.16 (0.28)
	Width Index	0.86 (0.48)
	Height Index	0.20 (0.29)
Fall 2019	Total Weight	0.07 (0.26)
	Meat Weight	0.57 (0.30)
	Meat Yield	0.26 (0.25)
	Width Index	0.33 (0.34)
	Height Index	0.17 (0.25)

6. Figures

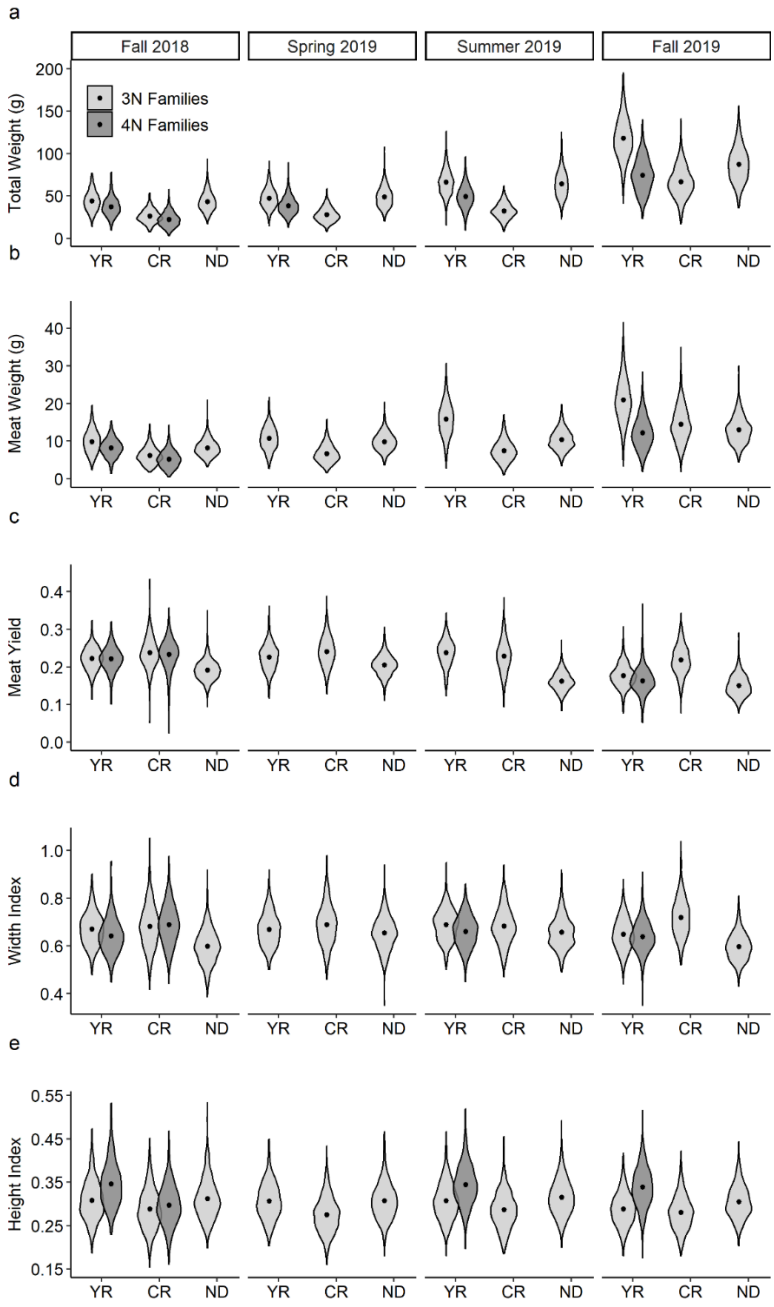


Figure 1. Violin plots of total weight (g), meat weight (g), meat yield, width index, and height index in triploid (3N, light gray) and tetraploid (4N, dark gray) families measured in the fall of 2018 and spring, summer, and fall of 2019 at the York River (YR), Choptank River (CR), and Nanda Creek (ND). Violins demonstrate the distribution of the data and are produced using kernel density estimation. Black dots represent the mean.

7. Appendix

Appendix 1: Genetic Correlations Among Sampling Periods

The format of the data below is as follows:

	Trait, Time 2			Trait, Time 3			Trait, Time 4		
Trait, Time 1	r_g	±	se	r_g	±	se	r_g	±	se
Trait, Time 2				r_g	±	se	r_g	±	se
Trait, Time 3							r_g	±	se

where genetic correlations (r_g) and standard errors (se) are specified for each row and column combination. Time 1, 2, 3, and 4 correspond to fall 2018, spring 2019, summer 2019, and fall 2019, respectively.

Triploid Families, York River

	Total Weight 2	Total Weight 3	Total Weight 4
Total Weight 1	1.03 ± 0.03	0.92 ± 0.06	0.81 ± 0.11
Total Weight 2		0.96 ± 0.06	0.79 ± 0.12
Total Weight 3			0.93 ± 0.06

	Meat Weight 2	Meat Weight 3	Meat Weight 4
Meat Weight 1	0.98 ± 0.06	0.92 ± 0.08	0.83 ± 0.16
Meat Weight 2		0.92 ± 0.09	0.72 ± 0.19
Meat Weight 3			0.87 ± 0.13

	Meat Yield 2	Meat Yield 3	Meat Yield 4
Meat Yield 1	1.03 ± 0.05	1.01 ± 0.07	0.93 ± 0.11
Meat Yield 2		0.95 ± 0.07	0.85 ± 0.11
Meat Yield 3			0.87 ± 0.09

	Width Index 2	Width Index 3	Width Index 4
Width Index 1	0.97 ± 0.06	0.92 ± 0.10	0.88 ± 0.15
Width Index 2		0.91 ± 0.09	0.73 ± 0.16
Width Index 3			0.98 ± 0.08

	Height Index 2	Height Index 3	Height Index 4
Height Index 1	1.10 ± 0.06	1.12 ± 0.08	1.15 ± 0.08
Height Index 2		1.11 ± 0.05	1.08 ± 0.05
Height Index 3			1.11 ± 0.06

Triploid Families, Choptank River

	Total Weight 2	Total Weight 3	Total Weight 4
Total Weight 1	0.88 ± 0.11	0.96 ± 0.07	0.98 ± 0.06
Total Weight 2		0.96 ± 0.07	0.89 ± 0.10
Total Weight 3			0.92 ± 0.07

	Meat Weight 2	Meat Weight 3	Meat Weight 4
Meat Weight 1	0.79 ± 0.19	1.19 ± 0.14	1.02 ± 0.12
Meat Weight 2		0.91 ± 0.13	0.75 ± 0.18
Meat Weight 3			0.85 ± 0.14

	Meat Yield 2	Meat Yield 3	Meat Yield 4
Meat Yield 1	0.97 ± 0.07	0.96 ± 0.08	0.89 ± 0.08
Meat Yield 2		0.98 ± 0.08	0.89 ± 0.08
Meat Yield 3			0.88 ± 0.09

	Width Index 2	Width Index 3	Width Index 4
Width Index 1	1.09 ± 0.10	0.88 ± 0.13	0.83 ± 0.16
Width Index 2		1.02 ± 0.13	0.55 ± 0.25
Width Index 3			0.70 ± 0.20

	Height Index 2	Height Index 3	Height Index 4
Height Index 1	0.84 ± 0.25	1.10 ± 0.15	1.08 ± 0.15
Height Index 2		1.12 ± 0.11	0.80 ± 0.18
Height Index 3			1.06 ± 0.06

Triploid Families, Nandua Creek

	Total Weight 2	Total Weight 3	Total Weight 4
Total Weight 1	0.98 ± 0.15	0.83 ± 0.13	0.62 ± 0.20
Total Weight 2		1.16 ± 0.12	1.06 ± 0.12
Total Weight 3			0.89 ± 0.09

	Meat Weight 2	Meat Weight 3	Meat Weight 4
Meat Weight 1	1.35 ± 0.38	1.11 ± 0.47	0.34 ± 0.53
Meat Weight 2		1.60 ± 0.75	0.87 ± 0.40
Meat Weight 3			0.90 ± 0.51

	Meat Yield 2	Meat Yield 3	Meat Yield 4
Meat Yield 1	1.10 ± 0.19	0.72 ± 0.22	1.03 ± 0.17
Meat Yield 2		1.03 ± 0.17	1.00 ± 0.14
Meat Yield 3			1.08 ± 0.11

	Width Index 2	Width Index 3	Width Index 4
Width Index 1	0.48 ± 0.49	1.67 ± 1.08	1.92 ± 1.11
Width Index 2		0.90 ± 0.48	0.82 ± 0.32
Width Index 3			1.38 ± 0.74

	Height Index 2	Height Index 3	Height Index 4
Height Index 1	1.25 ± 0.29	1.28 ± 0.15	1.01 ± 0.08
Height Index 2		1.40 ± 0.34	1.23 ± 0.26
Height Index 3			1.16 ± 0.10

Tetraploid Families at York River

Note: no data was collected for meat weight in tetraploids at YR for spring 2019 or summer 2019, or for height index or width index for spring of 2019.

	Total Weight 2	Total Weight 3	Total Weight 4
Total Weight 1	0.98 ± 0.05	0.94 ± 0.06	1.15 ± 0.20
Total Weight 2		0.93 ± 0.06	1.07 ± 0.14
Total Weight 3			NA

*NA = bivariate model (Total Weight 3 and Total Weight 4) failed to converge

	Meat Weight 4
Meat Weight 1	1.15 ± 0.36

	Meat Yield 4
Meat Yield 1	0.81 ± 0.15

*Family term excluded from model for Meat Yield (failed to converge)

	Width Index 3	Width Index 4
Width Index 1	1.10 ± 0.18	1.19 ± 0.29
Width Index 3		0.97 ± 0.12

	Height Index 3	Height Index 4
Height Index 1	0.91 ± 0.36	1.11 ± 0.34
Height Index 3		0.90 ± 0.13

Chapter 4: Simulations of breeding strategies for genetically improving triploid *Crassostrea virginica*

1. Introduction

Computer simulations can be used to evaluate alternative breeding strategies quickly and inexpensively. A breeding strategy is the design in which organisms are mated, phenotyped, and selected for the purpose of a breeding goal, such as to increase profitability of an industry. The outcomes of alternative breeding strategies may not be obvious from quantitative genetic theory alone because many interacting factors can affect the result. Measuring outcomes empirically would provide insight, but would also likely require running multiple breeding programs, which for many species would be impractical. Outcomes of breeding strategies can be evaluated relatively quickly and at low cost through simulations. Simulations have been used for assessing strategies for animal and plant breeding (e.g., Hickey and Gorjanc, 2012; Rutten et al., 2002; Wang et al., 2003), but likely not been applied to many species important to agriculture, possibly because designing and running these simulations has required expertise in computer programming, genetics, and breeding programs (Gaynor et al., 2020).

New technology has made simulating breeding programs more accessible. Gaynor et al. (2020) recently produced a package for R (R Core Team, 2019), AlphaSimR, with the intention of making simulations of breeding programs easier to design and execute. The package simulates stochastic biological processes, such as genetic recombination, and allows a wide range of breeding scenarios to be examined. Simulations in AlphaSimR can be designed piece by piece using R functions and users can interact directly with simulated populations in the R environment, both of which support flexibility and gradual learning. The package also has numerous functions and user-controlled specifications, making it flexible to various design

schemes and genetic qualities of simulated organisms, including the creation and mating of polyploids.

Polyploidy is an essential component of the selective breeding of the eastern oyster, *Crassostrea virginica*, in the Chesapeake Bay. Triploid *C. virginica* are the commercial aquaculture product — nearly all the hatchery-produced oysters in the region are triploid (Hudson, 2019) and are produced by mating diploids to tetraploids (e.g., Guo et al., 1996). Diploids and tetraploids are selectively bred by the Aquaculture Genetics and Breeding Technology Center (ABC) at the Virginia Institute of Marine Science (VIMS) with the goal of increasing the growth and survival of oysters cultured in the Chesapeake Bay (Allen et al., 2021). Given the popularity of triploids as the commercial product, genetically improving triploids made from mating diploids and tetraploids is a primary component of the breeding goal.

Multiple strategies exist for genetically improving triploids made from mating diploids and tetraploids. One option, currently used at ABC, is to select diploids and tetraploids from performance of diploids and tetraploids. Alternatively, performance of triploids could be incorporated into the breeding program, meaning diploids and tetraploids are selected based on the performance of diploids, tetraploids, and triploids. The options available for improvement of triploids are analogous to those in crossbreeding, wherein different lines or species are crossed with the intent of maximizing performance of the crossbred animal (e.g., Bijma and Van Arendonk, 1998; Hamilton et al., 2009; Lutaaya et al., 2001). Comparing strategies for crossbreeding has been the subject of previous studies (Bijma and Van Arendonk, 1998; Wei and Van Der Werf, 1994), however strategies for breeding polyploid animals have not been evaluated, likely because breeding polyploid animals is unique to oysters.

The strategies for improving triploids differ in their dependence on positive genetic correlations between ploidies. If traits are genetically correlated, then selectively breeding for change in one trait will also result in a change in the correlated trait. The current strategy at ABC of selecting diploids and tetraploids based solely on performance of diploids and tetraploids relies on genetic correlation – triploids will only improve if there is a positive genetic correlation between the traits in diploids and tetraploids and the traits in triploids. The alternative strategy of incorporating triploids into the breeding program can result in improvement of triploids regardless of the between-ploidy genetic correlation. Even at low or negative genetic correlations, diploids or tetraploids with high performing triploid relatives can be selected, resulting in diploids and tetraploids that are genetically improved for triploid production. The effect of the between-ploidy genetic correlation distinguishes the strategies, making the genetic correlation a critical factor in evaluating which strategy is better.

Breeding strategies for genetically improving crossbred animals have been simulated with genetic correlation as a variable (e.g., Bijma and Van Arendonk, 1998; Kerr et al., 2004; Wei and Van Der Werf, 1994). With livestock as the model, Wei and Van Der Werf (1994) compared a "pure-line selection" method (PLS), analogous to the strategy of selecting diploids and tetraploids based on their performance, with a "combined crossbred and purebred selection" method (CCPS), analogous to the single breeding strategy using information from diploids, triploids, and tetraploids for selection. From simulations, Wei and Van Der Werf (1994) found the PLS method only resulted in greater genetic improvement at genetic correlations greater than 0.70. Additionally, the CCPS method was increasingly better than the PLS as the genetic correlation decreased (Wei and Van Der Werf, 1994).

Based on quantitative genetic theory (Falconer and Mackay, 1996) and previous simulations in analogous scenarios (Wei and Van Der Werf, 1994), the breeding strategy resulting in greater genetic improvement of triploids is apparent at extreme values for the between-ploidy genetic correlation. At low or negative genetic correlations between ploidies, incorporating triploids into the breeding program would be the better strategy. Performance of diploids and tetraploids would not correlate well with triploid performance, and thus performance of triploid relatives would be essential in informing which diploids and tetraploids were genetically best for producing triploids. At very high genetic correlations (≥ 0.95), a parallel breeding program involving only diploids and tetraploids would be superior. Performance of diploid, triploid, and tetraploid relatives would be highly correlated, meaning the best diploids and tetraploids for triploid production could be accurately selected without also evaluating triploid performance. If not increasing accuracy in selection, producing triploids would be disadvantageous because it would allocate the resources of the breeding program away from producing diploids and tetraploids, the only candidates for mating. The higher the number of candidates for mating, the higher the intensity of selection that can be applied (Falconer and Mackay, 1996). At very high genetic correlations, the higher intensity of selection would cause a higher rate of genetic improvement in the parallel breeding program strategy (Falconer and Mackay, 1996; Wei and Van Der Werf, 1994).

Simulations would be valuable in comparing strategies for polyploid breeding at moderate values for the between-ploidy genetic correlations. It is unclear which strategy results in the greater improvement in triploids at moderate genetic correlations (0.30-0.90), and it is likely that the between-ploidy genetic correlations are moderate for many commercial traits in *C. virginica*. Genetic correlations were estimated between traits in tetraploid and triploid *C.*

virginica in Chapter 2 and 3 and spanned the range of moderate values, varying from 0.30 to greater than 1. Evaluating breeding strategies by simulation at moderate genetic correlations can serve as a guide for the long-term strategy of polyploid breeding in *C. virginica*, and perhaps other oyster species.

The objective of this study was to compare different strategies for selectively breeding diploid and tetraploid *C. virginica* for the genetic improvement of triploids. Two breeding strategies were evaluated through simulation: 1) separate diploid and tetraploid breeding programs, with no phenotyping of triploids, and 2) a single breeding program phenotyping diploids, triploids, and tetraploids. The primary aim was to determine the values of between-ploidy genetic correlation at which each strategy was superior in genetically improving triploids. Strategies were compared with genetic correlation as a variable, and family selection was used to imitate the breeding programs at ABC. It was hypothesized that parallel diploid and tetraploid breeding programs would be superior at high genetic correlations (0.65-0.80), and a single breeding program phenotyping diploids, triploids, and tetraploids would be superior at lower genetic correlations (< 0.60). Additionally, two other research aims were pursued: 1) how the result of the breeding strategies would compare if a combination of family and individual selection, or combined selection, was applied to diploid and tetraploid populations and 2) how the response to selection in diploids and tetraploids would compare in each breeding strategy. Comparisons of breeding strategies were based on the same resource allocation from the breeding program regarding the total number of families and individuals produced and measured each year. All simulations were carried out in AlphaSimR (Gaynor et al., 2020).

2. Methods

Breeding values and phenotypes were simulated using the software package AlphaSimR 0.13.0 (Gaynor et al., 2020). A breeding value is the transmittable (additive) genetic value of an individual that is the basis of selection for most animal breeding programs.

2.1 Founders and Simulation Parameters

The first step in designing the simulation was creating founder haplotypes and setting simulation parameters. Founder haplotypes are sets of DNA variations and they were used to produce the genomes of individuals in the first generation. Founder haplotypes were created with the runMacs function (details in Appendix A1). Next, simulation parameters were set (details in Appendix A1). Simulation parameters consisted of 1) the mean and variance of the breeding value of traits, 2) the heritability (always narrow-sense heritability) of traits, and 3) the genetic correlation among traits. Three traits were assigned to each individual: "diploid trait," "triploid trait," and "tetraploid trait." For all simulations, all three traits were assigned a mean breeding value of 0 and variance of 1. Founder populations of diploids and tetraploids were created from the founder haplotypes and simulation parameters. The same founder haplotypes were used for each simulation. Individuals of founder populations (founders) were mated to produce diploid, triploid, and tetraploid families.

2.2 Breeding Strategies

Simulations focused on testing two breeding strategies for genetic improvement of triploids. One strategy selected diploids and tetraploids in separate breeding programs, relying on the genetic improvement in each to improve the triploid – the "No Triploid Sibs" strategy.

Only diploids and tetraploids were produced and phenotyped in the No Triploid Sibs strategy; 74 diploid and 76 tetraploid full-sib families (150 in total) were produced and phenotyped each year. Full-sib families were produced using a partly factorial mating design (Gjedrem, 2005) in which each dam and sire were crossed with two different mates (Figure 1). The second strategy involved phenotyping diploids, tetraploids, and triploids in a single breeding program – the "Triploid Sibs" strategy. Fifty full-sib triploid, diploid, and tetraploid families were produced and tested each year in the Triploid Sibs strategy (150 total, same total as in the "No Triploid Sibs" strategy). Using a partly factorial mating design, each diploid dam was crossed to two tetraploid sires and each tetraploid sire was crossed to two diploid dams (Figure 2). Each diploid dam was also crossed to two diploid sires, and each tetraploid sire was crossed to two tetraploid dams (Figure 2). In result, triploids in the Triploid Sibs strategy were related to diploids and tetraploids as offspring, half-sibs, and more distant relatives.

2.3 Breeding Program Parameters

All simulations had the same family size, generation interval, and lifespan. Full-sib families comprised 50 individuals and were organized into year classes. A set of 150 full-sib families, produced as either 74 diploid and 76 tetraploid families (No Triploid Sibs strategy) or 50 diploid, 50 triploid, and 50 tetraploid families (Triploid Sibs strategy) made up a year class. The first three year classes were produced from the founder populations. Individuals were only used for mating once, and thus the first three year classes of the No Triploid Sibs strategy originated from 222 diploid and 228 tetraploid founders. The Triploid Sibs strategy originated from 225 diploid and 225 tetraploid founders. The generation interval in the simulation was one year. For example, individuals of year class 3 were candidates to produce year class 4. To

imitate a lifespan, only the three youngest year classes were candidates to produce the next year class. For example, sires and dams used to produce year class 5 could only be from year class 2, 3, and 4.

2.4 Phenotypes

Phenotypes were automatically generated for each trait in individuals. AlphaSimR first generated breeding values for each trait in individuals, then generated phenotypes for individuals based on the heritability (Gaynor, 2019) (example in Appendix A2). Phenotypes were used to estimate breeding values. In practice, the breeding value of an individual is unknown and is estimated based on phenotypes.

2.5 Pedigree

Typically, a phenotype alone is an unreliable indicator of the breeding value of an individual. Including phenotypes of relatives can increase the accuracy in estimation of breeding values because it provides more information on the genetic value of the individual. The degree phenotypes from relatives increase accuracy is based upon the heritability of the trait as well as the proportion of alleles shared between relatives, or the additive genetic relationship. Additive genetic relationships among individuals can be derived from the pedigree based on the degree of relationship by descent, or coancestry, and included in a statistical model as a covariance matrix, referred to as the additive relationship matrix. Typically a pedigree file, listing each individual and their sire and dam, is converted into an additive relationship matrix following rules for diploid organisms (Henderson, 1976). Coancestry among individuals is different in a pedigree of polyploids, especially with multiple ploidy levels, e.g., diploid, triploid, and tetraploid. For example, a triploid will be more closely related to its tetraploid parent than its

diploid parent. To have accurate relationship matrices for pedigrees with all levels of ploidy, matrices were generated using the R package polyAinv (Hamilton and Kerr, 2018), which applies rules that are appropriate for populations of any ploidy level, including populations of multiple ploidy levels.

2.6 Estimating Breeding Values

Breeding values were estimated from phenotypes and relationship matrices using linear individual models in ASReml (Gilmour et al. 2015). In an individual model (also referred to as animal model), breeding values are estimated from phenotypes and additive relationship matrices. Breeding values are estimated for every individual in the pedigree.

Breeding values were separately estimated for "diploid trait" and "tetraploid trait" for the No Triploid Sibs strategy. A pedigree of either the diploid population or tetraploid population was used with the following univariate individual model:

$$y = \mu + \text{Animal} + \varepsilon \qquad (1)$$

where y is a vector of phenotypes (e.g., phenotypes for "diploid trait"), μ is the mean of the phenotypes, Animal is a vector of breeding values (e.g., breeding values for "diploid trait" in diploids), and ε is the residual variation.

For the Triploid Sibs strategy, breeding values for "diploid trait," "triploid trait," and "tetraploid trait" were estimated simultaneously in all individuals using a three trait multivariate model with the same terms from model 1. A pedigree of all individuals – diploids, triploids, and tetraploids – was used with the model. With the multivariate model, a genetic covariance matrix was fitted among traits, or inter-trait covariance matrix. The inter-trait covariance matrix estimates genetic correlations among traits and incorporates the genetic correlations into the

estimation of breeding values. A non-zero genetic correlation indicates a breeding value for one trait is informative of the breeding value of another trait, and therefore, phenotypic information of one trait (e.g., "diploid trait") becomes informative of the breeding value of another (e.g., "triploid trait"). Thus, as long as between-ploidy genetic correlations are non-zero, phenotypic information of all traits from all relatives could be influential in estimating breeding values of "triploid trait" in the Triploid Sibs strategy.

2.7 Selecting Breeding Candidates

After the first three "years" in which founders were mated, individuals from diploid and tetraploid families were selected for mating. Individuals were selected based on the estimated breeding value of their family (Family EBV). Selecting based on Family EBV is called "between-family" selection. Family EBVs were calculated as the mean estimated breeding value of individuals within a full-sib family. For the No Triploid Sibs strategy, Family EBV for "diploid trait" was the criterion for selecting diploids and Family EBV for "tetraploid trait" was criterion for selecting tetraploids. In contrast, the criterion for selecting diploid and tetraploid families in the Triploid Sibs strategy was Family EBV for "triploid trait."

Individuals were selected from families as candidates for mating, herein referred to as breeding candidates. Selection of breeding candidates was a three-step process. First, best families were selected based on their Family EBV. Families in the 66^{th} percentile or greater for Family EBV were selected. For example, in the Triploid Sibs strategy, 50 of the 150 families available for selection (three youngest year classes, each with 50 families) were selected. Second, individuals with low phenotypic values were excluded to imitate culling of especially low performing individuals. Individuals with a phenotype of -1 were not breeding candidates,

which approximately represented the 30th percentile in the “diploid trait” and “tetraploid trait” phenotype before selection. Third, four individuals from each of the top families were randomly selected as breeding candidates. This process was the same for the No Triploid Sibs and Triploid Sibs strategies.

2.8 Mating Plans

Individuals were selected for mating by using the software package AlphaMate (Gorjanc and Hickey, 2018). From the list of breeding candidates, AlphaMate selected certain individuals and generated a mating plan. The mating plan was a list of pairs of individuals to cross and was generated from: a list of candidates and their estimated breeding values, an additive relationship matrix of the candidates, a desired mating scheme (Figure 1 or Figure 2), and a breeding target.

The breeding target refers to an objective related to genetic gain and genetic diversity, which are inversely related. Maximizing genetic gain involves selecting relatives because breeding values among relatives are correlated. Relatives contributing to the next generation increases the coancestry in the population, which leads to inbreeding and consequent loss of genetic diversity (Woolliams et al., 2015). A breeding target was set to balance genetic gain with loss of genetic diversity to imitate a breeding goal of genetic improvement over a long time period. For all simulations, the breeding target set in AlphaMate was 50% of the maximum coancestry, corresponding to 87% of the maximum genetic gain (Gorjanc and Hickey, 2018).

2.9 Comparisons

Simulations were used to compare the No Triploid Sibs and Triploid Sibs in three scenarios, explained below in subsections: 1) genetic correlations as a variable, 2) using a combination of family and individual selection, and 3) comparing the response to selection between diploids and tetraploids. Each scenario consisted of simulations replicated two to three times. A single simulation was carried out through six rounds of selection (diagram Figure 3). Replicate simulations had identical simulation parameters. The outcome for comparison was the true breeding value (provided by AlphaSimR) for the "triploid trait" in the diploid population and tetraploid population (i.e., the commercial brood stock). The mean breeding value for "triploid trait" was calculated within diploid and tetraploid populations after the sixth round of selection, when the populations consisted of individuals from year class 7, 8, and 9.

2.9.1 Scenario 1: Genetic Correlations and Breeding Strategies

The No Triploid Sibs and Triploid Sibs breeding strategies were compared as breeding methods to improve triploids with different genetic correlations between 1) "triploid trait" and "diploid trait" and 2) between "triploid trait" and "tetraploid trait." Simulations were replicated at each of the following genetic correlation values: 0.33, 0.45, 0.55, 0.65, 0.75, 0.80, and 0.90. The heritability of each trait was set based on empirical data for the underlying narrow-sense heritability of survival in *C. virginica*: 0.30 for "diploid survival" (Allen et al., 2021), 0.32 for "triploid survival" (Chapter 2), and 0.24 for "tetraploid survival" (Chapter 2).

2.9.2 Scenario 2: Combined Selection

Selecting specific individuals within top families was examined by simulation. Selecting specific individuals within families is called "within-family" selection. When used in combination

with between-family selection by selecting the best individuals from the best families, the method is referred to as “combined selection.” Combined selection was carried out in simulations by selecting the four individuals with the highest estimated breeding value within each top ranked family. Additionally, an estimated breeding value specific to the individual was provided to AlphaMate to generate mating plans.

Combined selection was used as a comparison with between-family selection. Breeding values from simulations of between-family and combined selection were compared for the Triploid Sibs and No Triploid Sibs breeding strategies. For the comparison, genetic correlations between “diploid trait” and “triploid trait” and between “triploid trait” and “tetraploid trait” were set at 0.75. The heritability was 0.30 for “diploid trait,” 0.32 for “triploid trait,” and 0.24 for “tetraploid trait.”

2.9.3 Scenario 3: Response to Selection Based on Ploidy

The No Triploid Sibs and Triploid Sibs breeding strategies were compared when heritability for all traits was equal and genetic correlations between traits were equal. By making heritability and genetic correlations the same for each trait, the effect of ploidy (diploid or tetraploid population) on the response to selection (breeding value for “triploid trait”) could be examined. Heritability for “diploid trait,” “triploid trait,” and “tetraploid trait” was set to 0.30. The genetic correlation between “diploid trait” and “triploid trait” and between “tetraploid trait” and “triploid trait” was 0.90.

2.10 Tracking Genetic Variation and Measuring Inbreeding

Genetic variation was tracked in diploid and tetraploid populations during the simulations of the No Triploid Sibs strategy from 2.9.3. "Diploid trait" was tracked in the diploid population and "tetraploid trait" was tracked in the tetraploid population because they were the traits subject to selection. Additive genic variation, additive genetic variation, and heritability were recorded for the founder population and for each year class (the code for extracting parameters is shown in Appendix A3). Additive genic variation was tracked because it represented the underlying additive genetic variation in the population. The additive genic variation is the additive genetic variation when the population is in gametic-phase equilibrium. A population is in gametic-phase equilibrium if alleles of different loci are randomly associated, that is, the frequency of allele associations (e.g., allele A from locus 1, allele B from locus 2) in gametes are a product of the frequency of each allele in the population (Lynch and Walsh, 1998). For example, two loci would be in gamete phase equilibrium if allele A at locus 1 had a frequency of 0.5, allele B at locus 2 had a frequency of 0.5, and allele A and B occur together in gametes at a frequency of 0.25.

The mean coefficient of inbreeding was recorded in the diploid and tetraploid populations in which genetic variation was tracked. The coefficient of inbreeding is the probability that two alleles at any locus in an individual are identical by descent, and for breeding purposes is an important indicator of loss of genetic diversity. The mean coefficient of inbreeding was calculated from the full pedigree of diploids and full pedigree of tetraploids in each replicate simulation (founders and year class 1 through 9). The coefficient of inbreeding was calculated using the polyAinv package (Hamilton and Kerr, 2018).

3. Results

3.1 Genetic Correlations and Breeding Strategies

The influence of genetic correlations on breeding values for "triploid trait" are depicted in Figure 4. The mean breeding value for "triploid trait" varied from 0.9 to 3.2 in diploids (Figure 4a) and from 0.8 to 2.3 in tetraploids (Figure 4b). At genetic correlations between 0.33 and 0.66, the Triploid Sibs (TS – open symbols) breeding strategy resulted in a higher mean breeding value in both diploids and tetraploids. The advantage of the Triploid Sibs strategy was greatest at the lowest genetic correlation of 0.33 — the mean breeding value was 2.5x greater in diploids and 2.0x greater in tetraploids. At genetic correlations higher than 0.66, the Triploid Sibs and No Triploid Sibs (NTS) strategies resulted in similar breeding values. The average difference between the breeding strategies at genetic correlations > 0.66 was 1% in diploids and 0.7% in tetraploids. The mean breeding value for "triploid trait" was consistently higher in diploids than tetraploids — on average, the mean breeding value was 1.2x greater in diploids. For diploids, breeding values increased with genetic correlation in a positive linear manner. In tetraploids, breeding value and genetic correlation resembled a logarithmic relationship because increases in breeding value got smaller as the genetic correlation increased.

3.2 Combined Selection

The response to selection from between-family and combined selection are shown in Figure 5. The mean breeding value was always higher with combined selection regardless of the breeding strategy or ploidy of the selected population (Figure 5). Relative to family selection, combined selection increased breeding values more in diploids than in tetraploids. In diploids, combined selection resulted in a 1.7x higher breeding value on average. In tetraploids, breeding

values were only 1.2x higher on average. The highest mean breeding value in diploids resulted from the No Triploid Sibs strategy and combined selection, which was 1.1x greater than the Triploid Sibs strategy with combined selection. In tetraploids, there was little difference between the Triploid Sibs and No Triploid Sibs strategy when combined selection was used (Figure 5).

3.3 Breeding Value Based on Ploidy

The response to selection was compared in diploid and tetraploid populations when all traits had the same heritability (0.30), and the genetic correlation between "triploid trait" and "diploid trait" as well as between "triploid trait" and "tetraploid trait" was set at 0.90. On average, the mean breeding value was 1.5x higher in diploids than tetraploids (Figure 6). The mean breeding value in diploids was similar between the Triploid Sibs strategy (3.0) and the No Triploid Sibs strategy (3.2). The mean breeding value in tetraploids was also similar between strategies — 2.4 for Triploid Sibs and 2.3 for No Triploid Sibs.

3.4 Genetic Parameters and Inbreeding

Genetic parameters were tracked in "diploid trait" in a diploid population and "tetraploid trait" in a tetraploid population during simulations with the No Triploid Sibs strategy. The simulations were from Scenario 3 (Methods 2.9.3), and thus each trait was set with the same heritability (0.30) and additive genetic variation (1.0) at the start of the simulation. Diploid trait in diploids and tetraploid trait in tetraploids were subject to selection in these simulations.

Additive genic variation – The additive genic variation is the additive genetic variation when the population is in gametic-phase equilibrium. Additive genic variation was higher in

diploid trait throughout the simulation (Figure 7a). In the founder populations, mean additive genic variation was 0.90 ± 0.02 (standard error) in diploid trait and 0.47 ± 0.04 in tetraploid trait. Additive genic variation changed little in each trait before the first selection (first selection occurred between year class 3 and year class 4). From year class 4 to the end of the simulation, mean additive genic variation decreased by 9% in diploid trait and 6% in tetraploid trait.

Additive genetic variation – In only tetraploid trait, additive genetic variation decreased sharply before selection (Figure 7b). In the founder population of tetraploids (f), tetraploid trait had an additive genetic variation of 1, but after one generation of random mating, the mean additive genetic variation was 0.61 ± 0.02. Additive genetic variation in tetraploid trait was higher in year class 2 (0.72 ± 0.04) and year class 3 (0.64 ± 0.07). Before selection, additive genetic variation in diploid trait ranged from 0.93 ± 0.04 to 1.07 ± 0.03 (Figure 7b).

After selection began, additive genetic variation in both diploid trait and tetraploid trait decreased sharply. Mean additive genetic variation in year class 4 was 0.75 ± 0.02 in diploids and 0.36 ± 0.001 in tetraploids, corresponding to a 19% decrease in diploids and 44% decrease in tetraploids between year class 3 and year class 4. From year classes 4 to 9, additive genetic variation decreased 12% in diploids and 0% in tetraploids.

Heritability – Heritability followed the same pattern as the additive genetic variation in each trait. In year class 1, 2, and 3, the heritability was on average 0.30 ± 0.01 in diploids and 0.22 ± 0.01 in tetraploids. From year class 4 through 9, the mean heritability was 0.23 ± 0.005 in diploids and 0.14 ± 0.002 in tetraploids.

The mean coefficient of inbreeding was recorded in the same diploid and tetraploid populations in which genetic parameters were tracked. By the end of the simulation, the diploid

population had an inbreeding coefficient of 0.027 ± 0.001, which was 2.1x larger than that in tetraploids (0.013 ± 0.002).

Discussion

4.1 Strategies and Genetic Correlation

Alternative strategies exist for breeding diploid and tetraploid oysters for the improvement of triploids and their effectiveness can vary depending on underlying genetic factors. Two breeding strategies were examined through simulation in AlphaSimR (Gaynor et al., 2020): 1) produce and phenotype only diploid and tetraploid families, "No Triploid Sibs," and 2) produce and phenotype diploid, triploid, and tetraploid families, "Triploid Sibs." A primary focus of this chapter was to evaluate the outcomes of these approaches by simulation. In the first scenario, the strength of genetic correlation between the trait targeted for improvement (triploid trait) and secondary traits (diploid trait and tetraploid trait) was examined for influence on breeding value for "triploid trait" in each strategy.

The Triploid Sibs strategy resulted in higher breeding values when the genetic correlation between the target trait (triploid trait) and secondary traits (diploid trait and tetraploid trait) was low (0.33 – 0.45) or moderate (0.55 – 0.66). The Triploid Sibs strategy was better at low genetic correlations, as hypothesized, because the response to selection in that strategy is less dependent on the genetic correlation. The Triploid Sibs strategy follows a direct response model (Falconer and Mackay, 1996). In the direct response model, genetic correlations between the target and secondary traits only influence the response to selection by affecting the accuracy of selection (r_{uY}, Table 1). As the absolute value of the genetic correlation increases, the accuracy (r_{uY}, Table 1) increases because secondary traits become more

informative of breeding values for the target trait. In the No Triploid Sibs strategy, selection is on secondary traits and therefore follows a correlated response model. Breeding values are estimated without information on the target trait, and thus the genetic correlation with the target trait does not affect the accuracy (r_{uX}, Table 1). However, the genetic correlation (rg_{YX}, Table 1) is multiplied directly to other terms deciding the response to selection (accuracy, intensity of selection, and standard deviation of breeding values, Table 1) and can therefore dictate the magnitude and sign (positive or negative) of the response. When the genetic correlation was low (< 0.66), the direct effect of the genetic correlation on the response to selection in the No Triploid Sibs strategy resulted in lower breeding values.

The Triploid Sibs and No Triploid Sibs strategy resulted in about the same breeding values when genetic correlations between the target and secondary traits were between 0.75 and 0.90. Breeding values likely converged because the advantages of each strategy were balanced at high genetic correlations. The advantage of the No Triploid Sibs strategy was that of the 150 families produced in each strategy, more diploid and tetraploid families were produced each year in the No Triploid Sibs strategy, causing a higher intensity of selection (*i,* Table 1) and thus increasing the response to selection. For the Triploid Sibs strategy, the advantage was that the accuracy (r_{uY}, Table 1) was expected to be higher than the product of the accuracy ($r_{uX,}$,Table 1) and genetic correlation ($rg_{YX,}$, Table 1) in the No Triploid Sibs strategy. When the genetic correlation was between 0.75 and 0.90, the advantages of each of these strategies canceled each other out, resulting in a similar outcome.

4.2 Selection Method

Combined selection was examined as a method to increase the response to selection in the Triploid Sibs and No Triploid Sibs strategy. In between-family selection, the method for the first set of simulations, breeding values are estimated for the family, not the individual. Individuals are randomly or haphazardly selected from the best families and may be selected from a wide distribution of breeding values because the genetics of full sibs can vary substantially. Full sibs share alleles but vary in breeding value because the alleles each individual receives from the parents are a matter of chance. Full sibs share *on average* 50% of the same alleles, but the actual proportion of shared alleles varies. For example, in humans, the 95% confidence interval for shared alleles ranges from 37% to 63% among full sibs (Speed and Balding, 2015). This variation in allele transmission, and the variance in breeding values resulting from it, means that half of all additive genetic variation in a population resides within families (Falconer and Mackay, 1996; Gjedrem, 2012). If selection on individual criteria (e.g., phenotypes) was used along with family criteria in the form of combined selection, the expectation would be higher accuracy in selecting the genetically best individuals, resulting in a higher response to selection.

The benefit of combined selection differed in diploids and tetraploids. For diploids, combined selection substantially increased the response to selection. Diploids selected with combined selection had a 1.7x greater response to selection compared to between-family selection. In tetraploids, combined selection had a relatively minor impact, likely because genetic variation in tetraploids was half that in diploids (explained in section 4.3). Interestingly, breeding values in diploids for the No Triploid Sibs strategy were higher than for the Triploid Sibs strategy under combined selection, but the same under between-family selection, at a genetic correlation of 0.75. The findings suggest the Triploid Sibs and No Triploid Sibs strategy were

equally accurate in selecting the top families (between-family selection), but the No Triploid Sibs strategy was more accurate in selecting the top individuals within families (within-family selection).

For diploids, within-family selection was more accurate with the No Triploid Sibs strategy because the data distinguishing sibs was directly informative of the individual breeding values. Individuals within families were differentiated by their phenotypes. In the No Triploid Sibs strategy, the phenotypes of diploids directly informed breeding values because selection was on the trait of the phenotype, "diploid trait." In the Triploid Sibs strategy, the phenotypes of diploids were less informative because selection was on "triploid trait." Using within-family selection, the difference in accuracy in a direct form (No Triploid Sibs) or indirect form (Triploid Sibs) depends on the genetic correlation between the traits of interest, in this case "diploid trait" and "triploid trait." Thus, within-family selection with the No Triploid Sibs strategy (direct form) was more accurate because "diploid trait" and "triploid trait" were not perfectly correlated (genetic correlation = 0.75).

Combined selection can result in a higher response to selection than between-family selection, yet for many traits in oysters, within-family selection is challenging. Within-family selection is simple if traits can be measured on live breeding candidates, such as total weight. Other commercial traits cannot be measured on live breeding candidates. In some cases, measurement requires killing the animal, which eliminates their possible use for mating. Also survival, when treated as a continuous variable (Falconer and Mackay, 1996), cannot be physically measured on candidates. One way within-family selection is possible for traits that are not measurable on breeding candidates is through a correlated response. If a trait measurable in live candidates, like total weight, and a trait not measurable in live candidates,

like meat weight, are genetically correlated, then selecting individuals for total weight will also result in selection for the genetically correlated trait. The effectiveness of the correlated response is dependent on the genetic correlation, and thus is a limited method of within-family selection.

Within-family selection can be directly applied to traits that cannot be measured on live candidates through genomic selection (Meuwissen et al., 2001). Genomic selection is using genome-wide genetic markers to estimate breeding values. The effect of each marker on the trait(s) of interest is estimated from a reference population of individuals that are genotyped and phenotyped. The breeding value, in this case 'genomic breeding value,' can then be estimated in individuals outside the reference population using their genotype and the estimated marker effects. Genomic selection has been integrated into breeding of livestock species (Meuwissen et al., 2016), and improvements in selection accuracy have been found with genomic selection for many aquaculture species (reviews by Hollenbeck and Johnston, 2018; Houston et al., 2020; Zenger et al., 2019), including oysters (*Crassostrea gigas*: Gutierrez et al., 2020, 2018). Incorporation of genomic selection for oysters, as well as many other aquaculture species with relatively small regional or local breeding programs, may depend on cost-effective genotyping strategies such as low density marker panels (Kriaridou et al., 2020; Lillehammer et al., 2013).

The within-family selection in the simulations in this chapter resembled genomic selection on diploid trait and tetraploid trait, but not genomic selection on triploid trait. Within-family selection was possible in the simulation because each individual had a phenotype. The phenotypes in diploids and tetraploids functioned as if marker effects for the diploid trait were applied to diploids and marker effects for the tetraploid trait were applied to tetraploids.

Genomic selection on triploid traits would have meant applying marker effects for the triploid trait to diploids and tetraploids and would be analogous to diploids and tetraploids having a phenotype for the triploid trait. Individuals only had phenotypes for the trait corresponding to their ploidy, and thus genomic selection on triploid traits was not simulated in this chapter. Genomic selection could enable direct selection on triploid traits in diploids and tetraploids with high accuracy and is therefore a useful subject for future simulation studies in polyploid breeding.

Polyploids may be challenging to select using genomic methods because of complexities in genotyping. The assignment of genotype classes at a locus in polyploids involves a higher risk of error (Grandke et al., 2016) because the difference among genotypes becomes more difficult to detect as ploidy increases. For polyploids, options exist to call genotypes as if they were diploids (0,1,2), assign polyploid genotypes (for tetraploid: 0,1,2,3,4), or use continuous genotype values, which can be applied as a ratio of sequencing depth of the alleles (de Bem Oliveira et al., 2019). de Bem Oliveira et al. (2019) used each genotyping strategy to estimate breeding values for traits in tetraploid blueberry *Vaccinium corymbosum*, and for most traits found little difference in accuracy among the methods, thus all may be worth evaluating for polyploid oysters.

4.3 Response to Selection in Tetraploids and Tracking Genetic Variation

Compared to diploids, tetraploids had a lower response to selection in all scenarios. The lower response to selection in tetraploids was the impetus to track characteristics of diploid and tetraploid populations during replicate simulations (Figure 7). To compare characteristics based on ploidy, traits in diploid and tetraploid founder populations were set with identical genetic

parameters (heritability, genetic variation, and mean genetic value) in the set of simulations (Scenario 3, Methods 2.9.3). Three features of the underlying genetic structure of the population were tracked: additive genic variation, additive genetic variation, and heritability. The additive genic variation represents the additive genetic variation if the population was randomly mating and was expected to largely determine the additive genetic variation in the population.

The additive genic variation was halved when tetraploids were originally made in AlphaSimR by the 'doubleGenome' function. The doubleGenome function imitates doubling of the genome, or somatic doubling, that can be induced in plants using colchicine (Blakeslee and Avery, 1937; Johnstone, 1939). Importantly, the doubling of the genome left the breeding values for each individual unaffected. AlphaSimR scales allelic dosages so that a homozygous diploid locus and homozygous tetraploid locus are considered to have the same additive genetic value (Gaynor, 2019; Gaynor et al., 2020). For example, the 0, 1, and 2 genotypes in diploids (e.g., aa, Aa, AA) are equivalent to the 0, 2, and 4 genotypes in tetraploids (e.g., aaaa, aaAA, AAAA). When the difference in breeding value between homozygous genotypes is equivalent in diploids and polyploids, as is the case with the genotype scaling in AlphaSimR, the additive genic variation will be reduced by a factor of 1/gametic ploidy level (2 for tetraploids) (Gallais, 2003). Gallais (2003) referred to this as a "dilution effect" that occurs because ploidy affects the distribution of genotypes. The additional sets of alleles in polyploids creates additional heterozygous genotypes, and in consequence, the genotype frequencies will be more closely centered around the mean (Figure 8). The more centered distribution in polyploids causes a lower variance in genotypes, and thus in this study, lower additive genic variation.

The relative additive genetic variation between diploid and tetraploid populations fluctuated due to gametic-phase disequilibrium. Gametic-phase disequilibrium caused the additive genetic variation to be higher than the additive genic variation in the tetraploid founders and tetraploids in year class 1,2, and 3. Gametic-phase disequilibrium is a non-random association of alleles at different loci, and these associations can increase or decrease the additive genetic variation in the population (Lynch and Walsh, 1998). A gametic-phase disequilibrium inflating the genetic variation in tetraploids may have been due to the somatic doubling producing a higher frequency of homozygous genotypes than expected based on allele frequency. Selection caused a negative gametic-phase disequilibrium in both diploids and tetraploids. Directional selection, which was applied in this study, is expected to generate a negative gametic-phase disequilibrium and thereby reduce the additive genetic variation, a phenomenon referred to as the Bulmer effect (Bulmer, 1971). Immediately after the Bulmer effect, in year class 4, the additive genetic variation in the tetraploids was roughly half of that in diploids, and the ratio remained that way for year class 5 through 9.

Lower additive genetic variation caused tetraploids to have a lower response to selection compared to diploids. Less additive genetic variation results in a lower potential response to selection because it means a lower variance in breeding values. The lower benefit from applying combined selection in tetraploids compared to diploids had a similar cause — lower heritability in tetraploids. Phenotypes of individuals were used for the within-family selection component of combined selection, and at low heritabilities, phenotypes are an inaccurate measure of breeding value. The low additive genetic variation caused low heritabilities in tetraploids and made the within-family component of combined selection relatively inaccurate, resulting in a low benefit from combined selection.

In practice, the dilution effect alone is not sufficient to expect that traits in polyploid oysters will have an especially low additive genetic variation. Additive genetic variation of a trait is based on genotype frequencies, which can be affected by ploidy, but also on the genotypic values, which will be trait specific. Traits in diploid, triploid, and tetraploid oysters are distinct because ploidy changes the biology of the animal. Additionally, polyploid animals have been estimated to have substantial additive genetic variation (oysters: Chapter 2, Chapter 3, salmon: Johnson et al., 2007).

4.4 Coefficient of Inbreeding Based on Ploidy

Despite the equivalent approach to selection in diploids and tetraploids, tetraploids had a much lower inbreeding coefficient. Selection in both populations ultimately occurred when mating plans were generated with AlphaMate (Gorjanc and Hickey, 2018). Tetraploids likely had a lower inbreeding coefficient because the effect of ploidy on coancestry was not accounted for in AlphaMate. The additional number of alleles at each locus in tetraploids lowers the coancestry, a phenomenon outlined by Gallais (2003). Coancestry was computed in AlphaMate based on additive relationship matrices, which are highly similar in diploids and tetraploids and therefore do not capture the lower coancestry in tetraploids. By overestimating the coancestry in the tetraploid population, the tetraploids were selected as if there was a higher emphasis on maintaining genetic diversity in tetraploids, resulting in lower inbreeding. The higher emphasis in maintaining genetic diversity in the selection of tetraploids was likely another reason why tetraploids had a lower response to selection throughout the simulations.

The different inbreeding rate between diploid and tetraploid populations demonstrates that tetraploids need to be selected differently to target a similar rate of genetic improvement.

Applying the same methods to generate mating plans for diploids and tetraploids, as in this study, will result in targeting a lower inbreeding rate and a lower rate of genetic improvement in the tetraploid population. To target the same balance between inbreeding and genetic improvement in a tetraploid population, the selection intensity must be higher in tetraploids (Gallais, 2003).

5. Conclusion

With between family selection, a single family breeding program phenotyping diploid, triploid, and tetraploid families is likely a better strategy for triploid improvement. Across a range of positive genetic correlations between ploidy (0.33 – 0.90), a single family breeding program phenotyping diploid, triploid, and tetraploid families resulted in higher or similar genetic improvement of triploids compared to phenotyping diploid and tetraploid families in separate family breeding programs. Applying combined selection would likely produce different results across the same range of genetic correlations. With combined selection at a between-ploidy genetic correlation of 0.75, the separate breeding program strategy produced diploids with a 1.1x greater breeding value for triploid traits. Tetraploids always had a lower response to selection compared to diploids, primarily due to a "dilution effect" caused by the additional heterozygote genotypes in tetraploids (Gallais, 2003) that led to tetraploids having lower additive genetic variation. Tetraploids may have also had a lower response to selection because diploids and tetraploids were selected at the same selection intensity. A higher selection intensity is likely required in tetraploid populations to achieve the same balance of inbreeding and genetic improvement as in diploid populations (Gallais, 2003).

6. Tables

Table 1: Equation for the direct response to selection (R) from the "Triploid Sibs" strategy and for the correlated response to selection (CR) from the "No Triploid Sibs" strategy that determined genetic gains in "triploid trait" (Y) in this study. Below equations are an explanation of terms and appropriate references (Ref).

Triploid Sibs: $R_Y = i \cdot r_{uY} \cdot \sigma_{AY}$ No Triploid Sibs: $CR_Y = i \cdot r_{uX} \cdot \sigma_{AY} \cdot rg_{YX}$			
Symbol	Name	Definition	Ref.
CR_Y	Correlated response to selection	Change in trait Y if selection on trait X	1
R_Y	Response to selection	Change in trait Y from selection on trait Y	1
Y	Target trait	In this study – triploid trait	1
X	Secondary trait	In this study – diploid trait or tetraploid trait	1
i	Intensity of selection	Inversely related to proportion of population selected	1
r_u	Accuracy of selection using measure u	Correlation between true breeding value and breeding value estimated using method u	2
u_X	Method of estimating breeding values for secondary trait	In this study, univariate model using Best Linear Unbiased Prediction (BLUP)	3
u_Y	Method of estimating breeding values for target trait	In this study, multi trait model using (BLUP) *	3
rg_{YX}	Genetic correlation	Correlation of breeding values for trait Y and trait X	1
σ_A	Square root of the additive genetic variance	Standard deviation of breeding values	1

1: Falconer, D.S., Mackay, T.F.C., 1996. Introduction to Quantitative Genetics, UK: Longman Group.
2: Walsh, B., 2012. Short-Term Selection Response: Breeder's equation. lecture notes, Uppsala EQG course, version 31 Jan
3: Isik, F., Holland, J., 2017. Genetic Data Analysis for Plant and Animal Breeding.

7. Figures

	2n dams	2n1	2n2	2n3	…	2n37
2n sires		x	x	x	…	x
2n1	x	201	202			
2n2	x		203	204		
2n3	x			205		
…	x				…	…
2n37	x	274				273

	4n dams	4n1	4n2	4n3	…	4n38
4n sires		x	x	x	…	x
4n1	x	401	402			
4n2	x		403	404		
4n3	x			405		
…	x				…	…
4n38	x	476				475

Figure 1: Spawning design for the “No Triploid Sibs” (NTS) breeding strategy, producing 74 diploid (2N) families (201-274) and 76 tetraploid (4N) families (401-476) using 74 diploid individuals (2n1-2n74) and 76 tetraploid individuals (4n1-4n76). Note each diploid and tetraploid are parents of two full-sib families.

					201,202	203,204	205,206	…	249,250
				2n sires	2n1,2	2n3,4	2n5,6	…	2n49,50
					x	x	x	…	x
				2n dams	2n1	2n2	2n3	…	2n25
	4n dams		4n sires		x	x	x	…	x
401, 402	4n1,2	x	4n1	x	301	302			
403, 404	4n3,4	x	4n2	x		303	304		
405, 406	4n5,6	x	4n3	x			305		
…	…	…	…	…				…	…
449, 450	4n49,50	x	4n25	x	350				349

Figure 2: Spawning design for the “triploid sibs” (TS) breeding strategy, producing 50 diploid (2N) families (201-250), 50 triploid (3N) families (301-350), and 50 tetraploid (4N) families (401-450) using 75 diploid individuals (2n1-2n75) and 75 tetraploid individuals (4n1-4n75). Note each diploid dam and tetraploid sire are parents of four full-sib families, while each diploid sire and tetraploid dam are parents of one full-sib family.

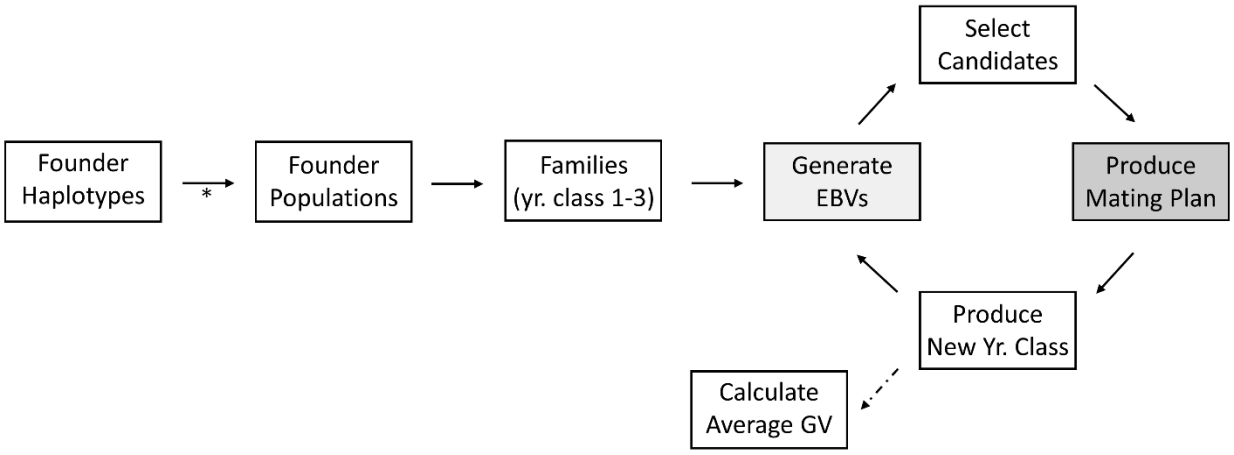


Figure 3: Flow diagram for a simulation. Boxes in white correspond to actions in AlphaSimR/R, light gray boxes correspond to actions in ASReml, and dark gray correspond to actions in AlphaMate. * Simulation parameters (see methods) were set to make founder populations from founder haplotypes. Dotted arrow (for calculating average genetic value, GV) indicates this action occurred after 6 rounds of selection.

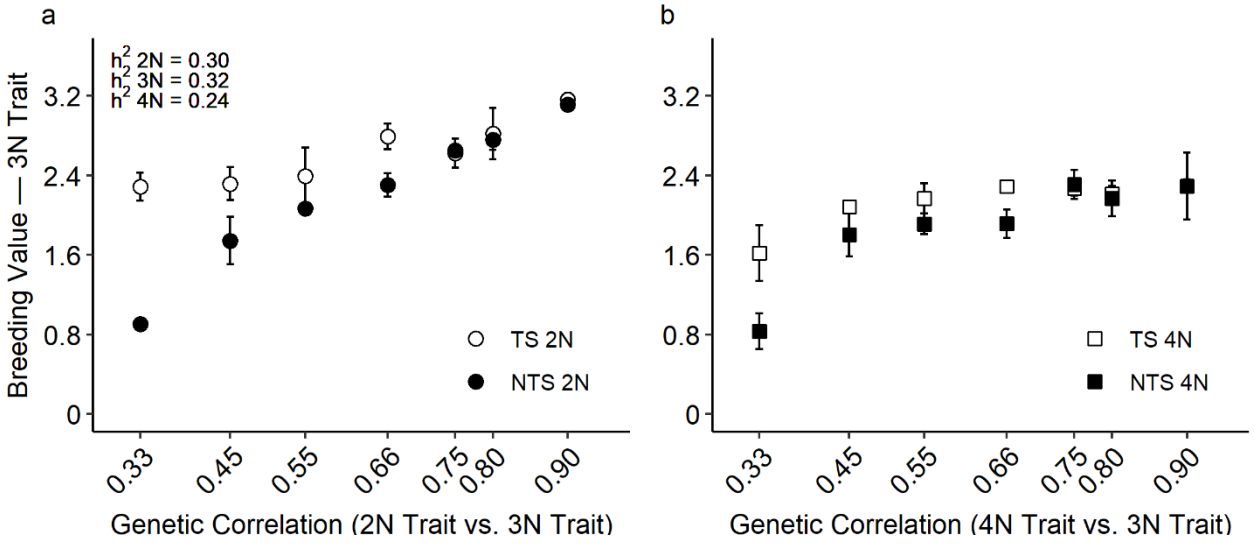


Figure 4: Mean breeding value for "triploid trait" for diploids (2N) and tetraploids (4N) selected using the Triploid Sibs (TS) or No Triploid Sibs (NTS) breeding strategy based on the genetic correlation between traits. a. Breeding value in diploids based on the genetic correlation between "diploid trait" and "triploid trait." b. Breeding value in tetraploids based on the genetic correlation between "tetraploid trait" and "triploid trait." Error bars represent standard error. h^2 2N, 3N, 4N = narrow-sense heritability for "diploid trait," "triploid trait," and "tetraploid trait," respectively.

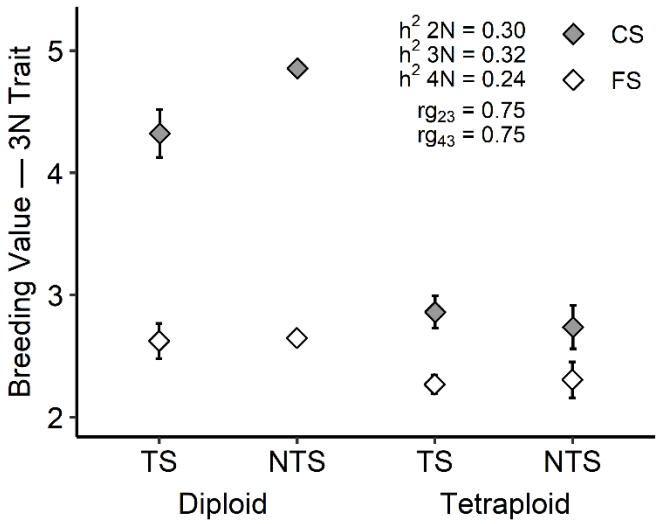


Figure 5: Mean breeding value for "triploid trait" for diploids (2N) and tetraploids (4N) selected using the Triploid Sibs (TS) or No Triploid Sibs (NTS) breeding strategy and combined selection (CS) or family selection (FS). Error bars represent standard errors. h^2 2N, 3N, 4N = narrow-sense heritability for "diploid trait," "triploid trait," and "tetraploid trait," respectively. rg23 = genetic correlation between "diploid trait" and "triploid trait." rg43 = genetic correlation between "tetraploid trait" and "triploid trait."

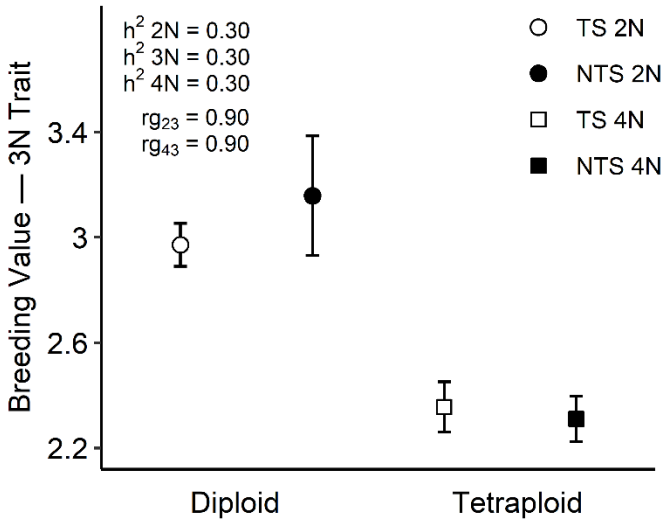


Figure 6: Mean breeding values from simulations examining the response to selection based on ploidy. a. Mean breeding value for "triploid trait" (3N Trait) for diploids (2N) and tetraploids (4N) using the Triploid Sibs (TS) or No Triploid Sibs (NTS) breeding strategy. h^2 2N, 3N, 4N = narrow-sense heritability for "diploid trait," "triploid trait," and "tetraploid trait," respectively. rg23 = genetic correlation between "diploid trait" and "triploid trait." rg43 = genetic correlation between "tetraploid trait" and "triploid trait."

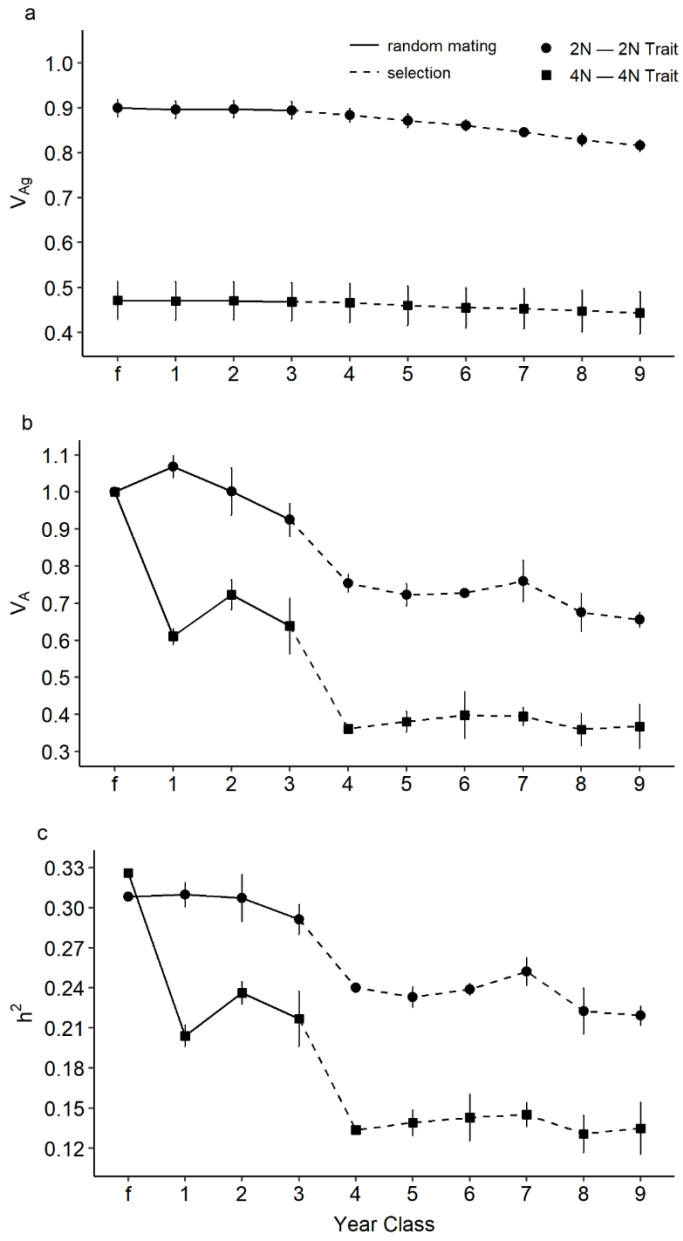


Figure 7: a. Additive genic variation (V_{AG}), b. additive genetic variation (V_A), and c. narrow-sense heritability (h^2) simulated for "diploid trait" in a diploid (2N) population and "tetraploid trait" in a tetraploid (4N) population over six rounds of selection using the No Triploid Sibs (NTS) strategy (see methods). Additive genetic variation was set to 1 and narrow-sense heritability was set to 0.30 for both traits in the diploid and tetraploid founder population (f). The first three year classes (1-3) were made from founders and thus represents random mating (solid line). Selection occurred between each subsequent year class (dashed line).

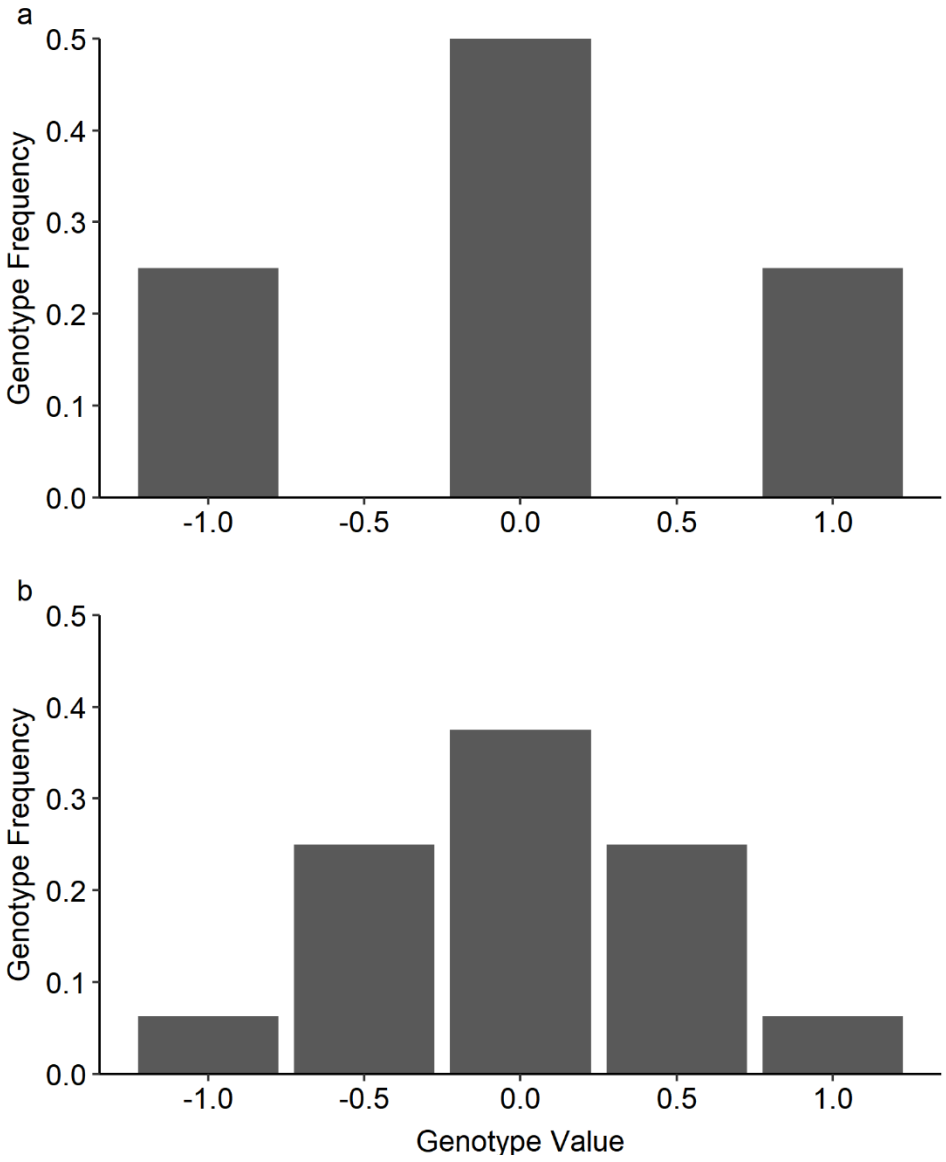


Figure 8: Frequencies of genotypes for a biallelic locus in Hardy-Weinberg Equilibrium and an allele frequency of 0.5 for a) diploids and b) tetraploids. In this example, the mean genotype value for diploids and tetraploids is 0, and the genetic variance is 0.50 for diploids and 0.25 for tetraploids. Calculation for genotype frequencies for tetraploids from Gallais (2003). In case of strictly additive effects in this study, genotype values and breeding values, as well as genetic variance and additive genetic variance, are synonymous.

8. Appendix

A1. Code used to generate founder haplotypes, simulation parameters, and founder populations for simulations in AlphaSimR 0.13.0.

1) Producing founder haplotypes

The following code was used to generate founder haplotypes (the same founder haplotypes were used for all simulations):

```
founderpop=runMacs(nInd=228,nChr=10,segSites=2000)
```

2) Simulation parameters

Simulation parameters varied among simulations and were set using the following sets of codes:

A) Code representing the mean genetic values, variance in genetic values, matrix of genetic correlations, and heritability of each trait.

```
mean234<-c(0,0,0)

var234<-c(1,1,1)

rg234=matrix(c(1,0.33,0.5,0.33,1,0.66,0.5,0.66,1),nrow=3,ncol=3)

h234<-c(0.3,0.32,0.24)
```

B) Code setting the "global simulation parameters" using the founder haplotypes and assigning QTLs for the additive traits based on the settings in A).

```
SP=SimParam$new(founderpop)

SP$addTraitA(nQtlPerChr=2000,mean=mean234,var=var234,corA=(rg234))

SP$setSexes("no")

SP$setVarE(h2=h234)
```

3) Founder populations

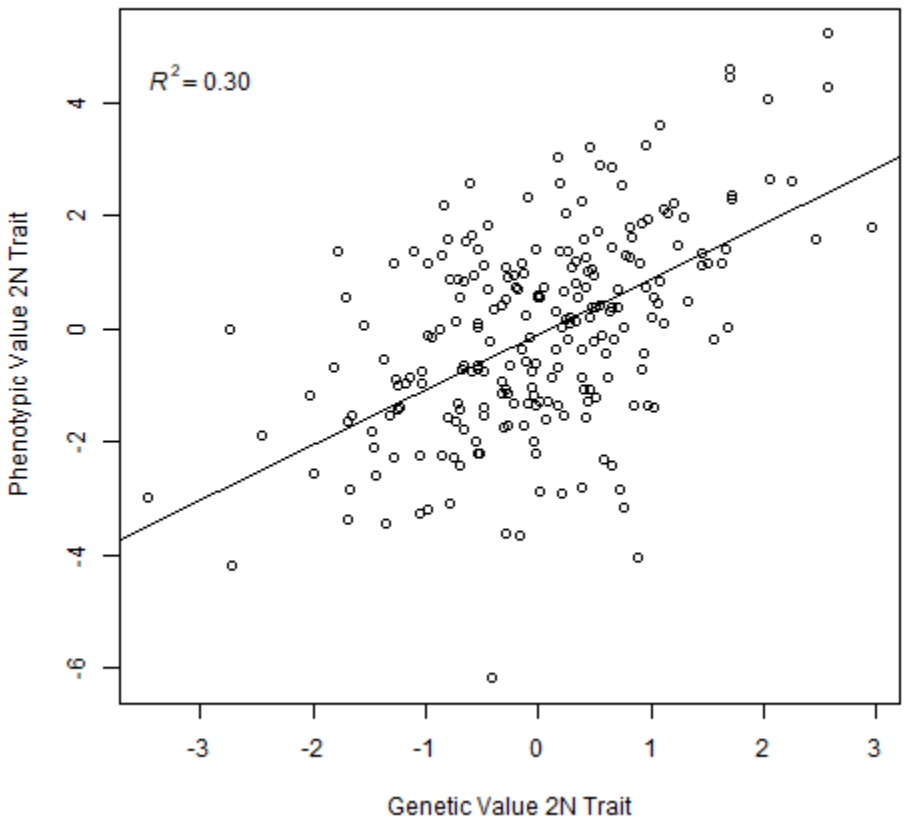
Code for generating a diploid founder population based on simulation parameters and for generating a tetraploid founder population. The tetraploid founder population was generated from a diploid founder population and the "doubleGenome" function.

```
pop2=newPop(founderpop, simParam = SP)

pop2B=newPop(founderpop, simParam = SP)

pop4= doubleGenome(pop2B, simParam=SP)
```

A2. Example of phenotypes and breeding values in a founder population.



FigureA1. Phenotypic values and breeding values of a founder population of diploids (2N) set with a genetic mean of 0, additive genetic variance of 1, and narrow-sense heritability of 0.30.

A3. Code used for output of additive genic variation, additive genetic variation, and phenotypic variation in AlphaSimR 0.13.0.

1) Additive genic variation

```
genicVarA(pop)
```

2) Additive genetic variation

```
varA(pop)
```

3) Phenotypic variation

```
varP(pop)
```

"pop" refers to population of interest

For additive genic variation, the output is a vector for each trait in the simulation. For additive genetic variation and phenotypic variation, it is a variance covariance matrix with dimensions equal to the number of traits in the simulation.

Chapter 5: Conclusion

Genetically improving triploids will likely be a goal of many oyster breeding programs because of the popularity of triploids as a commercial product. Culturing triploids may be preferred because of their fast growth rate (e.g., Dégremont et al., 2012; Harding, 2007), consistent meat quality (Allen, 1988), or as a way to inhibit genetic interactions with wild populations. Already, triploids are commercially produced in many countries, including the United States, China, Korea, Australia, New Zealand, and France. Triploids also make up a substantial portion of hatchery-produced *Crassostrea virginica* in the United States, including most of the cultured oysters in the northern Gulf of Mexico (Bodenstein et al., 2019; Wadsworth et al., 2019) and the Chesapeake Bay (Hudson, 2019). Despite the widespread use of triploids as a commercial product for oyster aquaculture, commercial traits in triploid oysters had never been analyzed using quantitative genetic methods before this dissertation, and thus the potential genetic improvement of triploids from selective breeding had not been assessed. Additionally, no prior research had been carried out to evaluate breeding strategies for triploid oyster improvement, which are multifaceted because triploids are reproductively sterile. The aim of this dissertation was to analyze the additive genetic variation in commercial traits in triploid *Crassostrea virginica* and examine strategies to selectively breed for their genetic improvement.

Methods of quantitative genetic analysis in this dissertation were different from those conventionally used in animal breeding. Traditionally, genetic relatedness within the population is estimated from a pedigree with the assumption that all individuals are diploid (Henderson, 1976). Unlike a pedigree containing all diploids, a pedigree with polyploids can contain disproportionate genetic contributions from parents to offspring, which would not be accounted

for using traditional methods. For example, a tetraploid sire contributes two thirds of the genome of a triploid offspring and one half of the genome of a tetraploid offspring. The presence of triploids and tetraploids in the pedigree in this dissertation required different methods to obtain appropriate estimations of genetic relatedness, and these were obtained using polyAinv, an R package developed by Hamilton and Kerr (2018). This dissertation represents the first application of this software to extant polyploid populations. Future quantitative genetic analyses in polyploid oysters will likely rely on applying polyAinv as demonstrated in this dissertation for best estimates of heritabilities, genetic correlations, and breeding values.

Commercial traits of the triploid population of *C. virginica* in this dissertation had substantial additive genetic variation. Heritability was often high (> 0.30) for survival, total weight, meat weight, meat yield, and shape traits. The high estimates of heritability indicate commercial aspects of triploids can be improved in future generations through selective breeding. The estimates for triploids were often higher than that reported for diploid *C. virginica* (Allen et al., 2021), however the estimates in this dissertation were based on many fewer families than the analysis of diploids in Allen et al. (2021). Additional field tests of triploid families will produce more robust estimates of heritability for triploids and make comparisons with diploids more sound.

Genetic correlations between ploidies were a major focus of this dissertation because they are critical to evaluating polyploid breeding strategies. Triploid *C. virginica* are commercially produced by mating diploids to tetraploids (e.g., Guo et al., 1996), and therefore improving triploid *C. virginica* means selectively breeding diploids and tetraploids. The original strategy at the Aquaculture Genetics and Breeding Technology Center (ABC) has been to select

diploid and tetraploid families based solely on their performance with the expectation that traits in triploids will change in parallel. The outcome of this strategy relies on the genetic correlations between the traits in diploids and tetraploids and the traits in triploids. Despite their importance to the effectiveness of polyploid breeding strategies, genetic correlations between oysters of different ploidy had never been estimated prior to this dissertation.

Genetic correlations estimated in this dissertation suggested selecting tetraploids based on their performance results in triploid improvement. Genetic correlations were estimated between traits in triploid and tetraploid oysters by testing triploid and tetraploid families related by tetraploid sire. The correlations were always positive, indicating that improving traits in tetraploids will result in a positive, correlated response in traits in triploids. Estimates of the correlations were also imprecise, usually having large standard errors (> 0.20). More precise estimates for genetic correlations can require a vast amount of data, especially if the traits involved in the correlation are measured on different sets of animals (Bijma and Bastiaansen, 2014). Even with precise estimates (i.e., ≤ 0.05) of genetic correlations between triploids and tetraploids, it may not be clear how best to genetically improve triploids. To better address the question of optimal polyploid breeding strategy, simulations were run to examine breeding strategies based on genetic correlation.

Simulations of breeding programs in AlphaSimR (Gaynor et al., 2020) indicated a breeding program phenotyping all ploidies results in the most genetic improvement of triploids across a wide range of positive genetic correlations. Simulations were designed to compare the original strategy at ABC to a strategy in which related diploid, triploid, and tetraploid families were phenotyped. Strategies were compared with genetic correlation of diploid vs. triploid and tetraploid vs. triploid as a variable. The strategy of phenotyping all ploidies resulted in higher

breeding values for the triploid trait at low genetic correlations (0.33 – 0.66), and the two strategies resulted in similar results at higher genetic correlations (0.75 – 0.90). Thus, in terms of genetic improvement of triploids, the breeding program phenotyping all ploidies was never worse and sometimes better at moderate positive genetic correlations (0.33-0.90), suggesting it may be the better option if there are no precise estimates of between-ploidy genetic correlations.

Multiple reasons exist for adopting a breeding program that phenotypes all ploidies to improve triploids. For one, the progress in improving triploids can be tracked closely because the genetic value of triploids will be measured as a part of annual genetic analyses of the breeding program to select the best diploid and tetraploid families. Second, the effectiveness of a breeding program that phenotypes all ploidies is less dependent on genetic correlations, which may change over time. For instance, genetic correlations may be in part caused by gametic-phase disequilibrium (Falconer and Mackay, 1996), which can decay over time through recombination. Finally, testing triploids may be required to examine commercial traits that are suspected to be unique to triploids. "Triploid mortality" (Guévélou et al., 2019; Matt et al., 2020) has been hypothesized as a trait specific to triploids, and was an impetus to produce the triploid families that were tested in this dissertation.

Triploid mortality, defined as mortality of greater than 20% of near market sized triploid oysters in late spring, likely has a strong genetic basis. This dissertation represents the first analysis of triploid mortality using quantitative genetic methods. A triploid mortality event occurred between May 7 and July 9 at Nandua Creek — three triploid families had survival less than 0.70, while most of the other triploid families had survival greater than 0.90. Survival in late spring was treated as a distinct trait for genetic analysis, and the heritability for this trait for

triploids at Nandua Creek was 1.06 ± 0.32 (standard error), suggesting much of the variation in survival during the mortality event was due to additive genetic effects. The high heritability in late spring survival indicates breeding can substantially reduce the risk of triploid mortality events in commercial oysters at Nandua Creek, and possibly at other sites.

It is unclear if selection at Nandua Creek would result in resistance to triploid mortality in all regions of the Chesapeake Bay. Over 20 farms in Virginia have reported late spring mortalities in triploids (K. Hudson, personal communication), including farms on the western shore of the Chesapeake Bay and the seaside shore of the Eastern Shore of Virginia. Are triploid mortality events across these sites influenced by similar genes? There are currently no data to address this question, such as genetic correlations of late spring survival among sites with triploid mortality events. A challenge in estimating these correlations is that triploid mortality events do not consistently occur at the same sites every year (Guévélou et al., 2019; Matt et al., 2020), and therefore multiple field tests may be required to measure a triploid mortality event at each site.

Selecting for resistance to triploid mortality may not effectively improve survival in triploids at sites that do not manifest "triploid mortality." Genetic correlations between survival in triploids at Nandua Creek, the only site with a triploid mortality event, and the two other sites with no history of triploid mortality events were low (0.46 ± 0.24 and 0.46 ± 0.22), suggesting different genes are responsible for survival of triploids at sites with and without triploid mortality. The low genetic correlations of survival are comparable to the genetic correlation in survival of diploid *C. virginica* at a mesohaline site and oligohaline site (0.57 ± 0.14) (Allen et al., 2021). ABC has chosen to separately select for diploids suited for low salinity zones and high salinity zones based primarily on this low genetic correlation in survival. A similar strategy of

selecting triploids separately based on the presence or absence of late spring mortalities may be warranted for maximum improvement of triploids across the Chesapeake Bay. On the other hand, the genetic parameters at which selecting separate populations is optimal has not been fully evaluated for family-based breeding, and thus may be worth simulating in future studies.

Future evaluations of breeding strategies to improve triploid oysters should incorporate genomic selection. A major benefit of genomic selection is that it enables direct estimation of breeding values for traits that cannot be measured in live selection candidates. For example, with genomic selection, a non-lethal collection of DNA of a selection candidate could inform the breeding value of the individual for destructive traits, like meat weight, or threshold traits, like survival. Triploid traits are analagous to such destructive and threshold traits because they cannot be phenotyped in live selection candidates — diploids and tetraploids cannot be phenotyped for triploid traits. Without genomics, diploid and tetraploid phenotypes are the only criteria to distinguish breeding values for triploid traits at the individual level. With genomic selection, beneficial alleles for triploid traits could be identified in diploid and tetraploid individuals, which would be highly advantageous for polyploid oyster breeding. Genomic selection has been found to increase the accuracy of selection in disease resistance in oysters (Gutierrez et al., 2020), however no polyploid oysters or polyploid animals have yet been subject to genomic selection.

Literature Cited

Allen Jr, S.K., 1983. Flow cytometry: assaying experimental polyploid fish and shellfish. Aquaculture 33, 317–328.

Allen, S.K., 1988. Triploid Oysters Ensure Year-round Supply. Oceanus 31, 58–63.

Allen, S.K., Bushek, D., 1992. Large-scale production of triploid oysters, *Crassostrea virginica* (Gmelin), using "stripped" gametes. Aquaculture 103, 241–251. https://doi.org/10.1016/0044-8486(92)90170-P

Allen, S.K., Downing, S.L., 1990. Performance of Triploid Pacific Oysters, *Crassostrea gigas*: Gametogenesis. Can. J. Fish. Aquat. Sci. 47, 1213–1222.

Allen, S.K., Downing, S.L., 1986. Performance of triploid Pacific oysters, *Crassostrea gigas* (Thunberg). I. Survival, growth, glycogen content, and sexual maturation in yearlings. J. Exp. Mar. Bio. Ecol. 102, 197–208.

Allen, S.K., Kube, P., Small, J., 2021. Genetic parameters for *Crassostrea virginica* and their application to family-based. Aquaculture.

Alleway, H.K., Gillies, C.L., Bishop, M.J., Gentry, R.R., Theuerkauf, S.J., Jones, R., 2019. The ecosystem services of marine aquaculture: valuing benefits to people and nature. Bioscience 69, 59–68.

Andrews, J.D., 1988. Epizootiology of the disease caused by the oyster pathogen *Perkinsus marinus* and its effects on the oyster industry., in: Amer. Fish. Soci. Special Publication. pp. 47–63.

Astaurov, B.L., 1969. Experimental Polyploidy in Animals. Annu. Rev. Genet. 3, 99–126.

https://doi.org/10.1146/annurev.ge.03.120169.000531

Barber, B.J., Mann, R., 1991. Sterile triploid *Crassostrea virginica* (Gmelin, 1791) grow faster than diploids but are equally susceptible to *Perkinsus marinus*. J.Shellfish Res. 10, 445–450.

Beck, M.W., Brumbaugh, R.D., Airoldi, L., Carranza, A., Coen, L.D., Crawford, C., Defeo, O., Edgar, G.J., Hancock, B., Kay, M.C., Lenihan, H.S., Luckenbach, M.W., Toropova, C.L., Zhang, G., Guo, X., 2011. Oyster reefs at risk and recommendations for conservation, restoration, and management. Bioscience 61, 107–116. https://doi.org/10.1525/bio.2011.61.2.5

Beckensteiner, J., Kaplan, D.M., Scheld, A.M., 2020. Barriers to Eastern Oyster Aquaculture Expansion in Virginia. Front. Mar. Sci. 7, 1–19. https://doi.org/10.3389/fmars.2020.00053

Benabdelmouna, A., 2014. Ploidy investigation on mortality related to OsHV-1 in spat Crassostrea gigas. Natl. Shellfish Assoc.

Bijma, P., Bastiaansen, J.W.M., 2014. Standard error of the genetic correlation: How much data do we need to estimate a purebred-crossbred genetic correlation? Genet. Sel. Evol. 46, 1–6. https://doi.org/10.1186/s12711-014-0079-z

Bijma, P., Van Arendonk, J.A.M., 1998. Maximizing genetic gain for the sire line of a crossbreeding scheme utilizing both purebred and crossbred information. Anim. Sci. 66, 529–542. https://doi.org/10.1017/S135772980000970X

Blakeslee, A.F., Avery, A.G., 1937. Methods of Inducing Doubling of Chromosomes in Plants. J. Hered.

Bodenstein, S.R., Stoeckel, J., Carmichael, R.H., 2019. Comparing responses of triploid and diploid Eastern oysters, *Crassostrea virginica*, to common farm stressors.

Bosch, D., Kuminoff, N., Stephenson, K., Miller, A., Pope, J., Harris, A., 2010. Evaluation of policy options for expanding oyster aquaculture in Virginia. Aquac. Econ. Manag. 14, 145–163.

https://doi.org/10.1080/13657301003776698

Bulmer, M.G., 1971. The effect of selection on genetic variability. Am. Nat. 105. https://doi.org/10.4314/sajas.v31i2.3836

Burreson, E.M., Ragone Calvo, L., 1996. Epizootiology of *Perkinsus marinus* disease of oysters in Chesapeake Bay, with emphasis on data since 1985. J. Shellfish Res. 15, 17–34.

Cox, E.S., Smith, M.S.R., Nell, J.A., Maguire, G.B., 1996. Studies on triploid oysters in Australia. VI. Gonad development in diploid and triploid Sydney rock oysters *Saccostrea commercialis* (Iredale and Roughley). J. Exp. Mar. Bio. Ecol. 197, 101–120. https://doi.org/10.1016/0022-0981(95)00147-6

de Bem Oliveira, I., Resende, M.F.R., Ferrão, L.F. V., Amadeu, R.R., Endelman, J.B., Kirst, M., Coelho, A.S.G., Munoz, P.R., 2019. Genomic prediction of autotetraploids; influence of relationship matrices, allele dosage, and continuous genotyping calls in phenotype prediction. G3 Genes, Genomes, Genet. 9, 1189–1198. https://doi.org/10.1534/g3.119.400059

de Melo, C.M.R., Divilov, K., Schoolfield, B., Langdon, C., 2019. Selection of group and individual traits of Pacific oysters (*Crassostrea gigas*) on the West Coast, US. Aquaculture 512, 734389. https://doi.org/10.1016/j.aquaculture.2019.734389

Dégremont, L., Bédier, E., Boudry, P., 2010. Summer mortality of hatchery-produced Pacific oyster spat (*Crassostrea gigas*). II. Response to selection for survival and its influence on growth and yield. Aquaculture 299, 21–29. https://doi.org/10.1016/j.aquaculture.2009.11.017

Dégremont, L., Ernande, B., Bédier, E., Boudry, P., 2007. Summer mortality of hatchery-produced Pacific oyster spat (*Crassostrea gigas*). I. Estimation of genetic parameters for

survival and growth. Aquaculture 262, 41–53. https://doi.org/10.1016/j.aquaculture.2006.10.025

Dégremont, L., Garcia, C., Allen, S.K., 2015. Genetic improvement for disease resistance in oysters: A review. J. Invertebr. Pathol. 131, 226–241. https://doi.org/10.1016/j.jip.2015.05.010

Dégremont, L., Garcia, C., Frank-Lawale, A., Allen, S.K., 2012. Triploid Oysters in the Chesapeake Bay: Comparison of Diploid and Triploid *Crassostrea virginica*. J. Shellfish Res. 31, 21–31. https://doi.org/10.2983/035.031.0103

Downing, S.L., Allen, S.K., 1987. Induced triploidy in the Pacific oyster, *Crassostrea gigas*: Optimal treatments with cytochalasin B depend on temperature. Aquaculture 61, 1–15. https://doi.org/10.1016/0044-8486(87)90332-2

Falconer, D.S., Mackay, T.F.C., 1996. Introduction to Quantitative Genetics, UK: Longman Group.

Frank-Lawale, A., Allen, S.K., Dégremont, L., 2014. Breeding and Domestication of Eastern Oyster (*Crassostrea virginica*) Lines for Culture in the Mid-Atlantic, Usa: Line Development and Mass Selection for Disease Resistance. J. Shellfish Res. 33, 153–165. https://doi.org/10.2983/035.033.0115

Gallais, A., 2003. Quantitative Genetics and Breeding Methods in Autopolyploid Plants. Paris.

Galstoff, P.S., 1964. The American Oyster *Crassostrea virginica* Gmelin. Fish. Bull. Fish Wildl. Serv. 64, 1–480.

Galtsoff, P.S., 1964. The American oyster, *Crassostrea virginica* gmelin. US Government Printing Office.

Gaynor, C., 2019. Traits in AlphaSimR 1–7.

Gaynor, R.C., Gorjanc, G., Hickey, J.M., 2020. AlphaSimR: An R-package for breeding program

simulations. bioRxiv. https://doi.org/10.1101/2020.08.10.245167

Gjedrem, T., 2012. Genetic improvement for the development of efficient global aquaculture: A personal opinion review. Aquaculture 344–349, 12–22. https://doi.org/10.1016/J.AQUACULTURE.2012.03.003

Gjedrem, T., 2005. Selection and Breeding Programs in Aquaculture. Springer, New York, NY.

Glude, J.B., 1975. A summary report of Pacific coast oyster mortality investigations 1965–1972, in: Proceedings of the Third US–Japan Meeting on Aquaculture at Tokyo, Japan. pp. 1–28.

Gorjanc, G., Hickey, J.M., 2018. AlphaMate: a program for optimizing selection, maintenance of diversity and mate allocation in breeding programs. Bioinformatics 34, 3408–3411. https://doi.org/10.1093/bioinformatics/bty375

Grandke, F., Singh, P., Heuven, H.C.M., de Haan, J.R., Metzler, D., 2016. Advantages of continuous genotype values over genotype classes for GWAS in higher polyploids: A comparative study in hexaploid chrysanthemum. BMC Genomics 17, 1–9. https://doi.org/10.1186/s12864-016-2926-5

Guévélou, E., Carnegie, R.B., Small, J.M., Hudson, K., Reece, K.S., Rybovich, M.M., 2019. Tracking Triploid Mortalities of Eastern Oysters *Crassostrea virginica* in the Virginia Portion of the Chesapeake Bay. J. Shellfish Res. 38, 101–113. https://doi.org/10.2983/035.038.0110

Guo, X., Allen, S.K., 1994a. Viable tetraploids in the Pacific oyster (*Crassostrea gigas* Thunberg) produced by inhibiting polar body 1 in eggs. Mol. Mar. Biol. Biotechnol. 3, 42–50.

Guo, X., Allen, S.K., 1994b. Reproductive potential and genetics of triploid Pacific oysters, *Crassostrea gigas* (Thunberg). Biol. Bull. 187, 309–318.

Guo, X., DeBrosse, G.A., Allen Jr, S.K., 1996. All-triploid Pacific oysters (*Crassostrea gigas* Thunberg) produced by mating tetraploids and diploids. Aquaculture 142, 149–161.

Guo, X., Ford, S.E., DeBrosse, G., Smolowitz, R., Sunila, I., 2003. Breeding and evaluation of eastern oyster strains selected for MSX, Dermo and JOD resistance. J Shellfish Res 22, 333–334.

Gutierrez, A.P., Matika, O., Bean, T.P., Houston, R.D., 2018. Genomic Selection for Growth Traits in Pacific Oyster (*Crassostrea gigas*): Potential of Low-Density Marker Panels for Breeding Value Prediction. Front. Genet. 9, 391. https://doi.org/10.3389/fgene.2018.00391

Gutierrez, A.P., Symonds, J., King, N., Steiner, K., Bean, T.P., Houston, R.D., 2020. Potential of genomic selection for improvement of resistance to ostreid herpesvirus in Pacific oyster (*Crassostrea gigas*). Anim. Genet. 51, 249–257. https://doi.org/10.1111/age.12909

Hamilton, M.G., Kerr, R.J., 2018. Computation of the inverse additive relationship matrix for autopolyploid and multiple-ploidy populations. Theor. Appl. Genet. 131, 851–860. https://doi.org/10.1007/s00122-017-3041-y

Hamilton, M.G., Kube, P.D., Elliott, N.G., McPherson, L.J., Krsinich, A., 2009. Development of a Breeding Strategy for Hybrid Abalone, in: Proceedings of the Association of the Advancement of Animal Breeding And Genetics. pp. 350–353.

Harding, J.M., 2007. Comparison of Growth Rates Bbetween Diploid DEBY Eastern Oysters (*Crassostrea virginica*, Gmelin 1791), Triploid Eastern Oysters, and Triploid Suminoe Oysters (C. ariakensis, Fugita 1913). J. Shellfish Res. 26, 961–972. https://doi.org/10.2983/0730-8000(2007)26[961:COGRBD]2.0.CO;2

Henderson, C.R., 1976. A Simple Method for Computing the Inverse of a Numerator Relationship Matrix Used in Prediction of Breeding Values. Biometrics 32, 69. https://doi.org/10.2307/2529339

Hickey, J.M., Gorjanc, G., 2012. Simulated data for genomic selection and genome-wide

association studies using a combination of coalescent and gene drop methods. G3 Genes, Genomes, Genet. 2, 425–427. https://doi.org/10.1534/g3.111.001297

Hollenbeck, C.M., Johnston, I.A., 2018. Genomic tools and selective breeding in molluscs. Front. Genet. https://doi.org/10.3389/fgene.2018.00253

Houston, R.D., Bean, T.P., Macqueen, D.J., Gundappa, M.K., Jin, Y.H., Jenkins, T.L., Selly, S.L.C., Martin, S.A.M., Stevens, J.R., Santos, E.M., Davie, A., Robledo, D., 2020. Harnessing genomics to fast-track genetic improvement in aquaculture. Nat. Rev. Genet. 21, 389–409. https://doi.org/10.1038/s41576-020-0227-y

Hudson, K., 2019. Results of the 2018 Virginia Shellfish Aquaculture Crop Reporting Survey. VIMS Mar. Resour. Rep. 2019–8.

Hudson, K., 2018. Virginia Shellfish Aquaculture Situation and Outlook Report. Results of the 2017 Virginia Shellfish Aquaculture Crop Reporting Survey, VIMS Marine Resource Report No. 2018-9. Virginia Sea Grant VSG-18-3.

Huvet, A., Normand, J., Fleury, E., Quillien, V., Fabioux, C., Boudry, P., 2010. Reproductive effort of Pacific oysters: A trait associated with susceptibility to summer mortality. Aquaculture 304, 95–99. https://doi.org/10.1016/j.aquaculture.2010.03.022

Huvet, A., Royer, J., Moal, J., Burgeot, T., Lapègue, S., Buolo, V., Nicolas, J.L., Lambert, C., Van Wormhoudt, A., 2008. Phenotypic characteristics of “R” and “S” oysters lines, selected for resistance or susceptibility to summer mortality, in: Summer Mortality of Pacific Oyster *Crassostrea gigas*. The Morest Project. pp. 197–241.

Isik, F., Holland, J., 2017. Genetic Data Analysis for Plant and Animal Breeding.

James, J.W., 1961. Selection in two environments. Heredity (Edinb). 16, 145–152. https://doi.org/10.1038/hdy.1961.17

Johnson, R.M., Shrimpton, J.M., Cho, G.K., Heath, D.D., 2007. Dosage effects on heritability and maternal effects in diploid and triploid Chinook salmon (*Oncorhynchus tshawytscha*). Heredity (Edinb). 98, 303–310. https://doi.org/10.1038/sj.hdy.6800941

Johnstone, F.E., 1939. Chromosome doubling in potatoes induced by colchicine treatment. Am. Potato J. 16, 288–304. https://doi.org/10.1007/BF02861918

Jouaux, A., Heude-Berthelin, C., Sourdaine, P., Mathieu, M., Kellner, K., 2010. Gametogenic stages in triploid oysters *Crassostrea gigas*: Irregular locking of gonial proliferation and subsequent reproductive effort. J. Exp. Mar. Bio. Ecol. 395, 162–170. https://doi.org/10.1016/j.jembe.2010.08.030

Keeling, S.E., Brosnahan, C.L., Williams, R., Gias, E., Hannah, M., Bueno, R., McDonald, W.L., Johnston, C., 2014. New Zealand juvenile oyster mortality associated with ostreid herpesvirus 1-an opportunistic longitudinal study. Dis. Aquat. Organ. 109, 231–239. https://doi.org/10.3354/dao02735

Kerr, R.J., Dieters, M.J., Tier, B., 2004. Simulation of the comparative gains from four different hybrid tree breeding strategies. Can. J. For. Res. 34, 209–220. https://doi.org/10.1139/x03-180

Kerr, R.J., Li, L., Tier, B., Dutkowski, G.W., McRae, T.A., 2012. Use of the numerator relationship matrix in genetic analysis of autopolyploid species. Theor. Appl. Genet. https://doi.org/10.1007/s00122-012-1785-y

Kihara, H., 1951. Triploid watermelons, in: Proc. Amer. Soc. Hort. Sci. pp. 217–230.

Koganezawa, A., 1975. Present status of studies on the mass mortality of cultured oyster in Japan and its prevention, in: Proceedings of the Third US-Japan Meeting on Aquaculture. pp. 29–34.

Kriaridou, C., Tsairidou, S., Houston, R.D., Robledo, D., 2020. Genomic Prediction Using Low Density Marker Panels in Aquaculture: Performance Across Species, Traits, and Genotyping Platforms. Front. Genet. 11, 1–8. https://doi.org/10.3389/fgene.2020.00124

Kube, P., Cunningham, M., Dominik, S., Parkinson, S., Henshall, J., Finn, B., Henshall, J., Bennett, R., Hamilton, M., 2011. Enhancement of the Pacific oyster selective breeding program.

Kube, P., Dove, M., Cunningham, M., Kirkland, P., Gu, X., Hick, P., O'Connor, W., Elliot, N., 2018. Genetic Selection for Resistance to Pacific Oyster Mortality Syndrome.

Leonhardt, J.M., Casas, S., Supan, J.E., La Peyre, J.F., 2017. Stock assessment for eastern oyster seed production and field grow-out in Louisiana. Aquaculture 466, 9–19. https://doi.org/10.1016/j.aquaculture.2016.09.034

Lillehammer, M., Meuwissen, T.H.E., Sonesson, A.K., 2013. A low-marker density implementation of genomic selection in aquaculture using within-family genomic breeding values. Genet. Sel. Evol. 45, 39. https://doi.org/10.1186/1297-9686-45-39

Love, D.C., Lane, R.M., Kuehl, L.M., Hudson, B., Harding, J., Clancy, K., Fry, J.P., 2020. Performance and conduct of supply chains for United States farmed oysters. Aquaculture 515, 734569. https://doi.org/10.1016/j.aquaculture.2019.734569

Lutaaya, E., Misztal, I., Mabry, J.W., Short, T., Timm, H.H., Holzbauer, R., 2001. Genetic parameter estimates from joint evaluation of purebreds and crossbreds in swine using the crossbred model. J. Anim. Sci. 79, 3002–3007. https://doi.org/10.2527/2001.79123002x

Lynch, M., Walsh, B., 1998. Genetics and analysis of quantitative traits. Sinauer Sunderland, MA.

Matt, J.L., Allen, S.K., 2021. A classification system for gonad development in triploid *Crassostrea virginica*. Aquaculture 532, 735994. https://doi.org/10.1016/j.aquaculture.2020.735994

Matt, J.L., Allen, S.K., 2014. Heteroploid mosaic tetraploids of *Crassostrea virginica* produce

normal triploid larvae and juveniles as revealed by flow cytometry. Aquaculture 432, 336–345. https://doi.org/10.1016/j.aquaculture.2014.05.015

Matt, J.L., Guévélou, E., Small, J.M., Allen, S.K., 2020. A field test investigating the influence of brood stock origin and ploidy on the susceptibility of *Crassostrea virginica* to "triploid mortality" in the Chesapeake Bay. Aquaculture 526, 735375. https://doi.org/10.1016/j.aquaculture.2020.735375

Meuwissen, T., Hayes, B., Goddard, M., 2016. Genomic selection: A paradigm shift in animal breeding. Anim. Front. 6, 6–14. https://doi.org/10.2527/af.2016-0002

Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819–1829.

Muff, S., Niskanen, A.K., Saatoglu, D., Keller, L.F., Jensen, H., 2019. Animal models with group-specific additive genetic variances: Extending genetic group models. Genet. Sel. Evol. 51, 1–16. https://doi.org/10.1186/s12711-019-0449-7

Mulder, H.A., Bijma, P., 2005. Effects of genotype x environment interaction on genetic gain in breeding programs. J. Anim. Sci. 83, 49–61. https://doi.org/10.2527/2005.83149x

Mulder, H.A., Veerkamp, R.F., Ducro, B.J., Van Arendonk, J.A.M., Bijma, P., 2006. Optimization of dairy cattle breeding programs for different environments with genotype by environment interaction. J. Dairy Sci. 89, 1740–1752. https://doi.org/10.3168/jds.S0022-0302(06)72242-1

Murray, T., Oesterling, M., 2010. Virginia Shellfish Aquaculture Situation and Outlook Report : Results of the 2009 Virginia Shellfish Aquaculture Crop Reporting Survey. Reports. https://doi.org/https://doi.org/10.21220/V5JX45

Nell, J.A., Cox, E., Smith, I.R., Maguire, G.B., 1994. Studies on triploid oysters in Australia. I. The

farming potential of triploid Sydney rock oysters *Saccostrea commercialis* (Iredale and Roughley). Aquaculture 243–255. https://doi.org/10.1046/j.1365-2109.1999.00296.x

Nyine, M., Uwimana, B., Blavet, N., Hřibová, E., Vanrespaille, H., Batte, M., Akech, V., Brown, A., Lorenzen, J., Swennen, R., Doležel, J., 2018. Genomic Prediction in a Multiploid Crop: Genotype by Environment Interaction and Allele Dosage Effects on Predictive Ability in Banana. Plant Genome 11, 170090. https://doi.org/10.3835/plantgenome2017.10.0090

Otto, S.P., Whitton, J., 2000. Polyploid Incidence and Evolution.

Paul-Pont, I., Evans, O., Dhand, N.K., Rubio, A., Coad, P., Whittington, R.J., 2014. Descriptive epidemiology of mass mortality due to Ostreid herpesvirus-1 (OsHV-1) in commercially farmed Pacific oysters (*Crassostrea gigas*) in the Hawkesbury River estuary, Australia. Aquaculture 422–423, 146–159. https://doi.org/10.1016/j.aquaculture.2013.12.009

Peeler, E.J., Allan Reese, R., Cheslett, D.L., Geoghegan, F., Power, A., Thrush, M.A., 2012. Investigation of mortality in Pacific oysters associated with Ostreid herpesvirus-1µVar in the Republic of Ireland in 2009. Prev. Vet. Med. 105, 136–143. https://doi.org/10.1016/j.prevetmed.2012.02.001

Piferrer, F., Beaumont, A., Falguière, J.C., Flajšhans, M., Haffray, P., Colombo, L., 2009. Polyploid fish and shellfish: Production, biology and applications to aquaculture for performance improvement and genetic containment. Aquaculture 293, 125–156. https://doi.org/10.1016/j.aquaculture.2009.04.036

R Core Team, 2019. R: A Language and Environment for Statistical Computing.

Ragone Calvo, L.M., Calvo, G.W., Burreson, E.M., 2003. Dual disease resistance in a selectively bred eastern oyster, *Crassostrea virginica*, strain tested in Chesapeake Bay. Aquaculture 220, 69–87. https://doi.org/10.1016/S0044-8486(02)00399-X

Ray, N.E., Maguire, T.J., Al-Haj, A.N., Henning, M.C., Fulweiler, R.W., 2019. Low greenhouse gas emissions from oyster aquaculture. Environ. Sci. Technol. 53, 9118–9127. https://doi.org/10.1021/acs.est.9b02965

Ritter, K., 2019. Fecundity of Triploid Eastern Oyster (*Crassostrea virginica*) as a Function of Tetraploid Lineage. Dissertations, Theses, and Masters Projects. Paper 1582642221.

Rutten, M.J.M., Bijma, P., Woolliams, J.A., Van Arendonk, J.A.M., 2002. SelAction: Software to predict selection response and rate of inbreeding in livestock breeding programs. J. Hered. 93, 456–458. https://doi.org/10.1093/jhered/93.6.456

Sae-Lim, P., Gjerde, B., Nielsen, H.M., Mulder, H., Kause, A., 2016. A review of genotype-by-environment interaction and micro-environmental sensitivity in aquaculture species. Rev. Aquac. 8, 369–393. https://doi.org/10.1111/raq.12098

Samain, J.-F., 2011. Review and perspectives of physiological mechanisms underlying genetically-based resistance of the Pacific oyster *Crassostrea gigas* to summer mortality. Aquat. Living Resour. 24, 227–236. https://doi.org/10.1051/alr/2011144

Samain, J.-F., Degremont, L., Soletchnik, P., Haure, J., Bédier, E., Ropert, M., Moal, J., Huvet, A., Bacca, H., Van Wormhoudt, A., others, 2007. Genetically based resistance to summer mortality in the Pacific oyster (*Crassostrea gigas*) and its relationship with physiological, immunological characteristics and infection processes. Aquaculture 268, 227–243.

Samain, J.F., McCombie, H., 2008. Summer mortality of Pacific oyster *Crassostrea gigas*: the Morest project. Eds Quae, Versailles, France.

Schulte, D.M., 2017. History of the Virginia Oyster Fishery, Chesapeake Bay, USA. Front. Mar. Sci. 4, 127. https://doi.org/10.3389/fmars.2017.00127

Segarra, A., Pépin, J.F., Arzul, I., Morga, B., Faury, N., Renault, T., 2010. Detection and

description of a particular Ostreid herpesvirus 1 genotype associated with massive mortality outbreaks of Pacific oysters, *Crassostrea gigas*, in France in 2008. Virus Res. 153, 92–99. https://doi.org/10.1016/j.virusres.2010.07.011

Shumway, S., 2011. Shellfish aquaculture and the environment. Wiley-Blackwell.

Speed, D., Balding, D.J., 2015. Relatedness in the post-genomic era: Is it still useful? Nat. Rev. Genet. 16, 33–44. https://doi.org/10.1038/nrg3821

Stanley, J.G., Allen Jr, S.K., Hidu, H., 1981. Polyploidy induced in the American oyster, *Crassostrea virginica*, with cytochalasin B. Aquaculture 23, 1–10.

Vu, S. V., Knibb, W., Nguyen, N.T.H., Vu, I. V., O'Connor, W., Dove, M., Nguyen, N.H., 2020. First breeding program of the Portuguese oyster *Crassostrea angulata* demonstrated significant selection response in traits of economic importance. Aquaculture 518. https://doi.org/10.1016/j.aquaculture.2019.734664

Wadsworth, P., Casas, S., La Peyre, J., Walton, W., 2019. Elevated mortalities of triploid eastern oysters cultured off-bottom in northern Gulf of Mexico. Aquaculture 505, 363–373. https://doi.org/10.1016/j.aquaculture.2019.02.068

Wang, J., Van Ginkel, M., Podlich, D., Ye, G., Trethowan, R., Pfeiffer, W., DeLacy, I.H., Cooper, M., Rajaram, S., 2003. Comparison of two breeding strategies by computer simulation. Crop Sci. 43, 1764–1773. https://doi.org/10.2135/cropsci2003.1764

Ward, R.D., Thompson, P.A., Appleyard, S.A., Swan, A.A., Kube, P.D., 2005. Sustainable Genetic Improvement of Pacific Oysters in Tasmania and South Australia. Fish. Res. Dev. Corp. Final report. Canberra, Aust. https://doi.org/10.1103/PhysRevLett.101.211102

Wei, M., Van Der Werf, J.H.J., 1994. Maximizing genetic response in crossbreds using both purebred and crossbred information. Anim. Prod. 59, 401–413.

https://doi.org/10.1017/S0003356100007923

Wolak, M.E., Reid, J.M., 2017. Accounting for genetic differences among unknown parents in microevolutionary studies: how to include genetic groups in quantitative genetic animal models. J. Anim. Ecol. 86, 7–20. https://doi.org/10.1111/1365-2656.12597

Woolliams, J.A., Berg, P., Dagnachew, B.S., Meuwissen, T.H.E., 2015. Genetic contributions and their optimization. J. Anim. Breed. Genet. 132, 89–99. https://doi.org/10.1111/jbg.12148

Zenger, K.R., Khatkar, M.S., Jones, D.B., Khalilisamani, N., Jerry, D.R., Raadsma, H.W., 2019. Genomic selection in aquaculture: Application, limitations and opportunities with special reference to marine shrimp and pearl oysters. Front. Genet. https://doi.org/10.3389/fgene.2018.00693