Summer 8-19-2016

Population Genetic Structure of Softshell Clams (Mya arenaria) with Regard to a Saxitoxin-resistant Mutation and Neutral Genetic Markers in the Gulf of Maine

Jennifer Phillips
University of Maine, couturejm4@gmail.com

POPULATION GENETIC STRUCTURE OF SOFTSHELL CLAMS (*MYA ARENARIA*) WITH REGARD TO A SAXITOXIN-RESISTANT MUTATION AND NEUTRAL GENETIC MARKERS IN THE GULF OF MAINE

By

Jennifer Michelle Phillips

B.S. University of Maine, 2008

A DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

(in Marine Biology)

The Graduate School

The University of Maine

August 2016

Advisory Committee:

Laurie Connell, Research Professor of Marine Science, Advisor

Brian Beal, Professor of Marine Ecology

Neal Pettigrew, Professor of Oceanography

Paul Rawson, Associate Professor of Marine Science

Rebecca Van Beneden, Professor of Biochemistry & Marine Science

DISSERTATION ACCEPTANCE STATEMENT

On behalf of the Graduate Committee for Jennifer Phillips I affirm that this manuscript is the final and accepted dissertation. Signatures of all committee members are on file with the Graduate School at the University of Maine, 42 Stodder Hall, Orono, Maine.

Laurie Connell, Research Professor Date

LIBRARY RIGHTS STATEMENT

In presenting this dissertation in partial fulfillment of the requirements for an advanced degree at the University of Maine, I agree that the Library shall make it freely available for inspection. I further agree that permission for "fair use" copying of this dissertation for scholarly purposes may be granted by the Librarian. It is understood that any copying or publication of this dissertation for financial gain shall not be allowed without my written permission.

Signature:

Date:

POPULATION GENETIC STRUCTURE OF SOFTSHELL CLAMS (*MYA ARENARIA*) WITH REGARD TO A SAXITOXIN-RESISTANT MUTATION AND NEUTRAL GENETIC MARKERS IN THE GULF OF MAINE

By Jennifer Michelle Phillips

Dissertation Advisor: Dr. Laurie Connell

An Abstract of the Dissertation Presented
In Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
(in Marine Biology)
August 2016

The softshell clam, *Mya arenaria*, is a commercially important bivalve species that is found in soft-bottom intertidal habitats throughout the Gulf of Maine, USA. This species is subjected to seasonal blooms of the toxic algae *Alexandrium* spp., and acts as a vector for paralytic shellfish poisoning (PSP) during harmful algal bloom (HAB) events. Some clams possess a naturally occurring genetic mutation of their voltage-gated sodium channels that grants them a resistance to the paralytic effects of saxitoxin (STX) produced by *Alexandrium* spp. The mutation allows these individuals to continue feeding during HABs, and greatly increases their tissue toxicity through bioaccumulation. This work describes the distribution of the resistant mutation in wild clam populations in the Gulf of Maine, and explores the population structure of *M. arenaria* with regard to the mutation, as well as neutral genetic markers. Analysis of neutral markers revealed no significant population structure within the Gulf of Maine, however *M. arenaria* does exhibit strong localized structure at the STX-resistant mutation locus. This structure is sustained by differential selective pressure exerted by *Alexandrium* spp. blooms, despite freely

occurring gene flow among clam populations. In Penobscot Bay, one area where the prevalence of the resistant mutation did not match the strength of selective pressure, it is likely that the resistant allele is maintained by gene flow through larval transport from other regions, rather than by seeding of hatchery stock carrying the mutation. This work can aid PSP monitoring efforts by identifying areas where risk is greatest to humans due to high numbers of resistant clams. In addition, distinguishing areas where one genotype is clearly favored over the others may be of interest to seeding programs trying to ensure that their stock is well suited for the location to which they will be transplanted.

DEDICATION

This work is dedicated to all of my teachers and professors, without whom none of this would have come to fruition.

ACKNOWLEDGEMENTS

I wish to thank the people who have collaborated with me on portions of this work, including Monica Bricelj (Rutgers University), and Robert Cerrato and Maren Mitch (Stony Brook University).

I would like to thank all of the Connell Laboratory technicians (Willis Beazley, Caleb Slemmons, Megan Altenritter, Katie Earle, Sarah Turner, Leslie Astbury, and Corey Hirn) for their assistance in the lab. I would also like to thank the Downeast Institute staff (George Protopopescu, Kyle Pepperman, and Cody Jourdet) for teaching me the fine art of spawning and raising clams.

Lastly I wish to thank all of my friends and family who have cheered me along the way, especially my husband, who has been a steadfast source of support and encouragement since the very beginning of this journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BIRSH.....Beals Island Regional Shellfish Hatchery

DEI..... Downeast Institute for Applied Marine Research & Education

EGoM..... eastern Gulf of Maine

EMCC .....eastern Maine coastal current

GTX .....gonyautoxin

MDI.....Mount Desert Island

NaV ..... sodium channel mutation locus in *Mya arenaria*

NeoSTX ..... neosaxitoxin

PB.....Penobscot Bay

PSP..... paralytic shellfish poisoning

PST.....paralytic shellfish toxin

RR .....resistant genotype

RS.....heterozygous genotype

SNE-NY ..... southern New England – New York

SS .....sensitive genotype

STX.....saxitoxin

TTX..... tetrodotoxin

VGSC.....voltage-gated sodium channel

WGoM ..... western Gulf of Maine

WMCC..... western Maine coastal current

CHAPTER 1
GENERAL BACKGROUND

1.1. *Alexandrium* spp. blooms and paralytic shellfish poisoning

The marine dinoflagellate genus *Alexandrium* comprises a cosmopolitan group of phytoplankton that are found throughout the world's oceans. Some species of *Alexandrium,* most notably the *Alexandrium tamarense* species complex (*A. tamarense*, *A. fundyense,* and *A. catanella*), produce a suite of neurotoxins that are collectively known as paralytic shellfish toxins (PSTs), including saxitoxin (STX), neosaxitoxin (NeoSTX), and gonyautoxin (GTX). When these harmful algae form large blooms (HABs) colloquially known as "red tides" in the coastal nearshore environment they become a concern to humans due to the risk of paralytic shellfish poisoning (PSP). Paralytic shellfish poisoning is an illness that occurs when humans consume shellfish that have been feeding on toxic *Alexandrium* spp. and bioaccumulating PSTs in their body tissues, and it constitutes a significant public health concern during blooms, as severe cases of PSP can be fatal. The shellfish themselves are also subject to the deleterious effects of PSTs. Consumption of *Alexandrium* spp. can affect their gill function, oxygen usage, feeding, metabolism, and burrowing abilities (Bricelj *et al.* 2005; MacQuarrie and Bricelj 2008).

1.2. Structure and function of voltage-gated sodium channels

Paralytic shellfish toxins induce toxicity by acting on the voltage-gated sodium channels (VGSCs) in excitable nerve and muscle cells. Voltage-gated sodium channels are

large membrane-spanning proteins that play a critical role in the generation of nerve action potentials. Under normal conditions, these channels respond to alterations in the cell membrane potential by undergoing a conformational change from a closed to an open configuration, allowing the transport of sodium ions through the channel and across the cell membrane. This influx of sodium ions is driven by an electrochemical gradient that is established by the active export of sodium from the cell via the sodium/potassium-ATPase pump (Catterall 1992). Sodium channels are essential for neuromuscular function in many organisms spanning a variety of phyla, and thus they are often the target of toxins such as PSTs. The outside surface of the sodium channel pore contains a binding site that binds both STX and tetrodotoxin (TTX) with high affinity. When either of those molecules binds to the sodium channel pore, it causes a reversible blockage of the channel, effectively acting like a cork in a bottle. Without the influx of sodium ions and the resulting depolarization of the cell membrane, it becomes impossible to generate nerve action potentials, and paralysis will shortly develop.

Voltage-gated sodium channels exist in three possible states: closed, open, and resting. During the depolarization phase, a change in cell membrane potential causes the channels to open, and sodium ions flood into the cell. This influx of sodium ions further depolarizes the membrane, triggering more channels to open (Anderson *et al.* 2005). The channels are then rapidly inactivated and enter the resting stage, during which time they cannot be activated again until the cell membrane is repolarized (Denac *et al.* 2000).

Voltage-gated sodium channels are comprised of an α subunit, and typically one or two β subunits. The β subunits have a large extracellular domain and a small intracellular domain, connected by a transmembrane segment (Catterall 2000). The α subunit contains

four domains (DI-DIV) connected by intracellular loops. Each domain contains six transmembrane segments (S1-S6) connected by extracellular and intracellular loops (Catterall 2000; Denac *et al.* 2000) (Figure 1.1).

The outer sodium channel pore contains both the selectivity filter and the binding site for TTX and STX. β-hairpins in the extracellular loops between the S5 and S6 transmembrane segments in each domain form a highly selective path, as well as a binding pocket for TTX and STX (Denac *et al.* 2000). The selectivity filter itself is formed by a group of four amino acid residues (aspartic acid, glutamic acid, lysine and alanine) known as the DEKA locus (Denac *et al.* 2000). The DEKA locus and a second ring of amino acid residues, the EEMD locus, form Toxin Site 1 where TTX and STX are bound (Anderson *et al.* 2005).

Tetrodotoxin and STX have proven to be useful in determining the shape and structure of VGSCs. These two toxins competitively bind to the previously mentioned binding site with high affinity, and studying the amino acid residues involved in binding of these toxin molecules has provided information about the amino acids that comprise the outer pore of the sodium channel and the toxin binding site (Catterall 1992; Denac *et al.* 2000). Models of TTX and STX binding have implicated the negatively charged glutamic acid residue (E) in Toxin Site 1 (equivalent to E945 in the rat sodium channel) (Auld *et al.* 1990) as being essential for binding of these toxins (Anderson *et al.* 2005), and it has been shown that mutations of this negatively-charged amino acid residue in the S5/S6 extracellular loop of DII can result in reduced binding affinity of STX (Noda *et al.* 1989; Pusch *et al.* 1991; Terlau *et al.* 1991).

The S4 segment in each domain functions as the voltage sensor. These segments contain positively charged amino acid residues at every third position (Anderson *et al.* 2005), and their transmembrane positioning makes them well suited for detecting changes in the cell membrane potential (Catterall 2000; Denac *et al.* 2000). Inducing mutations into the S4 segments by replacing the positively charged amino acid residues with neutrally charged residues reduces the activation response of the sodium channel, showing that these charged amino acid residues in the S4 segments play a crucial role in gating (Catterall 1992).

After the opening of the sodium channel and influx of sodium ions, the channel is then rapidly inactivated through a process that involves the intracellular loop between DIII and DIV. This loop contains the hydrophobic amino acid residues isoleucine, phenylalanine, and methionine (IFM) that bind to the connecting loops between the S4 and S5 segments in DIII and DIV to block the sodium channel pore from the internal side (Catterall 2000). Severing the connecting loop between DIII and DIV greatly slows the inactivation process, and inducing mutations into the IFM cluster completely halts inactivation, demonstrating the functional importance of these structures (Catterall 1992).

Figure 1.1 Sodium channel α subunit

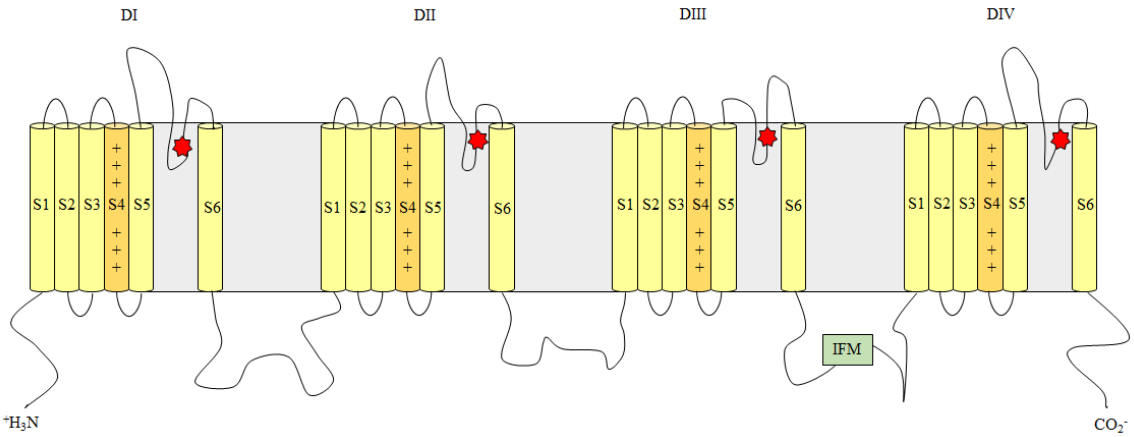


Diagram of the α subunit of a sodium channel showing the six transmembrane segments (S1-S6) in each of the four domains (DI-DIV), the voltage-sensing S4 transmembrane segment in each domain, and the IFM inactivation cluster between DIII and DIV. Red stars mark the location of Toxin Site 1 in the sodium channel pore where STX and TTX are bound.

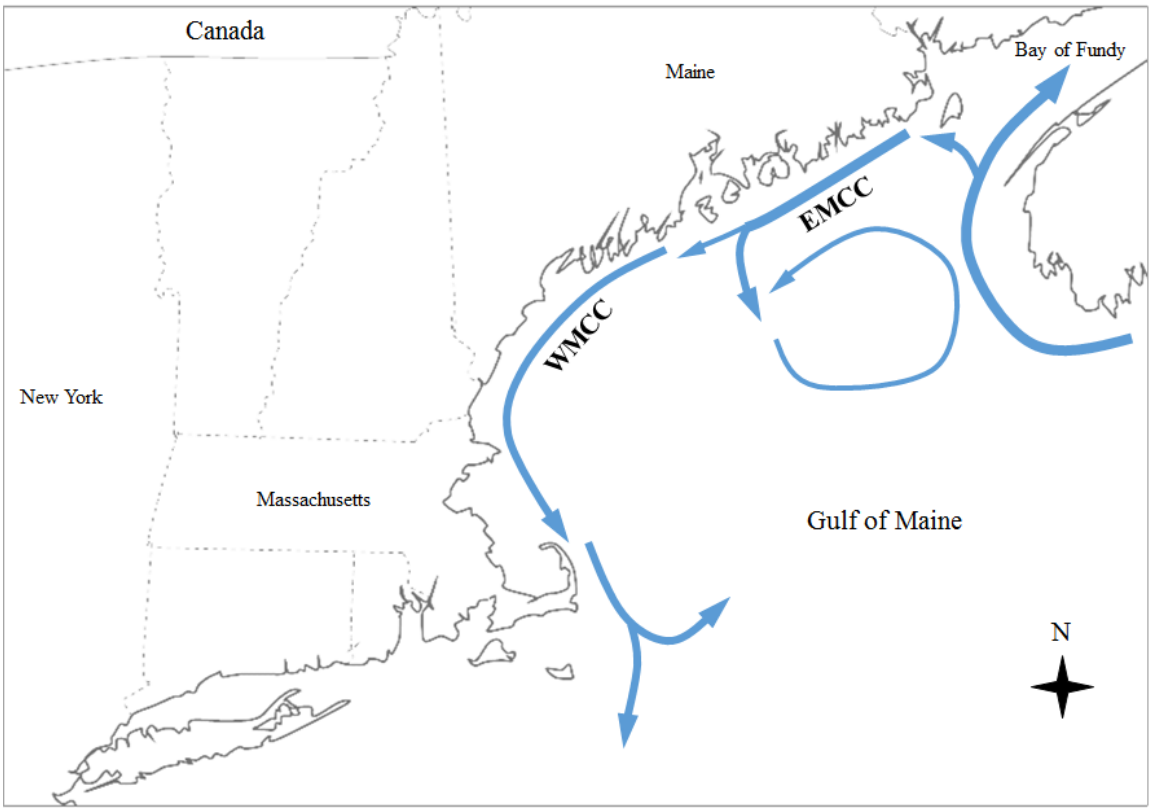
Since sodium channels are vital for neuromuscular function in many organisms, it is unsurprising that they are targeted by a number of toxins. In turn, many of the organisms that are routinely exposed to such toxins have developed sodium channel alterations that render them resistant or immune to the harmful effects of these compounds (Anderson *et al.* 2005).

1.3. *Alexandrium* spp. in the Gulf Of Maine

Algae from the genus *Alexandrium* are part of the native phytoplankton assemblage in the Gulf of Maine, USA, where they typically blooms in early to mid-summer every

year, following the spring diatom bloom. The distribution of *Alexandrium* spp. in the Gulf of Maine varies interannually, and depends on a combination of environmental factors such as the location and number of *Alexandrium* spp. cyst beds deposited by the previous year's bloom, and the oceanographic conditions in the Gulf of Maine in the early summer when blooms begin. At the initiation of a bloom, the annual circulation patterns of the Gulf of Maine are already established. These patterns are characterized by two major currents: the Eastern Maine Coastal Current (EMCC), which flows from the Bay of Fundy southwestward down the coast of Maine to Penobscot Bay where much of it turns offshore, and the Western Maine Coastal Current (WMCC), which flows from Penobscot Bay westward down the coast (Figure 1.2). The EMCC is typified by colder water temperatures and faster current speed, in contrast to the WMCC that tends to be slightly warmer and not as fast moving (Pettigrew *et al.* 2005). It is this disparity in speed between the two currents that causes the offshore veering of the EMCC, as it runs up against the WMCC source waters coming from Penobscot Bay (Pettigrew *et al.* 2005). Eastern Maine coastal current water recirculates in gyres in the Jordan and Georges Basins, and this circulation pattern can increase the residence time of these waters in the eastern Gulf of Maine (Pettigrew *et al.* 2005).

Figure 1.2 Major currents of the Gulf of Maine



Major current patterns of the Gulf of Maine, including the eastern Maine coastal current (EMCC) and the western Maine coastal current (WMCC). (Adapted from Pettigrew *et al.* 2005)

The degree of connectivity between the EMCC and WMCC varies both seasonally and from year to year. On a yearly timescale, this connectivity exhibits one of three possible scenarios: "door open," "door closed," or "door ajar." In a "door open" scenario, a weaker flow regime in the EMCC minimizes the effect of the convergent boundary near Penobscot Bay, and allows most of the EMCC to continue to flow down the coast and mix with the WMCC. A "door closed" scenario results from a very strong flow in the EMCC and a strong convergent boundary where it meets the WMCC, forcing the current to turn

offshore. Since the EMCC waters have a very distinct signature they can be tracked quite accurately, and hydrographic surveys during a "door closed" scenario showed almost no connectivity that year between the waters of the eastern Gulf of Maine and western Gulf of Maine, with the exception of some intermittent leakage in the nearshore environment (Pettigrew *et al.* 2005). A "door ajar" scenario falls somewhere between the previous two scenarios, and is the most common state for this system.

When growth conditions for *Alexandrium* spp. become optimal, resting cysts germinate from large cyst beds in Grand Manan basin at the mouth of the Bay of Fundy, and offshore in the western Gulf of Maine (White and Lewis 1982). The EMCC provides an ideal environment for *Alexandrium* spp. growth due to the ready availability of essential nutrients as a result of tidal mixing and a general lack of stratification (Townsend *et al.* 2001). The Bay of Fundy also provides highly favorable growth conditions due to large nutrient fluxes to the surface waters from strong tidal mixing in that region (Townsend *et al.* 2001). The WMCC presents a less favorable growing environment for *Alexandrium* spp. due to stratification caused by large plumes of freshwater input from several major river systems, including the Androscoggin, Kennebec, and Penobscot rivers. This density stratification keeps the surface waters nutrient-limited following the spring diatom bloom. Although a significant portion of the EMCC veers offshore during the summer months when the *Alexandrium* spp. bloom is normally at its peak, EMCC water can sometimes become subducted beneath the less-dense WMCC surface waters, causing a sub-surface *Alexandrium* spp. maximum that is occasionally detected at the pycnocline in the western Gulf of Maine. When *Alexandrium* spp. cells encounter unfavorable growing conditions they undergo a transformation into cyst form and sink to the benthos to await the return of

more favorable conditions, thereby re-seeding that area for the following year's bloom (Trainer *et al.* 2003). Maps of *Alexandrium* spp. cyst beds have been used in predictive modeling scenarios to try to forecast the spatial distribution and severity of the coming year's bloom, with varying degrees of success.

There are some areas within the Gulf of Maine that are more prone to high-toxicity *Alexandrium* spp. blooms than others. The eastern Maine and Bay of Fundy regions in particular are known for regularly occurring, very high-toxicity blooms that can produce toxicities almost an order of magnitude higher than in other parts of the Gulf of Maine (McGillicuddy *et al.* 2005). Parts of western Maine are also affected by high-toxicity blooms resulting from the transport of cells in buoyant plumes of estuarine water from river outflows (Franks and Anderson 1992). Wind can also affect patterns of toxicity through alongshore transport or through Ekman transport that can move offshore *Alexandrium* spp. cells into the nearshore where they can then affect shellfish beds (Franks and Anderson 1992; Townsend 2001). In Penobscot Bay, however, *Alexandrium* spp. blooms are relatively rare, a phenomenon that has been dubbed the "sandwich effect," since this area is lies between two regions where blooms are common (Hurst and Yentsch 1981; Shumway *et al.* 1988). The sandwich effect is considered to be the result of the EMCC transporting *Alexandrium* spp. cells offshore and away from Penobscot Bay during the summer months (Brooks and Townsend 1989; Pettigrew *et al.* 1998).

Evidence for the presence of *Alexandrium* spp. along the shores of current day New England and Atlantic Canada has been reported going back centuries, and recorded human PSP-related deaths in the Gulf of Maine region date from 1889 (Ganong 1889). The first monitoring program for PSP was initiated in Canada in 1945 and covered the Bay of Fundy

area (Townsend *et al.* 2001). The state of Maine began to monitor for PSP in 1957, but only in the eastern part of the state (Hurst 1975). The first major *Alexandrium* spp. bloom in the western Gulf of Maine occurred in 1972, after cells from the eastern Gulf of Maine were transported southwestward by the passage of Tropical Storm Carrie (often misidentified as a hurricane in the literature; N. Pettigrew, pers. comm.) (Twarog and Yamaguchi 1975; Franks and Anderson 1992; Anderson *et al.* 2005). This storm made landfall near Eastport, Maine and northeasterly winds associated with the storm moved *Alexandrium* spp. cells into the western Gulf of Maine where Ekman transport then carried those cells into the nearshore, resulting in the first large-scale PSP event in that region in recent history. Since then, large blooms of *Alexandrium* spp. have become a predictable annual occurrence throughout the Gulf of Maine, prompting Maine to expand its PSP monitoring program to cover the entire state coastline in 1974 (Hurst 1975). The U.S. Food and Drug Administration currently sets the regulatory limit for PSTs in shellfish at 80 μg STX eq 100 g^{-1} tissue.

In shellfish, both the rate of accumulation of PSTs and the level of toxins accumulated in the tissues are species-specific (Bricelj *et al.* 2005), and these factors have been demonstrated to correlate with particular species' sensitivity to VGSC inactivation by PSTs. For shellfish within the Gulf of Maine, blue mussels (*Mytilus edulis*) appear to be the least affected by PSTs, and can quickly accumulate these compounds in their body tissues when exposed to toxic strains of *Alexandrium* spp. (Twarog *et al.* 1972). The rapid toxication of this species during *Alexandrium* spp. blooms makes it well suited for use as a sentinel species in red tide monitoring programs. On the opposite end of the spectrum, eastern oysters (*Crassostrea virginica*) are known to be quite sensitive to PSTs, and begin

to experience harmful effects of VGSC inactivation at lower toxin concentrations than do mussels (Twarog *et al.* 1972). Softshell clams (*Mya arenaria*) fall somewhere between *M. edulis* and *C. virginica* in terms of their VGSC sensitivity and their susceptibility to the harmful effects of PSTs (Twarog *et al.* 1972). A reduced sensitivity to VGSC inactivation in *M. edulis* could explain why this species can accumulate such high toxin levels. Once a filter-feeding bivalve has consumed enough *Alexandrium* spp. to cause VGSC inactivation, they can no longer continue to feed due to the paralytic effects of the PSTs, and thus they can only reach a certain maximum threshold of toxicity. Since VGSC inactivation in *M. edulis* is more difficult to induce, this species can continue to feed and bioaccumulate toxins for a longer period of time than other shellfish during *Alexandrium* spp. blooms.

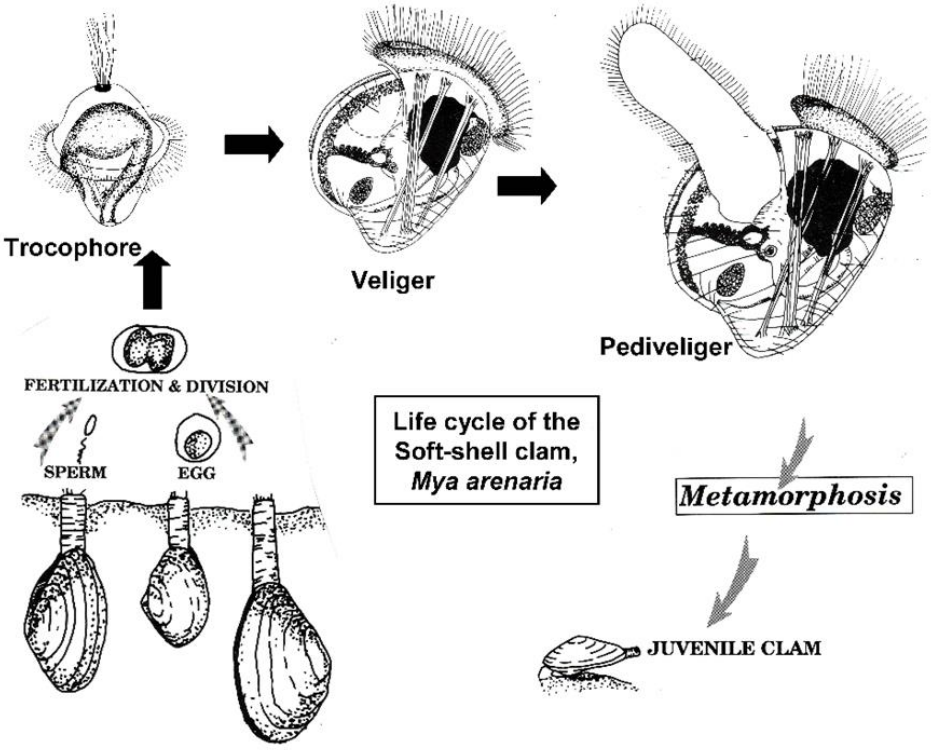
Large-scale *Alexandrium* spp. blooms can have substantial impacts on the shellfish industry in the Gulf of Maine, causing significant economic loss due to closures of harvesting areas, as well as upsurges in market price for shellfish and an increase in shellfish imported from outside the Gulf of Maine (Jin *et al.* 2008). The red tide that occurred in 2005 resulted in an estimated $15 million (USD) in economic losses in the state of Massachusetts alone (Anderson *et al.* 2005).

1.4. *Mya arenaria* life history

The softshell clam, *M. arenaria*, is a benthic, infaunal, filter-feeding bivalve that is found in soft bottom intertidal habitats throughout its native range in the northwest Atlantic Ocean, as well as numerous other locations around the globe where it has been artificially introduced. *Mya arenaria* reproduce through broadcast spawning, typically in the early summer between mid-May and mid-July (Caporale *et al.* 1997), although specific timing

can vary from region to region. Temperature cues are believed to trigger spawning, so variations in sea surface temperature can also influence spawning time from year to year, as well as by region. During spawning fertilization occurs in the water column, and the resulting trochophore, and then veliger larvae remain pelagic for four to six weeks (Caporale *et al.* 1997) (Figure 1.3). During this pelagic phase the larvae are at the mercy of the prevailing currents and can be transported large distances from their natal site. The hydrographic conditions at spawning sites can greatly influence the dispersal potential of these larvae and their connectivity with other populations. In a location where the water mass is incorporated into a gyre or eddy, or water movement is restricted by a shallow sill, the residence time of that water mass in that area can be large, and the potential for dispersal of larvae will be small. In a case like this, self-recruitment is an important factor in maintaining local populations. Conversely, if the spawning site experiences a strong directional current then larvae can be transported many kilometers from their natal site, and can provide an important source of recruits to other geographically distant populations. The length of time spent in the veliger stage is temperature-dependent, and when the larvae have developed to the point where they will soon be ready for settlement, they undergo a metamorphosis into the pediveliger stage that involves the development of a foot. At this point the larvae will be begin to settle on suitable soft-bottom habitat.

Figure 1.3 *Mya arenaria* life cycle



Stages of the life cycle of *M. arenaria* showing spawning and fertilization, the pelagic trochophore, veliger, and pediveliger larval stages, and settling and metamorphosis into the juvenile stage. (Image used with permission from the Downeast Institute.)

Newly settled clams can only burrow shallowly. These clams are vulnerable to biotic factors like predation from crustaceans and fish (Hunt and Mullineaux 2002) as well as abiotic factors like winter ice scouring until they are able to grow to a size where they can burrow more deeply (Beal 2004). Post-settlement mortality can be extremely high during the first few weeks (Hunt and Schiebling 1997), and predation in particular can have a profound effect on the survival rates of newly settled clams. Larval clams settle without

preference at all tidal heights within the intertidal (Beal *et al.* 2001), however predation pressure increases from the upper intertidal to the lower intertidal, which can affect clam distribution based on tide height. In the high-energy environment of the intertidal zone small juveniles (< 5 mm) can also be redistributed after settlement by tidal currents and sediment transport, affecting their eventual distribution as adults (Hunt and Mullineaux 2002). Abundance patterns of adult clams are therefore the result of the combination of predation pressure and sediment transport (Beal *et al.* 2001; Morse and Hunt 2013), rather than initial recruitment.

Temperature plays an important role in the maturation time of juvenile clams by directly affecting their metabolic rate, causing it to increase with increasing water temperature. While larval clams can tolerate a broad range of temperatures, including those as low as 10°C, the ideal temperature range for development and growth is 17-23°C (Stickney 1964). Regional differences in temperature between the eastern Gulf of Maine and the western Gulf of Maine as a result of differences in the EMCC and WMCC can affect the maturation rates of clams in those two areas. The length of time required for clams to grow to a legal harvest size of 2 inches (50.8 mm) depends on the water temperature, tidal height, and sediment type (Beal 2006). In the warmer waters of the western Gulf of Maine clams take an average of three years to reach legal harvest size, however in the eastern Gulf of Maine where sea surface temperatures are significantly cooler it can take clams up to eight years (Beal 2006). Tidal height is also an important factor for growth rate, because clams can only feed when they are submerged. The lower clams are in the intertidal, the longer they spend immersed during each tidal cycle and the more opportunity they have to feed and grow. On average, clams in the mid intertidal in

Maine take 4.5 to 6.5 years to attain legal harvest size (Beal 2006). Clams in the upper intertidal are only immersed for short periods of time and grow extremely slowly, taking upwards of eight years to attain legal harvest size (Beal 2006).

1.5. *Mya arenaria* fishery, management, and history of seeding

Mya arenaria is commercially fished in many areas, and in Maine the softshell clam fishery is one of the most economically valuable fisheries, averaging annual landings of 10.6 million pounds, with an estimated value of $14.6 million USD (Beal 2015). In recent history it has been ranked as the second or third most valuable fishery in the state behind the lobster and elver fisheries (Beal 2004), and it is the most commercially valuable bivalve fishery in the state (Congleton *et al.* 2006). This fishery uses a system of co-management between the state and local governments, with each municipality independently overseeing its own clam-flats and issuing licenses to harvesters under the larger oversight of the Maine Department of Marine Resources. At the local level, committees within each community devise an ordinance for their flats that typically includes the number of licenses that will be issued, a management plan for their shellfish resource, and the selection of a warden who will enforce the stipulations of the ordinance (Beal 2004; Beal 2006).

The state of Maine first authorized towns to begin regulating harvesting of clams through the issuing of licenses in 1901. The earliest clam management plan in Maine prohibited digging during the clams' summer spawning season from June 1st to September 15th (Beal 2004; Congleton et al. 2006). This restriction remained in place until the demand for live clams (rather than canned clams) began to rise in the 1930s, prompting some Maine counties to lift the ban on summer harvesting in 1937, and the remaining

counties followed suit in 1949 (Congleton et al. 2006). In 1935 a two-inch (50.8 cm) minimum harvesting size was enacted, however this restriction was repealed in 1963 to allow towns to establish their own size limits. The two-inch minimum harvesting size was reestablished in 1984, and remains in place today (Congleton et al. 2006). The number of clamming licenses being issued increased through the 1970s and the clam fishery recorded high landings during that time, however by the 1990s catch levels had fallen to historically low levels, despite the fact that licenses were still being issued in large numbers.

The serious decline of clam populations spurred exploration into the idea of seeding flats with cultured juvenile clams to attempt to bolster natural population numbers. Softshell clams were experimentally farmed as far back as the early 1900s (Beal 2004), but the first large-scale aquaculture effort began in 1987 with the founding of the Beals Island Regional Shellfish Hatchery (BIRSH) in Beals, Maine (Beal *et al*. 1995). BIRSH initially operated on a local scale with the goal of producing one million seed clams for the nearby towns of Machiasport, Roque Bluffs, Jonesboro, Jonesport, Beals, and Addison, but as the capacity of the facility grew additional towns began to participate in the seeding program (Beal 2004). BIRSH was given federal non-profit status in 1995, and in 2001 it was rebranded as the Downeast Institute for Applied Marine Research and Education (DEI). The facility relocated to Great Wass Island, ME in 2003, and grant funding enabled expansions of the DEI to include larger holding areas, algal culture rooms, a classroom, and a pier. Participation in the DEI's seeding program has grown to include over 50 towns in the states of Maine, Massachusetts, and Virginia.

The DEI's seeding program uses wild adult clams from the surrounding region as well as a hatchery-developed broodstock line to produce their seed clams. The broodstock

clams are conditioned in holding tanks at the DEI beginning in January, and after several months have passed and the clams have become ripe, spawning is induced through a thermal shock technique. The resulting larvae are pooled in large conical tanks where they are held until they develop to the pediveliger stage, at which point they are transferred onto floating screens where they can settle. By late fall these juvenile clams have reached a size at which they can be seeded, however fall seeding efforts result in high mortality of these individuals due to winter weather conditions and ice scour (Beal 2004). By overwintering the seed clams at the DEI and distributing them to towns for planting the following spring these losses can be avoided (Beal 2004). To mitigate the effect of predation on newly seeded clams the DEI recommends that seeded plots be protected with a flexible netting for the first year, until the seeded clams are able to grow to a size where they are less vulnerable to predators (Beal 2004).

1.6. *Mya arenaria* – green crab interactions

The softshell clam fishery in the Gulf of Maine has historically experienced steep declines as well as periods of recovery that are closely associated with changes in population numbers of the invasive European green crab, *Carcinus maenas*. *Carcinus maenas* was first reported in North America in 1817 in the mid-Atlantic United States, after likely being transported from its native Europe in ship ballast water. After becoming established in the U.S., the crabs steadily extended their range northward, with reported sightings on Cape Cod by the 1870s, in western Maine by the 1890s, in mid-coast Maine by the 1930s, and in eastern Maine and Canada by the 1950s (Scattergood 1952; Glude 1955).

Since the expansion of green crabs' range into the Gulf of Maine in the first half of the twentieth century, several studies have noted that green crab abundance strongly correlates with air and sea surface temperatures (Glude 1955; Welch 1969). A pattern of warmer, milder winters in the 1940s and 1950s was likely an important factor in green crabs' ability to expand their range into northern New England and Atlantic Canada (Ropes 1968). During this same time period the softshell clam fishery recorded a serious decline in landings, prompting the U.S. Bureau of Commercial Fisheries to issue an investigation that resulted in numerous studies citing green crabs as the cause of the decline (Glude 1955; Ropes 1968; Welch 1969). The late 1950s and 1960s brought a series of colder winters and decreased sea surface temperatures to the Gulf of Maine, causing significant mortality among green crabs. As green crab numbers markedly decreased, softshell clam populations rebounded (Welch 1969).

Softshell clam populations in the Gulf of Maine declined again over period of several decades between the late 1970s and 1990s, and from 1989 to 1999 the landings of softshell clams in Maine dropped by 75% (Beal *et al.* 1999). A second invasion of green crabs occurred on the Atlantic coast of Canada in the 1990s, and unlike their predecessors these green crabs originated from the northern end of their native European range and were therefore more cold-adapted than the first invaders (Roman 2006). The second invasion front quickly moved south into the Gulf of Maine, where it met and crossed the first invasion front. As a result, much of the Gulf of Maine now has populations of both warm and cold-adapted green crabs.

In recent years, increasing temperatures in the Gulf of Maine as a result of global climate change have been exacerbating the effects of green crab predation on softshell

clams. A warming trend in both the general climate and mean seawater temperatures has loosened the temperature constraint that has historically controlled the abundance of *C. maenas* in the Gulf of Maine, and as a result population numbers of green crabs have undergone dramatic growth. The predation effects of such a high number crabs have resulted in the loss of entire year classes of juvenile clams, and have severely depleted clam populations in the lower and mid-intertidal zones in many areas of Maine (Beal 2015). Dwindling clam populations are caused by a lack of newly settled spat and young clams recruiting to the fishery, and the impact of green crab predation on softshell clams is so strong that it is this intense predation, rather than recruitment, that drives patterns of clam distribution and abundance.

CHAPTER 2

SELECTION OF A SAXITOXIN-RESISTANT MUTATION IN *MYA ARENARIA* PRODUCES POPULATION GENETIC STRUCTURE

2.1. Introduction

Natural selective pressure imposed by an environmental stressor can have a significant effect on the genetic structure of a population of organisms, and softshell clams in the Gulf of Maine provide an excellent example of this phenomenon. Some softshell clams in areas prone to HABs of *Alexandrium* spp. have acquired a genetic mutation of their VGSCs that confers a resistance to the effects of PSTs produced by these algae, making this a highly beneficial mutation for clams that are commonly subjected to HAB conditions. Paralytic shellfish toxins can act as selective agents on populations of clams, favoring the individuals with this mutation.

Intraspecific differences in softshell clams' susceptibility to the effects of PSTs were first noted by Twarog *et al.* (1972), who surveyed nerve susceptibility to STX in eight different bivalve species. While some *M. arenaria* exhibit nerve block at moderate concentrations of STX (10^{-6} g/mL) (Twarog *et al.* 1972), other *M. arenaria* have nerve fibers that are much more resistant to the effects of STX, and exhibit only partial block even at high concentrations of STX ($< 10^{-4}$ g/mL) (Connell *et al.* 2007) (Table 2.1). While the genetic basis for these differential responses was not known in 1972, it was hypothesized that these variations could be the result of an alteration of the VGSCs in individuals that showed resistance to PSTs.

Table 2.1 Nerve sensitivity to STX among bivalve species

Species	Block of action potential by STX				
	10^{-8}	10^{-7}	10^{-6}	10^{-5}	10^{-4}
Mytilus edulis	0	0	0	0	0
Placopecten magellanicus	0	0	0	0	0
***Mya arenaria* (resistant)**	0	0	0	0	(+)
Mercenaria mercenaria	0	0	0	0	(+)
Modiolus demissus	0	0	0	0	+
Pecten irradians	0	0	0	(+)	+
***Mya arenaria* (sensitive)**	0	0	(+)	+	+
Crassostrea virginica	0	+	+	+	+
Elliptio complanata	0	+	+	+	+

Nerve sensitivity data from Twarog *et al.* (1972) and Connell *et al.* (2007) showing the concentration of STX in g/mL at which different bivalve species exhibit no block, partial block, or full block of action potential generation, as well as differences in nerve sensitivity between resistant and sensitive *M. arenaria*. (0 = no block, (+) = partial block, and + = full block).

Many organisms that are routinely exposed to these toxins have developed adaptations of their sodium channels that mitigate the harmful effects of these compounds (Anderson *et al.* 2005). For example, adaptive evolution through mutations of the VGSCs has been well characterized in a system of garter snakes (*Thamnophis sirtalis*) and poisonous TTX-producing newts (*Taricha* sp.), in which the garter snakes have acquired a resistance to TTX that allows them to prey on the newts (Brodie and Brodie 1999; Geffeney

et al. 2002; Feldman *et al.* 2009). Tetrodotoxin resistance in the garter snakes can result from one of several different amino acid substitutions in DIII and DIV of the sodium channel pore that neutralize the charged amino acid residues implicated in binding TTX, and cause a decrease in binding affinity of the TTX molecule by almost two orders of magnitude (Geffeney *et al.* 2002; Feldman *et al.* 2009). The different mutations in *T. sirtalis* form discrete spatial populations in these animals, and show that TTX resistance arose multiple times in different lineages of these snakes (Feldman *et al.* 2009).

Genetic adaptations in response to toxin exposure have also been documented in the copepod *Acartia hudsonica*, which displays differential sensitivity to STX between populations that are exposed to HABs and those that are not (Dan and Haley 2011). Saxitoxin resistance in *A. hudsonica* is not the result of a mutation that decreases binding affinity of the toxin molecule as seen in *T. sirtalis*, but instead results from a VGSC mutation that allows for "leakage" of sodium ions through the VGSC pore, even when the channel is blocked by STX (Chen 2010). The exact mechanism behind this method of resistance has proven to be complex, and is still being investigated.

Building on the results of Twarog *et al.* (1972), differences in softshell clams' responses to STX were described by Bricelj *et al.* (2002), who used a burrowing index to distinguish between STX-resistant and STX-sensitive clams. When saxitoxin-sensitive clams feed on *Alexandrium* spp. they are subject to paralytic effects that inhibit their ability to burrow, however STX-resistant clams are unaffected and retain the ability to burrow after feeding on *Alexandrium* spp. By exposing softshell clams to *Alexandrium* spp. in a laboratory setting and documenting the number of clams that are able burrow, the percentage of STX-resistant individuals can be determined. Bricelj *et al.* (2002) analyzed

clam populations from the Gulf of St. Lawrence, Bay of Fundy, Gulf of Maine, and Puget Sound, and found that sites with a history of PSP contained clam populations that were largely STX-resistant based on the burrowing index. Sites with no history of PSP had clam populations that were determined to be predominantly STX-sensitive by the burrowing index (Bricelj *et al.* 2002).

Research by Bricelj *et al.* (2005) focusing on *M. arenaria* described a base pair substitution (NaV) in the pore region of Domain II of the VGSCs in clams that exhibit resistance to PSTs that results in a single amino acid substitution. In these clams, the glutamic acid residue (E) in Toxin Site 1 (equivalent to E945 in the rat sodium channel) (Auld *et al.* 1990), which plays a critical role in binding STX, is replaced with an aspartic acid residue (D), causing a 1,000-fold decrease in the binding affinity of the STX molecule. While only one VGSC mutation is present in *M. arenaria*, this mutation appears to have arisen multiple times within populations of these animals (Connell *et al.* 2007). There are six possible nucleotide combinations that can occur at the NaV mutation locus (Table 2.2): three that cause the E → D amino acid substitution and result in a STX-resistant genotype (RR), one wild type that results in a STX-sensitive genotype (SS), and two that produce a heterozygous genotype (RS) with STX resistance between that of the RR and SS genotypes. The R allele is heritable, and is passed on to subsequent generations through simple autosomal Mendelian inheritance patterns (Hamilton 2009).

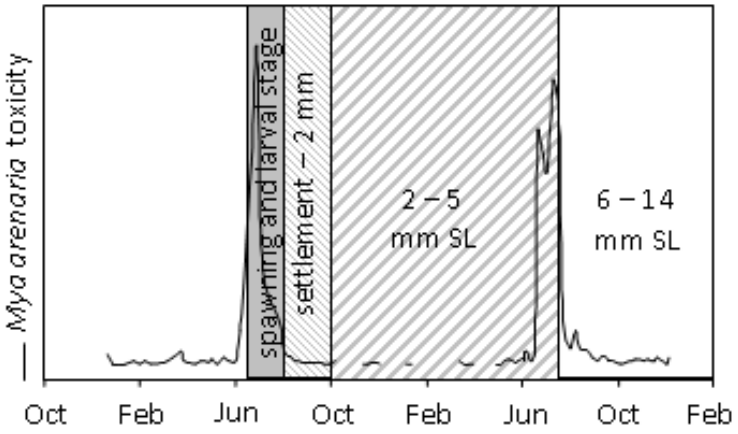
Table 2.2 Identified genotypes at the NaV mutation locus

Nucleotide	Amino Acid	Genotype
A/A	E/E	SS
A/C	E/D	RS
A/T	E/D	RS
C/C	D/D	RR
T/T	D/D	RR
C/T	D/D	RR

Six different nucleotide combinations observed at the NaV mutation locus, the encoded amino acid, and the resulting genotype exhibited in *M. arenaria.* SS indicates the STX-sensitive genotype, RR indicates the STX-resistant genotype, and RS indicates the heterozygous genotype. (Adapted from Bricelj *et al.* 2005).

When mixed populations of RR and SS *M. arenaria* were exposed to a toxic strain of *Alexandrium* sp. under laboratory conditions to simulate the effect of a HAB, the simulated bloom displayed the potential to be a strong agent of selection in favor of the RR clams (Bricelj *et al.* 2005; Bricelj *et al.* 2010). Clams with the SS genotype quickly became incapacitated after ingesting *Alexandrium* spp. and were unable to continue feeding or burrow into the substrate. These changes in behavior could lead to mortality in the wild, especially among recently settled clams in the young juvenile stage (~ 4 mm shell length). This life stage is the most vulnerable to the disabling effects of STX exposure (Bricelj *et al.* 2010), and is also the most vulnerable to the effects of predation, therefore the negative impact of HABs on these individuals is likely to be high (Figure 2.1)

Figure 2.1 Effect of *Alexandrium* spp. blooms on different life stages of *Mya arenaria*



Tissue toxicity in *M. arenaria* throughout the course of several years, superimposed on the timing of the spawning, larval, and young juvenile life stages. Selection from PSP events is strongest on 2-5 mm shell length juveniles. Larval stages are unaffected by PSP, since these individuals are too small to ingest *Alexandrium* spp. cells. (Figure from Bricelj *et al.* 2010; used with permission from M. Bricelj.)

Using information from a previous study that investigated burrowing (Bricelj *et al.* 2002), Connell *et al.* (2007) examined sensitivity to STX through an *in vitro* nerve assay in *M. arenaria* collected from several sites with a history of HABs, as well as from sites without a history of HABs, and found differential responses to STX in the nerves of clams from HAB and non-HAB sites. Comparisons between DNA sequences of the NaV mutation locus in the sampled clams and the results of the *in vitro* nerve assays demonstrated that this method is highly accurate at distinguishing between genotypes (Connell *et al.* 2007).

In softshell clams, behavioral responses to STX as indicated by burrowing index, and genotype at the NaV mutation site as indicated by DNA sequencing and *in vitro* nerve assay go hand in hand with the history of PSP at that particular location. This shows that selective pressure from *Alexandrium* spp. blooms drives population genetic structure with regard to STX resistance.

Adaptations that provide a benefit, such as the sodium channel mutation in *M. arenaria,* often come with an associated cost, as specific specializations can interfere with the organism's ability to perform other tasks (Brodie and Brodie 1999). In the case of the garter snake-newt system, the sodium channels of TTX-resistant garter snakes exhibit lower conductance, slower kinetics, slower generation of action potentials and changes in gating that are not present in the sodium channels of TTX-sensitive snakes (Brodie and Brodie 1999; Geffeney *et al.* 2002; Lee *et al.* 2011). As a result, physiological processes under neuromuscular control in TTX-resistant snakes are significantly slower than in the snakes without this VGSC mutation (Brodie and Brodie 1999).

In *M. arenaria*, clams that carry the STX-resistant mutation appear to be subjected to the same neuromuscular costs associated with resistance as the garter snakes do. These costs manifest in the form of slowed rates of burrowing (Walter unpublished), and negative effects on growth rate (Bricelj *et al.* unpublished; Connell *et al.* unpublished). Field plantings of juvenile clams conducted at a site where HABs were absent found that SS clams showed significantly higher shell growth rates than RR clams (Bricelj *et al.* unpublished). In addition, in laboratory experiments where RR and SS clams were fed a diet of non-toxic algae, SS clams displayed significantly higher shell and tissue growth rates after a period of three weeks (Bricelj *et al.* unpublished). These experiments used

clams originating from different source populations, however, so population-specific effects cannot be ruled out. Differences in growth rate between the resistant and sensitive genotypes are likely the result of differences in feeding rate, since feeding in *M. arenaria* is under neuromuscular control and therefore likely affected by the NaV mutation.

These growth rate effects mean that clams with the RR genotype spend a longer time in the small juvenile life stage where they are vulnerable to predation and other environmental effects. Thus, under non-HAB conditions, SS *M. arenaria* may have higher fitness than RR individuals. This implies that the resistant NaV mutation is only beneficial in places that have a high likelihood of experiencing a HAB during critical growth periods, and that in places where HABs are rare, or where bloom timing does not coincide with the juvenile life stage, it may be disadvantageous to possess the mutation.

In the wild, selective pressure exerted on *M. arenaria* populations by HABs will be spatially heterogeneous as a result of variability in occurrence, toxicity, and cell densities of blooms among different regions of the Atlantic coast of North America. In particular, there are well-described differences between HABs in the Gulf of Maine, and HABs in southern New England, as well as around Long Island, New York. Harmful algal blooms in the Gulf of Maine tend to be widespread in occurrence, and are typified by relatively low cell densities but high toxicities (Anderson *et al.* 1994). In contrast, HABs in southern New England and New York tend to occur as localized blooms that arise in a single embayment or estuary, and display high cell densities but much lower toxicities than those seen in Gulf of Maine HABs (Anderson *et al.* 1994; Hattenrath *et al.* 2010). Disparity in the toxicity of HABs between these two regions arises out of differences in the toxin compositions of their resident populations of *Alexandrium* spp. Gulf of Maine populations

of *Alexandrium* spp. produce more of the highly potent carbamate toxins than southern New England populations of *Alexandrium* spp., and these southern populations produce more of the weaker sulfamate toxins (Anderson *et al.* 1994). As a general trend, toxicity of *Alexandrium* spp. blooms on the Atlantic coast of North America tends to decline with decreasing latitude (Maranda *et al.* 1985).

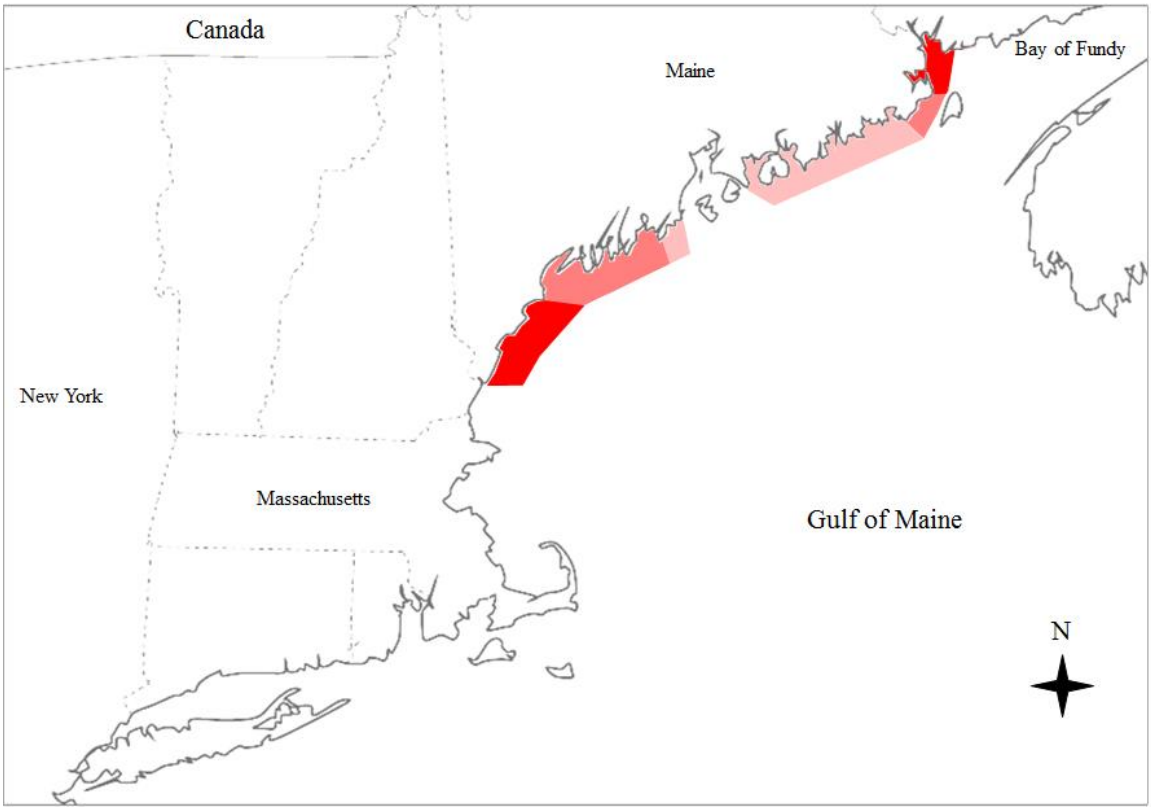
This work used DNA sequencing to survey genotypes for STX resistance in wild *M. arenaria* populations throughout the Gulf of Maine and southern New England, and compared the frequency of those genotypes to the history of HABs at the sampled locations to determine whether the strength of selective pressure from *Alexandrium* spp. correlated with genotype frequencies in the sampled clam populations. Based on the knowledge of the tradeoff between the benefits and costs of possessing the resistant NaV mutation, I hypothesized that the same selective pressure exerted by *Alexandrium* spp. in laboratory experiments is also taking place in *M. arenaria* populations in the natural environments of the Gulf of Maine and southern New England, and that patterns of STX-resistance among wild *M. arenaria* will reflect the patterns of *Alexandrium* spp. blooms in these locations. I predict that in areas where HABs are common *M. arenaria* populations will be largely comprised of individuals possessing the R allele, and that in areas where HABs are relatively rare populations will be largely comprised of individuals possessing the S allele.

2.2. Methods

Twenty-seven sample sites were clustered *a priori* into five different regions based on history of PSP in the Gulf of Maine (Pettigrew *et al.* 2005; Anderson *et al.* 2014) (Figure

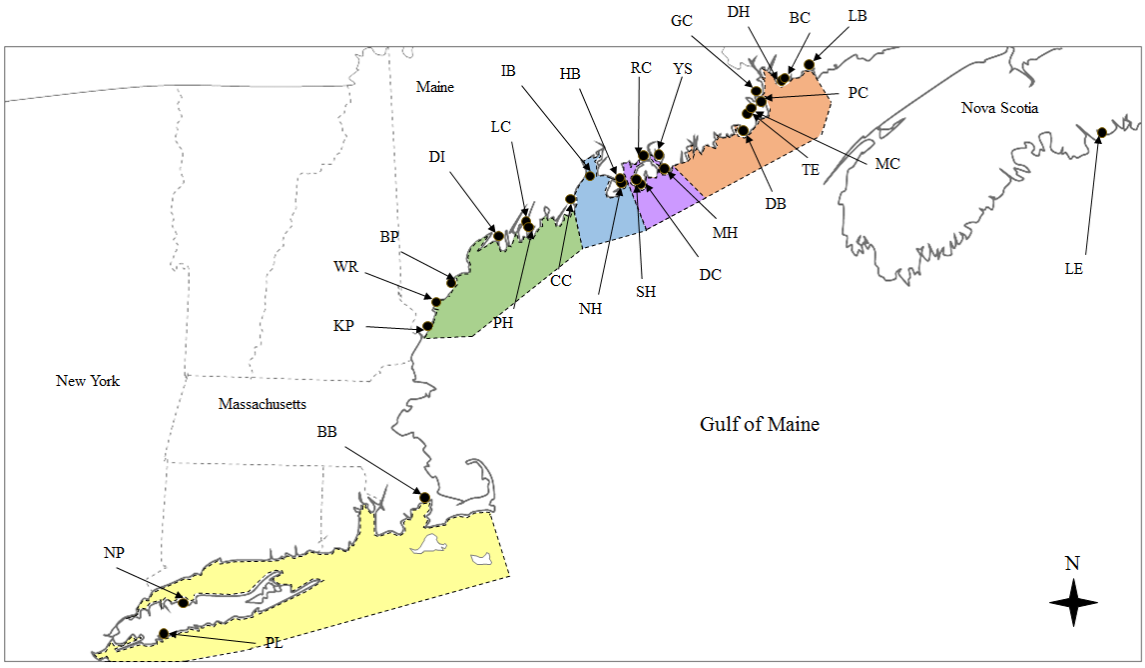
2.2). Adult *M. arenaria* were collected using a hand rake from the upper intertidal zone of each sample site (Figure 2.3).

Figure 2.2 PSP History in the Gulf of Maine



History of PSP in the Gulf of Maine based on maximum tissue toxicity data from the sentinel species *Mytilus edulis*, as a proxy for the approximate intensity of *Alexandrium* spp. blooms (Thomas *et al.* 2010). The darkest red shading indicates areas with tissue toxicity ranging from 1,600 to 10,000 µg STX eq 100 g^{-1}, the medium red shading indicates areas with tissue toxicity ranging from 700 to 3,000 µg STX eq 100 g^{-1}, and the light red shading indicates areas with tissue toxicity ranging from 10 to 1,000 µg STX eq 100 g^{-1}. In general, blooms are extremely toxic in eastern Maine, Cobscook Bay, and the far western Gulf of Maine, moderately toxic around the midcoast, weakly toxic around MDI, and absent in Penobscot Bay.

Figure 2.3 *Mya arenaria* sample sites



Pre-defined regions of the Gulf of Maine, southern New England, and Long Island, New York, where *M. arenaria* were collected. Red indicates the eastern Gulf of Maine (EGoM) region, purple indicates the Mount Desert Island (MDI) region, blue indicates the Penobscot Bay (PB) region, green indicates the western Gulf of Maine (WGoM) region, and yellow indicates the southern New England and New York (SNE-NY) region. Sample sites are shown within each region (site abbreviations are defined in Table 2.3).

Eastern Gulf of Maine (EGoM) was designated as the area from the Schoodic peninsula east into southern New Brunswick, including Cobscook Bay (Townsend *et al.* 2001). Mount Desert Island (MDI) was designated as the area between the Schoodic peninsula and the town of Brooklin on the eastern side of Penobscot Bay, including MDI. Penobscot Bay (PB) was designated as the area from Rockland on the western side of Penobscot Bay to Brooklin on the eastern side of the Bay, including the island of Islesboro.

Western Gulf of Maine (WGoM) was defined as the area from Penobscot Bay west to the Maine-New Hampshire border. Southern New England and New York (SNE-NY) was defined as an area encompassing the southern side of Cape Cod, and the north and south shores of Long Island. Clams were collected from eight sites in EGoM, five sites in MDI, four sites in PB, six sites in WGoM, and three sites in SNE-NY. One site, Lawrencetown Estuary on the eastern side of Nova Scotia, was left as a stand-alone site. The number of *M. arenaria* collected during sampling at each site ranged from 13 to 120, with an average of 40 animals per site. (Table 2.3).

Table 2.3 *Mya arenaria* sample site details

Site name	Abbreviation	Latitude & longitude	Year(s) sampled	Number of samples
Lawrencetown Estuary	LE	44.64914 -63.36816	2007	36
Lepreau Basin	LB	45.12852 -66.46151	2007	108
Beaver Cove	BC	45.07734 -66.73714	2009	48
Deadman's Harbour	DH	45.04906 -66.77008	2009	30
Prince Cove	PC	44.89589 -66.99306	2008	30
Gleason Cove	GC	44.97270 -67.05516	2008	30
Morong Cove	MC	44.85635 -67.10387	2010	47
Whiting Cove	TE	44.81096 -67.16156	2007	30
Duck Brook	DB	44.66274 -67.1879	2007	30
Mosquito Harbor*	MH	44.37432 -68.07194	2007, 2012	30, 20
Young's Shore	YS	44.48001 -68.11656	2007	30
Raccoon Cove*	RC	44.47189 -68.28261	2007, 2012	30, 66
Ship Harbor	SH	44.23005 -68.32362	2007	30
Duck Cove	DC	44.25485 -68.37732	2007	30
Naskeag Harbor*	NH	44.22959 -68.53542	2008, 2012	31, 50
Herrick Bay	HB	44.26508 -68.55417	2007	30
The Narrows	IB	44.31336 -68.90184	2015	30
Clam Cove	CC	44.13480 -69.09242	2015	30

Poorhouse Cove*	PH	43.88363 -69.55011	2009, 2010, 2011, 2012	100, 106, 277, 121
Lowes Cove	LC	43.93675 -69.57684	2009	21
Dingley Island	DI	43.81715 -69.88579	2012	53
Biddeford Pool	BP	43.43741 -70.36946	2009	34
Webhannet River	WR	43.32035 -70.57177	2010	48
Kittery Point	KP	43.10706 -70.66622	2012	19
Buttermilk Bay	BB	41.75067 -70.62692	2007	32
Northport Harbor*	NP	40.90103 -73.35386	2010, 2013	100, 105
Point Lookout	PL	40.59984 -73.56500	2009	40

Sites where *M. arenaria* were collected, including the location, year(s) sampled, and the number of clams that were sampled at each site. Sites marked with an asterisk indicate sites that were sampled multiple times.

Immediately upon collection each individual was measured using calipers and the shell length was recorded. DNA samples were obtained in a non-lethal manner by using a 1 mL Monoject™ Tuberculin syringe (Covidien) to extract 0.3 mL of hemolymph from the anterior adductor muscle of each animal. Following hemolymph extraction clams were typically returned to the flat. The hemolymph was transferred to individually numbered sterile tubes and placed on ice for transport to the Connell Laboratory at the University of Maine. The tubes of hemolymph were centrifuged (Eppendorf® Microcentrifuge 5418) at 10,000 rpm for 5 minutes to produce a cell pellet in each tube. The supernatant was pipetted off of each sample, and the cell pellet was resuspended in 50uL of a buffering

solution of 10mM Tris-HCl, pH 8.5. These hemolymph pellets can be used directly in PCR reactions as DNA templates (Hamilton and Connell 2009). The hemolymph was stored at -20°C.

At one site (Northport Harbor) sampled clams were frozen whole and shipped to the Connell Laboratory. A small piece of mantle tissue was excised from each of these clams, and DNA was isolated from that tissue using a DNeasy Plant Mini Kit (QIAGEN). DNA was stored at -20°C.

The hemolymph and tissue DNA were used in PCR reactions designed to amplify a 172 bp segment of the sodium channel gene containing the mutation site. Amplification was carried out in 25 μL total volume in illustra PuReTaq Ready-To-Go PCR Bead tubes (GE Healthcare Life Sciences) containing 22 μL of HyClone™ HyPure Molecular Biology Grade Water (GE Healthcare Cell Culture), 0.5 μL of each primer (oBTG-99F and oBTG-100R), and 2 μL of hemolymph (~50 ng of DNA) or DNA extract. A no-DNA control was also prepared with each set of PCR reactions. Cycling was done in a MJ Research PTC-200 Thermal Cycler (GMI Laboratory Instruments) and cycling conditions were set at (i) 95°C for 5 minutes, followed by 35 cycles of (ii) 95°C for 30s, (iii) 50°C for 30s, and (iv) 72°C for 120s, and finishing with (v) 72°C for 10 minutes.

The resulting PCR products and no-DNA controls were visualized on an I.D.$_{N.A.}$™ agarose gel to ensure that each sample had amplified successfully and was uncontaminated with non-sample DNA. The PCR products were purified using the Wizard© SV Gel and PCR Clean-Up System (Promega), and the DNA concentration of each sample was quantified using a NanoDrop© ND-1000 Spectrophotometer (Thermo Scientific). The purified PCR products were prepared for DNA sequencing in accordance with the

specifications of the University of Maine DNA Sequencing Facility, where they were sequenced using ABI BigDye™ Terminator chemistry. The raw data were analyzed using Sequencher® 5.0.1 software (Gene Codes Corporation), and the genotype of each individual was identified. The individual site data were pooled within each of the five regions, and homogeneity of the sample sites within each region was tested using a replicated goodness-of-fit test based on the G-statistic. Significantly heterogeneous sample sites in each region were identified with unplanned tests for homogeneity of replicates (Sokal and Rohlf 1981). Differences in the genotype frequencies between each of the regions was assessed with G-tests of independence (Sokal and Rohlf 1981).

Five of the original sample sites were resampled again after a period of three to five years (Table 2.3). The number of clams collected at each repeated site ranged from 18 to 109, with an average of 62 clams per site. Each clam was measured using calipers, and the shell length was recorded. A hemolymph sample was obtained from each individual using the same methods described above. The hemolymph was used in PCR reactions and the resulting products were sequenced using the same protocol described above. The genotype of each individual was determined, and G-tests of independence were used to ascertain whether observed differences in genotype proportions at each site over time were statistically significant. If significant changes in the genotype proportions were detected, the genotypes causing those changes were identified using Fisher's exact tests of independence (Sokal and Rohlf 1981).

At four of the five repeat sample sites clam shell length was used a proxy for age to ensure that the same year classes were not being resampled. At the Northport Harbor site a subset of the shells were sectioned by the Cerrato Laboratory at Stony Brook

University and used to calculate the age range of the sampled individuals at that site following the methods of Cerrato *et al.* (1991).

2.3. Results

Testing for homogeneity of genotype frequencies among the sample sites within each of the five regions determined that the sites within each region were homogeneously grouped, with the exception of two sites. The Morong Cove site in the EGoM region was identified as being significantly heterogeneous (G-value = 10.278, df= 2, p = 0.006), meaning that genotype frequencies at this site were significantly different from those at the other sites in this region. In addition, the Buttermilk Bay site in the SNE-NY region was significantly heterogeneous from the other sites in that region (G-value = 9.762, df = 2, p = 0.008) based on repeated G-tests of goodness-of-fit and unplanned comparisons analysis. Subsequent data analyses were conducted both with and without these two heterogeneous sites, and no differences in the general patterns of results were observed.

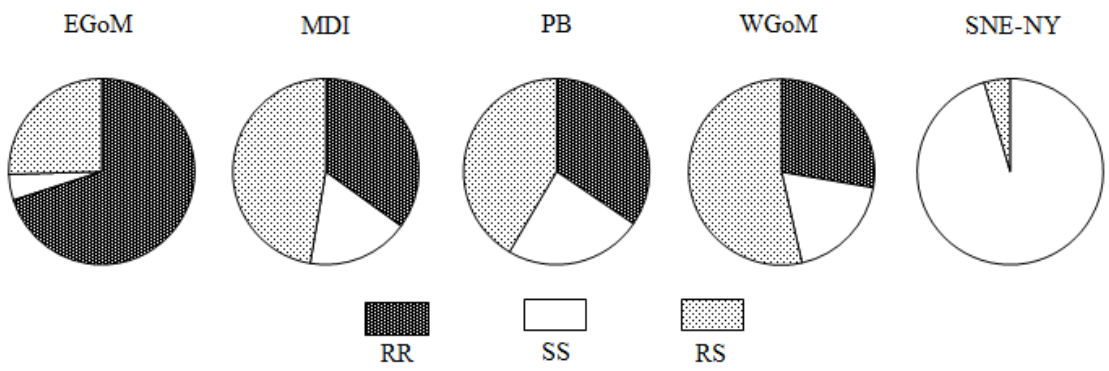
All of the three genotypes described with respect to STX resistance were found in each of the four Gulf of Maine regions. The EGoM region (n = 267) had the largest percentage of RR individuals at 70.0%, and also had the lowest percentage of RS individuals and SS individuals among the Gulf of Maine sites, at 25.5% and 4.5%, respectively. Populations in the MDI region (n = 186) were mostly comprised of the RS genotype at 47.3% of individuals, and also had a large percentage of RR individuals at 34.9%. The SS genotype comprised 17.7% of the individuals in the MDI region. The PB region (n = 142) had an approximately equal frequency of genotypes, with RR individuals comprising 34.5% of clams in that region, RS individuals comprising 41.5%, and SS

individuals comprising 23.9%. The WGoM region (n = 527) had the highest percentage of RS individuals of all the regions, comprising over half of the clams sampled at 53.5%. The WGoM region had the lowest percentage of RR individuals in the Gulf of Maine at 27.7%, and also had a low percentage of SS individuals at 18.8%. The Gulf of Maine regions showed a general trend of decreasing incidence of the RR genotype and increasing incidence of the RS genotype from east to west, while the SS genotype was most prevalent in Penobscot Bay.

Only two of the three genotypes were observed in the SNE – NY region (n = 176), with the RR genotype being absent in those sampled individuals. This region was almost entirely comprised of the SS genotype, with 95.5% of clams sampled having two S alleles. A small number of clams in this region displayed the RS genotype, with 4.5% of individuals sampled there being RS. (Figure 2.4).

The Lawrencetown Estuary site (n = 36) was primarily comprised of the SS genotype at 86.1% of the population, with a small number of clams possessing the RS genotype at 13.9%. The RR genotype was absent in clams sampled at this site.

Figure 2.4 Genotype proportions of *Mya arenaria* by region



The genotype frequencies with respect to STX resistance of sampled *M. arenaria* in each of the five regions of this study: eastern Gulf of Maine (EGoM), Mount Desert Island (MDI), Penobscot Bay (PB), western Gulf of Maine (WGoM), and southern New England – New York (SNE – NY). RR indicates the STX-resistant genotype, SS indicates the STX-sensitive genotype, and RS indicates the heterozygous genotype.

G-tests of independence performed among all five regions showed that genotype proportions were significantly different between the EGoM region and all other regions, and between the SNE-NY region and all other regions, but that genotype proportions among the MDI, PB, and WGoM regions did not significantly differ from each other (Table 2.4). These results did not change when the two heterogeneous sample sites (Morong Cove and Buttermilk Bay) were excluded from the analysis.

Table 2.4 Tests of independence for genotype frequencies between regions

	EGoM	MDI	PB	WGoM	SNE-NY
EGoM					
MDI	**1.02E-13**				
PB	**1.66E-13**	0.349			
WGoM	**2.77E-32**	0.084	0.023		
SNE-NY	**4.28E-113**	**2.25E-66**	**7.81E-51**	**9.31E-93**	

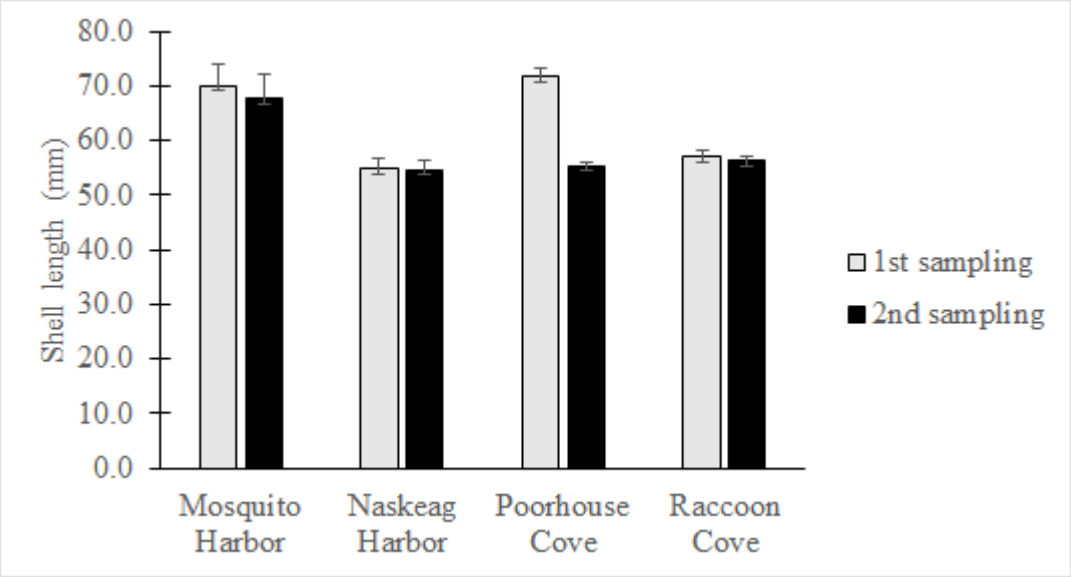
G-tests of independence based on the distribution of genotype frequencies among all five regions. Comparisons that are significant based on a Bonferroni-adjusted p-value of 0.005 are shown in bold. (EGoM = eastern Gulf of Maine, MDI = Mount Desert Island, PB = Penobscot Bay, WGoM = western Gulf of Maine, SNE-NY = southern New England-New York).

There were no statistically significant changes in genotype proportions over the period of time between samplings at four of the five repeatedly sampled sites. One repeated site, Naskeag Harbor, did show a significant change in genotype proportions over the four-year period between 2008 and 2012. A G-test of independence found that genotype proportions differed significantly between the first and second samplings at this site ($p = 0.00057$), and Fisher's exact test of independence determined that there was a significant decrease in the number of individuals displaying the RR genotype and a significant increase in the number of individuals displaying the RS genotype ($p = 0.017$). There was no significant change in the percentage of SS individuals.

Analysis of the size and age data taken from clams at each of the repeated sites showed that the repeat samples did not belong to the same year classes as the initial samples. At four sites the average shell length of individuals from the second sampling

taken three to five years after the first sampling was either the same or smaller than the average shell length of individuals from the first sampling, meaning that a lack of temporal changes in genotype proportions was not due to resampling individuals from the same year class (Figure 2.5).

Figure 2.5 Average shell length of *Mya arenaria* at sites sampled twice



The average shell length in millimeters of clams at four of the repeatedly sampled sites. Mosquito Harbor was sampled in 2007 and 2012, Naskeag Harbor was sampled in 2008 and 2012, Poorhouse Cove was sampled in 2009 and 2012, and Raccoon Cove was sampled in 2007 and 2012.

At the Northport Harbor site, clams from the first sampling in 2010 ranged in shell length from 38.2 to 83.6 mm, and in age from 1-7 years. This sample set included clams that experienced pre-bloom conditions at the time of recruitment, as well as clams recruited

after the first PSP event at this site in 2006. Individuals collected during the second sampling at this site in 2013 ranged from 38-42 mm in size and were 2-3 years old based on age results for shells of similar size collected in Northport Harbor in 2013. No differences in genotype frequency based on age were found across the age range sampled at this site.

2.4. Discussion

The results of this study provide a snapshot of the geographic distribution of the resistant NaV mutation in wild *M. arenaria* populations in the Gulf of Maine, as well as the variability of the incidence of the mutation among populations of clams in this environment. The resistant allele was observed as far south as the north shore of Long Island, where a very small proportion of the population sampled had the RS genotype. Since the population sampled on the south shore of Long Island was entirely comprised of the SS genotype, this appears to mark the limit of the southern extent of the resistant allele. The northern distribution of the NaV mutation extends into the Canadian Maritime Provinces, and while the scope of this study only ranged as far north as New Brunswick and Nova Scotia, the incidence of toxic HABs of *Alexandrium* spp. extends to the Bay of Fundy, the Gulf of St. Lawrence, and Newfoundland, Canada (Anderson *et al.* 1994). Clams from some locations in these regions have been examined using a burrowing index that has distinguished between resistant and sensitive genotypes with high accuracy (Bricelj *et al.* 2002), as well as through *in vitro* nerve assays and DNA sequencing (Connell *et al.* 2007). The results of these analyses indicated that the resistant allele is present in clam populations at sites in those regions that have a history of HAB exposure, and it is

probable that the resistant allele exists throughout *M. arenaria* populations that experience high-toxicity HABs on the Atlantic coast of North America. Connell and Bricelj (unpublished) also analyzed genotype frequencies of softshell clams from Brittany, France using DNA sequencing and found that the R allele was present in those populations.

The prediction that patterns of resistance in softshell clams would reflect the selective pressure exerted by the severity of toxin exposure from HABs was supported by the results from most of the regions in this study. In the EGoM region, high-toxicity HABs are a regularly occurring annual phenomenon, with cells reaching concentrations as high as 3.19×10^5 L^{-1}, and *M. arenaria* tissue toxicities being recorded in excess of 9000 μg STX eq 100 g^{-1} (Martin *et al.* 2014). This region had the highest percentage of the RR genotype seen in this study with 90.6% of the *M. arenaria* that were genotyped possessing at least one R allele, consistent with the hypothesis that HABs are exerting strong selective pressure on *M. arenaria* populations in the EGoM region. The fairly even frequency of genotypes among RR, SS, and RS in the MDI region suggest that HABs have a moderate impact on *M. arenaria* populations in this area. HABs of high toxicity are a regular occurrence in parts of the WGoM region, where tissue toxicities in the sentinel species *Mytilus edulis* can reach 300 μg STX eq 100 g^{-1} (Thomas *et al.* 2010). Eighty-one point two percent of *M. arenaria* sampled in the WGoM region possessed at least one R allele, although the majority of those individuals displayed the RS genotype. Even so, this still shows evidence that some degree of selection at the NaV locus is taking place in that region (Figure 2.6).

Another possible explanation is that RS genotype clams experience increased overall fitness compared to the other two genotypes. One R allele is enough to confer

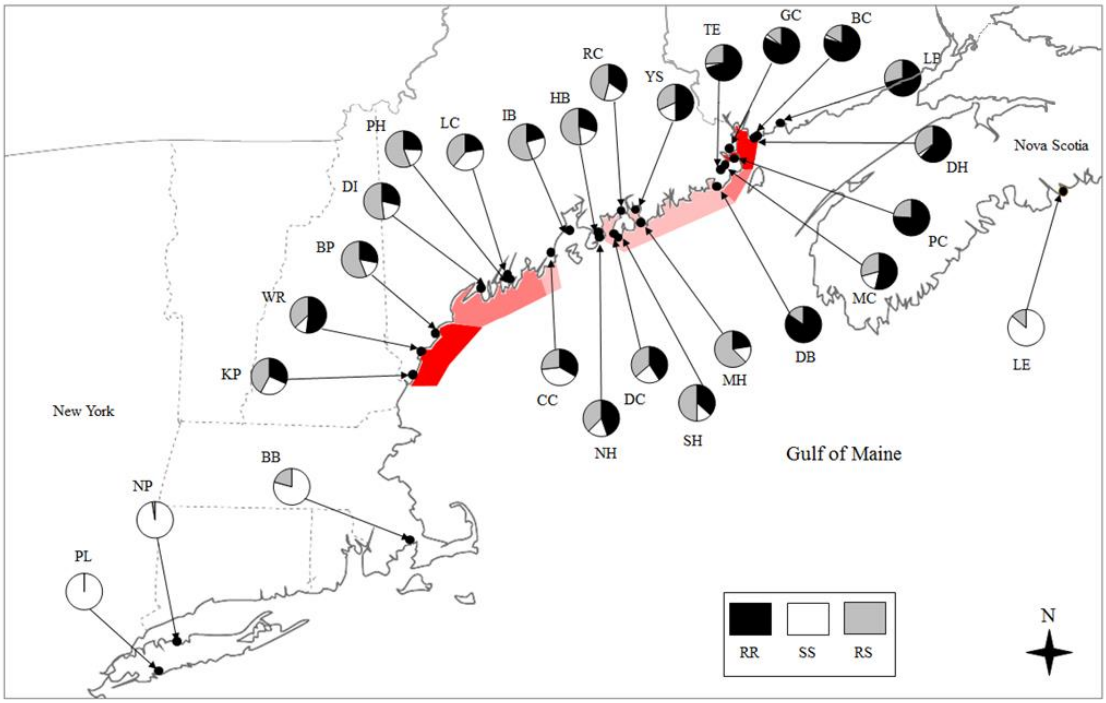
considerable resistance to STX, and RS clams do not begin to experience nerve block associated with STX exposure until they are exposed to the same toxin concentration that causes nerve block in the RR genotype, however RS clams exhibit a greater nerve block than the RR genotype at the same concentration of STX (Bricelj *et al.* unpublished). Costs associated with STX resistance have not been assessed in the RS genotype, so it is unknown whether these individuals experience burrowing and growth rate effects to the same degree as the RR genotype, or whether costs are less severe for these clams. If RS genotype clams see a lesser impact from the neuromuscular costs associated with resistance, then these individuals, which still have relatively high resistance to STX, could be the genotype with the greatest overall fitness for a broad range of habitats.

Harmful algal blooms are an infrequent phenomenon at the Lawrencetown Estuary site in eastern Nova Scotia, and those that do form tend to be small and localized in their distribution, as opposed to the widespread blooms that occur in the Gulf of Maine (Anderson *et al.* 1994). Genotype proportions in this population reflected a high incidence of the SS genotype at 86.1% of the population. The remainder of this population possessed the RS genotype, and none of the clams sampled at this site had the RR genotype. These genotype proportions are expected given the relative lack of selective pressure at this site, and the costs associated with the resistant allele.

Harmful algal blooms of *Alexandrium* spp. have become a regular event in the SNE-NY region within the last decade, although these HABs are of much lower toxicity than HABs in the Gulf of Maine due to differences in the toxin profiles of blooms in these two areas (Anderson *et al.* 1994; Hattenrath *et al.* 2010). Despite the high densities of these HABs, *M. arenaria* from the SNE-NY region were almost entirely comprised of the SS

genotype, and of the 4.5% that possessed an R allele, all of those individuals were the RS genotype. These results suggest that patterns of resistance in softshell clams are not driven solely by the presence or absence of HABs, or by *Alexandrium* spp. cell density, but that the maximum toxicity of those blooms plays an essential role in driving selection at the NaV locus, and determining the most advantageous genotype. The benefits associated with resistance must outweigh the associated neuromuscular costs to the clams in order for a genotypic shift to take place within a population, and when HABs are of low toxicity, as seen in the SNE-NY region, it may not be enough to tip the balance in favor of the RR and RS genotypes, even if the HABs are regularly occurring at high cell densities.

Figure 2.6 *Alexandrium* spp. bloom toxicity and genotype proportion by site



Genotype composition of populations superimposed over maximum tissue toxicity data from the sentinel species *Mytilus edulis*, as a proxy for the approximate intensity of *Alexandrium* spp. blooms (Thomas *et al.* 2010). The darkest red shading indicates areas with tissue toxicity ranging from 1,600 to 10,000 µg STX eq 100 g^{-1}, the medium red shading indicates areas with tissue toxicity ranging from 700 to 3,000 µg STX eq 100 g^{-1}, and the light red shading indicates areas with tissue toxicity ranging from 10 to 1,000 µg STX eq 100 g^{-1}. In general, blooms are extremely toxic in eastern Maine, Cobscook Bay, and the far western Gulf of Maine, moderately toxic around the midcoast, weakly toxic around MDI, and absent in Penobscot Bay. (See Table 2.3 for site abbreviations.)

HABs are relatively uncommon in PB because of oceanographic features in this region that inhibit the formation of blooms within the bay, and recorded tissue toxicities in *M. edulis* rarely rise above the regulatory threshold of 80 µg STX eq 100 g^{-1} (Thomas *et*

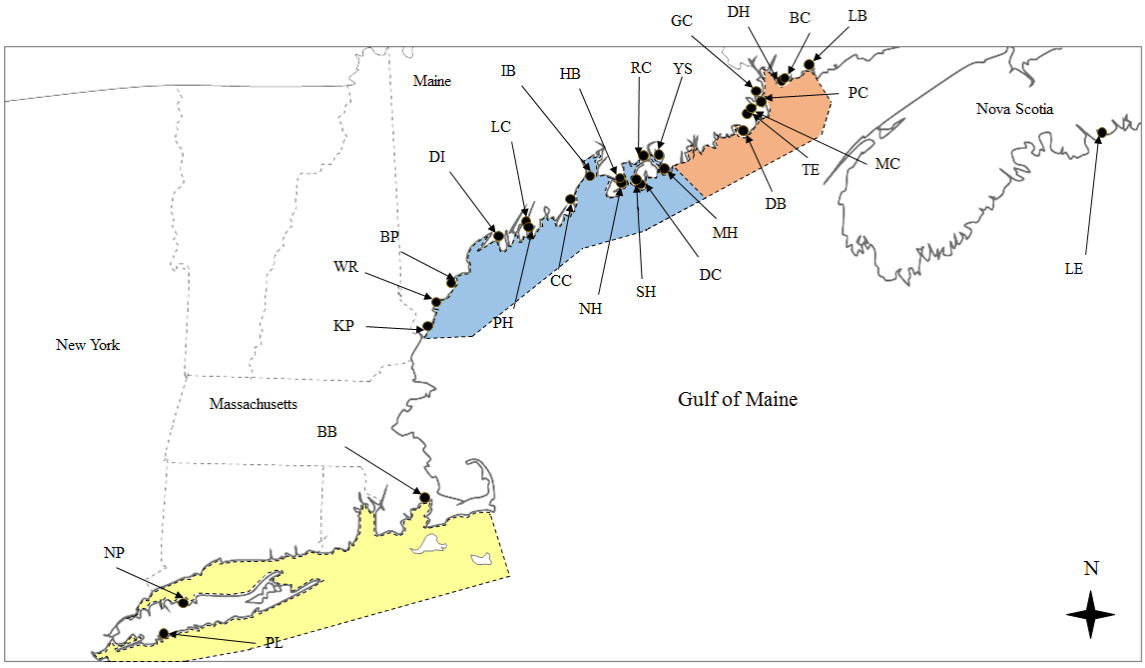
al. 2010). Based on this knowledge I predicted there would be a lower proportion of resistant individuals in the PB region. The results of this work, however, found that the PB region contained unexpectedly high percentages of clams displaying the RR and RS genotypes. During the summer months when *Alexandrium* spp. typically blooms, the convergent boundary between the EMCC and the WMCC near the mouth of Penobscot Bay causes the EMCC waters to turn offshore, resulting in a net offshore transport of water and effectively acting as a barrier that prevents *Alexandrium* spp. cells from entering the bay (Pettigrew *et al.* 2005). Given the rarity of HABs in the PB region, the neuromuscular costs associated with the R allele should make resistance disadvantageous for clams in this area, yet 78.0% of *M. arenaria* sampled in this region possessed at least one R allele.

It is possible that R allele is being maintained in *M. arenaria* populations in the PB region either by larval transport from an area where the R allele is common in the population, or by human seeding of juvenile hatchery clams that occurs in this region. If this theoretical introduction of the R allele was slowed or impeded it could result in a gradual change toward a more predominantly sensitive genotype in the population due to the costs associated with the R allele, and explain the genotype shifts that were observed at the Naskeag Harbor site over the duration of this study. It is also possible that interannual variation in the strength of the selective pressure exerted by *Alexandrium* spp. blooms affected genotype frequencies at Naskeag Harbor. Lastly, it is possible that predation by green crabs could play a role in genotype variations. While there is little localized data on green crab spatial distribution and population abundance in the Gulf of Maine, an overall trend of increasing numbers of these crabs has been observed in recent years (Tan and Beal 2015). The burrowing and growth rate effects associated with the resistant allele could

make RR individuals more vulnerable to predation, and cause a decrease in the prevalence of that genotype. Further work is needed to fully explore these possibilities and explain the anomalous presence of these RR genotype clams in the PB region, and the genotype shifts that occurred at Naskeag Harbor.

Results of the tests of independence performed among the genotype proportions of the five regions suggest that those predetermined regions could be redrawn into three, possibly four, discrete sections reflecting the distribution of genotypes. The first region encompasses EGoM, and is characterized by a very high proportion of RR genotype individuals. The second region encompasses MDI, PB, and WGoM, and is characterized by a more even distribution of the three genotypes. The third region encompasses Long Island, and possibly the south shore of Cape Cod, and is characterized by a majority of SS genotype individuals and a lack of RR genotype individuals (Figure 2.7). While the results of statistical testing within the SNE-NY region indicated that the Buttermilk Bay site was significantly heterogeneous, without more data it is difficult to say whether this is an indication that that particular site should belong to its own separate region. Harmful algal blooms rarely occur at this location, and those that do occur are only weakly toxic, resembling HABs that form near Long Island. Sampling more sites in this region would help to determine whether Cape Cod and Long Island are two distinct regions with regard to the resistant NaV mutation, or whether Buttermilk Bay is an anomalous site within an otherwise homogenous region (in the same manner as the Morong Cove site in the EGoM region).

Figure 2.7 New sample site groupings



Newly defined regions based on statistical testing for independence among the five originally defined regions of this study. The EGoM region is shown in red, the MDI, PB, and WGoM regions have been regrouped as one homogeneous region shown in blue, and the SNE – NY region is shown in yellow. (Site abbreviations are defined in Table 2.3).

The heterogeneity of the Morong Cove site in the EGoM region may be attributable to the patchy distribution of PSP events within Cobscook Bay where this site is found. Locations within Cobscook Bay can exhibit large differences in shellfish tissue toxicity levels, even between sites that are geographically close to one another. Some areas of Cobscook Bay remain free of harmful levels of PSP, even when other areas of the Bay are toxified (Shumway *et al.* 1988). A less severe history of PSP at the Morong Cove site resulting in lower selective pressure on these clams could explain the anomaly in genotype frequencies that was found at that location.

Analyzing the information from this study with a regression model could help identify how factors such as PSP history, sea surface temperature, predation pressure, and history of hatchery plantings affect genotype frequencies at a given site, however comprehensive data on several of these factors are not currently available. Shellfish toxicity data is one element that is available for many of the sample sites in this work, since it is sampled and recorded by the Maine Department of Marine Resources at sites throughout Maine during regular intervals each year. These data provide critical information for the Shellfish Sanitation Program, and represent an extensive source of PSP history that can be incorporated into a model. Hatchery plantings are another element for which there is a good source of data, since the DEI keeps records of towns that purchase seed clams from its hatchery. Precise planting sites within each town can be ascertained by contacting shellfish managers for those locations. Sea surface temperature is an element for which more records may be needed. Satellites and moorings provide wide-ranging data, but these approaches do not produce information on a fine enough scale to be able to ascertain precise temperature data for the individual sample sites in this study. Information regarding predation pressure is also incomplete. Green crabs are the primary predators of softshell clams in most regions of the Gulf of Maine, and while the expansion of these crabs' range into Maine in the second half of the twentieth century was well-documented (Scattergood 1952; Glude 1955; Roman 2006), precise estimates of population numbers and spatial distribution of these predators remains unknown. If more complete data become available in the future, a regression model approach could be a valuable addition to this work, but without further information constructing such a model remains unfeasible.

The results of this study have implications for locations that use seeded juvenile clams to attempt to bolster their declining softshell clam stocks. The high degree of genetic structure among softshell clams at the NaV mutation locus means that care should be taken to avoid planting seed clams into locations where their genetic makeup is poorly suited for that habitat, which could hinder those efforts. For example, if SS genotype clams were seeded in the EGoM those individuals would endure significant detrimental effects from the powerful HABs in that region, and likely suffer high mortality. Conversely, if RR clams were seeded in SNE-NY, the effects of the neuromuscular costs of resistance could possibly result in those individuals being outcompeted by locally adapted individuals. While genetically tailoring seed clams to fit localized environments is beyond the scope of current hatchery seeding efforts, selecting hatchery broodstock from the region that the seed clams are intended for, when it is possible to do so, could be a feasible strategy to ensure that the juvenile clams will be genetically suited for the environment they will be planted into in terms of the NaV resistant mutation.

These results also have implications for PSP monitoring in the Gulf of Maine. Since RR *M. arenaria* are unaffected by the paralytic effects of *Alexandrium* spp. ingestion, they are able to continue feeding during a HAB long past the point that SS genotype clams will become incapacitated and cease to feed. As a result RR clams are able to attain significantly higher tissue toxicities than their SS counterparts. In experiments conducted during a toxin accumulation study, SS genotype clams were able to attain tissue toxicities of 1700 to 3300 μg STX eq 100 g^{-1} before succumbing to the disabling effects of PSTs, while RR clams were able to attain tissue toxicities as high as 18000 to 20000 μg STX eq 100 g^{-1} after two weeks of exposure to *Alexandrium* sp. (MacQuarrie and Bricelj 2008).

Higher tissue toxicities in RR clams mean that these individuals present an increased risk for severe PSP, and therefore regions identified as having high numbers of RR clams should be closely monitored for PSP for the duration of the year when *Alexandrium* spp. blooms are possible in order to ensure public health and safety.

The spatial distribution of RR, RS, and SS clams in the Gulf of Maine and southern New England is the result of the interplay between the benefits of resistance to PSTs conferred to clams possessing the R allele, and the resulting neuromuscular costs that are associated with that allele. This evolutionary trade-off, juxtaposed on top of the environmental heterogeneity of the Gulf of Maine in terms of the prevalence and intensity of HABs causes the genetic mosaic that this study observed with respect to the R allele.

CHAPTER 3
NO EVIDENCE OF CHANGE IN GENOTYPE FREQUENCIES FOR SAXITOXIN RESISTANCE IN A POPULATION OF HATCHERY BROODSTOCK THAT HAVE UNDERGONE GROWTH SELECTION

3.1. Introduction

Some softshell clams are known to possess a mutation in DII of the α-subunit of their voltage-gated sodium channels (NaV) that grants them a resistance to the harmful effects of STX produced by the red tide dinoflagellates *Alexandrium* spp. This mutation causes a reduction in binding affinity of the STX molecule in the sodium channel pore, and prevents the onset of paralysis that is typically associated with STX ingestion (Bricelj *et al.* 2005). Three different genotypes for STX resistance have been identified at the NaV mutation locus: a STX-sensitive genotype (SS), a saxitoxin-resistant genotype (RR), and a heterozygous genotype (RS) that displays STX resistance that is intermediate between the SS and RR genotypes (Bricelj *et al.* 2005) (see Table 2.1). When the spatial distribution of these genotypes was surveyed in wild *M. arenaria* populations in the Gulf of Maine and southern New England, it was found that regions with a strong PSP history were predominantly comprised of clams with the RR and RS genotypes, and that regions with weak or sporadic histories of PSP were usually comprised of the SS genotype.

The resistant allele appears to come with associated costs for the clams that carry this mutation however, and these costs manifest in the form of slowed rates of burrowing and feeding (Walter unpublished; Bricelj *et al.* unpublished). During field experiments

where RR and SS genotype *M. arenaria* were planted into the intertidal zone at a site with no history of PSP, the SS genotype clams demonstrated higher shell growth rates than the RR genotype clams (Bricelj *et al.* unpublished). Growth differences between the RR and SS genotypes were also assessed in a laboratory setting, and after being fed a diet of non-toxic algae for 16 days, the SS clams had achieved higher shell and tissue growth rates than the RR clams (Bricelj *et al.* unpublished). Both the field plantings and the laboratory experiments used clams from different source populations however, so population-specific effects on growth rate cannot be ruled out in these data. Burrowing and growth rate data have not been examined in the RS genotype.

Hatcheries that produce aquacultured shellfish will often undertake selective breeding efforts to try to enhance desirable traits such as disease resistance or rapid growth rate in the offspring of their broodstocks. Selective breeding to enhance growth rate of offspring is a common strategy, since this can increase production and economic yield by producing animals that achieve legal harvesting size in a shorter timespan than would normally be required. In addition, some bivalves become less vulnerable to disease as they grow to larger sizes, as seen in the case of oysters *Crassostrea virginica* and juvenile oyster disease (Davis and Barber 1999).

The Downeast Institute for Applied Marine Research and Education (DEI) located in Beals, Maine, produces hatchery-reared juveniles of several commercially important bivalve species, including softshell clams that can be purchased by towns seeking to enhance their local wild stocks. Participation in the DEI's softshell clam seeding program currently encompasses over 50 towns in three different states. The DEI conducted experimental selective breeding to enhance the growth rate of a line of *M. arenaria*

hatchery broodstock by culling the slowest growing individuals from each successive generation of this line, leaving only the fastest growing individuals to contribute to the following generation. After two generations of selection for rapid growth this had resulted in a broodstock line that produced offspring with a significantly higher growth rate than offspring produced by wild individuals. In a field planting experiment conducted by the DEI comparing growth rates between the F1 generation and the F3 generation of this broodstock line, F3 clams demonstrated a mean shell growth rate that was 35.4% greater than the mean growth rate of the F1 clams ($p < 0.0001$), and an absolute growth rate that was 62% greater than that of the F1 clams (Brian Beal, pers. comm.).

This experiment presents an opportunity to investigate the effect of selection for growth rate on the distribution of genotypes for saxitoxin resistance from the F1 generation to the resulting F3 generation. Given the existing information on differences in growth rate between the RR and SS genotypes, I hypothesize that culling the slowest growing individuals from each generation will have resulted in a decrease in the frequency of the RR genotype through the successive generations from F1 to F3. This work could serve to complement existing data on differential growth rates based on genotype for STX resistance, and has the advantage of being conducted within the same population of animals.

3.2. Methods

Development of the *M. arenaria* broodstock line began in 2002 when the DEI collected wild adult clams from eastern Maine. These clams were taken to the DEI's

hatchery and induced to spawn through a thermal shock technique (Stickney 1964) to produce the F1 generation. F1 larvae were reared in the hatchery until they reached the juvenile stage (see Figure 1.3), at which point they were transferred to a field-based nursery for the summer and fall of 2002. At the end of the fall, F1 clams were returned to the hatchery for overwintering. Then in April 2003 these clams were planted into protected intertidal plots in the town of Beals, ME. Approximately two years later in June of 2005 the planted F1 clams were retrieved from those plots. The 30 largest F1 individuals were taken to the DEI hatchery where they were induced to spawn, and the resulting F2 clams were reared and overwintered in the same fashion as the F1 generation. In April 2006 the F2 clams were planted into protected intertidal plots in Beals, and remained there until they were retrieved in June 2008. The 30 largest individuals from the F2 generation were spawned in the DEI hatchery, producing the F3 generation that was reared in the same fashion as the previous generations.

The individuals that comprised the original F0 generation were not available to obtain DNA samples from. There were also no individuals from the F2 generation available to analyze. Juveniles from the F1 and F3 generations were frozen whole at -20°C. These individuals were transported to the Connell Laboratory at the University of Maine, where a small tissue sample was excised from each clam, and DNA was extracted from that tissue using a DNeasy Plant Mini Kit (QIAGEN**).** The extracted tissue DNA was stored at -20°C.

DNA samples were also obtained from adult F3 individuals at the DEI by using a 1 mL Monoject™ Tuberculin syringe (Covidien) to extract 0.3 mL of hemolymph from the anterior adductor muscle of each animal. The hemolymph was transferred to

individually numbered sterile tubes and placed on ice for transport to the Connell Laboratory. The tubes of hemolymph were centrifuged (Eppendorf® Microcentrifuge 5418) at 10,000 rpm for 5 minutes to produce a cell pellet in each tube. The supernatant was pipetted off of each sample, and the cell pellet was resuspended in 50uL of a buffering solution of 10mM Tris-HCl, pH 8.5. The hemolymph was stored at -20°C.

The hemolymph and tissue DNA samples were used in PCR reactions designed to amplify a 172 bp segment of the sodium channel gene containing the NaV mutation site. Amplification was carried out in 25 μL total volume in illustra PuReTaq Ready-To-Go PCR Bead tubes (GE Healthcare Life Sciences) containing 22 μL of HyClone™ HyPure Molecular Biology Grade Water (GE Healthcare Cell Culture), 0.5 μL of each primer (oBTG-99F and oBTG-100R), and 2 μL of hemolymph (~50 ng of DNA) or DNA extract. A no-DNA control was also prepared with each set of PCR reactions. Cycling was done in a MJ Research PTC-200 Thermal Cycler (GMI Laboratory Instruments) and cycling conditions were set at (i) 95°C for 5 minutes, followed by 35 cycles of (ii) 95°C for 30s, (iii) 50°C for 30s, and (iv) 72°C for 120s, and finishing with (v) 72°C for 10 minutes.

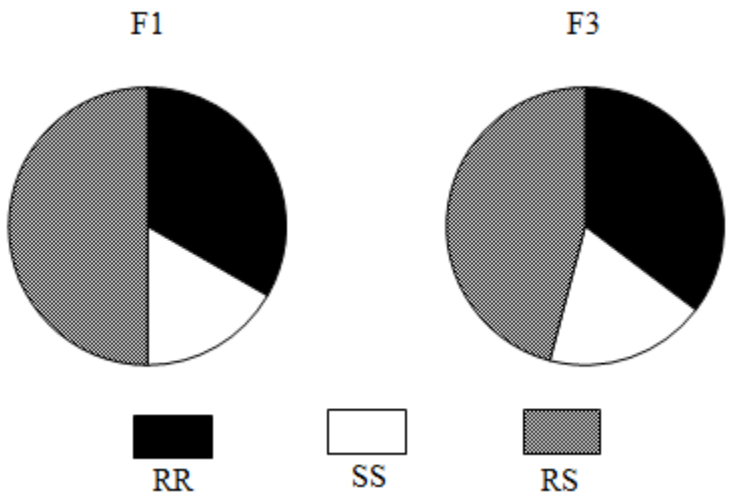
The resulting PCR products and no-DNA controls were visualized on an I.D.$_{N.A.}$™ agarose gel to ensure that each sample had amplified successfully and that there was no contamination from non-sample DNA. The PCR products were purified using the Wizard© SV Gel and PCR Clean-Up System (Promega), and quantified using a NanoDrop© ND-1000 Spectrophotometer (Thermo Scientific). The purified PCR products were prepared for DNA sequencing in accordance with the specifications of the University of Maine DNA Sequencing Facility, where they were sequenced using ABI BigDye™ Terminator chemistry. The raw data for each generation were analyzed using

Sequencher® 5.0.1 software (Gene Codes Corporation), and the genotype of each individual was identified. G-tests of independence were used to determine whether any genotype changes that were observed across generations were statistically significant.

3.3. Results

The results of the analysis demonstrated negligible change in genotype frequencies for saxitoxin resistance over the course of the development of this broodstock line. The F1 generation (n = 18) had genotype frequencies of 33.3% RR, 50.0% RS and 16.7% SS. The F3 generation (n = 67) contained a similar distribution of the three genotypes, at 35.3% RR, 45.9% RS, and 18.8% SS (Figure 3.1). Independence testing determined that there was no significant change in the frequency of any of the three genotypes between the F1 and F3 generation ($p = 0.92$).

Figure 3.1 Genotype frequencies in DEI F1 and F3 broodstock



Genotype frequencies with respect to STX resistance in clams from the F1 generation of DEI broodstock, and the F3 generation of DEI broodstock. RR indicates the STX-resistant genotype, SS indicates the STX-sensitive genotype, and RS indicates the heterozygous genotype.

3.4. Discussion

The theory of a cost of resistance associated with the NaV gene has been supported by the results of several previous studies, but all of these experiments used resistant and sensitive clams from different source populations, meaning that population- specific effects on growth rate could not be definitively ruled out as contributing to the observed differences between the RR and SS genotypes. This study examined individuals within a single line of broodstock, removing the possibility of confounding population effects.

This work examined the relationship between genotype and growth rate for the first time within a single population of clams. The results of this work did not provide any supporting evidence for the hypothesis of differences in growth rate among the three genotypes however, since selection for the trait of rapid growth resulted in no statistically significant change in genotype frequencies between the F1 and F3 generations despite significant differences in growth rate between these two generations.

Since these generations of clams were reared in field plots over periods of several years, the effect of selection due to HABs is a factor that must be considered when looking at genotype frequencies in this work. The field planting sites in Beals, ME are located in a region where shellfish tissue toxicity has historically been low (< 100 μg STX eq 100 g^{-1} in *Mytilus edulis*) (Thomas *et al.* 2010), so any selective pressure that these clams may have been exposed to would have been weak, and unlikely to affect the genotype frequencies of these planted individuals.

Given that the incidence of PSP in the region around Beals where the original F0 generation was collected from is generally low, this starting generation probably began with a fairly even distribution of the three genotypes. Sites from Chapter 2 with PSP histories that are similar to the Beals site typically have genotype frequencies are relatively uniformly distributed between RR, RS, and SS. If the F0 generation already contained moderate numbers of SS and RS genotype clams at the start of the experiment, then the effect of selection for growth rate may not have influenced genotype frequencies as dramatically as if the starting generation had been primarily comprised of the RR genotype. Repeating this experiment using an F0 generation with a large number of RR clams may result in a more significant change in genotype frequencies.

CHAPTER 4
MICROSATELLITE MARKERS SHOW LITTLE VARIATION AMONG *MYA ARENARIA* IN THE GULF OF MAINE

4.1. Introduction

Identifying genetic structure within populations of organisms can have a number of valuable applications, including for management of species, as well as for the development of marine reserves and protected areas. Understanding patterns of gene flow can help to identify populations that provide important sources of recruits to other populations as well as detect populations that are major sinks of these recruits. This can aid management entities in pinpointing populations that are of vital importance for conservation in at-risk species. Characterizing the strength of these genetic exchanges can also provide valuable information. Since genetic homogenization between two populations can be accomplished by a relatively low level of individual exchange (Hellberg 2009), genetic similarity between two populations does not necessarily imply a high rate of exchange of individuals. Identifying locations where exchange becomes inconsequential for the maintenance of population numbers and genetic structure can provide essential information about where boundaries of reserves and protected areas should be placed to safeguard targeted populations.

Knowledge of genetic population structure can also provide important information about organisms that are grown in aquaculture facilities. Such information can help hatcheries maintain healthy and productive broodstocks, and can also inform selective breeding efforts to produce offspring that display desirable traits such as disease resistance or rapid growth rate (Reece *et al.* 2004). Genetically informed selective breeding can also

be used to avoid or mitigate common problems that arise in hatchery stocks, namely the loss of genetic variability and reduction of effective population size. (Reece *et al.* 2004; Wang *et al.* 2014). In the marine environment, organisms with planktonic life stages have historically been assumed to have low levels of genetic variation across populations because of their wide capacity for dispersal and connectivity. However, more recent data demonstrate that this assumption does not always hold true. Mounting evidence indicates that planktonic larvae may disperse on smaller geographic scales than their potential would suggest (Hellberg 2009), and despite having great potential for dispersal larval self-recruitment plays an important role in maintaining many populations. Self-recruitment can offer the advantage of a high likelihood of encountering favorable habitat for recruitment and growth, and the possibility of benefiting from localized adaptations (Almany *et al.* 2007). Spatial environmental heterogeneity arising from differences in temperature, salinity, and other factors can favor such localized adaptations (Strasser and Barber 2009) resulting in population genetic structure on small scales, and it appears that such fine scale structure is much more common than was once thought (Saenz-Agudelo *et al.* 2009).

Genetic variation can arise out of significant differences between larval cohorts and their parent populations, a phenomenon known as chaotic genetic patchiness (Johnson and Black 1982). These differences between parents and offspring can result from factors such as the larvae being subjected to selection pressures while in the plankton stage (Johnson and Black 1982), or from differential reproductive success among the adults leading to only a few individuals contributing to the next generation (Hedgecock 1994). This "sweepstakes reproductive success" means that effective population sizes can be substantially smaller than actual population sizes (Hedgecock *et al.* 1982), and sweepstakes

reproduction has been confirmed to occur in a number of marine species, including bivalves such as Pacific oysters *Crassostrea gigas* (Li and Hedgecock 1998), European flat oysters *Ostrea edulis* (Hedgecock *et al.* 2007), and softshell clams *Mya arenaria* (St-Onge *et al.* 2015). A reduction in effective population size means that genetic drift, which has negligible effects when effective population size is large, can play a much larger role in generating variation in alleles frequencies than would be expected in the population genetic structure of organisms that experience sweepstakes effects (Hedgecock *et al.* 2007). Sweepstakes effects produce larval cohorts that exhibit reduced genetic variation compared to their parent population, stemming from the interaction between varying reproductive success and the unpredictable environmental changes that occur in the marine setting (Waples 1998). Spatial environmental heterogeneity means that sweepstakes effects may impact larval cohorts in one geographic region, while having negligible influence in another (Taris *et al.* 2009).

Intraspecific differentiation can follow geographical boundaries (past or present), and be delineated by spatially discrete habitats with their own locally adapted genotypes. Examples of such boundaries in the marine environment include salinity or temperature gradients, upwelling regions, strong current boundaries and historical changes in sea level (Hilbish 1985; Crandall *et al.* 2008). When these boundaries impede larval transport and gene flow between populations, intraspecific genetic divergence can develop that follows the boundaries of those barriers. Many marine species require specific habitat types, either as nurseries for their larvae, or for settlement as adults, and a fragmented distribution of suitable habitat can pose barriers to dispersal and connectivity, allowing localized variation to develop.

There are numerous methodologies using different types of markers for analyzing population genetic structure, with varying suitability for different scenarios. These markers include allozymes (Prakash *et al.* 1969), amplified fragment length polymorphisms (AFLPs) (Vos *et al.* 1995), and mitochondrial DNA markers (Avise *et al.* 1979a), among others, but some of the most commonly used tools for identifying population structure and describing gene flow are microsatellite markers (Hellberg 2009). Microsatellites, also known as simple sequence repeats (SSRs) or short tandem repeats (STRs) are small repeating motifs within the genome that mutate readily and rapidly, with individual locus mutation rates that range between 10^{-2} and 10^{-6} per generation (Dallas 1992). These sections of DNA are prone to slippage and proofreading errors during replication that result in the gain or loss of a repeat unit in the sequence, leading to alleles of variable lengths at these loci that can be distinguished from one another (Selkoe and Toonen 2006). The high mutation rates of microsatellites can give rise to remarkable genetic diversity, and this high degree of polymorphism provides the power to identify genetic structure on fine scales, making microsatellites particularly useful for population genetic studies.

Microsatellites are not without their drawbacks, however, and one that is particularly common in bivalves is the occurrence of null alleles. Null alleles are loci that fail to amplify during PCR, and can result in genotyping errors caused by heterozygotes mistakenly being classified as homozygotes. Microsatellite primers are designed to bind to the flanking regions of DNA surrounding microsatellite loci during PCR, and when polymorphisms occur in those flanking regions it can lead to failure of the primer to bind and lack of amplification. Less commonly, null alleles can be caused by PCR conditions that are not ideal for amplification of the loci of interest (Dakin and Avise 2004). While

null alleles occur in many organisms, they are an especially widespread phenomenon in bivalves (Hedgecock *et al.* 2004), and have led to difficulties in developing reliable microsatellite markers for species in this class (Brownlow *et al.* 2008). There are several strategies available to combat the difficulties posed by null alleles, although they are costly in terms of time and effort. One course of action is to identify the polymorphisms that are causing the problems with binding and amplification, and redesign the primers to account for those substitutions. Another approach that was illustrated by Lemer *et al.* (2011) is to use a set of multiple primers that each give different and complementary information concerning the loci of interest. This method was able to eliminate the effect of null alleles when applied to a population of the black-lipped pearl oyster *Pinctada margaritifera* after a single-primer approach failed to provide accurate data (Lemer *et al.* 2011). Other sources of error in microsatellite data include the existence of stutter peaks in chromatograms of the data that may be mistakenly identified as alleles, and large allele dropout, which occurs when alleles either fail to amplify during PCR, or they fall outside the expected size range of alleles at a particular locus and go undetected.

A number of descriptive values are typically calculated for each microsatellite locus and each sample population, including the number of alleles detected, the number of private alleles detected, the allelic richness, the expected and observed heterozygosities, and the coefficient of inbreeding. The number of alleles detected is simply a count of how many alleles are present at a specific locus or within a certain population. Private alleles are alleles that occur in only one population, and those alleles can be tallied as a function of locus or population. Allelic richness is a way of expressing the genetic diversity at a locus or within a population. Since differences in sample sizes can affect the results of this

calculation, allelic richness is typically assessed using rarefaction to offset biases caused by sample size. Expected and observed heterozygosity (H_E and H_O) are values that fall between 0 and 1 and reflect the degree of variability at a locus or within a population. Expected heterozygosity, also known as gene diversity, is the estimated probability that individuals will be heterozygous at a particular locus, while observed heterozygosity is the level of heterozygosity that is actually detected. When observed heterozygosity is lower than the expected heterozygosity it can be suggestive of phenomena such as inbreeding or population bottlenecks causing a loss of genetic variability. Inbreeding can be expressed as a coefficient of inbreeding (F_{IS}) that indicates the probability that alleles are derived from a common ancestor. As the value of F_{IS} increases, the degree of inbreeding becomes greater.

Microsatellite data are also typically assessed for linkage disequilibrium and adherence to the parameters of Hardy-Weinberg equilibrium. Testing for linkage disequilibrium is performed to ensure that alleles are associating randomly at all loci. If associations between alleles are found, those alleles are said to be linked. Hardy-Weinberg equilibrium models genotype and allele frequencies in a population, and asserts that these frequencies will remain in a state of equilibrium unless they are disturbed by outside forces, which can include mutations, selection, gene flow, genetic drift, and non-random mating. Departures from Hardy-Weinberg equilibrium can indicate the influence of one of these perturbing forces.

Microsatellite data can be analyzed using numerous strategies, each with their own advantages and constraints. One of the most commonly used methods for expressing the degree of variation between populations is the fixation index F_{ST}. In theory, F_{ST} is a value

between 0 and 1 that indicates the level to which two populations are genetically differentiated from each other, with 0 signifying panmixia and 1 signifying fixation for different alleles. F_{ST} can serve as a good "yardstick" measurement for making comparisons across populations and species, but it operates under several assumptions that are frequently violated in natural systems. F_{ST} assumes that migration between populations is constant and that migrants are exchanged equally, that effective population size is constant, that genetic drift and gene flow are at equilibrium, and that the markers used for analysis are not under selection (Whitlock and McCauley 1999). In addition, F_{ST} expresses the total variation across populations by allocating this value into variation within populations and variation between populations, meaning that as within-population variation increases, the amount of variation remaining to be attributed to between population variation decreases (Hellberg 2009). As a result, when within-population variation is high, the maximum possible value for between-population F_{ST} may be quite small, even if significant differentiation is present (Hellberg 2009). Despite these weaknesses, F_{ST} can still serve as a good tool for comparison provided that its limitations are taken into account.

Other strategies for examining gene flow among populations include assignment tests and parentage analyses. Assignment tests allocate individuals to the population or group that they are most likely to have originated by comparing each individual's genotype to the genotypes that are likely in the potential source groups based on allele frequencies in those groups. The Bayesian assignment method described by Rannala and Mountain (1997) typically performs with the highest accuracy when compared to other types of assignment tests (Saenz-Agudelo *et al.* 2009). Parentage analysis compares the genotype of each individual against a pool of potential parents, and tries to assign each individual to

a pair of parents using one of several strategies. These strategies consist of exclusion, categorical allocation, fractional allocation, full probability parentage analysis, parental reconstruction, and sibship (sibling) reconstruction (Jones *et al.* 2010). Factors such as whether it is possible to sample potential parents, whether known sibling relationships can be identified, and the size of the average family will indicate which method is most appropriate for a particular set of circumstances. Some methods such as exclusion are vulnerable to the presence of null alleles, and should be used with caution if null alleles may be present (Jones *et al.* 2010).

An estimate of the level of gene flow among populations can indicate which of the above-mentioned strategies will perform better under a particular scenario. In a case of low gene flow, assignment tests tend to provide the most accurate picture of population genetic structure, whereas parentage analysis has a tendency to over-estimate self-recruitment (Saenz-Agudelo *et al.* 2009). In a case of high gene flow the reverse is true, with parentage analysis being able to detect small deviations from panmixia whereas assignment tests are often unable to assign many individuals to any one specific group (Saenz-Agudelo *et al.* 2009). The power of parentage analysis to detect variation depends on the percentage of the population that is sampled however, and the inability to sample a sufficient number of organisms can diminish the accuracy of this strategy.

Softshell clams *M. arenaria* are sessile as adults, but have a planktonic larval stage that can last for up to six week (Caporale *et al.* 1997). The opportunity this presents for dispersal, combined with large effective population sizes in these broadcast spawners, has led to the belief that softshell clams are genetically homogenous across their native range. This idea has been supported by evidence from several studies. Research conducted by

Caporale *et al.* (1997) using ITS-1 ribosomal DNA to look for signs of genetic structure among softshell clams in New England found no significant evidence of heterogeneity, and proposed that high levels of gene flow between populations have resulted in genetic homogeneity among clams in this region. Strasser and Barber (2009) sequenced mitochondrial cytochrome oxidase I (COI) of clams sampled from the northwest Atlantic and came to a similar conclusion, stating that softshell clams displayed very low genetic variability across geographic subregions of the northwest Atlantic, likely as a result of recent recolonization following the last glacial period.

While these results support a hypothesis of panmixia, or near panmixia among softshell clams in the Gulf of Maine, the markers used in the above studies did not have the level of power that microsatellites have to resolve fine-scale population genetic structure. In particular, the ITS-1 ribosomal DNA method used by Caporale *et al.* (1997) is more often used to distinguish separation at the species level. There are numerous examples of microsatellites used to resolve such fine-scale genetic structure where previous methods had detected no appreciable variation. A study examining the HAB species *Alexandrium catanella* described previously unknown intraspecific diversity after analyzing populations of *A. catanella* using microsatellites (Masseret *et al.* 2009). Previous studies using ribosomal markers had only been able to detect distinct geographic clades (Adachi *et al.* 1996), but did not resolve the genetic diversity within those geographic clades that was detected using the microsatellite markers (Masseret *et al.* 2009). In the Suminoe oyster, *Crassostrea ariakensis*, previous analysis using restriction fragment length polymorphisms (RFLPs) found no genetic structure among these oysters in Asia. Microsatellite analysis, however, was able to characterize eight well-defined populations that fit an isolation-by-

distance model (Xiao *et al.* 2010), meaning that dispersal among these populations is constrained by geographic distance.

Mya arenaria displays distinct population genetic structure at the NaV mutation described as in Chapter 2, showing statistically significant differences in the frequency of genotypes for STX resistance among geographic regions in the Gulf of Maine and southern New England. The eastern Gulf of Maine is dominated by clams possessing the RR genotype, while the western Gulf of Maine displays a more mixed distribution of genotypes, and southern New England is almost entirely comprised of clams possessing the SS genotype. These patterns arise from selection pressure exerted on these populations by *Alexandrium* spp. blooms, but it is unknown whether these patterns are also present on a genetic level at markers not under selection. This work seeks to ascertain whether the population genetic structure that exists around the NaV mutation is mirrored at neutral microsatellite loci in populations from the Gulf of Maine. The results of this work can aid in estimating the level of selection on the NaV gene itself, which would complement existing data on selection based on the different genotypes for STX resistance.

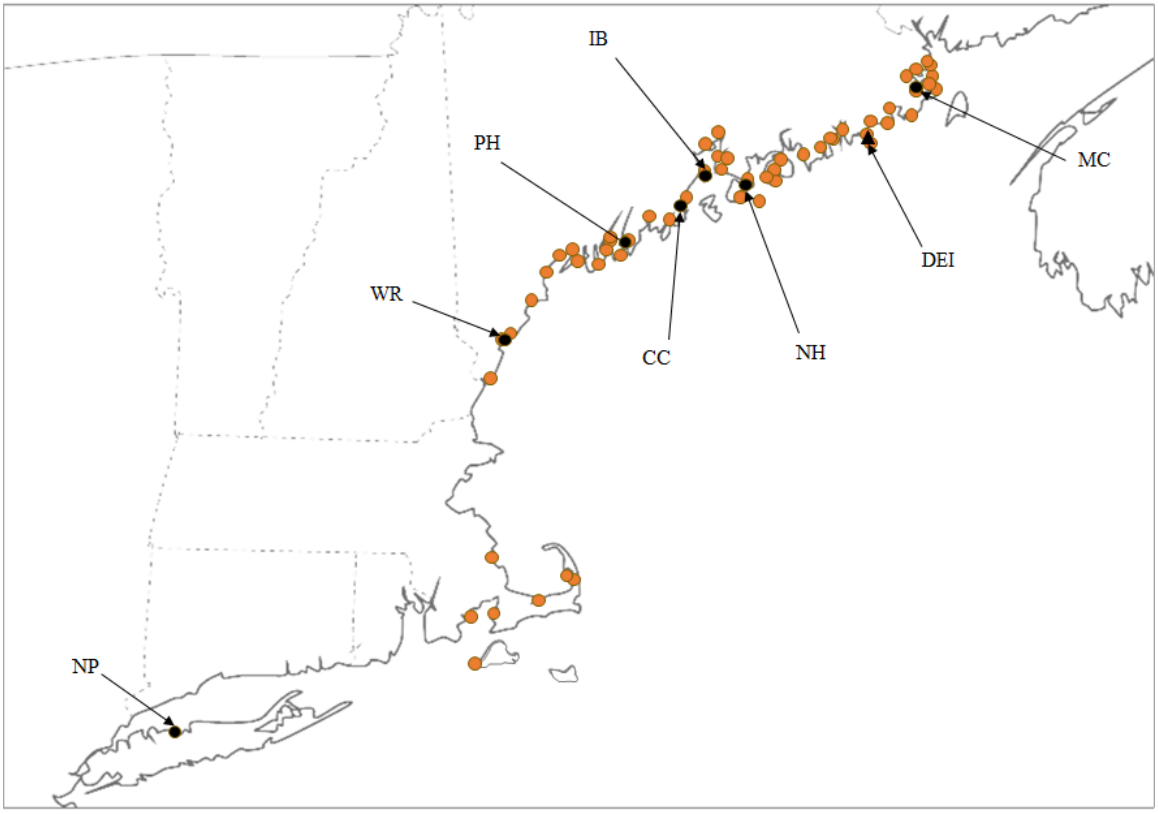
This work also examines the effect of a *M. arenaria* seeding program operating within the Gulf of Maine and southern New England. Seed clams produced by the DEI are purchased by municipalities throughout these areas for transplanting into their local stocks. Samples from the DEI hatchery broodstock will be incorporated into the microsatellite analysis to try to determine whether ubiquitous planting of seed clams throughout wild *M. arenaria* populations in the Gulf of Maine has influenced genotype frequencies in heavily seeded regions.

Analyzing the population genetic structure of *M. arenaria* can potentially provide important information for managing this fishery, particularly when that management strategy involves the introduction of seeded hatchery clams. Identifying any cryptic population subdivisions can help prevent the seeding of juveniles that are genetically diverged from the populations being seeded. Such a mismatch can cause disruption of the locally adapted genotypes if there is high survivorship among the seeded clams, or conversely, if the seeded clams are poorly suited for the localized environment, it can result in large-scale mortality of those individuals.

4.2. Methods

Hemolymph samples and extracted DNA samples from *M. arenaria* at seven of the sites included in Chapter 2 were used in this analysis (Figure 4.1) (Table 4.1). Six sites were chosen to span the area of the Gulf of Maine from east to west, and one site was chosen from Long Island, New York, to be an outgroup. Hemolymph that was collected from a line of broodstock clams at the DEI in Chapter 3 was also used in this analysis. The hemolymph and tissue DNA were used in PCR reactions designed to amplify eight different microsatellite markers for *M. arenaria* (St-Onge *et al.* 2011) (Table 4.2). Thirty of the same individuals that were used in the Chapter 2 analyses were used from each population, and amplification of all eight microsatellite markers was performed for each individual.

Figure 4.1 Sample populations for microsatellite analysis and municipalities receiving seed clams from the Downeast Institute



Labeled locations indicate populations used for microsatellite analysis. Six sites were picked to span the Gulf of Maine from east to west, one site from Long Island, New York was included as an outgroup, and broodstock clams from the DEI hatchery were also included in the analysis. (See Table 4.1 for population abbreviations.) Orange circles indicate municipalities throughout the Gulf of Maine and Cape Cod that purchased hatchery clams from the DEI seeding program between 2006 and 2012.

Table 4.1 Sample population details

Site name	Abbreviation	Latitude & longitude	Year sampled	Number of samples
Morong Cove	MC	44.85635 -67.10387	2010	30
Downeast Institute	DEI	44.48073 -67.59865	2011	30
Naskeag Harbor	NH	44.22959 -68.53542	2012	30
The Narrows	IB	44.31336 -68.90184	2015	30
Clam Cove	CC	44.13480 -69.09242	2015	30
Poorhouse Cove	PH	43.88363 -69.55011	2012	30
Webhannet River	WR	43.32035 -70.57177	2010	30
Northport Harbor	NP	40.60103 -73.35386	2013	30

Details of eight populations used for microsatellite analysis, including the population name, the population abbreviation, the location of each sample population (DEI broodstock were sampled at the hatchery), the year each population was sampled, and the number of samples per population.

Table 4.2 *Mya arenaria* microsatellite details

Locus	Forward & reverse primers (5' → 3')	Repeating Motif	Ta (°C)	GenBank accession
Mar1	F-AAGCCACGTCTCAAGCCTTA R-TATGCGTTCGTCCGTATGTG	$[(GC)_4(CA)_8GC(CA)_{12}]$	51	JN191327
Mar2	F-ATATGTGGGTAAATGGTTGGC R-TAATTTCCTGCTATTACTGAGG	$(TG)_{14}$	54	JN191328
Mar3	F-AGAAATGTAAGAGGAGATGC R-TCCGTAACATTTCACGTCCA	$(CA)_{12}$	51	JN191329
Mar4	F-TCAATGCCAAAACATTGGTTA R-ACCCCCAAGCCTTTACTAGC	$(AC)_{24}$	51	JN191330
Mar5	F-TTGGGTCAAAACGTTCACAA R-CATGGCCACTGGAAGTGTTA	$(GACA)_7$	56	JN191331
Mar6	F-CAATGCCCAACCCACTAAAC R-GCGATGTTGGTTGTGTTGAC	$(CA)_9$	51	JN191332
Mar7	F-TTGCAGGCGATGTTTGTATC R-ATACGGCATTCTTGGTCAGG	$(GT)_{23}$	56	JN191333
Mar8	F-CATGTGTGACGAAATGTTGATG R-CAGAGTCATACACTGTCCATTGC	$[(CA)_5CG(CA)_9]$	51	JN191334

Details of eight microsatellites markers for *M. arenaria* including primer sequences, the repeating motif within the genome, the annealing temperature (Ta) for each locus, and the GenBank accession number. (Adapted from St-Onge *et al.* 2011).

Amplifications were carried out in 15 μL total volume in illustra PuReTaq Ready-To-Go PCR Bead tubes (GE Healthcare Life Sciences) containing 9 μL of HyClone™ HyPure Molecular Biology Grade Water (GE Healthcare Cell Culture), 2.5 μL of Q-Solution from the Type-It Microsatellite PCR Kit (Qiagen), 0.5 μL each of forward and reverse primer for the marker being amplified, and 2.5 μL of hemolymph (~ 65 ng DNA). Cycling was done in a MJ Research PTC-200 Thermal Cycler (GMI Laboratory Instruments) and cycling conditions followed those specified in St-Onge *et al.* (2011).

The PCR products were purified using the Wizard© SV Gel and PCR Clean-Up System (Promega), and quantified using a NanoDrop© ND-1000 Spectrophotometer (Thermo Scientific). The PCR products were prepared for fragment analysis by being loaded into 96-well Half-Skirted PCR Microplates (Axygen), with each well containing between 0.5 and 5 μL of PCR product (to result in a final DNA concentration of 5-10 ng/μL), 0.5 μL of GeneScan™ 350 ROX™ dye Size Standard (Applied Biosystems), and enough Hi-Di™ Formamide to produce a total volume of 10 μL. Samples underwent size fragment length analysis at the University of Maine DNA Sequencing Facility.

The raw data were analyzed using GeneMarker® software (SoftGenetics®) to score the sample alleles and bin them into size classes. Each allele score was verified manually to ensure there were no scoring errors due to stutter peaks or artefacts. Once a multi-locus genotype had been compiled for each individual, the software MICROCHECKER 2.2.3 (Van Oosterhout *et al.* 2004) was used to examine the samples for null alleles, large allele dropout, and scoring errors due to stuttering.

FSTAT 2.9.3.2 (Goudet 2001) was used to evaluate the characteristics of each individual locus, including total number of samples, total number of alleles, mean number of alleles per site, number of private alleles, allelic richness, Nei's expected and observed heterozygosities (H_O and H_E), and Wright's coefficient of inbreeding (F_{IS}). The eight sample populations were also evaluated using FSTAT to identify individual population characteristics like the number of detected alleles, the mean within-loci allelic richness, the number of private alleles, Nei's expected and observed heterozygosities (H_O and H_E), and Wright's coefficient of inbreeding (F_{IS}). FSTAT was also used to determine whether the populations conformed to expected allele frequencies under Hardy-Weinberg equilibrium,

and whether there was significant linkage disequilibrium between any of the pairs of loci within and among populations. Pairwise values of F_{ST} were also assessed between all of the populations, without the assumption of Hardy-Weinberg equilibrium.

Assignment tests were conducted using GeneClass 2.0 (Piry *et al.* 2004) to assign each individual to the population that they were most likely to have originated from using a Bayesian method (Rannala and Mountain 1997) based on 10,000 simulated individuals and an algorithm using Monte-Carlo resampling (Paetkau *et al.* 2004) to calculate the probability of assigning each individual to one of the eight sample populations. The level of nominal significance was set at 5%. GeneClass was also used to assess the probability of first-generation migrants among all of the samples and populations, again using a Bayesian method (Rannala and Mountain 1997) based on 10,000 simulated individuals and an algorithm using Monte-Carlo resampling (Paetkau *et al.* 2004), and to estimate which of the populations they had most probably originated from. The level of nominal significance for this analysis was set at 5%.

Parentage analysis was also used to examine these data. Parentage analysis was carried out using COLONY 2.0.6.1 (Jones and Wang 2010) to search for full or half-sibling relationships among all of the sampled individuals using a maximum likelihood method. The analysis was conducted using COLONY's "long" run length, and the likelihood precision was set to "high." Individuals were grouped into family clusters based on the resulting data, and assignment to a specific cluster was based on maximum likelihood.

4.3. Results

PCR amplification was successful with most templates and primer sets, however the Mar2 locus amplified poorly in several of the sample populations and was dropped from further analyses. Analysis of the seven remaining microsatellite loci in MICROCHECKER found no evidence of scoring errors from stutter or large allele dropout, however it did reveal that six of the loci showed an excess of homozygote genotypes and a deficit of heterozygotes in at least one of the eight sample populations, possibly due to null alleles.

The number of individuals assessed at each of the seven loci ranged from 191 (Mar1) to 231 (Mar5), with an average of 215 samples per locus. The number of alleles at each locus ranged from 4 (Mar6) to 34 (Mar4) and averaged 21 alleles per locus. Mar6 showed the smallest mean number of alleles per site with 3.25, while Mar3 has the highest mean number of alleles per site with 17.5. All of the loci except for Mar6 had private alleles, ranging from 1 (Mar7) to 8 (Mar3). Allelic richness based on 15 diploid individuals ranged from 3.2 (Mar6) to 13.84 (Mar3), and averaged 10.01 across all loci. Observed heterozygosity H_O was lowest at Mar4 at 0.36, and highest at Mar3 at 0.83. The average H_O across all loci was 0.63. This was lower than the expected heterozygosity H_E that averaged 0.81 across all loci, and ranged from 0.53 at Mar6 to 0.92 at Mar3. The coefficient of inbreeding ranged from -0.001 (Mar6) to 0.556 (Mar4), and averaged 0.218 across all loci (Table 4.3).

Table 4.3 Microsatellite loci characteristics

	N	N_A	$_MN_A$ / site	N_{PA}	A	H_O	H_E	F_{IS}
Mar1	191	21	12.25	2	10.93	0.58	0.9	**0.351**
Mar3	206	31	17.5	8	13.84	0.83	0.92	**0.107**
Mar4	219	34	14.63	7	10.86	0.36	0.79	**0.556**
Mar5	231	16	10.13	4	8.4	0.67	0.78	**0.137**
Mar6	213	4	3.25	0	3.2	0.52	0.53	-0.001
Mar7	224	18	13.5	1	11.04	0.74	0.88	**0.162**
Mar8	219	24	14.25	3	11.77	0.69	0.87	**0.215**

Characteristics of the seven microsatellite loci used in this study including the number of samples N, number of alleles N_A, mean number of alleles per site $_MN_A$/site, number of private alleles N_{PA}, allelic richness A (based on 15 diploid individuals), observed heterozygosity H_O, expected heterozygosity H_E, and coefficient of inbreeding F_{IS} (significance was assessed through 500 randomizations in FSTAT and significant values are shown in bold).

The number of alleles detected in each population ranged from 71 (Downeast Institute) to 91 (Morong Cove), with an average of 85 alleles being detected among the eight populations. Mean within-loci allelic richness based on 15 diploid individuals was lowest in the Downeast Institute population at 8.35, and the highest allelic richness was seen in the Morong Cove population at 10.68. Private alleles were detected in all eight populations, and the number of private alleles ranged from four (Morong Cove and Poorhouse Cove) to one (Downeast Institute, The Narrows, and Clam Cove). Observed heterozygosity H_O among the eight sample populations ranged from 0.51 (Northport Harbor) to 0.64 (Downeast Institute), with an average observed heterozygosity of 0.57 across all populations. The observed heterozygosities in all of the populations were lower

than the expected heterozygosities H_E that ranged from 0.80 (Downeast Institute and Clam Cove) to 0.87 (Poorhouse Cove), with an average expected heterozygosity of 0.83 across all populations. Estimations of F_{IS} for each population ranged from 0.147 (Downeast Institute) to 0.310 (Northport Harbor) with an average F_{IS} of 0.228 among all of the populations (Table 4.4).

Table 4.4 Sample population characteristics

	N_A	A	N_{PA}	H_0	H_E	F_{IS}
Morong Cove	91	10.68	4	0.61	0.82	**0.160**
Downeast Institute	71	8.35	1	0.64	0.80	**0.147**
Naskeag Harbor	84	9.69	3	0.60	0.82	**0.213**
The Narrows	83	9.95	1	0.55	0.83	**0.257**
Clam Cove	83	9.72	1	0.55	0.80	**0.226**
Poorhouse Cove	89	10.63	4	0.55	0.87	**0.246**
Webhannet River	90	10.31	2	0.56	0.81	**0.261**
Northport Harbor	89	10.07	3	0.51	0.85	**0.310**

Characteristics of each of the eight sample populations including number of alleles detected N_A, mean within-loci allelic richness A (based on 15 diploid individuals), number of private alleles detected N_{PA}, Nei's observed heterozygosity H_O and expected heterozygosity H_E, and Wright's coefficient of inbreeding F_{IS} (significance was assessed through 1280 randomizations in FSTAT, and values that are significant based on the adjusted 5% nominal level are shown in bold).

Analyses conducted in FSTAT found that there were significant within-population deviations from the allele frequencies that would be expected under conditions of Hardy-Weinberg equilibrium as a result of a deficit of heterozygous genotypes. These

heterozygote deficiencies affected all eight populations. Testing for linkage disequilibrium between pairs of loci within each population and among all populations found that all there was no significant genotypic disequilibrium between any of the loci pairs.

Comparisons of pairwise F_{ST} values between all populations found significant genetic differentiation between the Downeast Institute population and all other sample populations with the exception of Morong Cove and Webhannet River, where p-values were close to the nominal level of significance but did not achieve it. Northport Harbor also showed significant levels of differentiation from two of the other populations, and the p-value for another comparison involving Northport Harbor was close to the nominal significance level. None of the other sample populations showed a significant level of genetic differentiation from the others (Table 4.5).

Table 4.5 Pairwise F_{ST} values and p-values between populations

	MC	DEI	NH	IB	CC	PH	WR	NP
MC		0.00357	0.13036	0.04464	0.025	0.03393	0.34107	**0.00179**
DEI	0.0181		**0.00179**	**0.00179**	**0.00179**	**0.00179**	0.00357	**0.00179**
NH	0.0034	0.0131		0.2625	0.14643	0.53214	0.43929	0.02143
IB	0.0198	0.0166	0.0084		0.12679	0.32679	0.1	0.24107
CC	0.0132	0.0293	0.0037	0.0152		0.04821	0.00357	0.00714
PH	0.0074	0.0166	0.0001	0.0083	0.0112		0.05714	0.1
WR	0.002	0.0172	0.0007	0.007	0.0125	0.0079		0.02857
NP	0.0303	0.0266	0.012	0.0129	0.0294	0.0023	0.0233	

Pairwise F_{ST} values among the eight sample populations are shown below the diagonal, and calculated p-values for each comparison are shown above the diagonal. The adjusted p-value for 5% nominal level is 0.00179, and significant p-values are shown in bold. Significance was assessed through 560 randomizations in FSTAT. (See Table 4.1 for population abbreviations.)

Results of the assignment test did not produce any meaningful patterns across any of the populations or samples. Most individuals were assigned either to Poorhouse Cove or to The Narrows. Across all individuals, Poorhouse Cove was assigned the most individuals with 81, and Downeast Institute was assigned the fewest individuals with four (Table 4.6). Only 20.4% (49 individuals) were correctly assigned to their source population. A chi-squared test of goodness-of-fit found that the results of the assignment test differed significantly from those that would be expected due to random chance ($p = 4.37 \times 10^{-32}$).

Table 4.6 Assignment test among populations

	MC	DEI	NH	IB	CC	PH	WR	NP
Morong Cove	10	-	1	2	1	10	4	1
Downeast Institute	5	4	-	3	-	12	4	2
Naskeag Harbor	3	-	-	6	1	14	6	-
The Narrows	2	-	1	9	2	9	4	3
Clam Cove	4	-	1	10	5	7	2	1
Poorhouse Cove	3	-	1	8	-	11	4	2
Webhannet River	6	-	3	6	1	7	4	3
Northport Harbor	1	-	-	10	1	11	2	5
All	34	4	7	54	11	81	30	18

The number of individuals from each population that were assigned to each of the eight sample populations based on probabilities generated by a Bayesian method (Rannala and Mountain 1997) using 10,000 simulated individuals and a Monte-Carlo resampling algorithm (Paetkau *et al.* 2004). The level of nominal significance was set at 5%. (MC = Morong Cove, DEI = Downeast Institute, NH = Naskeag Harbor, IB = The Narrows, CC = Clam Cove, PH = Poorhouse Cove, WR = Webhannet River, NP = Northport Harbor.)

Analysis of all samples for first generation migrants identified 57 individuals that were potential migrants to their population, however there was no clear pattern across all populations or individuals. The Downeast Institute population contained the most potential migrants with 12 individuals, and the Morong Cove population contained the fewest potential migrants with three individuals. Morong Cove migrants originated from Downeast Institute, Clam Cove, and Webhannet River. Downeast Institute migrants mostly originated from Naskeag Harbor and Webhannet River. Naskeag Harbor migrants mostly originated from Poorhouse Cove. The Narrows migrants mostly originated from Poorhouse Cove. Clam Cove migrants mostly originated from The Narrows. Poorhouse

Cove migrants originated from Naskeag Harbor, The Narrows, Webhannet River, and Northport Harbor. Webhannet River migrants mostly originated from Northport Harbor. Northport Harbor migrants mostly originated from The Narrows, Clam Cove, and Poorhouse Cove. Across all eight populations, Poorhouse Cove was the origin of the most migrants with 11, and Downeast Institute was the origin of the fewest migrants with three (Table 4.7). A chi-squared test of goodness-of-fit showed that these results did not differ significantly from those expected due to random chance ($p = 0.505$).

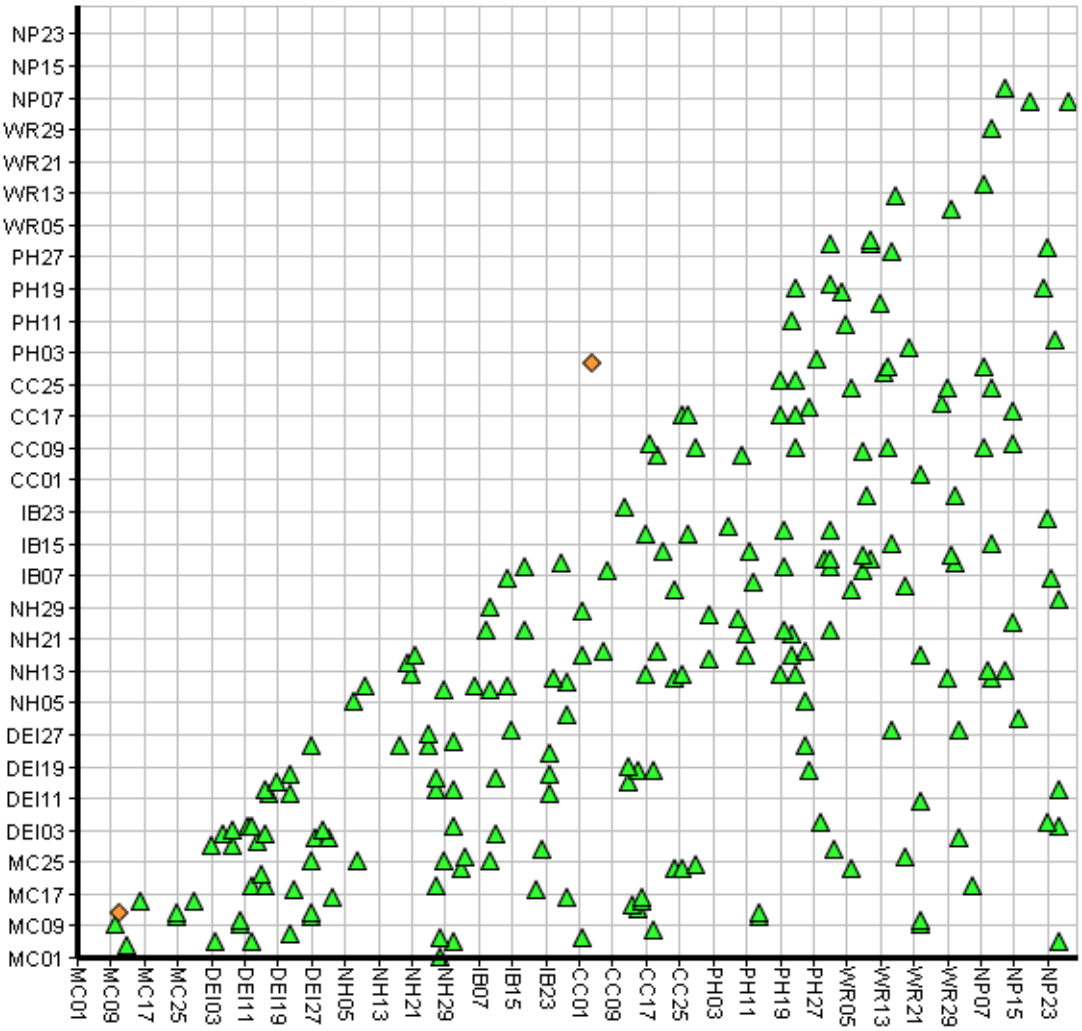
Table 4.7 First generation migrants among populations

	Source populations							
Number of migrants	MC	DEI	NH	IB	CC	PH	WR	NP
Morong Cove (n = 3)		1	-	-	1	-	1	-
Downeast Institute (n = 12)	2		3	-	-	2	3	2
Naskeag Harbor (n = 7)	1	-		1	2	3	-	-
The Narrows (n = 8)	1	1	-		1	3	2	-
Clam Cove (n = 6)	1	-	-	3		-	2	-
Poorhouse Cove (n = 4)	-	-	1	1	-		1	1
Webhannet River (n = 10)	1	1	2	-	2	1		3
Northport Harbor (n = 7)	-	-	-	2	2	2	1	
All (n = 57)	6	3	6	7	8	11	10	6

Individuals identified as first generation migrants in each of the eight sample populations, and the population that those migrants were mostly likely to have originated from. Probabilities were calculated using a Bayesian method (Rannala and Mountain 1997) with 10,000 simulated individuals and a Monte-Carlo resampling algorithm (Paetkau *et al.* 2004), and the level of nominal significance was set at 5%. (MC = Morong Cove, DEI = Downeast Institute, NH = Naskeag Harbor, IB = The Narrows, CC = Clam Cove, PH = Poorhouse Cove, WR = Webhannet River, NP = Northport Harbor.)

Tests for sibship among all individuals using parentage analysis methods also did not identify any meaningful patterns across any of the populations. COLONY identified two instances of full siblings in the data set, and 215 instances of half-siblings using maximum likelihood methods. Both sets of full siblings occurred between individuals within the same population, while the half-sibling relationships spanned individuals across all eight populations, and did not seem to follow any clear arrangement (Figure 4.2).

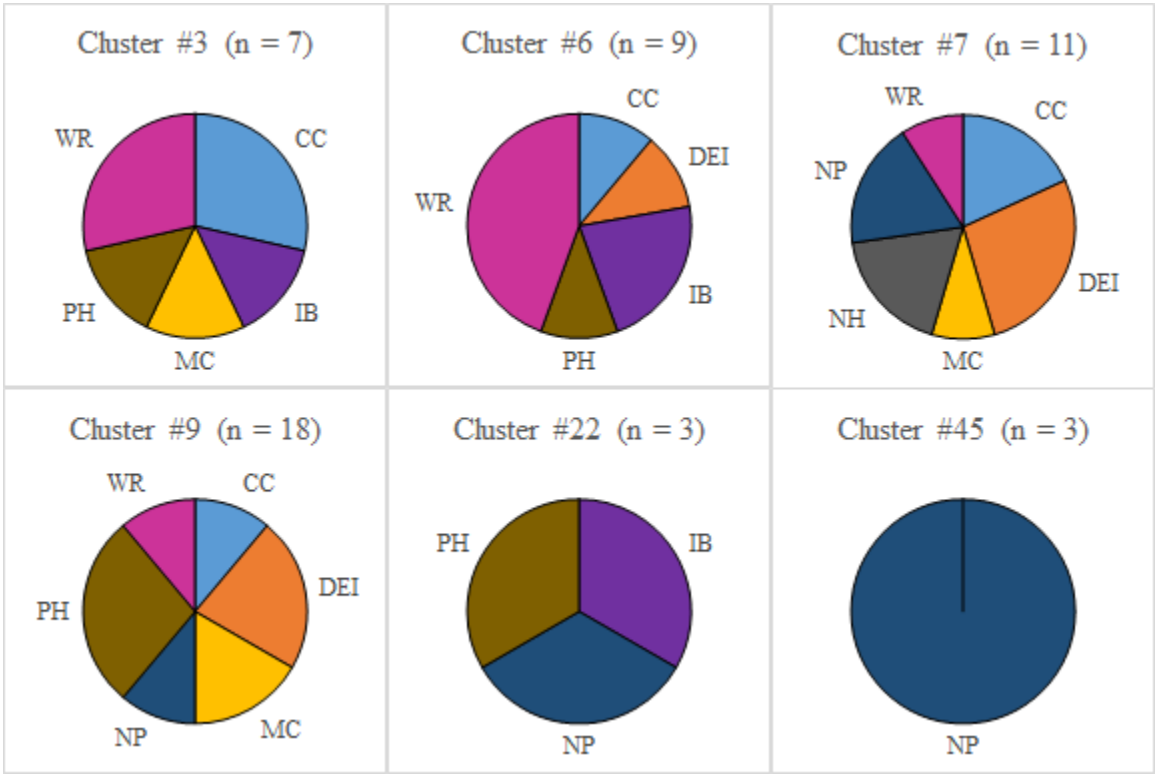
Figure 4.2 Identified sibling relationships across all sample populations



Full and half-sibling relationships identified among individuals from all sample populations. Full sibling relationships (n = 2) are indicated in orange and are shown in the upper diagonal of the figure. Half-sibling relationships (n = 215) are indicated in green and are shown in the lower diagonal. Relationships were calculated in COLONY using a maximum likelihood method based on individual genotype data of sampled *M. arenaria*.

Clustering of individuals into family groups also did not identify any meaningful patterns. This analysis produced 72 different clusters based on sibling relationships among sampled individuals. Fifty-nine of those clusters were comprised of only a single individual. Another seven clusters were comprised of two individuals apiece that were either full or half-siblings. Six clusters (#3, #6, #7, #9, #22 and #45) were comprised of three or more individuals. Cluster #45 contained three half-siblings that all originated from the same sample population. The remaining clusters contained individuals from a mixture of three to six different populations (Figure 4.3).

Figure 4.3 Family clusters based on identified sibling relationships



The six largest family clusters based on full and half-sibling relationships identified among sample individuals showing the population of origin of those individuals. Individuals were assigned to a family cluster based on a maximum likelihood method using COLONY.

4.4. Discussion

Population genetic structure can be influenced by many different factors, and in *M. arenaria* one of those drivers is differential selective pressure from HABs acting on the genetic trait of STX resistance (Bricelj *et al.* 2005; Connell *et al.* 2007). This work investigated whether patterns of differentiation at the NaV locus under selection were mirrored in selectively neutral microsatellite loci. The resulting picture of population genetic structure in *M. arenaria* indicates no similarities in the patterns of population

genetic structure between these two portions of the genome, and found no structure using the microsatellite markers.

PCR reactions were carried out for eight microsatellite loci in the populations, but a large number of individuals failed to amplify any alleles at the Mar2 locus and that marker had to be excluded from analysis. A study by St-Onge *et al.* (2013) that used these same markers also had difficulties with the Mar2 locus, and also dropped this locus from their study. These findings indicate that there may be problems with the Mar2 marker that should be investigated.

All eight of the populations in this study displayed significant deviations from Hardy-Weinberg equilibrium due to deficits of heterozygotes, however such deficits are a well-known and widespread occurrence in bivalves. Literature on population genetics work in bivalves documents many examples of heterozygote deficiencies in populations of these animals (Gaffney *et al.* 1990; Hare *et al.* 1996), and it appears to be a relatively common issue with a number of potential causes. These causes can be divided into two classes: those arising during processing and analysis of the data, and those arising from phenomena actually occurring in the population of interest.

The first category includes events such as incorrectly identifying artifacts, stutter peaks, or contamination as alleles, leading to the detection of false alleles. It also includes the possibilities of poor amplification of certain loci or the existence of null alleles, resulting in the failure to detect alleles that are present (Bonin *et al.* 2004; Selkoe and Toonen 2006). Another potential source of error during data analysis is homoplasy, or the existence of different alleles that are the same length. The methods used to distinguish between different alleles require the assumption that all alleles are of different sizes, so

homoplasy can result in the failure to detect alleles that may be present (Selkoe and Toonen 2006).

The second category of causes includes processes such as inbreeding, molecular imprinting, aneuploidy, selection against heterozygotes, and Wahlund effects on spatial or temporal scales (Gaffney *et al.* 1990). Inbreeding has long been known to cause a decrease in heterozygosity in affected populations, but was originally thought to be rare and of trivial importance for populations in the wild, especially for organisms like bivalves that tend to have large effective population sizes and planktonic larvae with a great capacity for dispersal. More recent research, however, has provided evidence that inbreeding and its subsequent effects such as inbreeding depression may be more common than originally thought, and that these effects can have considerable impacts on populations (Keller and Waller 2002). Molecular imprinting is a process by which alleles can be differentially expressed based on whether they are of maternal or paternal origin. The effects of molecular imprinting in wild populations can appear to be similar to the effects of inbreeding depression or null alleles (Chakraborty 1989), and will cause an apparent heterozygote deficit. Aneuploidy is the incidence of an abnormal number of chromosomes, typically arising from non-disjunction of chromosomes during cell division (Martin and Rademaker 1990). Hypoploidy, the lack of one chromosome in a homologous pair, is frequently observed in bivalves, especially after being exposed to certain chemical and environmental stressors (Bouilly *et al.* 2007), and can result in missing alleles during genetic analysis. Selection against heterozygote genotypes is possible if those genotypes have a lower fitness than homozygote genotypes. Such instances of such selection have been documented in oysters *Ostrea chilensis* (Toro and Vergara 1995), and are theorized

to occur in other bivalve species (Hare *et al.* 1996). The effect of selection can be detected by tracking heterozygosity in a cohort over time. Heterozygote deficiency in a population due to selection will manifest as a gradual decrease in heterozygosity with age, as opposed to other causes of heterozygote deficiency that will be evident in a larval cohort from the beginning (Toro and Vergara 1995). Wahlund effects can occur when there are cryptic subpopulations within a sample population. Pooling discrete subpopulations or cohorts will cause an apparent heterozygote deficiency if there are differences in the allele frequencies between those groups (Gaffney *et al.* 1990, Feldheim *et al.* 2011).

Given the numerous examples of null alleles in bivalves that occur in literature, and the fact that this is a known issue affecting microsatellite analyses in this class of organisms, it is likely that the heterozygote deficiencies observed in the eight *M. arenaria* populations in this study are the result of null alleles, although without further scrutiny other causes cannot be ruled out as also contributing to those deficits. Unfortunately, sequencing the microsatellite flanking regions of the sampled individuals and redesigning the primers to account for any polymorphisms was beyond the scope of this work, but ideally this would be the next step to attempt to ameliorate these heterozygote deficiencies.

All but one of the surveyed loci showed significant deviations from Hardy-Weinberg equilibrium as a result of reductions in observed heterozygosity as compared to expected heterozygosity, as well as statistically significant coefficients of inbreeding. The one exception was Mar6, which conformed to the allele frequencies that would be expected under conditions of Hardy-Weinberg equilibrium, and showed no reduction in observed heterozygosity or statistically significant level of inbreeding. Another study using these same markers found deviations from Hardy-Weinberg equilibrium at only three of the

seven loci (Mar1, Mar3, and Mar8) (St-Onge *et al.* 2013) rather than at six loci as was found in these results, and evidence of null alleles at only two of the loci (Mar 1 and Mar3) rather than at six as was seen in these results. It is likely that the results of my work have been confounded by the presence of null alleles at the affected loci that are causing perceived departures from Hardy-Weinberg equilibrium and artificially inflating estimates of inbreeding. It is also possible that movement of seed clams throughout the Gulf of Maine causes departures from Hardy-Weinberg equilibrium by introducing foreign alleles into populations in locations that are planted with seed clams.

The number of detected alleles and the allelic richness across the eight populations were high in comparison to the results seen in another study investigating *M. arenaria* (St-Onge *et al.* 2013), although the three populations in Penobscot Bay (Naskeag Harbor, The Narrows, and Clam Cove) had lower values than the others, and Downeast Institute was the lowest of all the populations. Statistical analysis showed that only the Downeast Institute population had significantly lower allelic richness compared to the others however, and that observed differences in the Penobscot Bay populations did not rise to the level of significance. Lower numbers of alleles and allelic richness are not surprising in the Downeast Institute population since this broodstock line was developed for rapid growth rate by culling slower growing individuals from each generation that may have resulted in the loss of some alleles from that population.

All eight populations had statistically significant inbreeding coefficients, which was unexpected given that *M. arenaria* typically has large effective population sizes and wide potential for dispersal. It is likely that these values are not a correct representation of inbreeding within sample populations, but rather that these results have been affected by

the high incidence of null alleles that was indicated in these populations. The one population that was expected to have a high level of inbreeding, Downeast Institute, did not have the highest inbreeding coefficient, despite several generations of selective breeding within this population without performing any outcrosses. This outcome is initially surprising, although it may also be skewed by artificially inflated results in the other populations.

Analysis of pairwise estimates of F_{ST} between all eight sample populations confirms the hypothesis that the DEI broodstock clams would be significantly diverged from the other populations due to the selective breeding that those individuals have undergone. This group was significantly different from all of the other populations except for Morong Cove and Webhannet River, where comparisons were close to the level of significance but did not attain it.

Northport Harbor, which was intended to be an out-group, was significantly different from two of the other populations (Morong Cove and Downeast Institute), but did not show significant differences from the remaining populations. Downeast Institute, as previously mentioned, is diverged from almost all of the other populations, so this result was expected. Morong Cove is the most geographically distant population from Northport Harbor in this study, and it is possible that the divergence between these two populations reflects some influence of isolation by distance. The fact that there were no significant divergences between any of the remaining populations despite some being separated by large distances does not necessarily mean that these populations are directly interbreeding, but rather that faraway populations can be connected by stepping stone populations between them and thus indirectly exchange genetic information (Hellberg 2009).

The assignment test produced results that were significantly different from those expected due to random chance, but those results do not appear to form any meaningful patterns. Most individuals were assigned either to Poorhouse Cove (39%) or The Narrows (23%). The miss-assignment of individuals sampled from populations located to the west of those two sites, like Webhannet River and Clam Cove, could be explained by recent larval transport along the prevailing currents, resulting in those individuals being assigned back to the populations upstream that they originated from. However the missassignment of individuals from populations east of Poorhouse Cove or The Narrows would seem to imply gene flow against the direction of the prevailing currents. Transport from southwest towards northeast against the prevailing currents is not impossible, and the existence of a latitudinal decrease in allelic richness among clam populations along the coast of North America from south to north suggests that this species expanded its range northward from southern refugia following the last glacial maximum (St-Onge *et al.* 2013). This movement would have been slow and incremental however, whereas the transport among populations indicated in this work seems to take place quite rapidly. Nearshore currents that occur close to the coastline are difficult to resolve spatially and temporally since they tend to vary over fine scales (Churchill *et al.* 2005), so these particular currents are not well described, but the possibility exists that one of these current systems could facilitate transport against the direction of the more prominent offshore currents.

Transfer of individuals between populations could explain instances where individuals were miss-assigned to a close neighboring population, but this was not the case for most of the individuals in this analysis. Further confounding the results, the Naskeag Harbor site, which is in the same region as The Narrows, was assigned almost no

individuals (3%). This great disparity between two sites that are so geographically close to each other is difficult to explain. The only population that was assigned fewer individuals than Naskeag Harbor was the Downeast Institute (2%), and the only individuals assigned to that group were individuals originally sampled there. Most of the individuals from the Downeast Institute population assigned elsewhere, mainly to Poorhouse Cove. This seems to imply that a large portion of the Downeast Institute broodstock were derived from Poorhouse Cove, which was not the case. Assignment testing tends to have difficulty assigning individuals to a population when genetic variability is low, and it is likely that this limitation is the cause of these confusing results. Low sample sizes can also affect the power of this analysis, since the potential source groups need to be distinct from one another, and a small sample size may not contain alleles that are rare within a certain group. Unfortunately it was not feasible to sample more than a small percentage of the population at each sample site.

Testing for recent migrants across all sample populations did not result in any clear patterns either. While 57 individuals (24%) were identified as recent migrants to their populations, statistical analysis found that the pattern of distribution of those migrants among the sample populations could be attributed to random chance. This seems to indicate that migration in the form of larval transport is occurring indiscriminately among the sample populations in this study, however this method is subject to the same weaknesses as the assignment test.

Parentage analysis revealed a number of sibling pairs among individuals in the data set, although those familial relationships do not seem to form any meaningful arrangements. Of the two full sibling pairs, both occurred between individuals within the

same sample population, which is the most logical scenario for full sibling relationships. The 215 half-sibling pairs that were identified were scattered within and across all eight of the sample populations, seeming to indicate shared parentage on a large scale throughout the area covered by this study. When individuals were grouped into family clusters based on these identified sibling relationships, those clusters often consisted of individuals from many different sample populations, seeming to indicate that offspring from the sampled areas are very widely dispersing. It seems unlikely, however, that individuals from such geographically distant locations as eastern Maine and Long Island could have a shared parent, and it is possible that null alleles could have affected this analysis. It is also possible that the level of genetic variation among the sample populations is simply too low to be able to accurately establish familial relationships. While parentage analysis tends to perform better than assignment testing under scenarios of low genetic variation, the sample populations in this study may be too homogeneous even for this particular method. In addition, the fact that it was not possible to sample more than a small percentage of each population may have handicapped the ability to identify relationships, since the precision of parentage analysis depends on the percentage of the population that is sampled.

As for the influence of softshell clam seeding efforts, it appears that either these seeded broodstock are not making significant contributions to the populations that they are being seeded into, or those seeded individuals were not sampled during any of the collections. The genetic homogeneity that is evident across the sample populations in this work therefore cannot be attributed to effects from seeding. Nevertheless, some of the results of this work may still be of interest to seeding programs, including the finding of reduced genetic diversity in the Downeast Institute population, and the fact that this

population has significantly diverged from wild clam populations in the Gulf of Maine. Planting seed clams that are genetically dissimilar to the local stocks they are intended to supplement can result in the seed clams being poorly suited for the environment they are being seeded into, and potentially lead to reduced fitness compared to the local wild stocks. In addition, the introduction of foreign genotypes into a population may disrupt the local balance of genotypes. To maximize the successful impact of seeding programs, care should be taken to ensure that hatchery stock do not significantly genetically diverge from the wild stocks that will be seeded.

In summary, it appears that wild *M. arenaria* populations in the Gulf of Maine can be characterized as one large, freely connected metapopulation, in which levels of gene flow among populations are sufficient to result in homogenization of those populations. No discernable population genetic structure could be identified at neutral microsatellite markers across the populations that were surveyed, and there are no existing patterns in relation to the population genetic structure that has been documented with regard to the NaV mutation.

CHAPTER 5
CONCLUSIONS

Mya arenaria populations in the Gulf of Maine exhibit striking differences in the occurrence of population genetic structure at the two different types of loci that were examined in these studies. These results provide information about the different factors that influence population genetic structure in this species, and the disparity between those results identifies the primary driver of that structure.

Mya arenaria exhibits a significant level of population structure at the resistant NaV mutation locus, as a result of the spatially heterogeneous formation of HABs and the variable levels selective pressure that those blooms exert on clam populations in different regions of the Gulf of Maine. Generally speaking, this selective pressure is most powerful in the eastern Gulf of Maine, strong in many areas of the western Gulf of Maine, moderate around Mount Desert Island, weak in southern New England, and largely absent in Penobscot Bay. This selection has produced significant differences in the distribution of genotypes for STX resistance at the NaV locus between discrete regions of the Gulf of Maine and southern New England. Populations in the eastern Gulf of Maine primarily contain individuals with the RR genotype, while populations in the western Gulf of Maine contain a fairly even mixture of genotypes, and populations in southern New England are almost exclusively made up of individuals with the SS genotype. This population genetic structure persists across yearly timescales, with genotype frequencies remaining relatively unchanged.

In sharp contrast to the strong population structure at the NaV mutation locus, neutral microsatellite markers reveal no evidence of any population genetic structure among *M. arenaria* in the Gulf of Maine. Extremely low levels of genetic variation at these loci failed to identify any meaningful patterns of differentiation among the wild populations that were sampled, although the results were confounded by the presumed presence of high numbers of null alleles in these analyses. These results suggest that gene flow is freely occurring between wild *M. arenaria* populations at a level sufficient to homogenize these populations at the neutral markers that were examined, and prevent the formation of localized variation. Even if populations are not directly interbreeding, they may be connected through stepping stone populations, or through migrants from a common pool of source individuals. The results of this work indicated gene flow occurring among populations across the statistically discrete regions for NaV locus genotypes, demonstrating that the neutral microsatellite markers do not display any of the same patterns that are observed for STX resistance.

The discrepancy between the results at the NaV locus and the microsatellite loci can be explained by the fact that changes at these loci occur at different rates. HABs can cause very rapid changes in the population genetic structure of *M. arenaria*, especially if the timing of a HAB coincides with the young juvenile stage of clams' development when they are most vulnerable to the harmful effects of PSTs (Bricelj *et al.* 2010). Local genotype shifts in response to HABs can occur over timescales as short as just a single generation. In contrast, the processes that result in population changes at neutral microsatellite loci happen over longer timescales, which is why it is possible to see no

evidence of structure at these loci despite the finding of a population structure at the NaV locus.

The differences in population genetic structure with respect to the NaV mutation are preserved by selective pressure stemming from regular annual HABs against a background of freely occurring gene flow among clam populations. For clams in the Gulf of Maine, genetic population structure is derived through differential selective pressure affecting a particular trait, rather than through geographical isolation of populations from each other and gradual divergence due to drift and lack of gene flow between them. This demonstrates a scenario in which structure can be maintained despite high levels of gene flow between populations.

Given the apparent commonness of widespread gene flow among populations and regions of the Gulf of Maine, it is probable that sites where the incidence of the R allele in the population does not agree with the history of HABs at that location, as observed at some of the sites in PB region in Chapter 2, are the result of larval transport from another region where the R allele regularly occurs in the population. The most likely candidate for the source of larval transport to the PB region would be the MDI region since it has a genotype distribution that is statistically equivalent to that of PB, and is situated in a location where the prevailing currents would facilitate transport from MDI towards PB. The degree of genetic sharing between populations in the PB region and the source populations providing the resistant allele-bearing larvae is substantial enough that the RR genotype is able to be retained in the PB region, despite the relative lack of selective pressure from HABs that should favor sensitive genotypes in that region because of the neuromuscular costs associated with resistance.

Temporal variations in the distribution of genotypes at the NaV mutation locus, such as those seen in the Naskeag Harbor population in Chapter 2, can be attributed to several factors. First, interannual variability in the occurrence, location, and severity of *Alexandrium* spp. blooms will cause differences in the strength of the selective pressure that is exerted on clam populations from year to year. Second, the degree of larval exchange between populations can fluctuate on a yearly basis, producing variations in the level of gene flow between those populations that could affect the abundance of genotypes with regard to the NaV mutation. The impact of HABs on shellfish in the Gulf of Maine can vary greatly on both temporal and spatial scales, as evidenced by large differences in tissue toxicity that have been observed across years and locations, including locations that are geographically close to one another (Shumway *et al.* 1994, Thomas *et al.* 2010). These differences can be attributed to the patchiness of *Alexandrium* spp. cell densities during blooms (Hurst and Yentsch 1981), and to factors affecting the intensity of transport of offshore cells into the nearshore environment where they can affect shellfish beds (McGillicuddy *et al.* 2005). Variations in water circulation patterns can influence temporal changes in genotype distributions by affecting the transport of larvae between populations, causing fluctuations in the degree of genetic sharing and connectivity between those populations. The nearshore environment is an extremely dynamic one, where great differences in tidal mixing, currents, wind, and other physical processes are common over short distances and timescales. (Largier 2003; Cowen and Sponaugle 2009). These processes can either enhance or inhibit the delivery of larvae to downstream populations and cause variability in the degree of genetic sharing between those populations. If the surplus of RR individuals in Penobscot Bay populations is dependent upon larval transport

from other populations, it is possible that interannual fluctuations in current patterns could cause changes in the strength of that transport, and contribute to the observed decrease in resistant individuals at the Naskeag Harbor site.

The results of this work may be of interest to seeding programs since they offer information about the genetic population structure of *M. arenaria* that could help improve the fitness of seeded broodstock for the environment they will be seeded into. While there appears to be no evidence to support a theory of genetically differentiated subpopulations of clams in the Gulf of Maine, the fact remains that there are significant differences among regions for the predominant genotypes for STX resistance, therefore a "one size fits all" approach for seeding juvenile clams may not be the best strategy. While the MDI, PB, and WGoM regions can be thought of as one homogenous area with regard to their genotype distributions, the EGoM region is distinctly different in its extremely high proportion of RR individuals, and presents an area where seeding should perhaps be treated differently. Given the common occurrence of HABs in the eastern Gulf of Maine and the very high levels of toxicity that accompany them, this is a place where one genotype is clearly more advantageous than the others, in contrast to the rest of the Gulf of Maine where the genotype proportions are distributed more equally. In addition, it has been shown that selectively breeding hatchery clams to encourage fast growth, as has been done with the DEI broodstock from Chapter 3, produces a shift in the distribution of genotypes for STX resistance towards more RS and SS individuals, and fewer RR individuals (Phillips unpublished). The offspring of such a broodstock line would be appropriate for seeding around Mount Desert Island, and in Penobscot Bay and western Maine where genotypes for STX resistance are evenly distributed within populations in those regions, but RS and

SS clams would be poorly suited for seeding in the eastern Gulf of Maine, and doing so would likely result in mortality among those individuals and hinder the effectiveness of seeding. For seeding in the eastern Gulf of Maine, a preferable strategy would be to use locally collected clams from that region as broodstock, thereby preserving the high incidence of the RR genotype in the resulting offspring.

One feature of possible concern in the DEI broodstock is the fact that these individuals have begun to significantly genetically diverge from some wild populations in the Gulf of Maine as evidenced by calculations of F_{ST}. Seeding wild populations with clams that are genetically different from those populations will either result in the disruption of local genotype frequencies if the seed clams are well adapted for the conditions, or large-scale mortality of those seed clams if they are poorly adapted for the local conditions. Another feature that may be of concern is this population's reduced genetic variation in comparison to the wild populations, as measured by allelic richness. This is an issue commonly encountered in hatchery broodstock lines, which tend to be derived from a small number of individuals and may be selectively bred to enhance certain desirable traits, resulting in the loss of alleles (Taris *et al.* 2006). Care must be taken to ensure that hatchery stock maintain an adequate level of genetic variability, since reduced variability can impair a population's ability to adapt and cope with disturbances (Taris *et al.* 2006). A strategy that would maintain diversity while still enabling selective breeding for enhanced growth would be to develop several lines of fast-growing broodstock and outcross those lines with each other at regular intervals (Sponenberg and Bixby 2007).

From a management perspective, this work can aid in identifying areas where there is an elevated risk of severe PSP, and areas in which harvesting closures due to HABs will

have the most serious economic effects. Resistant genotype clams accumulate substantially higher levels of PSTs in their body tissues than SS clams do during HABs because they remain unaffected by STX-induced paralysis, therefore RR clams pose a greater danger to humans who might accidentally consume them. In addition, it takes RR clams longer to depurate to a tissue toxicity level that is safe for consumption after a bloom has ended, meaning that harvesting areas comprised of predominantly RR clams will need to be closed for longer periods of time following HABs, exacerbating the economic effects of those closures. These issues will be of greatest concern in the eastern Gulf of Maine where clam populations are comprised, on average, of 70% RR individuals, and some populations are as high as 85% RR individuals. Eastern Maine has historically produced 45-65% of all softshell clam landings in the state of Maine (Beal 2002), so the impact of these issues could have notable implications for the state as a whole.

In summary, this work provides important information about the structure and connectivity of *M. arenaria* populations in the Gulf of Maine, and describes the spatial and temporal extent of a STX-resistant mutation in clams in the Gulf of Maine, southern New England, and Long Island, NY. These results provide insight into gene flow among clam populations in the Gulf of Maine, and recognize differential selective pressure on a specific trait as the driver of population structure in these animals. This work also identifies issues that may be of interest to hatchery seeding programs looking to maximize the success of their efforts, as well as to management entities that monitor for PSP in the Gulf of Maine.

REFERENCES

Adachi, M., Sake, Y., & Ishida, Y. (1996). Analysis of *Alexandrium* (Dinophyceae) species using sequences of the 5.8 S ribosomal DNA and internal transcribed spacer regions1. *Journal of Phycology*, *32*(3), 424-432.

Almany, G. R., Berumen, M. L., Thorrold, S. R., Planes, S., & Jones, G. P. (2007). Local replenishment of coral reef fish populations in a marine reserve. *Science*, *316*(5825), 742-744. doi: 10.1126/science.1140597

Anderson, D. M., Kulis, D. M., Doucette, G. J., Gallagher, J. C., & Balech, E. (1994). Biogeography of toxic dinoflagellates in the genus *Alexandrium* from the northeastern United States and Canada. *Marine Biology*, *120*(3), 467-478.

Anderson, D. M., Keafer, B. A., McGillicuddy, D. J., Mickelson, M. J., Keay, K. E., Libby, P. S., et. al. (2005). Initial observations of the 2005 *Alexandrium fundyense* bloom in southern New England: General patterns and mechanisms. *Deep Sea Research Part II: Topical Studies in Oceanography*, *52*(19), 2856-2876.

Anderson, P. A., Roberts-Misterly, J., & Greenberg, R. M. (2005). The evolution of voltage-gated sodium channels: were algal toxins involved?. *Harmful Algae*, *4*(1), 95-107.

Auld, V. J., Goldin, A. L., Krafte, D. S., Catterall, W. A., Lester, H. A., Davidson, N., & Dunn, R. J. (1990). A neutral amino acid change in segment IIS4 dramatically alters the gating properties of the voltage-dependent sodium channel. *Proceedings of the National Academy of Sciences*, *87*(1), 323-327.

Avise, J. C., Lansman, R. A., & Shade, R. O. (1979). The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. I. Population structure and evolution in the genus *Peromyscus*. *Genetics*, *92*(1), 279-295.

Beal, B. F., Lithgow, C. D., Shaw, D. P., Renshaw, S., & Ouellette, D. (1995). Overwintering hatchery-reared individuals of the soft-shell clam, *Mya arenaria* L.: a field test of site, clam size, and intraspecific density. *Aquaculture*, *130*(2), 145-158.

Beal, B. F., Bayer, R., Kraus, M. G., & Chapman, S. R. (1999). A unique shell marker in juvenile, hatchery-reared individuals of the softshell clam, *Mya arenaria* L. *Fishery Bulletin-National Oceanic and Atmospheric Administration*, *97*, 380-386.

Beal, B. F., Parker, M. R., & Vencile, K. W. (2001). Seasonal effects of intraspecific density and predator exclusion along a shore-level gradient on survival and

growth of juveniles of the soft-shell clam, *Mya arenaria* L., in Maine, USA. *Journal of Experimental Marine Biology and Ecology*, *264*(2), 133-169.

Beal, B. F. (2004). Softshell clam, *Mya arenaria*, mariculture in Maine, USA: opportunities and challenges. AAC Special Publications, 9, 41-44.

Beal, B. F. (2006). Biotic and abiotic factors influencing growth and survival of wild and cultured individuals of the softshell clam (*Mya arenaria* L.) in eastern Maine. *Journal of Shellfish Research*, *25*(2), 461-474.

Beal, B. (2015). *Interactive effects of tidal height and predator exclusion on growth and survival of wild and cultured juveniles of the soft-shell clam,* Mya arenaria *L., at two intertidal flats in southern Maine* (pp. 1-48). Downeast Institute for Applied Marine Research and Education. Retrieved from http://www.downeastinstitute.org/assets/files/manuals/Final%20Report%20-%20Wells%20and%20Portland%20clam%20project%20-%202014.pdf

Bonin, A., Bellemain, E., Bronken Eidesen, P., Pompanon, F., Brochmann, C., & Taberlet, P. (2004). How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, 13(11), 3261-3273.

Bouilly, K., Bonnard, M., Gagnaire, B., Renault, T., & Lapègue, S. (2007). Impact of diuron on aneuploidy and hemocyte parameters in Pacific oyster, *Crassostrea gigas*. *Archives of Environmental Contamination and Toxicology*, *52*(1), 58-63.

Bricelj, V. M., Cembella, A., Laby, D., Shumway, S. E., & Cucci, T. L. (1996). Comparative physiological and behavioral responses to PSP toxins in two bivalve molluscs, the softshell clam, *Mya arenaria*, and surfclam, *Spisula solidissima*. *Harmful and Toxic Algal Blooms, Intergovernmental Oceanographic Commission of UNESCO, Paris, Yasumoto, T. Oshima, Y., Fukuyo, Y.(Eds.)*, 405-408.

Bricelj, V. M., MacQuarrie, S. P., Twarog, B., & Trainer, V. (2002). Characterization of sensitivity to PSP toxins in North American populations of the softshell clam *Mya arenaria*. *Harmful Algae*, 172-174.

Bricelj, V. M., Connell, L., Konoki, K., MacQuarrie, S. P., Scheuer, T., Catterall, W. A., & Trainer, V. L. (2005). Sodium channel mutation leading to saxitoxin resistance in clams increases risk of PSP. *Nature*, *434*(7034), 763-767.

Bricelj, V. M., MacQuarrie, S. P., Doane, J. A. E., & Connell, L. B. (2010). Evidence of selection for resistance to paralytic shellfish toxins during the early life history of soft-shell clam, *Mya arenaria*, populations. *Limnology and Oceanography*, *55*(6), 2463-2475.

Brodie III, E. D., & Brodie Jr, E. D. (1999). Costs of exploiting poisonous prey: evolutionary trade-offs in a predator-prey arms race. *Evolution*, 626-631.

Brooks, D. A., & Townsend, D. W. (1989). Variability of the coastal current and nutrient pathways in the eastern Gulf of Maine. *Journal of Marine Research*, 47(2), 303-321.

Brownlow, R. J., Dawson, D. A., Horsburgh, G. J., Bell, J. J., & Fish, J. D. (2008). A method for genotype validation and primer assessment in heterozygote-deficient species, as demonstrated in the prosobranch mollusc *Hydrobia ulvae*. *BMC Genetics*, *9*(1), 1.

Caporale, D. A., Beal, B. F., Roxby, R., & Van Beneden, R. J. (1997). Population structure of *Mya arenaria* along the New England coastline. *Molecular Marine Biology and Biotechnology*, 6(1), 33-39.

Catterall, W. A. (1992). Cellular and molecular biology of voltage-gated sodium channels. *Physiological Reviews*, *72*(suppl 4), S15-S48.

Catterall, W. A. (2000). From ionic currents to molecular mechanisms: the structure and function of voltage-gated sodium channels. *Neuron*, *26*(1), 13-25.

Cerrato, R. M., Wallace, H. V., & Lightfoot, K. G. (1991). Tidal and seasonal patterns in the chondrophore of the soft-shell clam *Mya arenaria*. *The Biological Bulletin*, *181*(2), 307-311.

Chakraborty, R. (1989). Can molecular imprinting explain heterozygote deficiency and hybrid vigor?. *Genetics*, *122*(3), 713-717.

Chen, L. (2010). A molecular approach to the study of saxitoxin resistance in copepods. University of Connecticut: 165.

Churchill, J. H., Pettigrew, N. R., & Signell, R. P. (2005). Structure and variability of the western Maine coastal current. *Deep Sea Research Part II: Topical Studies in Oceanography*, *52*(19), 2392-2410.

Congleton, W. R., Pearce, B. R., Parker, M. R., & Causey, R. C. (2003). Mariculture siting-Tidal currents and growth of *Mya arenaria*. *Journal of Shellfish Research*, 22(1), 75-84.

Congleton Jr, W. R., Vassiliev, T., Bayer, R. C., Pearce, B. R., Jacques, J., & Gillman, C. (2006). Trends in Maine softshell clam landings. *Journal of Shellfish Research*, *25*(2), 475-480.

Connell, L. B., MacQuarrie, S. P., Twarog, B. M., Iszard, M., & Bricelj, V. M. (2007). Population differences in nerve resistance to paralytic shellfish toxins in softshell clam, *Mya arenaria*, associated with sodium channel mutations. *Marine Biology*, *150*(6), 1227-1236.

Cowen, R. K., & Sponaugle, S. (2009). Larval dispersal and marine population connectivity. *Annual Review of Marine Science*, *1*, 443-466.

Crandall, E. D., Frey, M. A., Grosberg, R. K., & Barber, P. H. (2008). Contrasting demographic history and phylogeographical patterns in two Indo-Pacific gastropods. *Molecular ecology*, *17*(2), 611-626.

Dakin, E. E., & Avise, J. C. (2004). Microsatellite null alleles in parentage analysis. *Heredity*, *93*(5), 504-509.

Dam, H. G., & Haley, S. T. (2011). Comparative dynamics of paralytic shellfish toxins (PST) in a tolerant and susceptible population of the copepod *Acartia hudsonica*. *Harmful Algae*, *10*(3), 245-253.

Dallas, J. F. (1992). Estimation of microsatellite mutation rates in recombinant inbred strains of mouse. *Mammalian Genome*, *3*(8), 452-456.

Davis, C. V., & Barber, B. J. (1999). Growth and survival of selected lines of eastern oysters, *Crassostrea virginica* (Gmelin 1791) affected by juvenile oyster disease. *Aquaculture*, *178*(3), 253-271.

Denac, H., Mevissen, M., & Scholtysik, G. (2000). Structure, function and pharmacology of voltage-gated sodium channels. *Naunyn-Schmiedeberg's Archives of Pharmacology*, *362*(6), 453-479.

Feldheim, K. A., Brown, J. E., Murphy, D. J., & Stepien, C. A. (2011). Microsatellite loci for dreissenid mussels (Mollusca: Bivalvia: Dreissenidae) and relatives: markers for assessing exotic and native populations. *Molecular Ecology Resources*, *11*(4), 725-732.

Feldman, C. R., Brodie, E. D., & Pfrender, M. E. (2009). The evolutionary origins of beneficial alleles during the repeated adaptation of garter snakes to deadly prey. *Proceedings of the National Academy of Sciences*, *106*(32), 13415-13420.

Franks, P. J. S., & Anderson, D. M. (1992). Alongshore transport of a toxic phytoplankton bloom in a buoyancy current: *Alexandrium tamarense* in the Gulf of Maine. *Marine Biology*, *112*(1), 153-164.

Franks, P. J. S., & Anderson, D. M. (1992). Toxic phytoplankton blooms in the southwestern Gulf of Maine: testing hypotheses of physical control using historical data. *Marine Biology*, *112*(1), 165-174.

Gaffney, P. M., Scott, T. M., Koehn, R. K., & Diehl, W. J. (1990). Interrelationships of heterozygosity, growth rate and heterozygote deficiencies in the coot clam, *Mulinia lateralis*. *Genetics*, *124*(3), 687.

Ganong, W. F. (1889). *The economic mollusca of Acadia* (No. 8). St. John, NB: Barnes.

Geffeney, S. L., Fujimoto, E., Brodie, E. D., & Ruben, P. C. (2002). Evolutionary diversification of TTX-resistant sodium channels in a predator–prey interaction. *Nature*, *434*(7034), 759-763.

Glude, J. B. (1955). The effects of temperature and predators on the abundance of the soft-shell clam, *Mya arenaria*, in New England. *Transactions of the American Fisheries Society*, *84*(1), 13-26.

Goudet, J. (2001). FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9. 3).

Hamilton, S. A. (2009). Investigating saxitoxin resistance in softshell clams (*Mya arenaria*): patterns of inheritance and improvements on methodology for tracking and identification (Masters Thesis, The University of Maine).

Hamilton, S., & Connell, L. (2009). Improved methodology for tracking and genetically identifying the softshell clam *Mya arenaria*. *Journal of Shellfish Research*, *28*(4), 747-750.

Hare, M. P., Karl, S. A., & Avise, J. C. (1996). Anonymous nuclear DNA markers in the American oyster and their implications for the heterozygote deficiency phenomenon in marine bivalves. *Molecular Biology and Evolution*, *13*(2), 334-345.

Hattenrath, T. K., Anderson, D. M., & Gobler, C. J. (2010). The influence of anthropogenic nitrogen loading and meteorological conditions on the dynamics and toxicity of *Alexandrium fundyense* blooms in a New York (USA) estuary. *Harmful Algae*, *9*(4), 402-412.

Hedgecock, D., Tracey, M. L., & Nelson, K. (1982). Genetics. In 'The Biology of Crustacea'. (Ed. LG Abele.) pp. 284-403.

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

Hedgecock, D., Li, G., Hubert, S., Bucklin, K., & Ribes, V. (2004). Widespread null alleles and poor cross-species amplification of microsatellite DNA loci cloned from the Pacific oyster, *Crassostrea gigas*. *Journal of Shellfish Research*, *23*(2), 379-386.

Hedgecock, D., Launey, S., Pudovkin, A. I., Naciri, Y., Lapegue, S., & Bonhomme, F. (2007). Small effective number of parents (N b) inferred for a naturally spawned

cohort of juvenile European flat oysters *Ostrea edulis*. *Marine Biology*, *150*(6), 1173-1182.

Hellberg, M. E. (2009). Gene flow and isolation among populations of marine animals. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 291-310.

Hilbish, T. J. (1985). Demographic and temporal structure of an allele frequency cline in the mussel Mytilus edulis. *Marine Biology*, *86*(2), 163-171.

Hunt, H. L., & Mullineaux, L. S. (2002). The roles of predation and postlarval transport in recruitment of the soft shell clam (*Mya arenaria*). *Limnology and Oceanography*, *47*(1), 151-164.

Hunt, H. L., & Scheibling, R. E. (1997). Role of early post-settlement mortality in recruitment of benthic marine invertebrates. *Marine Ecology Progress Series*, *155*, 269-301.

Hurst Jr, J. W. (1975). History of paralytic shellfish poisoning on the Maine coast 1958-1974. In *Proceedings of the First International Conference on Toxic Dinoflagellate B. 532* (p. 1975).

Hurst, J. W., & Yentsch, C. M. (1981). Patterns of intoxication of shellfish in the Gulf of Maine coastal waters. *Canadian Journal of Fisheries and Aquatic Sciences*, *38*(2), 152-156.

Jin, D., Thunberg, E., & Hoagland, P. (2008). Economic impact of the 2005 red tide event on commercial shellfish fisheries in New England. *Ocean & Coastal Management*, *51*(5), 420-429.

Johnson, M. S., & Black, R. (1982). Chaotic genetic patchiness in an intertidal limpet, *Siphonaria* sp. *Marine Biology*, *70*(2), 157-164.

Jones, A. G., Small, C. M., Paczolt, K. A., & Ratterman, N. L. (2010). A practical guide to methods of parentage analysis. *Molecular Ecology Resources*, *10*(1), 6-30.

Jones, O. R., & Wang, J. (2010). COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, *10*(3), 551-555.

Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, *17*(5), 230-241.

Largier, J. L. (2003). Considerations in estimating larval dispersal distances from oceanographic data. *Ecological Applications*, *13*(sp1), 71-89.

Lee, C. H., Jones, D. K., Ahern, C., Sarhan, M. F., & Ruben, P. C. (2011). Biophysical costs associated with tetrodotoxin resistance in the sodium channel pore of the garter snake, *Thamnophis sirtalis*. *Journal of Comparative Physiology A*, *197*(1), 33-43.

Lemer, S., Rochel, E., & Planes, S. (2011). Correction method for null alleles in species with variable microsatellite flanking regions, a case study of the black-lipped pearl oyster *Pinctada margaritifera*. *Journal of Heredity*, esq123.

Li, G., & Hedgecock, D. (1998). Genetic heterogeneity, detected by PCR-SSCP, among samples of larval Pacific oysters (*Crassostrea gigas*) supports the hypothesis of large variance in reproductive success. *Canadian Journal of Fisheries and Aquatic Sciences*, *55*(4), 1025-1033.

MacQuarrie, S. P., & Bricelj, V. M. (2008). Behavioral and physiological responses to PSP toxins in *Mya arenaria* populations in relation to previous exposure to red tides. *Marine Ecology Progress Series*, *366*, 59-74.

Maranda, L., Anderson, D. M., & Shimizu, Y. (1985). Comparison of toxicity between populations of *Gonyaulax tamarensis* of eastern North American waters. *Estuarine, Coastal and Shelf Science*, *21*(3), 401-410.

Martin, R. H., & Rademaker, A. (1990). The frequency of aneuploidy among individual chromosomes in 6,821 human sperm chromosome complements. *Cytogenetic and Genome Research*, *53*(2-3), 103-107.

Martin, J. L., LeGresley, M. M., & Hanke, A. R. (2014). Thirty years–*Alexandrium fundyense* cyst, bloom dynamics and shellfish toxicity in the Bay of Fundy, eastern Canada. *Deep Sea Research Part II: Topical Studies in Oceanography*, *103*, 27-39.

Masseret, E., Grzebyk, D., Nagai, S., Genovesi, B., Lasserre, B., Laabir, M., et. al. (2009). Unexpected genetic diversity among and within populations of the toxic dinoflagellate *Alexandrium catenella* as revealed by nuclear microsatellite markers. *Applied and Environmental Microbiology*, 75(7), 2037-2045.

McGillicuddy, D. J., Anderson, D. M., Lynch, D. R., & Townsend, D. W. (2005). Mechanisms regulating large-scale seasonal fluctuations in *Alexandrium fundyense* populations in the Gulf of Maine: results from a physical–biological model. *Deep Sea Research Part II: Topical Studies in Oceanography*, *52*(19), 2698-2714.

Morse, B. L., & Hunt, H. L. (2013). Impact of settlement and early post-settlement events on the spatial distribution of juvenile *Mya arenaria* on an intertidal shore. *Journal of Experimental Marine Biology and Ecology*, *448*, 57-65.

Noda, M., Suzuki, H., Numa, S., & Stühmer, W. (1989). A single point mutation confers tetrodotoxin and saxitoxin insensitivity on the sodium channel II. *FEBS Letters*, *259*(1), 213-216.

Paetkau, D., Slade, R., Burden, M., & Estoup, A. (2004). Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology*, *13*(1), 55-65.

Pettigrew, N. R., Townsend, D. W., Xue, H., Wallinga, J. P., Brickley, P. J., & Hetland, R. D. (1998). Observations of the Eastern Maine Coastal Current and its offshore extensions in 1994. *Journal of Geophysical Research: Oceans (1978–2012)*, *103*(C13), 30623-30639.

Pettigrew, N. R., Churchill, J. H., Janzen, C. D., Mangum, L. J., Signell, R. P., Thomas, A. C., et. al. (2005). The kinematic and hydrographic structure of the Gulf of Maine Coastal Current. *Deep Sea Research Part II: Topical Studies in Oceanography*, *52*(19), 2369-2391.

Piry, S., Alapetite, A., Cornuet, J. M., Paetkau, D., Baudouin, L., & Estoup, A. (2004). GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, *95*(6), 536-539.

Prakash, S., Lewontin, R. C., & Hubby, J. L. (1969). A molecular approach to the study of genic heterozygosity in natural populations IV. Patterns of genic variation in central, marginal and isolated populations of *Drosophila pseudoobscura*. *Genetics*, *61*(4), 841.

Pusch, M., Noda, M., Stühmer, W., Numa, S., & Conti, F. (1991). Single point mutations of the sodium channel drastically reduce the pore permeability without preventing its gating. *European Biophysics Journal*, *20*(3), 127-133.

Rannala, B., & Mountain, J. L. (1997). Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences*, *94*(17), 9197-9201.

Reece, K. S., Ribeiro, W. L., Gaffney, P. M., Carnegie, R. B., & Allen, S. K. (2004). Microsatellite marker development and analysis in the eastern oyster (*Crassostrea virginica*): confirmation of null alleles and non-Mendelian segregation ratios. *Journal of Heredity*, *95*(4), 346-352.

Roman, J. (2006). Diluting the founder effect: cryptic invasions expand a marine invader's range. *Proceedings of the Royal Society of London B: Biological Sciences*, *273*(1600), 2453-2459.

Ropes, J. W. (1968). The feeding habits of the green crab, *Carcinus maenas* (L.). *Fisheries Bulletin*, *67*(2), 183-203.

Ropes, J. W., & Stickney, A. P. (1965). Reproductive cycle of *Mya arenaria* in New England. *Biological Bulletin*, 315-327.

Saenz-Agudelo, P., Jones, G. P., Thorrold, S. R., & Planes, S. (2009). Estimating connectivity in marine populations: an empirical evaluation of assignment tests and parentage analysis under different gene flow scenarios. *Molecular Ecology*, *18*(8), 1765-1776.

Scattergood, L. W. (1952). *The distribution of the green crab,* Carcinides maenas *(L.) in the northwestern Atlantic*. Maine Department of Sea and Shore Fisheries.

Selkoe, K. A., & Toonen, R. J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, *9*(5), 615-629.

Shumway, S. E., Sherman-Caswell, S., & Hurst, J. W. (1988). Paralytic shellfish poisoning in Maine: monitoring a monster. *Journal of Shellfish Research*, *7*(4), 643-652.

Shumway, S. E., Sherman, S. A., Cembella, A. D., & Selvin, R. (1994). Accumulation of paralytic shellfish toxins by surfclams, *Spisula solidissima* (Dillwyn, 1897) in the Gulf of Maine: seasonal changes, distribution between tissues, and notes on feeding habits. *Natural Toxins*, *2*(4), 236-251.

Sokal, R. R., & Rohlf, F. J. (1981). Biometry (2nd edn). *New York: WH Feeman and Company*, *668*.

Sponenberg, D. P., & Bixby, D. E. (2007). *Managing breeds for a secure future: strategies for breeders and breed associations*. Albc.

Stickney, A. P. (1964). Salinity, temperature, and food requirements of soft-shell clam larvae in laboratory culture. *Ecology*, 283-291.

St-Onge, P., Parent, E, Sévigny, J., Tremblay, R., Rioux-Paré, R. (2011) Isolation and characterization of eight novel microsatellite markers for the softshell clam (*Mya arenaria*). *Molecular Ecology Resources Primer Database* mer-11-0122.

St-Onge, P., Sévigny, J. M., Strasser, C., & Tremblay, R. (2013). Strong population differentiation of softshell clams (*Mya arenaria*) sampled across seven biogeographic marine ecoregions: possible selection and isolation by distance. *Marine Biology*, *160*(5), 1065-1081.

St-Onge, P., Tremblay, R., & Sévigny, J. M. (2015). Tracking larvae with molecular markers reveals high relatedness and early seasonal recruitment success in a partially spawning marine bivalve. *Oecologia*, *178*(3), 733-746.

Strasser, C. A., & Barber, P. H. (2009). Limited genetic variation and structure in softshell clams (*Mya arenaria*) across their native and introduced range. *Conservation Genetics, 10*(4), 803-814.

Tan, E. B. P., & Beal, B. F. (2015). Interactions between the invasive European green crab, *Carcinus maenas* (L.), and juveniles of the soft-shell clam, *Mya arenaria* L., in eastern Maine, USA. *Journal of Experimental Marine Biology and Ecology*, *462*, 62-73.

Taris, N., Ernande, B., McCombie, H., & Boudry, P. (2006). Phenotypic and genetic consequences of size selection at the larval stage in the Pacific oyster (*Crassostrea gigas*). *Journal of Experimental Marine Biology and Ecology*, *333*(1), 147-158.

Taris, N., Boudry, P., Bonhomme, F., Camara, M. D., & Lapègue, S. (2009). Mitochondrial and nuclear DNA analysis of genetic heterogeneity among recruitment cohorts of the European flat oyster *Ostrea edulis*. *The Biological Bulletin*, *217*(3), 233-241.

Terlau, H., Heinemann, S. H., Stühmer, W., Pusch, M., Conti, F., Imoto, K., & Numa, S. (1991). Mapping the site of block by tetrodotoxin and saxitoxin of sodium channel II. *FEBS Letters*, *293*(1), 93-96.

Thomas, A. C., Weatherbee, R., Xue, H., & Liu, G. (2010). Interannual variability of shellfish toxicity in the Gulf of Maine: Time and space patterns and links to environmental variability. *Harmful Algae*, *9*(5), 458-480.

Toro, J. E., & Vergara, A. M. (1995). Evidence for selection against heterozygotes: Post-settlement excess of allozyme homozygosity in a cohort of the chilean oyster, *Ostrea chilensis* Philippi, 1845. *The Biological Bulletin*, *188*(2), 117-119.

Townsend, D. W., Pettigrew, N. R., & Thomas, A. C. (2001). Offshore blooms of the red tide dinoflagellate, *Alexandrium* sp., in the Gulf of Maine. *Continental Shelf Research*, *21*(4), 347-369.

Trainer, V. L., Eberhart, B. T. L., Wekell, J. C., Adams, N. G., Hanson, L., Cox, F., & Dowell, J. (2003). Paralytic shellfish toxins in Puget Sound, Washington state. *Journal of Shellfish Research*, *22*(1), 213-223.

Twarog, B. M., Hidaka, T., & Yamaguchi, H. (1972). Resistance to tetrodotoxin and saxitoxin in nerves of bivalve molluscs: a possible correlation with paralytic shellfish poisoning. *Toxicon*, *10*(3), 273-278.

Twarog, B. M., & Yamaguchi, H. (1975). Resistance to paralytic shellfish toxins in bivalve molluscs. In *Proceedings of the First International Conference on Toxic Dinoflagellate B. 532* (p. 1975).

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P., & Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, *4*(3), 535-538.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T., Hornes, M., et al. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, *23*(21), 4407-4414.

Wang, J., Qi, H., Li, L., & Zhang, G. (2014). Genome-wide survey and analysis of microsatellites in the Pacific oyster genome: abundance, distribution, and potential for marker development. *Chinese Journal of Oceanology and Limnology*, *32*, 8-21.

Waples, R. S. (1998). Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, *89*(5), 438-450.

Welch, W. R. (1969). Changes in abundance of the green crab, *Carcinus maenas* (L.), in relation to recent temperature changes. Maine Department of Marine Resources. Bureau of Commercial Fisheries Biological Laboratory.

White, A. W., & Lewis, C. M. (1982). Resting cysts of the toxic, red tide dinoflagellate *Gonyaulax excavate* in Bay of Fundy sediments. *Canadian Journal of Fisheries and Aquatic Sciences*, *39*(8), 1185-1194.

Whitlock, M. C., & McCauley, D. E. (1999). Indirect measures of gene flow and migration: FST≠ 1/(4Nm+ 1). *Heredity*, *82*(2), 117-125.

Xiao, J., Cordes, J. F., Wang, H., Guo, X., & Reece, K. S. (2010). Population genetics of *Crassostrea ariakensis* in Asia inferred from microsatellite markers. *Marine Biology*, *157*(8), 1767-1781.

APPENDIX: STATISTICAL DATA FROM CHAPTER 2

Table A.1 Within-cluster G-tests of goodness-of-fit for all sites

Site	RR	RS	SS	G-value	DF	p-value
Lepreau Basin	52	22	1	2.734	2	0.255
Beaver Cove	23	5	1	1.298	2	0.523
Deadman's Harbour	18	10	1	1.174	2	0.556
Gleason Cove	23	4	1	2.261	2	0.323
Prince Cove	19	6	0	2.398	2	0.302
Morong Cove	22	12	7	10.278	2	**0.00586**
Whiting Cove	19	7	1	0.043	2	0.979
Duck Brook	11	2	0	2.15	2	0.341
Mosquito Harbor	9	25	6	3.965	2	0.138
Young's Shore	11	7	4	2.53	2	0.282
Raccoon Cove	25	33	14	0.154	2	0.926
Ship Harbor	11	15	4	0.425	2	0.808
Duck Cove	9	8	5	1.109	2	0.575
Naskeag Harbor	25	21	10	2.688	2	0.261
Herrick Bay	8	14	5	1.195	2	0.55
Islesboro	6	16	7	3.057	2	0.217
Clam Cove	10	8	12	4.536	2	0.104
Lowe's Cove	4	7	7	3.948	2	0.139
Poorhouse Cove	90	216	70	3.092	2	0.213
Dingley Island	15	27	10	0.054	2	0.973
Biddeford Pool	7	14	4	0.139	2	0.933
Webhannet River	14	10	3	7.044	2	0.03
Kittery Point	6	8	5	1.104	2	0.576
Buttermilk Bay	0	6	23	9.762	2	**0.00759**
Northport Harbor	0	4	152	1.601	2	0.449
Point Lookout	0	0	35	3.223	2	0.2

Results of repeated G-tests of goodness-of-fit performed for all sites within each region defined in Chapter 2 to determine homogeneity within the *a priori* groupings. For each site, RR indicates the number of resistant genotype clams, RS indicates the number of heterozygous genotype clams, and SS indicates the number of sensitive genotype clams.

The calculated G-value and p-value for each site are shown, and p-values that are significant after applying a Bonferroni correction are shown in bold.

Table A.2 G-tests for temporal variation at repeated sample sites

Mosquito Harbor						
Year	RR	RS	SS	G-value	DF	p-value
2007	6	12	4			
2012	3	13	2	2.385	2	0.303

Naskeag Harbor						
Year	RR	RS	SS	G-value	DF	p-value
2008	18	7	4			
2012	7	14	6	14.941	2	**0.00057**

Northport Harbor						
Year	RR	RS	SS	G-value	DF	p-value
2010	0	1	51			
2013	0	3	101	0.468	2	0.792

Poorhouse Cove						
Year	RR	RS	SS	G-value	DF	p-value
2009	18	59	14			
2012	20	62	27	6.457	2	0.04

Raccoon Cove						
Year	RR	RS	SS	G-value	DF	p-value
2007	9	8	5			
2012	16	25	9	3.844	2	0.146

Data from the repeated sample sites in Chapter 2 showing the years each site was sampled, the number of resistant (RR) individuals, the number of heterozygous (RS) individuals, and the number of sensitive (SS) individuals. The G-values for each comparison were calculated using the initial genotype proportions from the first sampling as the expected genotype proportions for the second sampling (null hypothesis is that there is no change

between samplings). P-values that are significant after applying a Bonferroni correction are shown in bold.

BIOGRAPHY OF THE AUTHOR

Jennifer (Couture) Phillips was born in Concord, Massachusetts on January 15, 1986. She was raised in Litchfield, New Hampshire, and attended Campbell High School where she was a member of the National Honor Society and the Tri-M Music Honor Society. She graduated in 2004, and received a UMaine Presidential Scholarship to attend the University of Maine. While at the University of Maine she was a member of the Honors College, and defended an honors thesis as part of the requirements for a degree with honors. She graduated magna cum laude in 2008, receiving a Bachelor's degree in Marine Science. She returned to the University of Maine in the fall of 2010 and entered the Marine Biology graduate program. She has been a member of the International Society for the Study of Harmful Algae (ISSHA) since 2010, and has presented her research at national and international venues. Jennifer is a candidate for the Doctor of Philosophy degree in Marine Biology from the University of Maine in August 2016.