Stony Brook University

The role of developmental mode and population-level processes in shifting distributions: a study of calyptraeid gastropods

A Dissertation Presented

by

Abigail Eileen Cahill

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Ecology and Evolution

Stony Brook University

December 2014

Stony Brook University
The Graduate School

Abigail Eileen Cahill

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation.

Jeffrey S. Levinton, Dissertation Advisor
Distinguished Professor, Department of Ecology and Evolution

Dianna K. Padilla, Chairperson of the Defense
Professor, Department of Ecology and Evolution

Douglas J. Futuyma
Distinguished Professor, Department of Ecology and Evolution

Walter F. Eanes
Professor, Department of Ecology and Evolution

Frédérique Viard
Research Director, Centre National de Recherche Scientifique
Station Biologique de Roscoff (France)

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

The role of developmental mode and population-level processes in shifting distributions: a study of calyptraeid gastropods

by

Abigail Eileen Cahill
Doctor of Philosophy
in
Ecology and Evolution
Stony Brook University
2014

In benthic marine invertebrates with complex life cycles, dispersal happens primarily during the larval stage: sedentary, benthic adults produce larvae that live in the water column for a period of hours to months. The larval life stage is therefore critical to the ability of a species to expand or shift its range, including in response to climate change. Species with planktonic larvae are expected to show greater dispersal and gene flow than those with direct development. In contrast, reduced gene flow among populations of directly developing species increases the potential for local adaptation. Understanding not only larval traits, but population-level processes, is critical to predict how individual species may shift their ranges in response to climate change. The genus *Crepidula* (Gastropoda: Calyptraeidae) contains many species with different developmental modes.

In this dissertation, I used a planktonic developer (*C. fornicata*) and a direct developer (*C. convexa*) to assess genetic diversity and dispersal potential across larval types. Using microsatellite markers, I found that *C. convexa* has high levels of genetic structure within its native range, unlike *C. fornicata*. An introduced population of *C. convexa* showed high levels of genetic diversity, concurrent with previous results in *C. fornicata*. I used next-generation sequencing data to test the hypothesis that marginal populations of both species show reduced genetic diversity. I did not find a reduction in genetic variation, but found that marginal populations of *C. fornicata* are distinct from each other and from more central populations. I then focused on *C. fornicata* for a study of how larval settlement changes in response to adult density. I found that increased adult density leads to increased settlement, meaning that range edges may show low rates of settlement. The final chapter of this dissertation focused on the mechanism of settlement in *C. fornicata*. I conducted lab experiments to measure settlement in the presence of chemical cue from adult conspecifics. These experiments showed that larvae settled in response to two different cues from adults, a waterborne cue and one carried in adult pedal mucus, and that the waterborne cue could be partially inactivated by heat. Taken together, these studies represent an understanding of ecological and genetic processes that may affect the speed of climate-driven range shifts in these *Crepidula* species.

Dedication Page

To my parents, Dennis and Jeanine, who took this Ohio girl to see marine life as often as possible, to my brother Teddy, who had to come along, and to my grandmother Laverne, who told me I could do anything.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

The following abbreviations have been devised for this dissertation.

List of collecting sites, Chapter 2:

BA (Barnstable, Massachusetts)
DE (Lewes, Delaware)
JB (Jamaica Bay, New York)
NA (Nahant, Massachusetts)
NC (Beaufort, North Carolina)
PB (Padilla Bay, Washington)
VA (Gloucester Point, Virginia)
WM (Old Field, New York)

List of collecting sites, Chapter 3:

BA (Barnstable, Massachusetts)
NA (Nahant, Massachusetts)
NJ (Sandy Hook, New Jersey)
NL (Port Saunders, Newfoundland)
NS (Main-à-Dieu, Nova Scotia)
NY (Northport, New York)

List of factors inducing settlement, Chapter 5:

ACW (Adult-conditioned water)
PMG (Pedal mucus glasses)

Acknowledgments

This dissertation, seven years in the making, would not have been possible without many people and funding sources.

First, I owe thanks to my advisor, Jeffrey Levinton. On a rainy morning in February 2007, Jeff told me that we would find a way to get me to work in France, and we have managed that several times over. He has been patient with me as I stubbornly tried idea after idea, usually only to come around to what he suggested in the first place.

My committee has been a continuous source of support, encouragement, and motivation to do the best science that I can. Dianna Padilla has challenged me to be a better scientist and taught me much of what I know about marine invertebrates. Doug Futuyma's courses influenced me early on, and he has continued to direct my dissertation towards interesting and timely topics. Walt Eanes taught me population genetics and has gently edited my manuscripts to fix the things I forgot. Lastly, Frédérique Viard took me into her lab based on a lengthy email from an unknown American Ph.D. student begging to learn PCR – her generosity with her resources and her time have been critical in producing this dissertation, and she is the coauthor on the work presented as Chapter 2.

I received help in the lab and the field from a series of hardworking Stony Brook undergraduates: Khosnur Alam, W. Tom Bruno, Priyanga Jeyasri, Alia Juman, Aaron Pellman-Isaacs, Tom Ralicky, Kaitlyn Smith, Debbie Tan, John Tay, Annamaria Viazzo, and Annastacia Woytash.

I received sampling advice and help at various sites from Rachel Allee, Jim Browne, Daniel Duval, Ron Etter, Pat Gaffney, Matt Harke, Roger Mann, Dugald McGlashan, Jan Pechenik, Tim Rawlings, Paul Sargent, and Zach Siegrist.

Lab help for Chapter 2 was provided by Marion Ballenghien, Claire Daguin-Thiébaut, Daphné Grulois, Stéphane Mauger, Florentine Riquet, and the staff of the BioGenouest sequencing platform at the Station Biologique de Roscoff. Lab help for Chapter 3 was provided by the labs of Walt Eanes and Jackie Collier (especially Sabrina Geraci-Yee), as well as Omar Warsi. Sequencing was done at Cornell University's Institute for Genome Diversity, and I owe much to the help of all of the staff there, especially Sharon Mitchell.

Funding for Chapter 2 was provided by the American Malacological Society Melbourne R. Carriker Student Research Grant, the National Science Foundation International Research Experience for Students for work in the team Diversity and Connectivity in Coastal Marine Landscapes in Roscoff, and the Agence Nationale de la Recherche (project MIRAGE; no. ANR-05-BLAN-0001) to my coauthor, F. Viard.

Chapter 3 was funded by a National Science Foundation Doctoral Dissertation Improvement Grant. Resources from the iPlant Collaborative contributed to the research results reported in this chapter. The iPlant Collaborative is funded by a grant from the National Science Foundation (#DBI-0735191). URL: www.iplantcollaborative.org.

Additionally, my work has been supported by a National Science Foundation Graduate Research Fellowship, grants from the Conchologists of America, a Grant-in-Aid-of-Research from the Society of Integrative and Comparative Biology, and a Learner Grey Fellowship from the American Museum of National History.

Graduate school has not simply been the production of the science included here. I owe thanks to my labmate Pat Lyons, the Padilla lab (especially Paul Bourdeau, Sarah Gray, Alex Hooks, Mike McCann, Mica McCarty-Glenn, Laurie Perino), my cohort (Ben Greene, Xia Hua, Caitlin Karanewsky, Adam Laybourn, Allison Onstine, Jessie Stanton), and many friends and faculty members who provided scientific and emotional advice and encouragement through the years: Mary Alldred, Richard Allen, Stephen Baines, Karen Chan, Ryan Crim, Richard Emlet, M. Caitlin Fisher-Reid, Catherine Graham, Licinia Gouveia, Stacy Hadfield, Emily Herstoff, Aleya Kaushik, Erik Lavington, Joe Lachance, Joan Miyazaki, Dan Moen, Chris Morales, Marvin O'Neal, Dana Opulente, Josh Rest, Jim Rohlf, Jen Rollins, Emily Rollinson, Steve Sabatino, Fab Spagnolo, Richard Strathmann, Anna Webb, Beth Whitehill, John Wiens, and many more.

I have had the pleasure of collaborating particularly closely with several graduate students during my time at Stony Brook. J. Matt Hoch showed me how to conduct field research, how to submit a scientific paper, and generally how to enjoy working in the lab. Kelly O'Donnell suffered through the pedagogical literature and endless rounds of edits with me. Matt Aiello-Lammens helped me organize a large collaborative project and kept me sane while doing it. And Spencer Koury has been an invaluable collaborator and coauthor on the work that has become Chapter 5 of this dissertation.

Finally, I owe thanks to my family. My parents, Dennis and Jeanine, have supported my interest in marine biology since I was very small. My brother (Teddy) has been along for the whole ride, and my grandmother (Laverne) continues to encourage me to pursue my goals, wherever they take me. Many season passes to Sea World and trips to the Oregon Coast later, I am thrilled to dedicate this dissertation to them.

Statement of Permission for Copyrighted Work

I hereby state that permission has been granted by Springer-Verlag Berlin Heidelberg to reproduce the text of Chapter 2, "Genetic structure in native and non-native populations of the direct-developing gastropod Crepidula convexa," which I wrote with Frédérique Viard (who directed and supervised the research) and was published in *Marine Biology* in 2014.

Chapter 1: Introduction

As global climate continues to change, predictions of how populations and communities will respond are becoming more important to preserve biodiversity. Although there are correlations between species' responses (e.g., local extinctions, range shifts) and changing climate, there are few cases where species have been studied in enough detail to identify a mechanism linking climate change to these responses (reviewed in Cahill et al., 2013; Cahill et al., 2014). By studying population-level processes in more detail and understanding the mechanisms involved, scientists can better predict how populations, species, communities, and ecosystems are likely to change in response to climate over the coming decades.

One of the ways in which many species respond to climate change is to shift their geographic ranges, tracking their climatic niches. Many species responded to the climate variation during the Pleistocene with shifting ranges as glaciers advanced and retreated (Hampe and Petit, 2005; Hampe and Jump, 2011), and dozens of species are reported to be shifting their ranges in response to current anthropogenic climate change (e.g., Parmesan & Yohe, 2003; Hickling et al., 2006; Parmesan, 2006; Anderson et al., 2009; Thomas, 2010; Chen et al., 2011). Although the fundamental form of a bioclimatic niche includes only climatic variables (i.e., temperature and precipitation variables; Pearson and Dawson 2003), other environmental variables may be important determinants of distributions in both terrestrial (e.g., soil type) and marine (e.g., salinity) species. Understanding which of these variables may limit species' ranges can improve our ability to predict those changes.

Attempts to use species-level traits (e.g., body size and diet breadth in North American birds) to predict which species will be able to track their climatic niche have been only marginally successful (Angert et al., 2011). Understanding the mechanisms determining range

limits is important to more accurately predict shifts (Cahill et al., 2014). For instance, the ranges of several plant species have shifted their range towards warmer temperatures, indicating that precipitation patterns are more important than temperature in setting the geographic limits of these species (Crimmins et al., 2011; Harsch and Hill Ris Lambers, in press).

Another way in which species may respond to global change is by evolutionarily adapting to new conditions. A species' ability to do this will depend on things like additive genetic variation for relevant traits (Visser, 2008; Hoffmann and Sgrò, 2011), genetic correlations and constraints among traits (Etterson and Shaw, 2001), and the strength of selection relative to forces like genetic drift and migration from populations with alleles that are not beneficial under new conditions (Kremer et al., 2012). Another factor to consider is the strength and variation of selective pressures imposed by changing conditions. Climate variability is expected to increase (Thornton et al., in press), and responses to selection in a variable environment are expected to be different than those under constant selective pressure (e.g., Bennet et al., 1992). The two general processes of movement and adaptation will be of different relative importance in different species. For instance, a species with low dispersal ability is expected to shift its range more slowly, but may be more able to adapt to new conditions than a species with continuous gene flow from populations adapted to other conditions (Kawecki, 2008; Kremer et al., 2012).

In order to understand and predict range shifts in any given species, it is therefore important to understand many aspects of the population biology and natural history of that species. Although macroecological metrics like range size in plants (Slayter et al., 2013) or climate velocity (i.e., how fast climate changes in an area; Loarie et al., 2009) within assemblages of marine fish species can successfully identify patterns of range shifts in groups of

species, they often poorly predict range shifts of individual species (e.g., Pinsky et al., 2013). Morphological and life history traits that are not closely linked to mechanisms of response to climate change are also poor predictors (Angert et al., 2011). However, detailed species-level studies which include mechanisms thought to impact response to climate change do have high predictive power (e.g., thermal performance curves in *Mimulus*, Angert et al., 2011; additive genetic variation, Shaw and Etterson, 2012). Additionally, recent work suggests that using genetic data in addition to ecological information may improve predictions of range shifts with climate change (Fordham et al., 2014).

Benthic marine invertebrates with coastal distributions are interesting cases in which to test how range shifts or adaptation to climate change may occur. Many of these species are sessile or semi-sessile as adults. Therefore, dispersal and gene flow among populations takes place during larval or juvenile life stages. A general dichotomy exists between species with long-lived planktonic larvae and corresponding long-distance dispersal potential and those with direct-developing larvae lacking a planktonic stage and have correspondingly limited dispersal potential (Weersing and Toonen, 2009; Kelly and Palumbi, 2010; Selkoe and Toonen, 2011). Comparing closely related groups of species with different developmental modes and dispersal potential can elucidate the role of dispersal in determining genetic and ecological population differentiation. Many such comparisons of genetic differentiation exist in the literature (e.g., Hellberg, 1996; Collin, 2001; Kyle and Boulding, 2000; Lee and Boulding, 2009; Arndt and Smith, 2002; McMillan et al.; 1992).

One group of species with different developmental modes is the calyptraied gastropods, particularly within the genus *Crepidula*. Within the genus there are many examples of both planktotrophic and direct developing species (Collin, 2003). Species of *Crepidula* have also

served as the basis for a wide variety of scientific investigation, including studies of development (e.g., Collin 2003), population genetics (e.g., Dupont and Viard 2003, Daguin et al. 2007, Riquet et al., 2013), range limits (Rawlings et al., 2011), among other subjects (reviewed in Henry et al., 2010). Three species live in sympatry on the east coast of North America: the planktotrophic *C. fornicata* and *C. plana*, and the direct-developing *C. convexa* (Collin, 2001). There is evidence that *C. fornicata* is expanding its range northward over the last few decades due to changes in water temperature (Rawlings et al., 2011), and this northward expansion is predicted to continue (Saupe et al., 2014). Complementary historical and distribution information is not available for *C. convexa*.

This dissertation investigated the population genetics of *C. fornicata* and *C. convexa*, and larval settlement in *C. fornicata*, in an attempt to understand how evolutionary and ecological processes differ with developmental mode and may subsequently affect adaptation or range shifts in response to global change.

In chapter 2, I present a population genetic analysis of *C. convexa* populations over a wide geographic range based on five microsatellite loci. I demonstrated that seven native populations of this direct-developing species show a pattern of significant isolation-by-distance, consistent with previous results using other markers. These results are contrasted with the low population structure found in the planktotrophic *C. fornicata* over the same geographic range (Riquet et al., 2013). I also present population genetic data from an introduced population of *C. convexa* (Washington, U.S.A.) to evaluate the hypothesis that genetic diversity is reduced following an introduction and attempt to identify the source of introduction.

In chapter 3, I used high-throughput, next-generation sequencing data (SNPs) taken from *C. fornicata* and *C. convexa* to test the hypothesis that populations at the northern range margin will have reduced genetic variation relative to central ones. I also evaluated genetic structure in these species and interpreted it relative to the microsatellite results in Chapter 2 and results from other marker sets (Hoagland, 1985; Collin, 2001; Dupont et al., 2003; Riquet et al., 2013). *Crepidula fornicata* has not previously been sequenced at its northern range margin, nor has population structure been found in this species using other markers (i.e., all populations have the same genetic composition; Collin, 2001; Riquet et al., 2013). I found no support for the hypothesis of reduced variation at range margins, but marginal populations of *C. fornicata* are differentiated from each other and from more central populations.

I then investigated how larval settlement may impact expansion of range margins for *C. fornicata* in Chapter 4. This species is a protandrous hermaphrodite, has sessile adults, and has internal fertilization. Larvae must therefore settle out of the plankton and enter the benthic population near conspecifics for successful reproduction as adults. I manipulated the density of adult conspecifics in the field and measured larval settlement. I found that settlement increased with increasing adult density. Given this, changes of the range margin of *C. fornicata* may be limited by the ability of larvae to detect conspecifics using waterborne cues.

Finally, in Chapter 5 I focused on waterborne cues used by larvae to detect adults. Using a series of laboratory experiments, I designed a bioassay to measure larval settlement in response to cues from adult conspecifics. This study confirmed settlement in response to a conspecific waterborne cue (e.g. Pechenik and Lima, 1984) and discovered the additional effect of adult pedal mucus as a cue for settlement. I also found that the waterborne cue was partially inactivated by heat.

Taken together, these chapters comprise a better understanding of population-level processes in the two focal *Crepidula* species and how these processes may determine a species-level trait (i.e., the geographic range limit). Although these studies are not sufficient to predict the existence or extent of a northward range expansion in these species, the data they provide can be used to generate better predictions of changing species distributions as climate changes.

Chapter 2: Genetic structure in native and non-native populations of the direct-developing gastropod *Crepidula convexa*

2.1 Introduction

The spatial structure of genetic diversity can impact several ecological and evolutionary processes of a species, including its ability to adapt to local conditions. In most benthic marine invertebrates, larvae and juveniles are primarily responsible for dispersal because adults are sessile or have restricted movement (reviewed in Cowen and Sponaugle 2009). Therefore, the movement of genes among populations is mainly determined by the dispersal of early life-history stages in these species. Understanding how these early stages structure genetic diversity can illuminate how such diversity is created and maintained in marine invertebrates.

Due to the small size of most marine invertebrate larvae or juveniles, their frequently long planktonic larval durations (PLD), and the difficulty of following early stages in the ocean, most studies of dispersal in these species rely on proxies (but see Olson and McPherson 1987), including identification based on elemental fingerprinting (e.g., López-Duarte et al. 2012), genetic markers (e.g., Kinlan and Gaines 2003; Shanks 2009), and oceanographic models coupled with PLD (e.g., Siegel et al. 2003; Treml et al. 2012).

PLD varies widely in marine invertebrates, ranging from direct-developing species that spend no time in the plankton (PLD = 0 days) to species that live in the plankton for years (e.g., *Fusitriton oregonensis*, PLD = 4.5 years; Strathmann and Strathmann 2007). Species with longer PLDs are expected to disperse over longer distances, and PLD is often assumed to be correlated with spatial genetic structure, although with numerous exceptions (Shanks 2009). This structure is often measured as subdivision among populations, usually by using Wright's fixation index (F_{st}). Some authors have found a relatively strong negative correlation between

F_{st} and PLD among taxa (e.g., Siegel et al. 2003), while others have found a weak correlation, if any (e.g., Bradbury et al. 2008; Weersing and Toonen 2009; Kelly and Palumbi 2010; Selkoe and Toonen 2011). However, when comparing closely related taxa, direct-developing species have more highly structured populations than planktonic species (Weersing and Toonen 2009; Kelly and Palumbi 2010; Selkoe and Toonen 2011). Examples include solitary corals (Hellberg 1996), molluscs (*Crepidula* spp., Collin 2001; *Littorina* spp., Kyle and Boulding 2000; Lee and Boulding 2009), and echinoderms (*Cucumaria* spp., Arndt and Smith 2002; *Heliocidaris* spp., McMillan et al. 1992).

The gastropod genus *Crepidula* includes many planktotrophic and direct-developing species (Collin 2003) and has been the subject of many studies relating larval development mode to genetic differentiation within and among regions (e.g., Hoagland 1984, 1985; Collin 2001; Dupont et al. 2003; Riquet et al. 2013). The northern portion of the east coast of North America has three native, sympatric *Crepidula* species: the planktotrophic *C. fornicata* and *C. plana*, and the direct developing *C. convexa* (convex slipper limpet; Collin 2001). *C. fornicata* has a broader range than *C. convexa*, with the latter species' southern range limit occurring in the southeastern United States where it is parapatric with its lecithotropic sister species, *C. usutulatulina* (Collin 2002). As expected from its limited dispersal, *C. convexa* has more spatial genetic structure in a mitochondrial marker, cytochrome oxidase I (COI), than the planktotrophic species over a similar geographic range covering the distribution of the two species (Collin 2001), as predicted considering development mode and its effects on expected dispersal. Populations of *C. fornicata* analyzed using 17 microsatellites and 327 Amplified Fragment Length Polymorphism markers (AFLPs) have no detectable genetic structure along the east coast of North America (Riquet et al. 2013), although there are differences between Atlantic coast and

Gulf of Mexico populations as measured with COI (Collin 2001) as well as with faster-evolving markers like microsatellites and AFLPs (Riquet et al. 2013). However, *C. convexa* has not been analyzed with a larger marker set, or with markers from faster-evolving classes.

Spatial genetic structure and gene genealogies are also influenced by population size and demographic history (Marko and Hart 2011). Population genetic theory predicts that when a population experiences a major reduction in population size, it will show an overall reduction in genetic diversity with a concomitant increase in structure among sub-populations (Wright 1931; Nei et al. 1973). Since introduced species are assumed to go through founding events (Sakai et al. 2001), it is also expected that they will show a reduction in some measures of genetic diversity (e.g., allelic richness) relative to their native range (Barrett and Kohn 1991; Sakai et al. 2001; Allendorf and Lundquist 2003). Although reduced diversity has been found in several cases (e.g., the Chinese mitten crab, *Eriocheir sinensis*, Herborg et al. 2007), the observation is far from ubiquitous (Novak and Mack 2005; Roman and Darling 2007). In the case of aquatic invasions, genetic diversity is often not decreased in introduced populations (e.g., *Crepidula fornicata*, Riquet et al. 2013), and is sometimes maintained or even increased by admixture between colonizers originating from genetically distinct populations (Rius and Darling 2014; e.g., the gastropod *Cyclope neritea*, Simon-Bouhet et al. 2006).

The direct developing *C. convexa* has been introduced from the east coast of North America to a few bays on the west coast of North America: Boundary Bay, British Columbia (Carlton 1992), Padilla Bay, Washington (Wonham et al. 2005; Collin et al. 2006), and San Francisco Bay, California (Carlton 1992, McGlashan et al. 2008). The presumed vectors of multiple introductions were transplants of the eastern oyster, *Crassostrea virginica* (Carlton 1992; Wonham et al. 2005). These populations are hundreds of kilometers apart and have

remained isolated. Given the low dispersal distance of this species (Collin 2001), it is unlikely that gene flow has occurred among west coast subpopulations following the introduction(s).

The planktotrophic congener of *C. convexa*, *C. fornicata*, was first introduced to the west coast of North America (Hoagland 1977) and to Europe (Great Britain) from the east coast of North America in the late nineteenth century (Blanchard 1997). It has become the subject of many ecological (e.g., Thieltges et al. 2006 and references therein) and genetic studies (e.g., Hoagland 1985; Dupont et al. 2003; Viard et al. 2006; Riquet et al. 2013). Hoagland (1985) and Dupont et al. (2003) found no divergence between North American and European populations using allozymes, and no reduction in genetic diversity in the introduced populations. This observation was later reinforced with microsatellites and a genome scan using AFLPs (Riquet et al. 2013). *C. fornicata* was almost certainly introduced multiple times from different source populations (Hoagland 1985; Blanchard 1997) and its PLD of 2-4 weeks (Ament 1979; Henry et al. 2010) should be sufficient to allow immediate gene flow among introduced subpopulations.

This study examines the genetic structure of native populations of *C. convexa* and compares it to that found for *C. fornicata.* To expand the analysis from Collin (2001), which was done with only a single mitochondrial locus, we investigated the distribution of the genetic diversity of populations of *C. convexa* using six microsatellite loci. These findings are further discussed in light of the results obtained for *C. fornicata* by Riquet et al. (2013) with microsatellite loci to investigate the effects of developmental mode. We predicted that patterns distinct from those described for *C. fornicata* would be observed, particularly that native populations of *C. convexa* would show a significant pattern of isolation-by-distance along the east coast of North America. We also analyzed genetic diversity at these loci in an introduced

population of *C. convexa* from the west coast of North America and compared it to one population of *C. fornicata* introduced to a nearby site in the same region.

2.2 Methods

Crepidula convexa were collected from seven native populations along the Atlantic coast of the United States and from one introduced population in Padilla Bay, Washington, USA. Individuals were collected from many different substrates (e.g., razor clams, scallops, pebbles) at each site, but multiple slipper limpets were usually collected from each substrate. Collections were made by 1-2 people at each site, along at least a 100 m transect; the collecting area was larger in populations that were less dense. See Table 2.1 for further collection information and keys to locality abbreviations. Slipper limpets were removed from their substrate and preserved in 95 percent ethanol in the field for subsequent genetic analyses. The samples were collected in different years (Table 2.1), but given the lifespan of the snails (at least two years; Hendler and Franz 1971) and the strong spatial differentiation among populations, discrepancies in the time of collection are unlikely to have influenced the results. Direct-developing species are expected to show greater temporal stability in spatial genetic structure than planktonic ones at both neutral and non-neutral loci (Lee and Boulding 2009).

DNA was extracted using a NucleoSpin® 96 Tissue kit (Machery-Nagel) following the manufacturers' instructions. Samples were genotyped at six microsatellite loci, with amplification following PCR protocols detailed in Daguin-Thiebaut et al. (2009). These loci are known to be highly polymorphic in *C. convexa* (Daguin-Thiebaut et al. 2009) Forward primers were labeled with IRD700 or IRD800 infrared dyes and loci were analyzed using a 6.5% polyacrylamide gel with a LI-COR NEN Global IR2 DNA Analyzer (Daguin-Thiebaut et al. 2009).

Genotype data were analyzed for Hardy-Weinberg equilibrium within populations with Genepop 4.0 (Rousset 2008). Global and pairwise F_{st} among populations and isolation-by-distance (IBD) calculations were performed with Genepop on the Web (Rousset 2008). Isolation-by-distance calculations are performed by the program according to the procedure in Rousset (2000) based on Mantel tests between matrices of pairwise geographic differences and pairwise genetic differences among populations and the genetic distance is calculated using $F_{st}/(1-F_{st})$. The Padilla Bay population is introduced and thus cannot be connected to the other populations through natural dispersal. It was not included in the isolation-by-distance analysis. All the F_{st} value estimates were computed on the original dataset as well as a modified dataset with correction for null alleles. This latter dataset was obtained by estimating the frequency of null alleles in the dataset (i.e., missing data) using the Expectation Maximization (EM) algorithm of Dempster et al. (1977; cited in Chapuis and Estoup 2007), implemented in the software FreeNA (Chapuis and Estoup 2007); population genetic metrics were then recalculated based on the estimated dataset. Calculations of within-population gene diversity (a measure of the expected heterozygosity in a randomly mating population), allelic richness (a measure of the number of alleles in a population that is corrected for sample size using rarefaction; El Mousadik and Petit 1996), and F_{is} were conducted with FSTAT 2.9.3.2 (Goudet 2002).

The use of F_{st} with microsatellites has been questioned in the literature (Hedrick 1999, Meirmans and Hedrick 2011), particularly because its maximum value is reduced with highly polymorphic markers (Jost 2008, Whitlock 2011). Although F_{st} allows for a direct comparison with older literature (e.g., Weersing and Toonen 2009 and references therein), we also calculated Jost's D statistic (measuring the fraction of allelic variation among populations rather than deviations from panmixia; Jost 2008) using GenAlEx 6.5 (Peakall and Smouse 2012).

Genetic assignment tests, in which individuals are assigned probabilities of belonging to a population of origin using multilocus genotypes, were conducted using maximum likelihood methods (Rannala and Mountain 1997) with GeneClass 2.0 (Piry et al. 2004). Assignment tests are based on the alleles displayed by a given individual as compared to allelic frequencies of potential source populations, and can detect potential sources even in individuals that were introduced many generations before sampling. In these tests, an individual is assigned to a population based on its entire multilocus genotype; the test does not partition a genotype such that portions of it get assigned to different populations. A first analysis was conducted on only native (east coast) populations of *C. convexa* to verify the reliability of assignment tests for this species (i.e., individuals are assigned back to the populations in which they were collected with a high degree of confidence). A second analysis was conducted to assign individuals from the introduced PB population to native populations.

To investigate both the possibility of admixture from multiple sources in the PB population and the overall genetic structure without *a priori* assumptions about the population definition in the study species, we performed a clustering analysis with all of the data. However, the commonly-used program STRUCTURE performs poorly in situations with significant IBD (Pritchard et al. 2000). Instead, we conducted a clustering analysis without including the population of origin as a prior in the model, using a discriminant analysis of principal components (DAPC; Jombart et al. 2008, 2010) on all populations. DAPC was conducted with the ADEGENET 1.3-4 package in R 3.0.1 (R Core Development Team 2013).

2.3 Results

All loci were successfully amplified in all populations, except for CT5H8, which did not amplify in the NC population. Results are therefore presented for only five loci. The addition of

the sixth locus did not qualitatively change the results, and analyses of six loci in all populations except NC are presented in Table S.2.1. Populations that were collected at the same time did not show different patterns from the overall dataset (i.e., those collected in 2002 are not more closely related to each other than would be expected based on geography).

Genetic diversity estimates are summarized in Table 2.1. The average expected heterozygosity (H_e) across the five loci and seven native populations was 0.919 (95% confidence interval = 0.903 – 0.935, bootstrapped over all loci). The PB west coast population had an average H_e of 0.888 (95% confidence interval = 0.815 – 0.961). All populations and all loci showed deviations from Hardy-Weinberg equilibrium (HWE) due to an excess of homozygotes at all loci in all populations. The exception was the northernmost population, NA, which was in HWE at all loci (Table 2.1); it is worth noting that this is the population from which the microsatellites were developed and selected (including selection for HWE; Daguin-Thiebaut et al. 2009). The F_{is} values for all populations except NA were significantly greater than zero. The mean F_{is} for all native populations was 0.119 (95% confidence interval = 0.059 – 0.169). In the introduced PB population F_{is} = 0.101 (95% confidence intervals = 0.016 – 0.186), indicating that heterozygote deficiencies in this population are as large as those of the native populations, for instance due to a similar level of null alleles or to a Wahlund effect created by admixture between genetically differentiated source populations. The average richness (alleles per locus in each population, corrected for sample size) across all native populations and five loci was 13.24 (95% confidence interval = 12.36 – 14.12), and average richness of the PB population was 12.19 (95% confidence interval = 9.39 – 14.99). Both genetic diversity and allelic richness were highest in the central populations of the native range and decreased toward the marginal populations, especially at the northern edge (Table 2.1).

A significant overall genetic structure was observed among native populations (F_{st} = 0.017; p < 0.001; F_{st} = 0.018; p < 0.001 with correction for null allele frequencies; D_{est} = 0.422; p = 0.001). The average pairwise difference among native populations was F_{st} = 0.035; D_{est} = 0.422. As was true for the global F_{st}, taking into account a null allele class (based on Chapuis and Estoup 2007) changed the pairwise F_{st} values only slightly (data not shown). The pairwise matrices of corrected and uncorrected data are strongly correlated (Mantel test, r = 0.97, p < 0.001) and using the corrected dataset did not change the results of any statistical test. Only values obtained with the original dataset are thus provided in the following text. Pairwise comparisons of F_{st} of native population localities ranged from 0.006 (WM to JB) to 0.058 (BA to NC; Fig. 2.1a, Table S.2.2). This variation was partly explained by the geographic distance separating the populations: the native populations showed a significant pattern of isolation by distance (IBD slope = $2.80*10^{-5}$; Mantel test, r = 0.803, p = 0.007; Fig. 2.1a). Pairwise D_{est} values ranged from 0.096 (WM to JB) to 0.670 (BA to NC).

When including the Padilla Bay population in these genetic structure analyses, the global F_{st} value was slightly higher than only among native populations, and was significantly greater than zero (F_{st} = 0.021, p < 0.001; D_{est} = 0.458, p < 0.001) . The average pairwise F_{st} between the introduced PB population and any native population (0.049) was nearly as high as the largest pairwise F_{st} within the native range (Fig. 2.1a, Table S.1.2; average D_{est} between native populations and PB = 0.540). The same patterns were observed both comparing F_{st} values computed with the ENA method that takes null allele frequencies into account and comparing D_{est} values (Table S.2.3).

Self-assignment tests conducted in GeneClass showed an average of 96.8% successful self-assignment across all native populations, with no individuals rejected from their population

of origin (Table 2.2). In NA, JB, and VA, 100% of the individuals were correctly assigned to the population where they were collected. These tests also show that the proportion of individuals rejected from a tested source population with 95% probability increases when the spatial distance between the targeted population and the tested source increases in southern (values above diagonal in Table 2.2) or northern (values below diagonal in Table 2.2) directions in nearly all pairwise comparisons. In some cases, all individuals from a sampling location were rejected from a putative population of origin (i.e., a rejection value of 1.00; e.g., all the individuals collected in NC were rejected when tested for their assignment to NA or BA). This assignment analysis is in agreement with the significant genetic structure and isolation by distance observed among native populations. The introduced snails from PB were primarily assigned to the WM (40.0% of individuals) and DE (32.9% of individuals) populations (Fig. 2.2). Nearly all PB individuals were rejected from the populations located in the southern part of the range (VA and NC), as well as the two northernmost populations (NA and BA); these localities were thus excluded as the source for the PB populations (Fig. 2.2).

DAPC identified $k = 6$ as the most likely number of clusters in the dataset (Fig. 2.3). There was some clustering based on geography (e.g., the two populations from New York, WM and JB, show similar patterns; Fig. 2.3). The PB population, although different from all others, was more similar to the populations from New York (WM and JB). This corresponds to the GeneClass analysis that identified the WM population as one of the likely sources of introduction for the PB population.

2.4 Discussion

2.4.1 Isolation-by-distance patterns in *Crepidula* species

Native populations of *C. convexa* show a strong pattern of isolation-by-distance (IBD). The overall F_{st} in the dataset (0.018) is comparable to others found for gastropods measured with microsatellites, including both direct-developing and planktotrophic species (reviewed in Weersing and Toonen 2009). Marker type is known to be influential on F_{st} values, and microsatellites generally yield lower estimates of geographic differentiation than mtDNA or allozyme markers, making a within-marker comparison between close relatives the most valuable here (Weersing and Toonen 2009). A high level of spatial genetic structure was expected given the direct-developing life history of *C. convexa*. This pattern has been seen in other comparisons of closely related groups, including conspecific gastropods (Lee and Boulding 2009). Collin (2001) found that 76% of the total molecular variance in *C. convexa* COI sequences was found among populations, also corresponding to strong among-population structure.

In contrast, only 22% of total molecular variance was among populations in the planktotrophic *C. fornicata* (Collin 2001). When measured with 17 microsatellite markers, *C. fornicata* showed an F_{st} value of 0.011 (Fig. 2.1b; data from Riquet et al. 2013), less than that of *C. convexa*, over approximately the same geographic range (although this value increased to 0.026 when the marginal population from Long Boat Key (FL, Gulf of Mexico) was included in the analysis; Riquet et al. 2013). For a more direct comparison between species and because the maximum value of F_{ST} becomes quite small with polymorphic markers (Hedrick 1999, Jost 2008), Jost's D statistic was computed over the 5 populations of *C. fornicata* covering the northern portion of the range (populations 1-5 in Riquet et al. 2013), which overlap with the range covered in the present study. Since D is not bounded by maximum values below 1, the difference between the two species was increased: $D_{est} = 0.422$ ($p = 0.001$) and 0.023 ($p = 0.001$)

in *C. convexa* and *C. fornicata*, respectively. Note that the difference persists when including two genetically divergent populations of *C. fornicata* located in the southern part of the distribution range (from Florida, Fig. S.2.1; pop 6-7 in Riquet et al. 2013) with a $D_{est} = 0.071$ ($p = 0.001$).

The difference between the two species is reinforced by the IBD model. A clear and significant IBD pattern and high pairwise F_{st} values were observed in *C. convexa*, but no IBD and low pairwise *F*st values were observed in *C. fornicata* over the same range. A Mantel test between the two pairwise F_{st} matrices computed with populations sampled in close proximity for each of the two species (NA, BA, WM, DE, VA vs. MA (1), MA (2), NY (3), NJ (4), and VA (5) in Riquet et al. 2013; Fig. S.2.1) showed that the null hypothesis of no correlation between them could not be rejected ($p = 0.331$). In addition, we conducted a second IBD analysis of the *C. convexa* data eliminating the NC population to allow for complete concordance of sampling sites between the species. These results also showed significant IBD (IBD slope = 3.5×10^{-5}; Mantel test, $r = 0.809$, $p = 0.017$), indicating that the importance of IBD is not solely due to one highly divergent population. These results are consistent with previous comparisons of population genetic structure in *Crepidula* spp. using COI (Collin 2001) and allozyme (Ament 1978; Hoagland 1984) markers over the global distribution range. This study confirms that the distribution of the genetic diversity of the two species is different and more specifically, shows that the two species display different patterns of genetic structure over a smaller scale than previously examined (i.e., a few hundred km).

2.4.2 Hardy-Weinberg Equilibrium in *C. convexa*

All populations of *C. convexa* showed deviations from Hardy-Weinberg equilibrium due to a deficiency of heterozygotes ($F_{is} > 0$). This pattern is often found in populations of marine invertebrates (Addison and Hart 2005 and references therein). It is not clear what mechanism is

responsible for the deviations in *C. convexa*. However, problems with null alleles can be corrected for in this dataset. Taking null alleles into account did not change the results concerning genetic structure, suggesting that null alleles, if present, have not changed the results of our analyses. This confirms previous results from a microsatellite-based study of parent–offspring array in this species (Le Cam et al. 2014). Other mechanisms may thus play a role, for instance spatial Wahlund effects; dispersal of *C. convexa* may be less than the sampling scale, and snails are not found on all possible substrates at a site.

2.4.3 Marginal populations

Although gene diversity and allelic richness (corrected for sample size) were similar across all populations, there was a slight decrease in both measures in the two northern populations (NA and BA; Table 2.1). The NA population is very near the northern range edge of the species (Collin 2001). Theory predicts that populations at range edges will show reductions in genetic diversity due to small population sizes, reduced connectivity of populations, or selection for adaptation to marginal habitats (Kawecki 2008). Although it is not possible to explain the reduction in genetic diversity in marginal populations with this dataset, the IBD analysis shows that pairwise F_{st} between marginal and central populations is not greater than among central populations (i.e., connectivity is not greatly reduced). In addition, *C. convexa* is at least as abundant at northern sampling locations than more geographically central ones (AEC pers. obs.), indicating that population sizes are not greatly reduced at the northern range edge. However, confirming the observation of high abundance and its effects on maintaining genetic diversity requires further research.

The effect of reduced diversity in marginal populations may be more pronounced in *C. convexa* than in its planktotrophic congener, *C. fornicata*. Riquet et al. (2013) did not observe a

reduction in genetic diversity in one population of *C. fornicata* located at the southern range edge (i.e., Long Boat Key, Gulf of Mexico, Florida; Fig. S.2.1). A decrease in genetic diversity measures was not observed in the southernmost *C. convexa* population (NC) in this study, but this population is farther from the southern range edge than NA is from the northern edge (Collin 2001; 2002), where a reduction in diversity was observed. In addition, gene exchange at the southern margin with *C. ustulatulina*, the parapatric sister species of *C. convexa* (Collin 2002), may have occurred recently enough to obscure any effects of the range margin. Notably, one locus (CT5H8) did not amplify in the NC population. The microsatellite markers used were developed based on individuals from the northernmost populations (NA and BA; Daguin-Thiebaut et al. 2009), and this failure of amplification may be a result of genetic divergence across the species' range promoted by both historical and environmental factors. Strong genetic divergence towards the southern range edge has, for example, been found in *C. fornicata* (Collin 2001; Riquet et al. 2013). In particular, a genome-scan approach revealed eight outlier loci (i.e., loci showing a higher genetic divergence than expected based on neutral process) in *C. fornicata*, pointing to a strong divergence at the tip of the Florida Peninsula near a transition between temperate and subtropical marine zones (Riquet et al. 2013). Such a pattern is likely to have evolved from the coupling between endogenous (e.g., reproductive incompatibilities) and exogenous (e.g., temperature gradients) barriers at environmental boundaries (Bierne et al. 2011). Further work is needed to understand evolutionary and ecological dynamics at range margins in *Crepidula* species, for instance with a genome scan approach or candidate genes to examine selective processes at the margin.

2.4.4 Similar patterns of genetic diversity in introduced populations of *Crepidula* species

The overall genetic diversity, measured by allelic richness and H_e, in the introduced population of *C. convexa* from Padilla Bay, WA is equivalent to that within the native region (Table 2.1). This pattern of little to no reduction in diversity in introduced populations is a common feature in marine environments (reviewed in Roman and Darling 2007). In the case of *C. convexa* this is probably due to high genetic diversity in native source populations, as shown in this study, potentially coupled with multiple introductions from different sources. Despite the lack of direct evidence that *C. convexa* was introduced multiple times to Padilla Bay (Wonham et al. 2005; Collin et al. 2006), there is evidence for multiple introductions based on genetic assignment tests and clustering analysis (Fig. 2.2, 2.3). There were no clear signals of admixture from different source populations in the dataset, nor an increase in F_{is} in PB relative to native populations. However, F_{is} is different from zero, which supports the presence of a Wahlund effect due to the coexistence of genetically divergent and isolated sub-populations. Though *C. convexa* has been in Washington State for many decades (Townsend 1895) and has therefore had potential time to reach panmixia, *C. convexa* is not found on all possible substrates at this site. Mating among individuals on specific substrates coupled with juvenile dispersal that is less than the sampling scale may lead to this pattern of a Wahlund effect (see also Ordóñez et al. 2013). Given that we only sampled a single introduced population, however, we are unable to conclude that high genetic diversity is a general feature of introductions in *C. convexa*. More samples are needed to determine the generality of this result.

Average pairwise F_{st} values comparing the PB population to native populations are greater than the overall average pairwise F_{st} among native populations, but are roughly the same as the pairwise F_{st} between the northernmost and southernmost populations in the native range. The same patterns are found using D_{est}. This indicates that although the PB populations have

been isolated from their source on the order of several decades, their divergence does not exceed divergence now found over the entire native range.

Assignment tests show the most likely source populations in this dataset to be WM and DE (Fig. 2.2). These are located in Long Island Sound and Delaware Bay, respectively, both in the region that was the source of *Crassostrea virginica* transplants to the west coast of North America (e.g., a transplant from several bays in New Jersey and the Chesapeake Bay to Willapa Bay, Washington; Townsend 1895). A general problem of using assignment tests to identify source populations is that it is impossible to identify an unsampled population as a source in these analyses. The DAPC analysis shows that the PB population is distinct from either putative source population as well as the rest of the sampled populations (Fig. 2.3). However, the strong genetic structure observed in the native range as well as the concordance between known regions of oyster exports and the general regions identified as most likely source populations (Long Island Sound and Delaware Bay) support the hypothesis that *C. convexa* was introduced to Washington with imports of *Crassostrea virginica*.

Continuing the comparison between *C. convexa* and *C. fornicata*, high genetic diversity as found in the PB population of *C. convexa* was also found in an introduced population of *C. fornicata* located in nearby Mud Bay (Washington, USA; Riquet et al. 2013). This species was also presumably introduced to Puget Sound multiple times with eastern oyster (*Crassostrea virginica*) transplants during the same time period (i.e., late 19th and early 20th centuries; Blanchard 1997). These introduced populations of both *Crepidula* species share common genetic characteristics: 1) they display levels of genetic diversity which are high and similar to populations in the native range (this study, Riquet et al. 2013) and 2) they both show evidence of multiple introductions based on genetic assignment tests (Fig. 2.2, Fig. S.2.2). These shared

genetic properties are expected given the similar vectors and timing of introduction. In contrast, Riquet et al. (2013) found that more recently introduced populations of *C. fornicata* in Europe (e.g., France, Norway) are more genetically distinct from native populations, reflecting a different history of introduction. A more robust comparison of the two species within their overlapping introduced range would be valuable in understanding how introduction history and current dispersal patterns interact to form the currently observed patterns of genetic diversity in these two congeners.

2.5 Conclusion

These results show that the direct-developing *Crepidula convexa* has low gene flow among populations within its native range as evidenced by a strong pattern of isolation-by-distance and high values of self-assignment using multilocus genotypes. This contrasts with a pattern of weak genetic structure within the native range of *C. fornicata*, a planktotrophic congener. Genetic diversity is high in all populations with this set of markers, but there is a slight decrease in diversity at the northern range margin. An introduced population of *C. convexa* shows similarly high diversity, and assignment tests indicate that the most likely sources of introduction are from the mid-Atlantic coast (New York – Delaware), consistent with the proposed vector of introduction, eastern oysters. This pattern of high genetic diversity within an introduced population is also found in a nearby population of *C. fornicata*.

Table 2.1 Collection information. Collection and genetic diversity information for seven native populations and one non-native population of *Crepidula convexa*. Genetic diversity information is averaged across five microsatellite loci (data for locus CT5H8 not included, but see Appendix A for analyses with this locus). Native populations are arranged from north to south. **Bold** values are significant at $p < 0.01$. Values for the introduced population shown in *italics*.

Site	Location	Coordinates	Collection date	Sample size	Allelic richness	H_e	H_o	F_{is}
NA	Nahant, Massachusetts	42°26'11"N 70°56'20"W	2007	24	11.82	0.90	0.89	0.01
BA	Barnstable, Massachusetts	41°42'0 "N 70°17'58"W	2007	64	11.85	0.91	0.87	**0.05**
WM	Old Field, New York	40°56'13"N 73° 8'44"W	2009	83	14.60	0.94	0.79	**0.16**
JB	Jamaica Bay, New York	40°35'60"N 73°48'48"W	2009	15	14.78	0.94	0.83	**0.11**
DE	Lewes, Delaware	38°47'42"N 75° 5'53"W	2002	29	13.81	0.92	0.76	**0.17**
VA	Gloucester Point, Virginia	37°17'15"N 76°24'2"W	2002	32	12.76	0.90	0.79	**0.15**
NC	Beaufort, North Carolina	34°43'6"N 76°39'50"W	2002	31	13.04	0.93	0.75	**0.20**
PB	*Padilla Bay, Washington*	*48°32'11"N 122°31'46"W*	*2009*	*85*	*12.19*	*0.89*	*0.79*	***0.10***

Table 2.2 Assignment tests. Results of assignment tests of native populations of *Crepidula convexa* using five microsatellite loci. Each population listed in the left-hand column is the collection location for the individuals (sample size given in parentheses). The populations in the subsequent columns indicate possible assignment populations based on a maximum likelihood estimate. The top row in each pair indicates the proportion of individuals assigned to each population; second row indicates the proportion of individuals rejected from a given population with 95% probability (i.e. a value of 1.00 indicates that all individuals were rejected). Self-assignment values are shown in bold font. Numbers above the diagonal indicate probabilities of assignment to more southern populations; those below the diagonal indicate assignment to more northern populations. No individuals were rejected from all populations.

	NA	**BA**	**WM**	**JB**	**DE**	**VA**	**NC**
NA (24)							
	1.00	0.00	0.00	0.00	0.00	0.00	0.00
	0	0.083	0.042	0.417	0.417	0.542	0.792
BA (64)							
	0.016	**0.906**	0.031	0.016	0.016	0.00	0.016
	0.422	**0.00**	0.078	0.375	0.234	0.656	0.875
WM (83)							
	0.012	0.00	**0.940**	0.024	0.012	0.00	0.012
	0.843	0.663	**0.00**	0.446	0.470	0.759	0.687
JB (15)							
	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00
	0.867	0.667	0.267	**0.00**	0.600	0.667	0.800
DE (29)							
	0.00	0.00	0.034	0.00	**0.966**	0.00	0.00
	0.828	0.690	0.138	0.241	**0.00**	0.552	0.793
VA (32)							
	0.00	0.00	0.00	0.00	0.00	**1.00**	0.00
	0.781	0.906	0.750	0.188	0.250	**0.00**	0.719
NC (31)							
	0.00	0.00	0.00	0.032	0.00	0.00	**0.968**
	1.00	1.00	0.613	0.484	0.452	1.00	**0.00**

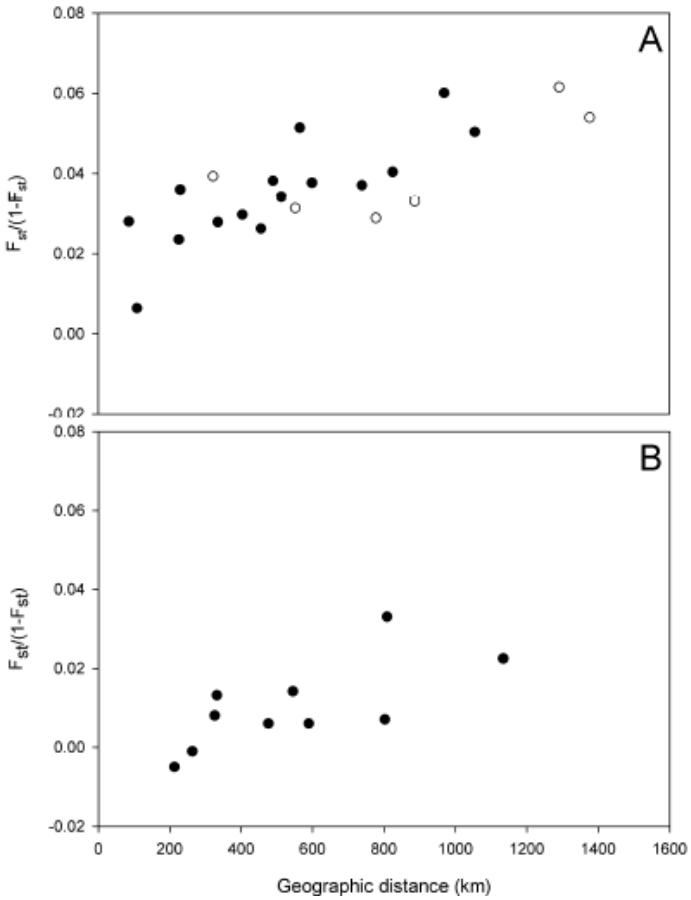

Fig. 2.1 Isolation by distance (IBD). A) Pairwise comparisons of seven native populations of *Crepidula convexa* (this study). Overall F_{st} = 0.018. IBD slope = 2.80×10^{-5}. Mantel test, r = 0.803, p = 0.007. B) Pairwise comparisons of five native populations of *C. fornicata* sampled over approximately the same latitudinal range (data from Riquet et al., 2013; pairwise F_{st} provided in Table S2 of Riquet et al., 2013). Overall F_{st} = 0.011. IBD slope = 2.67×10^{-5}. Mantel test, r = 0.569, p = 0.227. Note that the IBD in panel A is significant even when the most southern population (NC) and its associated comparisons are removed from the analysis (open symbols in panel A; for filled symbols only, overall F_{st} = 0.017, IBD slope = 3.5×10^{-5}. Mantel test, r = 0.809, p = 0.017).

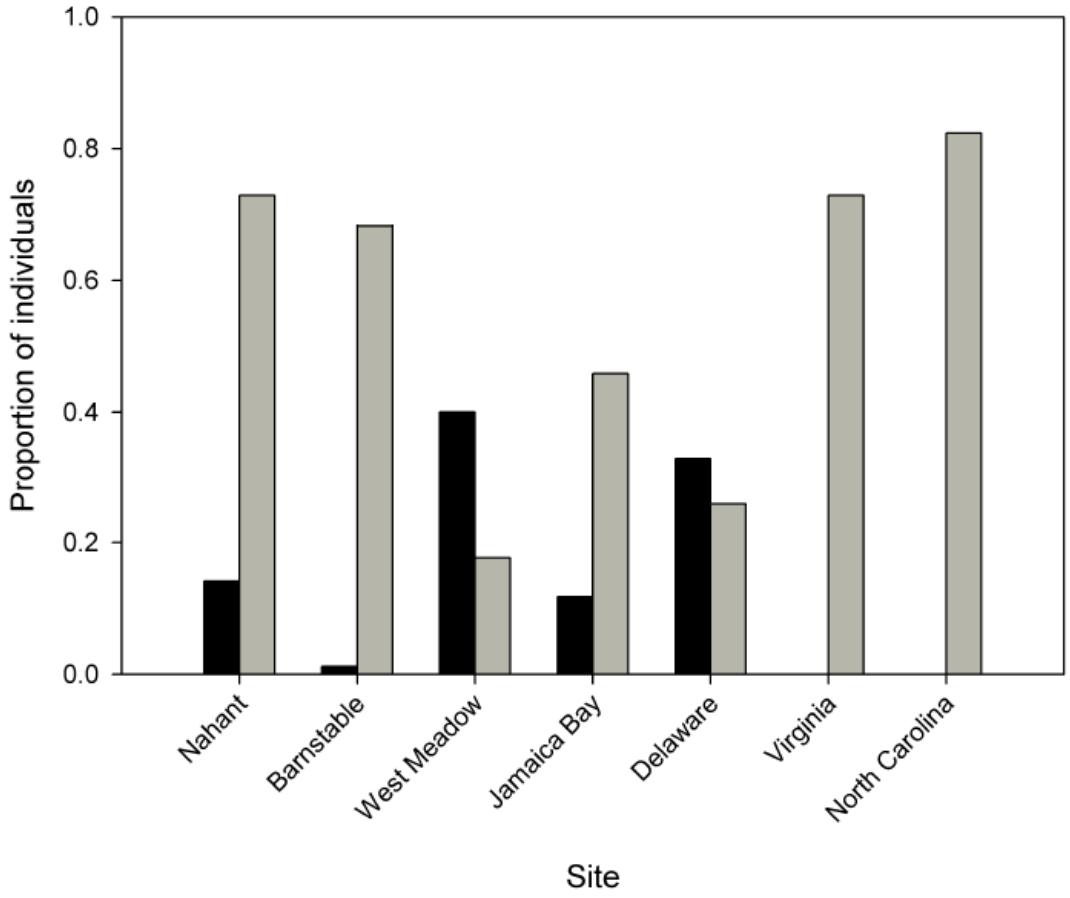

Fig. 2.2 Assignment tests. Proportion of individuals from the introduced (Padilla Bay) population of *Crepidula convexa* that were assigned to (dark bars) or rejected from (with 95% probability; light bars) each of the seven native populations using likelihood-based methods. Site information is listed in Table 2.1, and populations are arranged from north to south (left to right) on the x-axis.

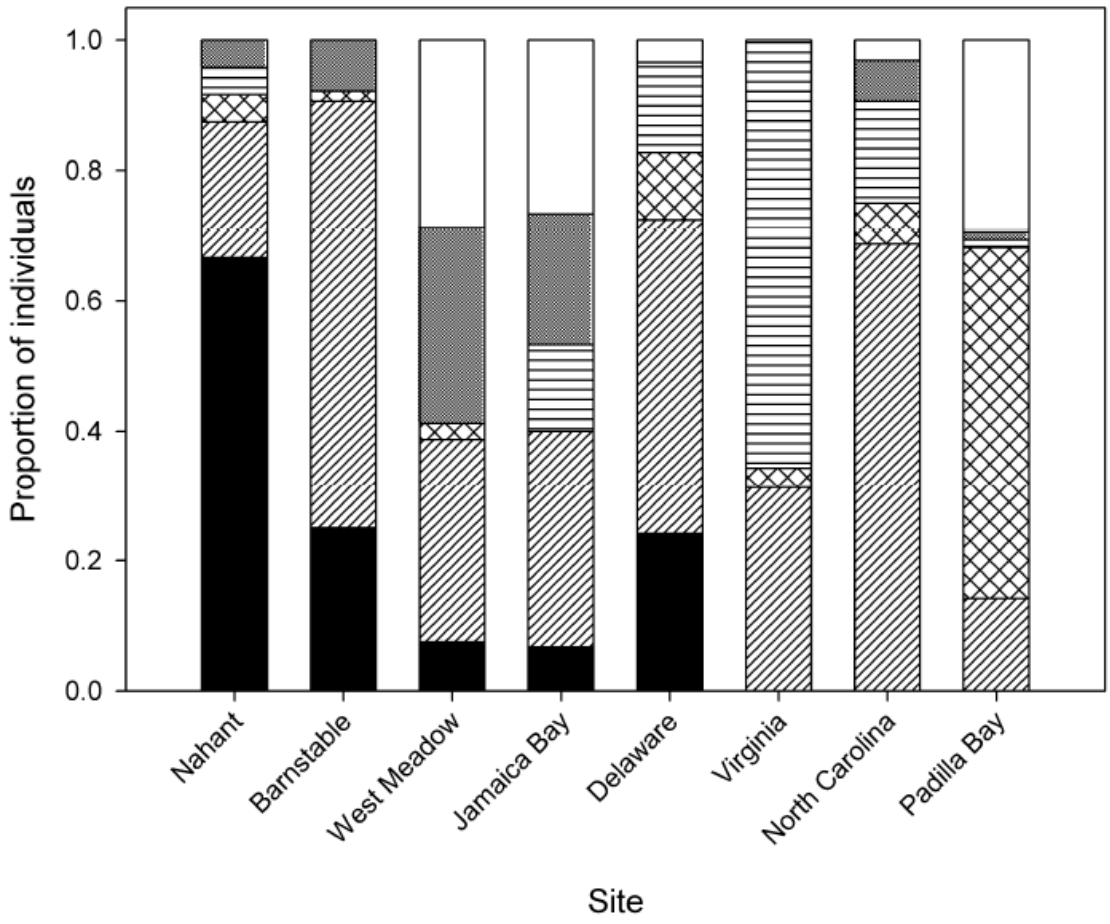

Fig. 2.3 DAPC analysis. Proportions of individuals from 8 populations of *Crepidula convexa* assigned to different genetic groups (k = 6) in a clustering analysis (discriminant analysis of principal components). The analysis groups individuals to get the smallest within-cluster genetic variance and the largest among-cluster genetic variance. Each of the six genetic clusters is indicated by a different color and pattern combination. Site information is listed in Table 2.1, and seven native populations are arranged from north to south (left to right) on the x-axis; the Padilla Bay population (Washington, U.S.A.) is introduced.

Chapter 3: High throughput sequencing of two species of gastropod with differing dispersal modes reveals genetic differentiation at northern range margins

3.1 Introduction

The limits of species' geographic ranges is one of both theoretical and applied concern in evolutionary ecology (Bridle and Vines, 2007; Kawecki, 2008; Sexton et al., 2009). Given that all species do not occur everywhere, range limits must have a cause. Understanding the mechanisms determining these range limits, as well as their underlying dynamics, is of current interest given that many species are experiencing range shifts in response to changing climate (e.g. Parmesan & Yohe, 2003; Perry et al., 2005; Parmesan, 2006; Thomas, 2010; Angert et al., 2011). There are many examples of proximate mechanisms that determine the location of range limits (reviewed in Gaston, 2009; Sexton et al., 2009; Sunday et al., 2012; Cahill et al., 2014), including interspecific competition (Gross and Price, 2000) or predation (Sievert and Keith, 1985). However, physiological limitations to abiotic factors are implicated more often than biotic factors in setting range limits at both warm-edge (i.e., low latitude, low altitude; Sexton et al., 2009; Cahill et al., 2014) and cold-edge (i.e., high latitude, high altitude; Sexton et al., 2009) limits.

Based on theoretical models, the role of genetics in limiting species' range shifts is a mix of too much versus too little gene flow or genetic variation. Large amounts of gene flow from a central population to a marginal one may cause the migration of alleles that are adaptive within the range but maladaptive at the margins (reviewed in Lenormand, 2002; Kawecki 2008). This swamping effect may prevent marginal populations from adapting to conditions at the range edge (Kirkpatrick and Barton, 1997). Very low amounts of gene flow, however, could result in small, isolated populations in the margins that do not receive adaptive alleles from other sites and are

prone to genetic drift (Bridle and Vines, 2007). Conversely, adaptation to environments at the margin of a species' distribution may allow for diversification at range margins (e.g., Bridle et al., 2014). The balance between gene flow and selection and the relative importance of these processes in determining species' latitudinal ranges has been frequently modeled (e.g., Holt and Gomulkiewicz, 1997; Kirkpatrick and Barton, 1997; Case and Taper, 2000; Price and Kirkpatrick, 2009; Bourne et al., 2014) and, to a lesser degree, investigated empirically (e.g., Sexton et al., 2011).

Marginal populations are expected show reduced genetic diversity relative to central populations due to fragmentation, small population size and a corresponding increase of genetic drift, and potentially strong adaptation to local conditions (Kawecki, 2008). This prediction assumes that a species has high population densities in the center of its range and low abundance at the margins, i.e., an abundant-center distribution (Lira-Noriega and Manthey, 2014). Genetic variation is indeed lower at range margins in many cases (reviewed in Gaston, 2003; Eckert et al., 2008; Kawecki, 2008; Lira-Noriega and Manthey, 2014), but support is far from universal (e.g., Moeller et al., 2011). A recent review by Lira-Noriega and Manthey (2014) found that a reduction in genetic diversity is stronger when considering the distance of a population to the center of the environmental niche, rather than its distance to the center of the geographic distribution. In this study, the environmental niche was calculated by making species distribution models and estimating the centroid of the environmental niche using a principal components analysis of environmental variables.

Coastal marine populations frequently exist in a linear series of populations, making the identification of populations at margins relatively easy. The connectivity and genetic differentiation of marine populations, particularly as this relates to larval development, is an area

of active study. For most marine invertebrates, adults are sessile or of relatively low mobility and dispersal occurs mostly in the larval or juvenile stage (Cowen and Sponaugle, 2009). This has led to the prediction that species with long-lived planktotrophic stages, which can live in the plankton for hours to years (e.g., Strathmann and Strathmann, 2007), will have lower levels of genetic differentiation than those with direct-developing larvae that do not live in the plankton (reviewed in Shanks, 2009). Although the predicted regression between larval duration and genetic structure (usually measured with F_{ST}) is not always observed (e.g., Weersing and Toonen, 2009; Kelly and Palumbi, 2010; Selkoe and Toonen, 2011), species with planktotrophic larvae do frequently have less geographically-based genetic differentiation than closely-related direct developers (e.g., in gastropods: Collin, 2001; Lee and Boulding, 2009; Chapter 2). This general pattern indicates that larval type may be related to the degree of gene flow to marginal populations, affecting both the amount of genetic diversity in marginal relative to core habitats and the degree of potential gene swamping (i.e., alleles that are beneficial in the core but detrimental in the margins).

Crepidula fornicata and *C. convexa* (Gastropoda: Calyptraeidae) are sympatric and have different dispersal ability. They are thus two species that can be used to compare diversity in populations at range margins based on dispersal ability. These two species are native to the east coast of North America, where they have largely sympatric ranges (Collin, 2001; Fig. 3.1). *Crepidula fornicata* is a planktotrophic developer, with a 2-4 week larval period (Collin, 2003) that allows for high dispersal potential. Populations of *C. fornicata* display no genetic differentiation from Massachusetts to the Atlantic coast of Florida when measured with molecular markers (allozymes: Hoagland, 1985; COI: Collin, 2001; microsatellites and AFLPs: Riquet et al., 2013), although populations of *C. fornicata* in the Gulf of Mexico are genetically

different from those in the Atlantic (Collin, 2001; Riquet et al., 2013). In contrast, *C. convexa* is a direct-developing snail with no planktonic larval stage. This lack of larval dispersal has led to relatively strong genetic differentiation among populations (Collin, 2001; Chapter 2, this dissertation). *Crepidula convexa* exhibits a strong pattern of isolation-by-distance (IBD), while *C. fornicata* has no significant IBD over the same geographic range (Chapter 2). Different amounts of dispersal may lead to different dynamics at the range margins between these two species.

The native range of *C. fornicata* extends from the Yucatan Peninsula (Mexico) in the south (Collin, 2001) to Newfoundland (Canada) in the north (Rawlings, 2011). The range of *C. convexa* is smaller, as predicted for a direct developer (Scheltema, 1986; Johannesson, 1988): it stretches from Georgia (United States) to Massachusetts (United States; Collin, 2001; Fig. 3.1). The northern range edge of *C. fornicata* may have expanded over the past decades, apparently tracking warming water temperature (Rawlings, 2011), and expansion is predicted to continue (Saupe et al., 2014). This expansion is mirrored in Europe, where *C. fornicata* was introduced and is now moving northwards with warmer temperatures (Thieltges et al., 2004).

Although southern range edges (low-latitude edge) are of interest due to potential climate-related range contractions (Cahill et al., 2014), the climate-related range expansions in *C. fornicata* mean that the northern edge will also be relevant to range shifts. Additionally, *C. convexa* has a cryptic sister species with a parapatric range to the south (*C. ustulatulina*; Collin, 2001, 2002), making it difficult to identify the southern range edge in this species. I therefore chose to study the northern range edge of these *Crepidula* species.

Here, I used high-throughput, next-generation DNA sequencing (genotype-by-sequencing, or GBS, Elshire et al., 2011; a variant on restriction enzyme associated DNA sequencing, or RAD-Seq, Baird et al., 2008) to investigate genetic diversity and structure in central and marginal populations of *Crepidula* species with differing dispersal ability. The large number of SNPs identified in RAD-Seq analyses have revealed cryptic population differentiation in other species when neutral markers such as microsatellites did not (e.g., in the herring *Clupea harengus*; Corander et al., 2013). By using GBS to generate a large library of SNPs, I was able to investigate genetic structure on a finer scale than previously achieved in these species, and found population differentiation in the heretofore undifferentiated *C. fornicata*. I also tested the hypothesis of reduced genetic diversity (measured with F_{is}, H_e, and H_o) in marginal populations relative to central ones, and related any reduction to the dispersal abilities of these two species. I expected to see a greater reduction of diversity in *C. convexa* due to its reduced connectivity among current-day populations, potentially leading to stronger genetic drift.

3.2 Methods

Population sample scheme and collection

Samples of *C. fornicata* and *C. convexa* were collected from four populations each along the Atlantic coast of the United States (Table 3.1, Fig. 3.1). Two central and two marginal populations were sampled for each species. Twenty to 28 individuals were collected from each population. Snails were collected from many different substrates at each site (e.g., rocks, bottles, clams, etc.), and multiple individuals were usually collected from the same substrate. Based on prior studies (Dupont et al., 2006; Le Cam et al., 2014), adults on a single substrate (e.g., one clam shell) are not closely related, so collecting multiple individuals from a single substrate is not expected to affect results. Collections were made by a single person at each site. With the

exception of the Newfoundland population of *C. fornicata*, animals were transported live to the laboratory at Stony Brook, New York, then removed from their substrates and placed directly in a freezer at -80°C. Samples from Newfoundland were removed from their substrates and placed in 95% ethanol before transport to New York.

DNA extraction and quality control

DNA was extracted from 0.2 g of tissue using a DNeasy® Blood & Tissue Kit (Qiagen®) according to the manufacturer's instructions. Cephalic tissue was used for extractions of *C. fornicata* and large *C. convexa*, and whole animals were used for small *C. convexa*. Following this extraction, mucopolysaccharides and other contaminants were removed using a DNeasy® column and reagents. First, unpurified samples were added to the supernatant of lysis buffer that had been chilled and centrifuged, and the solution was placed on a DNeasy® column. Wash buffer was then added to the mix (Qiagen® buffer AW1), and the reaction was centrifuged for 6 min at 5,000 rpm. A second wash buffer (Qiagen® buffer AW2) was added to the column, followed by 10 min of centrifugation at 5,000 rpm. Finally, elution buffer was added and centrifuged for 4 min at 5,000 rpm.

The concentrations of the cleaned DNA samples were measured using photometric dye (Quant-iT^{TM} PicoGreen®, Life Technologies) and a Mini-Fluorometer (TBS-380, Turner Biosystems) according to the manufacturers' protocols. A subset of the samples was digested with the restriction enzyme Sau3AI (New England Biolabs®), then run on a 1% agarose gel (95V for 50 minutes) and visually checked to verify DNA quality. Samples were concentrated to ≥20 ng/µl for sequencing.

Genotyping

Samples were shipped to the Institute for Genomic Diversity (IGD) at Cornell University, where they were analyzed using a restriction enzyme genotype-by-sequencing procedure (GBS; Elshire et al., 2011). Libraries were prepared for 20-26 individuals per population (Table 3.1). The IGD tested several common restriction enzymes on a subset of samples to see which produced the best library in terms of fragment size and number. Following this enzyme optimization, DNA from each individual was digested separately using PstI, a restriction enzyme with a six-base recognition site. DNA fragments were then ligated to a barcoded adaptor (a separate barcode for each individual) and a common adaptor. DNA fragments from each species were run on a separate 96-well plate with 95 wells each containing DNA from a different individual and one well serving as a blank for sequencing. The libraries were sequenced on an Illumina HiSeq 2500 with 96 samples sequenced per lane, and the reads generated were 100 base pairs (bp) long. The genome size of *C. fornicata* is unknown, but that of the congener *C. unguiformis* has been estimated at approximately 6.2 gigabases (Gb; Libertini et al., 2009). Assuming a 6 Gb genome size for both *Crepidula* species in my analysis, and a 200 million reads-per-lane output from the Illumina HiSeq, this amounts to an average < 0.05x coverage per individual.

Bioinformatics and analyses

The data were quality checked using FastQC version 0.10.1 (Andrews, 2010), a program that does quality control checks on high throughput sequence data to identify basepairs and sequences with high rates of error and low quality scores introduced by the sequencing technology. This program summarizes the quality scores (provided by the Illumina HiSeq during the analysis) for each base and sequence to make sure that the sequences are of sufficiently good quality to continue with downstream analyses.

Data were then analyzed using the nonreference pipeline Universal Network-Enabled Analysis Kit (UNEAK; http://www.maizegenetics.net/gbs-bioinformatics; Lu et al., 2013). The pipeline first trims the reads to 64 bp to remove the error-prone end of the sequence (i.e., the last base pairs sequenced), and then classifies identical reads as tags. A network analysis is used to find tags that differ by a single base pair (i.e., candidate SNPs; Lu et al., 2013). The pipeline was run with the default error tolerance rate (0.03, designed to minimize the chance that real tags are discarded as sequencing errors) and the default minimum minor allele frequency (0.05). The two species were analyzed with the same parameters in the same run of the pipeline. Both the FastQC and UNEAK analyses were conducted using the platform provided by the iPlant Collaborative (Goff et al., 2011).

Following the SNP discovery done with UNEAK, I separated the data by species, and all further analyses were conducted separately for each species. I then filtered the dataset to only those loci that were sampled at 75% or more of individuals (i.e., 72 individuals out of 95 total), as well as loci that were sampled at a mean coverage of ≥ 10X per individual (i.e., on average, each individual had 10 or more sequenced copies of the locus, allowing for accuracy in identifying heterozygote loci). I also removed all failed individuals, defined as those that were sequenced with less than 10% of the mean reads per sample for the species (eight individuals in *C. fornicata* and 10 individuals in *C. convexa*; see Table 3.2 for final sample sizes in each population). All subsequent analyses were done on this reduced dataset of SNPs only (discarding all other sequence data), and all SNPs used in these analyses therefore had at least 10X coverage.

Population structure of these SNPs was analyzed using discriminant analysis of principal components (DAPC; Jombart et al. 2008, 2010) on all populations. This is a clustering analysis

that first performs a principal components analysis on the multilocus genotypes of the samples, then a discriminant function analysis on the PC scores. The analysis minimizes within-group variation and maximizes between-group variation (Jombart et al. 2010). For each species, I ran two clustering analyses, one using the population of origin (i.e., collecting site) as a prior in the clustering analysis and one without this prior. The analysis run without collecting site as a prior relies on the Bayesian Information Criterion (BIC) to determine the best number of clusters in the data: the investigator selects the number of clusters that generates the lowest BIC. DAPC was conducted using the adegenet package, version 1.4-2 (Jombart et al., 2008) in R 3.0.1 (R Core Development Team 2013).

I calculated the expected heterozygosity (H_e) and observed heterozygosity (H_o) values and F_{is} for each SNP locus and population, as well as the overall population structure (measured with F_{st}) in the dataset, using the hierfstat package, version 0.04-10 (Goudet, 2005) in R 3.0.1. I used adegenet, version 1.4-2 (Jombart et al., 2008) to calculate the pairwise F_{st} values between all pairs of populations. H_o was calculated as the proportion of heterozygous samples in the data, and H_e was calculated with observed allele frequencies and assuming Hardy-Weinberg equilibrium (Goudet, 2005).

3.3 Results

Quality control and SNP calls

The FastQC results showed a mean per-base quality score (Phred score) in *C. convexa* of 34.64, and the first 74 bp of the sequences were consistently above 20 (corresponding to a 99% probability of an accurate base call; Ewing and Green, 1998). The mean per-sequence quality

score was 37. In *C. fornicata*, the mean per-base quality score was 35.10, and the first 94 base pairs were above a quality score of 20. The mean per-sequence quality score was 38.

After filtering the data to remove failed individuals (eight in *C. fornicata* and 10 in *C. convexa*; Table 3.2), SNPs with low mean coverage per individual (< 10X), and SNPs that were sampled in < 75% of individuals (i.e., < 72 individuals), there were 1903 loci remaining in the *C. fornicata* dataset and 309 loci remaining in the *C. convexa* dataset (Fig. 3.2). All subsequent analyses were conducted with these reduced datasets.

Interpopulation differentiation

In the DAPC analysis of *C. convexa*, when sampling site was used to inform the clustering, the analysis separated populations north and south of Cape Cod along the first principal components axis (Fig. 3.3A). Populations NJ and NY were mostly overlapping, and NA and BA overlapped slightly. The second principal component axis separated NA from BA. The three principal eigenvectors of the discriminant analysis were of lengths 132.83, 24.85, and 18.10, indicating that the first PC axis explained most of the variation in the data. When conducting the DAPC without using population of origin as a prior, the BIC with the lowest value corresponded to $k = 4$ genetic clusters. Each geographic population has a different proportion of individuals that belong to each cluster, and the four populations are relatively distinct (Fig. 3.4A).

The DAPC analysis of *C. fornicata* with sampling site as a prior supported four distinct genetic populations, again roughly corresponding to the four geographic populations from which I sampled. The NA and NY populations overlapped slightly on the first two principal component axes, with the populations from NS and NL very distinct both from each other and from the more southern populations (Fig. 3.3B). The three principal eigenvectors of the discriminant analysis

were of lengths 209.54, 97.91, and 33.12. When using DAPC to find the best-fit number of genetic clusters in the data, the lowest BIC value corresponded to $k = 5$ clusters. Again, the population from NL was very distinct from any other population, and there was a clear difference between more southern populations (NA and NY) and northern, marginal ones (NL and NS). In particular, there were genetic clusters that clearly corresponded to northern and southern groups of individuals, with little overlap between them (Fig. 3.4B).

Population genetic analyses

The overall expected heterozygosity (H_e) in the overall dataset for *C. convexa* was 0.314, with an overall observed heterozygosity (H_o) of 0.207. Each population also showed a lower observed frequency of heterozygotes than expected, with mean H_e values ranging from 0.269 – 0.364, and mean H_o ranging from 0.161 – 0.255 (Table 3.2). This deficiency in heterozygotes corresponded to an overall F_{is} value of 0.341, calculated across all populations and loci. Across populations, the mean F_{is} ranged from 0.312 – 0.563 (Table 3.2). However, all populations also showed a wide range of F_{is} values, ranging from -1 to 1 but biased towards positive values (Fig. 3.5A). The distributions of F_{is} values were different among populations ($F_{3,1161} = 19.79$, $p < 0.001$), with the central populations (NY and BA; $p = 0.07$) and the marginal populations (NA and NJ; $p = 0.97$) not different from each other based on post-hoc Tukey HSD tests; all other pairs of distributions were significantly different from each other. Not all loci were polymorphic in all populations. The proportion of polymorphic loci ranged from 68.8% (BA) to 86.7% (NJ) and corresponded to the class of loci with F_{is} values equal to 1 (Fig. 3.5A). No measures of genetic diversity were lowest in the northernmost population (NA; Table 3.2).

The overall F_{st} among all populations of *C. convexa* was equal to 0.018. Pairwise F_{st} values ranged from 0.015 (NA to BA) to 0.0286 (NA to NY), but comparisons of the sites on the same side of Cape Cod, north or south, yielded lower pairwise F_{st} values than comparisons across the Cape (Table 3.3). The average pairwise F_{st} was 0.023. There was no significant pattern of isolation-by-distance as determined by a Mantel test ($p = 0.32$). Instead, comparisons between the two central and two marginal populations (i.e., BA-NA and NY-NJ) yielded lower F_{st} values than comparisons of central to marginal populations (Table 3.3).

In *C. fornicata*, the overall H_e was 0.207, and the overall observed heterozygosity (H_o) was 0.124, calculated across all populations and loci. Within-population values of H_e were again higher than H_o, with mean H_e ranging from 0.176 – 0.241, and mean H_o ranging from 0.108 – 0.149 (Table 3.2). The overall F_{is} in the dataset, calculated across all populations and loci, was 0.400, with mean F_{is} across populations ranging from 0.278 – 0.591 (Table 3.2). As in *C. convexa*, there was a wide range of F_{is} values within all populations (-1 to 1) (Fig. 3.5B), and many loci were fixed within each population. The distributions of Fis were different among populations ($F_{3,6528} = 198.1$, $p < 0.001$), with only the central populations not different from each other (NS-NA, $p = 0.91$). All other pairwise comparisons of populations were different based on Tukey's HSD tests. Also as in *C. convexa*, the northernmost population (NL) did not show a reduction in any metric of genetic diversity relative to more southern populations.

The overall Fst among all populations of *C. fornicata* was 0.022, and the average pairwise F_{st} was 0.033. Pairwise F_{st} values ranged from 0.020 (NY to NA) to 0.044 (NA to NS). As with *C. convexa*, there was no significant pattern of isolation-by-distance (Mantel test; $p = 0.33$), but pairwise values comparing the two marginal populations (NS-NL) and two central

populations (NY-NA) were lower than any pairwise values comparing a central and a marginal population (Table 3.3).

3.4 Discussion

Genotype-by-sequencing of Crepidula *species*

Genotype-by-sequencing technologies, particularly the use of restriction enzymes to generate reduced representations of the genome for sequencing, were successful in creating a large set of SNPs for both *C. fornicata* and *C. convexa*. The number of SNPs generated was relatively small (e.g., as compared to > 90,000 for each of two species of *Iris* based on < 30 samples per species and using the same technique, Hamlin and Arnold, 2014), and most SNPs were not found in the majority of individuals. This is expected, given the very large genome of *Crepidula* species (approximately 6 Gb in *C. unguiformis*, Libertini et al., 2009), which reduced the average coverage per individual to < 0.05x. Despite this, after filtering for sufficient coverage per individual and per locus, the analysis still generated nearly 2000 SNPs in *C. fornicata* and over 300 SNPs in *C. convexa*, and these markers were able to distinguish among populations within both *C. convexa* and *C. fornicata*. This method is therefore promising for further development of population genetic markers in these and other non-model organisms with large genomes.

Population structure in Crepidula *species*

The DAPC analysis identified four distinct genetic clusters in *C. convexa*, which roughly corresponded to the four collecting sites used in this study (Fig. 3.4A). This corresponds to previous analyses of genetic structure in *C. convexa* (Collin, 2001; Chapter 2, this dissertation), and is expected given the direct development of this species and its correspondingly low dispersal distance. However, the snails collected in NY and NJ had genotypes that largely

overlap on the first two principal component axes (Fig. 3.3A). I did not sample the NJ population using microsatellites (see Chapter 2). PC1 separates the populations across Cape Cod. Cape Cod is known to be a biogeographic barrier for many marine invertebrates due to oceanic current structure, changes in temperature, and differences in available habitat (Wares, 2002), but I found no evidence that it is a dispersal barrier in *C. convexa* using microsatellites in Chapter 2 (Fig. 2.1A, 2.3). Although PC1 in Fig. 3.3A appears to separate populations that are on either side of Cape Cod, it is also consistent with populations that are closer together being more tightly clustered than those that are farther apart. I did not find a significant isolation-by-distance relationship among these four populations, but it is important to note that Mantel tests are relatively low-powered (Legendre and Fortin, 2010) and that I had few pairs of populations to compare. In order to see if Cape Cod really represents a barrier to gene flow in this species, more populations must be sampled.

The pattern observed in *C. fornicata* was quite different. The DAPC analysis identified five distinct genetic clusters within the data (Fig. 3.4B). In contrast to *C. convexa*, the central populations (NY and NA) were only slightly overlapping (Fig. 3.3B). This matches previous studies of *C. fornicata* with other marker sets (COI, Collin 2001; microsatellites and AFLPs, Riquet et al., 2013), which showed low levels of differentiation between these two sites (e.g., F_{ST} = 0.013 with 17 microsatellites or 0.012 with 327 AFLPs, Riquet et al., 2013).

However, the populations sampled near the range margin (NS and NL) were very different from the central ones (Fig. 3.3B, 3.4B). This is easily visible in Fig. 3.4B, where there was a sharp division in the genetic clusters present in northern versus southern populations (i.e., the dark grey cluster was only found in southern sites, and the light blue cluster was more prevalent in southern sites). In this case, PC1 (Fig. 3.3B) separated populations on the Atlantic

coast of North America (NY, NA, NS) from the NL population, which is located on the west side of the Cabot Strait.

The NS population, though distinct from those further south, did share some similarities with them (Fig. 3.4B). The differences may be largely explained by the geographic distance between populations. The NS collecting site is located approximately 1100 km from the next site south (NA). This distance is the same as that between NA and Chesapeake, Virginia, which displayed an F_{ST} of 0.022 when analyzed with 17 microsatellite markers (Riquet et al., 2013). Additionally, the oceanography and water movement between NA and NS, particularly circulation patterns in the Gulf of Maine and the Bay of Fundy (Miller et al., 1998), may make the effective distance between the two locations much larger than the geographic distance. However, further analyses and finer-scale sampling between NA and NS are required to assess the amount of genetic differentiation that can be explained by geographic distance (i.e., isolation-by-distance analyses). Other marine species have shown similar patterns of genetic structure (i.e., striking differences between Gulf and Atlantic coast populations) in the Gulf of St. Lawrence (e.g., halibut, Fairbairn, 1981; hard clams, Dillon and Manzi, 1992; calanoid copepods, Bucklin et al., 1996).

The NL population was distinct from all other populations of *C. fornicata* (Fig. 3.3B), though it did share genetic clusters with the other marginal population, NS (Fig. 3.4B). Although NL and NS are only 300 km apart, a distance over which *C. fornicata* normally displays very little genetic differentiation (Collin 2001, Riquet et al., 2013), the water circulation patterns in this area likely enhance isolation. The NS population is found on the northeast tip of Nova Scotia, to the east and south of the Cabot Strait on the Atlantic coast of the province. The NL population, from southwestern Newfoundland, is located to the northwest of this strait, inside

the Gulf of St. Lawrence. The Cabot Strait is a deep channel (> 200 m) of fast-moving water. At all times of the year and at all depths, the net motion of water is from west to east, out of the Gulf of St. Lawrence, and frequently at high velocities (Wu and Tang, 2011). It is very likely that the planktonic larvae of *C. fornicata* are unable to move from east to west across this current, and that gene flow would therefore be nearly unidirectional. This would account for both the distinctness of the NL population and the shared genetic clusters with NS, as there is probably some migration from NL to NS and points south.

The circulation within the Gulf of St. Lawrence is complicated (Rawlings, 2011; Wu and Tang, 2011), and might allow for genetic exchange between the NL population in this study and other populations within the Gulf. *Crepidula fornicata* populations are known from several locations in the Gulf, including Quebec, New Brunswick, Prince Edward Island, and the west coast of both Nova Scotia and Newfoundland, in areas where summer water temperatures reach 18°C (Rawlings et al., 2011). Further sampling should include individuals from these areas to understand genetic exchange within the Gulf of St. Lawrence. This might illuminate not only the distinctness of the NL population in this study (Figs. 3.3B, 3.4B), but also be particularly of interest for future range expansions of *C. fornicata*. The northernmost locations within the native range of *C. fornicata* are found in this area (Quebec).

Furthermore, the Gulf of St. Lawrence (and in particular, Prince Edward Island) is an area with high levels of shellfish aquaculture. Given that *C. fornicata* is known to be transported with oysters and mussels (Carlton, 1992), the distribution and spread of the species within the Gulf may be associated with human activity. It is even possible that the apparent northern spread of the species may be due not only to warming water, but also to human-assisted movement (Rawlings et al., 2011). If these populations are indeed distinct from those on the Atlantic coast,

as hinted at by the current study, any future attempts to identify standing genetic variation that may allow for adaptation to conditions at range margins should be conducted using these organisms.

Further work in these species should also take historical gene flow into account. Many marine invertebrates show a pattern of expansion from southern refugia following the Pleistocene glaciation (e.g., Wares and Cunningham, 2001). A phylogenetic analysis of COI in *C. fornicata* did not show a pattern consistent with post-Pleistocene expansion (i.e., northern haplotypes were not nested within more southern haplotypes; Collin, 2001). Both *Crepidula* species can exist subtidally, and therefore may show different historical patterns from species that only exist intertidally. However, evidence for current gene flow does not rule out past gene flow in structuring these populations, and should be investigated further.

Genetic variation in Crepidula *species*

Overall, the levels of genetic variation (H_e, H_o, proportion of polymorphic loci) were lower in *C. convexa* than in *C. fornicata* (Table 3.2). In *C. convexa*, there was no pattern of genetic diversity relative to central or marginal populations: NA (marginal) and NJ (central) showed the highest levels of H_o and the correspondingly lowest F_{is} values (Table 3). There was no reduction of genetic variation in NA, the northernmost population that has been identified in *C. convexa*. This is consistent with the lack of reduction in diversity relative to other populations (particularly those in New York) found in *C. convexa* using microsatellite markers (Chapter 2, Table 2.1). No other markers have been used to assess the genetic variation in the NJ population, so it was not possible to compare the high variation here to other data. Since the NJ population was the southernmost in the current dataset, it is also not possible to determine if high genetic

variation is only found at this site, or if variation is as high or higher in other central or southern populations.

The levels of genetic variation in *C. fornicata* revealed a similar pattern: the northernmost population (NL) showed the highest level of genetic diversity (H_e, H_o, or proportion of polymorphic loci), with NY (the southernmost population in this study) also showing high levels of variation (Table 3.2). There was therefore not a clear difference in the amount of variation between the central and marginal populations. This increase in genetic variation at the northern margin of the species could be due to several causes, including selection for the habitat found at the range margin.

Another explanation for this apparent increase in genetic variation is simply that the NL population is closer to Hardy-Weinberg equilibrium than other populations. Many marine invertebrates show deficiencies in heterozygotes leading to F_{is} values that are significantly different from 0 (Addison and Hart, 2005, and references therein). Significant deficiencies of heterozygotes have been found in both *C. convexa* (Chapter 2, Table 2.1), and *C. fornicata* (Riquet et al., 2013). The F_{is} values in *C. fornicata* are elevated above zero due to these heterozygote deficiencies (Table 3.2, Fig. 3.5A). The NL and NY populations may simply be closer to HWE than NS and NA, though the mechanism for this is unclear and further fine-scale analyses of the Gulf of St. Lawrence should be conducted.

F_{is} values in all populations of both species ranged from -1 to 1 (Fig. 3.5). The values less than 0 are probably due to small sample sizes; they represent loci that had higher heterozygosity than expected. The large number of loci in both species that had F_{is} values equal to 1 (i.e., were fixed; Fig. 3.5); they are polymorphic among populations of *Crepidula*, but are

fixed within populations. Future studies of these species should investigate these SNPs to better understand the processes that have resulted in population differences (e.g., low sample size, genetic drift, selection).

The distributions of F_{is} in both species are significantly different among populations, and in both cases the intermediate populations (i.e., NY and BA in *C. convexa*, NS and NA in *C. fornicata*) have higher average F_{is} values (Table 3.2, Fig. 3.5). These populations may be more prone to admixture than the marginal ones (NA in *C. convexa* and NL in *C. fornicata*) and therefore be more likely to experience a Wahlund effect, causing these elevated inbreeding values. However, the southernmost population for each dataset is not at the southern range margin and may also experience admixture from unsampled areas further south, so the relatively low F_{is} values in NJ and NY for *C. convexa* and *C. fornicata*, respectively, may not be explained by admixture among sampled populations.

Counterintuitively, the F_{st} values (both overall and pairwise) in *C. fornicata* were higher than those in *C. convexa*, indicating that populations of *C. fornicata* were more distinct than those of *C. convexa*. This contrasts with other results from these species (Collin, 2001; Riquet et al., 2013; Chapter 2), as well as expectations, but there are two possible explanations. First, *C. fornicata* populations were sampled at many more loci than *C. convexa* (1903 versus 309), so I had more power to distinguish among populations of *C. fornicata*. Second, the absolute distance spanned with the sampling of *C. fornicata* was much greater than that in *C. convexa* (Fig. 3.1), which may also increase population differentiation. The one pair of populations that was sampled for both species (NY and NA) showed a greater F_{st} value for *C. convexa* (0.0286) than for *C. fornicata* (0.0197). This is concordant with microsatellite analyses (Chapter 2) and

supports the idea that it is geographic distance, rather than marker number, which is responsible for the seemingly high F_{st} values in *C. fornicata*.

Lira-Noriega and Manthey (2014) show that the distance of a population from the climatic center of a species' range (i.e., optimal abiotic conditions) provides a better prediction of changes in genetic diversity than using the geographic center. This is unlikely to explain the difference in genetic variation in *C. fornicata*, as the northernmost population (NL) has the highest genetic variation (Fig. 3.1B) and is at the edge of the warm-water temperatures needed by *C. fornicata* to initiate reproduction and for larval development (Rawlings et al., 2011). More analysis of the environmental niche of both species is necessary to evaluate this idea.

In comparing these two congeners with different larval types and potential for dispersal, the most striking difference was that populations of the direct-developing *C. convexa* were differentiated over a relatively small spatial scale with a large marker set, supporting previous analyses (Collin, 2001; Riquet et al., 2013; Chapter 2). In contrast, *C. fornicata* shows very low differentiation in the center of its range (approximately the same geographic distance as the entire sampling area for *C. convexa*) and is much more differentiated near the northern margin. This contrasts with previous studies of this species, which have focused on central and southern populations (Collin, 2001; Riquet et al., 2013). The two marginal populations identified in this species were very different from each other (Fig. 3.2B, 3.3B). Gene flow between these divergent populations, each of which is different from more central ones may increase the genetic variation available to selection under conditions at the range margin. Gene flow among marginal populations of *Mimulus laciniatus* has been shown to help populations adapt to warm, dry conditions, whereas gene flow from central to marginal populations swamped the possibility of adaptation (Sexton et al., 2011). Future work to understand range expansion or adaptation in

marginal populations should investigate both the Gulf of St. Lawrence and the Atlantic coast of Canada and the potential for gene flow to connect them.

This work was all conducted at the northern range edge of these two species. Dynamics at the southern edge may look vastly different. A microsatellite analysis of *C. convexa* (Chapter 2) showed that the southernmost populations sampled (Virginia and North Carolina) were very different both from each other and from northern and central populations (Fig. 2.3). *Crepidula fornicata* populations in the Gulf of Mexico are different from Atlantic populations based on COI (Collin, 2001) and microsatellite and AFLP analysis (Riquet et al., 2013), as is true for other species (e.g., *Crassostrea virginica*, Reeb and Avise, 1990); this could be either due to the historical biogeography of the area or adaptation to the warmer temperatures of the Gulf of Mexico. Investigating these populations using a larger set of SNPs may help to predict how these populations are connected and may be able to adapt to warming water temperatures at the warm-edge limit of the species (i.e., in the Gulf of Mexico; Pinsky et al., 2013), which would provide a fuller picture of how genetics might affect the response of *Crepidula* species to climate change.

3.5. Conclusions

The use of genotype-by-sequencing technology generated a large library of polymorphic SNPs that could be used to analyze genetic variation and structure across populations of *C. convexa* and *C. fornicata*. Neither species showed a decrease in genetic variation at the range margin, and in fact, heterozygosity increased in the northernmost population of *C. fornicata*. Failing to find a decrease in genetic variation at range margins is not uncommon (Moeller et al., 2011; Lira-Noriega and Manthey, 2014). This prediction is based on the abundant-center hypothesis, where a species is expected to be most abundant in the middle of its geographic

range and decrease in abundance towards the edges. However, marine invertebrate species frequently do not show an abundant-center pattern (Sagarin and Gaines, 2002). Although robust abundance data are not available for either species of *Crepidula* analyzed here, it is unlikely that *C. convexa* shows an abundant-center distribution. The species is quite abundant at the NA site (pers. obs.), the northernmost site known for this species. *C. fornicata* may be more likely to show an abundant-center distribution: it is extremely abundant in NY (Hoch and Cahill, 2012), and less common in both NL and NS (though easily found within patches, A. Cahill and P. Sargent, pers. obs.). Despite this, I did not find a reduction in genetic diversity at the range margin.

Table 3.1. Collection information. Collection sites, dates, and sample sizes for four populations each of *Crepidula fornicata* and *C. convexa*, used for next-generation sequencing.

Species	Site	Center or Margin?	Location	Coordinates	Collection date	Sample size
C. fornicata	NL	M	Port Saunders, Newfoundland	50°38'49"N 57°16'34"W	September 2013	21
	NS	M	Main-à-Dieu, Nova Scotia	46°00'20"N 59°50'14"W	July 2013	26
	NA	C	Nahant, Massachusetts	42°26'11"N 70°56'20"W	July 2013	24
	NY	C	Old Field, New York	40°56'13"N 73°08'44"W	August 2013	24
C. convexa	NA	M	Nahant, Massachusetts	42°26'11"N 70°56'20"W	September 2012	25
	BA	M	Barnstable, Massachusetts	41°42'33"N 70°17'54"W	July 2013	24
	NY	C	Old Field, New York	40°56'13"N 73°08'44"W	August 2013	20
	NJ	C	Sandy Hook, New Jersey	40°26'50"N 73°59'44"W	October 2012	26

Table 2. Population genetic statistics. Average expected heterozygosity (H_e), observed heterozygosity (H_o), F_{is} values, and the proportion of SNP loci that are polymorphic for each population of *Crepidula fornicata* and *C. convexa*. Analyses were conducted separately for each species using SNPs that were sampled at more than 75% of individuals (i.e., 72 or more) and with an average of 10x coverage per individual. This equated to 1903 SNPs in *C. fornicata* and 309 SNPs in *C. convexa*. H_e, H_o, and F_{is} were averaged across all loci.

Species	Population	Number successfully sequenced	Mean H_e	Mean H_o	Mean F_{is}	Proportion polymorphic loci
C. fornicata	NL	21	0.176	0.149	0.278	0.789
	NS	24	0.223	0.113	0.591	0.613
	NA	21	0.241	0.108	0.581	0.700
	NY	21	0.187	0.126	0.343	0.825
C. convexa	NA	23	0.269	0.218	0.312	0.746
	BA	19	0.324	0.194	0.465	0.688
	NY	18	0.364	0.161	0.563	0.698
	NJ	25	0.300	0.255	0.325	0.867

Table 3. Pairwise F_{ST} values. Parwise F_{ST} values for four populations of *Crepidula fornicata* (top) and *C. convexa* (bottom) calculated using a set of 1903 SNPs (*C. fornicata*) or 309 SNPs (*C. convexa*). Sites are arranged from north to south.

	NL	**NS**	**NA**	**NY**
NL				
NS	0.0203			
NA	0.0364	0.0440		
NY	0.0356	0.0400	0.0197	

	NA	**BA**	**NY**	**NJ**
NA				
BA	0.0149			
NY	0.0286	0.0255		
NJ	0.0244	0.0251	0.0190	

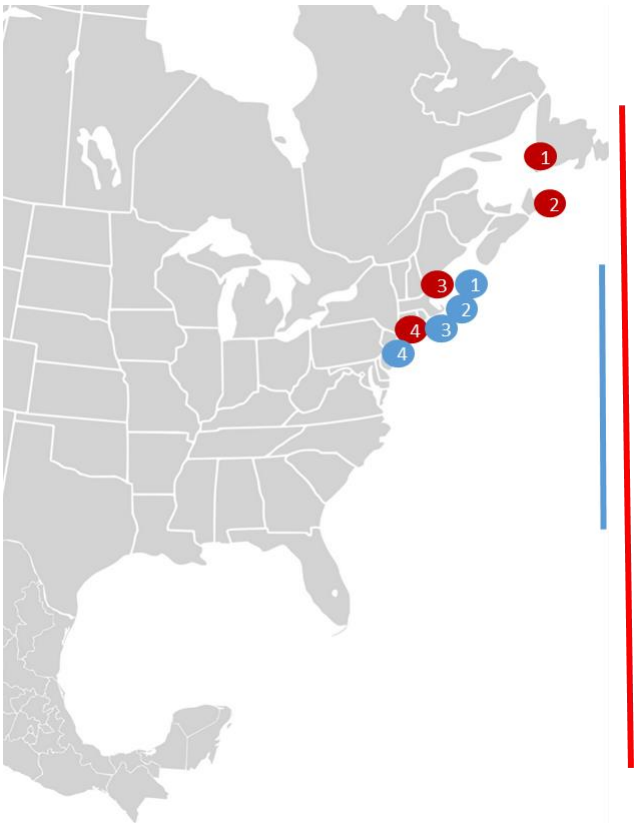

Fig. 3.1 Ranges of *Crepidula* spp. The geographic ranges of *Crepidula fornicata* (red) and *C. convexa* (blue), with sampling sites indicated. Red sites represent collecting locations of *C. fornicata*: 1 = Port Saunders, Newfoundland; 2 = Main-à-Dieu, Nova Scotia; 3 = Nahant, Massachusetts; 4 = Northport, New York. Blue sites represent collecting locations of *C. convexa*: 1 = Nahant, Massachusetts; 2 = Barnstable, Massachusetts; 3 = Northport, New York; 4 = Sandy Hook, New Jersey. Data on ranges from Collin (2001) and Rawlings et al. (2011).

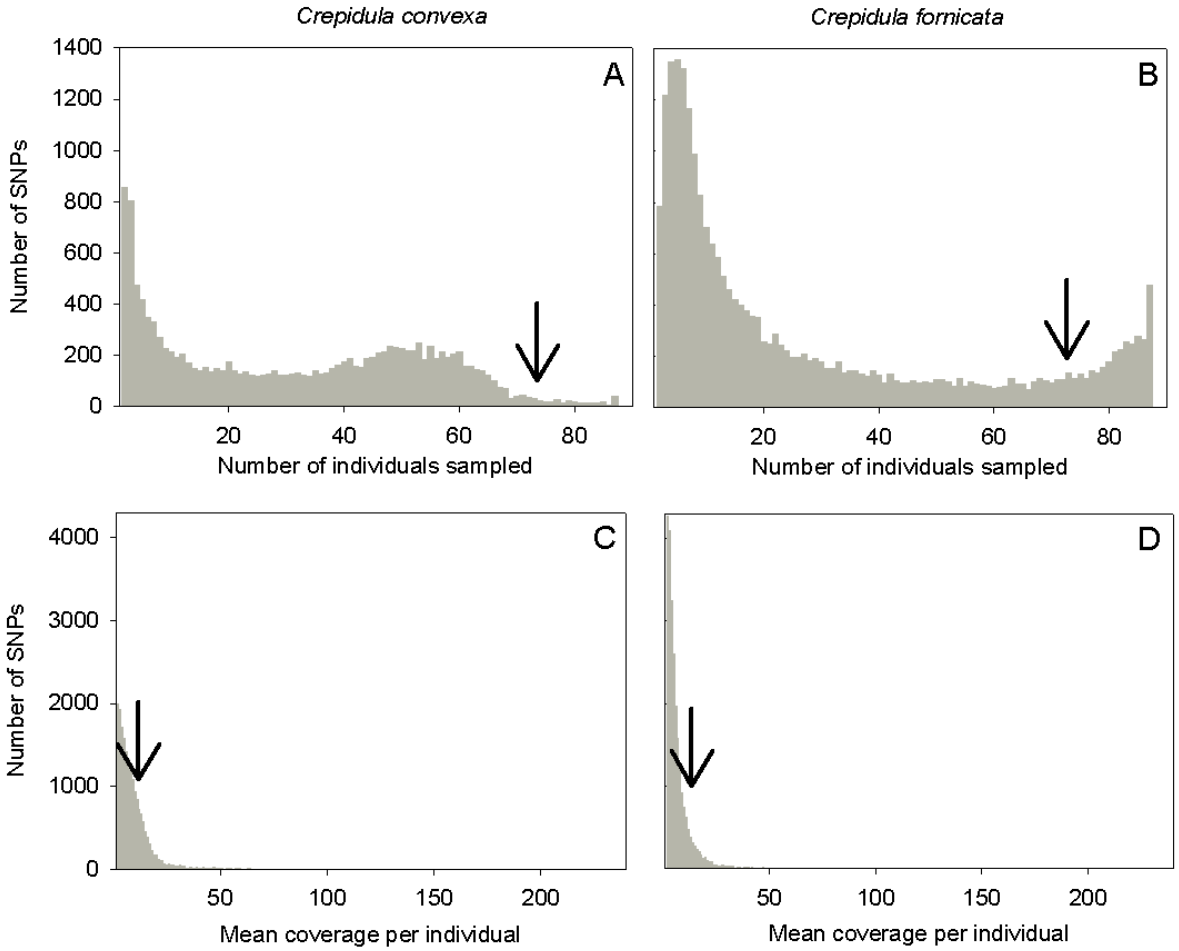

Fig. 3.2 Sampling coverage. Histograms showing sequencing coverage. Top panels show the number of individuals sampled per SNP in *Crepidula convexa* (A) and *C. fornicata* (B). Bottom panels show the distribution of mean coverage per individual in *C. convexa* (C) and *C. fornicata* (D). Arrows indicate the points at which the distributions were filtered; in each panel, all SNPs to the left of the arrows were discarded for further analyses.

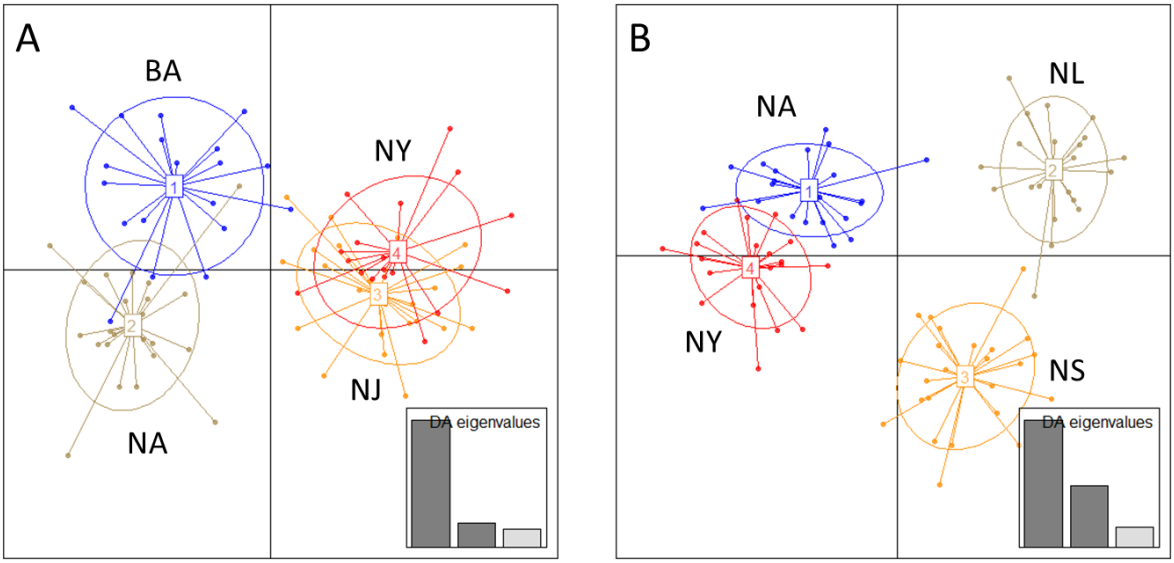

Fig. 3.3 Discriminant analysis of principal components. Scatter plot of a discriminant analysis of principal components for *Crepidula convexa* (A) and *C. fornicata* (B). Each color represents a different genetic cluster. The 67% inertial ellipses around each cluster represent the variance of the two PCs depicted. The insets represent the relative magnitude of the eigenvalues of the first three principal components. Site abbreviations for *C. convexa*: NA = Nahant, MA; BA = Barnstable, MA; NY = Northport, NY; NJ = Sandy Hook, NJ. Site abbreviations for *C. fornicata*: NY = Northport, New York; NA = Nahant, MA; NL = Port Saunders, Newfoundland; NS = Main-à-Dieu, Nova Scotia.

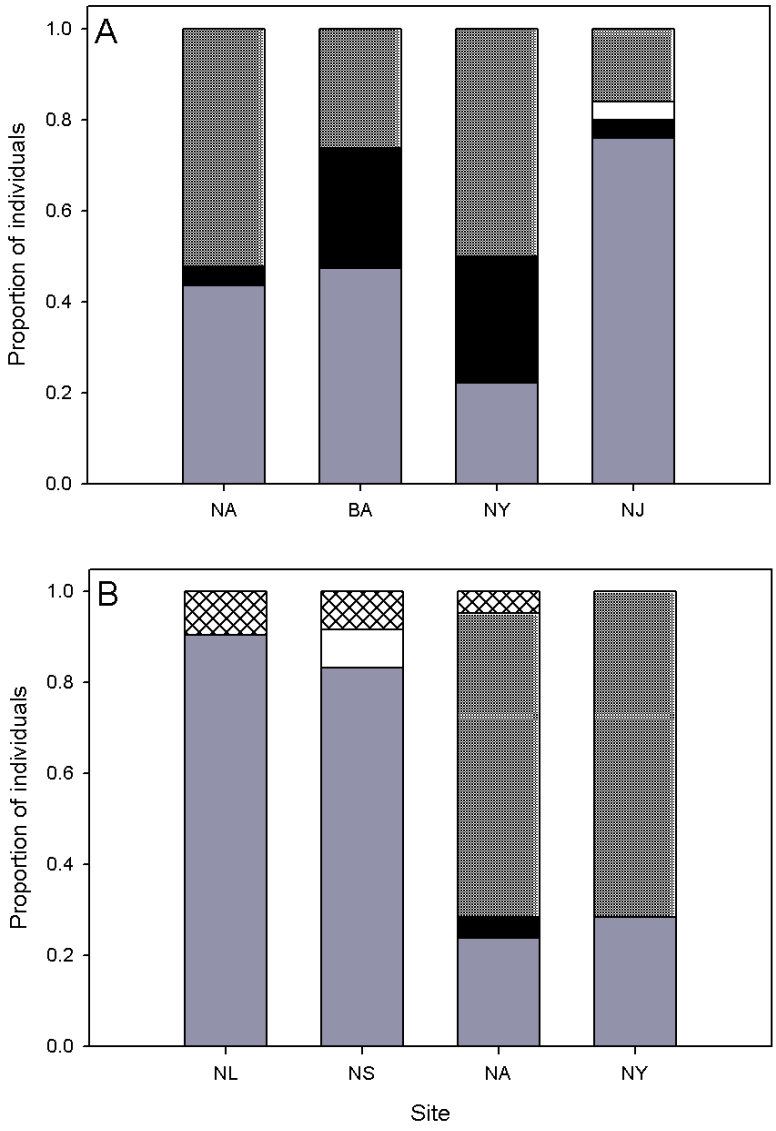

Fig. 3.4 Clustering of individuals. Proportions of individuals from 4 populations of *Crepidula convexa* (A) and *C. fornicata* (B) assigned to different genetic groups (k = 4 in panel A and k = 5 in panel B) in a clustering analysis (discriminant analysis of principal components). The analysis groups individuals to get the smallest within-cluster genetic variance and the largest among-cluster genetic variance. Each of the four genetic clusters within each species is indicated by a different color and pattern combination. Species were analyzed separately, so the groups in the two panels are independent of each other (i.e., a black color in panel A does not indicate an affinity with the black color in panel B). Sites are arranged from north to south along the x-axis, and abbreviations correspond to those in Table 3.1.

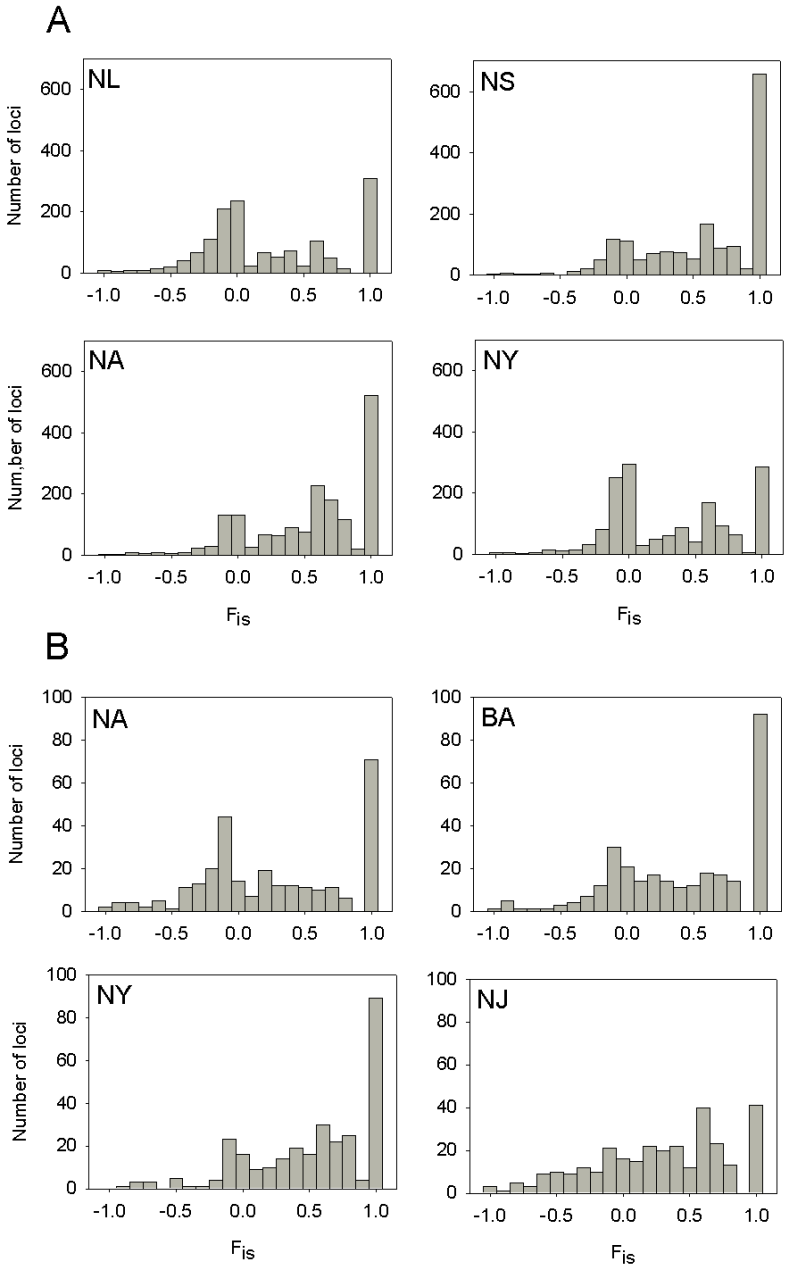

Fig. 3.5 Fis distributions. Histograms showing the distributions of F_{is} values within each population of *Crepidula fornicata* (panel A) and *C. convexa* (panel B). Y-axes represent the number of SNP loci, and site abbreviations correspond to those listed in Table 3.1.

Chapter 4: Adult density affects larval recruitment in the calyptraeid gastropod *Crepidula fornicata*

4.1. Introduction

For those invertebrate species with sessile, benthic adults that produce planktonic larvae, larval dispersal is the crucial first step in determining the distribution of adults (reviewed in Cowen and Sponaugle, 2009). Larvae disperse in the water column for a period of hours to years (Strathmann and Strathmann, 2007; Shanks, 2009), and then must shift from a planktonic habitat to the benthic habitat that they will occupy as an adult. A series of developmental and behavioral steps leads the planktonic larvae to the proximity of a suitable place for the adult to eventually grow and reproduce, although larval wastage to inappropriate habitats may be enormous (Thorson 1950). If the larva does reach an appropriate habitat, a final transition, known as settlement, involves final small-scale movements to a final location, and metamorphosis from a planktonic to a juvenile body plan (Pawlik, 1992).

The cues that invertebrates use to initiate metamorphosis and settlement are diverse and often poorly characterized (reviewed by Pawlik, 1992; Hadfield and Paul, 2001). However, many species settle in response to cues associated with adult conspecifics (e.g., the barnacle *Semibalanus balanoides*, Gabbott and Larman, 1987; the polychaete *Phragmatopoma californica*, Jensen and Morse, 1984; the oyster *Crassostrea virginica*, Zimmer-Faust and Tamburri, 1994).

Larvae of the calptraeid gastropod *Crepidula fornicata* (Linnaeus, 1758) have a 2-4 week planktonic phase (Collin, 2003), metamorphose, and then spend the rest of their life as sedentary, suspension-feeding adults (Collin, 1995). Adults of this species live in semi-permanent aggregations of multiple individuals, often referred to as stacks (Collin, 1995). *Crepidula*

fornicata is a protandrous species, meaning that all settling individuals must go through a male phase before reproducing as a female. Once settled, adults do not move among stacks (Collin, 1995) and mate within the stacks (Dupont et al., 2006; Le Cam et al., 2009) via internal fertilization (Collin, 1995). The sessile and gregarious lifestyle of *C. fornicata* means that the ability of a larva to locate a group of conspecifics should increase its potential fitness after metamorphosis, as an individual that is unable to locate a stack of conspecifics is dependent on subsequent settlers for reproduction.

Larvae of *Crepidula* spp. settle in response a water-borne cue from adult conspecifics (Pechenik and Heyman, 1987; McGee and Targett, 1989; Pechenik and Gee, 1993; Zhao and Qian, 2002), as well as congeners (McGee and Targett, 1989). Larvae also respond to other environmental cues, including biofilms (Pechenik and Gee, 1993; Zhao and Qian, 2002), dibromomethane (a compound made by coralline algae; Taris et al., 2010), and cues from other molluscan shells (e.g., *Busycon*; McGee and Targett, 1989). In studies that test multiple natural cues, the cues from conspecific adults consistently elicit the highest settlement response in the lab (McGee and Targett, 1989, Bohn et al., 2013b), though similar patterns are not always observed in the field (Bohn et al., 2013a,b).

The number of adults (i.e., amount of cue) required to elicit a response is not known, nor is the effect of adult density on larval responses. Although *C. fornicata* can reach very high densities (> 1000 individuals / m^2) in its introduced European range (Erhold et al., 1998) as well as its native range (Hoch and Cahill, 2012), densities can vary over two or three orders of magnitude both among and within sites (from < 10 ind/m^2 to > 1000 ind/m^2; Hoch and Cahill, 2012). The effect of patchy adult distributions on the ability of *C. fornicata* larvae to locate

conspecific adults is thus an important factor in understanding the ecology of this species, particularly population dynamics in both its native and introduced ranges.

Here, I used a manipulative field study to test the hypothesis that adult density influenced larval recruitment in *C. fornicata*. Based on the fact that adult conspecifics produce a waterborne settlement cue, I predicted that increasing adult density would lead to increasing larval recruitment.

4.2. Methods

I collected adult *Crepidula fornicata* from Crab Meadow Beach (Northport, New York, USA: 40°55'46"N, 73°19'38"W) 1-2 weeks before the experiments were deployed (dates below). I brought the animals to the lab, where they were held in a recirculating seawater system (salinity = 30 psu, temperature = 21°C) to be prepared for transfer to the field. During this time the animals were fed 1 liter of Shellfish Diet (Reed Mariculture, San Jose, California) (concentration = 10 million cells/ml) every day.

To test the effect of adult density on settlement, settlement arrays with 5 different densities of adult *C. fornicata* were used. Each of the five treatments, replicated five times, contained a different mass of living *C. fornicata* adults in multiple stacks of two – seven individuals each, attached to empty conspecific shells (i.e., the bottom member of each stack was attached to the empty shell of a dead *C. fornicata*). Although *C. fornicata* larvae preferentially settle on conspecifics, they will also settle on other surfaces. Therefore, in order to control for area available for settlement, each bag also contained artificial substrate in the form of plastic ping-pong balls (37 mm in diameter), which were sanded to create a rough surface and placed in the seawater system for approximately one week to develop a biofilm.

Settlement arrays were prepared using mesh bags constructed from plastic hardware net (30 cm x 30 cm, mesh size 1 cm^2). Stacks were left intact (i.e., were not reassembled from field-collected individuals). Mass was measured as wet tissue mass + shell mass, and large epibionts were removed from shells using a wire brush before weighing. Each ball was approximately equivalent to the surface area (SA) of 25g of adult snails calculated as a semi-ellipsoid (ping-pong ball SA = 4300 mm^2; average SA of 25 g snails = 5100 mm^2, n = 6 stacks, SD = 598 mm^2), and was full of air. The number of balls per bag varied inversely in proportion to the snail mass in the bag. Since all bags were the same size, varying adult snail mass was equivalent to varying adult density. Balls and snails were not fixed in place, but loose in the bags.

The five treatments were 0g snails (no snails present + 40 balls), approx. 25g snails (39 balls + an average of 8 snails, range = 7 – 11 snails), approx. 100g snails (36 balls + an average of 28 snails, range = 20 – 44 snails), approx. 500g snails (20 balls + an average of 134 snails, range = 115 – 160 snails), and approx. 1000g snails (0 balls + an average of 257 snails). The ranges in adult mass for each treatment were as follows: 25 g snails (range = 22.84 g – 30.09 g); 100 g snails (range = 95.01 g – 108.3 g); 500 g snails (range = 489.69 g – 510.13 g); 1000 g snails (range = 993.05 g – 1028.84 g). All snails < 5 mm in length were removed from stacks to avoid mistaking outplanted individuals for new recruits at the end of the experiment. Twenty-five bags were deployed in a randomized blocked design with five blocks of each of the five treatments. Within a block, bags were spaced 1 m apart (Fig. 4.1). Blocks were placed 5 m apart on a transect parallel to the shore. The substrate is sandy at the study site, making it unlikely that small recruits crawled among bags once they had metamorphosed.

Three replicates of the experiment were deployed at the Southampton Marine Park at the Ponquogue Bridge, Southampton, New York, USA (40°50'25"N, 72°29'56"W) on 1 August

2012, 20 June 2013, and 18 July 2013, with five blocks (25 bags) deployed on each date. In New York, *C. fornicata* begins brooding larvae in late April to mid-May, and although females can be found with eggs as late as October, most recruitment happens in June – August (pers. obs.).

The site was in Shinnecock Bay, an area known to have large *C. fornicata* populations (Hoch and Cahill, 2012), but the immediate vicinity of the arrays did not have large numbers of *C. fornicata*: the nearest population of adults was found approximately 0.5 km away. The experimental arrays were therefore the main source of *C. fornicata* cue at the site. The presence of *C. fornicata* larvae near the arrays was confirmed with a plankton tow each time the arrays were deployed. I found *C. fornicata* larvae in each plankton tow taken at the site, but the density of larvae was not quantified. Each mesh bag was attached to two cement bricks (23 cm x 11.5 cm x 6.5 cm) using cable ties. The bags were in contact with the bricks, but were not in contact with the substrate. They were placed at a tidal height of ca. -0.5 m (measured from mean low water). Spring low tide in the summer at the site is approximately -0.1 m (data from tidesandcurrents.noaa.gov), so the experimental animals were never exposed to the air. Bags were retrieved after two weeks following each deployment and returned to the lab, where all settled individuals (< 5 mm) were located by eye, removed from the substrate, and counted. The stacks and balls were not heavily fouled from their two weeks in the field, and settled individuals were easily visible.

I used ordinary least squares regression to test for an increase in settlement with increasing adult mass (density), with adult mass as the predictor variable and number of larvae settled per bag as the response variable. I used the total number settled per bag (i.e., the sum of settlement on balls and snails). The number of settled *C. fornicata* per bag was $\log_{10}$-

transformed to meet the assumption of normality. I used analysis of variance to test for differences among blocks within a deployment date (i.e., experimental run). In cases where there were no differences among blocks, I pooled data from all blocks within a single experimental run for regression analyses. These analyses were conducted in R, version 3.0.1 (R Core Team, 2013).

I tested to see if increased settlement on conspecific shell was explained by available surface area of shell or increased above this expectation (e.g., due to an attractive cue coming from a different, more concentrated source). For each block in the experiment, I used the total number of snails that settled on balls in the 0 g snail treatment (the negative control, with 40 balls) as the expectation for settlement in a bag if snails were only responding to surface area. I calculated the expected value for each treatment within a block using the following formula: $Expected_{snails} = (Settled_{control} / 40) * Balls_{equivalent}$, where $Settled_{control}$ is the total number of larvae settled on the 0 g snail treatment, and $Balls_{equivalent}$ is the number of balls equal to the surface area of snails in each bag (i.e., 1, 4, 20, or 40 for treatments with 25 g, 100 g, 500 g, and 1000 g snails, respectively).

Each bag therefore had an expected number of larvae settled on snails if larvae were responding purely to surface area, and an observed number of larvae that settled on snails. I used paired t-tests to test for a difference between the expected and observed values, with a null expectation of no difference between them. Data were combined across all runs of the experiment, and I conducted four two-sided, paired t-tests (one for each treatment).

To test for preferential recruitment to snails or balls within a bag (i.e., was substrate choice random on a small scale), I compared the number of juvenile snails on these two

substrates in the three treatments that contained both snails and balls. These treatments were 25 g snails : 39 balls (a 1:39 ratio by surface area, as the surface area of 25 g snails is approximately equivalent to that of a single ball), 100 g snails : 36 balls (1:9 by surface area), and 500 g snails : 20 balls (1:1 by surface area). For each treatment, I pooled data across all five blocks and three runs of the experiment and conducted three replicated-goodness-of-fit tests (one for each treatment; Sokal and Rohlf, 1995) to test for the fit of the observed data to the theoretical expectation that settlement would be proportional to the area of substrate present in each treatment (e.g., for the treatment with a 1:1 ratio the null hypothesis was equal settlement on both substrates).

Within bags, I calculated the excess percentage of larval settlement on snails using the following formula: Deviation = $(Snails_{observed} - Snails_{expected}) / (Snails_{observed} + Balls_{observed})$, where the denominator is equivalent to total settlement in a bag. $Snails_{observed}$ was the number of settlers on *C. fornicata* adults, $Balls_{observed}$ was the number of settlers on balls, and $Snails_{expected}$ was the number of expected settlers on *C. fornicata* adults. This deviation was calculated for each bag, and expected values were calculated using the 1 : 1, 1 : 9, and 1 : 39 ratios for each treatment. I then compared these deviations among treatments using a Kruskal-Wallis test due to failure to meet the assumption of normality.

4.3. Results

There were no significant block effects in any experimental run (Table 4.1, all $F_{1,23} < 0.3$, all $p > 0.6$), so data were pooled across blocks within each run for regression analyses. The OLS regressions for all three of the experimental runs were significant, indicating that *C. fornicata* recruitment increases with increasing adult density (August 2012: $y = 0.00048x + 0.5766$, $p =$

0.005; June 2013: $y = 0.00055x + 0.7716$, $p < 0.001$; July 2013: $y = 0.00038x + 1.037$, $p = 0.007$; Fig. 4.2). There were significant correlations between adult mass and the number of larvae settled for all three experimental runs (August 2012: $r = 0.569$, $p = 0.005$; June 2013: $r = 0.730$, $p < 0.001$; July 2013: $r = 0.533$, $p = 0.007$).

The number of snails settling on conspecifics was greater than expected based on surface area alone in treatments with 100 g snails (paired $T_{14} = 2.27$, $p = 0.040$), 500 g snails (paired $T_{13} = 5.28$, $p < 0.001$), and 1000 g snails (paired $T_{14} = 4.32$, $p < 0.001$). However, larvae in the treatment with 25 g snails did not settle on conspecifics more often than would be expected based on surface area (paired $T_{13} = 1.78$, $p = 0.10$; Fig. 4.3).

For all three treatments that contained both balls and snails, there was a significant excess in the number of settlers on snails versus balls within a bag (Fig. 4.4). The 1:39 (snails: balls) treatment had an average 18.7% excess above expected values (n = 14 bags, 33% standard deviation; overall fit $G_{13} = 169.8$, $p < 0.001$), the 1:9 treatment had an average 17.6% excess (n = 15 bags, 24% standard deviation; overall fit $G_{14} = 72.74$, $p < 0.001$), and the 1:1 treatment had an average 14.7% excess (n = 14 bags, 19% standard deviation; overall fit $G_{13} = 39.5$, $p = 0.002$). No trend of changing excess of settlers was seen among the different treatments ($\chi^2_2 = 0.860$, $p = 0.651$).

4.4. Discussion

Under field conditions, this experiment demonstrates that increasing density of *Crepidula fornicata* adults increased the amount of larval recruitment (Fig. 4.2). This pattern was consistent across multiple experimental runs in multiple years, despite potential seasonal and spatial variation in larval availability. Since larval density was not quantified throughout the

experiment, and therefore larval supply is unknown, it is not possible to compare the overall amount of settlement across experimental runs. This means that the relationship between adult mass and number of settlers is consistent despite potentially different amounts of available larvae in the bay. Correlation coefficients ranged from 0.533 – 0.73, indicating that only 28% to 53% of the variation in larval settlement was attributable to changes in adult density.

Larvae may prefer to settle on conspecific shells instead of plastic balls based on physical aspects of the substrate such as texture. However, the amount of settlement on conspecific shells was significantly higher than what would be expected based on the effect of surface area alone (Fig. 4.3). This shows that larvae are aggregating to conspecifics, consistent with the presence of a waterborne chemical cue produced by adult conspecifics. This cue attracts larvae to adults, induces larval settlement, and must be effective at least on the spatial scale of the size of the bag. There was no aggregation effect observed in treatments with 25 g snails but this effect was observed in the 100 g snails treatment, indicating that the threshold at which larvae are able to detect cue from conspecifics was reached between 25 g and 100 g adults per 0.09 m^2 (the density of adult snails in the experimental bags), or roughly 90 – 300 individuals m^{-2}.

In addition to the cue detected in this experiment that attracted snails to bags on the scale of the bag size to perhaps meters, preferential recruitment to conspecifics occurred on the scale of centimeters. Within bags that contained both snails and ping-pong balls, there was significantly higher settlement on snails over balls on a per-area basis (Fig. 4.4), consistent with both attraction to conspecifics by a chemical cue, or with post-settlement movement towards conspecifics. However, I was not able to eliminate the possibility that the difference in recruitment after two weeks was due to differential mortality between substrates, rather than differences in recruitment between substrates. No rigorous study exists of movement rates and

patterns in adult or juvenile *C. fornicata*. Collin (1995) saw low rates of movement between stacks for *C. fornicata*, although paternity analyses have found that putative fathers of larvae are not always found in the stack with the mothers (Dupont et al., 2006; Proestou et al., 2008; Le Cam et al., 2009), which could be due to either mortality or movement. Another possibility to explain higher settlement on snails than balls is that waterborne cue acts on a large scale to attract larvae to the benthic habitat, followed by another system that induces metamorphosis more effectively on snails than on balls. Distinguishing among these hypotheses will require fine-scale temporal measurements of settlement rates and observations of post-settlement movement.

Living on conspecifics will ultimately be important to *C. fornicata* adults. Since this sessile species is protandrous (the sex of an individual changes from male to female over time), and sperm transfer occurs by copulation, adults must be within reach of a mate in order to reproduce. Nearly all offspring are produced from matings of snails within stacks (Dupont et al., 2006, Le Cam et al., 2009). The results presented here are consistent with previous lab and field studies in *C. fornicata* that show preferential settlement on adult conspecifics and in response to conspecific cues (McGee and Target, 1989; Pechenik and Gee, 1993; but see Bohn et al., 2013a,b), though no studies have shown an effect of increased adult density on settlement or a localized attraction of conspecifics, relative to nearby bare substratum. Increased recruitment in the presence of conspecifics due to chemical cues has also been observed in other taxa (e.g. the barnacle *Semibalanus balanoides*, Gabbott and Larman, 1987; the polychaete worm *Phragmatopoma californica*, Jensen and Morse, 1984; the oyster *Crassostrea virginica*, Zimmer-Faust and Tamburri, 1994).

In this experiment, I cannot rule out the possibility that recruitment was affected by larvae released by the females in the experiment. In order for this to have happened, however, released larvae would have to have been retained in or near the mesh bags for the two week duration of the experiment, and to have completed the entirety of their planktonic larval development and metamorphosis within two weeks. This is unlikely to be the case. Firstly, the mesh size of the bags (1 cm^2) was many times larger than a newly-hatched veliger (approximately 250 µm). Secondly, although larvae can complete development in two weeks, this is at the lower end of their development time (i.e., 2-4 weeks; Collin, 2003), and it is unlikely that larvae would be released and also metamorphose during a single experimental run.

Increases in larval recruitment with increasing adult density may affect adult distributions in multiple ways. Larvae are attracted to a chemical cue when adult density is somewhere between 100 individuals m^{-2} and 300 individuals m^{-2} (Fig. 4.3), densities that are observed at many sites around Long Island (Hoch and Cahill, 2012). The process of increasing attraction to high densities of conspecifics may create a positive feedback loop such that areas with high adult densities become progressively denser through time. This may partially explain the very high densities (> 1000 ind. m^{-2}) seen in some areas of *C. fornicata*'s native (Hoch and Cahill, 2012) and introduced (Ehrhold et al., 1998) ranges. High densities, especially within semi-enclosed bodies of water, may also be due to factors acting at larger spatial scales, such as currents and larval retention (Cowen et al., 2002; Cowen and Sponaugle, 2009). However, my results show that feedback loops driven by selective settlement of larvae may be possible on the scale of meters, increasing the degree of aggregation (i.e., patchiness) within larger areas.

Selective settlement and recruitment may play a role in determining the range limits of *C. fornicata*. Areas with low adult density, below the threshold where larvae are attracted to a

chemical cue, (e.g., near geographic range margins) will be less likely to attract settlers via conspecific waterborne cues. Such selective settlement may slow the speed at which the species can expand its range, which is analogous to an Allee effect limiting the species' range (Keitt et al., 2001; Kubisch et al., 2014). The northern edge of the native range of *C. fornicata* is in Newfoundland, Canada, and has been moving northwards in the past decades (Rawlings et al., 2011). Recent distribution modeling within the native range predicts a continuing northward shift of suitable habitat for the species based on IPCC emissions projections and environmental variables like sea surface temperature and phytoplankton concentration (Saupe et al., 2014). The European range, where the species has been introduced, is also expanding northwards, apparently in response to warming temperatures (Thieltges et al., 2004; Bohn et al., 2012). The future rate of expansion is unclear, but will be determined in part by the ability of larvae to identify appropriate habitats and find conspecifics for mating.

The relative importance of Allee effects in limiting range expansion will be affected by the strength of a preference is for settlement on conspecifics. Since *C. fornicata* settles preferentially on conspecifics but does not require them for settlement (e.g., Bohn et al. 2013a,b; Fig. 4.2, this study), this effect is unlikely to strongly inhibit range expansion. Although *C. fornicata* are sessile as adults and copulate, their protandrous life history means that two individuals that settle together (e.g., a common response to another environmental cue) will ultimately be able to mate. Only 28% - 53% of the variance in larval settlement was explained by changes in adult density in this experiment. Other environmental cues may be important at range margins. For example, when tested in the laboratory, *C. fornicata* larvae show increased settlement in response to dibromomethane, a chemical produced by coralline algae. These algae are an important substrate within the species' introduced European range (Taris et al., 2010) and

also an abundant potential substrate in the northern parts of the native North American range of this species (pers. obs.).

The sex ratio of the adult population is another factor that may impact larval settlement behavior, but was not examined here. Since *C. fornicata* is a protandrous species, all settling individuals will be male before they reproduce as a female. This may make larvae or juveniles more likely to recruit to stacks with more females. The effect of adult sex ratio on *C. fornicata* settlement is unknown, and will require future study.

Larval settlement is not the only factor that determines which individuals will enter a population; post-settlement mortality is also important (Thorson, 1950; Connell, 1985). Although the results of the current study demonstrate that selective settlement can play a role in determining adult distributions of *C. fornicata*, the strength and drivers of post-settlement mortality in North America are unknown. Bohn et al. (2013a,b) found that post-settlement mortality was more important than selective settlement in determining distributions of adult *C. fornicata* in an estuary in Wales. However, the site used in these studies was in the intertidal zone while the site in the present study was in the subtidal zone. Organisms in the intertidal zone are subjected to a wider range of abiotic stressors than those in subtidal zones (Moran, 1999). Ruesink et al. (2014) found that post-settlement survival in different habitats, rather than larval habitat selection, determined the distribution of a non-native clam species. Post-settlement mortality may therefore be stronger in the intertidal zone, although mortality due to biotic factors, such as predation (Pechenik et al., 2010), may be greater in subtidal zone areas. A weaker attractive effect of conspecifics in the intertidal zone may also reflect the fact that those adults are exposed for a period of time every day, prohibiting larval recruitment during that time. Further study is necessary to understand the relative importance of selective settlement (in

particular, that directed by conspecific chemical cues) and post-settlement mortality (due to both abiotic and biotic factors) in determining adult distributions of *C. fornicata* under different ecological conditions and different locations.

4.5. Conclusions

Many marine invertebrate larvae, including *Crepidula fornicata*, use waterborne cues from conspecific adults to trigger larval settlement. This settling transition is critical in determining adult distribution. Larval *C. fornicata* settled in greater numbers when the density of adult conspecifics was higher in field experiments and an excess of settlement was attributed to settlement on live adults, as opposed to bare substratum. This relationship may cause feedback loops that affect the distribution of *C. fornicata* both within and among sites.

Table 4.1 Analyses of variance testing for block effects. Analyses of variance for each of the three experimental runs of the experiment (Month), testing for differences among blocks within each run.

Month	Source	SS	df	MS	F	p
August 2012	Block	0.1122	4	0.0281	0.207	0.931
	Error	2.4460	18	0.1359		
	Total	2.5582	22			
June 2013	Block	0.2799	4	0.0670	0.763	0.562
	Error	1.8348	20	0.0917		
	Total	2.1147	24			
July 2013	Block	0.467	4	0.1168	1.586	0.219
	Error	1.399	19	0.0736		
	Total	1.866	23			

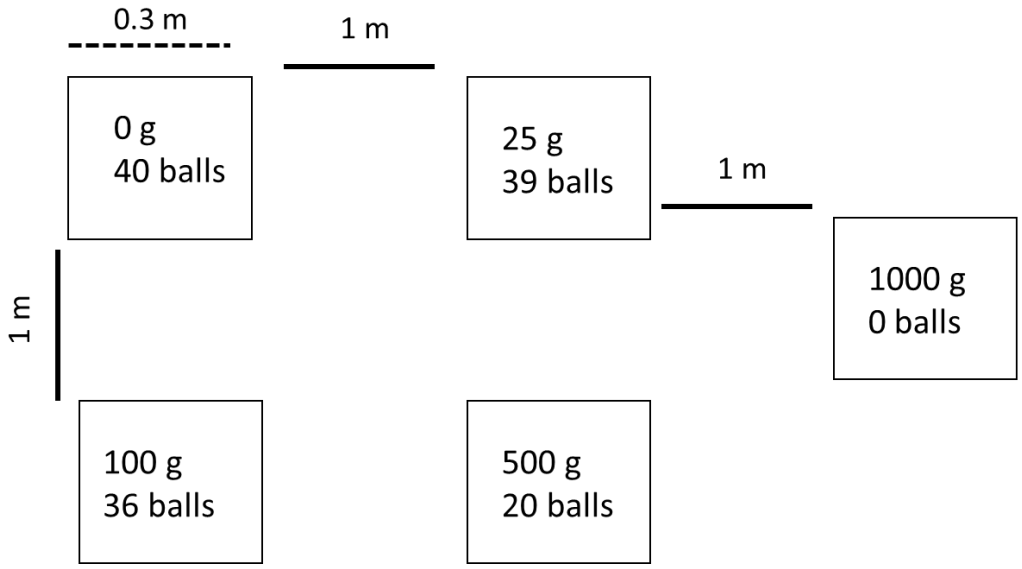

Fig. 4.1 Schematic of settlement arrays. Spatial array of bags containing different masses of snails and different numbers of ping-pong balls in a field experiment. Bags were made of plastic hardware cloth. Five replicated blocks of this set of five treatments were deployed at three different dates, and the placement of treatments with the array was randomized for each block.

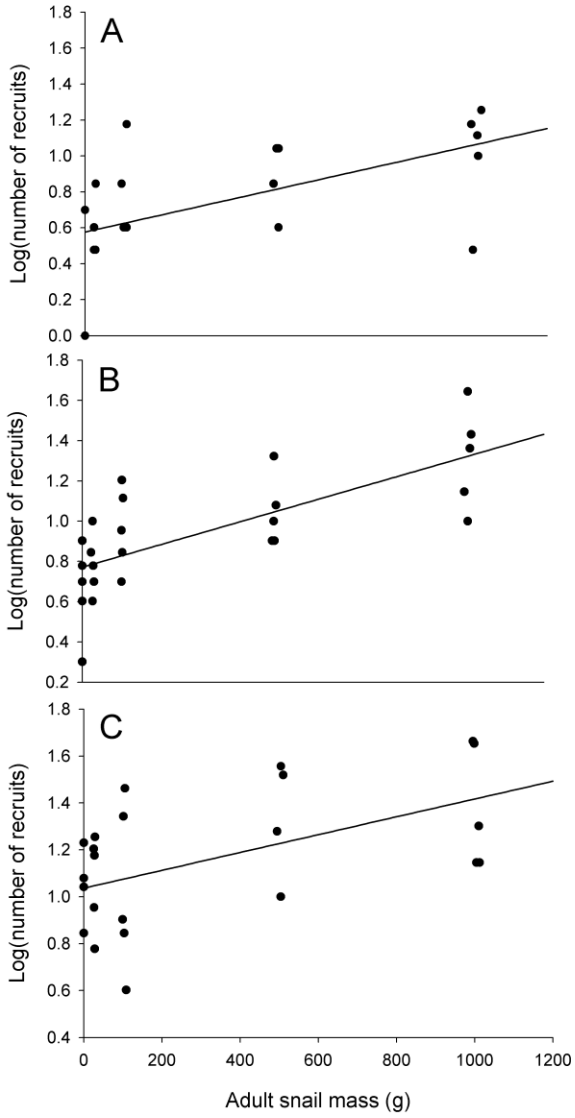

Fig. 4.2 Settlement regressions. Relationship between adult *Crepidula fornicata* mass and log-10-transformed number of *C. fornicata* recruits that settled. Each point represents a bag containing a given mass of adult snails and a number of sanded, biofilmed ping-pong balls to control the area available for settlement (see Methods). Best-fit lines represent OLS regressions. A) data from August 2012 ($y = 0.00048x + 0.5766$, $r = 0.569$); B) data from June 2013 ($y = 0.00055x + 0.7716$, $r = 0.730$); C) data from July 2013 ($y = 0.00038x + 1.037$, $r = 0.533$).

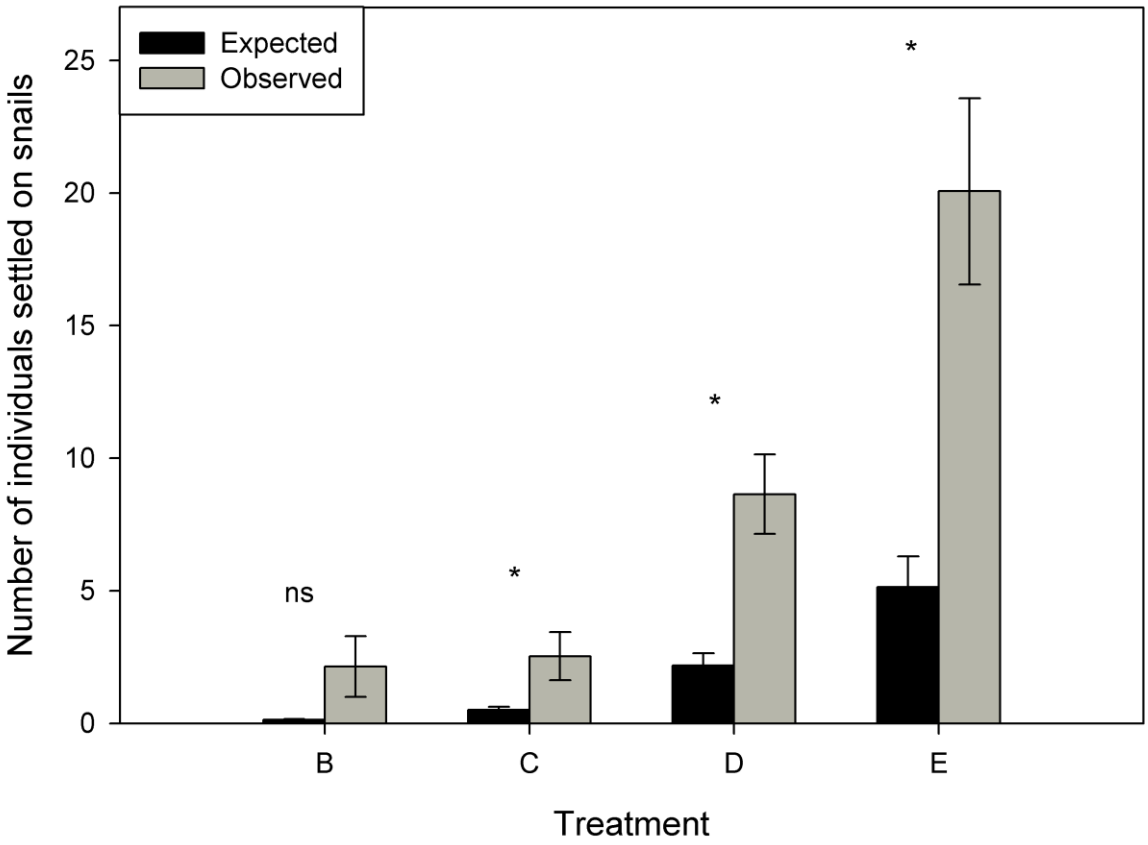

Fig. 4.3 Excess settlement on snails relative to control treatment. Number of individuals settled on conspecifics in each treatment. Dark bars represent the expectation for settlement on conspecifics given the amount of settlement on the negative control (0 g snail treatment); grey bars represent the observed number of individuals settled on conspecifics. Bars represent means for each treatment (± 1 S.E.). Data were combined across all experimental runs. Treatments where observed and expected values are different based on two-tailed paired t-tests are indicated with an *.

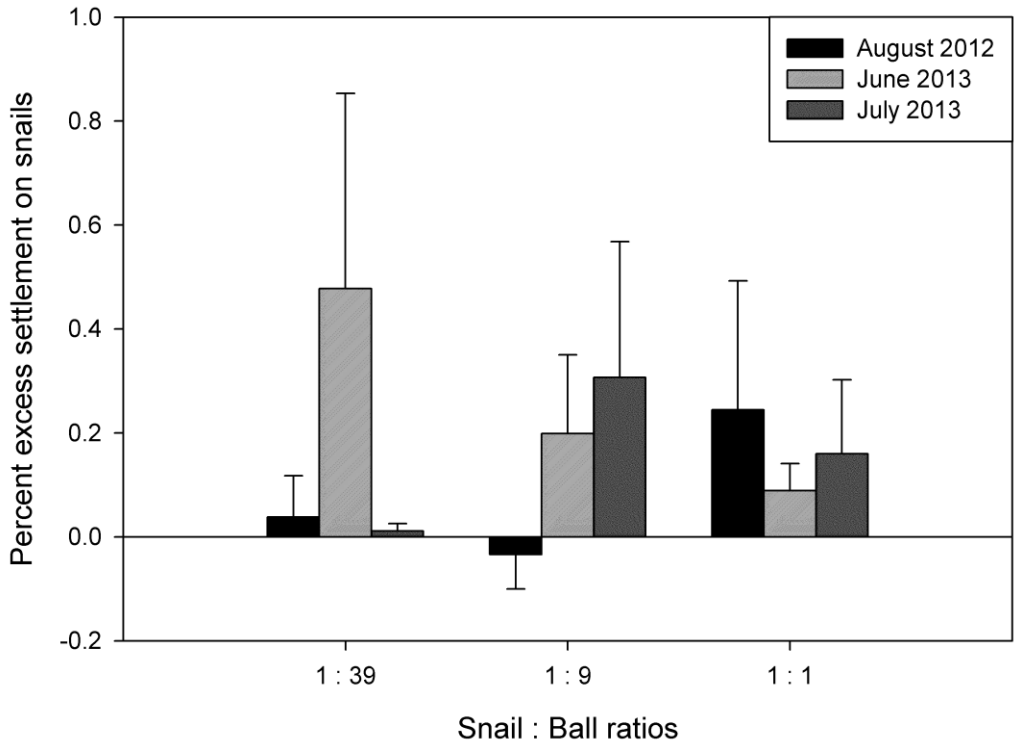

Fig. 4.4 Settlement on different substrates. Excess of settlement on snails in bags containing both adult conspecifics (snails) and ping-pong balls. Ratios displayed are the snail : ball ratios in the treatments. Bars represent mean deviations calculated for each treatment (± 2 S.E.). Data are displayed for each of three runs of the experiment.

Chapter 5: Larval settlement in *Crepidula fornicata* in response to multiple cues from conspecifics

5.1 Introduction

Marine invertebrates with complex life cycles have vastly different ecologies for the adult and larval stages. Planktonic larvae often disperse tens to hundreds of kilometers along a coastline, while the benthic adults are often sedentary and aggregated in spatially restricted habitats (reviewed by Cowen and Sponaugle, 2009). Understanding how broadly dispersed planktonic larvae successfully locate sites where they will survive as adults is an important challenge in marine ecology. The hypothesis of selective settlement states the planktonic-benthic ecological transition is mediated by physiological responses of larvae to physical or chemical cues associated with suitable sites for adults (e.g., Krug and Manzi, 1999) or from the adults themselves (e.g., Zimmer-Faust and Tamburri, 1994). To test this hypothesis with emphasis on cues from adults, I developed an optimized time-course bioassay and present evidence for the induction of settlement by two distinct settlement cues in the gregarious marine gastropod, *Crepidula fornicata* (Gastropoda: Calyptraeidae).

The terminology used to describe the process of transitioning from a larva to a juvenile is variable in the literature. Concomitant with the developmental change of metamorphosis are ecological transitions from planktonic to benthic habitats and from broad distributions to more clumped patterns. I use terms consistent with Pawlik (1992), where settlement refers to the entire process of transitioning from a planktonic larva to a benthic juvenile, while metamorphosis is the part of the process that includes irreversible developmental changes that prevent a larva from returning to its previous planktonic lifestyle.

The larvae of *C. fornicata* have a planktonic period of 2-4 weeks (Collin, 2003), which allows long dispersal distances for larvae and results in low genetic differentiation among populations connected by coastal currents (Collin, 2001; Riquet et al., 2013). Within a single location, larvae also have a wide spatial distribution, demonstrated by plankton tows within a single estuary that found larvae present in areas without adults (Rigal et al., 2010). In contrast, the sedentary adults are patchily distributed within intertidal and shallow subtidal habitats within their native range (Henry et al., 2010; Hoch and Cahill, 2012). Adults exhibit a clumped distribution due to their tendency to form large, semi-permanent mating groups called stacks (Collin, 1995). The presence of small juveniles aggregated on adults demonstrates recruitment to these stacks (McGee & Targett, 1989).

Larvae may settle indiscriminately but suffer selective post-settlement mortality in unsuitable patches (Thorson, 1950), and thus juveniles would be found only on stacks. Alternatively, larvae may use physical or chemical cues to settle selectively in appropriate habitats, avoid inappropriate habitats, or delay metamorphosis until appropriate cues are sensed (Thorson, 1950; Woodin, 1986; Pechenik and Eyster, 1989). Environmental (exogenous) cues that induce settlement may be associated with biofilms (e.g., *Crepidula onyx*, Zhao and Qian, 2002), other species that provide food or habitat (e.g., the hydroid *Proboscidactyla flavicirrata*, Donaldson, 1974; the soft coral *Alcyonium siderium*, Sebens, 1983), conspecifics (e.g., the barnacle *Semibalanus balanoides*, Gabbott and Larman, 1987; the polychaete worm *Phragmatopoma californica*, Jensen and Morse, 1984; the oyster *Crassostrea virginica*, Zimmer-Faust and Tamburri, 1994), or avoidance of species with negative impacts (e.g., the polychaete worm *Pseudopolydora kempi*, Woodin, 1985).

Both selective mortality and selective settlement of *C. fornicata* have been studied in the field. Although survivorship of juveniles can differ across substrate types and locations (Bohn et al., 2013a,b), other studies demonstrate that juveniles aggregate on adults and that increased adult density increases larval recruitment (McGee and Targett, 1989; Chapter 4). Additional support for selective settlement comes from laboratory studies. *Crepidula fornicata* larvae metamorphose in response to dissolved or suspended cues, including increased concentrations of KCl (Pechenik and Heyman, 1987), dibromomethane from coralline algae (Taris et al., 2010), and *C. fornicata* adult-conditioned seawater (Pechenik and Heyman, 1987; Pechenik and Gee, 1993; Bohn et al., 2013b). In some studies, adult-conditioned seawater is prepared in the same vessel in which *C. fornicata* larvae are tested (as in Penniman et al., 2013) confounding the effect of the adult-conditioned water with any potential effect of pedal mucus produced by adult snails. Molluscan pedal mucus has been shown to affect settlement rates in other marine invertebrate larvae, particularly barnacles, which can be repelled (e.g., Johnson and Strathmann, 1989) or attracted (Holmes, 2002) to the mucus. In the case of increased *Semibalanus balanoides* settlement in the presence of *Patella vulgata*, the attraction is due to physical properties of the mucus (Holmes, 2002). It is unclear if the reduction in settlement of *Balanus glandula* in the presence of predatory snails is in response to chemical cues or other mechanisms (e.g., altering the biofilm present on a substrate; Johnson and Strathmann, 1989).

The existence of a conspecific, waterborne cue that induces settlement in *C. fornicata* has been reported with a variety of experimental designs and assay conditions (e.g., Pechenik and Heyman, 1987; Pechenik and Gee, 1993; Bohn et al., 2013b), but the chemical nature of this cue remains unknown. Further characterization of the cue requires assays that control for biological, technical, and statistical sources of variation. Biological variability is due to genetic differences

among larvae, larval age, and differences among egg masses in survival, growth, and development (maternal effects; Hilbish et al., 1999). Technical variability stems from differing biotic and abiotic conditions known to affect larval settlement among previous experiments (e.g., Pechenik and Heyman, 1987; Pechenik and Gee, 1993). Statistical variability results from inferring a settlement rate using data collected at a single time point, as results can be sensitive to the time point selected (time course to metamorphosis varies under different conditions). To address these issues, I implemented a blocked experimental design that accounted for biological variability due to larval age, optimized biotic and abiotic conditions for settlement, and estimated rates of settlement using a single parameter that estimates the time course of settlement. I used an experimental design with time-course bioassay to test for the induction of settlement by adult-conditioned water and conspecific pedal mucus. I also used heat to attempt to experimentally reduce the activity of these settlement factors.

5.2 Methods

5.2.1 *Crepidula fornicata* collection and husbandry

Adult *Crepidula fornicata* were collected from Crab Meadow Beach (Northport, New York, USA: 40°55'46"N, 73°19'38"W) and returned to the lab the same day; collection occurred at spring tides in July, August, and September of 2013 for Experiments 1, 2, and 3, respectively. Adult females were removed from their substrates to check for incubating egg capsules. Capsules with larvae that were near hatching were selected and hatched by physically agitating them in a bowl of filtered seawater at room temperature. Larvae from multiple females that hatched at the same time were combined for larval rearing (two or three females per experiment). Larvae were reared in cultures of 800 ml of 1 µm-filtered seawater (FSW) at a concentration of one larva per four ml (i.e. cultures started with approximately 200 larvae each), and were fed

40,000 cells/ml of the alga *Isochrysis galbana* (clone T-Iso) daily. Seawater was collected from an underground well at Flax Pond Marine Laboratories, Old Field, New York (40°57'49"N, 73°08'26"W), and all water was filtered to 1 µm using a bag filter before use. Larval cultures were reared at 20°C and FSW was replaced via reverse filtration every three to four days. Larvae were tested every two days for competence (ability to metamorphose) once they developed shell brims (Pechenik, 1984) and were at least 750 µm long (Pechenik and Heyman, 1987). Competence was tested by placing 12-24 larvae (1-2 larvae from each culture) in 20 mM KCl solution for 8 hours. The bioassay was initiated within 24 hours of a group of larvae being designated as competent (75% of larvae metamorphosed in response to KCl; Pechenik and Heyman, 1987).

5.2.2 Bioassay

Prior to the start of the experiment, all glassware was acid-cleaned in 10% concentrated HCl, rinsed in deionized water, and autoclaved. To create adult-conditioned water (ACW) I placed 100 g of adult *C. fornicata* (shell and wet tissue mass) and one liter of FSW into a beaker without food, and aerated the water for twelve hours with an aquarium air pump. Large epibionts (e.g., barnacles, macroalgae) were removed from the shells, but shells were not otherwise treated. One liter of FSW was aerated as a control. After twelve hours, adults were removed and ACW and FSW were filtered to 40µm with a Nitex mesh filter. The ACW and FSW were then prepared for each specific experiment as described below. Temperature, dissolved oxygen, salinity, and pH were measured at all preparation steps. The parameters for the bioassay are listed in Table 5.1; these values were chosen based on optimization studies (Appendix B).

For the pedal mucus treatment (PMG), a single small (~15 mm) adult *C. fornicata* was added to all 60 ml glasses for 12 hours at the same time that the ACW was prepared. Glasses were filled with 35 ml FSW and covered to prevent evaporation and snail escape. During ACW preparation, glasses not receiving the PMG treatment had adults removed and were acid-cleaned and autoclaved to remove the mucus. Glasses with the PMG treatment had adults removed and were drained immediately before the experiment.

Each glass in the bioassay contained 20 ml of FSW or the test solution. Ten larvae were individually pipetted into the glass. Larval growth and development in many marine larvae, including *C. fornicata*, varies among rearing beakers, so this was accounted for by placing one larva from each 800 ml rearing beaker (ten total beakers) into each replicate glass. Therefore, rearing beaker and treatment glass were not confounded. The same set of rearing beakers was used in all blocks of each experiment.

Every 12 hours, the number of larvae metamorphosed in each glass was counted, and any mortality was recorded. Metamorphosed juveniles and dead larvae were removed from the trial at each time point. After 24 hours new solutions and glasses were prepared as described above for all replicates. Larvae were individually pipetted into the new glasses. The total time of the experiment was 48 h, which included five sampling time points and two different preparations of ACW and PMG. Larvae were not fed during the experiment.

5.2.3 Modeling Settlement Rates

I modeled larval settlement by predicting the proportion of larvae settled *(y)* at time *(t)* using the cumulative distribution function for the single-parameter exponential model:

$$y = 1 - e^{-\lambda t}$$

Given a constant probability of settlement, the waiting times for a single individual to settle *(t_i)* under a given treatment are exponentially distributed. Note that this model assumes that all larvae in the experiment are developmentally capable of settling (competent), although I began trials when 75% of larvae were competent. The overall results should not be affected if competence was equal in all treatments, a reasonable assumption given my random assignment of larvae to treatments. The exponential distribution is defined by the single parameter λ , which can be estimated as $\hat{\lambda} = n \,/\, \sum_{i=1}^{n} t_i$. However, because not all larvae settled during the first 48 hours, I calculated λ incorporating Type I censoring with the following equation:

$$\hat{\lambda} = \frac{r}{\sum_{i=1}^{r} t_i + \ T(n-r)}$$

where *n* is the total number of larvae tested and *r* is the number of the larvae that settle during the time course bioassay. Thus,$(n-r)$ is the number of non-metamorphosed larvae at time *T* which represents the end of the time course bioassay (fixed at 48 hours for all experiments).

The advantage of modeling larval settlement is that I was able to summarize the settlement rate for a replicate with a single value ($\hat{\lambda}$) that used time course data and also accounted for Type I censoring. Repeated measures ANOVA can also be used for time-course data, but requires more degrees of freedom than estimating a single parameter for the data because the degrees of freedom associated with a repeated-measures model is equal to the number of timepoints minus one, in addition to degrees of freedom used in treatment-by-time interactions (von Ende, 2001). By using lambda, I was able to conduct larger factorial experiments than would have been possible with a repeated-measures design. However, due to variation in larval batches and in cue preparation, results (values of lambda) are not comparable across experiments.

Statistical analysis of all experiments consisted of correcting the number of larvae tested (*n*) for mortality (average mortality per block = 2% (±1.8% SD), or approximately five larvae (±5 larvae); mortality was not statistically different among treatments) and then calculating $\hat{\lambda}$ for each replicate. I then analyzed $\widehat{\lambda}$ as a response variable in an analysis of variance framework as described below for each experiment. All statistics were conducted using JMPIN (Version 4.0.4, © SAS Institute 2001).

5.2.4 Experiment 1: Screening factors that induce settlement

The first experiment tested the effects of three factors on induction of larval settlement: ACW, PMG, and 20 mM KCl. I conducted a randomized complete block design with three blocks over ten days; all blocks used the same batch of larvae (derived from three females, with the three families mixed and reared in 10 beakers), therefore larvae in the later blocks were older. Each treatment combination had three replicate glasses for 24 glasses per block (72 total). All three factors (ACW, PMG, KCl) had two levels (present or absent) and were analyzed using a factorial design (eight possible treatments). The ACW and KCl treatments allowed validation of my time-course bioassay by comparison with previously published results demonstrating the inductive effects of these factors (e.g., Pechenik and Heyman, 1987). The treatment where all factors were absent was equivalent to FSW and served as a negative control in my analysis.

Statistical analysis of the modeled rate parameter (λ) was conducted with blocks as random effects and experimental treatment factors (KCl, ACW, and PMG) modeled as fixed effects. Planned comparisons were conducted of each factor against the control (H_0: KCl = FSW, ACW =FSW, and PMG = FSW). Significance of planned comparison was assessed with critical values calculated by Dunnett's procedure, which provides a correction for testing

multiple treatments against a common control (i.e., non-orthogonal comparisons; Montgomery, 2012).

5.2.5 Experiment 2: Inactivating factors that induce settlement

Using the same experimental design as Experiment 1, I tested whether heat would affect the induction of settlement. Larvae were derived from two females, with the two families mixed and reared in 10 beakers. The factors tested in this experiment were ACW, PMG, and heat, each with two levels (present or absent). The combination where all factors were absent was equivalent to FSW and served as a negative control. Test solutions (FSW and ACW) were microwaved to a rolling boil for two minutes. PMG were heated by microwaving for two minutes without water, bringing the mucus to a temperature of approximately 120°C. Glasses were then cooled before the addition of the test solutions. Following the application of heat, all experimental solutions were oxygenated for one hour with an aquarium pump and airstone, bringing all treatments to the same temperature and dissolved oxygen levels. As in Experiment 1, I used a randomized complete block design with three blocks over ten days from the same larval batch. The experiment had three replicate glasses per treatment for 24 glasses per block (72 total glasses per experiment). Additionally, three replicates of 20 mM KCl in FSW were run as a positive control for larval competence in each block.

Rate parameters (λ) for Experiment 2 were also analyzed as a randomized complete block design. As in Experiment 1, a mixed model with block as random effects and treatments as fixed effects was analyzed. However, for Experiment 2 I also tested whether heat affected factors that can induce settlement. Therefore, the orthogonal planned comparisons contrasted heated (H) and non-heated treatments (H_0: FSW = H*FSW, ACW = H*ACW, and PMG vs. H*PMG) using Fisher's Least Significant Differences (Montgomery, 2012).

5.2.6 Experiment 3: Localizing factors that induce settlement

To see if the cue was localized in the ACW or the pedal mucus, I also used heat as a level in a two-way ANOVA design. Larvae were derived from two females, with the two families mixed and reared in 10 beakers. Each factor, ACW and PMG, had three levels: absent (0), present unheated (+), and present heated (H). The treatment where both factors were absent was equivalent to FSW, the negative control. Heated ACW and PMG were prepared as described in Experiment 2. Experiment 3 was run as a single block with each treatment combination having three replicate glasses for 27 glasses total, plus three replicates of 20 mM KCl as a positive control. I therefore isolated the effects of heating ACW and PMG independently. Experiment 3 was analyzed as a two-way Model I ANOVA. Planned comparisons were conducted using Fisher's LSD procedure. I conducted comparisons of all levels of each factor to test the following hypothesis (H_0: FSW = H*FSW = H*ACW = PMG = H*PMG = ACW; H_1: FSW = H*FSW = H*ACW < PMG = H*****PMG < ACW). That is, induction would be equal to the negative control in the heated ACW treatment, that PMG and heated PMG would not differ (but would show higher settlement than the control), and that ACW would increase settlement most strongly.

5.3 Results

5.3.1 Bioassay development

I predicted the proportion of larvae expected to settle at twelve-hour intervals by modeling settlement with the rate parameter λ. The slope of the best-fit line of predicted and observed data was less than one (0.809), indicating that the model slightly underpredicted at most time points (Fig. 5.1A). The fit of λ to the cumulative proportions of larvae settled is illustrated with data from the first block of the first experiment (Fig. 5.1). The correlation of

predicted and observed values was high (overall r = 0.930) and consistent across treatments (Fig. 5.1B). Modeling settlement as the rate parameter λ was also a more informative statistical analysis than using proportions of larvae settled. The results from ANOVAs on arcsine-transformed proportions were different depending on the time point selected, such that the significance of both main effects and interaction terms depended on the time point selected for the analysis (Table 5.2, Appendix C). Using lambda avoided these inconsistencies. The only uncontrolled source of variation in this assay was due to larval age, which was modeled as a random (block) effect.

5.3.2 Experiment 1: Screening factors that induce settlement

All three factors (adult-conditioned water, ACW; pedal mucus glasses, PMG; and potassium chloride, KCl) showed an increased settlement rate (λ) relative to the filtered seawater (FSW) control (Fig. 5.2A). The linear model of the ANOVA using lambda as a response variable contained two statistically significant treatment effects (KCl, $F_{1,62} = 19.43$, $p < 0.001$; KCl*ACW $F_{1,62} = 14.57$, $p < 0.001$; Table 5.3), with block effects through time accounting for 24% of total variation.

Planned comparisons were performed with Dunnett's procedure for testing against a common control (FSW); at a joint significance level of 0.05 the critical difference was 0.0103. As a known artificial inducer of metamorphosis in *C. fornicata* (Pechenik and Heyman, 1987), KCl had the largest effect ($\overline{KCl}$- $\overline{FSW}$ = 0.0172, Fig. 5.2B). The strength of artificial induction was most likely responsible for non-additive induction effects (KCl*ACW, Table 5.3), as complete induction by KCl allowed for no additional effect of the ACW treatment. When tested in absence of KCl, ACW was a significant inducer of settlement ($\overline{ACW}$- $\overline{FSW}$ = 0.0115, Fig.

5.2C). However, the effect of PMG was not statistically significant ($\overline{PMG}$- $\overline{FSW}$ = 0.0065, Fig. 5.2D).

5.3.3 Experiment 2: Inactivating factors that induce settlement

The main effects of both ACW ($F_{1,62}$ = 54.77, $p < 0.001$) and PMG ($F_{1,62}$ = 22.08, $p < 0.001$) were significant, in contrast to Experiment 1 where PMG was not significant. The experimental application of heat significantly altered settlement rate for ACW, but not PMG or FSW (Fig. 5.3A). Block effects accounted for 7.4% of the variation in the linear model (Table 5.4). Planned orthogonal comparisons of heated and unheated treatments were conducted with a critical value of Fisher's Least Significant Difference (LSD) of 0.0042 (α =0.05). The reduction in settlement rate upon heating ACW was statistically significant ($\overline{ACW}$– $\overline{H * ACW}$ = 0.0045, Fig. 5.3C). However, the reduction of activity was not total, and the activity of heated ACW overlapped with PMG and heated PMG. Heating FSW had no effect on settlement rates relative to unheated FSW ($\overline{FSW}$– $\overline{H * FSW}$ = 0.0000, Fig. 5.3B), indicating that the boiling treatment itself did not increase settlement rates due to changing salinity or other factors. Similarly, heating PMG treatments did not affect settlement rates relative to unheated PMG ($\overline{FSW}$– $\overline{H * FSW}$ = 0.0000, Fig. 5.3B; $\overline{PMG}$– $\overline{H * PMG}$ = 0.0000, Fig. 5.3D).

5.3.4 Experiment 3: Localizing factors that induce settlement

Consistent with the results from Experiment 2, both main effects of ACW ($F_{2,28}$= 10.66, $p < 0.001$) and PMG ($F_{2,28}$= 4.88, $p < 0.02$) were significant (Table 5.5). Similarly, the statistical significance of planned comparisons using Fisher's LSD was consistent with Experiment 2: heating ACW significantly reduced settlement rates relative to unheated ACW ($\overline{ACW}$– $\overline{H * ACW}$ = 0.005), but heating PMG did not have a statistically significant effect relative to unheated PMG ($\overline{PMG}$– $\overline{H * PMG}$ = -0.001, Table 5.6).

In addition to replicating the planned comparisons of Experiment 2, I conducted additional contrasts permitted by the design of Experiment 3 (see Materials and Methods). Fisher's LSD is appropriate for these planned comparisons because of the statistical significance of main effect treatments in the ANOVA (Carmer and Swanson, 1973; Montgomery, 2012). Settlement rates of heated ACW were not significantly different from FSW ($\overline{H * ACW}$- $\overline{FSW}$ = 0.000), although heated PMG differed significantly from FSW ($\overline{H * PMG}$- $\overline{FSW}$ = 0.004, Fig. 5.4A, Table 5.6). Finally, there was a statistically significant difference in settlement rate in the presence of both settlement inducers (ACW and PMG) but with heat applied specifically to only one factor (Fig. 5.4C, Table 5.6). Unheated ACW combined with heated PMG increased settlement over the treatment of two heated factors (difference in means = 0.008), but unheated PMG combined with heated ACW did not show increased settlement over two heated factors (difference in means = 0.003).

5.4 Discussion

The first experiment detected settlement induction by KCl and ACW and suggested the inductive effect of pedal mucus in the absence of ACW (Fig. 5.2A). Induction of settlement in *C. fornicata* by KCl and ACW has been previously reported (Pechenik and Heyman, 1987; Pechenik and Gee, 1993). By reproducing previous results with strong statistical significance, Experiment 1 validated the use of an estimated rate parameter (λ) to evaluate factors that induce settlement in *C. fornicata*. The interaction of KCl*ACW was significant in the ANOVA, indicating that there was no additional effect of ACW in treatments that contained both KCl and ACW.

Experiment 2 revealed that the settlement cue in ACW was strongly reduced by heat (Fig. 5.3A,C). In this experiment, there was also a significant effect of pedal mucus on the induction

of settlement in *C. fornicata* (Fig. 5.3A, Table 5.4), in contrast to the results from Experiment 1. The discrepancy between experiments may be due to differences among larval batches and may also reflect a weak inductive effect of pedal mucus. Settlement due to mucus was always less than that due to ACW (Fig. 5.2A, 5.3A). This effect, however, was not reduced by heat application (Fig. 5.3A,D), indicating that settlement associated with ACW and PMG was not in response to the same stimulus. In all experiments the difference in treatment means of the ACW was consistently larger than that of the PMG, indicating greater induction in response to the waterborne cue than the mucus-based stimulus. However, the effect of heated ACW was equivalent to both PMG and heated PMG (and elevated above FSW), indicating that the cue in ACW was not completely inactivated by heat (Fig. 5.3A). There are multiple possible reasons for this. First, it is possible that despite preparing ACW separate from PMG, an inducer that is present in pedal mucus was released into the ACW treatment during preparation (e.g., mucus might have been detached from the beaker and entered the water when adults were removed prior to filtering the ACW) and was not affected by heat during the heated ACW treatment. Second, the chemical cue responsible for induction in ACW may be different from that in PMG, but incompletely inactivated by my heat treatment. Distinguishing between these possibilities will require future study.

To further dissect both induction by ACW and PMG, heat was applied independently to these two factors (Experiment 3; see Materials and Methods). As in Experiments 1 and 2, ACW and PMG significantly induced settlement, heating ACW significantly reduced settlement relative to unheated ACW, and heating PMG did not affect its inductive properties (Fig. 5.4A, Table 5.6). The additional reciprocal heating experiment (Figure 5.4D; unheated ACW + heated PMG compared to heated ACW + unheated PMG) showed that only ACW was inactivated by

heat, (Fig. 5.4C, Table 5.6). In this experiment, the effect of heated ACW on settlement was not different from FSW (i.e., FSW = HACW = PMG = HPMG; Table 5.6). The discrepancies in the effect of heating between Experiments 2 and 3 (i.e., partial inactivation by heat in Experiment 2 and complete inactivation by heat in Experiment 3) may be due to differences in larval sensitivity among batches or in cue preparations, but in both cases heating ACW did reduce larval settlement.

These experiments clearly demonstrate three points. First, *C. fornicata* larvae settle in response to a waterborne cue emitted from adult conspecifics. Second, this settlement cue is biologically active and its activity can be strongly reduced by heat. Third, *C. fornicata* larvae also settle in response to pedal mucus derived from adult conspecifics, but this effect is small and not altered by heating.

Because these settlement-inducing factors act additively (Fig. 5.1, Tables 5.4, 5.5) and have different heat sensitivities, I suggest there are two independent stimulus-response systems involved in the settlement process of *C. fornicata*: a waterborne, biologically active settlement cue that may be wholly or in part proteinaceous (e.g., small peptide or glycoprotein) and acts as a homing stimulus, and a second stimulus provided by pedal mucus and associated with a final site choice of settlement.

Known waterborne, chemical inducers of settlement and metamorphosis in other marine invertebrates range from carbohydrates (the sea urchin *Holopneustes purpurescens*, Williamson et al., 2000) to metabolites (the ascidian *Halocynthia roretzi*, Tsukamoto et al., 1999), to proteins and amino acids (the barnacle *Semibalanus balanoides*, Crips and Meadows, 1963; the oyster *Crassostrea virginica*, Zimmer-Faust and Tamburri, 1994). Experimental work on other

gastropods has also implicated a range of substances including carbohydrates (*Alderia modesta*, Krug and Manzi, 1999), metabolites (*Phestilla sibogae*, Hadfield and Pennington, 1990), volatile halogenated organic compounds (*Haliotis discus hannai*, Kang et al., 2004), and peptides (*Adalaria proxima*, Lambert et al., 1997). The bioassay approach developed here now allows me to further characterize the waterborne chemical inducer of settlement in *C. fornicata* and provides a standardized protocol and statistical analysis for the estimation of settlement rate under various chemical treatments. My experimental results may point toward a peptide or a protein that is partially inactivated by heat, but the diversity of known waterborne stimulants (reviewed in Pawlik, 1992; Hadfield and Paul, 2001) suggest that it is too early to conclude more than that two distinct stimulants affect metamorphosis in *C. fornicata*.

Previous investigators have suggested that contact cues may stimulate metamorphosis when competent larvae come into contact with adult conspecifics. Settlement in the barnacle *Semibalanus balanoides* increases after larvae touch basal plates of conspecifics (Knight-Jones, 1953; Crisp and Meadows, 1963). This stimulant is stable when subjected to heat (Knight-Jones, 1953). Other proteins have also been shown to induce settlement upon contact with both larvae and adults (Matsumurad et al., 1998).

Although I cannot rule out that the effect of pedal mucus is due to chemical cues (waterborne or contact) that are not affected by heat, the effect of gastropod pedal mucus in other systems is due to physical cues. Pedal mucus of *Patella vulgata* increases settlement in *Semibalanus balanoides* cyprid larvae, and the effect is due to the physical structure of the mucus (Holmes, 2002). However, until I conduct further experiments, I cannot eliminate the possibility that the effect of pedal mucus on larval settlement in these experiments was due to chemical components of the mucus that were not destroyed by the heat treatment (e.g.,

carbohydrates, which along with proteins make up a large component of gastropod mucus; Davies et al., 1990; Davies and Hawkins, 1998).

Likewise, I cannot eliminate the possibility that the waterborne settlement cue comes from adult-associated bacteria or other small biofouling organisms, but such organisms are likely to also be present in the PMG treatment and therefore ACW and PMG would not be expected to show different responses to heat. Taris et al. (2010) found that settlement in *C. fornicata* is induced by compounds from coralline algae that foul adult *C. fornicata* shells within its invasive European range. This was not a factor in the current experiment because these algae were not associated with *C. fornicata* at my collection site.

Settlement rates in this study were variable both among experiments and among blocks within experiments. Previous work on *C. fornicata* has found similar variability across experiments, and concluded that these differences could be explained by differences in larval growth rate due to food availability or temperature (Pechenik and Lima, 1984; Pechenik et al., 1996). However, all larvae were reared on the same diet and at the same temperature for this study. Variation in larval growth rate of *C. fornicata* is also influenced by sire (Le Cam et al., 2009) and maternal effects (Hilbish et al., 1999). By randomly mixing larvae from multiple broods, I spread unknown genetic variation evenly among all treatments and blocks within each experiment. In addition to variation among larvae, there was also variation among preparations of both the ACW and the PMG. Different adult animals were used for each preparation (two per block per experiment). I optimized and then standardized the abiotic and biotic parameters of the experiment to minimize differences among experiments associated with preparation of adult-conditioned water (Table 5.1, Appendix B). However, in the absence of a clearly identified chemical that induces settlement, I remain unable to control the exact concentration of cue

delivered to the larvae. Due to variation among batches of larvae (e.g., due to maternal effects, Hilbish et al., 1999), I also am unable to compare rates across experiments.

5.5 Conclusion

I developed a time-course bioassay to estimate settlement rates in *C. fornicata*. I found that conspecific adults emit a waterborne settlement cue that was experimentally reduced with heat. In addition to the waterborne cue, I discovered larval settlement was induced by a second, distinct stimulus present in adult pedal mucus. I propose that settlement in *C. fornicata* is a complex system involving at least two stimuli, one of which acts as a homing signal (waterborne), and another stimulus which may aid in small-scale site selection (mucus-based). My experimental design allowed for the detection of weak effects (e.g., heat inactivation and pedal mucus). Further characterization of this complex settlement system by purification of cue(s) as well as testing compound specific methods of cue inactivation are needed to further elucidate the nature of chemical cues that induce settlement in *C. fornicata*.

Table 5.1 Optimizations. Optimized values of abiotic and biotic parameters in the bioassay.

Variable	Range Tested	Optimum Value/Range
pH	7.9 – 10.2	8.0 – 8.3
Salinity	25 – 40	27-30
Trial length	8 h – 212 h	48 h
Time test factors were boiled	0 min – 10 min	2 min
Adult mass used for ACW (per liter filtered seawater)	4 g – 800 g	100 g
ACW preparation time	1 h – 24 h	12 h
Larvae per replicate 20 ml glass	5 – 20	10
Sampling time intervals	4 h – 24 h	12 h

Table 5.2 Changes in proportions of larval settlement demonstrating time-sensitivity of results calculated with proportions. Heat map indicating the significance of all main effects and interactions tested in Experiment 2, analyzed using ANOVAs on arcsine square root transformed proportions of larvae that settled; calculations were done at each time step. Dark grey: $p < 0.001$; light grey: $0.001 < p < 0.05$; white: $p > 0.05$.

Source of Variation	12 hours	24 hours	36 hours	48 hours
Adult Conditioned Water (ACW)	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**
Pedal Mucus Glass (PMG)	0.219	**0.002**	**< 0.001**	**< 0.001**
Heat Treatment (H)	0.072	**0.043**	0.082	0.082
ACW*PMG	0.265	0.562	0.946	0.946
ACW*H	0.755	0.169	0.070	0.070
PMG*H	0.830	0.449	0.527	0.527
ACW*PMG*H	**0.035**	**0.029**	0.575	0.575

Table 5.3 Experiment 1. Analysis of variance for Experiment 1, screening factors that induce settlement. KCL, ACW, and PMG, were applied in a factorial design. Analysis was conducted using λ as the response variable. Significant effects at p = 0.05 are highlighted in bold.

Source of Variation	df	SS	MS	F	p
Block Effect	2	1.74E-03	8.71E-04	-----	-----
Adult Conditioned Water (ACW)	1	6.71E-05	6.71E-05	1.29	0.26
Pedal Mucus Glass (PMG)	1	1.26E-05	1.26E-05	0.24	0.63
Potassium Chloride (KCl)	1	1.01E-03	1.01E-03	19.43	< **0.001**
PMG*ACW	1	1.07E-04	1.07E-04	2.04	0.16
PMG*KCl	1	1.26E-04	1.26E-04	2.43	0.12
ACW*KCl	1	7.59E-04	7.59E-04	14.57	< **0.001**
PMG*ACW*KCl	1	6.51E-06	6.51E-06	0.12	0.72
Error	62	3.23E-03	5.21E-05	0.00	
Total	71	7.06E-03			

Table 5.4 Experiment 2. Analysis of variance for Experiment 2, inactivating factors that induce settlement. ACW, PMG, and heat were applied in a factorial design. Analysis was conducted using λ as the response variable. Significant effects at p = 0.05 are highlighted in bold.

Source of Variation	df	SS	MS	F	p
Block Effect	2	2.28E-04	1.14E-04	-----	-----
Adult Conditioned Water (ACW)	1	1.08E-03	1.08E-03	54.77	< **0.001**
Pedal Mucus Glass (PMG)	1	4.37E-04	4.37E-04	22.08	< **0.001**
Heat Treatment (H)	1	4.71E-05	4.71E-05	2.38	0.13
ACW*PMG	1	1.82E-05	1.82E-05	0.92	0.34
ACW*H	1	4.67E-05	4.67E-05	2.36	0.13
PMG*H	1	7.03E-06	7.03E-06	0.36	0.55
ACW*PMG*H	1	7.48E-06	7.48E-06	0.38	0.54
Error	62	1.23E-03	1.98E-05		
Total	71	3.10E-03			

Table 5.5 Experiment 3. Analysis of variance for Experiment 3, localizing factors that induce settlement. ACW and PMG were present, absent, or heated in a factorial design. Analysis was conducted using λ as the response variable. Significant effects at p = 0.05 are highlighted in bold.

Source of Variation	df	SS	MS	F	p
Adult Conditioned Water (ACW)	2	1.44E-04	7.21E-05	10.66	< **0.001**
Pedal Mucus Glass (PMG)	2	6.61E-05	3.30E-05	4.88	**0.02**
PMG*ACW	4	2.64E-05	6.61E-06	0.98	0.44
Error	18	1.22E-04	6.76E-06		
Total	26	3.58E-04			

Table 5.6 Experiment 3 comparisons. Pairwise comparisons for Experiment 3, localizing factors that induce settlement, using Fisher's least square differences (LSD). Analysis was conducted using λ as the response variable. Significant differences at p = 0.05 are highlighted in bold.

Comparison	Differences in treatment means
ACW – FSW	**0.0052**
Heated ACW – FSW	0.0000
ACW – Heated ACW	**0.0052**
PMG – FSW	**0.0030**
Heated PMG - FSW	**0.0038**
PMG – Heated PMG	-0.0009
Heated ACW, PMG – ACW, Heated PMG	**-0.0052**
Heated ACW, PMG – Heated ACW, Heated PMG	0.0025
ACW, Heated PMG – Heated ACW, Heated PMG	**0.0077**

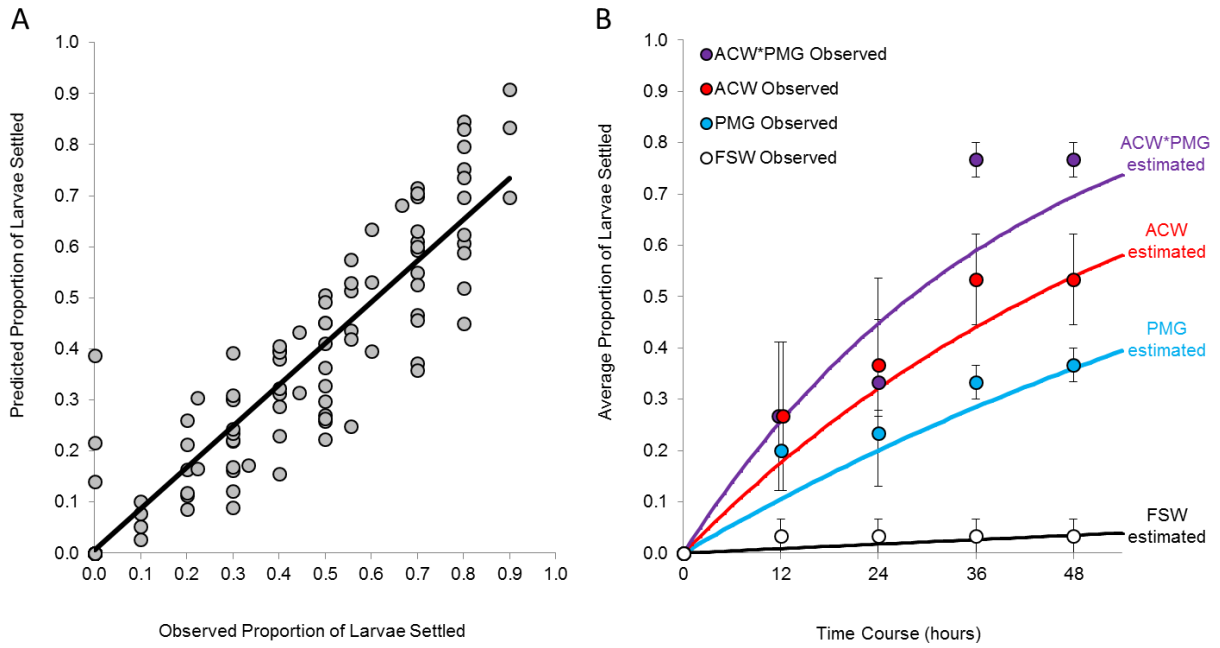

Fig. 5.1 Predicted and observed data. A) Plot of observed versus predicted proportions of larvae settled in each glass at each time point. Values calculated based on the first block of Experiment 1, screening factors that induce settlement. Shown is the best-fit line Ordinary Least Squares regression for the data (y = 0.809x + 0.006; r = 0.930). B) Observed larval settlement (filled symbols) and the values predicted by λ (colored lines) at each time point. Values calculated based on the first block of experiment 1. Error bars represent standard error. There are close correlations between observed and expected values for all treatments: ACW*PMG (r = 0.957), ACW (r = 0.977), PMG (r = 0.965), and FSW (r = 0.710).

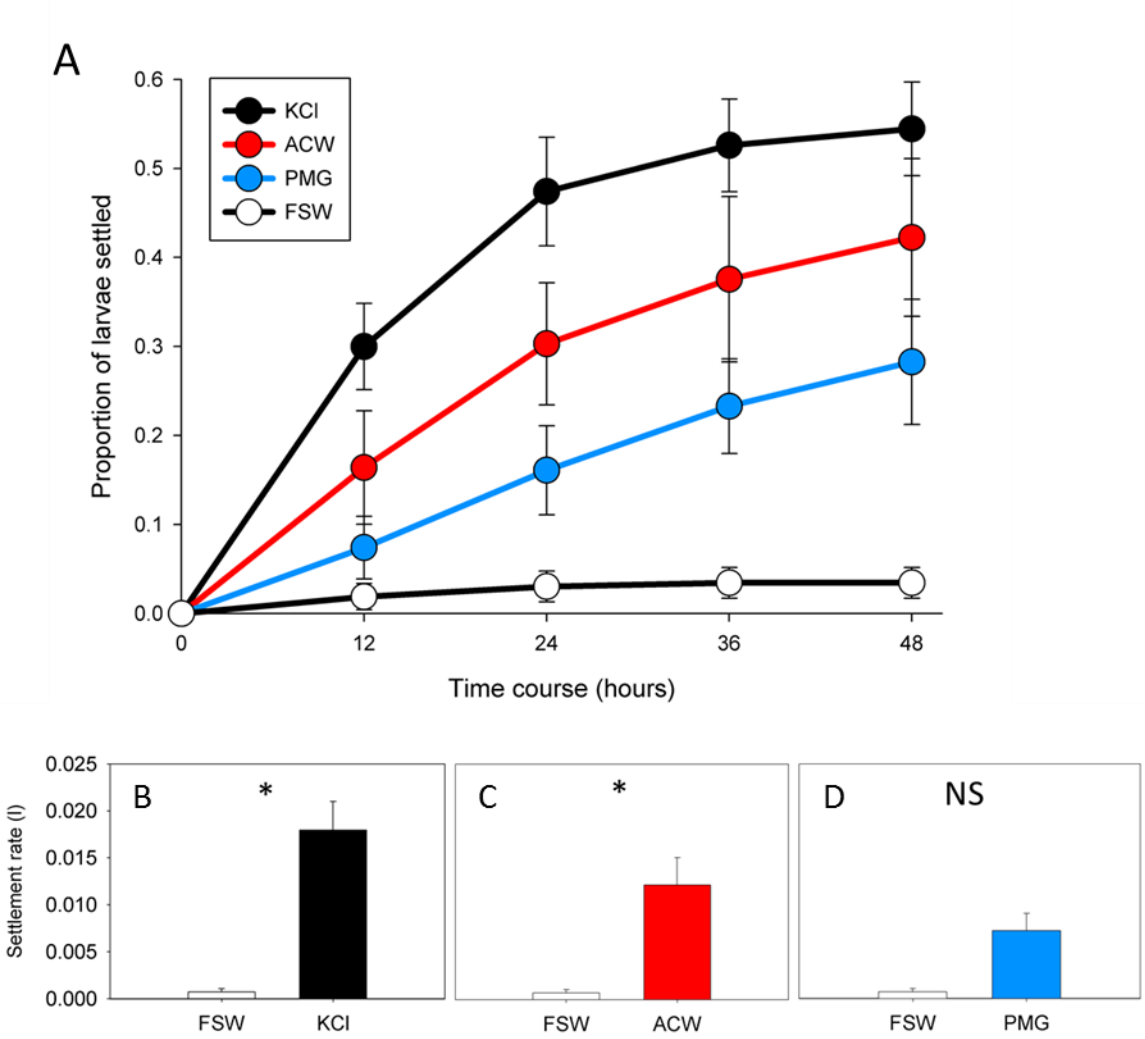

Fig. 5.2 Experiment 1. A) Screening factors that induce settlement, Experiment 1. Points represent running averages across all three experimental blocks; error bars represent 1 SE. B) Pairwise comparison of settlement rates for FSW (white) and KCl (black) treatments. N = 9 glasses. Error bars represent 1 SE. Significance at $p < 0.05$ as determined by Dunnett's test denoted by an *. C) Pairwise comparison of settlement rates for FSW (white) and ACW (red) treatments. N = 9 glasses. Error bars represent 1 SE. Significance at $p < 0.05$ as determined by Dunnett's test denoted by an *. D) Pairwise comparison of settlement rates for FSW (white) and PMG (blue) treatments. N = 9 glasses. Error bars represent 1 SE. Significance at $p < 0.05$ as determined by Dunnett's test denoted by an *.

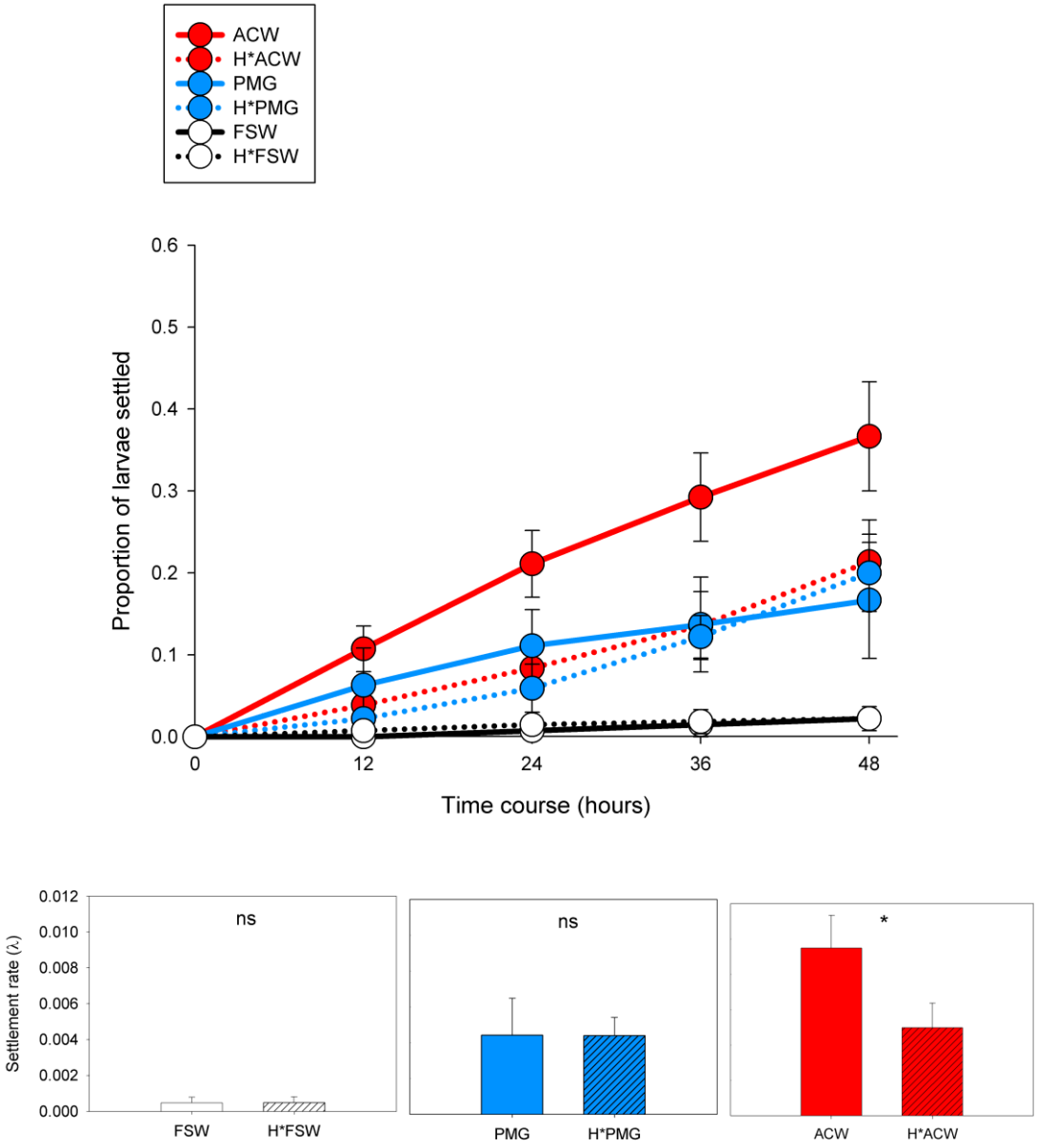

Fig. 5.3 Experiment 2. A) Inactivating factors that induce settlement, Experiment 2. Solid lines represent unheated factors; dashed lines represent heated factors. Points represent running averages across all three experimental blocks; error bars represent 1 SE. B) Pairwise comparison of settlement rates for unheated FSW (white) and heated FSW (hashed) treatments. N = 9 glasses. Error bars represent 1 SE. Significance at $p < 0.05$ as determined by Fisher's LSD denoted by an *. C) Pairwise comparison of settlement rates for unheated ACW (red) and heated ACW (hashed) treatments. N = 9 glasses. Error bars represent 1 SE. Significance at $p < 0.05$ as determined by Fisher's LSD denoted by an *. D) Pairwise comparison of settlement rates for unheated PMG (blue) and heated PMG (hashed) treatments. N = 9 glasses. Error bars represent 1 SE. Significance at $p < 0.05$ as determined by Fisher's LSD denoted by *.

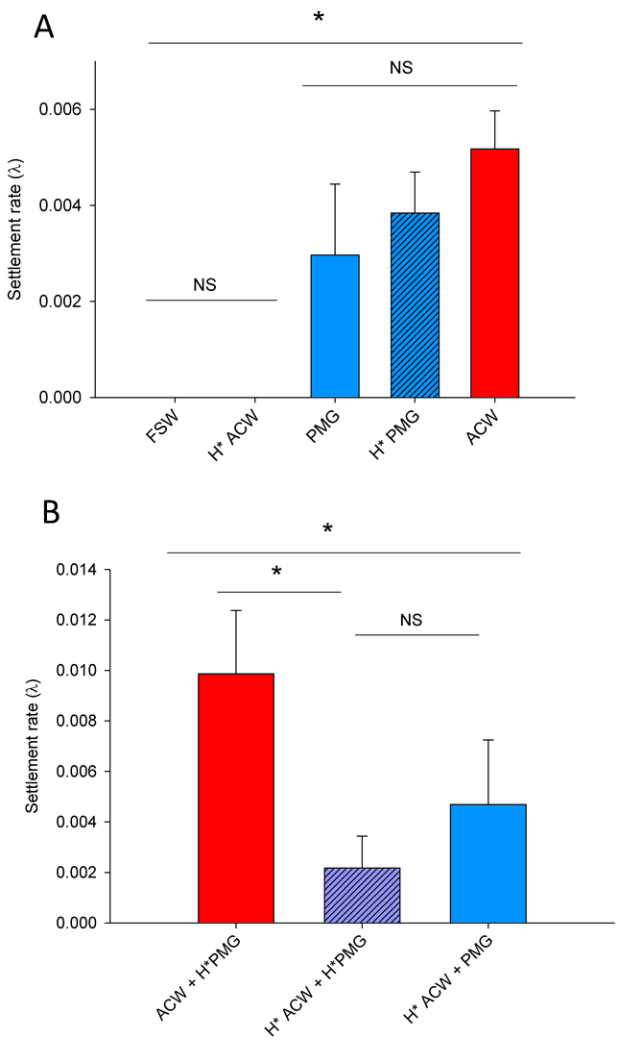

Fig. 5.4 Experiment 3. A) Localizing factors that induce settlement, Experiment 3. Settlement rates of heated and unheated factors. Error bars represent 1 SE. B) Post-hoc comparisons of ACW +, PMG H (unheated ACW, heated PMG), ACW H, PMG H (heated ACW, heated PMG), and ACW H, PMG + (heated ACW, unheated PMG). N = 3 glasses. Error bars represent 1 SE. Significance at $p < 0.05$ as determined by least significant differences denoted by an *.

Chapter 6: Conclusion

The ecology and evolution of species' range limits can be approached using nearly any subfield of ecology or evolution (e.g., physiology, population genetics, community ecology, etc.). Using a variety of approaches to understand extrinsic and intrinsic limits on species range will give a more thorough understanding of the factors that set these limits. Identifying the mechanisms determining species' range limits is of current concern given that many species are expanding or contracting their ranges in response to anthropogenic climate change (Parmesan and Yohe, 2003; Angert et al., 2011). We currently have few examples clearly identifying a proximate mechanism determining a range limit (Cahill et al., 2014), perhaps in part because identifying intrinsic and extrinsic drivers of range limits and eliminating other hypotheses is a complex problem.

To that end, I analyzed multiple population-level processes in two species of *Crepidula* with different developmental types. *Crepidula fornicata* has a planktotrophic larval form that lives in the water column for 2-4 weeks, giving this species a relatively high dispersal potential (Collin, 2003). In contrast, *C. convexa* has direct-developing larvae that crawl away from their mother as juvenile snails and thus have much lower dispersal potential (Collin, 2003). These two species are sympatric over the entire range of *C. convexa* (Collin, 2001), and *C. fornicata* may be moving northward with warmer water temperatures (Rawlings, 2011).

In chapter 2, I used a microsatellite analysis to examine population genetic structure in native populations of *C. convexa* and compared it to previous work using microsatellites and AFLPs in *C. fornicata* (Riquet et al., 2013). I found a strong pattern of isolation-by-distance using just five loci, which is expected given the low dispersal potential of *C. convexa*. This

contrasts with very low structure and no isolation-by-distance in *C. fornicata* (Riquet et al., 2013).

These analyses are concordant with the results in Chapter 3, where I used next-generation sequencing to generate a large (~12,000) library of SNPs of both *Crepidula* species to use for population genetic analyses. I did not find support for the hypothesis that marginal populations show reduced genetic variation relative to central ones. In *C. convexa*, the marginal populations had the same percentage of heterozygous SNPs per individual as two of the central populations, and in *C. fornicata*, the northernmost population showed increased heterozygosity relative to all other populations. Analyses of genetic structure were largely concurrent with those in chapter 2: *C. convexa* populations were differentiated, while *C. fornicata* populations in the center of the range were not. However, both marginal (Canadian) populations of *C. fornicata* were differentiated from the central populations and each other. Particularly of note is the strong difference between the population in Newfoundland, at the northern limit of this species, and a nearby population in Nova Scotia. It is likely that the circulation patterns in this area have caused divergence between these populations, although other explanations (such as selection) cannot be ruled out given the dataset.

This analysis showed that even in species with relatively high dispersal potential, marginal populations can be strongly diverged from more central ones. It also demonstrated the utility of using next-generation sequencing technology to answer genomic questions on non-model organisms, even ones with very large genomes (~6 GB). However, it identifies some interesting gaps in our knowledge of *C. fornicata*, a well-studied species and rising model system (Henry et al., 2010). First, no population genetic studies have examined population structure in between Massachusetts and Nova Scotia to understand how these populations may be connected

to ones further south. Second, the Newfoundland population belongs to a distinct genetic cluster from all previously studied populations. Given the circulation patterns in the Gulf of St. Lawrence, the many *C. fornicata* populations that exist there (Rawlings et al., 2011) may be different from populations on the Atlantic coast of North America. Third, no previous ecological or physiological studies have investigated the climatic niche, thermal tolerance, larval behavior, etc. of the populations in the Gulf of St. Lawrence. These populations are living in marginal thermal conditions and are the ones that have been expanding north (Rawlings, 2011), making them of particular interest to the study of range margins in this species.

Chapters 4 and 5 focused solely on *C. fornicata.* In chapter 4, I investigated the settlement behavior of *C. fornicata* larvae in relation to the density of adult conspecifics. Using a field experiment, I found that larval settlement increased with an increasing density of adults. Although relatively low amounts of variation in larval settlement were explained by adult density, this effect of preferential settlement on conspecifics may slow the expansion of a range margin, akin to an Allee effect (Keitt et al., 2001; Kubisch et al., 2014).

Based on the fact that larval *C. fornicata* are known to settle in response to waterborne cues from adult conspecifics in both lab (e.g., Pechenik and Gee, 1993) and field (Bohn et al., 2013a,b; Chapter 4, this dissertation) settings, I used a series of lab experiments to try to identify the responsible mechanism. I developed a bioassay using settlement rates and found that both a waterborne cue and pedal mucus from conspecific adults induced settlement in *C. fornicata*. Although the waterborne cue had been seen before, the inductive capacity of conspecific pedal mucus had not. These cues were different, as demonstrated when I partially inactivated the waterborne cue, but not the pedal mucus cue, using heat.

Adult conspecifics can therefore impact the distribution of larval settlement at different scales. At a larger scale (meters to tens of meters), the waterborne cue can serve as an attractant (chapters 4 and 5). Pedal mucus acts at a smaller scale. In my field experiment, I saw both scales of attraction at work: larvae settled more in treatments with more conspecifics, and within those treatments, they preferentially settled on conspecifics rather than control substrates. Although much further work is required to connect the lab and field experiments (e.g., is pedal mucus responsible for the preferential settlement on snails at small scales in the field, or is it the waterborne cue, or something unidentified?), these experiments test potential mechanisms to explain the distribution of *C. fornicata*, including at range limits. However, given the results of Chapter 3, these experiments should be repeated at the range margin.

To formulate a precise case for why the northern range limits of *C. fornicata* and *C. convexa* exist where they do, or to predict where these limits will shift under climate change scenarios, requires much more work than was possible in this dissertation. Such a body of work would require more information about population dynamics, temperature tolerances, biotic interactions, and many other factors. The work I have presented here includes evidence of ecological (settlement) and evolutionary (population genetic) processes that can affect range limits and population dynamics in these marginal populations. I have also identified holes in our knowledge of how marginal populations are connected and differentiated from more central populations. Further work in these species should work towards a more complete understanding of the factors limiting both northern and southern ranges in these common mollusc species.

References

Addison JA, Hart MW (2005) Spawning, copulation, and inbreeding coefficients in marine invertebrates. Biol Letters 1:450-453.

Allendorf FW, Lundquist LL (2003) Introduction: population biology, evolution, and control of invasive species. Conserv Biol 17: 24-30.

Ament AS (1978) Geographic variation in relation to life history in three species of the marine gastropod genus *Crepidula*. Dissertation, University of Pennsylvania.

Ament AS (1979) Geographic variation in relation to life history in three species of the marine gastropod genus *Crepidula*: growth rates of newly hatched larvae and juveniles. In: Stancyk SE (ed) *Reproductive Ecology of Marine Invertebrates*. University of South Carolina Press, Columbia, pp 61-76.

Anderson BJ, Akçakaya HR, Araújo MB, Fordham DA, Martinez-Meyer E, Thuiller W, Brook BW (2009) Dynamics of range margins for metapopulations under climate change. P Roy Soc B 276: 1415–1420.

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Available at http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ Accessed 5 August 2014.

Angert AA, Seth SN, Paul JR (2011a) Incorporating population-level variation in thermal performance into predictions of geographic range shifts. Intgr Comp Biol 51: 733-750.

Angert, AA, Crozier LG, Rissler LJ, Gilman SE, Tewksbury JJ, Chunco AJ (2011b) Do species' traits predict recent shifts at expanding range edges? Ecol Lett 14: 677-689.

Arndt A, Smith MJ (2002) Genetic diversity and population structure in two species of sea cucumber: differing patterns according to mode of development. Mol Ecol 7: 1053-1064.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 3: e3376.

Barrett SCH, Kohn JR (1991) Genetic and evolutionary consequences of small population size in plants: implications for conservation. In: Falk DA, Holsinger KE (eds) *Genetics and Conservation of Rare Plants*. Oxford University Press, Oxford, pp 3-30.

Bennett AF, Lenski RE, Mittler JE (1992) Evolutionary adaptation to temperature. I. Fitness responses of *Escherichia coli* to changes in its thermal environment. Evolution 46: 16-30.

Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why genome scans may fail to map local adaptation genes. Mol Ecol 20: 2044-2072.

Blanchard M (1997) Spread of the slipper limpet *Crepidula fornicata* (L. 1758) in Europe: current state and consequences. Sci Mar 61(S2): 109-118.

Bohn K, Richardson CA, Jenkins SR (2012) The invasive gastropod *Crepidula fornicata*: reproduction and recruitment in the intertidal at its northernmost range in Wales, UK, and implications for its secondary spread. Mar Biol 159: 2091-2013.

Bohn K, Richardson CA, Jenkins SR (2013a) The importance of larval supply, larval habitat selection and post-settlement mortality in determining intertidal adult abundance of the invasive gastropod *Crepidula fornicata*. J Exp Mar Biol Ecol 440: 132-140.

Bohn K, Richardson CA, Jenkins SR (2013b) Larval microhabitat associations of the non-native gastropod *Crepidula fornicata* and effects on recruitment success in the intertidal zone. J Exp Mar Biol Ecol 448: 289-297.

Bourne EC, Bodeci G, Travis JMJ, Pakeman RJ, Brooker RW, Schiffers K (2014) Between migration load and evolutionary rescue: dispersal, adaptation and the response of spatially structured populations to environmental change. P Roy Soc B 281: 20132795.

Bradbury IR, Laurel B, Snelgrove PVR, Bentzen P, Campana SE (2008) Global patterns in marine dispersal estimates: the influence of geography, taxonomic category and life history. P Roy Soc B 275: 1803-1809.

Bridle JR, Vines TH (2007) Limits to evolution at range margins: when and why does adaptation fail? Trends Ecol Evol 22: 140-147.

Bridle JR, Buckley J, Bodsworth EJ, Thomas CD (2014) Evolution on the move: specialization on widespread resources associated with rapid range expansion in response to climate change. P Roy Soc B 281: 20131800.

Bucklin A, Sundt RC, Dahle G (1996) The population genetics of *Calanus finmarchicus* in the North Atlantic. Ophelia 44: 29-45.

Cahill AE, Aiello-Lammens ME, Fisher-Reid MC, Hua X, Karanewsky CJ, Ryu HY, Sbeglia GC, Spagnolo F, Waldron JB, Wiens JJ (2014) Causes of warm-edge range limits: systematic review, proximate factors, and implications for climate change. J Biogeogr 41: 429-442.

Cahill AE, Aiello-Lammens ME, Fisher-Reid MC, Hua X, Karanewsky CJ, Ryu HY, Sbeglia GC, Spagnolo F, Waldron JB, Warsi O, Wiens JJ (2013) How does climate change cause extinction? P Roy Soc B 280: 20121890.

Carlton JT (1992) Introduced marine and estuarine mollusks of North America: an end-of-the-20th-century perspective. J Shellfish Res 11: 489-505.

Carmer SG, Swanson MR (1973) An evaluation of ten pairwise multiple comparison procedures by Monte Carlo methods. J Am Statist Assoc 68: 66-74.

Case RJ, Taper ML (2000) Interspecific competition, environmental gradients, gene flow, and the coevolution of species' borders. Am Nat 155: 583-605.

Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. Mol Biol Evol 24: 621-631.

Chen IC, Hill JK, Ohlemuller R, Roy DB, Thomas CD (2011) Rapid range shifts of species associated with high levels of climate warming. Science 333: 1024–1026.

Collin R (1995) Sex, size, and position: a test of models predicting size at sex change in the protandrous gastropod *Crepidula fornicata*. Am Nat 146: 815-831.

Collin R (2001) The effects of mode of development on phylogeography and population structure of North Atlantic *Crepidula* (Gastropoda: Calyptraeidae). Mol Ecol 10: 2249-2262.

Collin R (2002) Another last word on *Crepidula convexa* with a description of *C. ustulatulina* n. sp. (Gastropoda: Calyptraeidae) from the Gulf of Mexico and southern Florida. B Mar Sci 70: 177-184.

Collin R (2003) Worldwide patterns in mode of development in calyptraeid gastropods. Mar Ecol-Prog Ser 247: 103-122.

Collin R, Wonham MJ, Barr KR (2006) *Crepidula convexa* Say, 1822 (Caenogastropoda: Calyptraeidae) in Washington State, U.S.A. Am Malacol Bull 21: 113-116.

Connell JH (1985) The consequences of variation in initial settlement vs. post-settlement mortality in rocky intertidal communities. J Exp Mar Biol Ecol 93: 11-45.

Corander J, Majander KK, Cheng L, Merilä J (2013) High degree of cryptic population differentiation in the Baltic Sea herring *Clupea harengus*. Mol Ecol 22: 2931-2940.

Cowen RK, Sponaugle S (2009) Larval dispersal and marine population connectivity. Ann Rev Mar Sci 1: 443-466.

Cowen RK, Lwiza KMM, Sponaugle S, Paris CB, Olson DB (2002) Connectivity of marine populations: open or closed? Science 287: 857-859.

Crimmins SM, Dobrowski SZ, Greenberg JA, Abatzoglou JT, Mynsberge AR (2011) Changes in climatic water balance drive downhill shifts in plant species' optimum elevations. Science 331: 324-327.

Crisp DJ, Meadows PS (1963) Adsorbed layers: the stimulus to settlement in barnacles. P Roy Soc B 158: 364-387.

Daguin-Thiebaut C, Le Cam S, Viard F (2009) Isolation of 11 microsatellite markers in *Crepidula convexa* (Gastropoda, Calyptraeideae) for parentage analyses. Mol Ecol Resour 9: 917-920.

Davies MS, Hawkins J (1998) Mucus from marine molluscs. Adv Mar Biol 34: 1-71.

Davies MS, Jones HD, Hawkins SJ (1990) Seasonal variation in the composition of pedal mucus from *Patella vulgata* L. J Exp Mar Biol Ecol 144: 101-112.

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc B 39: 1–38.

Dillon RT, Manzi JJ (1992) Population genetics of the hard clam, *Mercenaria mercenaria*, at the northern limit of its range. Can J Fish Aquat Sci 49: 2574-2578.

Donaldson S (1974) Larval settlement of a symbiotic hydroid: specificity and nematocyst responses in planulae of *Proboscidactyla flavicirrata*. Biol Bull 147: 573-585.

Dupont L, Jollivet D, Viard F (2003) High genetic diversity and ephemeral drift effects in a successful introduced mollusc (*Crepidula fornicata*: Gastropoda). Mar Ecol-Prog Ser 253: 183-195.

Dupont L, Richard J, Paulet Y-M, Thouzeau G, Viard F (2006) Gregariousness and protandry promote reproductive insurance in the invasive gastropod *Crepidula fornicata*: evidence from assignment of larval paternity. Mol Ecol 15: 3009-3021.

Ehrhold A, Blanchard M, Auffret JP, Garlan T (1998) The role of *Crepidula* proliferation in the modification of the sedimentary tidal environment in Mont-Saint-Michel Bay (The Channel, France). CR Acad Sci II A 327 : 583-588.

Ekert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. Mol Ecol 17: 1170-1188.

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* (L.) Skeels) endemic to Morocco. Theor Appl Genet 92: 832–839.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotype-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6: e19379.

Etterson JR, Shaw RG (2001) Constraint to adaptive evolution in response to global warming. Science 294: 151-154.

Ewing B, Green P (1998) Base-calling of automated sequencer traces using *Phred*. II. Error Probabilities. Genome Res 8: 186-194.

Fairbairn DJ (1981) Biochemical genetic analysis of population differentiation in Greenland halibut (*Reinhardtius hippoglossoides*) from the northwest Atlantic, Gulf of St. Lawrence, and Bering Sea. Can J Fish Aquat Sci 38: 669 – 677.

Fordham DA, Brook BW, Moritz C, Nogués-Bravo D (2014) Better forecasts of range dynamics using genetic data. Trends Ecol Evol 29: 436-443.

Gabbott PA, Larman VN (1987) The chemical basis of gregariousness in cirripedes: a review. In: Southward AJ (ed) *Barnacle Biology*. A.A. Balkema, Rotterdam, pp. 377–388.

Gaston KJ (2003) *The structure and dynamics of geographic ranges*. Oxford University Press, Oxford.

Gaston KJ (2009) Geographic range limits: achieving synthesis. P Roy Soc B 276: 1395–1406.

Goff SA, Vaughn M, McKay S, Lyons E, Stapleton AE, Gessler D, et al. (2011) The iPlant collaborative: cyberinfrastructure for plant biology. Front Plant Sci 2: 1- 16.

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

Goudet J. (2005) Hierfstat, a package for R to compute and test hierarchical F-statistics. Mol Ecol Notes 5: 184-186.

Gross SJ, Price TD (2000) Determinants of the northern and southern range limits of a warbler. J Biogeogr 27: 869-878.

Hadfield MG, Pennington JT (1990) Nature of the metamorphic signal and its internal transduction in larvae of the nudibranch *Phestilla sibogae*. Bull Mar Sci 46: 455-464.

Hadfield MG, Paul VJ (2001) Natural chemical cues for settlement and metamorphosis of marine invertebrate larvae. In: McClintock JB, Baker BJ (eds) *Marine Chemical Ecology*. CRC Press, Boca Raton, pp. 431-461.

Hamlin JAP, Arnold ML (2014) Determining population structure and hybridization for two iris species. Ecol Evol 4: 743-755.

Hampe A, Jump AS (2011) Climate relics: past, present, future. Annu Rev Ecol Evol S 42: 313-333.

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. Ecol Lett 8: 461-467.

Harsch MA, Hille Ris Lambers J (In press) Species distributions shift downward across western North America. Glob Change Biol. doi: 10.1111/gcb.12697.

Hedrick PW (1999) Highly variable loci and their interpretation in evolution and conservation. Evolution 53: 313-318.

Hellberg ME (1996) Dependence of gene flow on geographic distance in two solitary corals with different larval dispersal capabilities. Evolution 50: 1167-1175.

Hendler G, Franz DR (1971) Population dynamics and life history of *Crepidula convexa* Say (Gastropoda: Prosobranchia) in Delaware Bay. Biol Bull 141: 514-526.

Henry JJ, Collin R, Perry KJ (2010) The slipper snail, *Crepidula*: an emerging lophotrochozoan model system. Biol Bull 218: 211-229.

Herborg L-M, Weetman D, van Oosterhout C, Hänfling B (2007) Genetic population structure and contemporary dispersal pattern of a recent European invader, the Chinese mitten crab, *Eriocheir sinensis*. Mol Ecol 16: 231-242.

Hickling R, Roy DB, Hill JK, Thomas CD (2005) A northward shift of range margins in the British Odonata. Glob Change Biol 11: 502-506.

Hilbish TJ, Sasada K, Eyster LS, Pechenik JA (1999) Relationship between rates of swimming and growth in veliger larvae: genetic variance and covariance. J Exp Mar Biol Ecol 239: 183-193.

Hoagland KE (1977) Systematic review of fossil and recent *Crepidula*. Malacologia 16: 363 – 420.

Hoagland KE (1984) Use of molecular genetics to distinguish species of the gastropod genus *Crepidula* (Prosobranchia: Calyptraeidae). Malacologia 25: 607–628.

Hoagland KE (1985) Genetic relationships between one British and several North American populations of *Crepidula fornicata* based on allozyme studies (Gastropoda: Calyptraeidae). J Mollus Stud 51: 177-182.

Hoch JM, Cahill AE (2012) Variation in size at sex-change among natural populations of the protandrous hermaphrodite, *Crepidula fornicata* (Gastropoda, Calyptraeidae). Mar Biol 159: 897-905.

Hoffmann AA, Sgrò CM (2011) Climate change and evolutionary adaptation. Nature 470: 479-485.

Holmes SP (2002) The effect of pedal mucus on barnacle cyprid settlement: a source for indirect interactions in the rocky intertidal? J Mar Biol Assoc UK 82: 117-129.

Holt RD, Gomulkiewicz R (1997) How does immigration influence local adaptation? A reexamination of a familiar paradigm. Am Nat 149: 563.

Jensen RA, Morse DE (1984) Intraspecific facilitation of larval recruitment: gregarious settlement of the polychaete *Phragmatopoma californica* (Fewkes). J Exp Mar Biol Ecol 83: 107-126.

Johannesson K (1988) The paradox of Rockall: why is a brooding gastropod (*Littorina saxatilis*) more widespread than one having a planktonic larval dispersal stage (*L. littorea*)? Mar Biol 99: 507-513.

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

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet 11: 94.

Johnson LE, Strathmann RR (1989) Settling barnacle larvae avoid substrata previously occupied by a mobile predator. J Exp Mar Biol Ecol 128: 87-103.

Jost L (2008) G_{st} and its relatives do not measure differentiation. Mol Ecol 17: 4015-4026.

Kang KH, Kim BH, Kim JM (2004) Induction of larval settlement and metamorphosis of the abalone, *Haliotis discus hannai* larvae using bromomethane and potassium chloride. Aquaculture 230: 249-259.

Kawecki TJ (2008) Adaptation to marginal habitats. Annu Rev Ecol Evol S 39: 321-342.

Keitt TH, Lewis MA, Holt RD (2001) Allee effects, invasion pinning, and species' borders. Am Nat 157: 203-216.

Kelly RP, Palumbi SR (2010) Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. PLoS ONE 5: e8594.

Kinlan BP, Gaines SD (2003) Propagule dispersal in marine and terrestrial environments: a community perspective. Ecology 84: 2007-2020.

Kirkpatrick M, Barton NH (1997) Evolution of a species' range. Am Nat 150: 1-23.

Knight-Jones EW (1953) Decreased discrimination during settling after a prolonged planktonic life in larvae of *Spirorbis borealis* (Serpulidae). J Mar Biol Assoc UK 32: 337-345.

Kremer A, Ronce O, Robledo-Arnuncio JJ, Guillaume F, Bohrer G, Nathan R, Bridle JR, Gomulkiewicz R, Klein EK, Ritland K, Kuparinen A, Gerber S, Schueler S (2012) Long-distance gene flow and adaptation of forest trees to rapid climate change. Ecol Lett 15: 378-392.

Krug PJ, Manzi AE (1999) Waterborne and surface-associated carbohydrates as settlement cues for larvae of the specialist marine herbivore *Alderia modesta*. Biol Bull 197: 94-103.

Kubisch A, Holt RD, Poethke H-J, Fronhofer EA (2014) Where am I and why? Synthesizing range biology and the eco-evolutionary dynamics of dispersal. Oikos 123: 5-22.

Kyle CJ, Boulding EG (2000) Comparative population genetic structure of marine gastropods (*Littorina* spp.) with and without pelagic larval dispersal. Mar Biol 137: 835-845.

Lambert WJ, Todd CD, Hardege JD (1997) Partial characterization and biological activity of a metamorphic inducer of the dorid nudibranch *Adalaria proxima* (Gastropoda: Nudibrancha). Invertebr Biol 116: 71-81.

Le Cam S, Pechenik JA, Cagnon M, Viard F (2009) Fast versus slow larval growth in an invasive marine mollusc: does paternity matter? J Hered 100: 455-464.

Le Cam S, Riquet F, Pechenik JA, Viard F (2014) Paternity and gregariousness in the sex-changing sessile marine gastropod *Crepidula convexa*: comparison with other protandrous *Crepidula* species. J Hered 105: 397-406.

Lee HJ, Boulding EG (2009) Spatial and temporal population genetic structure of four northeastern Pacific littorinid gastropods: the effect of mode of larval development on variation at one mitochondrial and two nuclear DNA markers. Mol Ecol 18: 2165-2184.

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

Lenormand T (2002) Gene flow and the limits to natural selection. Trends Ecol Evol 17: 183-189.

Libertini A, Vitturi R, Gregorini A, Colomba M (2009) Karyotypes, banding patterns, and nuclear DNA content in *Crepidula unguiformis* Lamark, 1822, and *Naticarius stercusmuscarum* (Gmelin, 1791) (Mollusca, Caenogastropoda). Malacologia 51: 111-118.

Lira-Noriega A, Manthey JD (2014) Relationship of genetic diversity and niche centrality: a survey and analysis. Evolution 68: 1082-1093.

Loarie SR, Duffy PB, Hamilton H, Asner GP, Field CB, Ackerly DD (2009) The velocity of climate change. Nature 462: 1052-1055.

López-Duarte PC, Carson HS, Cook GS, Fodrie FJ, Becker BJ, DiBacco C, Levin LA (2012) What controls connectivity? An empirical, multi-species approach. Integr Comp Biol 52: 511-524.

Lu F, Lipka AE, Glaubitz J, Elshire R, Cherney JH, Casler MD, Buckler ES, Costich DE (2013) Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. PLoS Genet 9: e1003215.

Marko PB, Hart MW (2011) The complex analytical landscape of gene flow inference. Trends Ecol Evol 26: 448-456.

Matsumurad K, Nagano M, Kato-Yoshinaga Y, Yamazaki M, Clare AS, Fusetani M (1998) Immunological studies on the settlement- inducing protein complex (SIPC) of the barnacle *Balanus amphitrite* and its possible involvement in larva-larva interactions. P Roy Soc B 265: 1825-1830.

McGee BL, Targett NM (1989) Larval habitat selection in *Crepidula* (L.) and its effect on adult distribution patterns. J Exp Mar Biol Ecol 131: 195-214.

McGlashan D, Ponniah M, Cassey P, Viard F (2008) Clarifying marine invasions with molecular markers: an illustration based on mtDNA from mistaken calyptraeid gastropod identifications. Biol Invasions 10: 51-57.

McMillan WO, Raff RA, Palumbi SR (1992) Population genetic consequences of developmental evolution in sea urchins (Genus *Heliocidaris*). Evolution 46: 1299-1312.

Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. Mol Ecol Res 11: 5–18.

Miller CB, Lynch DR, Carlotti F, Gentleman W, Lewis CVW (1998). Coupling of an individual-based population dynamic model of *Calanus finmarchicus* to a circulation model for the Georges Bank region. Fish Oceanogr 7: 219-234.

Moeller DA, Geber MA, Tiffin P (2011) Population genetics and the evolution of range limits in an annual plant. Am Nat 178: S44-S61.

Montgomery DC (2012) *Design and Analysis of Experiments*, 8th ed. Wiley.

Moran AL (1999) Size and performance of juvenile marine invertebrates: potential contrasts between intertidal and subtidal habitats. Amer Zool 39: 304-312.

Nei M, Maruyama T, Chakraborty R (1973) The bottleneck effect and genetic variability in populations. Evolution 29: 1-10.

Novak SJ, Mack RN (2005) Genetic bottlenecks in alien plant species. In: Sax DF, Stachowicz JJ, Gaines SD (eds) *Species Invasions: Insights into ecology, evolution, and biogeography*. Sinauer Associates, Inc., Sunderland, pp 201-228.

Olson RR, McPherson R (1987) Potential vs. realized larval dispersal: fish predation on larvae of the ascidian *Lissoclinum patella* (Gottschaldt). J Exp Mar Biol Ecol 110: 245-256.

Ordóñez V, Pascual M, Ruis M, Turon X (2013) Mixed but not admixed: a spatial analysis of genetic variation of an invasive ascidian on natural and artificial substrates. Mar Biol 160: 1645-1660.

Parmesan C (2006) Ecological and evolutionary responses to recent climate change. Annu Rev Ecol Evol S 37: 637–669.

Parmesan C, Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. Nature 421: 37–42.

Pawlik JR (1992) Chemical ecology of the settlement of benthic marine invertebrates. Oceanogr Mar Biol Annu Rev 30: 273-335.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28: 2537-2539.

Pearson RG, Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: are bioclimatic envelope models useful? Global Ecol Biogeogr 12: 361-371.

Pechenik JA (1984) The relationship between temperature, growth rate, and duration of planktonic life for larvae of the gastropod *Crepidula fornicata* (L.). J Exp Mar Biol Ecol 74: 241-257.

Pechenik JA (2006) Larval experience and latent effects – metamorphosis is not a new beginning. Integr Comp Biol 46: 323-333.

Pechenik JA, Eyster LS (1989) Influence of delayed metamorphosis on the growth and metabolism of young *Crepidula fornicata* (Gastropoda) juveniles. Biol Bull 176: 14-24.

Pechenik JA, Gee CC (1993) Onset of metamorphic competence in larvae of the gastropod *Crepidula fornicata*, judged by a natural and an artificial cue. J Exp Mar Biol Ecol 167: 59-72.

Pechenik JA, Heyman WD (1987) Using KCl to determine size at competence for larvae of the marine gastropod *Crepidula fornicata* (L.). J Exp Mar Biol Ecol 112: 27-38.

Pechenik JA, Lima GM (1984) Relationship between growth, differentiation, and length of larval life for individually reared larvae of the marine gastropod, *Crepidula fornicata*. Biol Bull 166: 537-549.

Pechenik JA, Estrella MS, Hammer K (1996) Food limitation stimulates metamorphosis of competent larvae and alters postmetamorphic growth rate in the marine prosobranch gastropod *Crepidula fornicata*. Mar Biol 127: 267-275.

Pechenik JA, Ambrogio OV, Untersee S (2010) Predation on juveniles of *Crepidula fornicata* by two crustaceans and two gastropods. J Exp Mar Biol Ecol 384: 91-98.

Penniman JR, Doall MK, Pires A (2013) Neural correlates of settlement in veliger larvae of the gastropod, *Crepidula fornicata*. Invertebr Biol 132: 14-26.

Perry AL, Low PJ, Ellis JR, Reynolds JD (2005) Climate change and distribution shifts in marine species. Science 308: 1912–1915.

Phillips BL, Brown GP, Webb JK, Shine R (2006) Invasion and the evolution of speed in toads. Nature 439: 803.

Pinsky ML, Worm B, Fogarty MJ, Sarmiento JL, Levin SA (2013) Marine taxa track local climate velocities. Science 341: 1239-1242.

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. J Hered 95: 536-539.

Price TD, Kirkpatrick M (2009) Evolutionarily stable range limits set by interspecific competition. P Roy Soc B 276: 1429-1434.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

R Core Team (2013) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. P Natl Acad Sci USA 94: 9197-9201.

Rawlings TA, Aker JM, Brunel P (2011) Clarifying the northern distributional limits of the slipper limpet *Crepidula fornicata* in the northwestern Atlantic. Am Malacol Bull 29: 105-119.

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

Rigal F, Viard F, Ayata S-D, Comtet T (2010) Does larval supply explain the low proliferation of the invasive gastropod *Crepidula fornicata* in a tidal estuary? Biol Invasions 12: 3171-3186.

Riquet F, Daguin-Thiébaut C, Ballenghien M, Bierne N, Viard F (2013) Contrasting patterns of genome-wide polymorphism in the native and invasive range of the marine mollusc *Crepidula fornicata*. Mol Ecol 22: 1003-1018.

Rius M, Darling JA (2014) How important is intraspecific genetic admixture to the success of colonising populations? Trends Ecol Evol 29: 233-242.

Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions. Trends Ecol Evol 22: 454-464.

Rousset F (2000) Genetic differentiation between individuals. J Evol Biol 13: 58–62.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol Ecol Resour 8: 103-106.

Ruesink JL, Freshley N, Herrold S, Trimble AC, Patten K (2014) Influence of substratum on non-native clam recruitment in Willapa Bay, Washington, USA. J Exp Mar Biol Ecol 459: 23-30.

Sagarin RD, Gaines SD (2002) Geographical distributions of coastal invertebrates: using one-dimensional ranges to test biogeographic hypotheses. J Biogeogr 29: 985-997.

Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA, Baughman S, Cabin RJ, Cohen JE, Ellstrand NC, McCauley DE, O'Neil P, Parker IM, Thompson JN, and Weller SG (2001) The population biology of invasive species. Annu Rev Ecol Syst 32: 305-32.

Saupe EE, Hendricks JR, Peterson AT, Lieberman BS (2014) Climate change and marine molluscs of the western North Atlantic: future prospects and perils. J Biogeogr 41: 1352-1366.

Scheltema RS (1986) On dispersal and planktonic larvae of benthic invertebrates: an eclectic overview and summary of problems. B Mar Sci 39: 290-322.

Sebens KP (1983) Settlement and metamorphosis of a temperate soft-coral larva (*Alcyonium siderium* Verrill): induction by crustose algae. Biol Bull 165: 286-304.

Selkoe KA, Toonen RJ (2011) Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. Mar Ecol-Prog Ser 436: 291-305.

Sexton JP, Strauss SY, Rice KJ (2011) Gene flow increases fitness at the warm edge of a species' range. P Natl Acad Sci USA 108: 11704-11709.

Sexton JP, McIntyre PJ, Angert AL, Rice KJ (2009) Evolution and ecology of species range limits. Annu Rev Ecol Evol S 40: 415–436.

Shanks AL (2009) Pelagic larval duration and dispersal distance revisited. Biol Bull 216: 373-385.

Shaw RG, Etterson JR (2012) Rapid climate change and the rate of adaptation: insight from experimental quantitative genetics. New Phytol 195: 752-765.

Siegel DA, Kinlan BP, Gaylord B, Gaines SD (2003) Lagrangian descriptions of marine larval dispersion. Mar Ecol-Prog Ser 260: 83-96.

Sievert PR, Keith LB (1985) Survival of snowshoe hares at a geographic range boundary. J Wildl Manage 49: 854-866.

Simon-Bouhet B, Garcia-Meunier P, Viard F (2006) Multiple introductions promote range expansion of the mollusc *Cyclope neritea* (Nassariidae) in France: evidence from mitochondrial sequence data. Mol Ecol 15: 1699-1711.

Slayter RA, Hirst M, Sexton JP (2013) Niche breadth predicts geographical range size: a general ecological pattern. Ecol Lett 16: 1101-1114.

Sokal RR, Rohlf FJ (1995) *Biometry*. W.H. Freeman and Company, New York.

Strathmann MF, Strathmann RR (2007) An extraordinarily long larval duration of 4.5 years from hatching to metamorphosis for teleplanic veligers of *Fusitriton oregonensis*. Biol Bull 213: 153-159.

Sunday JM, Bates AE, Dulvy NK (2012) Thermal tolerance and the global redistribution of animals. Nature Clim Change 2: 686–690.

Taris N, Comtet T, Stolba R, Lasbleiz R, Pechenik JA, Viard F (2010) Experimental induction of larval metamorphosis by a naturally-produced halogenated compound (dibromomethane) in the invasive mollusc *Crepidula fornicata* (L.). J Exp Mar Biol Ecol 393: 71-77.

Thieltges, DW, Strasser, M, Reise, K (2006) How bad are invaders in coastal waters? The case of the American slipper limpet *Crepidula fornicata* in western Europe. Biol Invasions, 8: 1673-1680.

Thieltges DW, Strasser M, van Beusekom JEE, Reise K (2004) Too cold to prosper – winter mortality prevents population increase of the introduced American slipper limpet *Crepidula fornicata* in northern Europe. J Exp Mar Biol Ecol 311: 375-391.

Thomas CD (2010) Climate, climate change, and range boundaries. Divers Distrib 16: 488–495.

Thornton PK, Eriksen PJ, Herrero M, Challinor AJ (In press) Climate variability and vulnerability to climate change: a review. Glob Change Biol. DOI: 10.1111/gcb.12581

Thorson G (1950) Reproductive and larval ecology of marine benthic invertebrates. Biol Rev Camb Philos Soc 25: 1-45.

Toonen R J, Pawlik JR (2001) Settlement of the gregarious tube worm *Hydroides dianthus* (Polychaeta: Serpulidae). II. Testing the desperate larva hypothesis. Mar Ecol-Prog Ser 224: 115-131.

Townsend CH (1895) The transplanting of eastern oysters to Willapa Bay, Washington, with notes on the native oyster industry. In: *United States Commision of Fish and Fisheries. Part XXI. Report of the commissioner for the year ending June 30, 1895*. 193-202.

Treml EA, Roberts JJ, Chao Y, Halpin PN, Possingham HP, Riginos C (2012) Reproductive output and duration of the pelagic larval stage determine seascape-wide connectivity of marine populations. Integr Comp Biol 52: 525-537.

Tsukamoto S, Kato H, Hirota H, Fusetani N (1999) Lumichrome: a larval metamorphosis-inducing substance in the ascidian *Halocynthia roretz*. Euro J Biochem 264: 785- 789.

Viard F, Ellien C, and Dupont L (2006) Dispersal ability and invasion success of *Crepidula fornicata* in a single gulf: insights from genetic markers and larval-dispersal model. Helgoland Mar Res 60: 144-152.

Visser ME (2008) Keeping up with a warming world; assessing the rate of adaptation to climate change. P Roy Soc B 275: 649-659.

von Ende CN (2001) Repeated-measures analysis: growth and other time-dependent measures. In: Scheiner SM, Gurevitch J (eds) *Design and Analysis of Ecological Experiments*. Oxford, Oxford University Press, pp. 134-157.

Wares JP (2002) Community genetics in the northwest Atlantic intertidal. Mol Ecol 11: 1131-1144.

Wares JP, Cunningham CW (2001) Phylogeography and historical ecology of the North Atlantic intertidal. Evolution 55: 2455-2469.

Weersing K, Toonen RJ (2009) Population genetics, larval dispersal, and connectivity in marine systems. Mar Ecol-Prog Ser 393: 1-12.

Whitlock MC (2011) G'_{ST} and D do not replace F_{ST}. Mol Ecol 20: 1083-1091.

Williamson J E, de Nys R, Kumar N, Carson D G, Steinberg P D (2000) Induction of metamorphosis of a marine invertebrate by the metabolites floridoside and isethionic acid from its algal host. Biol Bull 198: 332- 345.

Wonham MJ, O'Connor M, Harley CDG (2005) Positive effects of a dominant invader on introduced and native mudflat species. Mar Ecol-Prog Ser 289: 109-116.

Woodin SA (1985) Effects of defecation by arenicolid polychaete adults on spionid polychaete juveniles in field experiments: Selective settlement or differential mortality? J Exp Mar Biol Ecol 87: 119-132.

Woodin SA (1986) Settlement of infauna: larval choice? Bull Mar Sci 39: 401-407.

Wright S (1931) Evolution in Mendelian populations. Genetics 16: 97-159.

Wu YS, Tang CL (2011) Atlas of ocean currents in eastern Canadian waters. Canadian Technical Report of Hydrography and Ocean Sciences 271. Fisheries and Oceans Canada.

Zhao B, Qian P-Y (2002) Larval settlement and metamorphosis in the slipper limpet *Crepidula onyx* (Sowerby) in response to conspecific cues and the cues from biofilm. J Exp Mar Biol Ecol 269: 39-51.

Zimmer-Faust RK, Tamburri MN (1994) Chemical identity and ecological implications of a waterborne, larval settlement cue. Limnol Oceanogr 39: 1075-1087.

Appendices

Appendix A: Supplement to Chapter 2.

Table S.2.1 Genetic diversity information for six native populations and one non-native population of *Crepidula convexa* (native population NC not included). Genetic diversity information is averaged across six microsatellite loci. Native populations are arranged from north to south. Bold values are significant at $p < 0.01$. Values for the introduced population shown in *italics*.

Population	**Number genotyped**	**Allelic richness**	**H_e**	**H_o**	**F_{is}**
NA	24	12.08	0.899	0.882	0.018
BA	64	12.56	0.907	0.823	0.094
WM	83	15.66	0.940	0.796	0.152
JB	15	15.16	0.942	0.833	0.116
DE	29	14.74	0.929	0.767	0.179
VA	32	13.99	0.915	0.807	0.129
PB	*85*	*12.89*	*0.900*	*0.785*	*0.124*

Table S.2.2 Pairwise F_{st} values for seven native populations and one introduced population of *Crepidula convexa*, calculated using five microsatellite loci. Native populations are arranged from north to south. Values for the introduced population shown in *italics*.

	NA	**BA**	**WM**	**JB**	**DE**	**VA**	**NC**	***PB***
NA								
BA	0.0273							
WM	0.0359	0.0280						
JB	0.0346	0.0322	0.0057					
DE	0.0388	0.0357	0.0270	0.0233				
VA	0.0480	0.0567	0.0485	0.0271	0.0347			
NC	0.0512	0.0579	0.0314	0.0285	0.0304	0.0378		
PB	*0.0703*	*0.0626*	*0.0270*	*0.0351*	*0.0527*	*0.0362*	*0.0619*	

Table S.2.3 Pairwise D_{est} values for seven native populations and one introduced population of *Crepidula convexa*, calculated using five microsatellite loci. Native populations are arranged from north to south. Values for the introduced population shown in *italics*.

	NA	**BA**	**WM**	**JB**	**DE**	**VA**	**NC**	**PB**
NA								
BA	0.261							
WM	0.426	0.380						
JB	0.403	0.403	0.096					
DE	0.412	0.423	0.379	0.340				
VA	0.436	0.596	0.581	0.341	0.370			
NC	0.555	0.670	0.445	0.430	0.386	0.415		
PB	*0.620*	*0.630*	*0.357*	*0.380*	*0.527*	*0.646*	*0.621*	

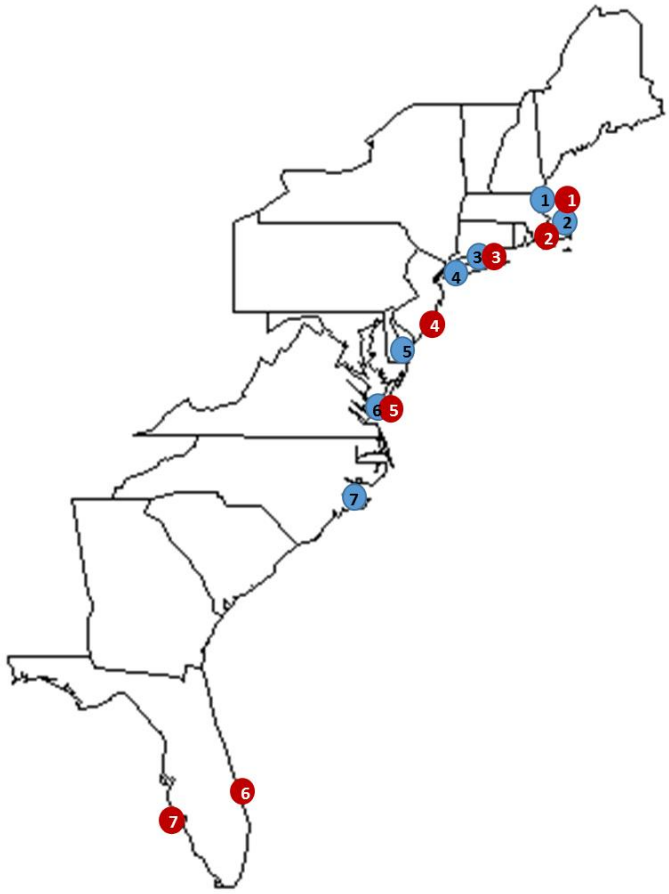

Fig. S.2.1 Map of collecting sites for *Crepidula convexa* (blue) and *C. fornicata* (red). Numbers of *C. fornicata* sites correspond to those in Riquet et al. (2013). *Crepidula convexa* sites: 1 = Nahant, Massachusetts; 2 = Barnstable, Massachusetts; 3 = Old Field, New York; 4 = Jamaica Bay, New York; 5 = Lewes, Delaware; 6 = Gloucester Point, Virginia; 7 = Beaufort, North Carolina. *Crepidula fornicata* sites: 1 = Nahant, Massachusetts; 2 = Fairhaven, Massachusetts; 3 = Long Island, New York; 4 = Somers Point, New Jersey; 5 = Chesapeake, Virginia; 6 = Fort Pierce, Florida; 7 = Longboat Key, Florida.

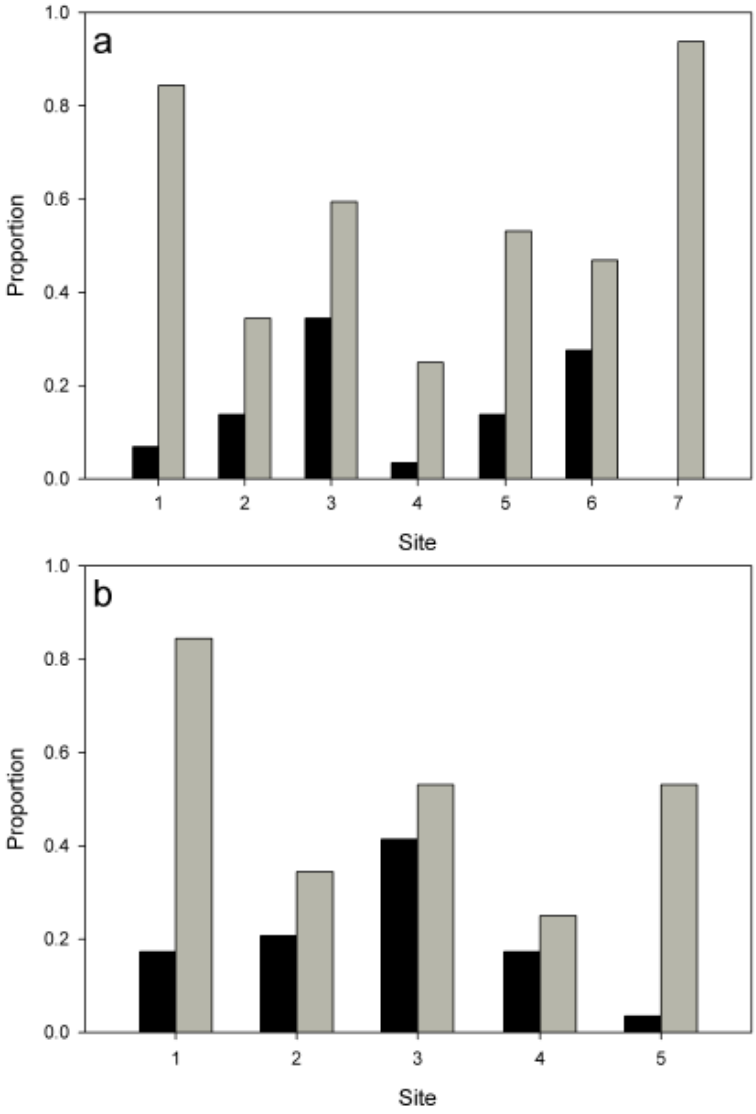

Fig. S.2.2 a) Proportion of individuals from the introduced (Mud Bay) population of *Crepidula fornicata* that were assigned to (dark bars) or rejected from (with 95% probability; light bars) each of the seven native populations studied in Riquet et al. (2013). Populations 6 and 7 are in a region not covered in the current study (i.e., Florida: western Atlantic (6) and Gulf of Mexico (7)). b) Removing these two populations does not qualitatively change the assignment results: the introduced population is assigned to several native populations.

Analyses were done using GeneClass as explained for *C. convexa* in the main text. Population numbers refer to notations from Riquet et al. (2013) and correspond to the red circles in Fig. S.2.1.

Appendix B: Supplement to Chapter 5. Optimization of experimental parameters.

A discussion of the results of each graph is provided. Each graph represents the cumulative proportion of larvae settled through time under different trial conditions. All values are treatment totals out of a sample size of 30 larvae (10 larvae per glass; three glasses). Apparent decreases in settlement through time are due to larval mortality during the trial. These graphs were visually inspected to determine the apparent optimal value of each parameter (i.e., highest settlement rate with lowest mortality); no statistical analyses were conducted. Only the variable being tested was altered in each trial; the selected value was then used for all other variables as described in Table 6 of the main text.

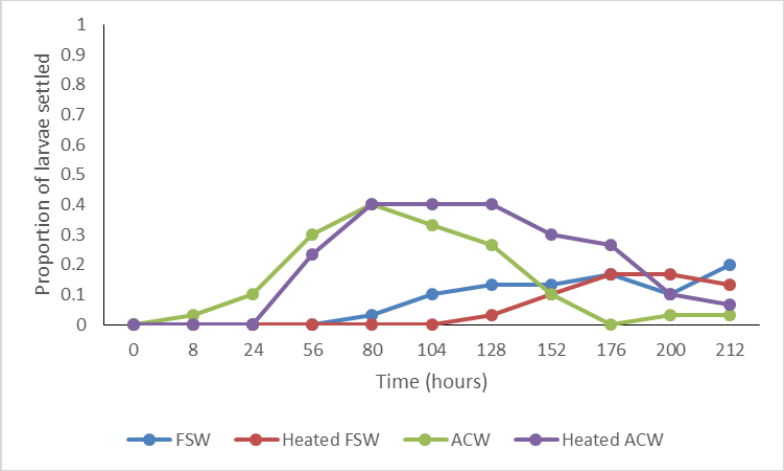

Fig. S.5.1 Trial duration. Data from a pilot experiment testing the interaction of ACW and heat. By 72 hours, mortality was greater than settlement in the ACW treatments, and spontaneous settlement began to increase in the FSW treatments. Forty-eight hours was selected as the optimal trial duration.

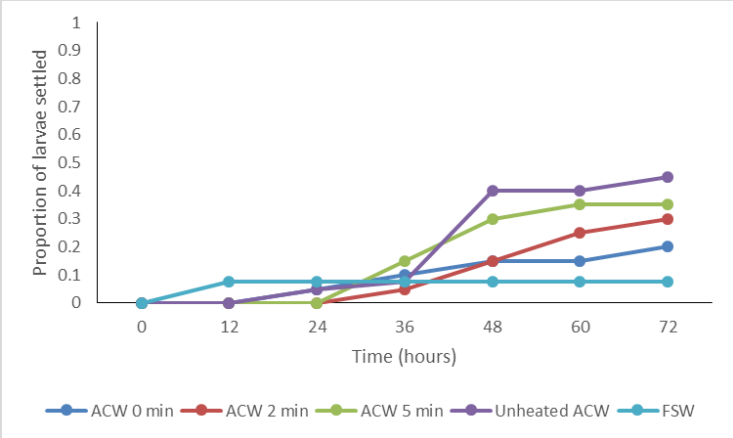

Fig. S.5.2 Boiling time. Data from a pilot experiment where ACW was brought to a boil for different amounts of time (0-5 minutes). Both a zero-minute boil (water removed from the microwave when it started to boil) and a two-minute boil depressed settlement below the level of unheated ACW. The five-minute boil did not depress settlement, and also increased salinity outside of the optimal range (i.e., 27-30 psu; Table 6 in main text), leading to increased mortality. A two-minute boil was chosen as the optimal value.

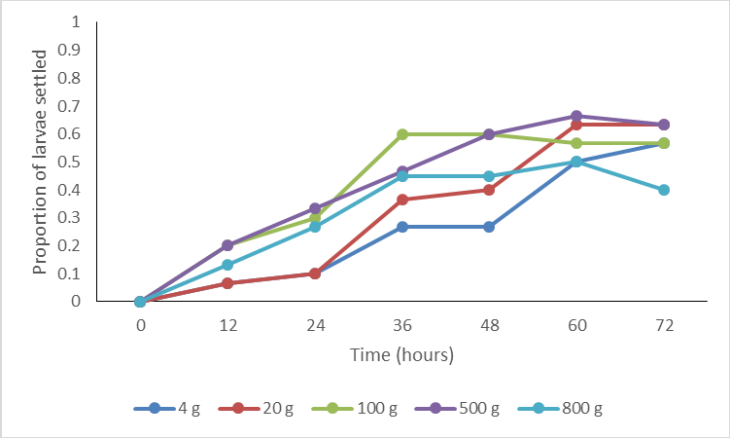

Fig. S.5.3 Adult mass used to prepare ACW. Data from a pilot experiment where all ACW was prepared for 12 hours with different adult mass (total wet mass = 4g, 20g, 100g, 500g, or 800g of adult mass per liter of FSW). The treatment made with 800 g of adult mass showed >75% mortality by the end of the experiment. 100 g and 500 g were comparable, but 100 g was chosen as the optimal value because it required fewer adults for preparation.

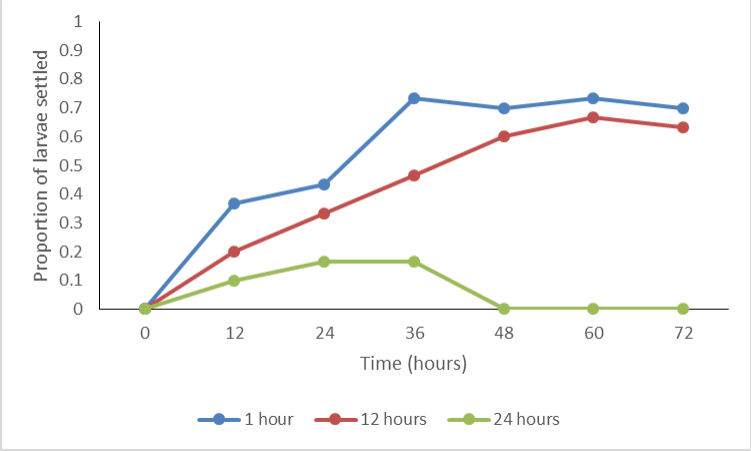

Fig. S.5.4 Cue preparation time. Data from a pilot experiment where all ACW was prepared with 100 g adult mass per liter of FSW for different durations of time (1 h, 12 h, 24 h). ACW prepared for 24 hours resulted in total larval mortality by 48 h during the settlement trial (presumably due to low oxygen and high bacterial and protozoan concentrations in the ACW in

this treatment). Although 1 hour and 12 hour preps were roughly comparable, 12 hours of preparation time was chosen for logistical reasons.

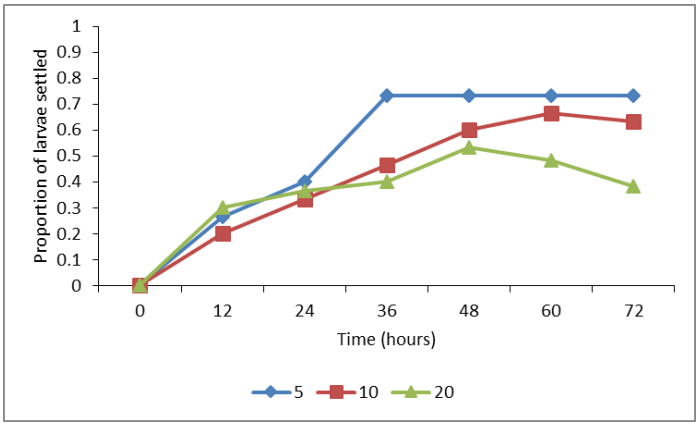

Fig. S.5.5 Larvae per glass. Data from a pilot experiment of 5, 10, or 20 larvae in each 20 ml glass of ACW prepared using the optimal values (Table 6). High mortality was seen with 20 larvae per glass, and individual larvae were difficult to examine. We therefore selected 10 larvae per glass.

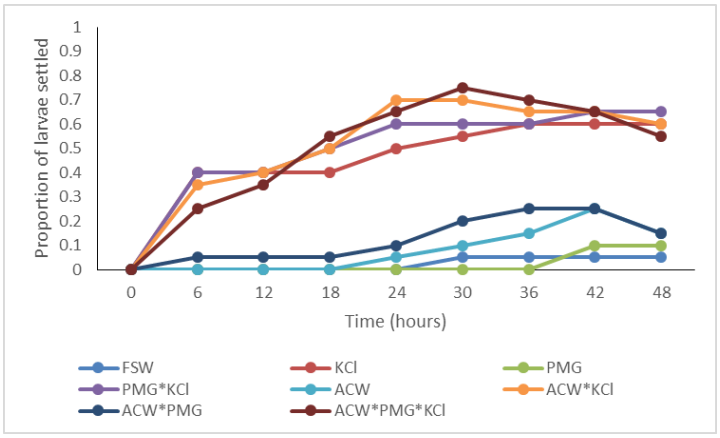

Fig. S.5.6 Time interval. Data from a pilot experiment of ACW, PMG, and KCl in a factorial design (8 total treatments). Data were collected every six hours for 48 hours. Based on this trial, data were ultimately taken every 12 hours for logistical reasons and to provide sufficient data to estimate lambda.

Appendix C: Supplement to Chapter 5. Settlement data as analyzed using proportions.

Settlement data from Chapter 5. The proportion of larvae settled at each timepoint from each experiment has been arcsine square root transformed to meet normality assumptions and analyzed using ANOVA, rather tha with the rate constant λ. The significance of both main effects and interactions changes through time.

Table S.5.1 Analyses of variance for Experiment 1 (screening factors that induce settlement). ANOVAs are conducted on the proportion of larvae settled at each of four time points (12, 24, 36, and 48 hours). Analyses conducted on arcsine square root-transformed data due to failure to meet the assumption of normality. p-values significant at < 0.05 are highlighted in bold.

12 hours:

Source of Variation	df	SS	MS	F	p
Block Effect	2	3808.75	1904.38	----	----
Adult Conditioned Water (ACW)	1	103.10	103.10	0.69	0.408
Pedal Mucus Glass (PMG)	1	0.96	0.96	0.01	0.936
Potassium Chloride (KCl)	1	5647.01	5647.01	37.98	< **0.001**
PMG*ACW	1	15.25	15.25	0.10	0.750
PMG*KCl	1	151.44	151.44	1.02	0.317
ACW*KCl	1	2318.35	2318.35	15.59	< **0.001**
ACW*PMG*KCl	1	241.34	241.34	1.62	0.207
Error	62	9218.95	148.69	0.00	
Total	71	21802.54			

24 hours:

Source of Variation	df	SS	MS	F	p
Block Effect	2	3561.12	1780.56	----	----
Adult Conditioned Water (ACW)	1	184.61	184.61	1.09	0.301
Pedal Mucus Glass (PMG)	1	0.70	0.70	0.00	0.949
Potassium Chloride (KCl)	1	8249.06	8249.06	48.69	< **0.001**
PMG*ACW	1	257.00	257.00	1.52	0.223
PMG*KCl	1	47.52	47.52	0.28	0.598
ACW*KCl	1	2176.13	2176.13	12.84	< **0. 001**
ACW*PMG*KCl	1	308.72	308.72	1.82	0.182
Error	62	10504.68	169.43	0.00	
Total	71	25628.40			

36 hours:

Source of Variation	df	SS	MS	F	p
Block Effect	2	3690.82	1845.41	----	----
Adult Conditioned Water (ACW)	1	1296.93	1296.93	10.11	**0.002**
Pedal Mucus Glass (PMG)	1	506.36	506.36	3.95	0.0514
Potassium Chloride (KCl)	1	4041.01	4041.01	31.51	**< 0.001**
PMG*ACW	1	532.14	532.14	4.15	**0.046**
PMG*KCl	1	1043.33	1043.33	8.13	**0.006**
ACW*KCl	1	3332.55	3332.55	25.98	**< 0.001**
ACW*PMG*KCl	1	280.06	280.06	2.18	0.145
Error	62	7952.36	128.26	0.00	
Total	71	22932.08			

48 hours:

Source of Variation	df	SS	MS	F	p
Block Effect	2	2981.98	1490.99	----	----
Adult Conditioned Water (ACW)	1	1167.98	1167.98	8.21	**0.006**
Pedal Mucus Glass (PMG)	1	531.98	531.98	3.74	0.058
Potassium Chloride (KCl)	1	4877.59	4877.59	34.29	**< 0.001**
PMG*ACW	1	737.22	737.22	5.18	**0.026**
PMG*KCl	1	1255.76	1255.76	8.83	**0.004**
ACW*KCl	1	3121.16	3121.16	21.94	**< 0.001**
ACW*PMG*KCl	1	407.50	407.50	2.87	0.096
Error	62	8818.34	142.23	0.00	
Total	71	24183.96			

Table S.5.2 Analyses of variance for Experiment 2 (inactivation of factors that induce settlement) at each time point. Analyses conducted on arcsine square root-transformed data due to failure to meet the assumption of normality. p-values significant at < 0.05 are highlighted in bold.

12 hours:

Source of Variation	df	SS	MS	F	p
Block Effect	2	794.73	397.36	----	----
Adult Conditioned Water (ACW)	1	1557.94	1557.94	14.90	< **0.001**
Pedal Mucus Glass (PMG)	1	161.10	161.10	1.54	0.219
Heat Treatment (H)	1	350.07	350.07	3.35	0.072
ACW*PMG	1	132.36	132.36	1.27	0.265
ACW*H	1	10.25	10.25	0.10	0.755
PMG*H	1	4.88	4.88	0.05	0.830
ACW*PMG*H	1	483.50	483.50	4.62	**0.035**
Error	62	6484.06	104.58		
Total	71	10188.03			

24 hours:

Source of Variation	df	SS	MS	F	p
Block Effect	2	498.44	249.22	----	----
Adult Conditioned Water (ACW)	1	3470.69	3470.69	29.18	< **0.001**
Pedal Mucus Glass (PMG)	1	1204.67	1204.67	10.13	**0.002**
Heat Treatment (H)	1	506.10	506.10	4.26	**0.043**
ACW*PMG	1	40.43	40.43	0.34	0.562
ACW*H	1	230.09	230.09	1.93	0.169
PMG*H	1	68.89	68.89	0.58	0.449
ACW*PMG*H	1	591.28	591.28	4.97	**0.029**
Error	62	7373.39	118.93		
Total	71	14221.83			

36 hours:

Source of Variation	df	SS	MS	F	p
Block Effect	2	1039.81	519.91	----	----
Adult Conditioned Water (ACW)	1	7986.16	7986.16	56.86	< **0.001**
Pedal Mucus Glass (PMG)	1	2839.19	2839.19	20.21	< **0.001**
Heat Treatment (H)	1	440.20	440.20	3.13	0.082
ACW*PMG	1	0.65	0.65	0.00	0.946
ACW*H	1	477.04	477.04	3.40	0.070
PMG*H	1	56.80	56.80	0.40	0.527
ACW*PMG*H	1	44.57	44.57	0.32	0.575
Error	62	8707.98	140.45		
Total	71	21873.32			

48 hours:

Source of Variation	df	SS	MS	F	p
Block Effect	2	1329.34	664.67	----	----
Adult Conditioned Water (ACW)	1	11332.15	11332.15	69.79	< **0.001**
Pedal Mucus Glass (PMG)	1	4227.40	4227.40	26.03	< **0.001**
Heat Treatment (H)	1	42.69	42.69	0.26	0.610
ACW*PMG	1	61.79	61.79	0.38	0.540
ACW*H	1	425.35	425.35	2.62	0.111
PMG*H	1	139.17	139.17	0.86	0.358
ACW*PMG*H	1	5.26	5.26	0.03	0.858
Error	62	10067.81	162.38		
Total	71	27955.72			

Table S.5.3 Analyses of variance for Experiment 3 (localization of factors that induce settlement) at each time point. Analyses conducted on arcsine square root-transformed data due to failure to meet the assumption of normality. p-values significant at < 0.05 are highlighted in bold.

12 hours:

Source of Variation	df	SS	MS	F	p
Adult Conditioned Water (ACW)	2	593.88	296.94	5.31	**0.015**
Pedal Mucus Glass (PMG)	2	389.66	194.83	3.48	0.053
PMG*ACW	4	412.61	103.15	1.84	0.165
Error	18	1007.28	55.96		
Total	26	2403.43			

24 hours:

Source of Variation	df	SS	MS	F	p
Adult Conditioned Water (ACW)	2	1972.60	986.30	17.41	< **0.001**
Pedal Mucus Glass (PMG)	2	672.40	336.20	5.94	**0.010**
PMG*ACW	4	603.63	150.91	2.66	0.066
Error	18	1019.58	56.64		
Total	26	4268.21			

36 hours:

Source of Variation	df	SS	MS	F	p
Adult Conditioned Water (ACW)	2	1701.96	850.98	11.16	**0.001**
Pedal Mucus Glass (PMG)	2	1406.15	703.08	9.22	**0.002**
PMG*ACW	4	583.98	145.99	1.91	0.152
Error	18	1372.84	76.27		
Total	26	5064.93			

48 hours:

Source of Variation	df	SS	MS	F	p
Adult Conditioned Water (ACW)	2	2092.66	1046.33	13.44	< **0.001**
Pedal Mucus Glass (PMG)	2	1454.22	727.11	9.34	**0.002**
PMG*ACW	4	584.06	146.02	1.88	0.159
Error	18	1401.33	77.85		
Total	26	5532.27			