THE POPULATION AND ECOLOGICAL GENETIC EFFECTS OF HABITAT FRAGMENTATION

A Thesis
Submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Master of Science
In Biology

By

Amy Battocletti, M.S.

Washington, DC
April 20, 2017
Acknowledgements

I would like to thank the members of the Hamilton and Wimp labs for their help in both the field and lab. I am grateful for the helpful ideas and discussion from Drs. Gina Wimp and Matthew Hamilton. I would also like to acknowledge and express thanks to the many students who helped. Monica Brady helped in both the lab and field on both the *T. minuta* and *S. patens* datasets, including the primer design and microsatellite marker development for *S. patens* and the spatial genetic structure analyses. Melissa Naugle helped in both the lab and field during the summer of 2016 and was instrumental in obtaining the 2015 *T. minuta* genetic dataset. Sunny Parmer aided in the lab work required to process the 2011 *T. minuta* samples. Sarah Qadir aided in both the lab and field work, particularly in the density counts of *T. minuta* and performed a greenhouse experiment to investigate the nature of *Pardosa* predation on *T. minuta*. My husband, Robert Baticletti, wrote the python code both to implement the methods of Sved et al. (2013), as described in Chapter 2, and to search the genomic sequencing reads for microsatellites.
THE POPULATION AND ECOLOGICAL GENETIC EFFECTS OF HABITAT FRAGMENTATION

Amy Battocletti, M.S.

Thesis Advisor: Gina Wimp, Ph.D.

ABSTRACT

Maintaining intraspecific variation is important for populations’ long-term success and is increasingly being recognized as an important conservation goal. Populations in anthropogenically fragmented habitats may lose variation rapidly via genetic drift, particularly in small fragments with a high ratio of edge to interior habitat. We studied the population and ecological genetic effects of habitat fragmentation on both a foundation plant, *Spartina patens*, and a dependent herbivore, *Tumidagena minuta*, using a naturally fragmented, salt marsh model system. We employed microsatellite marker analyses to estimate various measures of genetic variation, including allelic richness and heterozygosity, and to estimate the strength of genetic drift using estimates of effective population size ($N_e$). To achieve this, we developed a new program to estimate $N_e$ and developed new markers for *S. patens* from genome sequence data. We found lower *S. patens* genetic variation and lower $T. minuta$ $N_e$ near the *S. alterniflora* edges, indicating that *T. minuta* experience stronger genetic drift near edges. These findings reinforce the importance of habitat patch shape in influencing populations.
TABLE OF CONTENTS

Chapter 1: Introduction........................................................................................................1
Chapter 2: Development of Ne Estimation Program.........................................................19
Chapter 3: Population Genetic Patterns in *Tumidagena minuta* ..................................33
Chapter 4: *Spartina patens* Spatial Population Genetic Structure Analysis Using New
Microsatellite Markers ......................................................................................................47
Chapter 5: Conclusions ....................................................................................................58
Appendix A .......................................................................................................................61
Appendix B .......................................................................................................................68
Bibliography ....................................................................................................................79
CHAPTR 1: INTRODUCTION

Anthropogenic habitat destruction and fragmentation are leading causes of species extinctions and biodiversity loss (Didham et al. 2012). Habitat fragmentation is the process of converting continuous habitat into fragments, often via habitat loss. Species responses to fragmentation are highly variable as they are influenced by numerous factors, including the amount of habitat lost and time since fragmentation, patch and matrix characteristics, connectivity, and life-history traits (Laurance 2008). Patch area has been shown to exhibit a positive species-area relationship; however, the amount of core area, or interior, within a patch may be more important than the total area. As edges reduce the core area, this may effectively alter organismal area responses from what would be expected simply based on total area alone, resulting in non-linear relationships (Didham and Ewers 2012). Patch area effects are often governed by edge responses, particularly in patches with a high edge:interior habitat ratio, and this may alter responses that are expected based on a simple linear relationship to area or that assume a constant patch geometry. It is critical that we not only understand the effects of patch area, isolation and connectivity on species and ecosystems, but also that of edge:interior ratios and edge responses.

Although species show a variety of responses to edges, we can identify general patterns that are crucial for conservation management. Edge responses are often influenced by the spatial separation of resources and ecological flows between adjacent habitats, and species interactions and resource mapping (Ries et al. 2004). Edges were once thought to be beneficial to wildlife, but the increase in richness along edges may be
driven by generalists that are not spatially confined to one habitat, and can therefore use resources on both sides of the edge. These species tend to exhibit positive edge responses. In contrast, specialist species, including many rare species of conservation concern (Lande 1987, Laurance 2008), tend to exhibit negative responses to edges (Ries et al. 2004), particularly when the matrix habitat is inhospitable or anthropogenically disturbed (Martinson and Fagan 2014). For habitat specialists, the initial habitat loss is coupled with a population decline. Genetic variation may be lost in the initial population decline, and then lost thereafter via genetic drift associated with small population sizes. Edge effects may play an especially important role in small patches since patches are subject to the same forces affecting edge habitat at a scale determined by the edge:interior habitat ratio (Ewers and Didham 2007, Fletcher et al. 2007). Fragment area and geometry both affect the edge:interior ratio, and very small and complexly shaped fragments may even be entirely composed of edge habitat (Didham and Ewers 2012). For this reason, the edge:interior ratio is considered a major driver of area effects (Laurance et al. 2007), particularly for specialists (Matthews et al. 2014).

The affinity of a species for a particular habitat or matrix is critical to determining its sensitivity to fragmentation. Most vulnerable are species that find the matrix inhospitable and therefore tend to experience the habitat fragmentation process as a population fragmentation event. To conserve local or rare diversity, it is important to distinguish the species richness contributions of cosmopolitan and commensal species from those that are unique to a particular ecosystem or region. High mobility species that are capable of living in a variety of habitats, particularly disturbed habitats, may do well post-fragmentation (With and Crist 1995), eventually displacing local specialists.
General conclusions drawn from results that are biased by ubiquitous species may not be applicable to conservation management. There is a strong need to focus on the effects of fragmentation specifically on rare and regionally limited species, which tend to experience habitat changes as a population fragmentation event. **We aim to investigate the effects of habitat fragmentation on the population densities of a habitat specialist that exhibits negative edge responses.**

Conservation biologists are not only concerned with the loss of species diversity, but also the loss of intraspecific genetic variation. Maintaining intraspecific variation is an important conservation goal as it increases a population’s resiliency to disturbance and enables adaptation to a changing environment. In diverse populations, fitter phenotypes may compensate for less fit phenotypes. Phenotypes may respond differentially to environmental conditions, and their relative fitness may fluctuate over time. Low diversity populations, however, may lack the variety of phenotypes necessary for adequate compensatory dynamics during periods of stress and thereby suffer reduced resiliency. Additionally, they may lack the genetic variation necessary to respond to new selective pressures. In small populations, reduced fitness may lead to further declines in the population size and genetic variation loss, creating an “extinction vortex” (Fagan and Holmes 2006). For habitat managers, this means employing practices that maintain a high level of standing genetic variation in the managed system. This is especially critical when considering the fast rate of change and increased variability expected in many habitats due to climate change and human alteration (Burger and Lynch 1995).

Genetic variation may not only be lost in the initial population decline, but also thereafter via increased genetic drift and reduced gene flow, leading to increased
homogenization within populations (Zhang 2009). Genetic drift is a stochastic process; it is the random change in allele frequencies that occurs between the parental and offspring generations. Over time, drift acts to reduce genetic variation by ultimately bringing alleles to fixation or loss (Hamilton 2009). This process may particularly affect variation in small fragments, which may experience extinction/recolonization events more frequently than larger fragments. Drift also impacts a population’s ability to respond to selective pressures. In small populations, drift reduces the magnitude of the population’s response to selection and may eliminate it entirely. In anthropogenically fragmented systems, genetic variation may be lost even as the census size stabilizes, resulting in a “silent” loss of biodiversity. However, in well-connected meta-populations, gene flow may act to homogenize the allele frequencies between sub-populations, countering differentiation via drift (Hamilton 2009). We aim to investigate the effects of habitat fragmentation on the genetic variation of a habitat specialist that exhibits negative edge responses.

Habitat fragmentation research commonly focuses on species abundances and community composition; however, fragmentation may have the largest effect on ecological interactions. For example, patch size and connectivity has been shown to influence the complexity and presence of food web modules (Martinson and Fagan 2014), which may then cascade to further alter species presence and abundances. Niche-based competitive interactions have also been shown to contribute to fragmentation induced biodiversity loss (Bregman et al. 2015). In addition to altering communities via species presence/absence, we ask if landscape characteristics can alter ecological interactions via changes in intraspecific variation, which can then cascade to impact
higher trophic levels? Habitat fragmentation tends to reduce resource consumption in consumers with narrow niches, such as specialist herbivores that require a particular host plant for food or habitat (Martinson and Fagan 2014), indicating that changes in host plant intraspecific variation may particularly affect specialist herbivores. A prediction of ecological genetic theory is that gene expression in one individual can act as an environmental influence on other individuals, and intraspecific variation in plants has been found to influence diversity at higher trophic levels (Bangert et al. 2006). These genetic interactions are termed intra- or inter-specific indirect genetic effects depending on whether the two individuals are the same or different species (Whitham et al. 2006). Greater intraspecific variation typically entails a greater diversity of interspecific indirect genetic effects influencing the community (Whitham et al. 2006). For this reason, plant species’ intraspecific variation may be an especially important mechanism underlying biodiversity. **We aim to investigate the impact of genetic variation within a foundation plant species on the genetic variation within a dependent herbivore insect in a fragmented system.**

As wildlife species continue to be threatened by habitat loss and fragmentation, it is critical that we understand not only the effects on single species demographics, but also how fragmentation may affect intraspecific variation, and how those changes may cascade up to impact higher trophic levels. In this study, we ask 1) How do landscape characteristics influence the spatial population genetic structure of a habitat specialist? and 2) How do landscape characteristics affect the spatial population genetic structure of a dominant plant species? We investigate these questions using a salt marsh model system. The marsh is dominated by two habitat forming grass species in which
*Spartina patens* is dispersed throughout a matrix of *Spartina alterniflora*, forming a naturally patchy system. The specialist herbivore, *Tumidagena minuta*, lives isolated within patches of *S. patens*, forming a meta-population across the marsh. Both cordgrass species harbor a diversity of generalist and specialist arthropods, providing an ideal model for habitat fragmentation research.

**BACKGROUND AND MODEL SYSTEM**

*Population Genetics*

The effects of population fragmentation on genetic variation may be most strongly governed by the stochastic process of genetic drift, which is the random change in allele frequencies that occurs through sampling error during reproduction (Hamilton 2009). Essentially, the frequencies in the offspring generation are likely to differ from the parental generation simply by chance. The strength of drift is often assessed via estimates of the effective population size ($N_e$), which is the size of a Wright-Fisher population that would experience the same amount of drift as empirically observed. The Wright-Fisher model makes certain simplifying assumptions for modeling purposes, including a temporally constant population size, non-overlapping generations, equal sex ratios and family sizes, random mating, and an infinite gamete pool. Although genetic drift is a stochastic process, given the initial conditions, the changes in allele frequencies over time can be predicted using a probability distribution (Hamilton 2009), which may be used to estimate the rate at which genetic variation may be lost. This makes estimates of $N_e$ a powerful predictive tool for managers.
Within panmictic populations, the rate at which allelic loss occurs can be predicted using the equation $1 - (1/2N_e)$. The probability of fixation is equal to the allele’s initial frequency at any given point in time. Alleles present at high frequencies are most likely to become fixed, whereas alleles at intermediate frequencies are expected to segregate the longest, and rare alleles are likely to be lost first. Small patches may exhibit a deficit of rare alleles, whereas older, larger, or expanding populations may harbor an excess of rare alleles (Hamilton 2009).

There are multiple ways to estimate $N_e$, each yielding a unique interpretation. Methods using gametic disequilibrium, which is the non-random association of alleles at two loci, produce estimates of the effective number of breeders in the parental population. Due to the numerous underlying assumptions that eliminate most sources of disequilibrium (Figure 1), the disequilibrium present in a Wright-Fisher population is assumed to entirely occur through the random sampling of a finite number of gametes from the gamete pool. The magnitude of the allele frequency fluctuations between two generations is expected to increase as $N_e$ decreases. Temporally, $N_e$ can be estimated using the magnitude of the fluctuations between any two generations, which is equal to the standard error, $SE = \sqrt{(pq/2N_e)}$ for a di-allelic locus (Hamilton 2009). This yields an estimate that indicates the harmonic mean of the generations preceding the first and last generation used to calculate the $N_e$ (Waples 2005).

Estimates of the effective population size ($N_e$) represent the instantaneous strength of genetic drift in a Wright-Fisher model, thereby facilitating the distinction between instantaneous and historical processes (Rhodes and Chesser 1994, Wang et al. 2011). More specifically, genetic association between unlinked loci tends to decay over
time, and based on a recombination rate of 0.5, is predicted to approach zero after approximately seven generations. This means that any disequilibrium detected is assumed to have been initiated within that time period (Hamilton 2009). For the analysis of contemporary processes and prediction of trends, $N_e$ may serve as a useful currency for assessing the relative strengths of the evolutionary processes at any given location. Over long time scales, genetic variation is typically increased by mutation and gene flow, and lost via drift and selection. These processes are affected by both abiotic and biotic factors, such as intra and interspecific interactions, and their relative strengths are governed by $N_e$ (Hamilton 2009). Compared to the census size, the $N_e$ more accurately represents the behavior of a population in terms of the action of the various evolutionary processes, and is a valuable predictive tool for managers (Wang et al. 2011).

Maintaining populations’ adaptive capacity is an important conservation goal, and the strong drift associated with small $N_e$ can weaken a population’s ability to respond to selection. The relationship between drift and selection is characterized by the composite parameter $4N_e s$. When $4N_e s$ is greater than one, selection will tend to dominate either due to a large $N_e$ or very strong selection. If $4N_e s$ is roughly equal to one, then selection and drift are balanced. When $4N_e s$ is less than one, drift is the dominant force either due to a small $N_e$ or weak selection. In this case, drift may bring weakly deleterious alleles to fixation despite negative selection and beneficial alleles to loss despite positive selection (Hamilton 2009). Particularly vulnerable are quantitative traits, in which a phenotypic trait is influenced by multiple loci such that the selection exerted upon the phenotype is divided among the loci, resulting in weak selection on any given locus. As the number of
loci contributing to a trait increases, each individual locus tends to experience less selective pressure (Hamilton 2009).

**Meta-population Dynamics**

Island biogeography theory was developed to mechanistically explain how oceanic island size and isolation affects species richness (MacArthur 1967). Since then, it has become a useful model to explain meta-populations in fragmented habitats, in which habitat patches are disbursed throughout a matrix. Many meta-populations are dispersed throughout a matrix that is heterogeneous and somewhat traversable (Laurance 2008). Additionally, many natural systems, including the salt marsh, represent a finite island model.

In meta-populations, variation may be lost as sub-populations go extinct, leaving uninhabited patches. Often, patches are recolonized by immigrating individuals. When few individuals contribute to the new population, a founder effect occurs in which the new population represents only a sample of the variation present in the ancestral population, and therefore exhibits lower allelic richness. Following the initial colonization, reproduction is also associated with very strong drift due to the small population size. Rare alleles may be lost via drift, further reducing the allelic richness (Hamilton 2009).

Similar to a founder effect, a bottleneck results in a loss of genetic variation via high mortality. Similarly, the small number of surviving individuals often represents only a sample of the variation present in the pre-bottleneck generations. For populations that experience seasonal bottlenecks, $N_e$ can be estimated using the harmonic mean of the
census population size over time, $N_e = t / \Sigma(1/N_t)$ (Luikart et al. 1999, Hamilton 2009).

In meta-populations, low variation associated with founder and bottleneck events may be ameliorated by gene flow from neighboring sub-populations.

Isolation and low connectivity between sub-populations can result in differentiation primarily via drift, selection, and mutation. Over short timescales, variation may be maintained across the total population as allele frequency differences among sub-populations as different alleles fix in different sub-populations; however, drift will tend to reduce genetic variation within the sub-populations. Gene flow between sub-populations tends to homogenize allele frequencies. However, differential gene flow rates within a network can exacerbate pair-wise differentiation between sub-populations with low connectivity to each other relative to other patches, particularly if they are recently fragmented or isolated from each other. Differentiation is often assessed using estimates of the fixation index, $F_{ST}$ (Hamilton 2009). $F_{ST}$ compares the average Hardy-Weinberg expected heterozygosity ($H_E$) across subpopulations to the expected $H_E$ of the total population if it were panmictic. $F_{ST}$ is also useful for inferring the interplay between drift and gene flow. Using genetic data from two sub-population, $F_{ST}$ can be estimated and used to approximate the composite parameter $N_c m$, the effective migration rate, using the equation $F_{ST} = 1/(4N_c m + 1)$. Although this parameter is based on an infinite island model, whereas most natural populations exhibit a finite island model of population structure, $N_c m$ provides a good approximation. Large values of $N_c m$ indicate weak drift due to either a large $N_e$ or high gene flow. As $N_c m$ becomes smaller, drift becomes stronger. At values of $N_c m$ less than 1, drift is expected to lead to the differentiation of sub-populations (Hamilton, 2009).
Gene flow may come from elsewhere within the patch in the case patches that are large relative to dispersal, or from other habitat patches. Between patches, the isolation and connectivity of the patches will play a role in determining gene flow. The matrix may form a barrier preventing gene flow, or it may be somewhat hospitable to the organism, allowing some gene flow. The matrix may also be spatiotemporally heterogeneous, leading to variable rates of pairwise gene flow over space and time. Gene flow can increase fitness by reintroducing beneficial alleles that have been lost via drift (Alleaume-Benharira et al. 2006). In contrast, gene flow can have negative consequences by reintroducing deleterious alleles that have already been purged or causing outbreeding depression via admixture between demes that are locally adapted (Marsden et al. 2013).

**Isolation-by-distance**

Estimates made using the gametic disequilibrium method assume a continuous, panmictic population. However, when the habitat area is large relative to the mean dispersal distance, non-random mating occurs in which individuals are more likely to mate with nearby individuals, resulting in isolation-by-distance (IBD) and the formation of genetic neighborhoods (Wright 1943, 1946, Neel et al. 2013). In the presence of IBD, individuals will be more likely to mate with nearby individuals than distant ones. Over time, this may lead to heterogeneity throughout the population as individuals will tend to be more related to individuals within the vicinity (Wright 1943). In this case, any given estimate of \( N_c \) only indicates the effective number of breeders over the area of one genetic neighborhood (Neel et al. 2013).
In fragmented populations, the influence of edge effects likely differs between taxa based on how mobility influences IBD and genetic neighborhood dynamics. For example, within small habitat patches, mammals may exhibit spatially variable density at any given point in time simply based on random movement, while effectively existing as a single panmictic population with a spatially constant $N_e$ (Olson and Andow 2008). Essentially, the size of the patch would be equal to or smaller than the area typically traversed by a single individual. In this case, genetic data taken by a researcher at any given site within the patch would simply be a sample of the panmictic population and the population would behave as a single evolutionary unit. In insects, plants, and small vertebrates with low dispersal, there may be genetic neighborhoods forming as a result of IBD (Wright 1943). Essentially, the area of the patch is larger than the area typically traversed by a single individual within its lifetime. This area may be quantified using the organism’s dispersal, which is defined as the distance between the birth place of the parents and that of the offspring. The ‘neighborhood’ of a single individual is therefore equal to $4\pi\sigma^2 D$, in which $\sigma$ mean squared dispersal distance and $D$ is the census density (Neel et al. 2013).

Isolation-by-distance may also be described as the result of decreasing gene flow with increasing distance. In fragmented habitats, lower gene flow between edge and interior populations may enable differentiation, particularly if selective pressures or degree of reproduction isolation differ between edges and interiors. Selective pressures may exist due to differences in community composition or abiotic factors, or due to intraspecific differences that result in indirect ecological genetic effects. For example, genotypes of an herbivore may be differentially suited to host plant genotypes, leading to
spatial correlation between host and herbivore genotypes. In this research, we aim to investigate if differences in host plant genotypes or allelic variation between patch interiors and edges may correlate with differences in those of a dependent herbivore, providing empirical evidence of ecological genetic effects.

**The Core-Area Model**

By amplifying the effects of small patch sizes, edge effects reduce the effective area of a patch. This is most simply quantified by measuring the core area of the habitat, as defined by a certain distance from an edge (Laurance and Yensen 1991, Didham and Ewers 2012). More recently, the effective area model has expanded upon the core area concept by including species-specific responses, and GIS has advanced to provide useful tools for analysis (Sisk et al. 1997, Didham and Ewers 2012). There are numerous proximate reasons for edge effects, but ultimately, they tend to experience greater environmental extremes and higher variability than interior habitats (Laurance et al. 2007). Edges are also the interface between the patch and matrix habitats, and small differences in initial matrix conditions can set patches on divergent trajectories of change resulting from different interactions with abiotic and biotic matrix conditions. Spatiotemporal variability in edge effects can also be influenced by the presence of nearby fragment edges and with-in patch distance to the edge (Laurance et al. 2007).

**Model System**

The study system is located in the Great Bay Boulevard Wildlife Management Area in Tuckerton, New Jersey. This salt marsh covers approximately 2,000 hectares
Spartina patens, commonly called marsh hay, is a cordgrass that dominates in the high marsh. It is dispersed throughout a matrix of primarily S. alterniflora with areas of mudflats, salt pannes, and salt water ponds, forming a naturally fragmented system. Spartina alterniflora dominates in areas that are inundated with salt water regularly, whereas the less salt tolerant S. patens is the competitive dominant in more elevated areas that are inundated only during the highest tides. For low-mobility species that are restricted to S. patens patches, this naturally presents a finite island model of spatial population structure. The heterogeneity of the matrix also provides the opportunity to investigate the effects of matrix type and its influence on both connectivity and edge effects. The dynamics within this system can be generalized to plant-herbivore genetic effects in many heterogeneous environments.

A recent study indicates that large and small S. patens patches may be genetically differentiated from each other, even when highly connected, and that individuals in small patches may be less competitive than those in large patches. Individuals in large patches may also produce smaller genets, which are clumps of identical clones, than those in small patches (Holzapfel 2011). There is evidence that genotypes may differentiate over short distances, and it is possible that differences in abiotic conditions, such as salinity, may cause differences at a microclimate scale within patches as well (Silander and Antonovics 1979), particularly if edges experience higher salt and inundation stress.

Tumidagena minuta is a predominantly flightless planthopper species that lives within patches of Spartina patens. Individuals grow and reproduce within the same patch, have limited dispersal, and are likely to be affected by changes in their host plant (Robert F. Denno 1994, Peterson et al. 2001). Each patch supports a single T. minuta sub-
population within a meta-population, with the highest densities occurring in the patch interior and the lowest near the edges (Wimp et al. 2011). *Tumidagena minuta* experience seasonal fluctuations as a result of high winter mortality, and their effective population sizes are likely representative of the bottleneck they experience during the overwintering generation. They may also experience seasonal extinction and recolonization events resulting from intense winters. These characteristics make them ideal model organisms for studying habitat effects on short time scales.

Due to their sedentary nature (Robert F. Denno 1994), *T. minuta* are also ideal model organisms to study the effects of edges over small spatial scales. If there are genetic neighborhoods, then density differences may cause spatial heterogeneity in the relative strength of the four evolutionary processes, detectable via different $N_e$ estimates. In *T. minuta*, this enables the use of very small spatial and temporal scales to predict what is occurring at larger scales within longer-lived populations that would require much greater amounts of time and resources to study. Across species, the findings are likely generalizable at the scale of the area of a genetic neighborhood and the time period of one generation.

There are differences in morphological and physiological trait expression by *S. patens* individuals in large patches versus those in small patches (Holzapfel 2011). These differences have been attributed to genotype, but it is possible that they are also epigenetic in nature. If no genotypic differences are found between large and small patches using neutral genetic markers, then the physiological and morphological differences previously observed may be attributable to epigenetic modifications. If there are predictable differences in *T. minuta* genetic variation that correlates with patch size,
then this may indicate the presence of epigenetic indirect ecological effects, which would warrant further investigation.

The use of a salt marsh model system provides the additional benefit of generating system-specific knowledge that can aid in the conservation and management of this important ecosystem. Salt marshes are important coastal habitats that provide numerous ecosystem functions and services. Ecosystem functions are ecological processes that affect energy flow through the environment or the cycling of water, nutrients, and organic matter. Ecosystem services are the benefits that these functions provide to society (Cardinale et al. 2012). Salt marshes are particularly important for providing nutrient buffers that absorb pollution, such as nitrogen, from runoff before it enters the ocean, and acting as storm buffers that protect inland areas from storm surges and flooding. Marshes also provide critical nursery habitat for fishery species and provide recreational opportunities for fisherman and outdoor enthusiasts.

Research continues to show the importance of intraspecific variation in maintaining ecosystem functions and services (Grady et al. 2011, Cardinale et al. 2012). Biodiversity of genes, species and functional groups maintains the efficiency of primary production and nutrient cycling (Cardinale et al. 2012) and increases stability of ecosystem functions (Cottingham et al. 2001). Understanding salt marshes is especially important given the number of threats they face today, including sea level rise, coastal development, pollution, and invasion by the European reed Phragmites (Holzapfel 2011). Protecting the structural integrity of the system is the primary goal of many coastal restoration projects, and increasing intraspecific diversity will likely allow compensatory dynamics that may be adequate to sustain enough below-ground biomass for soil
stability. Increasing the structural resilience of coastal wetlands may save millions of dollars in future restoration efforts and coastal storm damages. Ecological restoration is increasingly important given the rapid rate of global change and human alteration of the environment. As ecological genetics theory continues to develop, and the underlying mechanisms controlling variation are revealed, this knowledge can be applied to restoring and maintaining important ecosystems.

**RESEARCH SIGNIFICANCE**

Habitat fragmentation is a major threat and management challenge confronting conservation biologists, making it critical that we understand the interplay of changes associated with habitat fragmentation, such as altered patch sizes and characteristics, connectivity, and matrix characteristics. Additionally, many species of conservation concern are habitat specialists that are especially sensitive to loss and fragmentation, and show negative responses to edges (Lande 1987, Ries et al. 2004, Martinson and Fagan 2014). It is important to understand the role edges play as a mechanism underlying the effects of habitat fragmentation on these populations. Moreover, most studies investigating the genetic impact of edges on populations are at range-wide scales, which confounds historical and contemporary processes. To understand the effect of anthropogenic habitat fragmentation requires investigation of the contemporary processes influencing populations within and among patch dynamics.

This research will also take a novel approach by investigating fragmentation effects through an ecological genetics framework. Rarely has the effect of genetic
diversity on species interactions been studied, and it is unknown how genetic diversity within a dominant plant species may impact meta-population dynamics at higher trophic levels. As many habitats are experiencing rapid change and alteration, it is critical that we understand this interplay between trophic levels. By including genetic diversity effects, such as interspecific indirect genetic effects (IIGE), this research has the potential to advance meta-population theory and improve its predictive power. Therefore, this research will both contribute to the field of ecological genetics and have immediate application in conservation biology and coastal wetland management.

The tools produced in this study will benefit the population genetics community. The microsatellite markers developed will be immediately disseminated via a published primer note, providing genetic markers for the community of scientists who study salt marshes. The markers may also have use in other Poaceae species, allowing researchers studying closely related species to potentially use these markers as well. The program implementing the methods of Sved (2013) will also be made available, providing an alternative method for dealing with null alleles and other sources of within locus disequilibrium. The development of microsatellite markers and a new program to estimate $N_e$ are tools that will benefit the population genetics community.
CHAPTER 2: Development of $N_e$ Estimation Program

The effective population size ($N_e$) is the size of a Wright-Fisher population that would experience the same amount of drift as empirically observed (Hamilton 2009). The Wright-Fisher model relies upon numerous underlying assumptions that eliminate most sources of disequilibrium, and therefore, the disequilibrium present in a Wright-Fisher population is assumed to occur entirely via drift. Genetic drift is associated with allele frequency fluctuations between generations, the magnitude of which is expected to increase as $N_e$ decreases. At small $N_e$ sizes, drift may overwhelm other evolutionary forces, such as natural selection, and lead to a loss of genetic variation as alleles randomly go to fixation or loss. For the analysis of contemporary processes, $N_e$ is critical for assessing the relative strengths of the evolutionary processes and is an invaluable tool for conservation managers.

$N_e$ may be estimated using gametic disequilibrium, which is the non-random association of alleles at two loci. Estimates of the effective population size ($N_e$) using gametic disequilibrium represent the instantaneous strength of genetic drift in a Wright-Fisher model, thereby facilitating the distinction between instantaneous and historical processes (Rhodes and Chesser 1994, Wang et al. 2011). More specifically, genetic association between unlinked loci tends to decay over time, and based on a recombination rate of 0.5, is predicted to approach zero after approximately seven generations. This means that any disequilibrium detected is assumed to have been initiated within that time period (Hamilton 2009). Compared to the census size, the $N_e$ more accurately represents
the behavior of a population in terms of the action of the various evolutionary processes (Wang et al. 2011).

Sved et al. (2013) introduced a new method of estimating effective population size (Ne), specifically a composite haplotype table that replaces other methods of inferring haplotypes from genotype data and a permutation correction factor that accounts for null alleles. The $r^2_{\text{permute}}$ correction factor simultaneously accounts for the within-locus disequilibrium caused by null alleles and the finite sample size. Additionally, Sved et al. (2013) suggested using a weighted average of $r^2_{ij}$ values to estimate $r^2_{lm}$ as an alternative to using a minimum allele frequency. I collaborated with Robert Battocletti to implement these methods in Python. The program produces estimates of Ne using both $r^2_{\text{comp}}$ and $r^2_{\text{delta}}$, and the output file contains a set of results for each. Ne estimates are accompanied by both delete one-locus jack-knife parametric and percentile confidence intervals. The $r^2_{\text{permute}}$ percentile confidence interval and the associated Ne values are also provided. Additionally, allelic richness, expected heterozygosity, and observed heterozygosity are produced for each input population. Tables containing intermediate results are provided to help the user evaluate deviations from underlying assumptions. The program can be downloaded from https://github.com/GeneHub/PYNE.

**Rationale**

Effective population size can be estimated using gametic disequilibrium; however, the presence of null alleles can lead to the underestimation of Ne via an increase in within-locus disequilibrium that increases the $r^2$ estimate. Sved et al. (2013) introduced a novel way of correcting for this bias using a new $r^2_{\text{permute}}$ correction factor
that quantifies the portion of $r^2$ attributable to both within-locus disequilibrium and finite sample size. Single-locus genotypes are permuted amongst individuals, maintaining allele and single-locus genotype frequencies constant while randomizing the association between loci. Over a large number of replicates, the average association between loci tends towards zero, theoretically leaving only the disequilibrium caused by within-locus disequilibrium and finite sample size.

**Data Import**

The program accepts data in the standard 2 and 3-digit allele genepop formats, and a single-population format that can be saved directly from Microsoft Excel as a csv document. The program will automatically recognize the input format. Data files must be placed into a designated input data folder that is automatically generated by the program. When opening the GIU, the input data folder location will be stated at the top of the page. The program will process all files located in this folder.

Each file processed will generate its own results folder, which will be named using the time that processing began. It is important to note that genepop format files may contain multiple populations, and the results for multiple populations within the same file will output to the same results folder. In this case, the user must provide a unique name to each population. If two populations within the same document have the same name, then the second population processed will overwrite the results for the first. For example, if two populations within the same genepop file are named “pop2”, then there will only be one result for “pop2” in the folder, and it will be for the most recent population “pop2” that was processed.
The programs recognize both the 2-digit and 3-digit allele identity Genepop formats, and can process files containing multiple populations. Genepop files need to be saved as .txt files for import into the program. A full explanation of the Genepop format can be found at http://genepop.curtin.edu.au/help_input.html. In general, the first line contains information identifying the dataset and can be anything the user wishes. Next, each following row should contain the name of a single locus, in the same order as they appear in the dataset. For example, the locus named in row 2 should correspond to the two alleles provided in columns 2 and 3, and so on. After the locus names, there should be a row that simply states “POP” to mark the beginning of population data. If the file contains data for multiple populations, then the populations just need to be separated by additional rows that state “POP”. The program simply recognizes the first column as the population name, so the comma following the name is not required.

The CSV file format is simpler than the Genepop format, however, a single file is limited to one population only. Each row represents an individual and each column represents a single allele. There is no way to input the name of the locus, but the program automatically combines alleles from the same locus and assigns numerical names to the loci beginning with 0. It is assumed that each pair of adjacent columns represent a single locus, for example, columns 1 and 2 represent locus 0, columns 3 and 4 represent locus 1, and so on.

**Output**

The output files will be placed in a results folder that is automatically generated when the app is launched. The location of this folder will be indicated at the top of the
GUI. There will be a text file with the Ne estimates, associated confidence intervals, and a table of locus-pair comparison values. Additionally, there will be a csv file that contains haplotype data for each pair-wise comparison. If the input data was in csv format, then the population name will simply be “dataset” because this format only permits a single population to be entered at a time. For populations input using the popgene format, the population names will be used as the name of the datafile since this format allows for multiple populations to be entered simultaneously.

METHODS

**Symbols used in the following equations**

Ne  
Indicates the first locus in a locus pair combination

\( l \)  
Indicates the second locus in a locus pair combination

\( i \)  
In a haplotype, indicates the allele at locus \( l \)

\( j \)  
In a haplotype, indicates the allele at locus \( m \)

\( D_{ij} \)  
Disequilibrium coefficient for the haplotype \( i,j \)

\( r_{lm}^2 \)  
“Within locus pair” correlation coefficient for the locus pair \( l,m \)

\( r^2 \)  
“Between locus pair” correlation coefficient; a weighted average of all \( r_{lm}^2 \) values

\( S \)  
Sample size, may be the input sample size, harmonic mean of \( S_{lm} \), or the mean of \( S_{lm} \)

\( S_{lm} \)  
Sample size for the locus pair \( l,m \)

**Composite Haplotype Table**

To estimate Ne from diploid genotype data using the linkage disequilibrium method, the genotypes must be converted to bigenic haplotypes. The program estimates haplotype frequencies by using a ‘composite haplotype table’ that is produced by converting the di-locus genotype for an individual, Aa Bb, to the four potential
haplotypes AB, Ab, aB, and ab. As a result, the composite haplotype will contain a number of haplotypes equal to 4S, given that there is no missing data. A table is produced for each unique pair-wise locus comparison (Sved et al. 2013).

The general equation for estimating allele frequencies, in which the number of individuals contributing haplotypes to the table is Sc, is as follows:

\[ f(a) = \frac{\text{count of } a}{2Sc} \]

Allele frequencies then yield the expected haplotype frequencies, estimated as:

\[ f(a,b) = f(a) \times f(b) \]

**Calculation of \( D_{ij} \)**

It is important to note that the allele frequencies are calculated independently for each pair-wise comparison from the individuals used to calculate \( D_{ij} \). Ideally, in a complete dataset with no missing data, the allele frequencies will be identical for a particular locus across pair-wise comparisons. If the dataset has missing genotypes; however, different allele frequencies may be used for the same alleles for different pairwise locus comparisons since individuals missing one of the genotypes are removed.

\( D_{ij} \) values for all haplotypes are estimated using the following equation in which \( M \) represents the count of haplotype i,j:

\[ D_{ij} = \frac{M}{4S} - \left( \frac{2n_{a}}{4S} \cdot \frac{2n_{b}}{4S} \right) \]  

(Sved et al. 2013)
In the program, 4S is not calculated directly as four times the sample size. To account for missing data, which will cause a deviation of the total count from 4S, it is simply the total count of all haplotypes.

**Missing Data**

For each pair-wise comparison, individuals missing entire genotypes are removed prior to populating the composite haplotype table. Since the composite haplotype table is populated using gametic haplotype combinations, we can introduce a new option that gives users the ability to retain individuals with one missing allele. This allows users to code potentially erroneous alleles as missing data while maintaining the two haplotypes produced by the unambiguous alleles. For example, in the case of genotype A0Bb, the haplotypes A,B and A,b will be added to the table. If there is uncertainty as to whether a product is a true allele or an artifact, then the user can now code the potentially erroneous alleles as missing data while retaining the 2 reliable haplotypes in the dataset. Retaining individuals with one known allele can improve the Ne estimate, particularly when S<<Ne, because the removal of all four potential haplotypes from the individual would alter the D values of the remaining haplotypes by differentially affecting the allele and haplotype frequencies.

**Minimum Threshold**

The user can choose to set a minimum allele frequency threshold in which only the D_{ij} values for haplotypes containing both alleles at or above the user defined minimum allele frequency are used to estimate the locus-pair r^2_{lm}. This reduces the bias
towards larger $r^2$ values introduced by rare alleles. The value of the minimum allele frequency threshold is input by the user.

**Within-locus pair correlation**

Each $D_{ij}$ value is converted to both an $r^2_{ij\text{(comp)}}$ and an $r^2_{ij\text{(delta)}}$ value. It is important to note that $D$ is not adjusted using Wier’s unbiased estimator, which may contribute to differences between these and other methods (Waples and Do, 2008). We use Equation 7 from Sved (2013); however, since we include a minimum frequency, we slightly modified the equation to normalize the weighting factor, $p_iq_j$, by the total $\Sigma p_iq_j$.

\[
\begin{align*}
    r^2_{ij\text{(comp)}} &= \frac{4D_{ij\text{(comp)}}^2}{p_i(1-p_i)q_j(1-q_j)} \\
    r^2_{ij\text{(delta)}} &= \frac{4D_{ij\text{(comp)}}^2}{[p_a(1-p_a)+(p_{aa}-p_a^2)][p_b(1-p_b)+(p_{bb}-p_b^2)]}
\end{align*}
\]

(Sved et al. 2013)

We apply the $r^2_{lm}$ equation to both the $r^2_{\text{(comp)}}$ and an $r^2_{\text{(delta)}}$, producing separate results for each.

\[
\begin{align*}
    r^2_{lm} &= \sum_{i=1}^{k_l} \sum_{j=1}^{k_m} \left[ r^2_{ij} \cdot \frac{p_iq_j}{\Sigma p_iq_j} \right]
\end{align*}
\]

(Sved et al. 2013)

The locus-pair correlation coefficient, $r^2_{lm}$, is a weighted average of the $r_{ij}$ values for haplotypes containing alleles above the minimum threshold. As suggested by Sved et al. (2013), each value is weighted by the allele frequencies, which is an alternative method to the minimum threshold for correcting the disproportionate effect of rare alleles. However, we also include a minimum allele frequency option. To account for the inclusion of this option, we slightly modified the equation to normalize the weighting.
factor, $p_iq_i$, by the total $\sum p_iq_i$. This allows the threshold and weighting to be used simultaneously. To further clarify, in Sved (2013), Equation 7, he weights the D values by $p_iq_i$, without normalization, because he assumes all potential haplotypes are used to estimate $N_e$. If all haplotypes are used, then $p_iq_i$ sums to one and the normalization would therefore require the division of $p_iq_i$ by one, which simply equals $p_iq_i$. In our program, however, we allow haplotypes to be excluded if they contain an allele that is below the minimum threshold. This means that when the user sets a minimum allele frequency, the sum of $p_iq_i$ for the haplotypes used will likely be less than one, and therefore, do need to be divided by the sum of $p_iq_i$ for normalization.

**Between locus-pair correlation**

The overall $r^2$ value is calculated as the weighted average of each pair-wise $r^2_{lm}$. For $L$ loci, there will $L(L-1)/2$ numbers of pair-wise comparison, each yielding values of $r^2_{lm}$. These values are weighted using the allelic richness at each locus and the $S_{lm}$ value, which is the effective ample size for a particular comparison, calculated as:

$$S_{lm} = \frac{(#\text{haplotypes in table})}{4}$$

The $r^2$ is calculated using equation 6 from Sved et al (2013):

$$r^2 = \sum_{l=1}^{L-1} \sum_{m=l+1}^{L} \left[ r^2_{lm} \cdot \frac{(s^2_{lm}(k_l-1)(k_m-1))}{\sum_{l=1}^{L-1} \sum_{m=l+1}^{L} (s^2_{lm}(k_l-1)(k_m-1))} \right]$$  
(Sved et al. 2013)

The weighting factor $S^2 (k_l -1)(k_m -1)$, in which $k_l$ and $k_m$ are the allelic richness at locus $l$ and $m$, was introduced by Waples and Do (2008), and gives more weight to loci that have higher allelic richness. Sved recommended replacing $S$ with $S_{lm}$, which differs
based on missing data (Sved et al. 2013). Weighting the values using the $S_{lm}$ allows loci with little or no missing data to weigh more influentially than loci with a relatively large amount of missing data.

**$N_e$ estimation**

Multiple $N_e$ estimates are produced, providing the user the option to choose the most appropriate method for their data. $N_e$ is estimated using both $r^2_{comp}$ and $r^2_{delta}$, and multiple correction factors. The general equation is:

$$N_e = \frac{1}{3(r^2 - \text{correction factor})}$$

When using a sample size correction factor, the equation becomes:

$$N_e = \frac{1}{3 \left( r^2 - \frac{1}{5} \left[ 1 - \left( \frac{1}{2S - 1} \right)^2 \right] \right)}$$

When using the permutation correction factor, the equation becomes:

$$N_e = \frac{1}{3 \left( r^2 - r^2_{\text{permute}} \right)}$$

**Sample-size Correction Factors**

The user may select from three sample size options for use, each of which produces a unique $N_e$ estimate that is provided in the output. The user may correct for sample size using either the input sample size ($S$), the mean $S_{lm}$, or the harmonic mean of
$S_{l,m}$. The simplest method is to use the number of individuals in the input file, termed the input sample size; however, this method does not take into account missing data. The other two methods use $S_{lm}$, which is the number of individuals contributing to the composite haplotype table for locus pair $l,m$. $S_{lm}$ is the sample size for a particular pairwise locus comparison, and will differ if there is missing or removed data. Ideally, if there is no missing data and the minimum threshold is zero, then each $S_{lm}$ will equal the input sample size. Missing data and rare alleles or high thresholds will decrease the mean $S_{lm}$ relative to the input sample size. $S_{lm}$ is calculated for each pairwise comparison by dividing the number of haplotypes in the final haplotype table by four.

$S_{lm} = (\# \text{ haplotypes in table})/4$

The sample size correction factor may use either the mean or the harmonic mean of $S_{lm}$.

Mean $S_{lm} = \frac{\sum S_{lm}}{L(L - 1)/2}$

Harmonic Mean $S_{lm} = \frac{\sum \frac{1}{S_{lm}}}{\frac{L(L-1)}{2}}$

It is important to note that when using the permutation correction factor, the sample size correction factor cancels out and is therefore omitted. In the results output, the Ne estimation will be given for each of the 3 sample size correction factors and the permutation correction factor separately. The equation for $r_{\text{delta}}^2$ contains a within-locus disequilibrium correction factor that accounts for excess homozygosity (Waples and Do 2008). It is therefore appropriate to use the $r_{\text{delta}}^2 N_e$ estimates with a sample size correction factor. If used in tandem with the permutation correction factor, it may
overcorrect for within locus disequilibrium and exhibit a bias towards slightly inflated $N_e$ estimates. In contrast, the $r^2_{\text{comp}} N_e$ estimate is best used in tandem with the permutation correction factor since the $r^2_{\text{comp}}$ equation does not contain the within-locus disequilibrium correction factor within the denominator. Without the permutation correction factor, $r^2_{\text{comp}}$ will likely produce an underestimate of $N_e$.

**Permutation Correction Factor**

An important feature of Sved’s method is the implementation of an $r^2$ permute correction factor that accounts for with-in locus disequilibrium. Most notably, the correlation produced via the increased homozygote frequencies associated with null alleles is confounded with other factors causing with-in locus disequilibrium, and can be corrected for using this method. The genotypes are randomly permuted within each locus to produce a new dataset. The permutation is performed for a large number of replicates, specified by the user, and the mean $r^2$ from the resulting distribution is used as the correction factor. By permuting the genotypes within each locus, the association between loci becomes completely random. When the association between loci is random, gametic disequilibrium should contribute zero disequilibrium to the $r^2_{\text{permute}}$ value. This means that the resulting mean $r^2_{\text{permute}}$ is caused by only the disequilibrium created via the finite sample size and the within-locus disequilibrium. By permuting the within-locus genotypes, the portion of the correlation coefficient due to both the finite sample size and within-locus genotype structure is isolated, allowing for it to be subtracted from the correlation coefficient originally obtained using the dataset. Most notably, the correlation produced via the increased homozygote frequencies associated with null alleles is
confounded with other factors causing within-locus disequilibrium, and can be corrected for using this method. Assuming that there is no linkage disequilibrium or population sub-structure, this leaves the portion of the correlation coefficient assumed to be due to true gametic disequilibrium, providing a more accurate estimate of the true population parameter.

The mean $r^2$ value of the replicate distribution is used as the correction factor. Since this value captures the disequilibrium produced by the finite sample size, there is no need to use a sample size correction factor as the sample size correction would simply cancel itself out. The corrected $r^2$ value is then used to calculate $N_e$.

**Confidence Intervals**

Presently, the program contains $N_e$ confidence intervals using delete one locus-pair jackknife parametric and percentile methods (Hill 1981, Waples and Do 2008). Also produced are the percentile confidence intervals associated with the $r^2_{[permute]}$ correction factor and the corresponding $N_e$ estimates.

*Jackknife Confidence Intervals:* The jackknife distribution is generated by eliminating one pair-wise locus combination at a time. The percentile jackknife confidence interval uses the $\alpha/2$ and the $1-\alpha/2$ values from the distribution. As suggested by Waples and Do (2008) and implemented in LDNe, we also included a parametric confidence interval in which the number of degrees of freedom is equal to the effective number of comparisons. This is based on the equations presented by Hill (1981), in which $n$ represents the number of independent comparisons. The distribution created using the delete one locus-pair jackknife method is used to estimate the $\text{Var}(\hat{r}^2)$. 

31
\[
\Phi = \frac{\text{Var}(\bar{r}^2)}{(\bar{r}^2)^2}
\]

\[
n \approx \frac{2}{\Phi}
\]

\[
n = \sum n_{i,j}
\]

*Permutation Correction Factor Confidence Intervals:* The distribution of \(r^2_{c[permute]}\) values yields a percentile confidence interval for the \(r^2_{c[permute]}\) correction factor. The correction factor values at the lower and upper confidence limits are then used to produce the associated \(N_e\) estimates.
CHAPTER 3: Population Genetic Patterns in *Tumidagena minuta*

Anthropogenic habitat loss and fragmentation are major threats to biodiversity. The threat occurs not only at the species level, but also within species. In fragmented populations, intraspecific variation can be lost in the initial population decline and then via genetic drift thereafter. This loss can impact the long-term persistence of a population and reduce its ability to adapt to future changes in the environment. Furthermore, there is a lack knowledge regarding the effects of edges specifically on arthropod communities, which may present ideal opportunities for study due to their short generation times (Murphy et al. 2016).

Populations that become fragmented may be influenced by numerous landscape characteristics, such as the size and shape of the patches and their connectivity. To predict the outcome for a population requires an understanding of many of these factors. In this study, we aim to understand how edge effects may affect fragmented populations that exhibit a negative response to edges, meaning that their density decreases with increasing proximity to the edge (Ries et al. 2004). In these scenarios, edge effects reduce the effective area of a patch, leaving a core area of habitat unaltered by the edge that is smaller than the total area of the patch (Laurance and Yensen 1991, Didham and Ewers 2012). Many species of conservation concern exhibit negative responses to edges (Ries et al. 2004), which may exacerbate the loss of diversity in small fragments beyond what is predicted simply by area (Sisk et al. 1997).

When the habitat patch is large relative to the mean dispersal distance, non-random mating occurs in which individuals are more likely to mate with nearby individuals,
resulting in isolation-by-distance (IBD) and the formation of genetic neighborhoods (Wright 1943, 1946, Neel et al. 2013). As a result of this IBD, spatial heterogeneity in density may lead to a corresponding heterogeneity in the strength of genetic drift. In populations that experience lower densities in edge habitat, this may lead to stronger genetic drift near edges. In patches that are small relative to the neighborhood size, this may result in both a faster rate of variation loss than would occur without the negative edge response, and a diminished ability to respond to selection. In patches that are large relative to the neighborhood size, this may result in polymorphism being maintained by heterogeneity in allele frequencies across the entire patch area, but the population would likely still experience a similar diminished ability to respond to selection as drift would be the dominant force at any given location (Neel et al. 2013).

An increase in the strength of drift in edge habitats may be detected via a comparison of $N_e$ estimates (Hamilton 2009). $N_e$ estimates are also useful because they can allow a comparison of the strength of drift to other evolutionary forces, such as selection, and may reveal how patch characteristics affect the predicted rate of allelic loss. If the strength of selection is known, then the strength of drift relative to selection can be assessed using the composite parameter $4N_es$. When $4N_es$ is less than 1, then drift is predicted to be the dominant force, reducing a population’s ability to respond to selection (Hamilton 2009).

The action of genetic drift in small, isolated populations may be countered by gene flow from populations in neighboring patches. This influx of variation may increase fitness by reintroducing beneficial alleles that have been lost via drift (Alleaume-Benharira et al. 2006). Gene flow between patches is influenced by multiple factors, most
notably, the connectivity of a patch to neighboring patches. The characteristics of the matrix may are also important. The matrix may be inhospitable to the organism, forming a barrier that prevents gene flow, or it may be somewhat hospitable, allowing gene flow. In the absence of gene flow, it would be predicted that in small patches with low \( N_e \) and strong drift, measures of variation, such as allelic richness and heterozygosity would be lower than in larger patches. If gene flow is adequate, however, variation lost via drift may be readily restored via reproducing migrants.

In this study, we seek to investigate the relationship between landscape characteristics in a fragmented habitat and the action of genetic drift in a species exhibiting a negative edge response. The planthopper *Tumidagena minuta* presents the ideal model organism for the study of this phenomenon. It has been shown to exhibit a negative response to edges, resulting in lower densities in edge habitat relative to interior habitats (Wimp et al. 2011). We hypothesize that: 1) *T. minuta* density will be lower in edge habitats relative to interior habitats; 2) *T. minuta* effective population size (\( N_e \)) will be lower in edge habitats relative to interior habitats; 3) *T. minuta* allelic richness and heterozygosity will be lower in edge habitats relative to interior habitats; 4) *T. minuta* allelic richness and heterozygosity will be positively correlated with patch area; and 5) *T. minuta* allelic richness and heterozygosity will be positively correlated with connectivity.
METHODS

Field collection

Samples of *T. minuta* have been collected for numerous years by the Wimp Lab Group from *S. patens* patches located within the Great Bay Boulevard Wildlife Management Area in Tuckerton, New Jersey (39° 30.8'N, 74° 19.0'W). This system represents a naturally fragmented system and is ideal for studies of fragmented landscapes. We collected samples of *T. minuta* from Tuckerton, New Jersey, during the summers of 2014, 2015, and 2016, from 38 locations within 22 patches. We estimated the densities for locations across all collection dates to assess the effects of habitat fragmentation on density. To investigate the effects of fragmentation on genetic variation and *N_e*, we genotyped individuals from four transects that include an *S. alterniflora* edge, a water bordering edge, and an interior location.

The *Spartina patens* habitat patches sampled range in size from 13 m² to 2037 m². These are divided into two general categories based on total area: small patches are less 200 m² and large patches are 200 m² or greater. The largest “small patch” is 137 m², and the smallest “large patch” is 267 m². Patches can be very large and oddly shaped, so they are further divided into subsections that constitute relatively continuous area. Some patches, in particular, patches 1 and 2, contain a series of sub-sections that are connected by small “isthmuses” of *S. patens*. The smallest patch sub-section sampled was 267 m² since the smallest patch was symmetrical and constituted a single section.

We used a D-Vac suction sampler to vacuum arthropods into sampling nets placed around the mouth of the hose, which has an aperture diameter of 21 cm. We used
eight 5-second placements of the D-Vac, placed as two rows of four, centered about the site coordinates. The nets were immediately closed to prevent escape of the arthropods. After closing each net, we euthanized the arthropods by placing the nets into plastic jars containing ethanol vapors. After several hours passed to ensure euthanasia, the contents of each net were transferred into a single glass jar containing 100% ethanol for storage.

We used a systematic collection pattern to reduce variation that could arise as a combination of low *T. minuta* dispersal and inconsistent collection patterns. If the sampling window is too large, then individuals from different neighborhoods may be mixed in a single sample. This can cause a Wahlund Effect, which is a perceived deficit of heterozygotes resulting from differences in Hardy-Weinberg expected heterozygosities across neighborhoods or sub-populations (Hamilton 2009). If the sampling window is too small, then there may be an excess of heterozygotes (Neel et al. 2013).

**Density counts**

To determine the census density, the jar contents were poured into a tray and the *T. minuta* counted under a dissecting microscope. We performed counts for adult males, adult females, and juveniles. We used the combined male and female count to estimate census density. We divided the total adult count by the area of the D-Vac head multiplied by the number of plunks to estimate the census density per meter$^2$. A limitation to estimating density via this method is that the D-Vac likely does not collect all of the planthoppers, particularly those hiding deep within the thatch layer, and we do not have an estimate of what portion is typically collected. The collection may also be biased towards individuals that are nearer to the blade tops due to their life stage or
mating status. However, this also means that the method most likely does not defaunate the location, which would influence future estimates.

**DNA Extraction**

We extracted DNA from each adult *T. minuta* individually using a micropestel to homogenize the tissue, and a DNA extraction method that employs chelex as a chelating agent. A 10% chelex solution is prepared containing 200 µl of TE buffer and 0.02 g of chelex per sample to be extracted. For each sample, 200 µl of 10% chelex solution is added to a 1.5 or 2 µl tube and a single planthopper. The individual is homogenized using a micropestel, 2 µl of Proteinase K is added, and the sample is vortexed at medium speed for 20 seconds. The sample is next incubated at 55°C for 30 minutes, and vortexed at medium speed 1 to 2 times during the incubation. After incubation, the samples are vortexed at maximum speed for 20 seconds and then placed in a microcentrifuge at maximum speed for 20 seconds. Next, they are incubated at 100°C for 8 minutes, and then vortexed again at maximum speed for 20 seconds. The samples are then centrifuged at maximum speed (14,000 rpm) for 2 minutes. The top 100 µl of supernatant, which contains the genomic DNA, is then pipetted into a new tube.

We additionally applied a precipitation step to further remove contaminants from the samples. We placed 2X 100% ethanol into each tube and then placed them into the freezer to precipitate overnight. The next day, we carefully poured out the ethanol, making sure not to disturb the pellet of gDNA at the bottom. We then left the tubes under the hood overnight to allow the ethanol to completely evaporate. The next day, we added 100 µl of TE buffer to re-suspend the gDNA.
**Locus amplification**

Polymorphic microsatellite markers have already been developed for *T. minuta* (Aggarwal et al. 2011). We amplified the DNA for nine loci using fluorescently tagged primers. The amplifications were performed using two multiplex reactions, the first containing loci TmAAC057, TmAAC062, TmAAC095, TmAAC165, and TmAAC252, and the second containing loci DdAAC053, DdAAC206, TmAAC097, and TmAAC236. The reactions each contained 6.25 µl of either Qiagen Multiplex Mix 2X or Quanta Super Mix 2X, 2 µl of DI H₂O, 1.25 µl of primer set mix containing 10 µm of each forward and reverse primer, and 2 µl of gDNA. For the first set, we used the following PCR protocol: denaturation at 95°C (5 min), followed by 28 cycles at 95°C (30s) / 60°C (1min 30s) / 72°C (30s), and then a final extension at 72°C (5 min). For the second set, we used the following PCR protocol: denaturation at 95°C (5 min), followed by 28 cycles at 95°C (30s) / 55°C (1min 30s) / 72°C (30s), and then a final extension at 72°C (5 min).

**Genotyping**

Individuals are genotyped via fragment analysis using the ABI 3100 Genetic Analyzer and Geneious software. We both visually scored alleles and scored using the binning function in Geneious. The alleles were recorded using the number of repeats present. Locus DD053 was omitted from the final dataset due to persistent problems with reliable amplification and scoring.
Estimations

We estimated the allele frequencies, and expected and observed heterozygosities using a python program. To test for significant heterozygote deficiencies, we performed Hardy-Weinberg exact tests using the Markov Chain method implemented in Genepop (1000 dememorizations, 100 batches, 1000 iterations per batch) (Rousset 2008). We tested for a significant difference between edge and interior allelic richness for each patch separately using paired t-tests.

We used multi-locus genotype data to estimate $N_e$ using two distinct gametic disequilibrium methods. The first is an $r_{\text{comp}}$ based method that implements a permutation correction factor specifically designed to account for the within locus disequilibrium caused by null alleles (Sved et al. 2013). This method was implemented using a Python program developed for this project, described in Chapter 2. The second is an $r_{\text{delta}}$ method that implements a minimum allele frequency of 0.05 to reduce the disproportionate effect of low frequency alleles on the linkage disequilibrium estimate, and a harmonic mean sample size correction factor (Waples and Do 2008). This method was implemented using NeEstimator, version 2 (Do et al. 2014).

To determine if a significant difference exists between the patch edge and interior densities, we performed a one-tailed paired t-test. To determine if a correlation exists between the observed allelic richness, heterozygosity, or $N_e$ and patch area or connectivity, we used simple linear regression. Connectivity refers to the spatial connectedness a focal patch has to other patches, and is typically calculated using both the distance to and area of the neighboring patches. We estimated connectivity using the incidence functional measure, $S_i = \sum_j A_j \exp (-\alpha d_{ij})$, in which $A$ is the area of the
neighboring patch \( j \), \( \alpha \) is the inverse of the mean dispersal distance, and \( d \) is the distance between the focal patch \( i \) and patch \( j \) (Martinson et al. 2012, Ziółkowska et al. 2014). For \( T. \) minuta, we used the \( \alpha \) value of 0.1 suggested by Martinson and Fagan (2014).

RESULTS

**Densities**

For patches with both edge and interior samples, densities in \( S. \) alterniflora edge habitats were significantly lower than densities in interior habitats \((p=0.002)\). However, densities in water-bordering edge habitats were not significantly different than in interior habitats \((p=0.188)\). To test for correlations between fragment density and patch characteristics, Frag 28 was omitted because it is an outlier that is located much further southeast on the marsh and likely experiences greater storm and tidal disruptions, and altered ecological interactions. In 2014, both small and large patches exhibited a positive relationship between area and adult density (Figures 2 and 3). In small patches, density was more strongly correlated with the total area (Figure 3), whereas in large patches, it is more strongly correlated with core area (Figure 2), indicating that patch geometry may become more relevant with increasing patch sizes. In 2015, the densities showed a very weak relationship to area across both small and large patches (Figures 2 and 3). This may be due to high mortality over the 2014-2015 winter, particularly due to the unusually cold temperatures in February (NOAA 2017).
**Polymorphism**

There was not a significant difference between interior and edge allelic richness for samples collected in 2011 (Patch 2A, p = 0.407; 2C, p = 0.441; Jose 1, p = 0.769), whereas the samples collected in 2015 did show a significant difference (Patch 1D, p = 0.003; 2A, p = 0.040; Jeremiah 2, p = 0.002). Estimates of interior allelic richness and expected heterozygosity were highly consistent across patches. Although the sample size is too small to confidently draw conclusions, the results suggest interior allelic richness and expected heterozygosity were not strongly influenced by connectivity (Figure 4) or area (Figure 5). We expect this relationship to differ for smaller patches, which may contain less within-patch gene flow and therefore rely more heavily upon gene flow from neighboring patches, leading to an increase in allelic richness and heterozygosity with increasing connectivity.

**Effective Population Size**

$N_e$ estimates were obtained for patches 2A, 2C, and Jose 1, collection date June 30, 2011; and patches 1D, 2A, and Jeremiah 2, collection dates July 14 and 17, 2015. Hardy-Weinberg exact tests performed for each population at each locus using the Markov Chain method implemented in Genepop (1000 dememorizations, 100 batches, 1000 iterations per batch) indicate that every locus exhibited a significant heterozygote deficit at one or more locations (Table 2). Although it is possible there is a biological cause, such as non-random mating, the results strongly suggests the presence of null alleles. Null alleles occur when an allele systematically evades detection and may result from a number of causes, such as a mutation in the flanking primer region that prevents
proper binding (Chapuis M.P. 2007). Size homoplasy can also introduce a systematic increase in observed homozygosity as separate alleles are mistaken for the same allele (Queloz V. 2010). If null alleles or homoplasy are present, then the \( r_{\text{comp}} \) method using the permutation correction factor should correct for the excess disequilibrium (Sved et al. 2013). Alternatively, an excess of homozygotes may also indicate that we used a sampling window that was too large, generating a Wahlund Effect, and that the field sampling strategy may need to be altered to account for very small genetic neighborhood sizes (Neel et al. 2013).

For each patch, the \( T. \ minuta \) \( N_e \) estimates for the edge populations were consistently smaller than the corresponding interior population (Tables 1). The finding that \( T. \ minuta \) exhibit variable \( N_e \) across small spatial scales within the same patch indicates that \( T. \ minuta \) exhibit sufficiently low dispersal that individuals breed within or near their natal site, leading to the formation of genetic neighborhoods via IBD (Wright 1943). The difference between edge and interior \( N_e \) is likely the result of small genetic neighborhood sizes coupled with differences in density. Numerous estimates, particularly the upper confidence interval limits, were infinite. Infinite estimates occur when the sample size is too small relative to the true \( N_e \), leading to a correction factor that is larger than the observed disequilibrium. This generally indicates that a larger sample size is necessary.
DISCUSSION

The density and $N_e$ results suggest that populations with lower densities in edge habitats will experience stronger genetic drift in edge habitats relative to interior habitats. The lack of a relationship between allelic richness or heterozygosity and area or connectivity suggests that gene flow between neighborhoods within the same patch is sufficient to maintain diversity in large patches.

Lower edge density will likely result in a faster loss of variation via drift in small patches that are relatively equivalent to the size of single or few genetic neighborhoods. As a result, *T. minuta* populations within small patches may exhibit lower allelic richness and heterozygosity than those within larger patches, with the exception of small patches that have high connectivity and therefore may experience high levels of gene flow. In large patches, polymorphism will likely be maintained over longer periods of time as drift acts on alleles differently in different sections of the patch. The magnitude of this effect will likely depend on the patch size relative to the neighborhood size (Neel et al. 2013). In either case, lower densities in edge habitats will most likely diminish the populations’ ability to respond to selection, as drift will be the dominant force.

The mechanisms underlying these differences between edge and interior habitats are unknown. Since *T. minuta* is a specialist herbivore that relies upon *S. patens* for food and habitat, it is possible that bottom-up forces may be influencing its densities. Ecological genetics theory predicts that gene expression in one individual can act as an environmental influence on another individual. When this occurs between individuals of different species, it is called an inter-specific indirect genetic effect (Whitham et al.)
Greater intraspecific variation is predicted to produce a greater diversity of interspecific indirect genetic effects influencing the community (Whitham et al. 2006), and there is evidence that the intraspecific variation in plants can influence diversity at higher trophic levels (Bangert et al. 2006).

In addition to bottom-up forces, *T. minuta* densities may be suppressed by top-down forces. In particular, the wolf spider, *Pardosa littoralis*, occupies edge habitat bordering *S. alterniflora* at higher densities than interior habitat (Wimp et al. 2011) and inhabit both the thatch and live blades extending above the thatch. Adults tend to travel about the blades and consume prey predominantly from the grazing food web, while *Pardosa* juveniles tend to stay in the thatch to avoid intraguild cannibalism, and consequently, primarily feed on the epigeic food web (Langellotto and Denno 2006). Demographic changes in *Pardosa* abundance and feeding behavior could impact *T. minuta* through differential predation rates that alter in strength both spatially and temporally, or via predator avoidance behavior (Finke and Denno 2006, Wimp et al. 2011, Bucher et al. 2015), possibly via chemical cues left by spiders (Bucher et al. 2015). A small experiment conducted by the Wimp Lab, led by Sarah Qadir, also revealed that *T. minuta* males may move about blades more readily than females (unpublished data). If *Pardosa* rely primarily on visual cues, then this excessive movement may make the males easier prey, leading to differential predation rates between males and females.

Differences in abundances between *S. alterniflora* and water edges may provide evidence that *Pardosa* is suppressing *T. minuta* densities, warranting the pursuit of manipulative experimentation to investigate the influence of *Pardosa* on *T. minuta* dynamics.
Since *T. minuta* form a metapopulation across the marsh, this system may also provide the opportunity to study how bottom-up and top-down forces may impact metapopulation dynamics. For example, these dynamics within arthropod herbivore populations may be impacted by the genetic diversity within associated plant species. Since an increase in plant genetic diversity is expected to increase habitat quality through increased phenotypic and IIGE variation, it is possible that these source-sink dynamics may be amplified or diminished depending on the variation within the foundation species constituting each patch.

The knowledge gained from this study may have immediate practical application in restoration and conservation management, which is becoming increasingly important in the face of rapid global change and human alteration of the environment (Keller et al. 2015). For example, absolute species richness often depends on both patch geometry and habitat composition, and the inclusion of these factors, along with species-specific edge responses, has been shown to improve model performance (Sisk et al. 1997). Changes in composition near edges can alter richness from what is predicted by simpler models using patch area alone (Sisk et al. 1997). However, the impact of patch geometry is still a challenge for managers, particular those designing reserves since very large habitat fragments tend to exhibit a wide variety of shapes (Didham and Ewers 2012). This research focuses on the impact of both total and core area in the salt marsh system on genetic variation, and may help to shed light on this issue, providing knowledge that may further aid in reserve design.
CHAPTER 4: *Spartina patens* Spatial Population Genetic Structure Analysis Using New Microsatellite Markers

*Spartina patens*, commonly called marsh hay, is a high marsh cordgrass native to eastern North America (A. Baumel et al. 2002). Patches of *S. patens* tend to be dispersed throughout a matrix of the congeneric *S. alterniflora*, which is dominant in areas that are typically covered by the high tide. It is a perennial that reproduces both sexually and through vegetative growth via rhizomes (A. Baumel et al. 2002). Through rhizomal spread a single genotype can form a clump of blades called a genet, although the extent to which this occurs relative to sexual reproduction is not well understood. The dense mats of rhizomes formed by both *S. patens* and *S. alterniflora* are important in engineering and maintaining the integrity of the marsh (Kirwan 2009). For example, when this underground growth has been compromised by nutrient influx, marsh dieback has been observed (Darby F.A. 2008). Given how critical salt marshes are for the ecological and economic services they provide, understanding how the population genetics of *S. patens* influences its phenotypic characteristics and ecological impacts will likely aid in the restoration and maintenance of these important systems. This is becoming increasingly important in the face of anthropogenic habitat and climate changes that could lead to rapidly rising sea level and increasing salt water inundation (Kennish 2014).

There are differences in morphological and physiological trait expression by *S. patens* individuals in large patches versus those in small patches (Holzapfel 2011). A recent study indicates that large and small *S. patens* patches may be genetically differentiated from each other, even when highly connected, and that individuals in small patches may be less competitive than those in large patches. Individuals in large patches
may also produce smaller genets, which are clumps of identical clones, than those in small patches (Holzapfel 2011). Within the same patch, there is evidence that genotypes may differentiate over short distances. It is possible that differences in abiotic conditions may cause differences at a microclimate scales (Silander and Antonovics 1979). Since S. patens borders both water and S. alterniflora, this provides the additional opportunity to study the impact of edge and matrix type (Ries et al. 2004) on genetic differences between patches or over small scales within the same patch.

In order to perform this study, we aim to develop polymorphic microsatellite markers for use as genetic markers in S. patens. Although polyploidy and aneuploidy are common in the Spartina genus, S. patens has been reported to be a diploid (2n=40) (A. Baumel et al. 2002). However, we are not aware of any microsatellite markers that have been developed for this species. Microsatellites, also called short tandem repeats, are short, repeating segments of DNA, typically between 2 to 9 nucleotides in length. They are useful genetic markers because they are neutral, co-dominant, and have high mutation rates (Hamilton 2009). They are also relatively cheap to amplify, making them ideal for population level studies. Microsatellites are commonly used to produce estimates of genetic variation such as allelic richness and heterozygosity. Heterozygosity is a useful metric for inferring variation since Hardy-Weinberg expected heterozygosity (H_E) tends to increase with increasing polymorphism (Hamilton 2009). Empirical deviations from H_E are also informative. Higher than expected homozygosity in polymorphic neutral markers can indicate that phenotypic traits are also influenced by high homozygosity resulting from non-random mating or selfing. Inbreeding can significantly affect plant physiology, particularly traits affecting resource capture, and reduce fitness and survival.
Inbreeding in plants can also effect other species via ecological interactions such as habitat or food requirements, or trophic cascades (Kittelson et al. 2015).

In this study, we aim to investigate the effects of habitat edges on the spatial genetic structure of the foundation plant species *S. patens*. We hypothesize that there will be less variation in edge habitats relative to interior habitats. To perform this study, we collected samples from patches at the Great Bay Boulevard Management Area in Tuckerton, New Jersey, during the summers of 2014 through 2016. Since markers do not exist for *S. patens*, we developed new microsatellite markers from genome sequencing data.

METHODS

**DNA Sequencing**

DNA was extracted from a single *S. patens* individual collected in the summer of 2013 from the Great Bay Boulevard Wildlife Management Area in Tuckerton, New Jersey. Plant tissue from the blade was homogenized using liquid nitrogen and a mortar and pestle, and then the DNA extracted using a Qiagen DNeasy plant extraction kit. Genome sequencing was performed using an Illumina Hi-Seq 2500 instrument (Zalapa et al. 2012) in December, 2013. Paired-end reads 151 nucleotides in length were produced.

**Microsatellite identification**

An in-house python program written by Robert Battocletti was used to screen the single reads for tri-nucleotide repeats. Tri-nucleotide repeats are easier to score than di-
nucleotide repeats, reducing genotyping errors, while still minimizing issues associated with longer motif lengths, such as embedded di- or tri-nucleotide repeats. The program may be downloaded from https://github.com/GeneHub/PySTR.

**Primer Development**

We designed flanking PCR primers for the candidate loci using the Primer3 plug-in within Geneious software. We designed primers between 18 and 22 base pairs in length that contained a G-C clamp on the 3’ end. We targeted annealing temperatures between 55°C and 60°C and avoided primers that may form dimers or hairpins.

**Testing for polymorphism**

We empirically tested the primer pairs for successful locus amplification and polymorphism using *S. patens* individuals collected from Tuckerton, New Jersey, and Narragansett, RI, during the summer of 2014. We collected single grass blades by hand, and individually placed them into a small bag with an ID number. Samples were immediately brought to the lab and stored at -80°C. We homogenized plant tissue using a bullet blender with 1.6mm steel beads and extracted DNA using the chelex method described in Chapter 2. The 20.3 ul PCR reaction contained 4.0 ul of TaqOne GC-Rich Buffer, 0.4 ul of 100mM MgSO4, 2.0 ul of 2 mM each dNTP, 0.2 ul of TaqOne polymerase, 0.8 ul of 10 uM forward primer, 0.8ul of 10 uM reverse primer, 10.0 ul DI H20, and 2.5 ul of 1/40th dilution gDNA (250 to 450ng/ul undiluted). We performed PCR denaturation at 95°C for 5 minutes, followed by 30 cycles of 95°C for 1 minute, the annealing temperature for 1 minute, then 72°C for 2 minutes, and then a final elongation
step of 72°C for 10 minutes. PCR products were visualized under a UV light using 2% agarose gel and ethidium bromide.

We empirically tested primer pairs for 52 microsatellite loci for successful amplification using untagged primers and gel electrophoresis. Forty-one primer pairs produced PCR products, and 37 of these produced heterozygotes, indicating polymorphism. Primer pairs that consistently amplified more than two fragments were discarded. For those that were successful, we genotyped individuals using M-13 tagged primers to test for adequate polymorphism with more resolution (Schuelke 2000). The M-13 tagged primers may be used with dye for fragment analysis in the ABI 3100, allowing scoring to the nucleotide length level. The 6 ul PCR reaction contained 3.125 ul of 2X Quanta Tough Mix, 0.1 ul of 10 uM Forward M13 tagged primer, 0.4 ul of 10 uM reverse primer, 0.2 ul of M13 tagged 6-FAM dye, 1.55ul of DI H2O, and 1.875 ul of 1/40th dilution gDNA. The PCR protocol used was 94°C for 5 minutes, followed by 30 cycles of 94°C for 45s, the annealing temperature for 45s, then 72°C for 45s, followed by 8 cycles of 94°C for 30s, 53°C for 45s, and 72°C for 45s, and then a final elongation step of 72°C for 5 minutes. We genotyped individuals for fragment length using the ABI 3100 Genetic Analyzer and Geneious software. We genotyped 12 individuals from 3 populations to obtain general polymorphism and allelic richness data.

**Development of multiplex sets**

We selected a subset of the microsatellites that successfully produced fragments and displayed adequate polymorphism using the M-13 tagged primers (Schuelke 2000) to develop into dye-labelled multiplex sets. We tested co-amplification of primer pairs
using the M-13 tagged primers to determine which loci to amplify within the same set. We then ordered a single primer from each pair in a single set with either the NED, 6-FAM, PET, or VIC tag. The multiplex PCR products were analyzed using a ABI3100 genetic analyzer with McLabs LIZ 500 orange size standard, and the fragments scored using Geneious software.

**Spatial population genetic structure**

We collected samples of *S. patens* from Tuckerton, New Jersey, during the summers of 2015 and 2016. The sampling locations overlap those of *T. minuta*, enabling analysis of potential ecological genetic effects of *S. patens* on the genetic variation of *T. minuta*. The grass blades were collected individually by hand, placed into a small bag and given an ID number. Blades were collected systematically in 1 meter intervals in either a 5-m wide transect or full grid layout, depending on the dimensions of the patch. A label identifying the position within the grid or transect was placed into the bag with each blade. The sampling locations overlap those of *T. minuta*, enabling a correlational analysis of genetic variation measures. Samples are brought directly to the lab and stored at -80°C. Plant tissue was either homogenized using liquid nitrogen and a mortar and pestle, or chopped and then homogenized in the Bullet Blender with 1.6mm stainless steel beads. The DNA was extracted using either a Qiagen DNeasy plant extraction kit or the chelex method described in Chapter 2. The DNA was amplified for the microsatellite markers using PCR and fluorescently labeled primers. Individuals were genotyped for fragment length using the ABI 3100 Genetic Analyzer and Geneious software. Individuals collected in 2014 were used to obtain general polymorphism and allelic
richness data for each patch. Multi-locus genotypes were used to assess the spatial genetic structure within each patch, with particular attention to clonality and differentiation between edge and interior habitats.

*Patch Characteristics:* We estimated the total area, edge habitat area, *S. alterniflora* edge habitat area, and interior core area of each patch using Google Earth. Edge habitat is defined as any area within the *S. patens* patch that is 1 meter or less from an edge. Since large patches often exhibit asymmetrical geometry, we also quantified the patch characteristics for the sub-sections in which samples were taken.

*Edge-Interior Effects:* We assigned multilocus genotypes to their position within the transect or grid to obtain the spatial genetic structure. This spatial structure analysis indicates whether certain genotypes or alleles occur more frequently in edge or interior habitats than would be randomly expected.

**RESULTS**

**Microsatellite development**

We have tested primer pairs for 52 microsatellite loci. Forty-one have produced PCR products, and 37 produced heterozygotes, indicating polymorphism. We are presently testing loci for adequate polymorphism using M-13 tagged primers, and 5 loci have successfully been amplified in sets using dye-labelled primers (Table 3). We used these microsatellite markers to make estimates of critical evolutionary genetic measures, including levels of heterozygosity and allelic richness, in patches collected in 2014, and to determine spatial genetic structure in patches that were collected in 2016 using grids.
and transects. Across the populations collected in 2014, the single locus $H_O$ ranged from 0.17 to 0.83, and the $H_E$ from 0.16 to 0.74 (Table 4). The average observed heterozygosity was $0.53 \pm 0.27$ for Frag 28, $0.48 \pm 0.27$ for Frag A4, and $0.49 \pm 0.20$ for Narrow River. The total allelic richness for the loci ranged from 3 to 6, with an average of $4.40 \pm 1.52$ alleles. Based on these results, these five microsatellite markers displayed adequate polymorphism for use in the spatial genetic structure analysis.

**Spatial Population Genetic Structure**

The transect for 2A extended 32 meters from a *Spartina alterniflora* edge to a water edge. It was five meters wide, with the patch continuing on both sides. The *S. alterniflora* edge was dominated by a single genotype that was detected 19 meters into the interior of the patch. This single genotype encompassed 64 one m$^2$ grid cells, approximately 40.76% of the transect area. This genotype appears to have formed a large genet. Within the interior of the patch, defined as a grid cell found at least one meter from an edge, 24 distinct genotypes were discovered. These genotypes tended to form much smaller genets and even appeared to be intermixed in some areas. On average, these covered $3.04 \pm 2.97$ m$^2$ each, equivalent to 1.04% of the transect area. The water edge resembled the interior of the patch in terms of genet size and intermixing. Four distinct genotypes were found, covering $5.00 \pm 3.56$ m$^2$ each, equivalent to 3.18% of the transect area. It is important to note that since this transect only sampled a width of 5 meters, the genotypes could cover larger areas on either side that was not sampled.

Since frag A5 is much smaller than the patch 2A, we were able to collect samples in a grid pattern across the entire patch. Ten unique multi-locus genotypes were detected
in Frag A5, but it is dominated by 3 genets. In contrast to 2A, the largest genet extends from the water edge and covers 38 m², equivalent to approximately 48.10% of the patch area. The other two genets both extend from the *S. alterniflora* edge. The second largest genet covers 22 m², or 27.85% of the area. The third largest genet covers 9 m², or 11.39% of the area. The remaining seven genotypes are each only found in one or two grid cells, which covers only 1.27 to 2.53% of the patch area. Six of these seven genotypes are found within 2 meters of a water edge.

A limitation of this study is that we were only able to determine the spatial genetic structure for a single large patch (n=1) and a single small patch (n=1). To determine if a significant difference exists between edges and interiors, or patches of various sizes, will require the genotyping of additional samples collected in 2015 and 2016. During these summers, we collected either transects or grids from 11 patches. Producing the genetic data for these samples will greatly help to shed light on trends in *S. patens* spatial genetic structure and likely allow for stronger conclusions to be reached.

**DISCUSSION**

In the large patch (2A), a single genotype dominated the *S. alterniflora* edge (Figure 6a). With the exception of this genet, the findings are consistent with a previous study that suggested genets tended to be larger in small patches relative to large patches (Holzapfel 2011). At the *S. alterniflora* edge, it is possible that abiotic stress outweighs biotic factors, such as competition for resources. This genotype may be better adapted to higher salinities, and therefore able to outcompete other individuals in this environment.
or simply survive where other genotypes cannot. Salt (Niranjan Baisakh 2006) and heat (Niranjan Baisakh 2008) stress have been found to alter gene expression in S. alterniflora, and similar genetic responses could be occurring within S. patens. The measurement of abiotic conditions may be useful in determining if this genotype tracks a particular environmental condition.

Within the interior of the patch and towards the water edge, far more unique genotypes were detected and they tended to be more interspersed with much smaller genets. Rhizomes still extended reasonable distances to sprout new blades, as single genotypes were detected meters away, but they did not form single large genets (Figure 6a). These differences in genotype and genet size both between patches and within the same patch could have cascading effects by influencing other abiotic conditions and biotic interactions, such as nutrient availability and the success of other plant species (Proffitt C.E. 2005).

Other landscape characteristics, including patch area and connectivity, may influence S. patens spatial genetic structure and genetic variation, and thus warrant further investigation. Populations in larger patches may tend to have higher allelic richness and heterozygosity relative to populations in smaller patches. This could be revealed through analysis of total and private allelic richness (A and R), and observed and expected observed heterozygosity (H₀ and Hₑ).

Habitat fragmentation has been found to decrease genetic variation in common and rare plants via genetic drift in small, isolated populations (Honnay and Jacquemyn 2007, Aguilar et al. 2008). Heterozygosity may be lost via allele frequency changes associated with drift, or via non-random mating in which individuals are more likely to
mate with relatives due to lower densities. Inbreeding depression can negatively affect plant physiology, particularly traits affecting resource capture, and reduce fitness and survival (Kittelson et al. 2015). A reduction in resource quality or diversity can have cascading effects on the community, particularly when these losses occur in a foundation species.

Since gene flow can reintroduce alleles lost via drift, connectivity may play an important role in determining the amount of variation found within a patch. A recent study of spatial genetic structure in the congeneric hybrid, *S. foliosa* X *S. alterniflora*, suggests that seeds and/or pollen typically disperse within 200 meters or less of the parent (Sloop et al. 2011). Gene flow from neighboring patches may increase variation in patches with high connectivity. This model system presents an ideal opportunity to study the effects of patch size and connectivity on the genetic variation and spatial genetic structure of a foundation plant species.

Foundation plant species are critical to maintaining the habitat that supports higher trophic levels. This research will illuminate the effects of fragmentation on a critical foundation species, and by using a salt marsh model system, will provide system-specific knowledge critical for maintaining this important ecosystem. This research will both contribute broadly to the field of ecological genetics and have immediate application in conservation biology and wetland management.
Habitat fragmentation is major challenge confronting conservation biologists, making it critical to understand how changes in landscape characteristics and biotic interactions affect fragmented populations. In this study, we investigated the effects of habitat fragmentation on the genetic variation of both a foundation plant species, *Spartina patens*, and a dependent herbivore, *Tumidagena minuta*. We found that within *S. patens* patches, *T. minuta* tended to exhibit lower effective population sizes ($N_e$) in edge habitats bordering *S. alterniflora* relative to interior or water edge habitats. This suggests that populations may experience stronger genetic drift in *S. alterniflora* edge habitats. In large *S. patens* patches, our findings suggest that *S. patens* may exhibit lower variation in habitats bordering *S. alterniflora* due to the prevalence of a single or few genets, possibly because of the strong abiotic stress encountered. Within the large patch, in addition to the single genet that dominated the *S. alterniflora* edge, twenty-nine additional genotypes exhibiting smaller genet sizes and more intermixing were detected within the interior and water edge habitats. Within the small patches, 3 genets dominated the patch, with only 7 additional genotypes detected. The impact of the *S. alterniflora* edge on both *T. minuta* and *S. patens* reinforce the importance of matrix type and habitat patch shape in influencing fragmented populations.

In the larger *S. patens* patch, the *S. alterniflora* edge was dominated by a single genotype (Figure 6a). This lack of variation at the *S. alterniflora* edge could be affecting both the density and genetic variation found within *T. minuta*, resulting in lower estimates at that edge relative to the interior and water edge. In addition to low densities,
if IIGE’s are occurring between S. patens and T. minuta, then this dominance of a single genotype could be favoring particular genotypes within T. minuta. If T. minuta genotypes track changes in the underlying S. patens genotypes, then similar clumping of T. minuta functional genotypes may be occurring, leading to positive assortative mating due to proximity. This scenario could result in different long-term outcomes for T. minuta variation than simple changes in density within a well-mixed population, and warrants further investigation of differences in non-neutral, expressed phenotypes that may result from underlying genetic or epigenetic differences.

It is also possible that there are differences within the same genet due to epigenetic differences or spatial heterogeneity in gene expression resulting from underlying environmental conditions. This could result in phenotypic differences across different grass blades within the same dominant genets. If evidence of epigenetic modifications were observed, these large genets may present a unique opportunity to investigate the influence of epigenetic modifications on ecological interactions, community dynamics, and interspecific indirect genetic effects.

Along the water edge of large patches, both T. minuta and S. patens populations seemed to resemble the interior populations more so than the S. alterniflora edge populations. It is possible that both species are affected similarly and independently by other abiotic or biotic conditions. It is also possible that the S. patens spatial genetic pattern is a remnant of the historical extent of the patch. A phenomena associated with sea level rise called high marsh “ponding” may be occurring in which lower elevations within the large S. patens patches have accumulated water over time, turning into a shallow pond (Gedan et al. 2009). If this is occurring, then the water edges may have
been patch interiors prior to ponding. If ecological genetic effects are occurring, then it is possible that *T. minuta* are being influenced by the genetic variation within the *S. patens*. To understand these spatial genetic patterns will require further investigation.

The model system offers the opportunity to take a novel approach to studying habitat fragmentation by not only investigating contemporary processes at small spatial scales, but by also including the ecological genetic effects of fragmentation. It is unknown how genetic diversity within a dominant plant species may impact meta-population dynamics at higher trophic levels within fragmented systems. Because many habitats are experiencing rapid change and alteration, it is critical that we understand this interplay between trophic levels. Investigating the effects of edges on genetic variation, ecological interactions, and meta-population dynamics, will provide knowledge that may further aid in reserve design. This research may also contribute to the emerging field of landscape genetics, which seeks to integrate population and community genetics with landscape features and spatial structure, but still lacks a developed body of theory (Manel and Holderegger 2013). This research may both contribute to the fields of ecological and landscape genetics, and the system-specific knowledge gained may have immediate application in conservation biology and wetland management.
Appendix A: Tables

Table 1. Effective population size (Ne) estimates for Tumidagena minuta interior and edge locations within patches located in the Great Bay Boulevard Management Area, Tuckerton, NJ. Samples were collected in the summers of 2011 and 2015. The gametic disequilibrium methods employed include an $r^2_{\text{comp}}$ method with a minimum allele frequency of zero and the $r^2_{\text{perm}}$ permutation correction factor proposed by Sved (2013), and an $r^2_{\text{delta}}$ method with a minimum allele frequency of 0.05. Confidence intervals were obtained using a delete one locus jackknife percentile confidence interval method.

<table>
<thead>
<tr>
<th>Location</th>
<th>Collection Date</th>
<th>Density (adults/m²)</th>
<th>n</th>
<th>$N_e$ Estimate</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
<th>$N_e$ Estimate</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2A ES</td>
<td>30-Jun-11</td>
<td>653.5</td>
<td>96</td>
<td>160.2</td>
<td>144.7</td>
<td>190.1</td>
<td>149.4</td>
<td>82.4</td>
<td>444.9</td>
</tr>
<tr>
<td>2A Int</td>
<td>30-Jun-11</td>
<td>3318.3</td>
<td>102</td>
<td>173.3</td>
<td>152.8</td>
<td>217.8</td>
<td>147.9</td>
<td>132.9</td>
<td>Infinite</td>
</tr>
<tr>
<td>2C ES</td>
<td>30-Jun-11</td>
<td>693.0</td>
<td>96</td>
<td>255.5</td>
<td>220.9</td>
<td>311.6</td>
<td>455.1</td>
<td>143.6</td>
<td>149.3</td>
</tr>
<tr>
<td>2C Int</td>
<td>30-Jun-11</td>
<td>3256.3</td>
<td>96</td>
<td>346.8</td>
<td>278.6</td>
<td>415.8</td>
<td>Infinite</td>
<td>178.6</td>
<td>Infinite</td>
</tr>
<tr>
<td>Jose1 ES</td>
<td>30-Jun-11</td>
<td>788.7</td>
<td>96</td>
<td>169.1</td>
<td>144.6</td>
<td>231.3</td>
<td>274.5</td>
<td>91.6</td>
<td>Infinite</td>
</tr>
<tr>
<td>Jose1 Int</td>
<td>30-Jun-11</td>
<td>1797.2</td>
<td>96</td>
<td>228.4</td>
<td>187.9</td>
<td>300.0</td>
<td>391.9</td>
<td>107.8</td>
<td>Infinite</td>
</tr>
<tr>
<td>1D-A ES</td>
<td>14-Jul-15</td>
<td>158.5</td>
<td>35</td>
<td>82.3</td>
<td>68.1</td>
<td>120.0</td>
<td>122.3</td>
<td>43.3</td>
<td>Infinite</td>
</tr>
<tr>
<td>1D-A EW</td>
<td>14-Jul-15</td>
<td>457.7</td>
<td>48</td>
<td>140.9</td>
<td>96.7</td>
<td>297.9</td>
<td>207.9</td>
<td>67.5</td>
<td>Infinite</td>
</tr>
<tr>
<td>1D-A Int</td>
<td>14-Jul-15</td>
<td>1056.3</td>
<td>60</td>
<td>Infinite</td>
<td>Infinite</td>
<td>Infinite</td>
<td>Infinite</td>
<td>163.4</td>
<td>Infinite</td>
</tr>
<tr>
<td>1D-B Int</td>
<td>14-Jul-15</td>
<td>1059.9</td>
<td>60</td>
<td>180.2</td>
<td>141.0</td>
<td>295.3</td>
<td>734.3</td>
<td>109.5</td>
<td>Infinite</td>
</tr>
<tr>
<td>2A ES</td>
<td>17-Jul-15</td>
<td>288.7</td>
<td>48</td>
<td>Infinite</td>
<td>Infinite</td>
<td>Infinite</td>
<td>331.9</td>
<td>64.6</td>
<td>Infinite</td>
</tr>
<tr>
<td>2A EW</td>
<td>17-Jul-15</td>
<td>771.1</td>
<td>48</td>
<td>Infinite</td>
<td>Infinite</td>
<td>Infinite</td>
<td>Infinite</td>
<td>118.3</td>
<td>Infinite</td>
</tr>
<tr>
<td>2A Int</td>
<td>17-Jul-15</td>
<td>521.1</td>
<td>60</td>
<td>Infinite</td>
<td>Infinite</td>
<td>Infinite</td>
<td>4145.5</td>
<td>118.3</td>
<td>Infinite</td>
</tr>
<tr>
<td>Jeremiah_2 ES</td>
<td>14-Jul-15</td>
<td>-</td>
<td>30</td>
<td>114.8</td>
<td>81.7</td>
<td>Infinite</td>
<td>2738.0</td>
<td>57.6</td>
<td>Infinite</td>
</tr>
<tr>
<td>Jeremiah_2 EW</td>
<td>14-Jul-15</td>
<td>-</td>
<td>48</td>
<td>98.9</td>
<td>79.8</td>
<td>133.4</td>
<td>179.1</td>
<td>61.4</td>
<td>Infinite</td>
</tr>
<tr>
<td>Jeremiah_2 Int</td>
<td>14-Jul-15</td>
<td>503.5</td>
<td>60</td>
<td>189.5</td>
<td>151.6</td>
<td>693.5</td>
<td>Infinite</td>
<td>127.5</td>
<td>Infinite</td>
</tr>
</tbody>
</table>
Table 2. Allelic richness (A), observed and expected heterozygosity (H_0 and H_E) for all loci for both edge and interior locations within all patches. Asterisks indicate a significant heterozygote deficit as indicated by the Hardy-Weinberg Exact Test implemented in Genepop (p<0.05) (Rousset 2008).

a)

<table>
<thead>
<tr>
<th>Locus</th>
<th>Edge A</th>
<th>H_0</th>
<th>H_E</th>
<th>Interior A</th>
<th>H_0</th>
<th>H_E</th>
</tr>
</thead>
<tbody>
<tr>
<td>TmAAC252</td>
<td>10</td>
<td>0.61*</td>
<td>0.79</td>
<td>11</td>
<td>0.67*</td>
<td>0.79</td>
</tr>
<tr>
<td>TmAAC095</td>
<td>11</td>
<td>0.73*</td>
<td>0.82</td>
<td>13</td>
<td>0.75*</td>
<td>0.83</td>
</tr>
<tr>
<td>TmAAC062</td>
<td>15</td>
<td>0.64</td>
<td>0.61</td>
<td>13</td>
<td>0.55</td>
<td>0.53</td>
</tr>
<tr>
<td>TmAAC165</td>
<td>8</td>
<td>0.64*</td>
<td>0.70</td>
<td>7</td>
<td>0.65</td>
<td>0.61</td>
</tr>
<tr>
<td>TmAAC057</td>
<td>22</td>
<td>0.80*</td>
<td>0.90</td>
<td>21</td>
<td>0.80*</td>
<td>0.92</td>
</tr>
<tr>
<td>DdAAC206</td>
<td>9</td>
<td>0.72</td>
<td>0.73</td>
<td>11</td>
<td>0.74</td>
<td>0.75</td>
</tr>
<tr>
<td>TmAAC097</td>
<td>10</td>
<td>0.66*</td>
<td>0.76</td>
<td>12</td>
<td>0.67*</td>
<td>0.72</td>
</tr>
<tr>
<td>TmAAC236</td>
<td>9</td>
<td>0.56*</td>
<td>0.66</td>
<td>10</td>
<td>0.59*</td>
<td>0.66</td>
</tr>
</tbody>
</table>

b)

<table>
<thead>
<tr>
<th>Locus</th>
<th>Edge A</th>
<th>H_0</th>
<th>H_E</th>
<th>Interior A</th>
<th>H_0</th>
<th>H_E</th>
</tr>
</thead>
<tbody>
<tr>
<td>TmAAC252</td>
<td>10</td>
<td>0.51*</td>
<td>0.73</td>
<td>9</td>
<td>0.63*</td>
<td>0.76</td>
</tr>
<tr>
<td>TmAAC095</td>
<td>11</td>
<td>0.74*</td>
<td>0.85</td>
<td>13</td>
<td>0.83</td>
<td>0.84</td>
</tr>
<tr>
<td>TmAAC062</td>
<td>10</td>
<td>0.54</td>
<td>0.57</td>
<td>8</td>
<td>0.30</td>
<td>0.29</td>
</tr>
<tr>
<td>TmAAC165</td>
<td>6</td>
<td>0.55*</td>
<td>0.64</td>
<td>7</td>
<td>0.59</td>
<td>0.61</td>
</tr>
<tr>
<td>TmAAC057</td>
<td>25</td>
<td>0.91*</td>
<td>0.93</td>
<td>34</td>
<td>0.82*</td>
<td>0.92</td>
</tr>
<tr>
<td>DdAAC206</td>
<td>9</td>
<td>0.63*</td>
<td>0.66</td>
<td>8</td>
<td>0.55</td>
<td>0.59</td>
</tr>
<tr>
<td>TmAAC097</td>
<td>12</td>
<td>0.73</td>
<td>0.78</td>
<td>12</td>
<td>0.59*</td>
<td>0.67</td>
</tr>
<tr>
<td>TmAAC236</td>
<td>11</td>
<td>0.60</td>
<td>0.63</td>
<td>11</td>
<td>0.63</td>
<td>0.64</td>
</tr>
</tbody>
</table>
### Jose 1 - 20110630

<table>
<thead>
<tr>
<th>Locus</th>
<th>Edge</th>
<th>Interior</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>H₀</td>
</tr>
<tr>
<td>TmAAC252</td>
<td>12</td>
<td>0.54*</td>
</tr>
<tr>
<td>TmAAC095</td>
<td>12</td>
<td>0.69*</td>
</tr>
<tr>
<td>TmAAC062</td>
<td>11</td>
<td>0.53</td>
</tr>
<tr>
<td>TmAAC165</td>
<td>9</td>
<td>0.76</td>
</tr>
<tr>
<td>TmAAC057</td>
<td>22</td>
<td>0.80*</td>
</tr>
<tr>
<td>DdAAC206</td>
<td>9</td>
<td>0.60*</td>
</tr>
<tr>
<td>TmAAC097</td>
<td>12</td>
<td>0.67</td>
</tr>
<tr>
<td>TmAAC236</td>
<td>11</td>
<td>0.57</td>
</tr>
</tbody>
</table>

### 1D - 20150714

<table>
<thead>
<tr>
<th>Locus</th>
<th>Edge-S</th>
<th>Edge-W</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>H₀</td>
</tr>
<tr>
<td>TmAAC252</td>
<td>7</td>
<td>0.38*</td>
</tr>
<tr>
<td>TmAAC095</td>
<td>12</td>
<td>0.66*</td>
</tr>
<tr>
<td>TmAAC062</td>
<td>7</td>
<td>0.45</td>
</tr>
<tr>
<td>TmAAC165</td>
<td>6</td>
<td>0.48</td>
</tr>
<tr>
<td>TmAAC057</td>
<td>19</td>
<td>0.91</td>
</tr>
<tr>
<td>DdAAC206</td>
<td>6</td>
<td>0.74</td>
</tr>
<tr>
<td>TmAAC097</td>
<td>10</td>
<td>0.71</td>
</tr>
<tr>
<td>TmAAC236</td>
<td>8</td>
<td>0.60*</td>
</tr>
</tbody>
</table>
### e) 1D - 20150714

<table>
<thead>
<tr>
<th>Locus</th>
<th>Interior A</th>
<th></th>
<th></th>
<th>Interior B</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>H₀</td>
<td>Hₑ</td>
<td>A</td>
<td>H₀</td>
<td>Hₑ</td>
</tr>
<tr>
<td>TmAAC252</td>
<td>8</td>
<td>0.71*</td>
<td>0.78</td>
<td>10</td>
<td>0.70*</td>
<td>0.82</td>
</tr>
<tr>
<td>TmAAC095</td>
<td>13</td>
<td>0.64*</td>
<td>0.87</td>
<td>13</td>
<td>0.79*</td>
<td>0.86</td>
</tr>
<tr>
<td>TmAAC062</td>
<td>12</td>
<td>0.54</td>
<td>0.54</td>
<td>10</td>
<td>0.49</td>
<td>0.48</td>
</tr>
<tr>
<td>TmAAC165</td>
<td>6</td>
<td>0.62</td>
<td>0.63</td>
<td>7</td>
<td>0.66</td>
<td>0.72</td>
</tr>
<tr>
<td>TmAAC057</td>
<td>18</td>
<td>0.73*</td>
<td>0.90</td>
<td>23</td>
<td>0.89</td>
<td>0.93</td>
</tr>
<tr>
<td>DdAAC206</td>
<td>9</td>
<td>0.64</td>
<td>0.71</td>
<td>8</td>
<td>0.60</td>
<td>0.62</td>
</tr>
<tr>
<td>TmAAC097</td>
<td>10</td>
<td>0.73</td>
<td>0.66</td>
<td>10</td>
<td>0.63</td>
<td>0.69</td>
</tr>
<tr>
<td>TmAAC236</td>
<td>13</td>
<td>0.66</td>
<td>0.69</td>
<td>10</td>
<td>0.62</td>
<td>0.74</td>
</tr>
</tbody>
</table>

### f) 2A - 20150717

<table>
<thead>
<tr>
<th>Locus</th>
<th>Edge-S</th>
<th></th>
<th></th>
<th>Edge-W</th>
<th></th>
<th></th>
<th>Interior</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>H₀</td>
<td>Hₑ</td>
<td>A</td>
<td>H₀</td>
<td>Hₑ</td>
<td>A</td>
<td>H₀</td>
<td>Hₑ</td>
</tr>
<tr>
<td>TmAAC252</td>
<td>7</td>
<td>0.14*</td>
<td>0.79</td>
<td>7</td>
<td>0.16*</td>
<td>0.79</td>
<td>9</td>
<td>0.62*</td>
<td>0.76</td>
</tr>
<tr>
<td>TmAAC095</td>
<td>12</td>
<td>0.74*</td>
<td>0.83</td>
<td>8</td>
<td>0.65*</td>
<td>0.79</td>
<td>13</td>
<td>0.76*</td>
<td>0.86</td>
</tr>
<tr>
<td>TmAAC062</td>
<td>8</td>
<td>0.44*</td>
<td>0.55</td>
<td>12</td>
<td>0.54</td>
<td>0.56</td>
<td>14</td>
<td>0.53</td>
<td>0.49</td>
</tr>
<tr>
<td>TmAAC165</td>
<td>7</td>
<td>0.56</td>
<td>0.58</td>
<td>7</td>
<td>0.64*</td>
<td>0.70</td>
<td>6</td>
<td>0.64</td>
<td>0.66</td>
</tr>
<tr>
<td>TmAAC057</td>
<td>19</td>
<td>0.73*</td>
<td>0.90</td>
<td>19</td>
<td>0.78*</td>
<td>0.91</td>
<td>23</td>
<td>0.75*</td>
<td>0.91</td>
</tr>
<tr>
<td>DdAAC206</td>
<td>8</td>
<td>0.64</td>
<td>0.64</td>
<td>8</td>
<td>0.63</td>
<td>0.65</td>
<td>9</td>
<td>0.60</td>
<td>0.68</td>
</tr>
<tr>
<td>TmAAC097</td>
<td>9</td>
<td>0.68</td>
<td>0.71</td>
<td>12</td>
<td>0.73*</td>
<td>0.74</td>
<td>15</td>
<td>0.81</td>
<td>0.74</td>
</tr>
<tr>
<td>TmAAC236</td>
<td>9</td>
<td>0.62</td>
<td>0.66</td>
<td>8</td>
<td>0.69</td>
<td>0.66</td>
<td>9</td>
<td>0.71</td>
<td>0.78</td>
</tr>
<tr>
<td>Locus</td>
<td>Edge-S</td>
<td></td>
<td>Edge-W</td>
<td></td>
<td>Interior</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>--------</td>
<td>----------</td>
<td>--------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>(H_O)</td>
<td>(H_E)</td>
<td>A</td>
<td>(H_O)</td>
<td>(H_E)</td>
<td>A</td>
<td>(H_O)</td>
<td>(H_E)</td>
</tr>
<tr>
<td>TmAAC252</td>
<td>8</td>
<td>0.57*</td>
<td>0.80</td>
<td>10</td>
<td>0.49*</td>
<td>0.81</td>
<td>9</td>
<td>0.66*</td>
<td>0.78</td>
</tr>
<tr>
<td>TmAAC095</td>
<td>10</td>
<td>0.76</td>
<td>0.83</td>
<td>11</td>
<td>0.77</td>
<td>0.84</td>
<td>11</td>
<td>0.78</td>
<td>0.85</td>
</tr>
<tr>
<td>TmAAC062</td>
<td>11</td>
<td>0.50</td>
<td>0.47</td>
<td>11</td>
<td>0.53*</td>
<td>0.62</td>
<td>13</td>
<td>0.54</td>
<td>0.57</td>
</tr>
<tr>
<td>TmAAC165</td>
<td>6</td>
<td>0.59</td>
<td>0.65</td>
<td>6</td>
<td>0.48</td>
<td>0.54</td>
<td>8</td>
<td>0.66</td>
<td>0.66</td>
</tr>
<tr>
<td>TmAAC057</td>
<td>19</td>
<td>0.73*</td>
<td>0.90</td>
<td>16</td>
<td>0.82*</td>
<td>0.91</td>
<td>20</td>
<td>0.81*</td>
<td>0.92</td>
</tr>
<tr>
<td>DdAAC206</td>
<td>6</td>
<td>0.83</td>
<td>0.66</td>
<td>8</td>
<td>0.68</td>
<td>0.67</td>
<td>10</td>
<td>0.72</td>
<td>0.68</td>
</tr>
<tr>
<td>TmAAC097</td>
<td>8</td>
<td>0.67*</td>
<td>0.70</td>
<td>8</td>
<td>0.60</td>
<td>0.66</td>
<td>9</td>
<td>0.65</td>
<td>0.67</td>
</tr>
<tr>
<td>TmAAC236</td>
<td>9</td>
<td>0.90</td>
<td>0.67</td>
<td>11</td>
<td>0.75</td>
<td>0.70</td>
<td>12</td>
<td>0.62*</td>
<td>0.76</td>
</tr>
</tbody>
</table>
Table 3. Characteristics of the primer pairs developed for the 5 microsatellite loci used to assess population genetic structure in *Spartina patens*. Provided for each pair are the forward and reverse primer sequences, repeat motif, the fragment size (bp) from the initial sequence read, the annealing temperature used for single PCR amplification, and the size ranges across all 5 populations tested.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Primer sequences (5' - 3')</th>
<th>Repeat motif</th>
<th>Allele size (bp)</th>
<th>Size Range (bp)</th>
<th>Size Range (repeats)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sp001</td>
<td>F: ACTTATGTCTCGCGCTAGCA</td>
<td>TAA</td>
<td>111</td>
<td>72 - 84</td>
<td>10 - 14</td>
</tr>
<tr>
<td></td>
<td>R: GCGTTGTCTTCGCGTTC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sp002</td>
<td>F: GGATAAGTGGAGCTTCATACCAGC</td>
<td>TAT</td>
<td>83</td>
<td>68 - 83</td>
<td>8 - 13</td>
</tr>
<tr>
<td></td>
<td>R: CATCAGGCTAAGATCGGTGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sp003</td>
<td>F: AACGTCCACCTCTATATTATCGG</td>
<td>CAA</td>
<td>124</td>
<td>113 - 128</td>
<td>5 - 10</td>
</tr>
<tr>
<td></td>
<td>R: GTGGTTTCATGTCTAGTCTACCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sp004</td>
<td>F: GTGTTCCAGGGACATTGTTGCC</td>
<td>CTT</td>
<td>96</td>
<td>75 - 99</td>
<td>4 - 12</td>
</tr>
<tr>
<td></td>
<td>R: AGGCTCATAAACCTTCGTTTGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sp005</td>
<td>F: CAAAGGCATCAAACATGATCACC</td>
<td>TTA</td>
<td>94</td>
<td>73 - 91</td>
<td>3 - 9</td>
</tr>
<tr>
<td></td>
<td>R: TCCTTGGGTATTGTCAAGGTACG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Characteristics of the 5 microsatellite markers developed for *Spartina patens* and used to assess population genetic structure. Provided are the allelic richness (A), expected heterozygosity (H<sub>E</sub>), and observed heterozygosity (H<sub>O</sub>) across 3 locations collected in 2014. The frag 9 and frag 28 populations are located in the Great Bay Boulevard Wildlife Management Area in Tuckerton, New Jersey. The Narrow River population is located in Narragansett, RI.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Size Range (bp)</th>
<th>Size Range (repeats)</th>
<th>Narrow River (n=8)</th>
<th>Frag A4 (n=12)</th>
<th>Frag 28 (n=12)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td>HO</td>
<td>HE</td>
</tr>
<tr>
<td>Sp001</td>
<td>81 - 114</td>
<td>13-24</td>
<td>2</td>
<td>0.63</td>
<td>0.49</td>
</tr>
<tr>
<td>Sp002</td>
<td>68 - 80</td>
<td>08-12</td>
<td>2</td>
<td>0.25</td>
<td>0.22</td>
</tr>
<tr>
<td>Sp003</td>
<td>113 - 119</td>
<td>05-07</td>
<td>2</td>
<td>0.43</td>
<td>0.46</td>
</tr>
<tr>
<td>Sp004</td>
<td>81 - 96</td>
<td>06-11</td>
<td>5</td>
<td>0.75</td>
<td>0.74</td>
</tr>
<tr>
<td>Sp005</td>
<td>76 - 88</td>
<td>04-08</td>
<td>3</td>
<td>0.38</td>
<td>0.32</td>
</tr>
</tbody>
</table>
Appendix B: Figures

Figure 1. Conceptual representation of the assumptions underlying effective population size estimations. The yellow circle represents potential sources of observed disequilibrium.
a) [Graph showing the relationship between Total Area (m²) and Adult Density (#/m²)]

b) [Graph showing the relationship between Core area (m²) and Adult Density (#/m²)]
Figure 2. a) For large patches, the relationship between *T. minuta* interior adult density and patch sub-section total area on July 8, 2014 (blue, $r^2 = 0.13895$) and June 30, 2015 (orange, $r^2 = 0.01425$). b) For large patches, the relationship between *T. minuta* interior adult density and patch sub-section core area on July 8, 2014 (blue, $r^2 = 0.28272$) and June 30, 2015 (orange, $r^2 = 0.00058$). c) For large patches, the relationship between *T. minuta* interior adult density and patch sub-section non-*S. alterniflora* edge habitat area on July 8, 2014 (blue, $r^2 = 0.47159$) and June 30, 2015 (orange, $r^2 = 0.00053$).
a) 

![Graph showing Adult Density vs. Total Area](image)

b) 

![Graph showing Adult Density vs. Core Area](image)

c)
Figure 3. a) For small patches, the relationship between *T. minuta* patch adult density and total area on July 8, 2014 (blue, $r^2 = 0.7838$) and June 30, 2015 (orange, $r^2 = 0.60463$). b) For small patches, the relationship between *T. minuta* patch adult density and core area on July 8, 2014 (blue, $r^2 = 0.49075$) and June 30, 2015 (orange, $r^2 = 0.76717$). c) For small patches, the relationship between *T. minuta* patch adult density and non-*S. alterniflora* edge habitat area on July 8, 2014 (blue, $r^2 = 0.32705$) and June 30, 2015 (orange, $r^2 = 0.00053$).
Figure 4. a) The relationship between connectivity and interior observed allelic richness for adult *T. minuta* collected from patches 2A, 2C, and Jose 1 from the Great Bay Boulevard Management Area, Tuckerton, NJ, on June 30, 2011 (blue: $b_0 = -6 \times 10^{-5}$, $b_1 = -12.622$, $r^2 = 0.12954$) and for patches 1D, 2A, and Jeremiah 2 collected on July 14-17,
2015 (orange: \( b_0 = 0.0003, b_1 = 11.197, r^2 = 0.46228 \)). b) The relationship between connectivity and interior expected heterozygosity for adult *T. minuta* collected from patches 2A, 2C, and Jose 1 from the Great Bay Boulevard Management Area, Tuckerton, NJ, on June 30, 2011 (blue: \( b_0 = -1 \times 10^{-6}, b_1 = 0.7054, r^2 = 0.00351 \)) and for patches 1D, 2A, and Jeremiah 2 collected on July 14-17, 2015 (orange: \( b_0 = -1 \times 10^{-6}, b_1 = 0.7329, r^2 = 0.02533 \)).
Figure 5. a) The relationship between patch area and interior observed allelic richness for adult *T. minuta* collected from patches 2A, 2C, and Jose 1 from the Great Bay Boulevard Management Area, Tuckerton, NJ, on June 30, 2011 (blue: \(b_0 = -0.0005, b_1 = 13.029, r^2 = 0.83687\)) and for patches 1D, 2A, and Jeremiah 2 collected on July 14-17, 2015 (orange: \(b_0 = -0.0005, b_1 = 13.029, r^2 = 0.83687\)).
b = 0.0005, b_1 = 10.835, r^2 = 0.28723).  b) The relationship between patch area and interior expected heterozygosity for adult T. minuta collected from patches 2A, 2C, and Jose 1 from the Great Bay Boulevard Management Area, Tuckerton, NJ, on June 30, 2011 (blue: b_0 = 4X10^{-5}, b_1 = 0.6529, r^2 = 0.44227) and for patches 1D, 2A, and Jeremiah 2 collected on July 14-17, 2015 (orange: b_0 = 2X10^{-5}, b_1 = 0.7074, r^2 = 0.99048).
Figure 6. Spatial genetic structure for a) the transect of 2A, and b) the grid of Frag A5. Genotypes bordering the *Spartina alterniflora* edge are shaded light grey, and those bordering the water edge are shaded dark grey. The top of the figure A represents the *S. alterniflora* edge, and the bottom is the water edge. In figure b, both the top and upper right side represent the *S. alterniflora* edge, and all other sides are bordered by water. Genotypes were arbitrarily assigned a letter. The letters do not indicate relatedness and do not represent the same genotype in different patches. Dashes indicate that a multi-locus genotype could not be assigned due to missing data.


Rousset, F. F. 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular ecology resources 8:103-106.


