THE RELATIONSHIP BETWEEN SPECIES DIVERSITY AND GENETIC DIVERSITY IN A FRAGMENTED FOREST SYSTEM: A CASE STUDY OF *CAREX LEPTONERVIA* IN THE MONTÉRÉGIE, QUÉBEC

Catherine Polcz
Master of Science
Department of Plant Science
McGill University
Montreal, Quebec

December 2014

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

DEDICATION

For Catharine Parr Traill

ABSTRACT

The fragmentation of native temperate forests is a defining issue of rural and urban development in southern Québec. From a forest plant perspective, as natural forest habitat is reduced and subdivided into smaller, more isolated patches, environmental conditions are altered, population size is decreased, and species dispersal may be interrupted. Thus, fragmentation may cause reduced species and genetic diversity in plant systems. Traditionally, ecological and population genetic studies have assessed the impacts of fragmentation on the two measures of diversity separately. In this thesis, we consider both fundamental aspects of biodiversity in a fragmented system. We chose to compare the genetic diversity (GD) of *Carex letponervia* with species diversity (SD) at two levels: study site and community level. Study site SD measurements approximate the diversity of all plant species growing in a forest, and community SD measurements describe the diversity of plant species growing with *C. leptonervia*. In chapter 2 of this thesis we describe the development of seven simple sequence repeat (SSR) markers we developed for *C. leptonervia*. The markers we used came from three different sources: a set of putative loci identified in a partial genome sequence of *Carex lupulina*; SSR sequences that were previously developed for *Carex scoparia*; and SSR sequences previously developed in our lab for *Carex limosa*. Fifteen percent of loci x population tests revealed departures from Hardy-Weinberg equilibrium. Linkage equilibrium tests revealed that two loci were significantly linked across all populations. In chapter 3 we describe our overall goal of relating landscape structure, genetic diversity, and species diversity. We hypothesize that site area, connectivity, and environment will similarly impact SD of *C. leptonervia* communities and GD of *C. leptonervia* and that SD and GD are positively correlated.

To address this, we assessed overall plant species richness in 20 forest patches ranging from six to over 2000 hectares. Within each forest, we also measured species diversity of the plant community associated with *C. leptonervia* by centering seven 4 m^2 quadrats on focal *C. leptonervia* plants. Thirty individuals of *C. leptonervia*, including one from each of the focal quadrats, were sampled at each site to assess genetic diversity.

Genetic diversity was characterized using six of the SSR loci that we developed in chapter 2. We found that species richness (SR) at the *forest patch level* shows a strong positive trend with patch area, but not with connectivity among patches. We found that at the *point-community level*, high SR related to high values of SR at the overall site and that Shannon diversity was positively correlated to *community extent* (a measure that calculated by dividing SR of point-communities by SR of sites). We found no evidence that soil characteristics (pH and percent organic matter) impact values of point-community diversity. Principal coordinates analysis and canonical correspondence analysis ordinations revealed that *C. leptonervia* is ecologically flexible and grows with a large number of species and in a variety of habitat types. To assess patterns in genetic diversity, we ran tests of isolation-by-distance, Nei's G_{ST}, and a Bayesian clustering analysis. These results all show that *C. leptonervia* has low population structure and that there is gene flow throughout the Montérégie. To test for a species diversity-genetic diversity correlation (SGDC), we conducted Pearson correlation tests between all GD measures and SD measures (site and point-communities). We found that gene diversity exhibited a positive significant correlation to species richness at the site level but was not significantly correlated to the diversity of *C. leptonervia* point-communities. These results indicate that processes important at the level of site are also important in maintaining GD. We conclude that for an ecologically flexible and well-dispersed species, it is likely that the point-community diversity as estimated by quadrat sampling underestimates the actual diversity of species co-existing with *C. leptonervia,* and that the diversity of site may provide a more accurate approximation. We conclude that the parallel process driving the correlation is likely study site area. Surprisingly, we found that landscape connectivity was not an important determinant of SD or GD. In conclusion, we propose that linking population genetics and community ecology can further an understanding of natural systems and also that isolated forests are important in maintaining diversity in the Montérégie.

RÉSUMÉ

Les milieux forestiers tempérés du sud du Québec sont de plus en plus fragmentés et isolés les uns des autres. Les conditions environnementales sont grandement altérées, la taille des populations est réduite et la migration et dispersion des espèces sont perturbées et même parfois interrompues. La fragmentation des forêts réduit donc la diversité des espèces et la diversité génétique. Cette thèse combine ces deux aspects fondamentaux de la biodiversité puisque les études antérieures les ont analysés séparément. L'espèce focale utilisée pour étudier la diversité génétique fût *Carex leptonervia* et ce, sur deux échelles : à l'échelle des lots boisés isolés ainsi qu'à l'échelle des communautés immédiates de nos plantes focales.

Le chapitre 2 explique les sept marqueurs de séquences répétées (SSR) développés pour *Carex leptonervia.* Ils proviennent d'un ensemble de loci identifiés du génome de *Carex lupulina,* de séquences SSR développées pour *Carex scoparia* et de séquences développées dans notre laboratoire pour *Carex limosa.* Nous avons observé que 15% des loci utilisés pour les analyses de populations pour HWE étaient instables. De plus, nous avons identifié deux loci liés de façon significative à travers toutes les populations.

Le chapitre 3 définit nos objectifs généraux quant à la compréhension des facteurs qui lient la structure du paysage, la diversité génétique, la diversité des espèces et énonce les questions qui suivent. Selon les perspectives de génétique des populations et de macroécologie, les processus sont-ils similaires en comparant les modèles récurrents de diversité dans une région donnée? Quelles sont les relations entre la diversité génétique et la diversité des espèces dans ce système? Pour adresser ces questions, nous avons recensés le nombre d'espèces végétales dans 20 milieux forestiers de dimensions variant entre six et deux milles hectares. Dans chaque lot boisé, nous avons recensé la diversité de la communauté directement associée au *Carex leptonervia* en définissant un quadrat de 4 mètres carrés autour des plants. Trente individus, dont sept plantes focales, furent échantillonnés à chaque site afin de

quantifier la diversité génétique de chaque site. Celle-ci fut caractérisée en utilisant six loci SSR développés au chapitre 2.

Nos résultats révèlent que la richesse des espèces d'un lot est liée de façon significative à sa dimension et non pas à sa connectivité avec les sites boisés avoisinant. Dans les communautés immédiates des plantes focales, la richesse des espèces (SR) est proportionnelle à celle des lots. L'indice de diversité Shannon (SHAN) est quant à lui relié au niveau de dispersement des *C. leptonervia* à travers leur site. Ce niveau de dispersement (*community extent*), est calculée en divisant *SR* des communautés immédiates par *SR* des sites.

Les caractéristiques du sol telles que le pH et le pourcentage de matière organique n'ont pas influencés de manière significative la diversité des communautés immédiates. L'analyse en coordonnées principales et l'analyse des ordinations de correspondance canoniques (*CCA*) montrent que *C. leptonervia* est une espèce aux conditions environnementales flexibles et qui interagit avec une grande variété d'espèces. Nous avons procédé aux tests suivant pour déterminer la redondance de la diversité génétique : *isolation-by-distance*, *Nei's* G_{ST} et *Bayesian clustering analysis*. Le *Carex leptonervia* est vastement dispersé à travers la Montérégie et ne semble pas avoir de structure de population bien définie.

Nous avons effectué un test de corrélation Pearson avec chaque variables de diversité (H_E, H_O et *AR* pour la diversité génétique; *SR Site*, *SR Quadrat* et *SHAN quadrat* pour la diversité des espèces) pour trouver la corrélation entre la diversité génétique et la diversité des espèces. La diversité génétique (H_E) est positivement reliée à la diversité des espèces (*SR*) au niveau des sites mais non au niveau des communautés immédiates des plantes focales. Ces résultats indiquent que les processus majeurs à l'échelle des sites sont aussi importants au maintien de la diversité génétique.

Considérant que le *C. leptonervia* est une espèce écologiquement flexible et bien répandue à travers le territoire, nous concluons que l'échelle du site est plus représentative que l'échelle des communautés directement associées aux plants pour analyser la diversité des espèces d'une communauté donnée. Selon nos analyses, cette

corrélation est principalement influencée par la superficie du site. En conclusion, nous proposons de combiner la génétique des populations et l'écologie des communautés afin d'améliorer notre compréhension des systèmes naturels et des forêts isolées afin de préserver la diversité montérégienne.

ACKNOWLEDGEMENTS

Many people have helped me complete this thesis over the past two years. Thanks to my supervisor, Marcia Waterway for her guidance and expertise. Thanks to my academic committee, Elena Bennett and Jaswinder Singh for their useful comments. Thanks again to the Bennett lab for accepting me into their lab group and for their amazing helpfulness and willingness to share field sites and GIS resources with me. Thanks to Raphaelle Frechon for her help in the field, for running the soil pH and organic matter analyses, for contacting the landowners in the Montérégie, translation of the abstract, easy laughter, and friendship. Thanks to David Maneli for helping us organize our stay at the Gault Nature Reserve. I also send my gratitude to the many landowners in the Montérégie who welcomed us into their forests. Thanks to Eric Roalson (Washington State University) for sharing his unpublished data on *Carex lupulina* SSR loci. Thanks to Rhea Dsilva and Lesley Seto for sharing their expertise in the lab. Thanks to Shuningbo Ye for her tireless assistance in helping me run gels. Thanks to Rosalie Frechette and the McGill University and Génome Québec Innovation Centre for providing sequencing and genotyping services. Thanks to Guillaume Larocque for his patience and guidance in all things statistics, and GIS. Thanks to Ellie Goud, Nadia Cavallin, and Alberto Prado for their enthusiasm in talking botany in the hallways. And finally, thanks to my support network beyond McGill, for seeing me through this!

I would also like to thank my funding sources. This research was funded by an NSERC Discovery Grant awarded to Dr. Waterway. I also received a Graduate Excellence Fellowship from McGill and an Excellence Award from the Québec Centre for Biodiversity Science that allowed me to travel and present my results at two scientific meetings.

PREFACE

This thesis is manuscript-based and consists of a literature review, two manuscript chapters, and a chapter on overall conclusions. The first chapter provides a literature review and general introduction to the thesis. The second chapter describes the development of seven *Carex leptonervia* microsatellite loci. The microsatellite loci we describe in the second chapter are used in our analysis of genetic diversity in the third chapter. Chapter three describes a study that compares species diversity of *Carex leptonervia* communities to the genetic diversity of *C. leptonervia* in the Montérégie region of Southern Québec. The second and third chapters are both in preparation for submission to *Molecular Ecology Resources* and *Ecography* respectively. References are included at the end of the thesis. Referencing follows American Psychological Association (APA) style.

CONTRIBUTIONS OF AUTHORS

Marcia Waterway is the second author of chapter 2 and chapter 3. For chapter 2, she assisted in the selection of final markers and in deciding upon a microsatellite screening protocol. I ran all of the analyses and conducted the lab work. For chapter 3, the initial ideas were my own and Dr. Waterway helped me further refine my ideas and to create a sampling protocol. I decided upon the statistical analyses and ran the tests myself. The writing is my own, but greatly improved with the helpful and thorough edits provided by Dr. Waterway.

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

1 LITERATURE REVIEW AND GENERAL INTRODUCTION

1.1 General Introduction

Forest fragmentation has been identified as a major threat to conservation throughout many of the world's forest systems, and threatens many different taxa, ecosystems, and ecosystem services. In response, the impacts of fragmentation are widely studied. In addition to their conservation focus, fragmentation studies are also conceptually rich and have offered researchers the opportunity to test many ecological theories. Most recently, fragmented systems have provided a system in which parallel ecological and evolutionary processes may be examined. In this thesis we characterize the impact that forest fragmentation in the Montérégie in southern Québec has on the genetic diversity of *Carex leptonervia,* and the species diversity of *C. leptonervia* communities. We also test the hypothesis that similar forces drive species diversity and genetic diversity and we test for a correlation between the two diversity levels. Chapter 2 of this thesis describes the development of microsatellite markers that we use to characterize the genetic diversity of *Carex leptonervia* in chapter 3.

1.2 Research objectives

In addition to contributing ecological data to the Montérégie region conservation efforts, my project increases knowledge of *Carex* ecology and genetics. There exist already many studies that consider how habitat fragmentation *individually* affects genetic variation and species diversity, yet, fewer studies that look at how fragmentation impacts the two diversity measures simultaneously. Furthermore, there is little consensus on the causes of positive correlations between species diversity and genetic diversity. To my knowledge, our study is the first to compare genetic diversity to *two* levels of species diversity. Thus, we hope to examine the role of both spatial and deterministic community assembly processes in species diversity and genetic diversity correlations.

1.3 Conceptualizing fragmented populations

Although naturally occurring "patchiness" in a landscape may exist (for example: discrete wetlands in the subarctic Vellend & Waterway, 1999), fragmentation is most often a term used to describe the results of anthropogenic disturbance. Habitat fragmentation creates a novel configuration of habitat in a region, where habitat fragments are smaller, more isolated, and separated by non-habitat areas, usually agricultural or urban areas.

In order to understand the ecology of fragmented landscapes, ecologists have applied the theory of island biogeography (MacArthur & Wilson, 1967; Vellend, 2003) and metapopulation theory (Verheyen et al., 2004; Honnay et al., 2005; see review of concept in Hanski & Gilpin, 1991) as conceptual frameworks. Both of these theories emphasize the interaction between discrete populations or habitat patches. The defining parameters of community assembly in island biogeography theory are: immigration rate, extinction rate, distance from mainland source areas, and island size. In the translation of island biogeography to a fragmented landscape, *islands* translate into *habitat fragments*. Metacommunity theory itself stems from island biogeography theory, and re-adapts it to non-island habitats (Gonzales, 2004). In both of these theories, species richness in any given habitat is a function of area, isolation, and habitat diversity. Therefore characteristics of any given community are a product of local processes and regional processes and are largely determined by the qualities of neighbouring fragments (Leibold et al., 2004). Though a useful tool, the adoption of the island biogeography framework in the aid of conservation decisions is not without its critics, most notably Yrjö Haila, who claims that the real pattern most commonly observed in fragmented landscapes are simple species-area relationships (Haila, 2002).

The role of area is established in many ecological patterns, but the role and definition of connectivity in island or fragmented systems appears less consistent. Measurement of connectivity takes into consideration both the *distance* of a focal patch to neighbouring patches, and also the *size* of those forest patches. Tischendorff & Fahig (2000) define connectivity as "the degree to which a landscape facilitates or impedes

movement of organisms among resource patches." Thus, landscape connectivity describes a term only meaningful in a species-specific context, since different species have differing abilities to move through non-habitat landscapes and different dispersal potential (Wiens & Milne, 1989; Tischendorf & Fahrig, 2000).

1.4 Forest fragmentation and the physical environment

In addition to increased spatial isolation of populations, fragmentation also alters numerous environmental variables. The break-up of continuous habitat into small units increases the perimeter-to-area ratio of the patches, thus subjecting a greater patch area to the differing environmental conditions of the edge (Murcia, 1995). Reported edge effects include changes in microclimate, increased light penetration, increased air and soil temperature, decreased humidity (Matlack, 1993; Chen et al., 1993; Honnay et al., 2005) and greater chemical inputs from adjacent agricultural fields (Murcia, 1995). These alterations at the edge may then alter patterns of herbaceous layer regeneration and interspecies competition, as well as plant-animal interactions (Honnay et al., 2005). Furthermore, these impacts may penetrate as far as 50 m into the forest interior (Matlack, 1993).

In addition to edge effects, forests fragments may exhibit a narrow range of environmental conditions. For example, in New England, preferential clearing of land left disproportionate amounts of slope and well-drained sand plains (Hall et al., 2002). Forest patches may also be swampy, or may experience disturbance associated with timber and maple syrup production (personal observation in the Montérégie of S. Québec). As a result, forest fragments may not be good approximations of the environmental conditions found in large continuous forests.

1.5 A population genetics approach to habitat fragmentation

Genetic diversity (GD) describes the evolutionary potential of a species, where populations with high GD have greater evolutionary potential (Taberlet et al., 2012). Processes that govern genetic diversity in plant populations are: mutation, genetic drift, migration, and selection. Other determinants include population size and breeding

system. GD is measured as the variation in a set of genes, and is commonly used in assessment of inbreeding, migration, and population differentiation studies.

Traditional theoretical models of genetics in fragmented systems include: Wright's infinite island model (1931), the stepping stone model (Kimura, 1953; Kimura & Weiss, 1964), metapopulation theory (Levins, 1969), Wright's (1943) isolation-by-distance (IBD) and recent landscape context models (see review in Manel et al., 2003). The earliest of these models, Wright's infinite island model, describes a system where islands (or any discrete population) freely exchange propagules with each other, and all exhibit equal rates of migration and genetic drift. Kimura's stepping stone model states that migration happens “in-steps” and that exchange in each generation is restricted between adjacent colonies. Metapopulation gene flow models build upon Kimura's model and assume unequal gene flow in populations in a system dominated by extinction and colonization dynamics. These models all seek to explain differences in the amount of genetic diversity between sites. IBD was originally conceptualized to describe the probability of mating between individuals as a function of the geographic distance between them in continuous populations (Wright, 1943). Recently, this term has been adapted to describe the negative correlation between gene exchange and geographic distance between discrete populations (e.g. Broquet et al., 2006). Finally, recently developed landscape context models, consider more explicitly the landscape mosaic and emphasize the role of barriers to gene flow, and thus require a detailed characterization of the landscape.

The reality of asexual reproduction in many plant species means that their patterns of genetic diversity may not always follow those predicted by theory that assumes outcrossing. The breeding system of plants plays an important role in measures of genetic diversity, since species may either exclusively outcross, self-pollinate, exhibit clonal growth, or have a mixed breeding system (may outcross and self-pollinate). For example, a study of three herbs in an agricultural landscape (Berge, 1998) found that two outcrossing species *Festuca ovina* and *Lychnis viscaria,* had more genetic diversity than the mostly inbreeding species *Arabidopsis thaliana.*

1.6 Fragmentation causes losses in genetic diversity

Losses in GD of plants in fragmented populations have been well documented (see review in Young et al*.,* 1996, and Aguilar et al., 2008). GD is initially lost when continuous habitat is cleared and individuals are lost, GD is then further reduced via the effects of increased inbreeding and genetic drift (Vellend et al. 2003; Young et al*.,* 1996). Fragment isolation may also inhibit gene flow, causing isolated habitat fragments to receive fewer propagules thus further accelerating genetic drift. Admixture (the mixing of divergent genetic lineages) occurs when propagules migrate to or from isolated populations. Although admixture is normally thought to be beneficial by increasing genetic diversity, recent studies have also found otherwise, as admixture may cause a loss of locally adapted genotypes at the landscape scale (Rius & Darling, 2014). Furthermore, the negative impact fragmentation has on GD is felt more strongly in each subsequent generation (Aguilar et al., 2008).

For plants, reduced genetic diversity can lead to reduced germination, seed, pollen and ovule production, the fixation of deleterious alleles and changes in plant physiological traits (summarized in Thrall et al., 2000). In addition to these changes, reduced genetic diversity increases extinction risk as populations may become less resilient to changes in environmental conditions, storms, disease events and changing selection pressures (Young et al., 1996).

There are several life history aspects that determine how rapidly and to what extent fragmentation will impact intraspecific genetic variation. Since the effects of inbreeding and drift become more pronounced over generations, genetic erosion is more quickly evident in short-lived species (Young et al., 1996). Plants that reproduce asexually or clonally appear buffered from a loss of genetic diversity when observed over a short duration (Young et al., 1996, Schaal & Leverich, 1996). However when observed on longer time scales, erosion of clonal diversity also leads to reduced genetic diversity in rhizomatous species, as found in *Carex lasiocarpa* (McClintock & Waterway, 1994). Increased distance and isolation from neighbouring populations greatly disrupts gene flow in species that exploit wind or gravity for seed dispersal and pollination (Aguilar et

al., 2008). The vulnerability or resilience of a species is also linked to its prevalence throughout the landscape. Naturally rare species by definition have small populations and may be more resilient genetically to fragmentation than more common species with high pre-fragmentation levels of GD (Aguilar et al., 2008). Apomictic species (those that produce seeds without sexual reproduction) may suffer no losses in GD, and are often able to sustain high levels of intrapopulation genetic variation (Schaal & Leverich, 1996).

Factors that increase connectivity have demonstrated positive effects on genetic diversity (Vellend and Geber, 2005); thus connectivity among patches has been a major focus of previous studies. Corbit et al. (1999) found that both remnant and regenerated hedgerows act as dispersal corridors for a number of forest herbaceous plants, and that the width of the hedgerow and the proximity of the hedgerow to forests are important factors in their ability to act as corridors. A study in southern Québec similarly found a large number of forest herbs growing in hedgerows, including *Carex leptonervia* (Roy & de Blois, 2006). Even low levels of gene flow among populations via pollen or seeds can alleviate loss of genetic diversity by dampening effects of genetic drift (Aguilar et al., 2008).

In addition to decreasing levels of genetic diversity, isolation of patches in fragmented environments increases the genetic differentiation of populations (Schaal & Leverich, 1996; Honnay et al., 2005). Patches that are well connected and in close proximity likely have higher gene flow and are expected to have a high proportion of shared alleles.

1.7 The use of genetic markers in population genetics

The field of population genetics may best be understood as having three levels of inquiry: the genotypic analyses of individuals; analyses of gene flow and migration history of populations; and phylogenetic and speciation analyses (Sunnocks, 2000). As a result of these differing areas of inquiry, many types of molecular markers exist, each exhibiting differing rates of mutation and sensitivity to the timescale of the inquiry. For questions at an ecological time scale, the use of microsatellites, non-coding 1-6 nucleotide base tandem repeats, has become increasingly common. In addition to being

highly polymorphic, microsatellites (SSR's) exhibit high success rates for cross-amplification among congeners, and are co-dominant, allowing assessments of heterozygosity. The accepted mechanism causing high mutation rates and thus polymorphism in microsatellites is DNA replication slippage. This is an error in DNA replication and occurs when the nascent and template strands "realign out of register" (Schlötterer, 2000). Different microsatellite alleles thus differ in the number of tandem repeats, as DNA replication slippage mutations act either to add or delete tandem repeats (Schlötterer, 2000). Microsatellites are useful in revealing two things: allele identity and the differences in number of tandem repeats. Differences in the size of microsatellites alleles (number of tandem repeats) may yield more nuanced examinations of population divergence. Furthermore, for population genetic studies, it is essential that molecular markers be selectively neutral. Neutral markers provide measures of genetic drift and migration and are not subject to selection. This is essential, as markers susceptible to selection cannot distinguish selection from drift or migration. Although some studies suggest that microsatellites may be under selection (Kashi & King, 2006; Gemayel et al., 2010) they are commonly used to estimate levels of neutral genetic diversity in population genetic studies. Single nucleotide polymorphisms (SNPs) are another unilocus technique growing in popularity. SNPs have lower mutation rates than microsatellites and are most often diallelic (Guichoux et al., 2011), making them less useful for many population genetic studies. Allozymes are another unilocus co-dominant marker that have recently fallen out of common use. They were among the first markers used, but have been replaced by other markers with higher mutation rates and resolution (Sunnocks, 2000). Formerly popular, multilocus techniques include randomly amplified polymorphic DNA (RAPDs), and amplified fragment length polymorphisms (AFLPs). Cited drawbacks to multilocus techniques include variation that may be non-heritable or derived from other organisms, and also the lack of co-dominance (Sunnocks, 2000).

1.8 Measurements of population differentiation, isolation-by-distance, and genetic structure

There are two main classes of analyses that can be employed to assess genetic differentiation among populations: F-statistics and R-statistics. F-statistics were developed to assess genetic differentiation of populations when divergence is a product of genetic drift. The first F_{ST} evaluation was developed by Wright (1969) for biallelic markers. Nei (1973) generalized F_{ST} for use of markers with multiple alleles. Nei's evaluation is sometimes reported as Nei's coefficient of gene variation, or G_{ST}. G_{ST} is calculated in the formula $1-H_S/H_T$, where H_S is the mean heterozygosity within populations, and H_T is the total heterozygosity of combined populations. G_{ST} values range from 0 to 1. Values close to 0 indicate that most of the genetic variation is within populations, and that many populations have the same alleles, with low differentiation of populations. High G_{ST} indicates that most of the variation is between populations and that populations are highly differentiated from one another. F-statistics are also hierarchical and can be assessed globally, for populations, and for individuals. Slatkin's R_{ST} analysis is modelled on Wright's F_{ST} but accounts for a stepwise mutation model (Balloux & Logan-Moulin, 2002). The stepwise mutation model (SMM) declares that the DNA replication slippage mechanism either adds or deletes tandem repeats of the microsatellite locus, and thus alleles that are closer in length are more recent derivations than those with greater differences in length. It is this accounting for a SMM in R-statistics that makes it an appropriate measure for analysis of microsatellite markers, as F-statistics are unable to distinguish drift from mutation. However, an oft-cited drawback of R-statistics, is high variance, and to achieve results of a similar power (to F-statistics), more individuals and populations need to be sampled.

Isolation-by-distance (IBD) analyses test whether populations geographically closer together have more shared alleles (see review in Meirmans, 2012). To compute IBD, a Mantel test is used to test correlation between a pairwise genetic distance matrix and a geographic distance matrix. If there is no correlation it may be inferred that dispersal is not limiting at the geographic scale of study.

Bayesian evaluation of population structure (Pritchard et al., 2010) provides information about how many distinct genetic clusters underlie a population or meta-population under study. The popular program STRUCTURE (Pritchard et al., 2010) seeks to identify genetic clusters based upon differences in allele frequency and co-occurrence. These techniques are most often employed when a geographic or latitudinal barrier is suspected of disrupting gene flow. The characterization of population structure provides a useful aid in the interpretation of F_{ST} and IBD.

1.9 Plant species in secondary forests

Forest fragments are best understood as being of two types: remnant forests and secondary forests. Remnant forests are defined as areas of pre-fragmented landscape that have been disturbed (but not previously cleared), whereas secondary forests have regrown after the abandonment of agricultural lands (Rackham, 1980). The species composition of each is expected to be distinct because of these different origins.

Secondary forests may also have the unique feature of having lost all ancestral plants and diaspores (Hermy & Veheryen, 2007; Jacquemyn & Brys, 2008). Therefore, for secondary forests, diversity is largely determined by dispersal and recolonization ability of herbaceous forest species (Bellemare et al., 2005; Vellend et al*.,* 2007). Species that lack morphological adaptations for dispersal or those dependent on myrmecochory (ant dispersal) have much lower secondary forest abundances than species with seeds adapted for wind, bird, or mammal dispersal (Bellemare et al*.,* 2001). Furthermore, it may be assumed that species that cannot colonize hedgerows between patches would also exhibit lower abundances in secondary regrowth. Temperate forest herbs are best understood as being adapted to a stable environment, as evidenced by low annual seed production, long pre-reproductive growth phases, and absence of long-term persistent soil seed banks (Bellemare et al., 2001). A lack of adaptations required for rapid recolonization of disturbed environments ensures the absence of such species in secondary forests. Peterken & Game (1984) found that the composition of secondary forests remain distinct from primary forests for hundreds of years post-regrowth. Thus, some changes in composition post-fragmentation may be irreversible. Furthermore, in

comparison to remnant forests, secondary forests exhibit an impoverishment in species diversity (Matlack, 1994; Singleton et al., 2001; Bellemare et al., 2002; Vellend 2004).

1.10 Plant species in remnant forests

Unlike secondary forest stands, remnant forests were not previously cleared of all forest vegetation. Yet, the resultant conditions of habitat fragmentation—habitat loss, smaller habitat areas, increased edge effects, genetic erosion, and reduced dispersal—directly or indirectly cause population losses in many plants in remnant forests. A rich literature supports this supposition that fragmentation causes reduced species diversity of forest plants (Dumortier et al., 2002; Falk et al., 2008; Honnay et al., 2005; Vellend, 2004); yet, species extirpation may not be immediate, and there are many confounding factors that alter the population dynamics and fitness of plant species. Larger patches are more likely to have greater soil heterogeneity in addition to a greater diversity of tree species; and thus a greater selection of microhabitats for herbaceous species (Honnay et al., 2005). Edge effects, as mentioned previously, change the forest microclimate, and allow for increased chemical inputs from agricultural fields (Murcia, 1995), while altering the dynamics of interspecies competition (Honnay et al., 2005; Murcia, 1995). Edge effects thus act to further decrease the forest area that is habitable for many forest herbs. Fragmentation has also been reported to decrease insect abundance and diversity, putting plants dependent on insect pollination at a disadvantage (Kearns et al., 1998). Extirpated species may also have fewer re-colonization events if dispersal is interrupted. Thus, forest fragments that exhibit connectivity to, or are in close proximity to other patches may receive propagules from neighbouring sites, thus renewing species diversity. Therefore, connectivity between fragments may ameliorate declines in species richness. Yet, the difficulties in making meaningful measurements of connectivity make this trend difficult to establish in empirical studies (Lindborg & Eriksson, 2004; Dumortier et al., 2002; Kolb & Dieckmann, 2004; Vellend, 2004).

There are few empirical studies showing that forest fragmentation may act to preserve species diversity. Kellman (1996) proposed that isolation may reduce the

spread of pathogens in a landscape and that in cases of high compositional turnover between fragments, fragmentation may preserve regional diversity.

1.11 Plant communities at different observational scales

Throughout the past century, the study of community-assembly has been a major theme in ecology (Ricklefs, 2008). Contemporary ecological theory describes species co-occurrence as a product of chance, speciation, dispersal, environmental factors, and biotic interactions (Götzenberger et al*.,* 2011). Though not mutually exclusive in operation, different drivers of assembly dominate different scales of observation (Götzenberger et al*.,* 2011), and appropriately, different theoretical frameworks have been established to understand assembly at different scales. The theory of island biogeography and the metacommunity concept are primarily characterized by neutral and spatial processes and thus important to an "island" or habitat fragment, and may not translate to finer scales of observation. Forests are heterogeneous environments and allow for the existence of numerous communities. A community is traditionally defined as a group of species co-occurring in space and time, often sharing similar environmental tolerances. Examples include species growing on limestone outcrops in forests, in well-drained slope communities, or species associated with vernal pools. We illustrate the differences in community assembly at the forest site and community level in Figure 1.1. In Figure 1.1, we define community as point-community, an estimation of community based on quadrat sampling (Ricklefs, 2008).

The relationship between regional and local species composition and diversity is partly explained by the species pool hypothesis, which states that diverse communities also have diverse species pools (Zobel, 1997). A species pool is a set of species that may potentially be found in a community, as determined by species range, environmental requirements, and history (Zobel, 1997). The concept can further be specified as the *actual species pool*: a set of species present in a community or continuous habitat (Herben, 2000). Many past studies have demonstrated positive correlations between local community diversity and SD at larger spatial scales (Caley & Schluter, 1997; Šímová, et al., 2013). Positive correlations between local and regional species diversity have

been interpreted as indicating the impact of regional factors on local communities; whereas, the absence of correlation has been interpreted as indicating the importance of ecological determinants of local diversity (Cornell & Lawton, 1992). However, it has also been argued that studies correlating the two diversity measures are flawed, and that a correlation is a by-product of sampling, since regional diversity will always be greater than local diversity (Herben, 2000).

1.11.1 Forest fragments

At the level of the forest patch, species diversity (SD) and community composition are mainly governed by dispersal to the patch, size of the patch and history (Vellend, 2004). Environmental heterogeneity within the patch also plays a role in community assembly and SD, but can be reduced to being a function of patch size, with larger areas having greater potential for environmental heterogeneity. Hubbel (2001) hypothesizes that community drift also acts at the site level and is a determinant of community composition. Ecological drift is analogous to genetic drift and describes random changes in community composition.

Dispersal to any given locality is dependent upon numerous conditions. Tischendorf and Fahrig (2000) outline migration rate in forest fragments as being a function of 1) the area of forest around a focal patch, 2) the number of emigrants from surrounding forests and 3) how non-forest barriers impede movement and cause mortality. Although competition and biotic interactions occur in forest fragments, they are generally not deemed important predictors of diversity at this scale.

1.11.2 Point-communities

The concepts of *community* and *niche* are central concepts in community ecology (Ricklefs, 2008). The term *community* has long been used to describe a group of species coexisting in space and time, with a scale that fits the question. Yet, the extent to which such a community exists in nature and can be used to define, describe, or delineate a *niche* is unresolved (Ricklefs, 2008). To emphasize the role of sampling in approximating communities, Ricklefs (2008) proposes that the term *point-community* be used to describe species coexisting and sampled in the same area. For clarification in

this paper, I will use community to describe a conceptualized community, and point-community in reference to communities in empirical studies.

Distribution of species in a continuous site depends on dispersal or movement within site, environmental filtering and biological interactions, including competition. Flinn et al. (2010) found that distribution of species with adaptations for short-distance dispersal depended on spatial patterns whereas long-distance dispersers were found along more narrow environmental gradients. Factors important at the community level for plant growth include: soil type, pH, soil moisture, light availability, and disturbance (Beatty, 2003). At this scale, productivity may also play a role. In a study of plant traits across a fertility gradient in a forest, higher species richness was found at more fertile sites (Naaf & Wulf, 2012). A rich literature also exists around the idea that competition prevents similar species from coexisting (Weiher, Clarke & Keddy, 1997). Therefore, not all potential species in a species pool are likely to be found together.

1.12 Unified biodiversity

Speculation that similar forces maintain species diversity and genetic diversity was first penned by Antonovics (1976), and has recently experienced a resurgence (Vellend 2004, Vellend & Geber, 2005, Vellend, 2005, Vellend, 2010). Vellend (2010) proposes that the three fundamental processes in population genetics: genetic drift, mutation, and gene flow, are analogous to the macroecological processes of: ecological drift, speciation, and dispersal. Many modern theories of species diversity, including island biogeography (Wilson & MacArthur, 1967), the metapopulation concept (Levins, 1969), neutral theory (Hubbel, 2001) and the individualistic concept of species association (Gleason, 1926) were inspired by fundamental concepts in population genetics, notably, the island model (Wright, 1931) and the neutral theory of molecular evolution (Kimura, 1983) which states that most of the genetic variation in a population is determined by mutation and genetic drift rather than selection. Vellend (2010) argues that ecology has been slow to embrace neutral processes, and that the field would benefit from a fundamental restructuring.

The synthesis of ecology and population genetics can give structure to ecological theory (Lawton, 1999) while allowing for a greater understanding of biodiversity (Vellend, 2010). The number of studies merging the two fields has recently increased, and of particular prominence are studies looking at species-genetic-diversity-correlations, or SGDCs (Vellend, 2003). A survey of recent SGDC results reveal: positive significant correlations (Vellend, 2004: He et al*.,* 2008; Lamy et al., 2013; Cleary et al., 2006; Papadoupolou et al., 2011; Struebig et al., 2011, Blum et al*.,* 2012; Wei & Jiang 2012); negative correlations (Wehenkel et al., 2006, Puscas et al, 2008) and no correlation (Odat et al*.,* 2004; Derry *et al.,*2009; Silvertown et al*.,*2009, Taberlet et al.,2012, Wei & Jiang, 2012). Since spatial processes are believed to generate correlations, the patterns are best observed in fragmented (natural or anthropogenic) or island scenarios where dispersal is limited (Vellend 2004; Cleary et al., 2006; Struebig et al., 2011; Blum et al., 2012; Sei et al.,2009; He et al.,2008; He & Lamont, 2010; summarized in Vellend, 2014). Positive correlations in SGDCs have been argued to be evidence of neutral processes (Etienne & Oloff, 2004; Vellend, 2005; Vellend & Geber, 2005; Lamy et al., 2013).

There is much to be gained from an increased understanding of the interplay between population genetics and community ecology, and also the unification of evolutionary and community ecological theory. As biodiversity increasingly becomes a focus of global conservation efforts (Cardinale et al*.,* 2012), an increased knowledge of the relationship between genetic diversity and species diversity allows for a more in-depth study of these two primary units of biodiversity. If the measures are correlated it would suggest that one can be a stand-in for the other in conservation contexts where sampling is costly and time-consuming (Vellend et al., 2014).

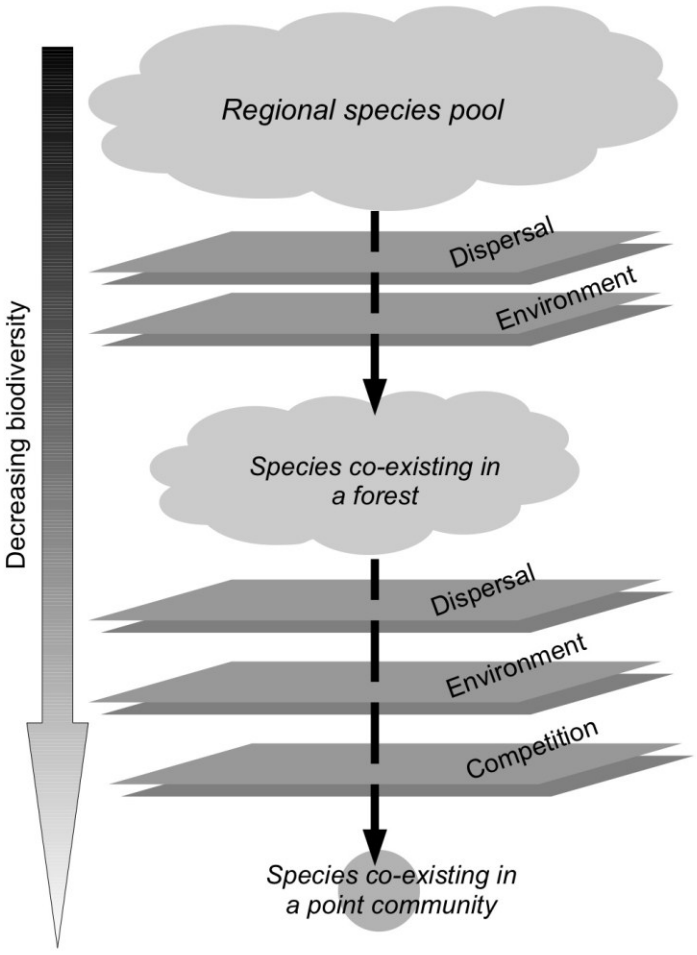


Figure 1.1: Diagram illustrating the factors affecting community composition at different levels.

CHAPTER 2

2.1 Introduction

Carex leptonervia (Fernald) Fernald is a common forest plant native to eastern North America with a range spanning from Québec to North Carolina. In addition to *C. leptonervia, Carex* section *Laxiflorae* contains 13 other species of forest plants native to eastern North America (Bryson & Naczi, 2002). Relative to the abundance and importance of species in section *Laxiflorae* to deciduous forest ecosystems, little has been published on their genetics, and beyond a study of *C. blanda* (Finch & Alexander, 2011), little has been published on their ecology, and life-history. The impressive diversity, similar physiology, shared ancestral history and distinct environmental preferences of species in section *Laxiflorae* further adds to the potential of these species to act as model systems in many ecological and evolutionary studies.

Simple sequence repeats (SSRs), commonly known as microsatellites, have previously been developed for a number of *Carex* species, including: *C. extensa* (Escudero et al*.,* 2010)*, C. kobomugi* (Ohsako & Yamane, 2007)*, C. limosa* (unpublished, our lab)*, C. macrocephala* (King & Roalson, 2009; King et al*.,* 2009)*, C. moorcroftii* (Liu et al*.,* 2009)*, C. rugulosa* (Ohbayashi et al*.,* 2008)*, C. scabrifolia* (Hodoki et al., 2009)*,* and *C. scoparia* (Hipp et al., 2009). These markers have been developed mainly for studies of migratory history and for assessment of breeding systems. Some of these loci have been tested for cross-amplification in distantly related species with success (Escudero et al*.,* 2010; King & Rolason, 2009; Hipp et al., 2009), yet none have been tested or developed for section *Laxiflorae*. Our objective in this paper is to develop microsatellite loci for *C. leptonervia* for use in a population genetic study (chapter 3 of this thesis).

2.2 Methods

The 82 microsatellites loci tested in this study came from three different sources. Since the SSR loci in this study were developed for other species, we first had to establish whether any SSR sequences were conserved in *C. leptonervia.* We tested loci that were previously unpublished, and also high-quality loci that were fully developed for two other *Carex* species. For use in our study described in chapter 3, each locus

developed had to cross-amplify in *C. leptonervia,* exhibit polymoprhism, amplify consistently, and have easily identifiable peaks in fragment analysis. Loci that were difficult to amplify or had ambiguous chromatogram peaks were not analysed further. SSR loci that met the above criteria were also assessed for Hardy-Weinberg Equilibrium (HWE) and linkage disequilibrium (LD).

2.2.1 Cross-amplification of *Carex lupulina, Carex limosa*, and *Carex scoparia* SSR loci

We assessed 68 putative SSR loci isolated in a partial genome sequence of *Carex lupulina* (454 Life Sciences, Branford, Connecticut), that had either di, tri, tetra, of penta-nucleotide repeats, and at least six tandem repeats in *C. lupulina* (E.R. Roalson, Washington State University, unpublished data). We also tested cross-amplification for the following 11 microsatellite loci developed previously for *C. scoparia:* S082, S180, S245, S047, S128, S175, S087, S119, S181, S102 and S177 (Hipp, et al., 2009); and for three microsatellite loci developed in our lab for *C. limosa:* CL88, CL101, CL113 (Ouborg & Waterway, personal communication; CL101 and CL113 are used in Escudero et al*.,* 2010).

All loci from the *C. lupulina* partial genome sequence were tested for amplification in four *Carex* species: *C. laxiflora*, *C. leptonervia*, *C. limosa*, and *C. lupulina*. We selected the first three species because they are of interest in our lab, and we used *C. lupulina* as a control. The *C. scoparia* loci (Hipp et al., 2009) and *C. limosa* primers were only tested for cross-amplification in *C. leptonervia.* To increase the probability of successful cross-amplification, we amplified at three different annealing temperatures (T_A): 54ºC, 58ºC, and 61ºC. After initial amplification success, further optimization of T_A was then determined using gradient PCR (see Table 1.1 for final T_A). DNA was amplified in 25µL reactions with the following reagents: 1xPCR buffer (New England Biolabs, $MgCl_2$ free), 0.13 mM of each dNTP, 2 mM $MgCl_2$, 0.18 µM of each forward and reverse primer, and 0.75U of *Taq* DNA polymerase (New England Biolabs), and 15-30 ng of genomic DNA. Thermal cycling was performed on a Mastercycler Pro S (Eppendorf) and also on a Gene Amp PCR System 9700 (Applied Biosystems) as follows: 95ºC for 4 min followed by 35 cycles of 94ºC for 45 s, T_A for 45 s, and 72ºC for 45 s, and a final elongation step of

72ºC for 10 min. To confirm presence of an SSR sequence in *C. leptonervia* we sent one or two amplified DNAs from each locus for sequencing. All sequencing was performed at the McGill University and Génome Québec Innovation Centre, Montréal, Canada. Chromatograms were visualized using the program Xplorer (dnaTools, Fort Collins, CO).

2.2.2 First round polymorphism testing using M13

We assessed polymorphism in 18-20 *Carex leptonervia* individuals from each of 20 populations in southern Québec, collected in June 2013 (details of collection described in chapter 3). DNA extraction methods were as follows: 50 mg of silica dried leaf tissue was ground using zirconium beads in the AutoGen Autogrinder48 (AugoGen, Holliston, Massachusetts). Ground tissue was then dissolved in a CTAB buffer (described by Doyle & Doyle 1987, but without β-mercaptoethanol) and purified using the Autogen 850 alpha DNA purification system (AutoGen, Holliston, Massachusetts) following the plant DNA protocol supplied by Autogen. We added the fluorophore 6-FAM to all amplified sequences following the M13-tail protocol outlined in Schuelke (2000). Addition of fluorescently labeled M13 tails is an economical two-stage process used to incorporate fluorophores into amplified DNA. We made a modified PCR cycle that had two different annealing temperatures to work with three different primers: a 6-FAM labeled M13 sequence, a forward primer with an M13 sequence tail, and an unmodified reverse primer. Throughout the first stage of the modified PCR, the SSR sequence was amplified, and in the second stage the M13 label was added to the forward primer sequence. In most cases the M13 sequence was added to the forward primer, but when addition of an M13-tail was predicted to cause secondary structure and primer-dimers, it was added to the reverse primer. To further avoid secondary structures and primer-dimers, we alternatively used two M13 sequences as needed for each locus: M13(-20): GTAAAACGACGGCCAGT and M13(-40): GTTTTCCCAGTCACGAC.

Samples were amplified in 37µL reactions with the following reagents: 1xPCR Buffer (New England Biolabs, $MgCl_2$ free), 2 mM $MgCl_2$, 0.12 mM each dNTP, 0.001 µM of M13 labeled forward primer, 0.03 µM M13-FAM labeled primer, 0.04 µM of reverse primer, and 15-30 ng of genomic DNA. The thermal cycling profile was as follows: initial

denaturation for 4 minutes at 94ºC; 94ºC for 45 s; T_A for 45 s; 72ºC for 45 s for 30 cycles, followed by 8 cycles of 94ºC for 45 s; 53ºC for 45 s; 72ºC for 45 s, with a final extension of 8 minutes at 72ºC (T_A values are given in Table 2.2).

Amplification products were dissolved in formamide and run on an ABI 3730 XL DNA Analyzer at the McGill University and Génome Québec Innovation Centre. Fragment analysis was conducted using STRand (Toonen & Hughes, 2011).

2.2.3 Testing for Hardy-Weinberg Equilibrium and linkage disequilibrium

We used fluorescently labeled forward primers, 6-FAM and VIC (Applied Biosystems) for each polymorphic locus (Table 2.3). Amplification procedures follow those outlined in the cross-amplification trials (for T_A see Table 2.3). Samples amplified with different fluorophores were pool-plexed (samples were combined to be run in the same lanes) to reduce genotyping costs. Genotyping procedures follow those outlined in the previous section.

We then tested seven promising loci for deviations from Hardy-Weinberg Equilibrium (HWE) and linkage disequilibrium (LD). We performed HWE and LD on a total of 467 individuals of *C. leptonervia* with 23-27 plants from each of the same 20 populations (Appendix 1). The LD test used is a test of composite linkage disequlibrium (Weir, 1996). We then used Fisher's method (Fisher, 1932) to test for global linkage for each pair of loci across all samples. LD was calculated using Genepop on the web: http://genepop.curtin.edu.au/ (Raymond & Rousset, 1995; Rousset, 2008).

We used an exact test of Hardy-Weinberg equilibrium with the assumption of heterozygote deficiency (Rousset & Raymond, 1995) and used a Markov Chain (MC) algorithm to estimate the p-value (Guo & Thompson, 1992). The parameters we used for the MC algorithm were 1000 dememorization steps, 100 batches and 1000 iterations per batch. HWE was calculated using Genepop on the web: http://genepop.curtin.edu.au/ (Ramond & Rousset, 1005; Rousset, 2008).

2.3 Results

2.3.1 Cross-amplification in *C. leptonervia*

In total, 29 of the 82 SSR sequences had successful cross-amplification in *C. leptonervia.* Twenty-five of the 68 loci identified in *C. lupulina* amplified in *C. leptonervia* as compared to 34 in *C. laxiflora,* 19 in *C. limosa,* and 25 in *C. lupulina* (Table 2.1). Fifteen of the loci that amplified in *C. leptonervia* possessed an SSR sequence (Table 2.2). For the 11 *C. scoparia* loci, three were polymorphic and cross-amplified in *C. leptonervia* (Table 2.3). Only one of the three *C. limosa* primers exhibited polymorphism and had a conserved SSR sequence (Table 2.3).

2.3.2 Hardy-Weinberg equilibrium and linkage disequilibirum

After Bonferroni corrections (modified p=0.0036), 22 out of 140 tests (7 loci x 20 populations) showed significant deviation from HWE and heterozygote deficiency (Appendix 2). Locus CL88 shows the most deviation from HWE with 9 populations out of HWE. Locus S175 shows 6 populations that are out of HWE. Some pairs (loci x population) could not be tested because loci were not polymorphic. For LD, after a Bonferroni correction (modified p=0.0036), only two loci (S175 and S082) were significantly linked across all populations (Appendix 3).

2.4 Conclusions

In total, after polymorphism testing, seven loci were found to have 2-5 alleles across the 20 populations of *C. leptonervia* (Table 2.3). Deviation from HWE may indicate inbreeding or genetic neighbourhoods. This finding fits with previous studies that found a high rate of selfing and evidence of inbreeding in *Carex* (Whitkus, 1988; King & Roalson, 2008). Inbreeding has also been found to predominate in self-compatible caespitose species (Bruederle et al., 2008).

The cross-amplification success of many distantly related *Carex* species shows potential for these loci for use in other species in the *Laxiflorae* section and potentially throughout the genus.

Locus	Forward primer sequence	Reverse primer sequence	*C. lupulina*			T_A			
			Motif	#SSR'S	# repeats	*lax*	*lep*	*lim*	*lup*
LUP1	GCTCCGACTCCTGTGATTGT	AAACTTCGACGAGGGCCTAT	AT	3	13	59	59	0	59
LUP5	CAAGGCATCGGCAAAATAAT	GGGAGAGACAGTTGGCATTT	CA	1	13	58	58	58	58
LUP6	AAGGACGTCAGGTTTTGTGG	GCAGTCTAGCTGCCTCTGCT	AT	1	13	0	54	0	0
LUP7	TCCGCAAGGGTACCTTTTTA	CAAGAGACGAGGGTGGTTGT	AT	1	13	54	0	0	0
LUP8	CCCGCATTGCAATTAAATCT	CAACTCACCGTAGCCCATCT	AT	1	13	0	0	0	58
LUP10	CCTCGCTACTGTCTGGGAAC	AACGATTTGAGGAGCTACCG	AT	1	13	58	58	0	58
LUP11	CAGTCCTTGTACACTGGGCA	TTGGTTTTGCTGCTGCTATG	AT	1	13	0	0	58	58
LUP12	ATTTCATCACCACCTGCCTC	TTTGTGTCAAGGATGCATTTC	AT	1	13	57	0	0	0
LUP13	TTGCGCTATCACAAAAGCAC	TGTTCAGCGGATTGTTTTCA	AT	1	13	0	54	0	0
LUP15	GTACAGGGACCACATTGCCT	GGATCAAATATGCGCTCGAT	AT	1	13	58	55	0	55
LUP16	TAGCTTTGAACCATCCTGGG	TTTAACTGGATTCCCGTGCT	TC	2	14	54	54	54	0
LUP17	TTGCCATTAAAGGGGTCAAG	CAGCTCTTTCCACGGTTCTC	TG	2	14	0	0	58	0
LUP18	ACACGTGTCATGCGTGAGTT	CCCAACGTGCTTAATAATGGA	AT	2	13	54	0	0	0
LUP19	AAAAATTGGAAAGTTCGGGC	TCCAATTCCTCTGCTTTGCT	AT	1	13	55	55	0	55
LUP21	GGCCAGAATTTACGCACTCT	TCACCTCCTCTATGCTTGTGG	GT	2	14	57	0	57	57
LUP22	AATTCAACATCCTCCCTCCC	TTTCATTTTACATGCCGCAC	TC	2	14	0	0	58	58
LUP25	GCTCACACGCTTCTTTTCCT	AAAGCGCTACAACGGTTCAG	TC	1	14	57	0	0	0
LUP27	GGCTTGTACAGCCCAAACA	ATGGGGTTGGATTTAGTGGA	AG	1	14	58	0	0	0
LUP28	CCAGTCCACTCTCAAGCAAA	GGTGGTGGTGACACATTCAA	AG	1	14	0	0	58	58
LUP29	CGATTTGTGCAGTCAGGTGT	AAGGGCCTCAAATGAAGACTC	AT	1	14	58	0	54	54
LUP30	CAAGCATGGGAAGACAGACA	CCTCACCACTTAGCCAGCTC	GA	2	14	59	0	0	59
LUP31	AGGCCCTTGCAAATGAAATA	ATACAGCACGTGATTGGCAC	TG	1	15	0	0	58	58
LUP32	TCATTTTTCAGCATCTCCCC	AGGAGGGGTAACGGTCCTAA	TG	1	15	0	0	0	58
LUP33	GTTGGCCACAAGTTCCAAAT	ATAAGAGGAGCTCCCATCGC	AAT	1	14	59	59	59	59
LUP34	TTTTACCACTATCCGCTCCG	TGGCATATCACCTTCCATGA	AAT	2	14	54	54	54	54
LUP35	CATTTTCCCGTGCAATTCTT	GCCAGACCTTCACCTTGTTC	AC	3	14	58	0	0	0
LUP36	CGTAAGATATCACGGCAGCA	CTCAAGCTCAAGCCCATTTC	AG	1	16	58	54	0	58

Table 2.1 a): Results of cross-amplification of 68 *C. lupulina* microsatellite (SSR) loci in *C. leptonervia* (lep), *C. laxiflora* (lax), *C. limosa* (lim), and *C. lupulina* (lup). A T_A value of 0 indicates that cross-amplification was not successful. The forward and reverse primer sequences, motif, #SSR's and # of repeats data comes from the *C. lupulina* genome scan (E.R Roalson, Washington State University, unpublished data). The table continues onto the next page.

Locus	Forward primer sequence	Reverse primer sequence	*C. lupulina*			T_A			
			Motif	#SSR'S	# repeats	*lax*	*lep*	*lim*	*lup*
LUP38	CTTCTCACGAAAGGACTGGG	ACGGAGGGGCTCTATGCTAT	ATAC	1	16	54	0	0	58
LUP39	TTTTACCTCCCGGTGAGAAG	CCGGAGGTACTTCGCATTTA	AAT	1	16	54	54	0	0
LUP40	GCCACCGTACTTGACCTCAT	TGAAGAAGGTGGTTTGGAGG	CAC	3	15	59	59	59	0
LUP41	ACAAGCGCATAATCCTACCG	CAAAACCCTAGAGGCTCCCT	AAT	1	18	54	54	54	0
LUP42	GCTGACAGAAGCTGCCAGTA	CACCCCATTCAGAAAGGAAA	TG	1	18	0	0	0	55
LUP49	TCCTGGACTCGTAAATGTTGG	AGTTTGATTCGTTTGCAGCC	TTA	1	21	0	0	0	58
LUP50	AATGGGACATTGAGCTCTCC	AATCTTACCTTTGACCCGGC	TTA	1	19	54	0	0	0
LUP52	AGAGAGGAAGAGGCAAAGCC	ATCTGTACCCCGCTATTCCC	AAT	1	11	58	0	0	0
LUP55	ATGTTGTGCCAAACTAGCCC	TGTATCGGCCTTTTGGTAGC	AG	1	11	58	0	58	0
LUP56	ACTATACATCAAACCCGCCG	GGTCAACGCGAATACAAGGT	TC	1	12	61	61	0	61
LUP59	GCCTCCTGCATGACCTACAT	AGTTCCCTACGGGTGATGTG	TG	1	12	58	58	58	0
LUP63	GAGAACCTGGACCTAAACCAA	CTGCTCTCATTGCCTCCATT	TC	1	12	58	58	0	0
LUP68	CGGGCAGAGTTATGGTGTTT	TTAATTGTTGCTGTCGCTGC	TA	1	12	0	0	58	0
LUP71	GAATCAGGTATGGACGCAGC	CCCTTATCCGCTGCATTTTA	CT	1	12	58	0	0	0
LUP74	TTTTCCTCTCTCTTGGCAGAA	CGGACTTTCCAAATTGTGCT	GA	1	12	0	0	0	58
LUP78	CAAAAGGGGTCCATAGGGTT	GATCACGGCTTGATGGAGAT	CA	1	12	58	58	58	0
LUP82	GATTCTAATCCCACCCAGCA	CATGCATGCCTCTGTTTGAC	AT	1	12	58	0	0	0
LUP83	CTTAATCCGACGCTCCATGT	CGCAAATCACAAAGATTCCA	AT	1	12	58	58	58	58
LUP84	CTACGGCCATGTTGGAAAAA	TTGGTTGATTGCACCAATGT	AT	1	12	0	0	0	58
LUP97	AGGGGTGTGTGTAGCCAAAG	GCACGTGCACCGTATCTCTA	GA	1	13	58	58	58	58

Table 2.1 b): Continuation of table 2.1 a). Results of cross-amplification of 68 *C. lupulina* microsatellite (SSR) loci in *C. leptonervia* (lep), *C. laxiflora* (lax), *C. limosa* (lim), and *C. lupulina* (lup). A T_A value of 0 indicates that cross-amplification was not successful. The forward and reverse primer sequences, motif, #SSR's and # of repeats data comes from the *C. lupulina* genome scan (E.R Roalson, Washington State University, unpublished data).

Locus	Genotyping results	#		% heterozygotes	T_A
		sent	alleles		
CL88*	good	17	6	0	65
LUP5*	good	24	4	0.13	58
LUP10*	difficulty scoring	18	N/A	N/A	61
LUP15*	difficulty scoring	18	N/A	N/A	63
LUP16*	difficulty scoring	28	2	0.31	54
LUP21*	difficulty scoring	18	N/A	N/A	61
LUP30*	low polymorphism, all homozygotes	16	2	0	58
LUP33*	low polymorphism, all homozygotes	12	1	0	58
LUP34*	amplification difficulty	18	3	0	54
LUP52*	good	24	4	0.95	57
LUP59*	amplification difficulty	18	N/A	N/A	56
LUP62*	difficulty scoring	18	5	0.11	58
LUP63*	good	20	5	0.32	58
LUP78*	difficulty scoring	18	N/A	N/A	57
LUP80*	amplification difficult	25	3	0.05	58
LUP89*	difficulty scoring	13	N/A	N/A	58
LUP97*	low polymorphism	21	3	0.06	58
S082	good	25	5	0.09	59
S102	good	27	5	0.17	54
S119	low polymorphism, difficulty scoring	25	2	100	59
S175	good	25	3	0.73	59
S245	difficulty scoring	18	N/A	N/A	57

Table 2.2: SSR loci sent for polymorphism testing in *Carex leptonervia*. An * indicates SSR was confirmed by sequencing in *C. leptonervia*. Genotyping results are classified as: "good"; "difficulty scoring"; "low polymorphism"; "all homozygotes"; and "amplification difficulty". The # of alleles is based on number of individuals in "# sent" column. The % of heterozygotes is based upon the total individuals that were successfully genotyped. The listed T_A is final optimized annealing temperature for these loci.

SSR	Motif	Primer sequences 5'-3'	# alleles	T_A (in °C)	GenBank Accession	Reference
S082	GAT	F: TGAGAACCCTAGGCAGATGG	9	59	EU369640	Hipp et al., 2009
		R: GGGGAAACAAGGTCGTTTAGA				
S102	ACAT	F: CGGAAAGAGGTAGCACAAGC	8	54	EU369642	Hipp et al., 2009
		R: AATCTGCTGATGCAACAATTTA				
S175	CTT	F: TATTGGGTGTGCGATTGAGA	9	59	EU369646	Hipp et al., 2009
		R: TCAGATCAGCCAAGTCATCG				
LUP5	CA	F: CAAGGCATCGGCAAAATAAT	9	58	KM215212	This paper
		R: GGGAGAGACAGTTGGCATTT				
LUP52	AAT	F: AGAGAGGAAGAGGCAAAGCC	6	58	KM215215	This paper
		R: ATCTGTACCCCGCTATTCCC				
LUP63	TC	F: GAGAACCTGGACCTAAACCAA	12	58	KM215214	This paper
		R: CTGCTCTCATTGCCTCCATT				
CL88	TC	F: GCTCAGTAGCTGATGCCAA	6	65	KM215213	This paper
		R: TGGAAAGCATCTCGTAGGAAC				

Table 2.3: Final SSR loci. The number of alleles reported in column 4 is based on sample sizes ranging from 15-24 individuals from each study site (20 in total). This was the same group sampled for HWE and LD (Appendix 1).

CONNECTING STATEMENT

In chapter 2 we described the development of seven polymorphic microsatellite markers. In chapter 3, we apply six of these markers to a case study examining the effects of forest fragmentation on species diversity and genetic diversity in the Montérégie in Southern Québec. We discarded the microsatellite loci LUP63 in our chapter 3 analysis due to the difficult interpretation of chromatograms during fragment analysis, which was not evident to us in our chapter 2 study. We apply the microsatellite markers to assess various measures of genetic diversity, G_{ST}, isolation-by-distance and population structure.

CHAPTER 3

3.1 Introduction

The unification of ecology and population genetics holds the potential to advance a more holistic and nuanced understanding of biodiversity. Groundwork has been laid for the conceptual unification of the two disciplines (Antonovics, 1976; Amarasekare, 2000; Agrawal, 2003; Vellend, 2003; Chave, 2004; Vellend & Geber, 2005) and unification has also been investigated empirically in many different ecological systems (see review in Vellend, 2010). In Vellend's influential 2004 paper, the relationship between the genetic diversity of *Trillium grandiflorum* and the species diversity of herbaceous forest communities was examined in upstate New York, and a positive correlation (SGDC) between species diversity (SD) and genetic diversity (GD) was first proposed. In the decade since this paper was published, numerous SGDC studies have been described for many different taxa and in many different systems, and many SGDC studies report positive correlations between the two diversity measures (Vellend, 2004; Cleary et al., 2006; He et al., 2008; Papadoupolou et al., 2011; Struebig et al., 2011; Blum et al., 2012; Wei & Jiang, 2012; Lamy et al*.,* 2013). In the SGDC context, correlations have been used to infer parallel processes occurring at the two levels of diversity.

Many SGDC studies are examined in systems affected by habitat fragmentation. Habitat fragmentation has been documented to cause both losses in species diversity (Haila, 2002) and genetic diversity (Young et al., 1996; Aguilar et al., 2008) and is considered a major threat to biodiversity (Fahrig, 2003). Habitat fragmentation disrupts spatial processes and creates discrete habitat units each with a specific *area* and degree of *isolation*. The relationship between habitat area and biodiversity is a long studied pattern in ecology (Arrhenius, 1921; Gleason, 1922; review in Drakare, Lennon & Hillebrand, 2006) and is often used to explain the loss of species diversity in fragmented landscapes. Furthermore, the loss of habitat, associated with fragmentation, is hypothesized to reduce the population sizes of many species. Smaller populations are at greater risk of genetic erosion, more vulnerable to catastrophic events and

consequently face higher risks of extirpation (Young et al., 1996; Honnay & Jacquemyn, 2007). The interruption of dispersal associated with habitat fragmentation further compounds the losses of genetic diversity in fragmented populations, as gene flow is reduced. The disruption of dispersal also has implications for species diversity, as repopulation of extirpated populations is less likely. Processes that act to decrease dispersal and gene flow between fragments in fragmented environments increase the *isolation* of fragments and processes that increase immigration and gene flow between fragments increase the *connectivity* of fragments. Many SGDC studies are undertaken in fragmented environments since fragment area and isolation can be easily characterized and are hypothesized to impact both fundamental units of biodiversity.

In this study, we sought to characterize the impacts of forest fragmentation in southern Québec on the genetic diversity of the forest sedge *Carex leptonervia* (Fernald) Fernald and to examine the relationship of GD to the SD of both the closely associated plant community and full forest stand. Our estimation of associated community is a close approximation to the diversity of the species actually growing in close proximity to *C. leptonervia*, whereas forest stand SD is an approximation of the overall collection of species found at forest sites inhabited by *C. leptonervia.* To our knowledge, we are the first to compare the GD of an organism to two levels of SD. The importance of different factors as drivers of SD and community assembly at different observational scales has been long established as a major paradigm in community ecology. At larger scales of observation, environmental variables in addition to spatial characteristics like site area and connectivity are critical, whereas at smaller scales the drivers of diversity are attributed to deterministic processes like environmental filtering and competition (Figure 1.1). Interestingly, relatively few SGDC studies have considered the role of the environment (Cleary et al., 2006; He et al., 2008; Sei, Lang & Berg, 2009; He & Lamont, 2010) and few of these have found the environment to be a driver of a positive correlation (Cleary et al., 2006; He et al., 2008).

In this study, we characterize the patterns of genetic diversity in *C. leptonervia,* a species that has never previously been the focus of a genetic study. We also examine the impact of forest fragmentation in the Montérégie on the diversity of forests where *C. leptonervia* grows. Finally, we test the hypotheses that GD is positively correlated to SD of the forest fragment as a whole as well as to the plant community directly associated with *C. leptonervia* and examine both spatial and environmental variables as potential drivers of the correlation.

3.2 Study system

We chose *C. leptonervia* (section *Laxiflorae*) as a model species to assess diversity in the Montérégie region in southern Québec (Figure 3.1). *Carex* is an important genus in North American deciduous forest communities and *Carex* is also very diverse in the Montérégie. For example, at Mont Saint-Hilaire there are records of as many as 60 *Carex* species (2010 Mont Saint-Hilaire species list, see Flinn et al., 2010). *Carex* section *Laxiflorae* is prominent in eastern North American deciduous forests with 14 species occurring from Québec to North Carolina. *Carex leptonervia* grows in mesic to wet-mesic forests (Bryson, 1980; Haines 2011), thrives in disturbed areas (Voss, 1972) and occurs frequently throughout the Monteregian forests.

The main mechanism for dispersal in *C. leptonervia* is gravity (Flinn et al., 2010), but there is also evidence that deer and birds infrequently ingest and disperse *Carex* perigynia (Bryson, 1980; Myers et al., 2003; Williams et al. 2008). Within *Laxiflorae*, Friedman and Barrett (2009) report that *C. laxiflora* is predominatly self-fertilized and that *C. plantaginea* has a predominantly mixed-mating system. Similar findings were also found throughout the genus in other caespitose species (Bruederle et al., 2008). Furthermore, like other forest plants, forest species of *Carex* are adapted to a stable forest ecosystem, and consequently have low seed production, and no specialized means of long distance dispersal (Honnay et al*.,* 2005). Thus, we believe that our characterization of *C. leptonervia* in the Montérégie can provide useful insight into how forest plants respond to fragmentation.

We sampled 20 forested sites in the Montérégie region of southern Québec (Figure 3.1) in the summer of 2013. The Montérégie is an administrative region to the south and east of Montréal, spanning from the St. Lawrence River to the US-Canadian border. The Montérégie is named for the Monteregian hills, a series of plutonic intrusions that formed 130-90 million years ago when the North American plate passed over the New England hotspot (Feininger & Goodacre, 1995). Upon European settlement in the 17th century, a majority of the arable land was deforested and repurposed for agricultural use, the majority of which is still in production. Today, about 30% of area is forested (Bélanger and Grenier, 2002) and 75% of those forests are < 25 ha in size, leaving little refuge for native plants and animals.

The Montérégie is part of the St-Lawrence lowlands forest region (Rowe, 1972), a region previously dominated by continuous deciduous and mixed forests. In a study of Haut-Saint-Laurent, 100 km west of our study area, Bouchard et al. (1989) report that a disproportionate amount of lowland forests have been cleared for agricultural use and that upland and morainic ridges remain forested. In Haut-Saint-Laurent Bouchard et al., (1989) described upland forests as dominated by sugar maple (*Acer saccharum*), bitternut hickory (*Carya cordiformis*), basswood (*Tilia americana*), American beech (*Fagus grandifolia*), ironwood (*Ostrya virginiana*) and white ash (*Fraxinus americana*) and the lowland forests by red ash (*Fraxinus pennsylvanica*), red maple (*Acer rubrum*), and white cedar (*Thuja occidentalis*). One field site, Mont-Saint-Hilaire (MSH) is well studied, and is considered the largest old-growth forest in the St. Lawrence River valley (Arii, Hamel & Lechowicz, 2005). MSH has a canopy dominated by *Acer saccharum*, *Fagus grandifolia, Tilia americana, Ostrya virginiana, Quercus rubra, Betula paprifera, Betula alleghaniensis, Quercus rubra,* and *Pinus strobus* and canopy composition on MSH has been found to be determined by soil moisture and nutrient regimes (Arii, Hamel & Lechowicz, 2005).

The overall goal of this study was to capture the diversity of *C. leptonervia* habitat in the region. The 20 forest stands we sampled were from a 1050 km² area

spanning N-S from Saint-Ours to Saint-Jean-sur-Richelieu (Figure 3.1). In addition to 18 mesic forest stands <500 ha, we also sampled two Monteregian hills, Mont-Saint-Hilaire (MSH) and Rougemont. These sites are distinct in their geology and are significantly larger than the other forests sampled in this study. MSH is also unique in having portions of old growth forest—characterized by having never been logged and distinct in having tall white pines and red oaks.

3.3 Methods

We visited 20 the forest stands at least twice in the summer of 2013 for the collection of three datasets: GD, SD site, and SD point-community. For the rest of this paper, we will use the term "point-community" to describe community; this term is defined by Ricklefs (2008) as an estimation of community based on quadrat sampling. All forest stands were first visited in late May and June 2013 for collection of *C. leptonervia* leaf tissue for later DNA extraction (GD dataset). We started the site flora lists at this time (SD site dataset). Upon completion of leaf tissue collection, we revisited sites in June and July to sample quadrats (SD point-community dataset) and complete the flora lists.

3.3.1 Selecting field sites and measuring site spatial characteristics

We defined a field site as a discrete forest stand that had a population of *Carex leptonervia* more or less throughout, with >50% canopy cover and with boundaries delineated by non-forest vegetation (usually agricultural land) or large roads. The forest canopy in each site was dominated by species in the genus *Acer* and forest composition followed local patterns in hydrology and environmental conditions. The canopy composition varied both within and across sites. Sampling across different forest types and environmental gradients allowed us to best estimate the diversity of *C. leptonervia* habitat throughout the Montérégie. Brief descriptions of each study site are provided in Appendix 4.

We used resources freely provided by the Système d'information écoforestière (SIEF) to identify potential forest stands in our study. SIEF is a comprehensive database

of forestry and ecological information collected by the government of Québec in the 1990's (Québec Ministère des Resources naturelles et de la Faune, 2003). We used SIEF data to classify forest age and area. Age class in the SIEF is based on the age of the oldest tress. We avoided younger forests that were completely cleared and that have regrown within the last 80 years by using only forest stands that had at least part of the stand with a SIEF classification of "indeterminate age >80 years". This is the oldest SIEF age classification aside from "old growth". We used this as an indication that all sampled forests may be forest remnants, at least in part, and have the potential for an ancestral population of *C. leptonervia.*

Since assessment of spatial characteristics is central to our study, we wanted to sample sites over a large range of size and connectivity measures (Table 3.1). We calculated size and connectivity prior to the field season by importing SIEF data into ArcGIS 10.1 (ESRI, Redlands, California) and the program Fragstats v.4 (McGarigal et al., 2012). Forest fragment area was calculated using ArcGIS10.1 using data provided by the SIEF. Connectivity is a measure of the connectedness of a focal forest site to other forests in the region. Measures of connectivity take into account the distance to other forests within a specified search radius and also the size of those forests. The Proximity Index (abbreviated as PROX; Gustafson & Parker, 1992) as calculated in the program FRAGSTATS v.4 was used to approximate site connectivity. PROX is a value ≥ 0; PROX equals 0 when there are no forests within the specified search distance. The specified search distance we used in the computation was 1 km, based on the potential for long-distance dispersal events of forest herbs by white-tailed deer (Myers et al., 2003).

3.3.2 GD dataset: tissue collection, DNA extraction, and SSR selection

Plant tissue for the population genetics study was collected in late May and June 2013 when *C. leptonervia* was flowering or had immature fruit so the identification could be readily confirmed. We collected 41-91 plants from each study site. The number of individuals collected depended upon the abundance of *C. leptonervia* at each site. Individuals were collected as they were encountered with a minimum distance of

2.4 m from one another. In areas where plants were especially abundant, individuals would be collected every 50 steps (approximately 25 m). During tissue collection, effort was made to cover the full geographical extent of each site. For small sites (<100 ha) it was possible to walk a majority of the area and forest edge. For sites >500 ha, sites were entered at different entry points to optimize sampling. Sampling effort (time spent at each site) was dependent on site size. Approximately, 4-6 hours were spent at sites <100 ha; 6-9 hours for sites 100-400 ha; and up to 24 hours spent at sites >400 ha. GPS points were recorded for every plant collected and used to relocate individuals for quadrat sampling.

We transferred fresh tissue samples to silica gel for optimal DNA preservation, and collected voucher specimens of flowering culms. We randomly selected 30 individuals from each study site for DNA extraction. DNA extraction methods follow those described in chapter 2 of this thesis.

We used six microsatellite (SSR) loci to assess genetic diversity: CL88, LUP5, LUP62, S175, S082, and S102 (Table 2.3). Microsatellite development was completed in our lab and is outlined in chapter 2. We also assessed the SSR LUP63 described in chapter 2, but removed this from our analysis because fragment analysis was difficult and inconsistent, resulting in >30% missing data for this locus.

3.3.3 SD point-community dataset: quadrat surveys and environmental data

To approximate the heterogeneity of the *C. leptonervia* community at each forest stand, we randomly selected seven *C. leptonervia* plants from which we had previously collected leaves for DNA extraction to act as centres for 4 m^2 circular quadrats for vegetation sampling, with the restriction that each plant came from one of seven approximately equal-sized fragments specified on a gridded map of the forest stand. When *C. leptonervia* was absent from one of the seven gridded segments, we randomly chose an individual from one of the other six segments. This methodology allowed us to sample *C. leptonervia* habitat throughout the extent of the site to capture the underlying environmental heterogeneity of each forest stand.

As most forest diversity is in the ground layer, we estimated percent cover of all herbaceous species regardless of height and all woody plants <1 m tall (Roberts-Pichette & Gillespie, 1999). Trees >1 m tall were recorded, but percent cover was not taken. We collected quadrat data in June and July when forest herb communities are at their highest diversity (Vymazalová et al., 2012) and during the fruiting phase of *C. leptonervia*. We estimated percent cover for each species within a 4 m^2 quadrat with the aid of a small circular grid representing 4% of the quadrat area. To prevent bias in estimates, the final estimates were arrived at by consensus of two botanists. We also performed a prism sweep for each quadrat to quantify the basal area of the most common trees growing in *C. leptonervia* communities. A wedge prism with a 2x factor was used, and methods followed Lee et al. (1998).

Soil pH and organic content are expected to be important variables determining community assembly. Accordingly, we collected ~250 g of soil from each of the seven quadrats at each site. Soil was sampled beneath the surface litter from the approximated *Carex* rooting depth (2-10 cm) from 6-10 locations around the focal plant and pooled for analysis. Soils were then analyzed for pH, and percent organic matter (%OM) using a loss-on-ignition (LOI) protocol (Bell, 1964) and for soil pH. For LOI analysis, soils were sieved in a 2 mm sieve, then dried at 106°C overnight. Samples were weighed and then ashed in an oven at 360°C for 4 hours before the final weight was measured. To analyze pH we made a 1 g/ml dilution slurry mix of soil and distilled water (Kaira, 1995) and measured pH using a Thermo Scientific Orion 9165BNWP combination Sure-Flow pH electrode (Fisher Scientific) on an Accumet pH Meter 915 (Fisher Scientific).

3.3.4 Site SD dataset: flora sampling

Cumulative flora lists were compiled to analyze the species composition of the forest stands in which *C. leptonervia* occurs. The presence/absence lists were started during the first visit to each site in May or early June and completed a month later in July when we re-visited sites to conduct the quadrat surveys. Spring ephemerals were

captured in our May and June surveys and later-blooming species were apparent during the July surveys. Using aerial photos collected from Google Earth and a GPS to track our progress, we were able to survey the majority of deciduous forest types (identified via air photo interpretation) at a site. A flora checklist of 194 common local forest species was used to make data recording efficient.

At least some individuals of *C. leptonervia* were consistently found near the forest edge so we prioritized sampling within 50 m of the forest edge at each site. We also made multiple transects through the interior of each forest as we surveyed the *C. leptonervia* population and re-located individuals using a GPS for quadrat sampling. For small forests <200 ha we sampled the whole forests thoroughly, aided by the aerial photos and GPS. For large forests we spent multiple days sampling and we re-entered the forest sites on different days from different entry points to ensure we covered the greatest area possible. We thoroughly investigated unique site features, including silver-maple swamps, and topographically diverse areas.

The sampling method was time-based and the time spent at each site was dependent on the size of the site, with more time spent at larger sites. We calculated sampling effort by dividing the time spent at each site by the size of each site (Table 3.1; Appendix 5). Unknown plants were collected for more detailed examination and identification. When possible, plants were identified to species, and when this was not possible we identified specimens to genus. For the small proportion of individuals we could not confidently identify to species, we recorded them as unknown and gave them unique identifiers (example: “unknown species 1”). Species nomenclature follows VASCAN (Brouillet et. al., 2010).

3.3.5 Data Analysis

3.3.5.1 Assessing GD and genetic structure

Genetic diversity of *C. leptonervia* populations for all study sites was measured as allelic richness (AR), observed heterozygosity (H_O) and gene diversity (H_E) and was calculated using R package ‘mmod’ (Winter, 2012). Since sample size biases calculations

of AR (Leberg, 2002), we used a standard 21 individuals per population, removing individuals to minimize missing data per locus to <10%. Assessments of all populations for Hardy-Weinberg equilibrium (HWE) using an exact test (Raymond & Rousset, 1995) and linkage disequilibrium (following Weir, 1996) have already been presented in chapter 2.

We ran three different tests to detect genetic structure and population differentiation of *C. leptonervia*. We assessed genetic differentiation using Nei's G_{ST} (Nei, 1973) in the R package 'adegenet' (Jombart, 2008). We assessed isolation-by-distance (Wright, 1943) by testing for correlation between genetic distance and geographic distance at each study site. To do this, we tested for correlation between a genetic distance matrix (using Nei's distance; Nei, 1972) and the geographic distance between sites (in km) using a Mantel test (Mantel, 1967) in R using the package 'ade4' (Dray et al., 2007). The geographic distance matrix was calculated in R using the program 'fields' (Fields Development Team, 2006). We tested genetic structure of *C. leptonervia* in the Montérégie using a Bayesian clustering method in the program STRUCTURE version 2.3.4 (Pritchard et al., 2000). STRUCTURE implements a Markov Chain Monte Carlo (MCMC) method to estimate K (the number of genetic clusters; Pritchard et al., 2000), where genetic clusters are characterized by a unique set of alleles. We used an admixture and correlated allele frequency model to estimate K, and ran tests for K=1 to K=n +1 (n=number of study sites). We ran 10 independent runs for each value of K (100, 000 burn-in, 1,000,000 MCMC runs) and used the ΔK Evanno method to select the optimal value of K (Evanno et al., 2005) in the program Structure Harvester v0.6.93 (Earl & von Holdt, 2011). We also calculated the admixture proportions of each study site belonging to each K cluster by averaging the results of 10 independent runs.

3.3.5.2 Species diversity and community assembly

We used only species richness (SR) to measure SD at the site level. For point-communities we used SR and the Shannon diversity index (SHAN), which also

incorporates abundance estimates, to measure SD. We concatenated the raw abundance data from the seven individual quadrats to give one value of Shannon diversity and species richness per site.

To analyse species composition, we used two ordination methods: principal coordinate analysis (PCoA) and canonical correspondence analysis (CCA) using R package 'vegan' (Oksanen et al., 2013). In both cases, quadrat data was first Hellinger-transformed to dampen the influence of abundant species (Legendre & Gallagher, 2001). A Bray-Curtis dissimilarity matrix was then calculated based upon this Hellinger-transformed data. PCoA allowed us to observe differences in species composition and also the most dominant species for both sites and quadrats. We also tested whether soil characteristics could explain the canopy composition of *C. leptonervia* using a CCA. We constrained the CCA with the two soil characteristics we measured for each quadrat: pH and %OM.

3.3.5.3 Linking diversity to spatial and environmental factors

To examine the role of spatial and environmental processes in determining diversity, we ran a series of multiple regressions. For the analysis of GD we ran two series of three independent multiple regressions using the explanatory variables log (ha) and log (PROX) for each measure of GD (H_E, H_O, and AR). For the GD analyses, we ran two regression analyses: one with all 20-sites and a second with 16-sites. In the 16-site analysis we removed sites that did not follow the well-documented pattern of increasing genetic variation with increasing population size (reviewed in Leimu et al., 2006). In this study, we used forest stand area to approximate *C. leptonervia* population size at each site.

For the analysis of site SD, we ran a backward multiple regression with the variables log (ha) and log (PROX) and included a measure of sampling effort at each site (Table 3.2). All 20 sites were included in SD regressions. For the analysis of point-community SD, we ran a backwards multiple regression on the SD of concatenated quadrats with the explanatory variables: CV (%OM), CV (pH), log (PROX), log (ha), site SR,

and community extent. Community extent is a variable we created to describe how representative *C. leptonervia* communities were of the forest stand as a whole. We calculated this value by dividing the SR of concatenated quadrats by the total SR for the site flora list. A site with a higher index of community extent has *C. leptonervia* occurring with a greater proportion of plant species in the forest stand, whereas in a site with a low index of community extent, *C. leptonervia* may be restricted to a relatively small portion of the site and co-occurs with a small number of the overall species (Table 3.2). A low index of community extent may also occur if a forest had species-rich areas different from the areas of the site where *C. leptonervia* grows. We included the coefficient of variation (CV) of pH and %OM as indices of the heterogeneity of the *C. leptonervia* habitat at each site. For each regression, we used an Akaike information criterion (AIC) approach to determine the best regression model. For each final model, we also ran a Kolmogorov-Smirnov test to confirm that residuals had normal distributions.

3.3.5.4 Testing for correlation between SD and GD

We used a Pearson correlation to compare the three GD measures (H_E, H_O, AR) to the three SD measures (SR site, SR quadrat, and SHAN quadrat). In this analysis, all quadrat SD measures are for concatenated quadrats. We ran two series of pairwise tests (three GD measures x three SD measures): a 20-site series where all sites were maintained in the analysis, and a 16-site series. In the 16-site analysis we removed four sites from the analysis after data exploration and the identification of sites that did not follow the pattern of increasing GD and forest area exhibited by the other 16 sites in the GD study. We identified and removed Site 8 (Mont Saint-Hilaire), Site 4, Site 3 and Site 26, which did not conform to a pattern of increasing GD with increasing site area (discussion in 3.3.5.3).

3.4 Results

3.4.1 Genetic Diversity

The degree of polymorphism varied across GD measures and study sites (Table 3.3). The mean H_E across sites was 0.3189 (SD ± 0.07449), the mean H_O was 0.2569 (SD ± 0.09503) and the mean AR was 17.35 (SD ± 3.021). A linear regression of all 20 populations revealed no significant relationship between spatial variables and GD measures (Table 3.4; Figure 3.2a).

Upon the removal of four sites (3, 4, 8 and 26) that did not conform to the common pattern of increasing GD with increasing population size (estimated via forest area), a relationship between forest area and H_E was revealed. A linear regression of area and H_E showed that forest area explained 46% of the variation in H_E, and that sites with larger areas exhibited higher H_E (p=0.004; Table 3.4; Figure 3.2b). Forest area also explained 20% of variation in AR (p=0.082; Table 3.4; Figure 3.10b). There was no significant relationship between H_O and forest area (r^2=0.033, Table 3.4, Figure 3.10b). The inclusion of the variable PROX did not improve the fit of the regression models and was removed from the final model.

Nei's G_{ST} was equal to 0.167 (Nei, 1973). This is a low value and reveals that many of the same alleles are found in many populations. The IBD test revealed that geographic distance between sites (Appendix 8) does not explain the genetic distance (Appendix 7) of the sites (Monte-Carlo test, p=0.366, 999 replicates). The analysis of genetic clustering (Pritchard et al., 2000) revealed the presence of three genetic clusters as determined by the ΔK method (ΔK=39.46; Figure 3.3; Evanno et al., 2005). The proportion of each study site (Table 3.5) belonging to each genetic cluster is illustrated in Figure 3.4. We also found evidence of admixture in individual plants. The bar plot in Figure 3.4 depicts the proportion of each individual in each site belonging to each genetic cluster.

3.4.2 Species diversity of point-communities

In the assessment of *C. leptonervia* point-communities, we found 209 species in a total of 140 quadrats across 20 forest stands. The SR ranges from 4-31 for each individual quadrat and 42-60 for concatenated quadrats from each site. Apart from *C. leptonervia,* the most common forest plants in the quadrats were *Rubus pubescens, Onoclea sensibilis, Maianthemum canadense, Impatiens capensis, Athryium filix-femina, Acer saccharum, Prunus virginiana* and *Aralia nudicaulis*. Other common *Carex* species growing with *C. leptonervia* include *C. intumescens, C. arctata* and *C. gracillima.* To explain variation in SR and Shannon diversity (SHAN) of the concatenated quadrats, we ran two independent backward stepwise multiple regressions with the explanatory variables log (PROX), log (ha), community extent, CV (pH) and CV (%OM). We ran two regressions with the same input variables to explain variation in SR and Shannon diversity (SHAN) of concatenated quadrats. The best model to explain SHAN only contained the variable community extent, and was significant (p=0.004), explaining 38% of the variation in SHAN (Figure 3.7). When we removed the two outliers MSH and Rougemont, we found that log (SR) explained 40% of the variation in SR of concatenated quadrats (Table 3.4; Figure 3.7).

The PCoA of the Hellinger-transformed matrix of quadrat data (Figure 3.8) shows spread along both primary and secondary axes. Interpretation of spread however is complicated by the large number of quadrats (140). Many sites are heterogeneous and have quadrats with different local environmental conditions, so we found little clustering of quadrats from within each site. Analysis of the position of species on the ordination plot shows an apparent moisture gradient along the first axis. Common species of well-drained mesic forests, including *Acer saccharum* and *Prunus virginiana* are positioned at the positive end of the primary axis. Towards negative values of the primary axis are species more common to wetter mesic forests, including *Onoclea sensibilis, Arisaema triphyllum,* and *Maianthemum canadense.* Species that occurred more infrequently clump together in the middle of the plot. The first axis explains 10%

and the second 7% of variation in community composition among the 4 m^2 quadrats (Appendix 6).

Carex leptonervia grows in many different forest types and with 36 different tree species (Appendix 9). The canopy of point-communities was dominated by *Acer saccharum, Acer rubrum*, *Fraxinus pennsylvanica*, *Abies balsamea* and *Ulmus americana*. More infrequently, we also found *C. leptonervia* growing in wetland swamps dominated by *Fraxinus nigra* and *Acer saccharinum.* We performed a canonical correspondence analysis (CCA) to assess the influence of pH and %OM in the differentiation of canopy communities (Figure 3.9). Moisture-tolerant species such as *Betula alleghaniensis, Fraxinus nigra, Tsuga canadensis, Acer rubrum,* and *Abies balsamea* are found in the direction of increasing %OM. In contrast, dry-fresh forest species such as *Ostrya virginiana, Acer saccharum,* and *Carya cordiformis* are found in the direction of decreasing %OM. We were unable to discern a logical pattern of assembly along the pH gradient. In addition to these groupings of species there are many outlier species with infrequent occurrences. Together, the two environmental gradients explained only 2.18% of the total inertia of the CCA (Appendix 6).

3.4.3 SD at the site level

A total of 418 species were observed across 20 forest stands. Site 4 was the least diverse with 93 species and MSH (Site 8) was the most diverse site with 216 species (Table 3.1; for site description see Appendix 1). We observed strong relationships between species richness and forest area using our site flora lists (Table 3.4; Figure 3.5). The regression results showed that larger sites had greater SR, and that connectivity was not an important determinant of site SR.

The PCoA ordination plot (Figure 3.6) of the flora data shows that MSH (Site 8) and Rougemont (Site 42) form a cluster and are clearly differentiated from the St. Lawrence lowlands forests. Sites 4, 18, and 20 are small well-drained mesic forests <75 ha dominated by *Acer saccharum* and have the highest positive values on the first axis. Sites 40, 34, and 30 are small wet-mesic forests dominated by *Acer rubrum* with

networks of swamps throughout and share low negative values on the first axis. Site 7 is positioned away from the other sites and was also unique in not having any *Acer saccharum*. The ordination generally illustrates a pattern of small well-drained mesic forests, small wet-mesic forests, and also many large heterogeneous forests with both wet and well-drained portions (Sites 5, 26, and 31). The first 5 axes are important in explaining 60.26% of the variation (Appendix 6).

3.4.4 Correlation between SD and GD

In the 20-site analysis we found no significant correlation between GD and SD (Figure 3.10a). Correlation results between GD and SD measures are reported in Table 3.6. However, in the 16-site analysis (Figure 3.10b), where Sites 3, 4, 8, and 26 were excluded we found two significant positive correlations between site SR and H_E ($R=0.822$, $p<0.001$, Figure 3.10b) and site SR and AR ($R=0.653$, $p=0.006$, Figure 3.10b). We also found similar positive trends in five other pairwise comparisons of SD x GD (H_O x SR site, SR quadrat x H_E, SR quadrat x H_O, SHAN quadrat x AR and SHAN quadrat x H_O). In every case, our p-values were consistently lower in the 16-site analysis (Table 3.6).

3.5 Discussion

Overall, we found *C. leptonervia* to be an "ecologically flexible" (term by Kellman, 1996) and common species in the Montérégie. We base this on the impressive diversity of associated species, the large range of environmental conditions across which it grows, and its common presence in forest stands. Forest fragmentation may even create conditions favourable to *C. leptonervia,* as demonstrated by its common occurrence near the forest edge (personal observation). The genetic analysis similarly depicts *C. leptonervia* as a species that is mobile throughout the Montérégie, as evidenced by low genetic differentiation among sites and by the assessment of population structure.

3.5.1. Characterization of *Carex leptonervia* genetic diversity in the Montérégie

This study was the first to assess patterns of genetic diversity in *C. leptonervia* and revealed many interesting findings. Here we compare the patterns of genetic diversity found in *C. leptonervia* to other genetic studies in the genus *Carex.* In an

analysis of previous *Carex* allozyme studies, Bruederle et al. (2008) reported a low mean H_E (0.043) for inbreeding, caespitose species and high genetic differentiation among populations. They also suggested that rhizomatous *Carex* species, on average, have higher heterozygosity than do caespitose species. Surprisingly, we did not find the same patterns in the caespitose species *C. leptonervia* using microsatellite markers, but these discrepancies may be due to differences in molecular markers used. Measures of heterozygosity in other *Carex* microsatellite studies are summarized in Table 3.7. The average observed heterozygosity we report in this study (H_O=0.26) is higher than in the few other *Carex* microsatellite studies where this measure is reported. Average gene diversity in *C. leptonervia* populations (H_E=0.32) is also high compared to other *Carex* species (Table 3.7). Carex *leptonervia* had higher gene diversity than two rhizomatous species *C. macrocephela* (King & Roalson, 2008) and *C. moorcroftii* (Liu, Wei & Dong, 2009), in contrast to the patterns based on allozyme data noted above. Differences between gene diversity in *C. leptonervia* and *C. macrocephela* can probably be attributed to the rarity of *C. macrocephela,* which is considered at-risk throughout most of its range (King & Roalson, 2008). Additional data on microsatellite variation in a larger selection of *Carex* species are needed to better assess patterns of differentiation.

The relationship between population size and genetic variation is well noted in literature (reviewed in Leimu et al., 2006). We used the forest stand area to approximate *C. leptonervia* population size, with the assumption that larger forests had greater potential to have larger *C. leptonervia* populations. This base assumption, thought convenient for our comparison of GD to SD, may not have been accurate and may explain anomalies observed in our data. In this study, we found that in the majority of sites, H_E, AR, and H_O were positively correlated with forest area. Thus, we found evidence that most populations in smaller forest stands had impoverished genetic diversity, which is in line with other findings of plant species genetic diversity in fragmented forest areas (review in Young et al., 1996; Aguilar et al., 2008). However, we

also found a number of sites with unexpectedly high and low values of GD that did not follow this overall trend.

Only the 16-site regression analysis found a positive relationship between GD and forest area (Table 3.4). A discussion of the unique attributes of the four sites we removed in this analysis follows. Site 3 has the highest values of H_E (0.503) and AR (25) while being intermediate in size (40.28 ha). Site 3 is also only 1.69 km away from Site 4 which also has high values of H_E (0.396) and AR (25), given its small size of 6.30 ha. Other studies have found that the presence of a persistent seed bank may mitigate the loss of GD in small and isolated plant populations (Honnay et al., 2008). Land-use history plays a major role in determining the presence of a seed bank, and it has been found that forests that were previously ploughed in agricultural use have impoverished seed banks (Singleton et al., 2001; Vellend, 2004). Since sites 3 and 4 are in close proximity, it may be that they share a similar land-use history that has resulted in genetically diverse and preserved seed banks. Unfortunately, we can only hypothesize this explanation, since we do have data on the specific farming practices at these two forests.

We also identified populations in large forests that had uncommonly low values of GD. We were surprised to find a consistently low GD at MSH (Site 8) and Site 26 given their large size. MSH, though a nature reserve, has very sparse groundcover and a correspondingly low value of point-community SR (Table 3.2). We hypothesize that the intensive white-tailed deer activity at this site is keeping the groundcover sparse and the *C. leptonervia* population small. Though *Carex* perigynia are not thought of as a choice food for deer, they have been found in deer scat (Myers et al 2004; Williams et al., 2008). We also observed a comparatively small *C. leptonervia* population at Site 26 and suggest that this population may also be kept small by deer herbivory and various disturbances including timber production and sugar-bush activity.

3.5.2 Interacting *Carex leptonervia* populations

Comparison of the values we obtained for G_{ST}, isolation by distance (IBD) and population structure to other *Carex* microsatellite studies is difficult because only a few

studies report similar analyses. The G_{ST} value for 20 populations of *C. leptonervia* in the Montérégie is 0.16. Liu et al. (2009) reported a comparatively high G_{ST} of 0.66 for nine populations of the rhizomatous species *Carex moorcroftii* in the Qinghai-Tibet plateau. The area sampled in the Quinghai-Tibet plateau was considerably larger than where we sampled in the Montérégie, making comparison difficult.

Traditional IBD models assume a neutral island model of equally exchanging populations (Meirmans, 2012) and patterns of IBD are typically found in fragmented forests when the matrix equally impedes dispersal to all fragments (Culley & Grubb, 2003). Yet there are also many exceptions to this pattern, most notably when there is a hierarchical population structure in the landscape. Hierarchical population structure occurs when not all populations equally exchange propagules. Unequal exchange is often a result of historical or geographic factors, and may be a result of river and watershed boundaries (Maki, Michiko & Inoue, 1996; Ellis, Weis & Gaut, 2007) or differing patterns of postglacial recolonisation (Taberlet et al. 1998; Tribsch, Schönswetter & Steussy, 2002). The structure analysis (Figure 3.4; Table 3.5) reveals three intermixed genetic clusters in the landscape, and thus a pattern of IBD is unlikely since there is geographic isolation of these genetic clusters (with the exception of cluster "C"). The northern portion of the study area is dominated by genetic cluster "B" (Figure 3.4), whereas the southern half is dominated more by genetic cluster "A". Genetic cluster "C" plays a relatively smaller role in the landscape but dominates Site 6 and Site 3, which are not close together. Each genetic cluster is present at each forest stand but the proportion of each population belonging to each cluster varies.

Differences in the genetic make-up of each site may be attributed to genetic barriers in the landscape and genetic drift. Sites closer together should have a more similar make-up and this is mostly what we see, with the notable exception of sites 3 and 6, which are dominated by genetic cluster "C". There are many explanations that may account for the unusual pattern exhibited at sites 3 and 6. The two sites though geographically far from one another have a similar species composition (Figure 3.6) and

are swampy sites dominated by *Acer rubrum* and *Fraxinu*s species (Appendix 4). It is possible that some of the SSR alleles dominant in cluster "C" are linked to phenotypes that are adapted to swampy deciduous forests. Furthermore, had it been possible to sample more forests in the region, it is possible that a more nuanced pattern of the genetic variation of *C. leptonervia* in the Montérégie would have been revealed.

3.5.3 Plant species diversity in the Montérégie

The forests we sampled in the Montérégie have high species diversity. In total, we found 418 species across twenty sites, with SR ranging from 93-216 per forest stand. Another recent study of plant diversity in the Montérégie (Larouche, 2013) found only 177 plant species across 52 forest stands but only three 10 x 10 m plots were sampled at each site in that study. We identified 56 tree species across all of our field sites and found 35 trees growing in the *C. leptonervia* communities (Appendix 9). Similar values were found by Ziter (2013) who recorded 52 trees in a tree diversity study in the Montérégie. A species list previously compiled for MSH from past detailed hectare-by-hectare diversity surveys (Maycock, 1961; Bell, Lechowicz & Waterway, 2001; Bell, Lechowicz & Waterway, 2006) revealed a total of 620 vascular plant species (Flinn, unpublished; see also Flinn et al, 2010). We spent six days sampling MSH and found 216 plant species (35% of 2010 species list). Yet, our value is similar to what was previously found in other studies at MSH that used different sampling methods. A study sampling upland quadrats found 215 species (Gilbert & Lechowicz, 2005) and a study of wet-mesic quadrats at MSH reported 280 species (Flinn, Lechowicz, & Waterway, 2008). The Montérégie also appears to be more species rich than other regions with a similar land-use history. Vellend (2004) sampled forest herbs in upper New York State and found that SR values calculated from a combined 750 m^2 area and multiple transects only ranged from 13-53 in each forest.

In the assessment of point-communities, we found that SR varied from 4-31 for individual 4 m^2 quadrats, and 42-60 for concatenated quadrats (representative of a 28 m^2 area). This range in SR is typical of forests in the region. Marchand & Houle (2006)

sampled the forest edge of forests in Southern Québec and found SR ranged from 8-11 species in 2 m^2 quadrats, while Gilbert & Lechowicz (2005) sampling at MSH found SR of 50 m^2 plots ranged from 7-44. This also suggests that the diversity of *C. leptonervia* communities is typical of other understory communities in the region. Surprisingly, the SR of point-communities at MSH was among the lowest values we found across all 20 forest stands (mean SR=11.43). Relative to the size of this site, the population of *C. leptonervia* was small and relatively sparse at MSH, and in general the understory layer was depauperate.

The ordination plots depict variability in the species composition at both the site and point-community level (Figure 3.8; Figure 3.9). It is possible that seven quadrats at each site were too few to capture the diversity of *C. leptonervia* communities within each forest stand, and that more quadrats would reveal *C. leptonervia* growing in proximity to more species. The stratified random selection design we used may also have increased the variance among quadrats at each site. The total number of species we recorded in point-communities was 45% of the total species we recorded across all sites. The total number of trees we recorded in the *C. leptonervia* communities captured 65% of the diversity we found at the site level. We consistently found *C. leptonervia* growing along moisture and light gradients and found it in forest edges, in well-drained mesic forests, and in swamps. Roy and de Blois (2006) also reported *C. leptonervia* in treed hedgerows. A similar diversity of habitat was reported for *Carex blanda,* a close relative of *C. leptonervia,* also in *Carex* section *Laxiflorae.* In a study on *C. blanda* habitat preferences in Kansas, Finch and Alexander (2011) reported *C. blanda* growing in equal densities in woodland, edge, and grassland habitat. These patterns depict *C. leptonervia,* like other species in *Laxiflorae,* as an ecologically flexible species that can grow in a variety of Monteregian forests communities.

We ran a CCA of the point-community canopy data to test whether tree community composition occurred along an environmental gradient. Unexpectedly, we found that the constraining variables we input (pH and %OM) were weak in explaining

variation in species composition (2% of variation explained). These findings are contrary to many studies that found the environment to be a strong predictor of community composition (Gilbert & Lechowicz, 2004; Arii et al., 2005; Jones et al., 2008; Flinn et al., 2010). It is possible that the inclusion of additional abiotic and biotic variables (Weiher, Clarke & Keddy, 1998) may improve the amount of variation in community composition explained in our analyses. Other abiotic predictors that may affect community composition include topography, solar radiation, disturbance, and soil fertility (Gilbert & Lechowicz 2004; Arii et al., 2005; Flinn et al., 2010; Naaf & Wulf, 2012). Land-use history of sites may also explain forest community composition (Hibbs, 1983; Bellemare, Motzkin & Foster, 2001; Vellend, 2004).

3.5.4 Forest Isolation

We expected isolation to have a greater influence on patterns of GD in *C. leptonervia* in the Montérégie. The minimal role of isolation in GD studies in fragmented systems has previously been attributed to extinction debt (Honnay et al., 2005). Extinction debt describes a scenario where the distribution of alleles reflects the historical landscape configuration; in other words, not enough time has passed for local extirpations to occur in habitat fragments and patterns of diversity may not reflect current connectivity. The Montérégie region has experienced ongoing deforestation since the 17th century, so it is unlikely that our results are due to an extinction debt. An alternative hypothesis is that white-tailed deer and various birds may be moving *C. leptonervia* propagules (and other forest herbs) around the Montérégie. Two recent studies in the north-eastern United States report the presence of *Carex* perigynia in deer scat and even report germination of excreted *Carex* perigynia (Myers et al., 2004; Williams et al., 2008). At the turn of the twentieth century deer were nearly driven to extinction in eastern North America, but with the introduction of conservative hunting limitations and extensive forest regeneration from abandoned agricultural fields, deer populations have increased to a point where they are very common and have pest status in southern Québec (Richer et al., 2005). Thus, even if deer infrequently move *C.*

leptonervia around the landscape, it may be enough to prevent genetic drift and to maintain genetic diversity in spatially disjunct populations. In addition to deer, Bryson (1980) includes bobwhite quail, American turkey, and various songbirds as dispersers of *Carex*, but the frequency of these dispersal events is unknown. Since we found that isolation had little impact on GD and SD in the Montérégie, we conclude that isolated forests play an integral role in preserving forest plant biodiversity.

3.5.5 Correlations between SD and GD in fragmented systems

Our initial correlation tests between GD and SD did not reveal any significant correlations between the two diversity measures (Table 3.6; Figure 3.10a). It is only upon exploration of these findings and the removal of four populations that exhibited an inconsistent relationship to forest area and GD (discussed in section 3.5.1.1; Table 3.6 Figure 3.10b) that a positive correlation between seven pairwise GD and SD measures was revealed.

We found that in the 16-site analysis, the SD of the *forest stand* has a significant positive correlation to GD. Interestingly, in both the 20-site and 16-site analysis, the SD of the *point-communities* was not a strong correlate with GD (Table 3.6; Figure 3.10a; Figure 3.10b). There are multiple interpretations of this result. The site flora list may have been a more accurate approximation of *C. leptonervia* community diversity than the point-community list. It is also likely that seven quadrats were not enough to fully capture the point-community diversity. Alternatively, it may be that we did not find a strong correlation because environmental and biotic processes that are commonly hypothesized to determine SD at the point-community level were not important in determining GD. In the SGDC literature, the role of the environment as a driver of SGDCs is ambiguous; few studies consider it and some have identified it as a driver of positive SGDCs (He et al., 2008) and some have not (He & Lamont, 2010; Sei, Lang & Berg, 2009). We did not find any SGDC studies that explicitly examined the role of biotic factors like competition.

We found a significant relationship between forest stand SR and H_E and non-significant positive correlations between site SR and H_O and site SR and AR (Table 3.6, Figure 3.10b) in the 16-site analysis. Forest area is likely to be the driving factor influencing these positive correlations. Forest area explained 64.8% of the variation in site species richness ($p<0.001$; Table 3.4) and 46.3% of variation in H_E ($p=9.004$; Table 3.4), and forest area also weakly explained values of H_O and AR in each population (Table 3.4). Relationships between area and species diversity are well documented (review in Drakare, Lennon & Hillebrand, 2006) as is the relationship between genetic diversity and population size (review in Leimu et al., 2006). Similar results identifying area as a parallel driver have been found in other SGDC studies in fragmented or island systems (Vellend, 2004; Struebig et al., 2011), but not all; a study of sand dune plants in Australia did not report dune area to be a driver of the positive SGDC (He et al., 2008).

It is possible that the inclusion of other factors may have strengthened the correlation of GD and SD in the 20-site and 16-site analyses. History is a factor often considered in SGDC studies in fragmented systems, and land-use and glaciation history have been found to be important drivers of SGDCs in a study of stream fishes in Florida (Blum et al., 2012). In our study it is possible that land-use history (i.e., forest age and disturbance) may be a factor contributing to the positive SGDC we report in this study.

Notably, we did not find that isolation played an important role in the SGDC analysis, and the role of habitat isolation as a driver of SGDC in similar fragmented systems remains inconclusive. Many studies found that isolation does not correlate with SD and GD (Struebig et al., 2011, Vellend et al., 2004, Blum et al., 2012), and a single study found it was important (Sei, Lang & Berg, 2009).

3.6 Summary and conclusion

The results of this field study have greatly broadened what is known about *C. leptonervia* ecology and genetics, and in particular, how *Carex* has responded to anthropogenic changes in the landscape. Furthermore, our genetic analysis revealed patterns of *C. leptonervia* dispersal, and indicated that animal vectors are moving

perigynia around the landscape. This finding suggests that cryptic dispersal may also occur for other forest *Carex* species for which little is known about their dispersal. We also found that forest stand isolation does not play a significant role in determining SD or GD in the Montérégie. Thus, isolated forests are also important in the preservation of biodiversity in fragmented forest landscapes, and in addition to large continuous forests, they should be considered in management and conservation contexts.

In our SGDC study, we sought to understand how the examination of community at different spatial scales influenced the strength of correlations. To achieve this, we took a novel approach and compared GD to two measures of SD. We over-approximated community to be all species co-occurring at a site and also took a more traditional approach and sampled the point-community, or species growing in close proximity to *C. leptonervia*. We found that GD was more strongly correlated with the SD of the study site. SGDC studies regularly attribute positive correlations as evidence of parallel processes (Vellend & Geber, 2005; Vellend, 2010; Vellend, 2014). Our results suggest that the spatial processes operating at the level of study site SR are also determining measures of GD (notably, area). The absence of correlations in SGDC studies may in part be a result of sampling methods and the occurrence of site-specific peculiarities in ecology and demography. Some of these peculiarities may include patterns of animal migration, levels of herbivory, levels of local adaptation, and even forest management, all of which may be difficult to quantify and are not uniform across all study sites. A comprehensive study of all possible influencing factors in SGDC studies is difficult to achieve but a disproportionate focus on spatial attributes, though helpful in understanding diversity, may overshadow other unique processes that occur on the landscape.

Site	Area (ha)	PROX	SR	SR sampling effort	Latitude	Longitude
3	40.28	196.26	129	0.037	45º42'46.26''N	73º13'11.72''W
4	6.30	5.96	93	0.328	45º41'58.56''N	73º12'33.50''W
5	224.28	732.89	138	0.015	45º37'55.09''N	73º10'12.92''W
6	67.59	22.32	134	0.037	45º27'14.77''N	73º14'33.76''W
7	11.34	12.90	119	0.13	45º28'47.06''N	73º11'00.67''W
8	1645.92	121.39	216	0.004	45º33'00.72''N	73º09'13.68''W
10	37.8	14.72	125	0.079	45º27'46.63''N	73º13'59.25''W
14	66.99	159.03	115	0.030	45º37'58.40'N	73º06'42.74''W
18	9.81	74.05	111	0.255	45º44'49.51''N	73º09'57.98''W
19	9.36	0.07	120	0.145	45º47'15.82''N	73º08'13.29''W
20	75.43	15.19	114	0.027	45º46'42.93''N	73º14'43.04''W
22	209.52	7.67	152	0.010	45º49'16.94''N	73º13'03.99''W
26	493.92	29.05	167	0.009	45º45'05.80''N	73º13'03.53''W
27	360	489.57	144	0.008	45º30'46.01''N	73º12'03.99''W
30	100.89	63.23	133	0.020	45º53'00.05''N	73º07'38.35''W
31	164.52	1079.81	128	0.012	45º31'26.83''N	73º11'04.90''W
33	8.37	133.97	109	0.182	45º23'40.86.''N	73º06'38.76''W
34	77.58	594.38	121	0.026	45º20'36.20''N	73º06'39.70''W
40	49.41	84.27	107	0.061	45º27'27.97''N	73º12'11.13''W
42	2175.3	74.30	205	0.002	45º27'27.97''N	73º12'11.13''W

Table 3.1: Each site has an associated value of area measured in hectares (ha), connectivity (PROX), species richness (SR) and sampling effort. Sampling effort was measured by dividing the number of hours spent at a site by the total number of hectares. Latitude and longitude are recorded for the approximate centre of each study site.

Site	SR	SHAN	community extent	CV (pH)	CV (%OM)
3	47	2.59	0.36	0.06	1.11
4	38	2.82	0.409	0.08	0.53
5	52	2.70	0.38	0.08	0.56
6	42	2.54	0.31	0.1	0.34
7	43	2.65	0.36	0.04	0.51
8	36	2.87	0.17	0.13	0.47
10	47	2.60	0.38	0.1	0.32
14	46	3.06	0.40	0.08	0.6
18	44	2.82	0.40	0.09	0.33
19	52	2.88	0.43	0.14	0.57
20	50	3.10	0.44	0.12	1.26
22	52	2.92	0.34	0.12	0.57
26	57	2.74	0.34	0.07	0.79
27	46	2.54	0.32	0.11	1.03
30	60	2.93	0.45	0.12	0.85
31	48	2.83	0.38	0.06	0.54
33	42	2.98	0.39	0.17	0.72
34	53	3.00	0.44	0.09	0.84
40	42	2.46	0.40	0.12	0.81
42	48	3.07	0.23	0.1	0.38

Table 3.2: Point-community level characteristics. Measures of species richness (SR) and Shannon diversity (SHAN) are reported for concatenated quadrats. Community extent is calculated as the proportion of species found in the concatenated quadrats as compared to the full site flora list. We also calculated the coefficient of variation (CV) for pH and %OM. Coefficient of variation (CV) calculations are based on the seven independent quadrat measurements we took at each site and are reported for both soil pH and percent organic matter (%OM).

Site	AR	H_E	H_O	N
3	25	0.513	0.431	23
4	19	0.396	0.312	25
5	13	0.268	0.172	23
6	18	0.381	0.257	23
7	16	0.276	0.249	25
8	15	0.253	0.205	25
10	17	0.295	0.205	24
14	15	0.292	0.324	25
18	17	0.244	0.124	24
19	18	0.342	0.331	24
20	21	0.305	0.379	24
22	19	0.387	0.452	25
26	14	0.198	0.164	21
27	17	0.353	0.287	23
30	17	0.327	0.258	22
31	17	0.333	0.222	25
33	15	0.246	0.147	23
34	15	0.260	0.162	25
40	15	0.270	0.130	23
42	24	0.441	0.331	24

Table 3.3: Genetic diversity at each site. We measured genetic diversity as AR, H_O, and H_E for each population. Sample size for AR was standardized to 21 individuals per population. The sample size for H_O and H_E is recorded in column N and varied from 21-25 individuals.

Explanatory Variable	SR (site)		SHAN (quadrats)		SR (quadrats)		AR		H_E		H_O	
	r^2	p	r^2	p	Pr^2	P	r^2	p	r^2	p	r^2	p
log(ha)	0.648	**<0.001***	(-)		(-)		0.1999	0.082	0.4633	**0.004***	0.033	0.244
log(PROX)	(-)		(-)		(-)		(-)		(-)		(-)	(-)
CV (pH)	N/A		(-)		(-)		N/A		N/A		N/A	
CV (%OM)	N/A		(-)		(-)		N/A		N/A		N/A	
community extent	N/A		0.380	**0.004***	(-)		(-)		(-)		(-)	
SR (site)	N/A		(-)		0.400	**0.003***	N/A		N/A		N/A	
	Kolmogorov-Smirnov test											
p-value	0.352		0.320		0.237		0.547		0.270		0.429	

Table 3.4: Results of the regression models explaining the greatest amount of variation in each diversity measure. The diversity measures reported are species richness (SR) of site, SR and Shannon diversity (SHAN) of point-communities, allelic richness (AR), gene diversity (H_E) and observed heterozygosity (H_O). Explanatory variables included in the regressions are: the logarithm of area (ha) connectivity (PROX), the coefficient of variation (CV) of pH, the CV of percent organic matter (%OM), community extent and SR of site floras. N/A indicates that the explanatory variable was not included in the multiple regression, whereas (-) indicates that the variable was removed in the final regression equation. An * indicates a significant p-value. The results of Kolmogorov-Smirnov tests for normality of residuals are also reported.

Site	A	B	C
3	0.112	0.081	0.808
4	0.570	0.306	0.123
5	0.447	0.489	0.066
6	0.176	0.216	0.608
7	0.341	0.608	0.051
8	0.204	0.687	0.110
10	0.614	0.335	0.051
14	0.670	0.277	0.053
18	0.145	0.673	0.182
19	0.155	0.575	0.270
20	0.391	0.511	0.098
22	0.149	0.570	0.281
26	0.276	0.647	0.077
27	0.825	0.071	0.104
30	0.139	0.554	0.306
31	0.354	0.397	0.248
33	0.349	0.578	0.073
34	0.203	0.671	0.126
40	0.268	0.555	0.178
42	0.536	0.176	0.288

Table 3.5: Partitioning of each study site into genetic clusters A, B, and C. These results were obtained by averaging the results of 10 runs for K=3 in the program STRUCTURE.

SD	GD	16 site analysis		20 site analysis	
		R	p-value	R	p-value
SR Site	**H_E**	0.822	<0.001*	0.077	0.748
SR Site	**AR**	0.653	0.006	0.121	0.612
SR Site	**H_O**	0.438	0.090	0.083	0.729
SR Quadrat	**H_E**	0.4737	0.474	-0.050	0.833
SR Quadrat	**AR**	0.107	0.695	-0.013	0.956
SR Quadrat	**H_O**	0.331	0.211	0.130	0.588
SHAN Quadrat	**H_E**	0.092	0.718	-0.071	0.769
SHAN Quadrat	**AR**	0.373	0.1544	0.132	0.580
SHAN Quadrat	**H_O**	0.398	0.127	0.235	0.319

Table 3.6: Pairwise Pearson correlations for each measure of species diversity (SD) and genetic diversity (GD). Two treatments are reported: outlier study sites removed and all sites. SD measures reported are species richness (SR) and Shannon diversity (SHAN). SR and SHAN are reported for concatenated quadrats, and only SR is reported for site. GD measures reported are gene diversity (H_E), allelic richness (AR) and observed heterozygosity (H_O). An * indicates a significant correlation.

Species	Growth form, habitat, and location	# SSR's	GD	HWE	IBD	G_{ST}/F_{ST}	Citation
C. extensa	caespitose salt marsh Mediterranean	4	H_E=0.10 H_O=0.04	Departure from HWE	N/A	N/A	Escudero et al., 2010
C. kobomugi	rhizomatous sand dunes Japan	14	H_E=0.545	N/A	No	N/A	Ohsako & Yamane, 2007
C. leptonervia	caespitose forest Québec	6	H_E=0.32 H_O=0.26	Excess homozygotes	No	G_{ST}=0.16	This study
C. macrocephala	rhizomatous sand dunes British Columbia	11	H_O=0-0.2	Departure from HWE	N/A	N/A	King & Roalson, 2009
C. moorcroftii	rhizomatous Tibetan plateau China	10	H_E=0.10	N/A	No	G_{ST}=0.66	Liu, Wei & Dong, 2009
C. scabrifolia	rhizomatous salt marsh Japan	9	H_E=0-0.679 H_O=0-1	Excess heterozygotes	N/A	Pairwise F_{ST} =0.234-0.631	Hodoki et al., 2009

Table 3.7: Survey of *Carex* population genetic studies that use microsatellite loci. All measures of genetic diversity (GD) reported in the study are in column 4, with the exception of allelic richness which was usually reported as a range per locus. Column 6 reports the results of isolation-by-distance (IBD) tests. "No" means that no pattern was found and N/A means that a correlation between genetic distance and geographic distance was not found.

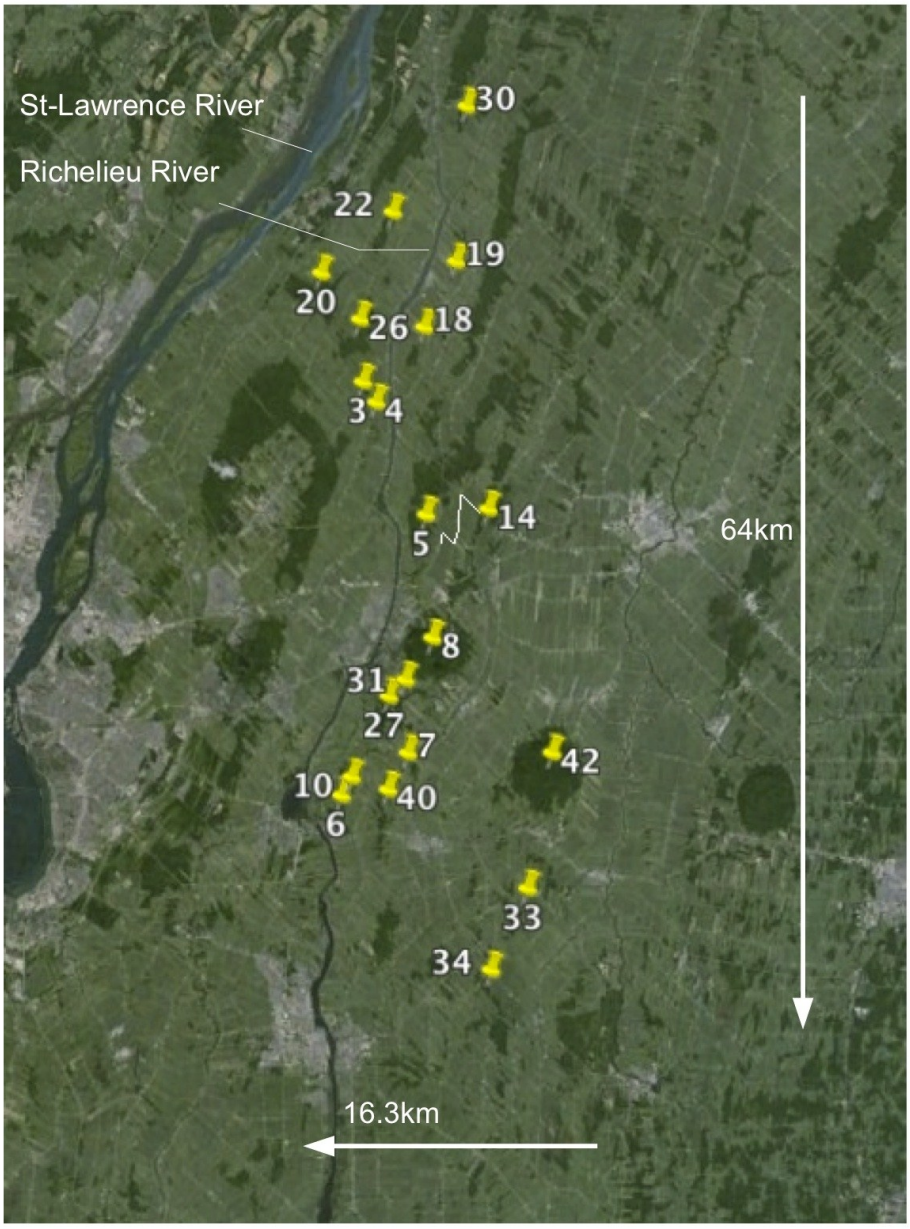


Figure 3.1: A Google Earth map of the Montérégie region with 20 study sites labelled. Site 8 and Site 42 are the Monteregian hills Mont-Saint Hilaire, and Rougemont, respectively. The Richelieu River bisects the field sites. Site 30 is the northern-most study site in Saint-Ours and Site 34 is the southern-most site in Saint-Jean-sur-Richelieu.

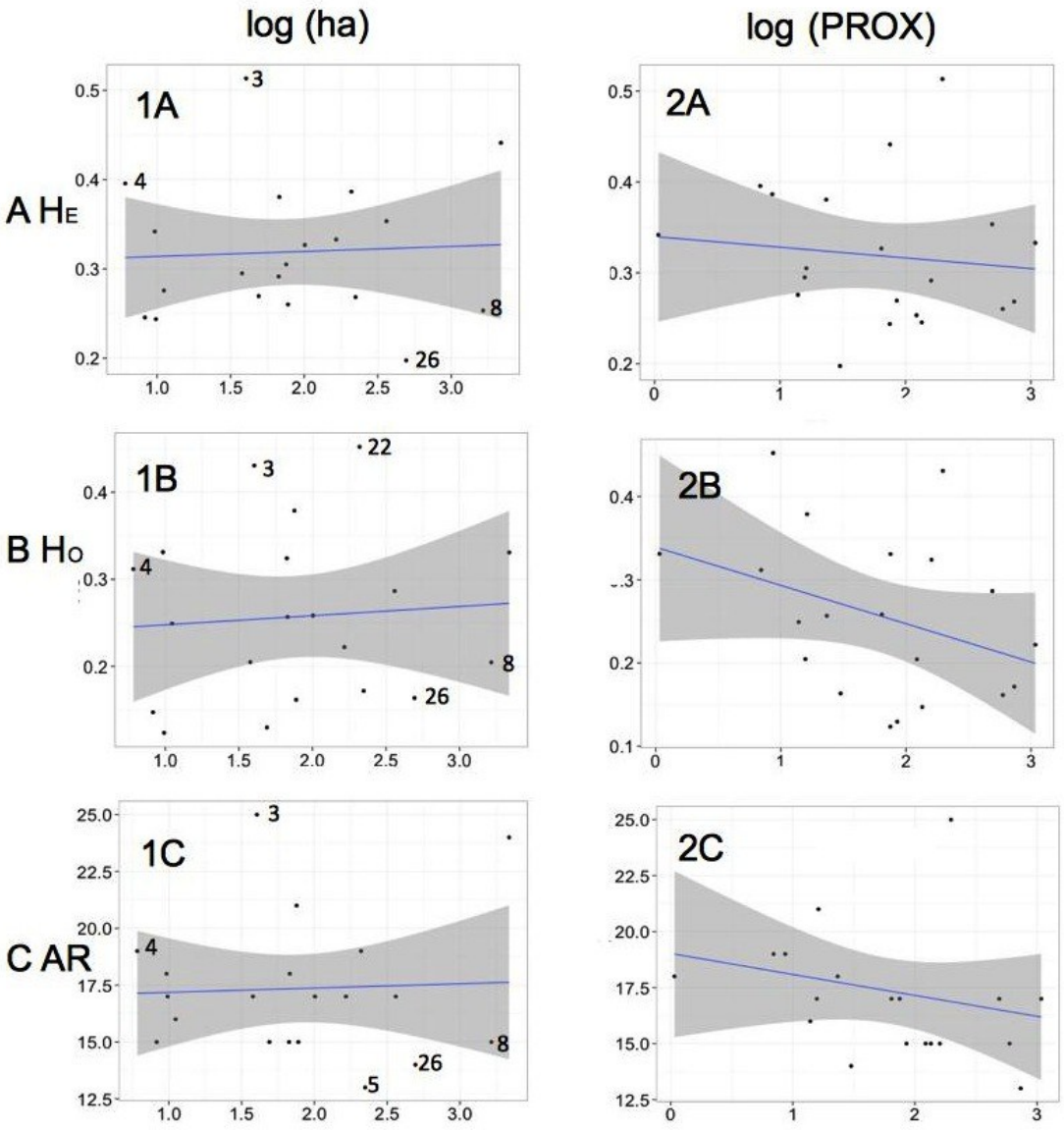


Figure 3.2 a: Scatter plots of pairwise tests of the spatial variables area [log (ha)] and connectivity [log (PROX)] and genetic diversity measures in the 20 site-analysis. The genetic diversity measures are gene diversity (H_E), observed heterozygosity (H_O) and allelic richness (AR). The shaded grey area shows 95% confidence interval and the slope is the slope of a linear regression. For regression results between these variables please see Table 3.4. Study sites that deviate from the trend of increasing GD with increasing area are identified by number.

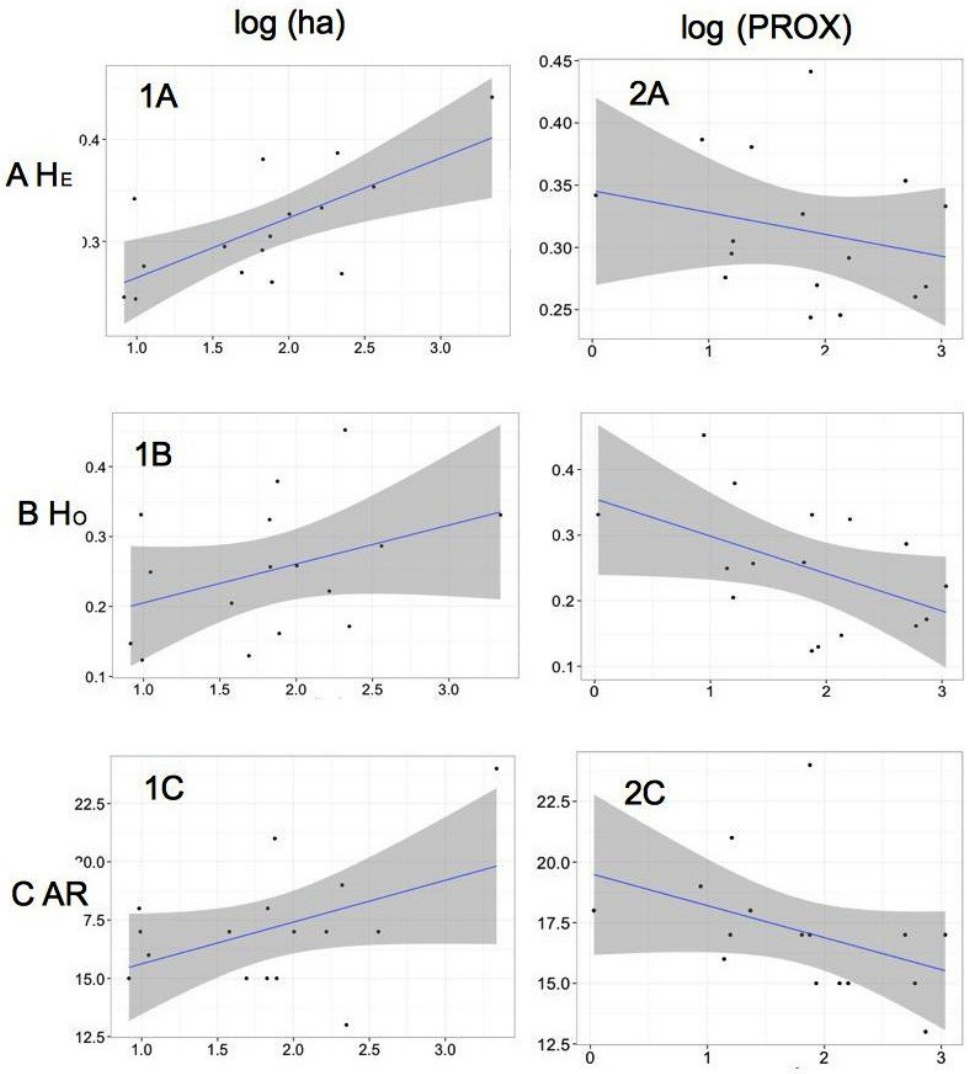


Figure 3.2 b: Scatter plots of pairwise tests of the spatial variables area [log (ha)] and connectivity [log (PROX)] and genetic diversity measures in the 16-site analysis. The genetic diversity measures are gene diversity (H_E), observed heterozygosity (H_O) and allelic richness (AR). The shaded grey area shows 95% confidence interval and the slope is the slope of a linear regression. For regression results between these variables please see Table 3.4.

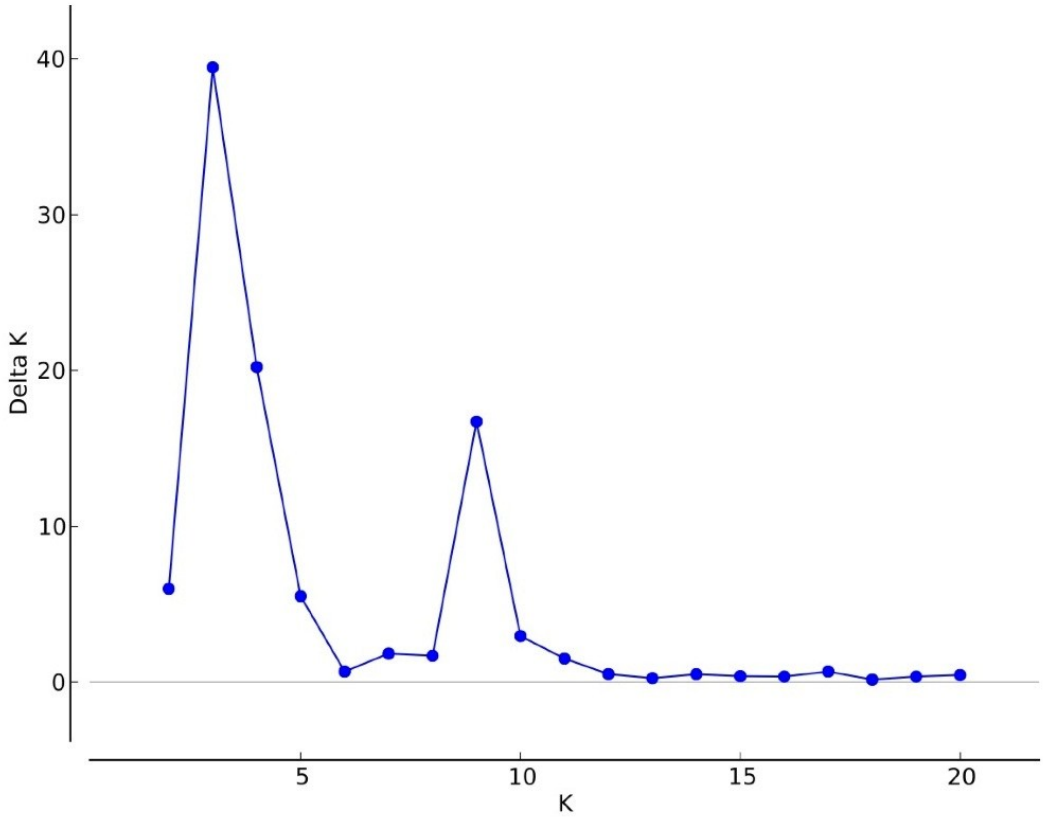


Figure 3.3: Structure Harvester output. Following the methods of Evanno et al., (2005) the results plot reveals that ΔK is highest when K=3. We define K as the number of genetic clusters as defined in the program STRUCTURE. The line plot starts at K=2, since ΔK cannot be calculated when K=1.

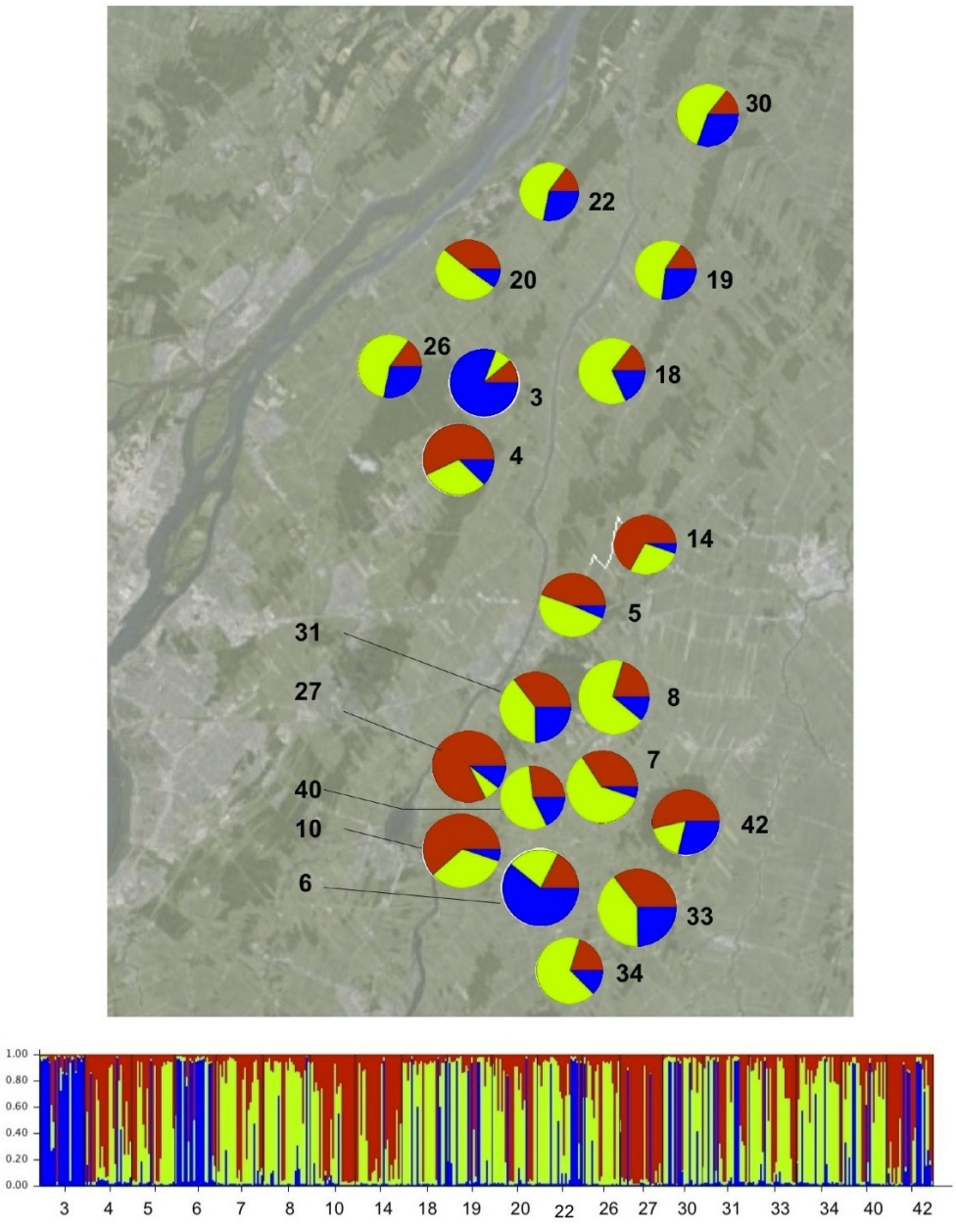


Figure 3.4: A map depicting the proportion of each study site into genetic cluster A, B, and C. Genetic cluster A=red, B=green, C=blue. The bar plot is output from the program STRUCTURE and shows the proportion of each individual in each population belonging to each genetic cluster. For exact proportions of each genetic cluster at each study site, please see Table 3.5.

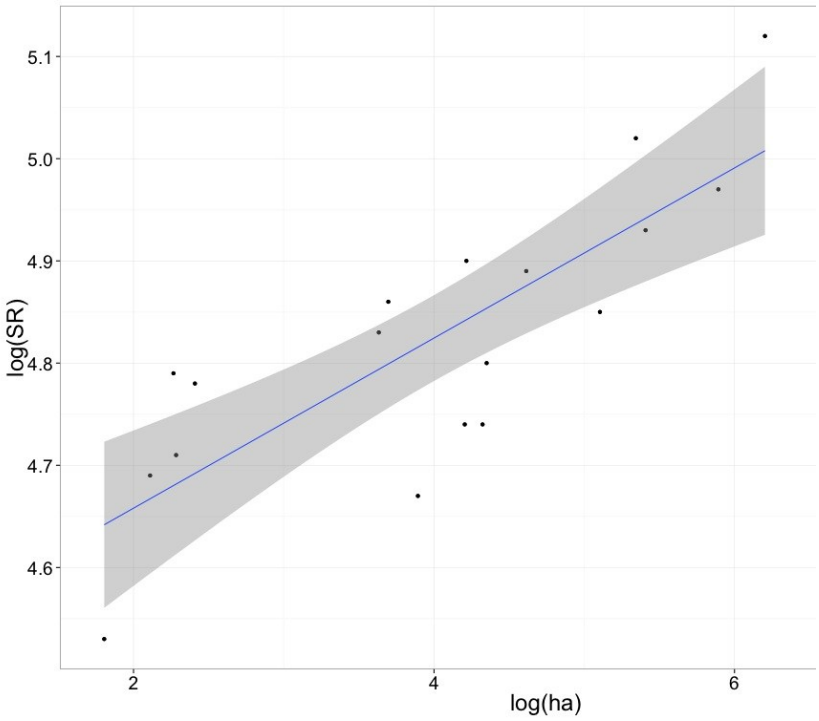


Figure 3.5: Scatter plot of area [log (ha)] x species richness [log (SR)] with regression line, and 95% confident intervals in dark grey.

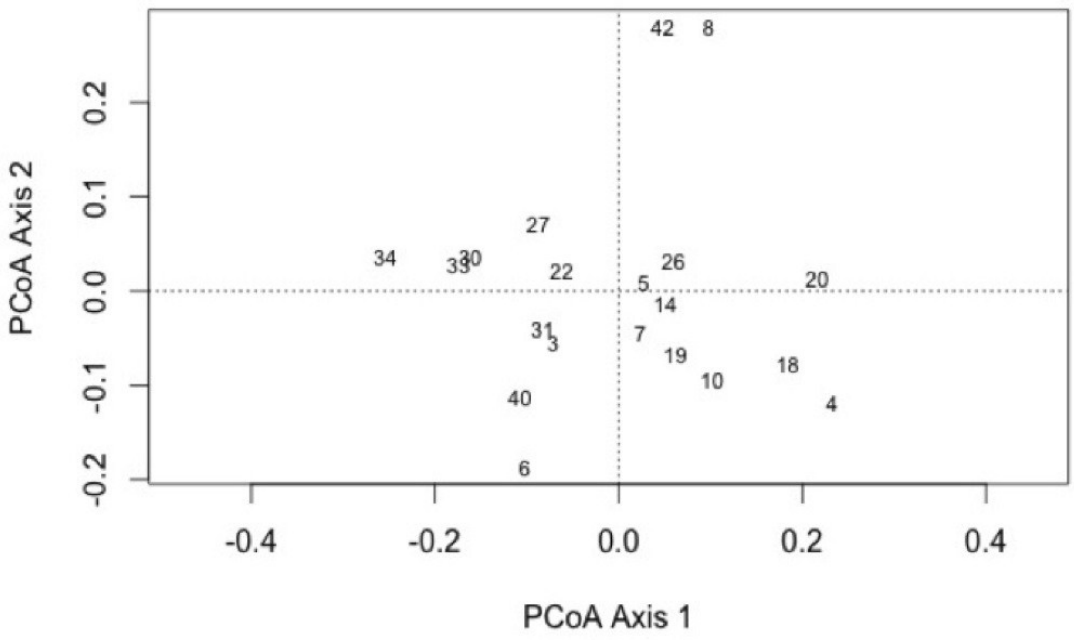


Figure 3.6: Principal coordinate analysis (PCoA) of site flora lists. Site 8 is Mont Saint-Hilaire and Site 42 is Rougemont. PCoA Axis 1 explains 20% of variation and PCoA Axis 2 explains 20% of variation.

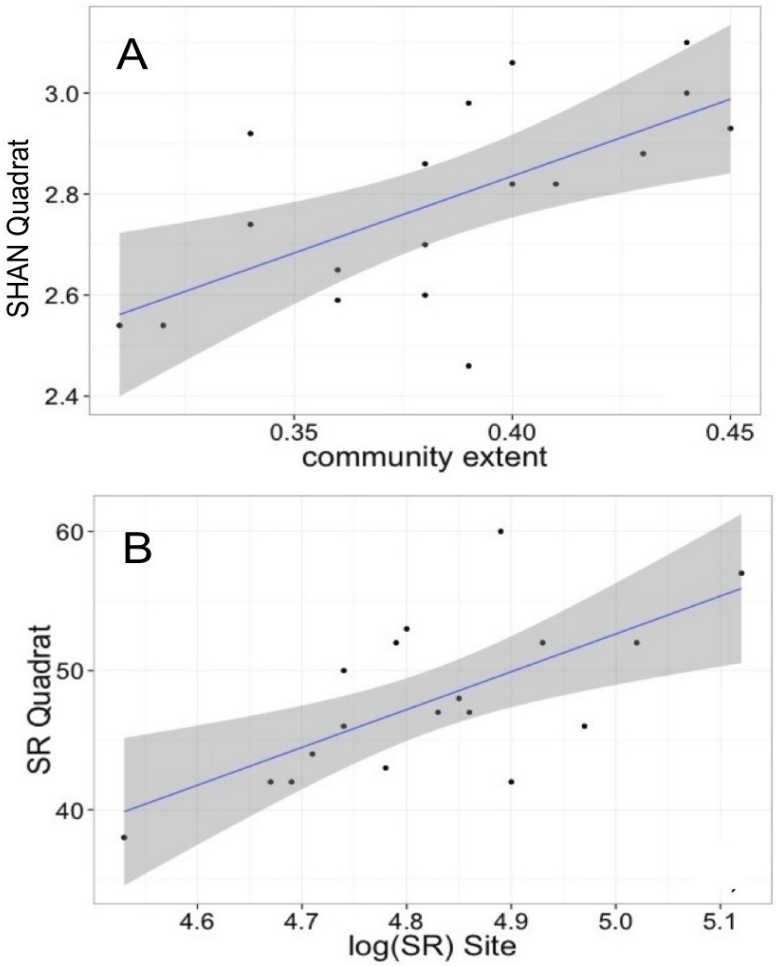


Figure 3.7: Scatter plots for quadrat species diversity and explanatory variables. A depicts a scatter plot of community extent x Shannon diversity (SHAN). Figure B depicts a scatter plot of site species richness (SR) and SR of quadrats. The grey regions represent the 95% confidence interval of a linear regression.

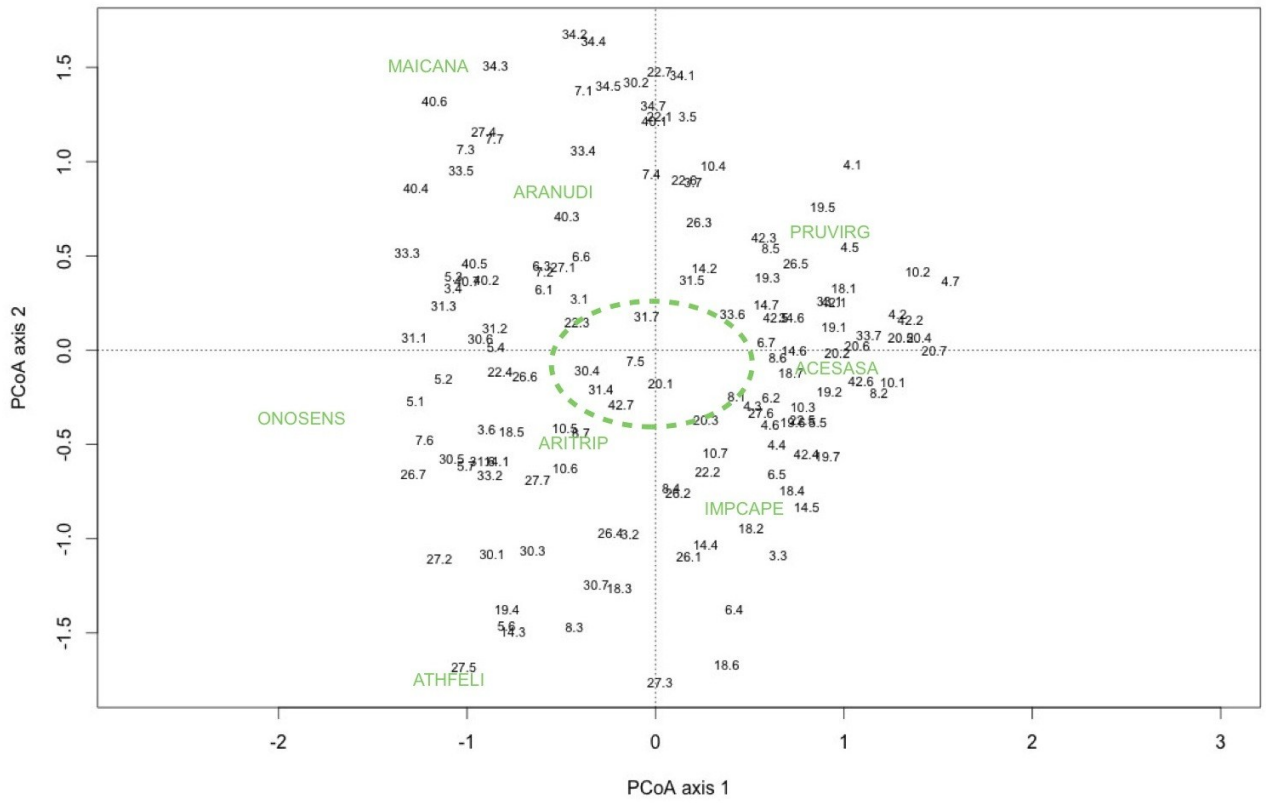


Figure 3.8: Principal coordinate analysis (PCoA) of individual quadrats. The most dominant species are represented by forestry codes (first three letters from genus + first four letters from species epithet). Translation of the codes is as follows: ACESASA= *Acer saccharum,* ARANUDI= *Aralia nudicaulis,* ARITRIP= *Arisaema triphyllym,* ATHFELI= *Athyrium filix-femina,* IMPCAPE= *Impatiens capensis,* MAICANA= *Maianthemum canadense,* ONOSENS= *Onoclea sensibilis,* PRUVIRG= *Prunus virginiana.* All other species are in the region delineated by the broken ellipse in the centre of the plot. Axis 1 explains 10% of variation and axis 2 explains 7% of variation.

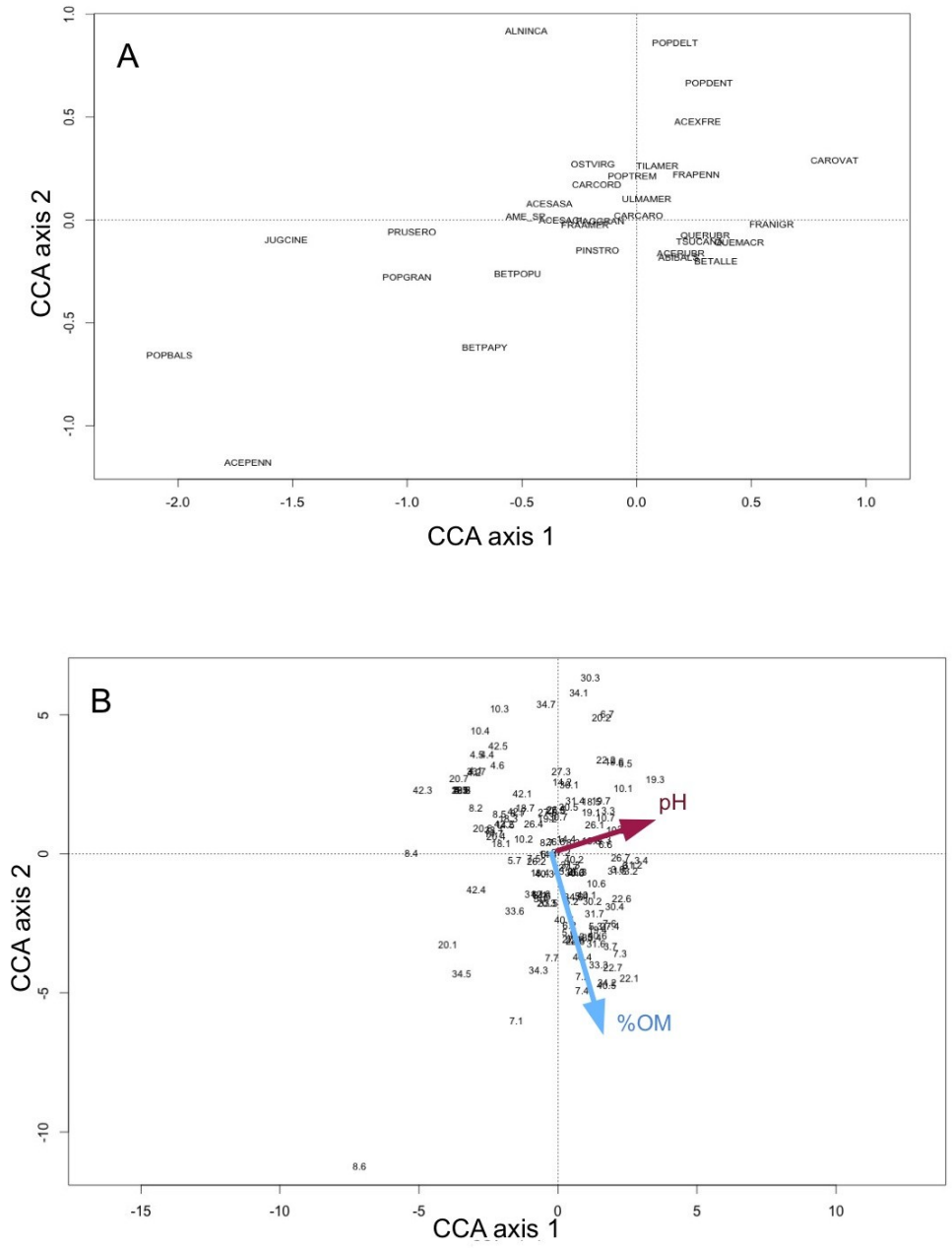


Figure 3.9: Canonical Correspondence analysis of Hellinger-transformed canopy composition data. Part A depicts the species only and B shows the biplot with quadrat labels (species omitted). For translation of species codes reported in A please see Appendix 8.

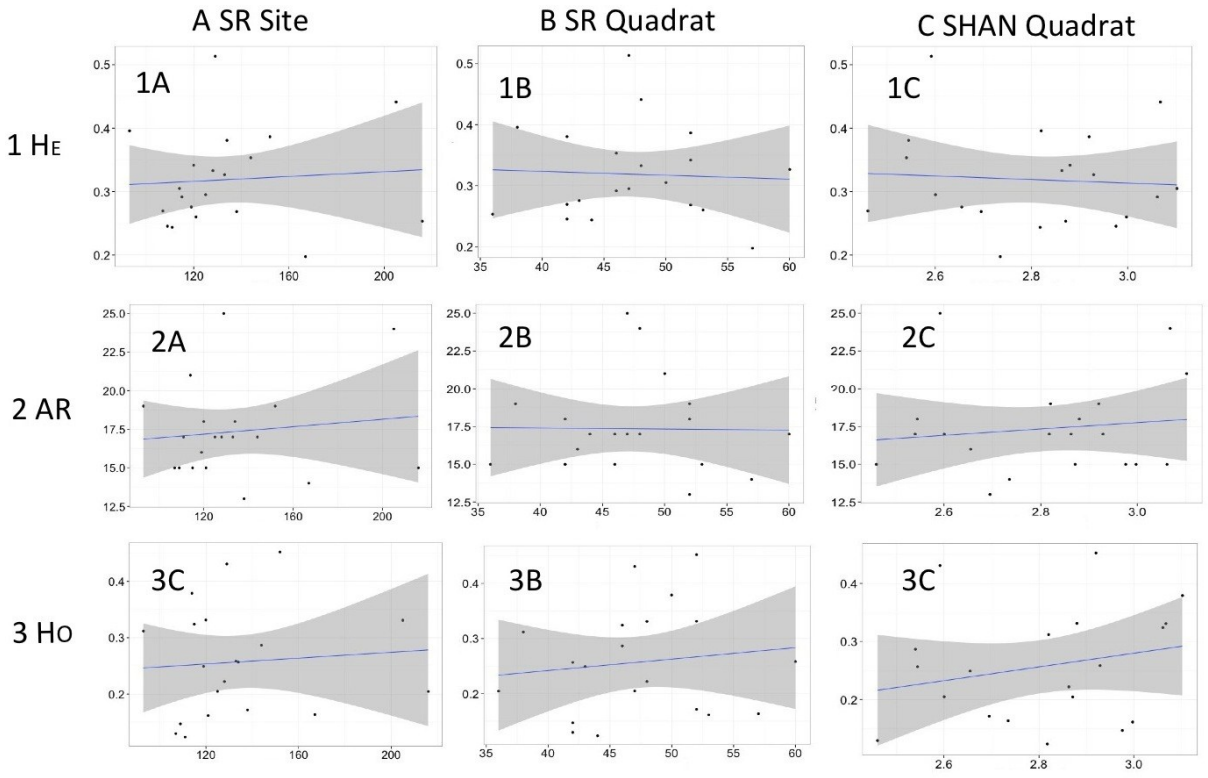


Figure 3.10 a: Pairwise scatterplots of all genetic diversity measurements (H_E, AR, H_O) by species diversity measurements (SR Site, SR Quadrats, and SHAN quadrats) in the 20-site analysis. Genetic diversity measures reported are gene diversity (H_E), allelic richness (AR) and observed heterozygosity (H_O). Species diversity measures used are species richness (SR) and Shannon diversity (SHAN).

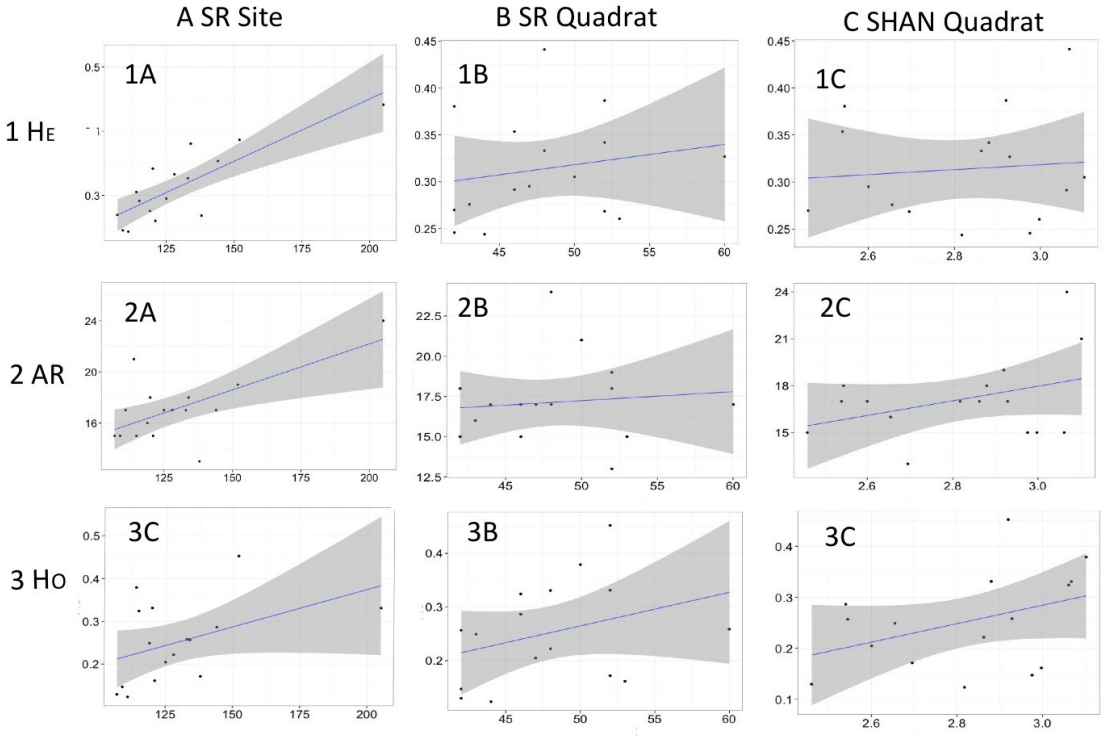


Figure 3.10 b: Pairwise scatterplots of all genetic diversity measurements (H_E, AR, H_O) by species diversity measurements (SR Site, SR Quadrats, and SHAN quadrats) in the 16-site analysis. Genetic diversity measures reported are gene diversity (H_E), allelic richness (AR) and observed heterozygosity (H_O). Species diversity measures used are species richness (SR) and Shannon diversity (SHAN). Results of Pearson correlations for each pairwise test are reported in Table 3.6.

GENERAL CONCLUSIONS

The field and lab findings of this study greatly broaden what is known about both the genetics and ecology of *C. leptonervia.* In addition to the development of reliable and high-quality microsatellites for use in *C. leptonervia,* these microsatellite markers show potential to be used throughout much of the genus as evidenced by successful cross-amplification in distantly related species. We screened a majority of the microsatellite loci we tested on three other species (*C. laxiflora, C. limosa,* and *C. lupulina*). These results provide a starting point for further development of these loci for use in these species and their relatives. We also identified lower levels of inbreeding and population differentiation and higher rates of heterozygosity than we had predicted based on previous studies of caespitose *Carex* species. We propose that it would be of interest to know if similar species in *Carex* section *Laxiflorae* behave similarly to *C. leptonervia,* as this could add nuance and detail to the understanding of caespitose *Carex* species.

From an ecological perspective, we were able to characterize *C. leptonervia* as a species with flexible ecological tolerances and as a species that thrives in many different vegetation communities. Overall, we attribute the high occurrence of *C. leptonervia* in the Montérégie to these broad environmental tolerances, the ability to thrive in edge and disturbed habitat, and the ability to disperse. In this study we were unable to empirically identify the mechanism of *C. leptonervia* long-distance dispersal but propose that the cryptic migration of *Carex* and other forest herbs is an interesting avenue for future research.

We thoroughly surveyed 20 forest fragments in the Montérégie, the majority of which have no record of past full flora surveys (with the exception of Mont Saint-Hilaire and Rougemont). In this survey we also identified 41 species of *Carex* across the twenty study sites. Although it is a common and diverse genus in this region, *Carex* is also a notoriously difficult genus to identify, and it is often identified only to the genus level. Therefore, in addition to our identification of 418 species, our characterization of *Carex*

in the Montérégie contributes uniquely to biodiversity studies in the region. These records provide an important baseline for future studies that seek to document change in the Montérégie, a region that remains under great urban and agricultural development pressure.

Surprisingly, we found that forest connectivity played at most a small role in our assessment genetic diversity and species diversity in the Montérégie. We view this as an unexpected but positive finding in our study and conclude that isolated forests are as important as highly connected forests in the maintenance of biodiversity in the Montérégie. Since the Montérégie contains many isolated forests, we suggest that the importance of these forests should not be overlooked in future management and conservation decisions. We also reiterate the importance of preserving large continuous forest fragments as they contribute greatly to regional species diversity and genetic diversity.

LITERATURE CITED

Agrawal, A.A. (2003). Community genetics: new insights into community ecology by integrating population genetics. *Ecology, 84,* 543-544.

Aguilar, R., Quesada, M., Ashworth, L., Herrerias-Diego, Y. & Lobo. J. (2008). Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Molecular Ecology 17*, 5177-5185.

Amarasekare, P. (2000). The geometry of coexistence. *Biological Journal of the Linnaean Society, 71,* 1-31.

Antonovics, J. (1976). The input from population genetics: 'the new ecological genetics'. *Systematic Botany, 1,* 233-245.

Arii, K., Hamel, B.R. & Lechowicz, M.J. (1999). Environmental correlates of canopy composition at Mont St. Hilaire, Quebec, Canada. *Journal of the Torrey Botanical Society, 132,* 90-102.

Arrhenius, O. (1921). Species and area. *Journal of Ecology, 9,* 95-99.

Balloux, F. & Lugon-Moulin, N. (2002). The estimation of population differentiation with microsatellite markers. *Molecular Ecology, 11*, 155-165.

Beatty, S.W. (2003). Habitat heterogeneity and maintenance of species in understory communities. In F.S. Gillam & M.R. Roberts (Eds.), *The herbaceous layer in forests of eastern North America* (177-197). New York: Oxford University Press.

Bélanger, L. & Grenier, M. (2002). Agriculture intensification and forest fragmentation in the St. Lawrence valley, Québec, Canada. *Landscape Ecology, 17,* 495-507.

Bell, D.F. (1964). Loss-on-ignition as an estimate for organic matter and organic carbon in non-calcareous soils. *Journal of Soil Science, 12,* 84-92.

Bell, G., Lechowicz, M.J., & Waterway, M.J. (2001). The precision of adaptation in forest plants. In J. Silverton & Antonovics J. (Eds.) *Integrating ecology and evolution in a spatial context*. Oxford: Blackwell Science.

Bell, G., Lechowicz, M.J., & Waterway, M.J. (2006). The comparative evidence relating to functional and neutral interpretations of biological communities. *Ecology, 87*, 1378-1386.

Bellemare, J., Motzkin, G., & Foster, D.R. (2001). Legacies of agricultural past in the forested present: an assessment of historical land-use effects on rich mesic forests. *Journal of Biogeography, 29*, 1401-1420.

Berge, G., Nardal, I. & Hestmark, G. (1998). The effect of breeding systems and pollination vectors on the genetic variation of small plant populations within an agricultural landscape. *Oikos, 81,* 17-29.

Bierzychudek, P. (1982) Life histories and demography of shade-tolerant temperate forest herbs: a review. *New Phytologist, 90*, 757-776.

Blum, M.J., Bagley, M.J., Walters, D.M., Jackson, S.A., Daniel, F.B., Chialoud, D.J., & Cade, B.S. (2012). Genetic diversity and species diversity of stream fishes covary across a land-use gradient. *Oecolgia, 168,* 83-95.

Borcard, D., Gillet, F., & Legendre, P. (2011). *Numerical ecology with R*. New York: Springer.

Bouchard, A., Dyrda, S., Bergeron, Y., & Meilleur, A. (1989). The use of notary deeds to estimate the changes in the composition of 19th century forests, in Haut-Saint-Laurent, Quebec. *Canadian Journal of Forest Research, 19,* 1146-1150.

Brouillet, L., Desmet, P., Coursol, F., Meades, S.J., Favreau, M., Anions, M., Bélisle, P., Gendreau, C., Shorthouse, D., and contributors. (2010+). Database of vascular plants of Canada (VASCAN). Online at: http://data.canadensys.net/vascan.

Broquet, T., Ray, N., Petit, E., Fryxell, J.M., & Burel, F. (2006). Genetic isolation by distance and landscape connectivity in the American marten (*Martes americana). Landscape Ecology, 21,* 877-889.

Bruederle, L.P, Yarbrough, S.L., & Fehlberg, S.D. (2008). Allozyme variation in the genus *Carex* … 15 years later: 1986-2001. In Naczi, F.C. & B.A. Editor (Eds.). *Sedges: uses, diversity, and systematics of the Cyperaceae* (187-196). St. Louis: Missouri Botanical Garden Press.

Bryson, C.T. & Naczi, F.C. (2002). *Carex* section *Laxiflorae.* In Flora of North America editorial committee, (Eds). *Flora of North America North of Mexico*, Vol. 23 (431-448). New York: Oxford University Press.

Caley, J.M., & Schluter, D. (1998). The relationship between local and regional diversity: reply. *Ecology, 79,* 1827-1829.

Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings C., Venail, P. (...) & Naeem, S. (2012). Biodiversity loss and its impact on humanity. *Nature, 486,* 49-67.

Chave, J. (2004). Neutral theory and community ecology. *Ecology Letters, 7,* 241-243.

Chen, J., Franklin, J.R., & Spies, T.A. (1993). Contrasting microclimates around clearcut, edge, and interior of old-growth douglas-fir forest. *Agricultural and Forest Meteorology, 63,* 219-237

Cleary, D.F., Fauvelot, C., Genner, M. J., Menken, S., & Mooers, A. (2006). Parallel responses of species and genetic diversity to El Niño Southern oscillation-induced environmental destruction. *Ecology Letters, 9,* 304-319.

Corbit, M, Marks, P.L., & Gardescu, S. (1999). Hedgerows as habitat corridors for forest herbs in Central New York USA. *Journal of Ecology, 87,* 220-232.

Cornell, H.V., & Lawton, J.H. (1992). Species interactions, local and regional processes, and limits to the richness of ecological communities: a theoretical perspective. *Journal of Animal Ecology, 61,* 1-12.

Culley, T.M., & Grubb, T.C. (2003). Genetic effects of habitat fragmentation in *Viola pubescens* (Violaceae), a perennial herb with chasmogamous and cliestogamous flowers. *Molecular Ecology, 12,* 2919-2930.

Derry, A.M., Arnott, S.E., Shead, J.A., Hebert, D.N., & Boag, P.T. (2009). Ecological linkages between community and genetic diversity in zooplankton among Boreal Shield lakes. *Ecology, 90,* 2275-2286.

Drakare, S., Lennon, J.J., & Hillebrand, H. (2006). The imprint of the geographic, evolutionary and ecological context on species-area relationships. *Ecology Letters, 9,* 215-227.

Dray, S. and Dufour, A.B. (2007): The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software, 22*, 1-20.

Dumortier, M., Butaye, J., Jacquemyn, H., Van Camp, N., Lust, N., & Hermy, M. (2002). Predicting vascular plant species richness of fragmented forests in agricultural landscapes in central Belgium. *Forest Ecology and Management, 158*, 85-102.

Earl, D.A., & von Holdt, B.M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources, 4,* 359-361. Online at taylo0.biology.ucla.edu/structureHarvester/.

Ellis, A.G., Weis, A.E., & Gaut, B.S. (2007). Spatial scale of local adaptation and population genetic structure in a miniature succulent, *Agryderma pearsonii. New Phytologist, 174,* 904-914.

Escudero, M, Vargas, P., Arens, P., Ouborg, N.J., & Luce, M. (2010). The east-west-north colonization history of the Mediterranean and Europe by the coastal plant *Carex extensa* (Cyperacease). *Molecular Ecology, 19,* 352-370.

Etienne, R.S., & Oloff, H. (2004). A novel genealogical approach to neutral biodiversity theory. *Ecology Letters, 7,* 170-175.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology, 14,* 2611-2620.

Feininger, T., & Goodacre, A.K. (1995). The eight classical Monteregian hills at depth and the mechanism of their intrusion. *Canadian Journal of Earth Science, 32,* 1350-1364.

Fields Development Team. (2006). fields: tools for spatial data. National Center for Atmospheric Research, Boulder, CO. Online at: http://www.cgd.ucar.edu/Software/Fields

Finch, D.D., & Alexander, H.M. (2011). Variation in plant distributions, plant traits and disease levels across a woodland/grassland ecotone. *American Midland Naturalist, 166,* 309-324.

Fisher, R.A. (1932). *Statistical methods for research workers* (4th ed). London: Oliver and Boyd.

Friedman, J. & Barrett, S.C. (2009). The consequences of monoecy and protogyny for mating in wind-pollinated *Carex. New Phytologist, 181,* 488-497.

Flinn, K.M., Lechowicz, M.J., & Waterway, M.J., (2008). Plant species diversity and composition of small wetlands within upland forests. *American Journal of Botany, 95,* 1216-1224.

Flinn, K.M., Gouhier, T.C., Lechowicz, M.J. & Waterway, M.J. (2010). The role of dispersal in shaping plant community composition of wetlands within an old-growth forest. *Journal of Ecology, 98,* 1292-1299.

Gemayel, R., Vinces, M.D., Legendre, M., & Verstrepen, K.J. (2010). Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annual Review of Genetics, 44,* 445-477.

Gleason, H.A. (1922). On the relation between species and area. *Ecology, 3,* 158-162.

Gleason, H.A. (1926). The individualistic concept of plant association. *Bulletin of the Torrey Botanical Club, 53,* 7-26.

Gilbert, B. & Lechowicz, M.J. (2004). Neutrality, niches and dispersal in a temperate forest understory. *Proceedings of the National Academy (USA), 101,* 7651-7656.

Gotelli, N.J. & Ellison A.M. (2004). *A primer of ecological statistics* (2nd ed). Sunderland, Massachusettes: Sinauer Associates Inc.

Götzenberger L., de Bello, F., Bråthen, K.A., Davison, J., Dubuis, An., Guisan, A., Leps, J., Lindborg, R., Moora, M., Pärtel, M., Pellissier, L., Pottier, J., Vittoz, P., Zobel, K., & Zobel, M. (2011). Ecological assembly rules in plant communities—approaches, patterns and prospects. *Biological Reviews, 87*, 111-127.

Guichoux, E., Lagache, L., Wagner, S., Chaumeil, S.P., Léger, P., Lepais, O., Lepoitevin, G., Malusa, T., Revardel, E., Salin, F., & Petit, R.J. (2011). Current trends in microsatellite genotyping. *Molecular Ecology Resources, 11*, 591-611.

Guo, S.W., & Thompson, E.A. (1992). Performing the exact-test of Hardy-Weinberg proportion for multiple alleles. *Biometrics, 48,* 361-372.

Gustafson, E. J., & Parker, G. R. (1992). Relationships between landcover proportion and indices of landscape spatial pattern. *Landscape Ecology, 7*, 101–110.

Haines, A. (2011). New England Wild Flower Society's Flora Novae Angliae: A manual for the identification of native and naturalized higher vascular plants of New England. New Haven: Yale University Press.

Hanski, I., & Gilpin, M. (1991). Metapopulation dynamics: brief history and conceptual domain. *Biological Journal of Linnaean Society, 42,* 3-16.

Haila, Y. (2002). A conceptual genealogy of fragmentation research: from island biogeography to landscape ecology. *Ecological Applications, 12,* 321-334.

He, T., Lamont, B.B., Krauss, S.L., Enright, N.J. & Miller, B.P. (2008). Covariation between intraspecific genetic diversity and species diversity within a plant functional group. *Journal of Ecology, 96,* 956-961.

He, T., & Lamont, B.B. (2010). Species vs. genotypic diversity of nitrogen-fixing plant functional group in a metacommunity. *Population Ecology, 52,* 337-345.

Helm, A., Oja, T., Saar, L., Takkis, K., Talve, T., & Pärtel, M. (2009). Human influence lowers plant genetic diversity in communities with extinction debt. *Journal of Ecology, 97,* 1329-1336.

Herben, T. (2000). Correlation between species richness per unit area and the species pool cannot be used to demonstrate species pool effect. *Journal of Vegetation Science 11,* 123-126.

Hermy M. & Verheyen, K. (2007). Legacies of the past in the present-day forest biodiversity: a review of past land-use effects on forest plant species composition and diversity. *Ecological Research, 22,* 361-371.

Hibbs, D.E. (1983). Forty years of forest succession in central New England. *Ecology, 64,* 1394-1401.

Hipp, A.L., Kettennign, K.M., Feldheim, K.A., & Weber, J.A. (2009). Isolation of 11 polymorphic tri- and tetranucleotide microsatellite loci in a North American sedge (*Carex scoparia*: Cyperaceae) and cross-species amplification in three additional *Carex* species. *Molecular Ecology Resources, 9,* 625-627.

Hodoki, Y., & Ohbayashi, K. (2008). Genetic analysis of salt-marsh sedge *Carex scabrifolia* Steud. populations using newly developed microsatellite markers. *Conservation Genetics, 10,* 1361-1364.

Hoderegger, R., Buehler, D., Gugerli, F. & Manel, S. (2010). Landscape genetics of plants. *Trends in Plant Science, 15,* 676-683.

Honnay, O., Jacquemyn, H., Bossuyt, B., & Hermy, M. (2005). Forest fragmentation effects on patch occupancy and population viability of herbaceous plant species. *New Phytologist, 166,* 723-736.

Honnay, O. & Jacquemyn, H. (2007). Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. *Conservation Biology, 21,* 823-831.

Honnay, O., Bossuyt, B., Jacquemyn, H., A. Shimono, & Uchiyama, K. (2008). Can a seed bank maintain the genetic variation in the above ground plant population? *Oikos, 117,* 1-5.

Hufford, K.M., & Mazer, S.J. (2003). Plant ecotypes: genetic differentiation in the age of ecological restoration. *Trends in Ecology & Evolution, 18,* 147-155.

Jacquemyn H., & Brys, R. (2008). Effects of stand age on the demography of a temperate forest herb in post-agricultural forests. *Ecology, 89*, 3480-3489.

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

Jones, M.M., Tuomistor, H., Clark, D.B., & Olivas, P. (2006). Effects of mesoscale environmental heterogeneity and dispersal limitation on floristic variation in rain forest ferns. *Journal of Ecology, 94,* 181-195.

Kaira, Y.P. (1995). Determination of pH of soils by different methods: collaborative study. *Journal of AOAC International, 78,* 310-321.

Kashi Y. & King D. (2006). Simple sequence repeats as advantageous mutators in evolution. *Trends in Genetics, 22,* 253-259.

Kearns, C.A., Inouye, D.W. & Waser, N.M. (1998). Endangered mutualisms: the conservation of plant-pollinator interactions. *Annual Review of Ecology and Systematics, 29,* 83-112

Kellman, M. (1996). Redefining roles: plant community reorganization and species preservation in fragmented systems. *Global Ecology and Biogoegraphy Letters, 5,* 111-116.

Kimura, M. (1953). "Stepping stone" model of population. *Annual Report of the National Institute of Genetics, 3,* 62-63.

Kimura, M., & Weiss, G. (1964). The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics, 49,* 561-576.

Kimura, M. (1983). *The neutral theory of molecular evolution*. London: Cambridge University Press.

King, M. & Roalson, E. (2009). Isolation and characterization of 11 microsatellite loci from *Carex macrocephala* (Cyperaceae). *Conservation Genetics, 10,* 531-533.

Kolb, A., & Diekmann, M. (2004). Effects of environment, habitat configuration and forest continuity of the distribution of forest plant species. *Journal of Vegetation Science, 13,* 199-208.

Lamy, T., Jarne, P., Laroche, F., Pointer, J.-P., Huth, G., Segard, A., & David, P. (2013). Variation in habitat connectivity generates positive correlations between species and genetic diversity in a metacommunity. *Molecular Ecology, 22,* 4445-4456.

Larouche, M. (2013). Effects of past land use and landscape context on plant species composition and richness in woodlots of an agricultural landscape in Québec (Unpublished Master Thesis). McGill University, Montreal.

Lawton, J.H. (1999). Are there general laws in ecology? *Oikos, 84,* 177-192.

Leberg, P.L. (2002). Estimating allelic richness: effects of sample size and bottlenecks. *Molecular Ecology, 11*, 2445-2449.

Lee, H., Bakowsky, W., Riley, J., Bowles, J., Puddister, M., Uhlig, P., & McMurray, S. (1998). Ecological land classification for southern Ontario: first approximation and its application. Ontario Ministry of Natural Resources, Southcentral Science Section, Science Development and Transfer Branch. SCSS Field Guide FG-02.

Legendre, P., & Anderson, M.J. (1999). Distance-based redundancy analysis: testing multi-species responses in multi-factorial ecological experiments. *Ecological Monographs, 69*, 1-24

Legendre, P., & Gallagher, E.D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia, 129*, 271-280

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M. & Gonzalez, A. (2004). The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters, 7,* 601-613.

Leimu, R., Mutikainen, R., Koricheva, J., & Fischer, M. (2006). How general are the relationships between population size, fitness and genetic variation? *Journal of Ecology, 94,* 942-952.

Levins, R., (1969). Some demographic and genetic consequences of environmental heterogeneity. *Bulletin of the Entomological Society of America, 15,* 237-240.

Lindborg, R., Eriksson, O. (2004). Historical landscape connectivity affects present plant species diversity. *Ecology, 85,* 1840-1845.

Linhart, YB., & Grant, M.C. (1996). Evolutionary significance of local genetic differentiation in plants. *Annual Review of Ecological Systematics, 21,* 237-277.

Liu, W., Wei, W., & Dong, M. (2009). Clonal and genetic diversity of *Carex moorcroftii* on Qinghai-Tibet plateau. *Biochemical Systematics and Ecology, 37,* 370-377.

MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton, New Jersey: Princeton University Press.

Maki, M., Masuda, M., & Inoue, K. (1996). Genetic diversity and hierarchical population structure of a rare autotetraploid plant, *Aster kantoenisis* (Asteraceae). *American Journal of Botany, 83,* 296-303.

Manel, S., Schwartz, M., Luikart, G., & Taberlet, P. (2003). Landscape genetics: combining landscape, ecology, and population genetics. *Trends in Ecology and Evolution, 18,* 189-197.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research, 27,* 209-220.

Marchand, P., & Houle, G. (2006). Spatial patterns of plant species richness along a forest edge: What are their determinants? *Forest Ecology and Management, 223,* 113-124.

Matlack, G.R. (1994). Plant species migration in a mixed-history forest landscape in Eastern North America. *Ecology, 75,* 1491-1502.

Maycock, P. (1961). Botanical studies on Mont St. Hilaire, Rouville county, Quebec: general description of the area and a floristic survey. *Canadian Journal of Botany*, 39, 1293-1325.

McClintock, K., & Waterway, M. (1994). Genetic differentiation between *Carex lasiocarpa* and *C. pellita* (Cyperacae) in North America. *American Journal of Botany, 8,* 224-231.

McGarigal, K., Cushman, S.A., & Ene, E. (2012). FRAGSTATS v4: Spatial pattern analysis program for categorical and continuous maps. University of Massachusetts, Amherst. Online at: http: www.umass.edu/landeco/research/fragstats/fragstats.html

Meirmans, P.G. (2012). The trouble with isolation by distance. *Molecular Ecology, 21,* 2839-2846.

Ministère des resources naturelles et de la faune. (2003). Système d'information écofrestière (SIEF). Tiles 31h06, 31h11, 31h14. Scale 1:20,000 [computer file]. Québec, Canada

Myers, J.A., Vellend, M., Gardescu, S., & Marks, P.L. (2004). Seed dispersal by white-tailed deer: implications for long distance dispersal and migration of plants in eastern North America. *Oecolgia, 139,* 35-44.

Naaf, F., & Wulf, M. (2012). Plant community assembly in temperate forests along gradients of soil fertility and disturbance. *Acta Oecologica, 39,* 101-108.

Nei, M. (1972). Genetic distance between populations. *American Naturalist, 106,* 283-292.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America, 70*, 3321-3323.

Odat, N., Jetschke, G., & Hellwig, F.H. (2004). Genetic diversity of *Ranunculus acris* L. (Ranunculaceae) populations in relation to species diversity and habitat type in grassland communities. *Molecular Ecology, 13,* 1251-1257

Ohsako, T., & Yamane, K. (2007). Isolation and characterization of polymorphic microsatellite loci in Asiatic sand sedge, *Carex kobmugi* Ohwi (Cyperaceae). *Molecular Ecology Notes, 7,* 1023-1025.

Ouborg, N.J., Vergeer, P., & C. Mix. (2006). The rough edges of the conservation genetics paradigm for plants. *Journal of Ecology, 94,* 1233-1248.

Papadopoulou, A., Anastasiou, I., Spagopoulou, F., Stalimerous, M., Terzopoulous, S., Legakis, A., & Vogler, A.P. (2011). Testing the species-genetic diversity correlation in the Aegean Archipelago: toward a haplotype-based macroecology. *The American Naturalist, 178,* 241-255.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H. & Wagner, H. (2013). vegan: community ecology package. R package version 2.0-10. http://CRAN.R-project.org/package=vegan.

Peterken, G.F., Game, M. (1984). Historical factors affecting the number and distribution of vascular plant species in the woodlands of Central Lincolnshire. *Journal of Ecology, 72,* 155-182.

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155,* 945-959.

Puşcaş, M., Taberlet, P., & Choler, P. (2008). No positive correlation between species and genetic diversity in European alpine grasslands dominated by *Carex curvula. Diversity and Distributions, 14,* 852-861.

Rackham, O. (1980). *Ancient Woodland: its history, vegetation, and uses in England*. London: Edward Arnold.

Raymond M., & Rousset, F. (1995). GENEPOP(version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity, 86,* 248-249.

Richer, M.C., Oullet, J.P., Lapoint, L., Crête, M, & Huot, J. (2005). Impacts of white-tailed deer grazing in hay fields in Southern Québec. *Wildlife Society Bulletin, 33,* 1274-1281.

Ricklefs, R.E. (2008). Disintegration of the Ecological Community. *The American Naturalist, 172,* 741-750.

Rius, M., & Darling. J.A. (2014). How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology and Evolution, 29,* 233-242

Roberts-Pichette, P., & Gillespie, L. (1999). Terrestrial vegetation biodiversity monitoring protocols, EMAN Occasional Papers Series, Report No. 9. Burlington, ON: Ecological Monitoring Coordination Office.

Rousset, F. (2008). Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources, 8,* 103-106.

Rowe, J.S. (1972). Forest regions of Canada. Canadian Forestry Service, Department of Fisheries and the Environment. Ottawa, Canada.

Roy, V., & de Blois, S. (2006). Using functional traits to assess the role of hedge row corridors as environmental filters for forest herbs. *Biological Conservation, 130,* 592-603.

Schaal, B.A. & Leverich, .J. (1996). Molecular variation in isolated plant populations. *Plant Species Biology, 11*, 33-49.

Schlötterer. (2000). Evolutionary dynamics of microsatellite DNA. *Chromosoma*, *109,* 365-371.

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

Sei, M., Lang, B.K., & Berg, D.J. (2009). Genetic and community similarities are correlated in endemic-rich springs of the northern Chihuahuan desert. *Global Ecology and Biogeography, 18,* 192-201.

Silvertown, J., Biss, P.M., & Freeland, J. (2009). Community genetics: resource addition has opposing effects on genetic and species diversity in a 150-year experiment. *Ecology Letters, 12,* 165-170.

Šímová, I., Li, Y.M., & Starch, D. (2013). Relationship between species richness and productivity in plants: the role of sampling effect, heterogeneity and species pool. *Journal of Ecology, 101,* 161-170.

Singleton, R., Gardescu, S., Marks, P.L., & Geber, M.A. (2001). Forest herb colonization of post-agricultural forests in central New York State, USA. *Journal of Ecology, 89,* 325-338.

Struebig, M.J., Kingston, T., Petit, E.J., Le Comber, S.C., & Zubaid, A. (2011). Parallel declines in species and genetic diversity in tropical forest fragments. *Ecology Letters, 14,* 582-590.

Sunnocks, P. (2000). Efficient genetic markers for population biology. *Trends in Ecology and Evolution, 15,* 199-203.

Taberlet, P. (1998). Biodiversity at the intraspecific level: the comparative phylogeographic approach. *Journal of Biotechnology, 64,* 91-100.

Taberlet, P., Zimmermann, N.E., Englisch, T., Tribsuh, A., Holderegger, R., Alvarez, N. & Gugerli, F. (2012). Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. *Ecology Letters, 15,* 1439-1448.

Tischendorf, L., & Fahrig, L. (2000). On the usage and measurement of landscape connectivity. *Oikos, 90:* 7-19.

Thrall, P.H., Burdon, J.J. & Murray, B.R. (2000). The metapopulation paradigm: a fragmented view of conservation biology. In Young, G. & Clarke, G.M. (Eds.) *Genetics, demography and viability of fragmented populations* (75-95). United Kingdom: Cambridge University Press.

Toonen, R.J., & Hughes, S., (2011). Increased throughput for fragment analysis on ABI Prism 377 automated sequencer using a membrane comb and STRand software, *Biotechnologies, 31,* 1320-1324.

Tribsch, A., Schönswetter, P. & Steussy, T.F. (2002). *Saponaria pumila* (Caryophyllaceae) and the ice age in the European alps. *American Journal of Botany, 89,* 2024-2033.

Vellend, M. (2003). Island biogeography of genes and species. *The American Naturalist, 162,* 358-365.

Vellend, M. (2004). Parallel effects of land-use history on species diversity and genetic diversity of forest herbs. *Ecology, 85,* 3043-3055.

Vellend, M. (2010). Conceptual synthesis in community ecology. *Quarterly Review of Biology, 85,* 183-206.

Vellend, M., Myers, J.A., Gardescu, S., & Marks, P.C. (2003). Dispersal of *Trillium* seeds by deer: implications for long-distance migration of forest herbs. *Ecology, 84,* 1067-1072.

Vellend, M., & Geber, M. (2005). Connections between species diversity and genetic diversity. *Ecology Letters, 8,* 767-787.

Vellend, M., Verheyen, K., Flinn, K. M., Jacquemyn, H., Kolb, A., Van Calster, H., (...) & Hermy, M. (2007). Homogenization of forest plant communities and weakening of species–environment relationships via agricultural land use. *Journal of Ecology, 95,* 565-573.

Verehyen, K., Vellend, M., van Calster, H., Peterken, G., & Hermy, M. (2004). Metapopulation dynamics in changing landscapes: a new spatially realistic model for forest plants. *Ecology, 85,* 3302-3312.

Voss, E.G. (1972). Michigan flora: Part 1 gymnosperms and monocots. Bloomfield Hills: Cranbrook Institute Science Bulletin 55 and University of Michigan Herbarium.

Vymazalová, M., Axamanová, I., & Tichý. (2012). Effect of intra-seasonal variability on vegetation data. *Journal of Vegetation Science, 23,* 978-984.

Wehenkel, C., Bergmann, F., & Gregorious, H. (2006). Is there a trade-off between species diversity and genetic diversity in forest tree communities? *Plant Ecology, 185,* 151-161.

Wei, X., and Jiang, M. (2012). Contrasting relationships between species diversity and genetic diversity in natural and disturbed forest tree communities. *New Phytologist, 193,* 779-786.

Weiher, E., Clarke, G.D., & Keddy, P.A. (1998). Community assembly rules, morphological dispersion, and the coexistence of plant species. *Oikos, 81,* 309-322.

Weir, B. (1996). *Genetic data analysis II.* Sunderland: Sinauer Associates.

Whitkus, R. (1988). Experimental hybridization among chromosome races of *Carex pachystachya* and related species *C. macloviana* and *C. preslii* (Cyperaceae). *Systematic Botany, 13,* 146-153.

Wiens, J.A. & Milne, B.T. (1989). Scaling of 'landscapes' in landscape ecology, or landscape ecology from a beetle's perspective. *Landscape Ecology, 3,* 87-96.

Winter, D.J. (2012). mmod: an R library for the calculation of population differentiation statistics. *Molecular Ecology Resources, 12,* 1158-1160.

Williams, S.C., Ward, J.S., & Ramakrishnan, U. (2008). Endozoochory by white-tailed deer (*Odocoileus virginianus*) across as suburban/woodland interface. *Forest Ecology and Management, 255,* 940-947.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics,* 97-159.

Wright, S. (1943). Isolation by distance. *Genetics, 28,* 114-138.

Wright, S. (1969). *Evolution and the genetics of populations*. Chicago: University of Chicago Press.

Young, A., Boyle, T. & Brown, T. (1996). The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology and Evolution, 11,* 413-418.

Ziter, C. (2014). The effect of forest fragmentation on aboveground carbon stocks and tree diversity: a case study of the Monteregie, Quebec. (Unpublished Master's Thesis). McGill University, Montreal.

Zobel, M. (1997). The relative role of species pools in determining plant species richness: an alternative explanation of species coexistence. *Trends in Ecology and Evolution, 12,* 266-269.

APPENDICES

Site	SSR loci						
	CL88	**LUP5**	**LUP52**	**LUP63**	**S082**	**S102**	**S175**
3	23	18	19	17	22	20	21
4	23	21	24	19	24	24	21
5	22	19	20	18	23	22	21
6	21	15	22	19	21	21	18
7	23	22	20	20	22	22	22
8	25	20	19	18	25	25	23
10	21	20	21	20	23	22	22
14	25	24	20	21	24	25	23
18	23	21	20	21	23	22	24
19	20	22	22	17	23	21	22
20	24	22	21	23	24	24	20
22	23	19	20	12	22	23	24
26	20	17	15	13	19	19	20
27	22	23	22	17	23	23	21
30	20	17	19	15	21	20	21
31	25	21	20	20	22	25	24
33	24	22	18	22	22	24	23
34	21	21	21	19	24	21	23
40	21	19	20	16	21	21	23
42	21	23	19	20	24	21	24

Appendix 1: Sample sizes for HWE and LD tests. Sample sizes reflect the number of individuals that genotyped successfully size for each loci/pop ranges from 15-25.

	SSR loci						
Site	**CL88**	**LUP 5**	**LUP52**	**LUP 63**	**S102**	**S082**	**S175**
	P-val	P-val	P-val	P-val	P-val	P-val	P-val
3	0.0135	1.0000	0.9916	0.0303	**0.0000***	0.3781	**0.0000***
4	0.9148	**0.0001***	0.3594	**0.0000***	0.0010	0.0652	0.0239
5	0.0167	N/A	0.9889	N/A	0.0234	N/A	**0.0000***
6	0.0705	N/A	0.0387	N/A	0.0039	N/A	0.2705
7	1.0000	N/A	0.1586	N/A	1.0000	N/A	0.0003
8	1.0000	N/A	0.1097	**0.0000***	0.0223	0.0117	**0.0001***
10	**0.0001***	N/A	0.0249	**0.0001***	N/A	1.0000	0.0263
14	1.0000	N/A	1.0000	**0.0001***	0.0460	N/A	0.0593
18	N/A	0.0259	0.0102	0.0005	0.0014	1.0000	**0.0000***
19	1.0000	N/A	1.0000	**0.0000***	0.0101	0.0020	0.0645
20	1.0000	1.0000	0.6630	**0.0000***	0.0781	1.0000	**0.0003***
22	1.0000	1.0000	1.0000	0.0441	0.0703	N/A	0.0545
26	N/A	N/A	0.6692	0.0399	0.0275	N/A	0.1383
27	0.0284	0.0242	**0.0000***	N/A	0.0016	N/A	0.8294
30	1.0000	N/A	0.0123	**0.0000***	**0.0000***	N/A	0.0015
31	0.0149	N/A	0.1585	0.0026	0.0004	N/A	**0.0001***
33	0.0018	N/A	1.0000	N/A	0.0219	N/A	0.0183
34	**0.0002***	0.0733	1.0000	**0.0000***	0.0160	1.0000	0.0018
40	1.0000	N/A	0.9996	0.0036	**0.0002***	1.0000	0.4382
42	0.0825	0.0015	0.0077	**0.0000***	0.0020	N/A	0.1622

Appendix 2: Hardy-Weinberg equilibrium exact test results. Significant values indicate deviation from HWE in the direction of heterozygote deficiency. An * indicates significant deviation after Bonferroni correction (modified p= 0.00036). N/A indicates that test could not be performed because locus was not polymorphic.

Locus pair		chi²	df	p-value
LUP5	LUP63	19.41106	22	0.61983
LUP5	CL88	23.02395	18	0.18967
LUP63	CL88	36.41131	34	0.35708
LUP5	S102	24.31616	18	0.14494
LUP63	S102	31.44164	34	0.59364
CL88	S102	33.45888	36	0.59006
LUP5	LUP52	14.64422	16	0.55083
LUP63	LUP52	31.77329	30	0.37815
CL88	LUP52	36.23983	36	0.45746
S102	LUP52	28.26739	36	0.81754
LUP5	S175	22.00022	22	0.45988
LUP63	S175	32.70808	32	0.43205
CL88	S175	49.65327	38	0.09765
S102	S175	25.69494	38	0.93605
LUP52	S082	69.07279	40	0.00291
LUP5	S082	6.07863	18	0.99587
LUP63	S082	18.22691	20	0.57246
CL88	S082	15.60139	22	0.83517
S102	S082	12.70653	20	0.88962
LUP52	S082	23.33447	22	0.38305
S175	S082	Infinity	22	Highly sign*

Appendix 3: Results for global Fisher's test of Linkage-Disequilibrium test. An * indicates significant deviation from linkage equilibrium after Bonferroni correction (modified p= 0.00239).

APPENDIX 4

Site descriptions. For each site we recorded general disturbances, evidence of maple-syrup production and sustainable timber production, dominant forest types, unusual topography, presence of rivers, surrounding habitat type and any other unusual conditions.

Site 3 is a very rural site bounded on all sides by agricultural land. The site is moist and dominated by *Acer rubrum* and *Fraxinus nigra*.

Site 4 is a small, dry site dominated by *Acer saccharum*, and also *Ostrya virginiana*. A large proportion of the site is managed as a sugar bush with sparse ground cover. The site extends to road Rang des Trente, while the other three edges of the site are bounded by agricultural land.

Site 5 is a very large rural site owned by many different landowners. It has a very hilly topography with many steep slopes. It had a large swampy area but was mostly dry and not very disturbed. Many landowners had cabins in the woods where they harvested timber or had sugar-bushes. The site was dominated by *Acer saccharum* and *Acer rubrum.*

Site 6 is a more urban site, bordering roads on two sides, and backing onto agricultural fields. A large portion of the site is swampy and dominated by *Acer rubrum* and *Fraxinus pennsylvanica*. The drier half of the site still had a lot of *Acer rubrum* and Acer *saccharum*. There were some trails through the site, and some regrowth that is very weedy and dense. There was no maple syrup production at this site.

Site 7 is small and in a very rural area, bounded on all sides by agricultural land. It is fairly dry and dominated by *Acer rubrum.* The site is largely undisturbed, with no maple syrup or timber production. In our surveys we surprisingly found no *Acer saccharum.*

Site 8 is Mont Saint-Hilaire, one of the Monteregian Hills. It has a very unique topography, compared to other sites and with a much higher elevation. It is mostly very dry and dominated by *Acer saccarhum* and *Fagus grandifolia*. There is a small lake, Lac Hertel which has small streams flowing into it. Mont-Saint-Hilaire is distinguished by having many rocky outcrops. Disturbances at this site include many trails, and also an overabundance of white tailed deer. There are many buildings and some roads on a portion of this site associated with McGill University's Gault Nature Reserve. In recent years, white-tailed-deer have devastated the site, and herbaceous ground cover is very low. The population of *Carex leptonervia* was small compared to the size of the site.

Site 10 was a small dry site in a rural setting bounded by agricultural fields. The site was dominated by *Acer rubrum* and *Fraxinus pennsylvanica*. Disturbances were few, and there was no sugar-bush. The site had one small swampy area in it, and was generally flat.

Site 14 is fairly dry and dominated by *Acer saccharum* with also abundant *Acer rubrum* and *Carya cordiformis*. This site is flat and dry and had a very active sugar-bush.

Site 18 is a small dry flat site dominated by *Acer saccharum*, *Acer rubrum*, and *Fagus grandifolia*. It is in a rural area and was bounded on all sides by agricultural fields. There was an extensive sugar-bush throughout it.

Site 19 is a small site in a very rural area surrounded on all sides by agricultural land. It has a path running throughout and is bisected by a small stream, but is overall dry. It is dominated by *Acer saccharum*, *Acer rubrum*, and with abundant *Tilia americana*.

Site 20 is a forest patch split by a road, so we sampled the smaller portion where we had landowner permission. The property is owned by many people, and has a high level of timber and maple-syrup production. The forest has some small streams running through it, but was mostly dry and had a canopy dominated by *Acer saccharum, Acer rubrum,* and *Fraxinus pennsylvanica.* In areas of heavily managed sugar bush, the ground cover was sparse.

Site 22 is a large site dominated by *Acer rubrum*, *Acer saccharum*, *Tilia americana* and *Fraxinus americana,* with numerous *Acer saccharinum* swamps throughout. This site has numerous landowners and some paths running throughout. It was bordered on one end by a road and surrounded by agricultural fields on the other sides.

Site 26 is a very large site with many different landowners. It is in a very rural area surrounded by agricultural land. It has a very large river cutting through it, and we sampled on both sides of the river. Since it was so large it had a very heterogeneous canopy, but was moly dominated by *Acer saccharum*, *Acer rubrum*, and *Fagus grandifolia*. There were cabins throughout the site, some trails, and signs warning of hunting, yet obvious disturbance appeared minimal.

Site 27 is an urban site very close to Mont-Saint-Hilaire. It had a large mown hydro line cutting through it, and an edge that was very disturbed with a lot of invasive species and trails. The interior was mostly moist forest with some standing water. Canopy was dominated by *Acer rubrum*, *Acer saccharum*, *Tilia americana* and had areas of *Acer sacharinum* swamp. It was owned by many different people but did not have apparent signs of maple syrup or timber production.

Site 30 is a site in a rural area bounded by agriculture on all sides. The site had minimum disturbances in it, but was bisected by a large field and thicket. The site was dominated by *Acer rubrum*, *Tilia americana*, and *Fraxinus pennsylvanica*. The site has no apparent sugar bush disturbances.

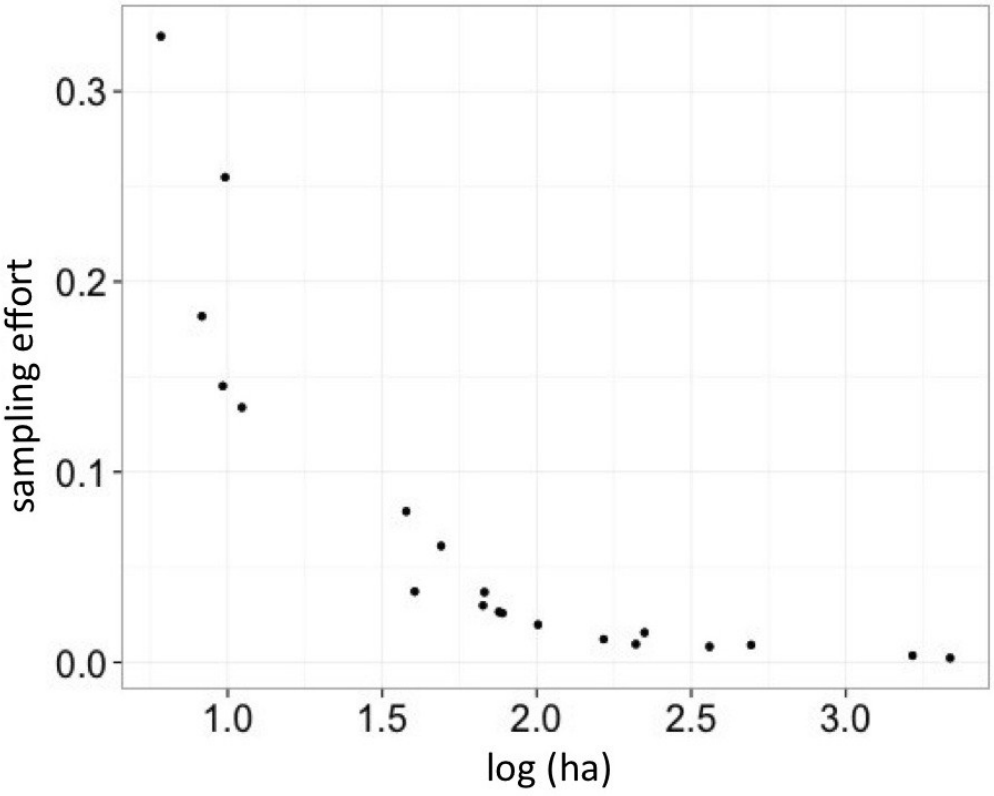
Site 31 is a very urban site very close to Mont-Saint-Hilaire. There were signs of camping at this site, but there were no sugar-bushes or cabins. One edge of the site is bordered a quarry and large lake and there were many gravelly paths near the quarry boundary. This large site has both *Fraxinus nigra* swamps and dry forests dominated by *Acer rubrum* and *Ulmus americana*. Though the site was big, the *Carex leptonervia* population was small and plants were scarce.

Site 33 is a very small site in a rural area, nearly backing onto highway 10. It had a large swampy area with standing water near the highway but is mostly dry and dominated by *Acer saccharum* and *Acer rubrum*. A small stream and culvert separate the dry portion of the forest from the wet. In the dry forest, ground-cover was sparse and we found extensive networks of groundhog burrows. There is a small trail going throughout the site, and aside from that anthropogenic disturbance appears minimal.

Site 34 is bordered by roads on two sides and has many anthropogenically-disturbed areas. There are trails running throughout the site and some apparent timber production. There are many invasive species at the edges. The site is mostly dry but has some small streams running through it. The canopy is dominated by *Acer rubrum*, with *Populus tremuloide*s and *Betula papyrifera*.

Site 40 is rural but borders onto a number of residential properties. There are numerous small roads and trails throughout the site. The forest is mostly dry and dominated by *Acer rubrum* and *Fraxinus americana* but the site had a swampy interior. The *Carex leptonervia* population extended throughout the site. There was one large sugar-bush with extensive tubing.

Site 42 is Rougemont, the second of our Monteregian Hill sites. It has a topography similar to Mont-Saint-Hilaire, and is relatively undisturbed, except for a well paved trail leading to the summit and an extensive trail and sugar-bush system operated by monks of the nearby Abbaye Notre-Dame-de-Nazareth. *Carex leptonervia* was throughout the site. The site was mostly dry and had a great diversity in forest type and canopy.



Appendix 5: Sampling effort vs forest stand area. Sampling effort is specific to each forest and was calculated by dividing the time spent at each site by the size of each site. Log (ha) is specific for each forest stand.

Axis	PCoA site	PCoA quadrats	CCA prism sweep
pH			
eigenvalue	N/A	N/A	0.10934
proportion explained	N/A	N/A	0.01652
cumulative proportion	N/A	N/A	0.01652
%OM			
eigenvalue	N/A	N/A	0.03517
proportion explained	N/A	N/A	0.0531
cumulative proportion	N/A	N/A	0.02183
1			
eigenvalue	0.3258	5.82419	0.52238
proportion explained	0.2022	0.09699	0.07892
cumulative proportion	0.2022	0.09699	0.10075
2			
eigenvalue	0.2292	4.12381	0.42004
proportion explained	0.1422	0.0687	0.06346
cumulative proportion	0.3444	0.16566	0.16421
3			
eigenvalue	0.71	3.20834	0.3804
proportion explained	0.1061	0.05343	0.05747
cumulative proportion	0.4505	0.21908	0.22168
4			
eigenvalue	0.10277	2.66536	0.3521
proportion explained	0.08824	0.04438	0.05319
cumulative proportion	0.53877	0.21908	0.27487
5			
eigenvalue	0.10277	2.65153	0.3319
proportion explained	0.06378	0.04416	0.05014
cumulative proportion	0.60255	0.30762	0.32501

Appendix 6: Eigenvalues for the first five ordination axes. The PCoA is for Hellinger-transformed individual quadrats. For the CCA axes 1-5 are values of the unconstrained axes. The constrained axes are pH and %OM.

	3	4	5	6	7	8	10	14	18	19	20	22	26	27	30	31	33	34	40	42
3	0																			
4	1.69	0																		
5	9.8	8.12	0																	
6	28.82	27.39	20.39	0																
7	26.11	24.56	16.98	4.78	0															
8	18.82	17.18	9.19	12.34	8.18	0														
10	27.84	26.41	19.44	0.99	4.3	11.52	0													
14	12.24	10.61	4.55	22.02	17.94	9.77	21.15	0												
18	5.66	6.26	12.82	33.03	29.79	21.94	32.06	13.39	0											
19	10.78	11.57	17.84	38.18	34.78	26.78	37.22	17.66	5.34	0										
20	7.58	9.23	17.33	36.14	33.62	26.4	35.15	19.25	7.08	8.51	0									
22	12.08	13.57	21.41	40.9	38.13	30.6	39.92	22.54	9.19	7.15	5.22	0								
26	4.32	5.83	13.82	33.14	30.38	22.97	32.16	15.57	4.04	7.61	3.69	7.77	0							
27	22.32	20.81	13.48	6.98	3.93	5.56	6.08	15.07	26.23	31.32	29.79	34.38	26.62	0						
30	20.47	21.57	28.24	48.56	45.19	37.17	47.6	27.89	15.56	10.41	15.12	10.18	16.46	41.72	0					
31	21.19	19.63	12.06	8.65	4.94	3.77	7.79	13.37	24.86	29.89	28.72	33.19	25.45	1.8	40.29	0				
33	36.43	34.8	26.82	11.6	11.04	17.64	12.21	26.52	39.47	44.11	44	48.23	40.6	14.92	54.41	15.52	0			
34	42	40.39	32.46	15.57	16.2	23.26	16.38	32.23	45.15	49.82	49.57	53.86	46.2	20.13	60.12	20.93	5.71	0		
40	28.43	26.93	19.56	2.35	2.88	10.99	2.42	20.75	32.34	37.4	35.87	40.49	32.73	6.13	47.8	7.52	10.07	14.63	0	
42	29.12	27.44	19.35	14.35	10.24	11.2	14.23	17.75	31.13	35.3	36.58	40.23	32.97	12.2	45.33	11.49	10.44	15.78	12	0

Appendix 7: Pairwise matrix of geographic distance between study sites. Measurement is in km, and was measured using the great-circle distance that accounts for the Earth's sphericity.

	3	4	5	6	7	8	10	14	18	19	20	22	26	27	30	31	33	34	40	42
3	0																			
4	0.53	0																		
5	0.58	0.13	0																	
6	0.33	0.32	0.39	0																
7	0.57	0.24	0.14	0.42	0															
8	0.54	0.27	0.17	0.41	0.11	0														
10	0.45	0.33	0.29	0.44	0.27	0.26	0													
14	0.45	0.26	0.25	0.39	0.23	0.26	0.13	0												
18	0.56	0.29	0.19	0.44	0.1	0.05	0.25	0.25	0											
19	0.54	0.22	0.15	0.39	0.15	0.14	0.28	0.24	0.16	0										
20	0.5	0.29	0.23	0.44	0.19	0.17	0.13	0.15	0.17	0.17	0									
22	0.43	0.27	0.25	0.36	0.22	0.23	0.18	0.13	0.23	0.18	0.14	0								
26	0.56	0.29	0.19	0.43	0.08	0.09	0.27	0.25	0.07	0.18	0.2	0.24	0							
27	0.36	0.34	0.4	0.31	0.41	0.43	0.32	0.3	0.43	0.44	0.4	0.34	0.41	0						
30	0.53	0.16	0.14	0.31	0.18	0.18	0.33	0.28	0.21	0.13	0.25	0.24	0.21	0.4	0					
31	0.51	0.12	0.19	0.24	0.27	0.29	0.39	0.32	0.31	0.24	0.34	0.3	0.3	0.36	0.13	0				
33	0.43	0.3	0.36	0.23	0.41	0.41	0.43	0.39	0.43	0.43	0.45	0.41	0.41	0.29	0.34	0.26	0			
34	0.42	0.32	0.39	0.21	0.44	0.42	0.46	0.42	0.45	0.44	0.48	0.43	0.43	0.33	0.35	0.27	0.07	0		
40	0.58	0.25	0.14	0.44	0.06	0.09	0.27	0.24	0.08	0.11	0.17	0.22	0.1	0.44	0.18	0.28	0.44	0.46	0	
42	0.54	0.15	0.13	0.39	0.16	0.23	0.27	0.2	0.23	0.18	0.22	0.21	0.22	0.36	0.19	0.22	0.38	0.42	0.17	0

Appendix 8: Pairwise matrix of Nei's genetic distance between sites (calculation follows Nei, 1972).

Code	Species
ABIBALS	*Abies balsamea*
ACEPENN	*Acer pennsylvanicum*
ACERUBR	*Acer rubrum*
ACESACI	*Acer saccharinum*
ACESASA	*Acer saccharum*
ACEXFRE	*Acer x freemanii*
ALNINCA	*Alnus incana*
AME_SP	*Amelanchier sp.*
BETALLE	*Betula alleghaniensis*
BETPAPY	*Betula papyrifera*
BETPOPU	*Betula populifolia*
CARCARO	*Carpinus caroliniana*
CARCORD	*Carya cordiformis*
CAROVAT	*Carya ovata*
FAGGRAN	*Fagus grandifolia*
FRAAMER	*Fraxinus americana*
FRANIGR	*Fraxinus nigra*
FRAPENN	*Fraxinus pennsylvanica*
JUGCINE	*Juglans cinerea*
OSTVIRG	*Ostyra virginiana*
PINSTRO	*Pinus strobus*
POPBALS	*Populus balsamifera*
POPDELT	*Populus deltoides*
POPGRAN	*Populus grandidenta*
POPTREM	*Populus tremuloides*
PRUSERO	*Prunus serotina*
QUEMACR	*Quercus macrocarpa*
QUERUBR	*Quercus rubra*
THUOCCI	*Thuja occidentalis*
TILAMER	*Tilia americana*
TSUCANA	*Tsuga canadensis*
ULMAMER	*Ulmus americana*
ULMRUBR	*Ulmus rubra*

Appendix 9: Trees growing in *Carex leptonervia* communities. Species codes used in the canonical correspondence analysis (CCA) are in the first hand column and full species names are reported in column 2.