LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AD	Anno Domini
AMOVA	Analysis of molecular variance
AR	Allelic richness
cpDNA	Chloroplast DNA
CS	Clone size
CZ	Continuous zone
DBH	Diameter at breast height
DNA	Deoxyribonucleic acid
DZ	Discontinuous zone
EWC	Eastern white cedar
FERLD	Forêt d'enseignement et de recherche du lac Duparquet
F_{ij}	Kinship coefficient
F_{is}	Inbreeding coefficient
FRQNT	Fonds de recherché du Québec- Natures et technologies
FSC	Forest Council Stewardship
F_{st}	Population genetic differentiation
GHG	Greenhouse gas

HCVFs	High Conservation Value Forests
H_e	Expected heterozygosity
H_o	Observed heterozygosity
H_s	Gene diversity
HWE	Hardy-Weinberg equilibrium
IAM	Infinite allele model
IBD	Isolation by distance
IPCC	Intergovernmental Panel on Climatic Change
LD	Linkage disequilibrium
MCMC	Markov chain Monte Carlo
MLGs	Multi-locus genotypes
MR	Multi-ramet genets
mtDNA	Mitochondrial DNA
MZ	Marginal zone
N_a	Number of alleles
ncDNA	Nuclear DNA
N_e	Number of effective alleles
NJT	Neighbor-joining tree
NSERC	Natural Sciences and Engineering Research Council of Canada

PA	Private allele
PCA	Principal coordinates analysis
PCR	Polymerase chain reaction
R_{el}	Relatedness
R_{elc}	Corrected relatedness
SGS	Spatial genetic structure
SMM	Stepwise mutation model
SR	Single-ramet genets
TPM	Two-phase model

RÉSUMÉ

Le cèdre blanc de l'Est (CBE, *Thuja occidentalis* L.) est une espèce de fin de succession de la forêt boréale du Canada. L'objectif à long terme de cette étude était de mieux comprendre l'influence du climat et des feux sur la dynamique et la structure génétique des populations de CBE. Nous avons utilisé des marqueurs moléculaires pour: 1) étudier la diversité génétique des populations de CBE suivant un gradient latitudinal, 2) mettre en évidence l'importance du mode de régénération le long d'une succession après feu et évaluer l'effet de l'historique des perturbations sur la structure génétique du CBE, 3) estimer les effets des caractéristiques du paysage sur sa diversité génétique et la valeur des peuplements résiduels non brûlés pour la conservation de la diversité. Nous avons observé une augmentation du taux de consanguinité dans les populations marginales et discontinues. Toutefois, nous n'avons noté aucun effet latitudinal sur la diversité génétique (H_s), la richesse allélique (RA), ou la différentiation des populations (F_{st}) le long d'un gradient allant de la forêt boréale mixte à la forêt coniférienne du Nord du Québec. L'isolement plus enlevé des populations au nord de l'aire de répartition continue de l'espèce n'est apparemment pas lié à un effet décelable sur la diversité génétique. De façon générale, les populations marginales fragmentées de CBE semblent être protégées de l'érosion génétique. Notre étude a mis en évidence un patron de régénération dynamique le long d'un gradient de succession de 250 ans après feu. Le taux de régénération asexuée augmente légèrement avec l'accroissement du temps écoulé depuis le dernier feu (1916, 21.8%; 1823, 27.0%; 1760, 30.9%). Le site le plus jeune (1916) reçoit une proportion plus importante de flux de gènes (82.4%) en provenance des deux autres sites. La structure génétique spatiale à l'échelle du peuplement (SGS) est significative chez les gaules (hauteur moyenne: 60± cm) et plus faible chez les arbres adultes. La régénération végétative chez les gaules augmente la SGS sur de courtes distances. La distance des sources de graines, le temps requis pour que l'arbre arrive à maturité et produise des graines ainsi que la présence de microsites favorables sont parmi les facteurs contrôlant l'abondance du CBE le long de la succession après feu. La comparaison de paysage plus ou moins fragmentés incluant des petits îlots préservés, des vielles forêts et des peuplements situés sur les îles du lac Duparquet a mis en évidence une dynamique source-réservoir associée à la présence d'un taux élevé de flux de gènes entre les peuplements de CBE. Des différences significatives ont été notées entre les trois types de paysage pour la richesse allélique (RA) et la différentiation entre les populations (F_{st}). Comparativement aux autres sites, les peuplements résiduels non brûlés ont une RA (5.06) et un nombre d'allèles privés (AP = 5) plus faibles et le niveau de différenciation (F_{st} = 0.052) le plus élevé. Aussi bien le vent que les caractéristiques du paysage façonnent la structure génétique du CBE à travers les différents paysages. Des mesures appropriées devraient être prises pour protéger les vieux massifs continus de CBE avant qu'ils ne forment des îlots résiduels de petites dimensions caractérisés par une variation génétique plus faible.

Mots clés: Forêt boréale, peuplements résiduels après feu, flux de gènes, diversité génétique, gradient latitudinal, paysage lacustre, marqueur microsatellite, limite nordique, succession après feu, *Thuja occidentalis* L.

ABSTRACT

Eastern white cedar (EWC, *Thuja occidentalis* L.) is an important late-successional tree species in the Canadian boreal forest. The long-term objective of this study was to improve the understanding of influence of climate and fire upon the dynamics and genetic structure of EWC populations. Specifically, we used molecular markers to: 1) examine latitudinal effects on its genetic diversity, 2) investigate the relative importance of the mode of regeneration along a post-fire succession as well as the effect of disturbance history on its genetic structure, 3) estimate the effects of landscape features on its genetic diversity, and the conservation value of fire residuals. We observed high inbreeding in marginal and discontinuous populations. There were no significant latitudinal effects on gene diversity (H_s), allelic richness (AR), or population differentiation (F_{st}) along a gradient from the boreal mixed-wood to northern coniferous forest in Quebec. Increased population isolation apparently did not correlate with a detectable effect on genetic diversity. Overall, the fragmented populations of EWC appear to be well-buffered against effects of inbreeding on genetic erosion. Our work revealed a dynamic regeneration pattern along a 250-year-long post-fire successional gradient. The percentage of asexual regeneration slightly increased with stand development (1916, 21.8%; 1823, 27.0%; 1760, 30.9%). The youngest site received a large proportion of gene flow (82.4%) from the two older sites. Fine-scale spatial genetic structure (SGS) was high and significant in saplings (mean height: 60± cm), and weaker in adult trees. Clonal growth in saplings increased SGS over short distances. The distance from seed sources, the time needed to produce seed-bearing trees and abundance of suitable microsites are among the factors controlling EWC abundance along post-fire succession. Our results showed a source-sink dynamic associated with the presence of a high level of gene flow between EWC stands including fragmented small EWC fire skips, naturally fragmented EWC islands and mainland EWC old forests. There were significant differences in allelic richness and population differentiation among the three landscape types with small fire skips having the lowest AR (5.06) and highest F_{st} (0.052), as well as fewest private alleles (PA = 5). Both wind and landscape features shaped the genetic structure of EWC among different landscapes. Appropriate measures should be taken to protect mainland EWC old forests before they turn to small patches that were characterised by reduced genetic variation.

Key words: Boreal forest, fire residuals, gene flow, genetic diversity, latitudinal gradient, lacustrine landscape, microsatellite marker, northern edge, post-fire succession, *Thuja occidentalis* L.

CHAPTER I

GENERAL INTRODUCTION

1.1 Problem Statement

Climate strongly shapes species range, demography and genetic diversity (Hewitt 2000; Thomas *et al.* 2004; Hoban *et al.* 2010). Both natural and anthropogenic causes, including the massive emission of greenhouse gas (GHG) and its accumulation in the atmosphere, drive climate change (IPCC 2007). Climate warming now occurs unequivocally, as global observation of widespread increase in temperature is recorded. This occurs, to even a greater extent, at higher northern latitudes (IPCC 2007). Consequently, one can envisage species range expansions, which will be marked most particularly for populations located at the limit of their range, which are more sensitive to rapid changes in climate (Davis & Shaw 2001; Iverson *et al.* 2004; Mimura & Aitken 2010). Marginal populations thus play an important role in the response to climate change when they hold evolutionary potential (Thomas *et al.* 2004; Hampe & Petit 2005). They may serve to originate further range expansion (Hunter Jnr & Hutchinson 1994).

Boreal trees species distribution usually becomes increasingly sporadic as one moves northward along a latitudinal gradient. Human activities superimpose upon climatic factors, contributing to population fragmentation at landscape scale (Lesica & Allendorf 1995; Eckert *et al.* 2008). The populations at the edge of their range are subjected to marginality which may affect population genetic diversity due to decreased gene flow, increased fragmentation (Diniz-Filho *et al.* 2009; Tollefsrud *et al.* 2009; Hoban *et al.* 2010), and to relatively small population size (Lönn & Prentice 2002).

Lack of reproduction may cause a decline in tree population. Sexual reproduction usually initiates re-colonisation following stand-replacing disturbances, *e.g.,* forest fires. However,

both climatic and biotic factors can limit sexual reproduction (Sirois 2000; Dorken & Eckert 2001; Frenne *et al.* 2012). These factors result in the reduction of survival rates at various stages in the life cycle (*e.g.*, pollination, seed dispersal, seed germination, seedling, establishment) (Woodward 1987). A progressive decline in success of sexual reproduction can lead to reduced capacity in these species to sustain themselves, as well as to colonise new sites (Tremblay *et al.* 2002; Messaoud *et al.* 2007).

Climate change especially affects the boreal forest, due to the direct impact of climate change upon species regeneration and such change's indirect effect on disturbance regime (Bergeron 1998; Kneeshaw *et al.* 2011). In Canada, wildfire has been recognised as the principal agent of disturbance in the boreal forest and it is characterised by high-intensity crown fires which initiate secondary succession processes in the burned areas (Heinselman 1981; Van Wagner 1983; Payette 1992; Bergeron 2000). Climate warming, which the accumulation of greenhouse gas induces, might triggers an increase in fire frequency (Bergeron *et al.* 2010). Previous study has reported that late successional species could invade burnt areas immediately following fire (Bergeron & Charron 1994) while their further recruitment is delayed (Simard *et al.* 1998).

Large-scale disturbances by fires also create a mosaic boreal landscape that comprises the most the burned and the little remaining forest lands. Tree stands that have escaped consecutive forest fires often are located in a specific habitat, (*e.g.*, protected from lake and watercourses, steep slope or humid habitats). These form fire residuals or patches of swamp forest that have attained old growth stages which early successional species have come to dominate. These old forest remnants, usually late-successional forest patches, provide refuges for disturbance-sensitive boreal species (Segerström 1997), and they serve as seed source for forest reestablishment (Asselin *et al.* 2001).

This work examines the relative impact of climate, disturbances, fragmentation, and the mode of recruitment on the genetic structure of eastern white cedar, a late-successionnal and fire-sensitive species in the boreal mixed-woods. This study integrates a reliable basic knowledge of disturbance history with the genetic data.

1.2 Literature Review

1.2.1 The Species

Eastern white cedar (EWC, *Thuja occidentalis* L.), native to North America, is a wind-pollinated conifer with a broad natural distribution (Figure 1.1) (Little 1971). The primary range extends from James Bay in the North to Tennessee and North Carolina in the South, and from the Gulf of St. Laurence in the East to the southeastern part of Manitoba to the West (Johnston 1990). Several states in United States and, within Canada, the province of Nova Scotia list it as an endangered species (Newell 2005; USDA 2013).

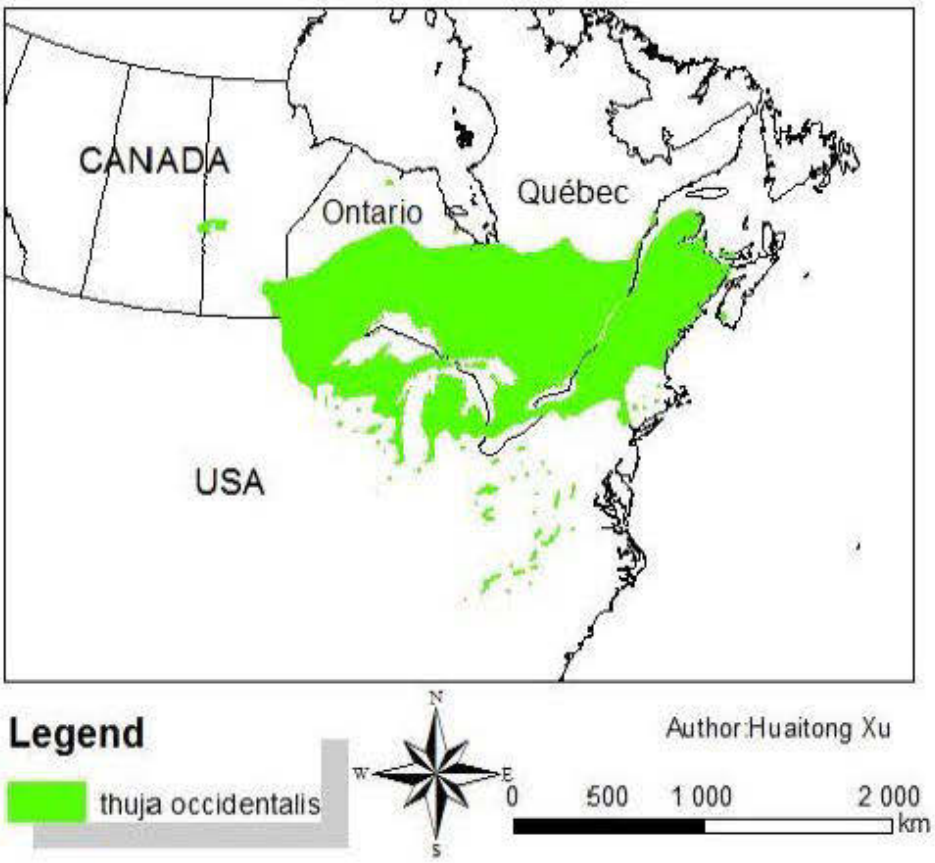


Figure 1.1 Distribution of Eastern white cedar (*Thuja occidentalis* L.)

The EWC is an evergreen monoecious and wind-pollinated tree species belonging to the cypress family (Cupressaceae) (Johnston 1990). It is a shadow-tolerant (Bakuzis & Hansen

1959), slow-growing conifer which may require about 80 to 100 years to reach a diameter of 18 cm on wet organic soils (Godman 1958). It can grow up to 15 to 38 meters high with conical crowns (Chambers 1993). The EWC has a chromosome number of 2n = 22 (Chambers 1993). Female and male flowers, which may occur separately on the same tree, usually develop in late April or May, with cone formation starting in late June; the cones mature in August; but they may remain attached to the tree until the following spring (Johnston 1990). The EWC's seed has two flat lateral wings with two cotyledons (Schopmeyer 1974). Seed production usually occurs early and it peaks after 75 years, with abundant seed crops produced every three to five years (Johnston 1990). Seed dispersal usually begins in September and ends in November, being disseminated mainly by wind. Some seeds continue to fall throughout the winter, and germination normally begins in late May or June (Johnston 1990).

1.2.2 Ecology

The EWC was considered at one time to be a short-lived tree species (Collingwood 1940). However, Archambault and Bergeron (1992) discovered old living EWC with age ranging from 500 to 900 years in Québec. The Niagara cliff-dwelling EWC are of old-growth forests, with trees exceeding 1000 years of age (Larson & Kelly 1991). Most of these individuals are small and stunted (Matthes-Sears & Larson 1995). The EWC seedlings have relatively high survival rates at low moisture conditions which makes them more able to tolerate extreme cliff conditions (Bartlett *et al.* 1991). In addition, specific portions of the root system link to specific portions of the shoot system (Larson *et al.* 1993). These characteristics contribute to their stress-tolerance to harsh environments.

Long-life characteristics (exceeding 1000 years of age) (Larson 2001) make EWC valuable for dendrochronological studies (Kelly *et al.* 1992), and for regional climate analyses (Kelly *et al.* 1994). The EWC's radial growth correlates with summer precipitation and temperature (Archambault & Bergeron 1992). Furthermore, Tardif and Stevenson (2001) have reported that radial growth negatively correlates with maximum August temperature. This result indicates that climate does not strongly influence EWC's radial growth at the limit of its western range in Manitoba. Buckley *et al.* (2004) have revealed a weakening relationship between climate and growth using a 2787 years-old tree-ring chronology from cliff-dwelling

EWC in Niagara, Ontario. The EWC that grows in mesic and xeric sites responds similarly to climate in Québec (Tardif & Bergeron 1997).

The EWC can be either an early or a late successional species. For example, it was found to colonise initially upland old fields (well-drained) and pasture lands in Maine (Curtis 1944). Scott and Murphy (1986) have reported EWC to be an early successional species colonising dunes in Lake Michigan. In the Canadian boreal forest, it is found more frequently at intermediate successional stages, in association with balsam fir (*Abies balsamea* [L.]Mill.), white spruce (*Picea glauca* [Moench] Voss) and with black spruce (*Picea mariana* [Mill.]B.S.P.), trembling aspen (*Populus tremuloides* Michx.) and white birch (*Betula papyrifera* Marsh.), and it generally dominates old-aged stands with balsam fir and white birch (Bergeron 2000; Chen & Popadiouk 2002).

Earlier findings suggest the existence of ecotypic differentiation among different sites (mesic, xeric) (Habeck 1958; Musselman *et al.* 1975). Matthes-Sears and Larson (1990) did not confirm this observation; they found a similar level of net photosynthetic rates in response to drought in both mesic and cliff environments. Further confirming their conclusion was the absence of differences in the level of nutrient, productivity, photosynthetic response to shading, and of allozymic variation between cliff and swamp sites (Matthes-Sears & Larson 1991; Matthes-Sears *et al.* 1991). Small architectural variations and the lack of difference in seed morphology were observed between lowland (wet) and upland (dry) sites (Briand *et al.* 1991; Briand *et al.* 1992).

1.2.3 Regeneration

Vegetative regeneration is common in EWC (Curtis 1946), which reproduces vegetatively by layering and more rarely by coppice and root suckers (Johnston 1990). Sexual reproduction is observed more frequently in stands where EWC is at low-density (Lamy *et al.* 1999).

The lack of recruitment has raised concern regarding EWC stands' sustainability in western Nova Scotia (Ringius 1979). In this region, EWC regeneration may become difficult due to high deer browsing pressure (Caulkins 1967). A large increase in deer population reduces EWC recruitment or regeneration (Heitzman *et al.* 1997). Deer, moose, and hare browsing on

it are detrimental to EWC regeneration in Québec (De Blois & Bouchard 1995; Bouchard & Domon 1997; Laroche *et al.* 2006). Late frosts, desiccation of seedlings, and yearly variation in seed production are all factors identified as detrimental to EWC regeneration in lowland forest terrain in the Lake States (Rooney *et al.* 2002).

The regeneration of EWC increased exponentially with the percentage of hummock (up to 70%) on wetland sites in Michigan (Chimner & Hart 1996). Seedling survival is higher on decaying logs than on bare mineral soil in the boreal mixed-woods in Québec (Simard *et al.* 2003). Successful seedling establishment correlates with the presence of dead and decaying, coarse, woody material in Minnesota in the mixed conifer-hardwood forest (Cornett *et al.* 1997).

Decaying wood is an excellent microsite for EWC seedling establishment. It facilitates adequate temperature, moisture, and fungal associations (Cornett *et al.* 2000). Water is the most important factor for safe seed beds (Cornett *et al.* 2000). The regeneration of EWC is highest on leaf litter seedbeds, followed by sphagnum moss and burn-type bryophytes (fire-succession mosses) (Davis *et al.* 1998). The presence of bryophyte on moderately decayed stumps acts as seed traps and helps to keep the moisture high for germination (Holcombe 1976). Heavy shade is detrimental to EWC seedling establishment (Davis *et al.* 1998), and so much shade can cause high mortality due to low light condition (Cornett *et al.* 2001).

1.2.4 Genetics

Genetic studies on EWC have reported a low level of population differentiation that associates with the presence of extensive gene flow between populations (Perry *et al.* 1990; Lamy *et al.* 1999). High selfing within populations also has been observed with populations essentially in Hardy–Weinberg equilibrium (Perry & Knowles 1990; Perry *et al.* 1990; Lamy *et al.* 1999). No allozymic variation has been observed between cliff and swamp sites (Matthes-Sears *et al.* 1991). Perry and Knowles (1991) have observed no significant difference in outcrossing rates among early, intermediate and late germinates in northern EWC populations. In contrast, Lamy *et al.* (1999) have reported a substantial level of genetic sub-structuring and an excess of homozygotes within small, isolated populations in southwestern Québec. The average genetic distances and level of differentiation between

populations were higher in southwestern Quebec than for populations in northern Ontario (Lamy *et al.* 1999). In the Canadian eastern range, no significant difference in genetic diversity has been found between peripheral and core populations, but fine-scale spatial genetic structure was higher in the core population (Pandey & Rajora 2012b; Pandey & Rajora 2012a).

1.3 Objective of This Study

The long-term objective of this study has been to improve understanding of the relative influence of climate and disturbances upon the dynamic and genetic structures of eastern white cedar in the boreal forest.

Specifically, the objectives were: (1) to examine the impact of population fragmentation along a latitudinal gradient upon the genetic diversity of EWC towards the northern edge of its range; (2) to understand the mode of colonisation and of invasion by EWC into boreal forest stands following forest fire; and (3) to investigate the genetic conservation value of EWC in fire residuals, and to understand better the effects of landscape features on its genetic structure.

1.4 Experimental Approach Used in the Research

In the study, we used polymorphic microsatellite markers to analyse EWC genetic diversity and structure.

In the first paper (chapter II), the study area divides into three bioclimatic zones on the basis of the abundance of EWC. The site occupation rates by EWC along the gradient toward its northern range limit have been estimated to be at 55%, 9%, and 3% in the continuous, discontinuous, and marginal zones, respectively. We selected a total of 24 populations covering the three zones (Continuous: CZ1 to CZ8; discontinuous: DZ1 to DZ7; marginal: MZ1 to MZ9). Population sizes range from less than one hundred individuals, in marginal and discontinuous zones, to thousands of individuals, in the continuous zone. Between 15 and 30 trees were randomly selected in each site, for a total of about 180 trees per zone.

In the second paper (chapter III), we selected three fire-initiated sites (1916 AD, 1823 AD, and 1760 AD) covering a 250 year-long post-fire succession. One single-hectare plot (100 x 100 m) was established in each site, from which we selected 10 EWC regeneration patches. Our study sampled a maximum of 100 saplings (up to 10 saplings per patch) in each site. For mature trees, we sampled EWC with DBH $\geqslant$ 35cm in 1760 (145 trees), DBH $\geqslant$ 25cm in 1823 (95 trees), and all trees in 1916 (14 trees, mean DBH = 16.8 cm).

In the third paper (chapter IV), we selected nine sites covering three different types of landscape. We selected three small EWC fire skips located in an area that last burned in 1944 to represent sites fragmented in a forest landscape. One of these sites (f443) has not burned since 1717, and the other two (f441, f442) have escaped forest fires since 880 cal. BP. The number of EWC trees ranged from 20 (f442) and 25 (f441) to <100 (f443). We selected three mainland old sites (f60, f97, f23) which had from 500 to more than 1000 EWC trees per hectare, which last burned in 1760, 1797, and 1823, respectively, in order to represent sites with different time elapsed since fire occurred in a non-fragmented forest landscape. Three naturally fragmented islands (> 100 EWC trees) – is39, is42, and is134 – last burned in ca 1889, ca 1825, and after ca 1825 (precise year unknown), respectively, were selected to represent a lacustrine landscape. We randomly selected between 20 and 30 EWC trees at each site.

In the fourth paper (Appendix I), we initially developed 117 microsatellite markers for EWC using shotgun pyro-sequencing, a next-generation sequencing technology performed on 454 GS-FLX Titanium platform. We then validated 16 out of 48 designed markers on a panel of 24 EWC individuals collected from 24 sites across northern Québec (one individual per site). We tested co-amplifications of all multiplexed primers on two populations (of 30 trees each) sampled from islands in Lake Duparquet, northwestern Québec.

I list the first three papers below in the order of presentation in the thesis.

1. Xu, H., Tremblay, F., Bergeron, Y., Paul, V. and Chen, C. (2012) Genetic consequences of fragmentation in "*arbor vitae*," eastern white cedar (*Thuja*

occidentalis L.), toward the northern limit of its distribution range. *Ecology and Evolution*, 2(10), 2501-2515. (Impact Factor 2012: 1.184)

In the said chapter we examined the impact of population fragmentation on the genetic diversity of EWC towards the northern edge of its range, and we tested the effect of latitudinal gradient on genetic diversity in EWC populations. H. Xu did the lab work, analysed the data, and wrote the manuscript. Drs. Tremblay and Bergeron conceived the study and edited the manuscript. V. Paul collected the samples. Dr. Chen gave some suggestions.

2. Xu, H., Tremblay, F. and Bergeron, Y. Sexual or asexual regeneration, which is more important for a late successional species, eastern white cedar (*Thuja occidentalis* L.)? (submitted to Annals of Botany) (IF 2012: 3.449)

In this chapter we examined the effect of disturbance history on the development of genetic structure in EWC along a chronosequence that covers a 250-year-long post-fire succession, and clarified the mode of colonisation and invasion by eastern white cedar into boreal forest stands following wildfires. H. Xu collected the samples, performed the laboratory work and the data analysis, and led the writing of the manuscript. Dr. Tremblay and Dr. Bergeron conceived the ideas, and provided suggestions during the manuscript preparations.

3. Xu, H., Tremblay, F. and Bergeron, Y. Importance of large old-growth forest patches in maintaining the genetic diversity of "*arbor vitae*", the eastern white cedar (*Thuja occidentalis* L.), in boreal fire-dominated landscapes (submitted to Conservation Biology) (IF 2012: 4.355)

In this chapter, we investigated the genetic diversity of eastern white cedar in its forest remnant which has survived successfully following consecutive wildfires. We also took advantage of the co-existence of three types of landscape (naturally fragmented EWC islands, fragmented small patches (fire skips), and non-fragmented mainland old EWC forests) to abet our understanding of effects of landscape features on genetic structure. H. Xu did the sampling in the field and genotyping using microsatellites in the laboratory, analysed the

data, and wrote the manuscript. Dr. Tremblay and Dr. Bergeron set the framework of the research and of the experimental approach and they edited the manuscript.

I list the fourth paper as Appendix I, which reports on the development of microsatellite markers for EWC.

4. Xu, H., Tremblay, F. and Bergeron, Y. (2013) Development and multiplexed amplification of SSR markers for *Thuja occidentalis* (Cupressaceae) using shotgun pyrosequencing. *Applications in Plant Sciences,* 1(5), 1200427. (IF: due 2015) (formerly *American Journal of Botany - Primer Notes and Protocols in the Plant Sciences,* IF 2011: 2.664)

In this article, we developed and validated sixteen microsatellite markers for EWC. H. Xu did sampling, performed the laboratory work and the data analysis, and wrote the manuscript. Dr. Tremblay and Dr. Bergeron made suggestions during the preparation of the manuscript.

CHAPTER II

GENETIC CONSEQUENCES OF FRAGMENTATION IN "*ARBOR VITAE*", EASTERN WHITE CEDAR (*THUJA OCCIDENTALIS* L.), TOWARD THE NORTHERN LIMIT OF ITS DISTRIBUTION RANGE

Article published in 2012

in *Ecology and Evolution*, 2(10), 2501-2515.

2.1 Résumé

Nous avons testé l'hypothèse suivant laquelle les populations marginales fragmentées de Cèdre blanc de l'Est (CBE, *Thuja occidentalis* L.) seraient génétiquement isolées en raison d'un flux de gènes réduit entre les populations. Les modèles prédisent une diminution de la diversité génétique intra-population et une augmentation de la différenciation inter-populations du centre vers les limites de l'aire de répartition d'une espèce. Dans cette étude, 24 populations de CBE ont été échantillonnées le long d'un gradient latitudinal allant de la forêt boréale mixte à la forêt résineuse nordique. Des analyses de génotypage ont été conduites à l'aide de marqueurs microsatellites. Les populations marginales et discontinues se caractérisent par des indices de fixation (F_{is}) positifs indiquant la présence d'un déficit d'hétérozygotes (populations marginales: F_{is} = 0.244; P_{HW} = 0.0042 et discontinues: F_{is} = 0.166; P_{HW} = 0.0042). Les populations au sud du gradient, dans la zone de répartition continue du CBE, étaient proches de l'équilibre de Hardy-Weinberg (F_{is} = -0.007; P_{HW} = 0.3625). Nous n'avons pas observé d'effet significatif de la latitude sur la diversité génétique (H_s), la richesse allélique (AR) ou sur la différenciation entre les populations (F_{st}). Des analyses bayésiennes et de groupement (neighbour-joining tree) ont montré que la structure génétique des populations était en partie expliquée par leur origine géographique. Les conséquences de la fragmentation sur la structure génétique des populations se traduisent par des coefficients de consanguinité positifs, deux fois plus élevé en moyenne pour les populations de la zone marginale en comparaison avec la zone discontinue. Ces résultats suggèrent la présence d'un plus fort taux de reproduction végétative au sein des populations fragmentées de CBE, ainsi que des échanges géniques plus fréquents entre des arbres avoisinants apparentés. L'accroissement de l'isolement des populations ne semblait pas avoir d'effet sur la diversité génétique des populations. Les populations marginales de CBE semblent donc être bien protégées contre les effets de la consanguinité sur la perte de diversité génétique.

2.2 Abstract

We tested the hypothesis that marginal fragmented populations of eastern white cedar (EWC) are genetically isolated due to reduced pollen and gene flow. In accordance with the central-marginal model, we predicted a decrease in population genetic diversity and an increase in differentiation along the latitudinal gradient from the boreal mixed-wood to northern coniferous forest. A total of 24 eastern white cedar populations were sampled along the north-south latitudinal gradient for microsatellite genotyping analysis. Positive F_{is} values and heterozygote deficiency were observed in populations from the marginal (F_{is} = 0.244; P_{HW} = 0.0042) and discontinuous zones (F_{is} = 0.166; P_{HW} = 0.0042). However, populations from the continuous zone were in HW equilibrium (F_{is} = -0.007; P_{HW} = 0.3625). There were no significant latitudinal effects on gene diversity (H_s), allelic richness (AR), or population differentiation (F_{st}). Bayesian and NJT (neighbour-joining tree) analyses demonstrated the presence of a population structure that was partly consistent with the geographic origins of the populations. The impact of population fragmentation on the genetic structure of EWC is the presence of a positive inbreeding coefficient, which was two times on average in marginal zone of that of populations from the discontinuous zone. This result indicated a higher occurrence of selfing within fragmented EWC populations coupled with a higher degree of gene exchange among near-neighbour relatives, thereby leading to significant inbreeding. Increased population isolation was apparently not correlated with a detectable effect on genetic diversity. Overall, the fragmented populations of EWC appear well-buffered against effects of inbreeding on genetic erosion.

2.3 Introduction

Climate is among the most important ecological processes that strongly shape the range and genetic diversity of a species (Hewitt 2000; Thomas *et al.* 2004; Sexton *et al.* 2009; Hoban *et al.* 2010; Provan & Maggs 2011). The well-documented central-marginal model (Diniz-Filho *et al.* 2009), which is also referred to as the abundant-centre model (Sagarin & Gaines 2002; Sagarin *et al.* 2006), predicts geographical variation in population genetic structure across a species' range (Loveless & Hamrick 1984; Yakimowski & Eckert 2008). Populations at the edge of their distribution range are subject to ecological marginality, which may affect population genetic diversity due to harsher environmental conditions (e.g., limited resources for growth and mating), isolation and fragmentation (Diniz-Filho *et al.* 2009; Tollefsrud *et al.* 2009; Hoban *et al.* 2010). Fragmented populations may be prone to genetic loss and increased genetic differentiation through drift (Ellstrand & Elam 1993; Young *et al.* 1996; Aguilar *et al.* 2008). However, these responses are unlikely to be universal. Long-lived plant species, such as trees, may be buffered against genetic effects for decades or centuries (Templeton & Levin 1979; Cabin 1996; Piotti 2009). Tree species combine life-history traits that promote a high level of gene flow between populations, the maintenance of a high within-population gene diversity and low population differentiation (Hamrick *et al.* 1992). Thus, the genetic consequences of recent alterations to mating systems in remnant fragments are sometimes not detectable for a long time (Gamache *et al.* 2003).

In the boreal forests of Canada, many tree species reach their continuous distribution range at the transition between the southern mixed-wood forests which are dominated by balsam fir (*Abies balsamea* [L.] Miller), and the northern coniferous forest which is dominated by black spruce (*Picea mariana* [Miller] BSP). The present-day transition between these two boreal zones is controlled by both climate and fire (Bergeron *et al.* 2004). Mixed-wood forests are characterised by smaller and fewer severe fire events than are coniferous forests (Hély *et al.* 2001). Large and severe fires induce high tree mortality that results in a disadvantage to mixed-wood forest species, which generally need survivor seed trees to reinvade burned areas (Asselin *et al.* 2001; Bergeron *et al.* 2004; Albani *et al.* 2005).

Paleoecological records indicate that the presence of mixed-wood forest species in the coniferous forest possibly represents the remnants of formerly larger populations and would thus result from the fragmentation of those initial populations. In western Québec, postglacial colonisation occurred rapidly after the retreat of proglacial Lake Ojibway (8400 cal. BP) and involved all of the tree species that are presently found within the area (Richard 1980; Liu 1990; Carcaillet *et al.* 2001). Since 7000 cal. BP, balsam fir and black spruce have dominated the mixed-wood and coniferous forests, respectively (Garralla & Gajewski 1992; Gajewski *et al.* 1996; Carcaillet *et al.* 2001). The decline in the number of mixed-wood forest species could be related to the climatic shift that characterized the beginning of the Neoglacial period and the establishment of cooler and drier summers coincident with an increase in fire frequency in the coniferous forest 3000 cal. BP (Carcaillet *et al.* 2001; Ali *et al.* 2009).

(a) Eastern white cedar (b) Foliage and cones of eastern white cedar

Figure 2.1 Eastern white cedar (*Thuja occidentalis* L.)

Eastern white cedar (*Thuja occidentalis* L.) (Fig. 2.1) is ill-adapted to fire and needs a protected area to reinvade burned areas. This species does not regenerate easily after fire, and population fragmentation following such a disturbance greatly limits its natural distribution. Eastern white cedar (EWC) reaches its northernmost distribution limit in the James Bay region of Québec at the ecotone of the mixed-wood and coniferous forest, at which point its

distribution becomes increasingly sporadic as one moves northward along a latitudinal gradient.

In this study, we examined the impact of population fragmentation on the genetic diversity of EWC towards the northern edge of its range. Previous genetic studies have been based on studying allozymes and showed contrasting results. Lamy et al. (1999) reported the presence of a substantial level of genetic sub-structuring (F_{st}=0.073) within six EWC populations. In contrast, Perry and others showed that six northern populations were not differentiated (F_{st}=0.016) and, indeed, were in Hardy-Weinberg equilibrium (Perry & Knowles 1989; Perry & Knowles 1990; Perry *et al.* 1990; Perry & Knowles 1991).

Our main hypothesis is that populations of EWC are more genetically isolated toward the northern edge of this species' range, due to reduced pollen and gene flow between populations. We tested whether population differentiation increases and genetic diversity decreases from continuous to discontinuous and to the peripheral part of the species' distribution range. Understanding the genetic structural pattern of ecotonal populations is important because remnant marginal stands that have been eroded from larger populations that were present during the early Holocene might be at the forefront of range expansion driven by climatic changes. The amount and structure of genetic variation within these remnant populations will likely affect their potential to respond to climatic changes.

2.4 Methods

2.4.1 Study Area and Materials

Eastern white cedar, which is native to North America, is a wind-pollinated, monoecious, evergreen conifer species (Fowells 1965). An abundant seed crop occurs every 3-5 years, with cones opening in the autumn, but seeds may continue to fall throughout winter. Sexual maturity is generally reached at an early age, but effective seed dispersal is observed after age 20 years. Most seeds are disseminated by wind, with seed dispersal distances with estimates ranging from 45 to 60 m (Fowells 1965). Eastern white cedar is a long-lived species, which can live up to 800 years in Quebec (Archambault & Bergeron 1992).

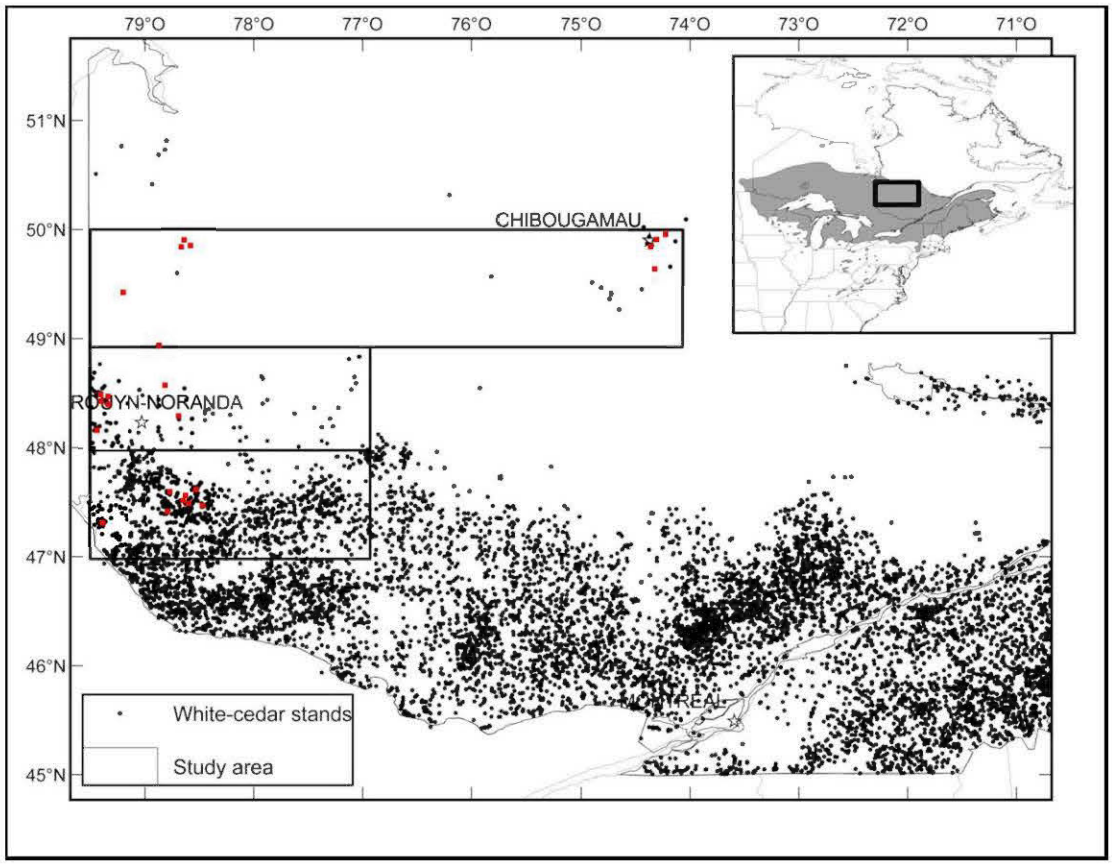


Figure 2.2 Distribution of Eastern white cedar in Quebec and North America (shaded region in the map inset) and sampling sites were doted in red; the study area is divided according to bioclimatic zone (Paul 2011).

The study area is located in the Abitibi-Témiscamingue and Nord-du-Québec regions of Quebec and is divided into three bioclimatic zones based on the abundance of EWC (Fig. 2.2). The continuous zone falls into the balsam fir (*Abies balsamea* [L.] Mill.) and yellow birch (*Betula alleghaniensis* Britton) bioclimatic domain and represents an area where eastern white cedar is common. The discontinuous zone is in the balsam fir and white birch (*Betula papyrifera* Marsh.) bioclimatic domain and marks the northern edge of the continuous distribution, where eastern white cedar becomes less common in the forest matrix. The marginal zone is in the black spruce (*Picea mariana* [Mill.] B.S.P.) and feather moss bioclimatic domain, where only a few isolated populations are found. The site occupation rates by EWC along the gradient were estimated to 55%, 9% and 3% in the continuous, discontinuous and marginal zones, respectively (Paul 2011).

A total of 24 populations were selected: eight in the continuous zone (Témiscamingue, CZ1 to CZ8), seven in the discontinuous zone (Abitibi, DZ1 to DZ7) and nine in the marginal zone (Chibougamau, MZ1 to MZ4; James Bay, MZ5 to MZ9) (Table 2.1). Population sizes range from less than one hundred individuals in marginal and discontinuous zones to thousands of individuals in the continuous zone, with the exception of two marginal populations (MZ6, MZ2) that had 8 and 11 trees, respectively. The distance between one population and its nearest neighbour ranges from about 2 to 70km, except for populations in Chibougamau (MZ1-MZ4), which were located about 300km from others in the marginal zone. Between 15 and 30 trees were randomly selected in each site, for a total of about 180 trees per zone; we retained marginal populations MZ6 and MZ2 in the analysis. Foliage was collected from individual trees in each population and used for DNA analysis.

Table 2.1 Genetic variability in 24 Easter White Cedar populations

Location	Pop	Latitude	Longitude	*N*	*AR*	N_a	N_e	H_o	H_e
Marginal zone (MZ)									
Chibougamau	MZ1	49.8754	-74.3928	21	6.1	8.0	4.6	0.655	0.756
	MZ2	49.90916	-74.3226	11	6.4	7.0	5.2	0.727	0.794
	MZ3	49.95351	-74.2291	20	6.6	8.0	6.1	0.463	0.826
	MZ4	49.64176	-74.3341	18	6.9	8.3	6.4	0.639	0.840
James Bay	MZ5	48.92772	-78.8858	30	6.6	9.5	6.2	0.661	0.815
	MZ6	49.42317	-79.211	8	5.3	5.3	4.0	0.813	0.740
	MZ7	49.85853	-78.6072	20	6.3	8.3	5.2	0.638	0.797
	MZ8	49.88349	-78.6461	25	6.7	10.0	5.1	0.680	0.798
	MZ9	49.85609	-78.6449	24	6.1	8.8	4.9	0.740	0.781
	Mean	-	-	20	6.3	8.1	5.3	0.668	0.794
	Pooled	-	-	177	11.5	11.5	7.8	0.657	0.867
Discontinuous zone (DZ)									
Abitibi	DZ1	48.5402	-78.6419	30	6.0	8.3	4.9	0.683	0.789
	DZ2	48.47015	-79.4524	24	6.1	8.3	4.8	0.625	0.789
	DZ3	48.47979	-79.4368	25	5.4	7.5	4.2	0.720	0.753
	DZ4	48.43161	-79.4018	28	5.4	8.0	4.1	0.759	0.743
	DZ5	48.26296	-78.5748	25	6.0	8.3	5.1	0.620	0.759
	DZ6	48.43101	-79.3842	25	6.4	8.5	5.3	0.630	0.805
	DZ7	48.20132	-79.4191	19	5.2	6.5	4.0	0.882	0.728
	Mean	-	-	25	5.8	7.9	4.6	0.703	0.767
	Pooled	-	-	176	11.8	11.8	6.3	0.697	0.834
Continuous zone (CZ)									
Témiscamingue	CZ1	47.42922	-78.6785	30	5.6	7.8	4.4	0.842	0.771
	CZ2	47.41669	-78.6821	27	5.2	6.8	3.9	0.796	0.712
	CZ3	47.39557	-78.7316	26	4.6	6.0	3.5	0.827	0.714
	CZ4	47.34505	-79.3926	15	4.6	5.0	3.6	0.850	0.721
	CZ5	47.3111	-78.5155	23	5.5	7.3	4.3	0.870	0.744
	CZ6	47.45395	-78.5877	30	6.2	8.8	5.6	0.883	0.801
	CZ7	47.41894	-78.6784	29	6.9	9.5	6.4	0.793	0.826
	CZ8	47.41579	-78.7117	18	6.1	8.0	5.0	0.778	0.780
	Mean	-	-	25	5.6	7.4	4.6	0.830	0.759
	Pooled	-	-	198	12.5	13.5	6.3	0.831	0.823

N_a, average number of alleles per locus; N_e, average number of effective alleles per locus; H_o, observed heterozygosity; H_e, expected heterozygosity; *AR*, allelic richness; *N*, number of individuals genotyped per population.

2.4.2 DNA Extraction, Microsatellite Loci Amplification and Genotyping

Foliage samples were ground, and genomic DNA was extracted using the GenElute Plant Genomic DNA Miniprep Kit (Sigma-Aldrich, St.-Louis, MO, USA). Amplification was performed by a gradient polymerase chain reaction (PCR) in a total volume of 10 µL using a 96-well GeneAmp PCR System 9700 (Applied Biosystems, California, USA). Each reaction mixture contained 2.5 µL of DNA extract, 2.5 mM $MgCl_2$, 1 pmol each of forward and reverse primers, 0.2 µL of 10 mM dNTP Mix, 1 µL 10X NovaTag Hot Start Buffer and 0.25 U NovaTag Hot Start DNA Polymerase (Novagen PCR Kit, Madison, WI, USA). The best results were obtained by performing a touchdown PCR that decreased the annealing temperature by 0.2°C every other cycle. At the end of each cycle, we added a final 72°C extension step. Loci developed by O'Connell and Ritland (2000) for *Thuja plicata* and by Nakao *et al.* (2001) for *Chamaecyparis obtusa* were utilized for microsatellite genotyping. Four loci exhibited high polymorphism (Table S2.1). Prior to electrophoresis, 0.5 µL of fluorescent dye-labelled PCR products were mixed with 0.25 µL of internal standard (MapMarker-1000) and 10 µL of deionized formamide. The loading products were heat denatured at 95°C for 3 min, immediately placed on ice for 5 min, and separated using capillary electrophoresis on an ABI Prism 3130 Genetic Analyzer (Applied Biosystems). Microsatellites were sized and genotyped using GeneMapper 3.7 (Applied Biosystems).

2.4.3 Descriptive Statistics

Micro-Checker software (Van Oosterhout *et al.* 2004) was used to detect null alleles and large allele dropouts at each locus for each population. We used the program FreeNA to estimate the frequencies of putative null alleles [r] and genetic differentiation [F_{st}] with and without ignoring the null alleles at each locus (Chapuis & Estoup 2007). Allele frequency, allele number and genetic estimates within populations including the average number of alleles per locus [N_a], average number of effective alleles per locus [N_e], observed heterozygosity [H_o], and expected heterozygosity [H_e] were calculated using GenAlex v. 6.2 (Peakall & Smouse 2006). We also calculated allelic richness [AR] using rarefaction and the inbreeding coefficient [F_{is}] at each locus. The calculations were performed using FSTAT v. 2.9.3 (Goudet 2001). We also calculated the aforementioned genetic estimates on pooled samples for each zone. Hardy-Weinberg equilibrium was tested in each population. We also

ran a global test of Hardy-Weinberg equilibrium for pooled samples from three distribution zones and for all pooled samples as a group. Bonferroni correction (Rice 1989) was applied when testing the significance of heterozygosity deficit and heterozygosity excess. All of the HW equilibrium tests were performed in FSTAT v. 2.9.3 (Goudet 2001).

2.4.4 Latitudinal Effects on Genetic Estimates

We tested for latitudinal effects by comparing differences in population genetic estimates among the three zones (marginal, discontinuous, and continuous). The genetic estimates that we compared included *AR*, H_o (Nei 1987), gene diversity [H_s] (Nei 1987), F_{is} (Weir & Cockerham 1984), F_{st} (Weir & Cockerham 1984), relatedness [R_{el}], and corrected relatedness [R_{elc}]. We applied Hamilton's (1971) measure of relatedness, which was calculated using an estimator that was strictly equivalent to the one proposed by Queller and Goodnight (1989). To avoid bias in relatedness when inbreeding exists, we applied the corrected relatedness of Pamilo (1984; 1985). All calculations and subsequent comparisons using a permutation procedure (10,000 iterations) were performed using FSTAT v. 2.9.3 software followed the statistics of its documentation (Goudet 2001).

2.4.5 Population Genetic Structure

To reveal genetic structure, and test if the samples could be clustered according to their respective distribution zones, we used STRUCTURE v. 2.3.2 software (Pritchard *et al.* 2000). Individuals were assigned to a number of assumptive clusters (assumptive groups) (K) ranging from 1 to 15 with an admixture model and the option of correlated allele frequency (Falush *et al.* 2003). All parameters were set following the user's manual. To choose an appropriate run length, we performed a pilot run that showed that burn-in and MCMC (Markov chain Monte Carlo) lengths of 300,000 each were sufficient to obtain consistent data. Increasing the burn-in or MCMC lengths did not improve the results significantly. Ten replicate runs for each value of K were carried out. The most likely value of K was selected by plotting ΔK following the ad hoc statistics (Evanno *et al.* 2005). The STRUCTURE results were graphically displayed using DISTRUCT (Rosenberg 2004). A neighbour joining tree analysis (Saitou & Nei 1987) was also used to analyse the genetic structure of our samples. The neighbour joining tree was visualised using TreeView software (D.m.page

1996) based on Nei's standard (Nei 1987) genetic distance, Ds, calculated using POPULATIONS v. 1.2.30 (http://bioinformatics.org/~tryphon/populations/). The neighbour joining tree was bootstrapped 1,000 times.

We determined the overall level of genetic differentiation using analysis of molecular variance (AMOVA) (Excoffier *et al.* 1992). The genetic distance matrix based on pairwise F_{st} (Weir & Cockerham 1984) was used to carry out the AMOVA using Arlequin v. 3.11 (Excoffier *et al.* 2005), with 10000 permutations. AMOVA was performed without grouping populations, with grouping populations by assigning them to three geographic zones, and with grouping populations by assigning them to a number of genetic groups that were identified by STRUCTURE v. 2.3.2 (Pritchard *et al.* 2000). We also performed a separate AMOVA on data from each of the three distribution zones. The geographic distance matrix was calculated using PASSaGE2 software (Rosenberg 2011). A Mantel test (Mantel 1967) was applied to analyse the correlation between the geographic distance and Nei's standard genetic distance (Nei 1987). All Mantel tests were performed using GenAlex v. 6.2 (Peakall & Smouse 2006).

2.4.6 Population Genetic Bottleneck

We tested for a recent population genetic bottleneck using the program BOTTLENECK v. 1.2.02 (Piry *et al.* 1999). An infinite allele model (IAM) and one-step stepwise mutation model (SMM) were applied in the bottleneck program (Cornuet & Luikart 1996). Because all loci were in-between, we finally used the option of a two-phase model (TPM) (Di Rienzo *et al.* 1994) with 95% SMM and 5% IAM and a variance of 12, as recommended by Piry *et al.* (1999). Wilcoxon's test, which is better adapted to a dataset with few polymorphic loci (our case), has a robustness similar to the sign test and is as powerful as the standardised differences test, was used to test the significance of the heterozygosity excess (Piry *et al.* 1999). A graphical descriptor was also used to distinguish between stable and bottlenecked populations (Luikart *et al.* 1998). We complemented the results of heterozygosity excess and mode-shift tests with Bayesian MSVAR (Beaumont 1999; Storz & Beaumont 2002; Girod *et al.* 2011). MSVAR assumes that microsatellite data evolve by a stepwise mutation model and it relies on Markov Chain Monte Carlo (MCMC) simulation to estimate the posterior

distribution of parameters that describe the demographic history (Beaumont 1999). The parameters of interest in our study were current population size (N_0), ancestral population size at the time population started to decline or expand (N_1), and time (in generations) since population started to decline or expand (T). The change in population size was determined by the ratio r ($r=N_0/N_1$) where $r < 1$ indicates decline, $r = 1$ indicates stability, and $r > 1$ indicates expansion (Beaumont 1999). As the generation time for EWC is unknown, we used a value of 20 years, given that its effective seed dispersal is observed after age 20 (Fowells 1965). The exponential model was applied. The length of run for chains was determined by Raftery-Lewis statistic (Raftery & Lewis 1992; Raftery & Lewis 1995). Two-hundred million iterations were sufficiently long for each chain to converge, with every 10 000th sample points being stored. The first 10% of data points were discarded from chains as burn-in to achieve stable simulations. The output was analysed with CODA 0.14-7 package implemented in R version 2.15.0 (http://cran.r-project.org/).

2.5 Results

2.5.1 Descriptive Statistics

The number of alleles per locus ranged from 10 (Locus TP10) to 17 (Locus TP12) (Table S2.1). Our results showed that all four loci were highly polymorphic (Table S2.2). The number of alleles per locus ranged from 8 at locus TP10 in the populations from the discontinuous distribution zone to 17 at locus TP12 in populations from the continuous distribution zone (Table S2.2). All loci exhibited positive F_{is} except for locus TP10 (Table S2.1). MICRO-CHECKER detected the presence of null alleles at loci TP9, TP11 and TP12, and there was no evidence for large allele dropout or scoring errors due to stuttering. Null alleles occurred at very low frequencies, and similar levels of genetic differentiation (F_{st}) were obtained when either excluding or not excluding the null alleles (Table S2.1).

At the population level, *AR* averaged 5.9 and ranged from 4.6 (CZ3, CZ4) to 6.9 (MZ4, CZ7). N_a ranged from 5.3 (MZ6) to 10.0 (MZ8), with an average of 7.8. The mean N_e was 4.9, with lowest value being 3.5 (CZ3) and the highest being 6.4 (MZ4, CZ7). H_o had a mean value of

0.7 and was lowest in population MZ3 (0.463) and highest in population CZ6 (0.883). The mean H_e was 0.77, ranging from 0.712 (CZ2) to 0.826 (MZ3, CZ7) (Table 2.1).

When populations were pooled, *AR* was quite similar among the three distribution zones (11.5, 11.8, and 12.5), as was N_a. H_o showed an increase from the marginal zone (0.657) to the discontinuous zone (0.697), further, to the continuous zone (0.831) (Table 2.1). The populations from the continuous distribution zone had the highest proportion of rare alleles (frequency < 1%; 0.148) and the highest total number of alleles (54) across the loci; the populations with the second highest proportion were from the discontinuous distribution zone (0.106; 47), and the populations with the least were from the marginal distribution zone (0.065; 46) (Table S2.2). Only populations from the continuous distribution zone had private alleles (one at locus TP10 and TP12) (Table S2.2).

2.5.2 Latitudinal Effects on Genetic Estimates

Among the 24 populations, seven (four marginal: MZ3, MZ4, MZ5, MZ7; three discontinuous: DZ2, DZ5, DZ6) showed a significant deficiency of heterozygotes and a departure from Hardy-Weinberg equilibrium (data not shown). None of the populations from the continuous distribution zone exhibited significant departure from HW equilibrium (data not shown). When populations were pooled, the global HW test revealed a significant departure from equilibrium and a slight heterozygote deficiency (F_{is}=0.145; P_{HW}=0.0125). Positive F_{is} values and heterozygote deficiency were also observed in populations from the marginal (F_{is}=0.244; P_{HW}=0.0042) and discontinuous (F_{is}=0.166; P_{HW}=0.0042) distribution zones. However, populations from the continuous zone were in HW equilibrium (F_{is}=-0.007; P_{HW}=0.3625) (Table 2.2).

Table 2.2 Hardy-Weinberg equilibrium test

Region	F_{is}	Heterozygosity deficit	Heterozygosity excess	*P* value
Marginal zone	0.244	*	N/A	0.0042
Discontinuous zone	0.166	*	N/A	0.0042
Continuous zone	-0.007	N/A	ns	0.3625
Global	0.145	*	N/A	0.0125

N/A, not applicable; ns, not significant; * $P < 0.05$; Bonferroni corrections were applied.

The difference in H_o among the populations from the three zones was highly significant (P=0.003), as were differences for F_{is} (P=0.002) and R_{elc} (P=0.005). We did not find any significant differences for AR, H_s, F_{st} and R_{el} among the populations from the three zones (Table 2.3).

Table 2.3 Comparisons of genetic estimate differences among populations of *Thuja occidentalis* from three zones

	AR	H_o	H_s	F_{is}	F_{st}	R_{el}	R_{elc}
Marginal zone	6.334	0.657	0.823	0.202	0.060	0.096	-0.505
Discontinuous zone	5.805	0.697	0.786	0.112	0.070	0.119	-0.253
Continuous zone	5.589	0.831	0.777	-0.070	0.066	0.132	0.130
P values	ns (0.072)	** (0.003)	ns (0.153)	** (0.002)	ns (0.926)	ns (0.702)	** (0.005)

ns, not significant; *$0.01 \leq P < 0.05$, significant; ** $P < 0.01$, highly significant; P values were obtained after 1000 permutations; AR, allelic richness; H_o, observed heterozygosity; H_s, gene diversity. F_{is}, inbreeding coefficient; F_{st}, population differentiation; R_{el}, relatedness; R_{elc}, corrected relatedness.

Further comparisons revealed that the difference in H_o was not significant between populations from the marginal and discontinuous zones. It was significantly different between the discontinuous and continuous zones (P=0.010) and between the marginal and continuous zones (P=0.001) (data not shown). Similarly, the differences between populations for F_{is} and R_{elc} were only significant between the discontinuous and continuous zones (F_{is}, P=0.027; R_{elc}, P=0.052) and between the marginal and continuous zones (F_{is}, P=0.001; R_{elc}, P=0.001) (data not shown).

2.5.3 Genetic Structure Patterning

Bayesian analysis demonstrated the presence of population structure. The three clusters detected by STRUCTURE (Fig. S2.1) are displayed in orange, yellow, and blue. The largest cluster (yellow) includes 14 populations crossing the three zones (MZ5, MZ6, MZ7, MZ8, MZ9, DZ1, DZ2, DZ3, DZ4, DZ5, CZ5, CZ6, CZ7 and CZ8). The cluster depicted in blue includes five populations: four from southern sites in Témiscamingue (CZ1 to CZ4) and one from the discontinuous zone (DZ7). The cluster depicted in orange includes four populations from the northern sites (MZ1 to MZ4) and DZ6 in the discontinuous zone (Fig. 2.3). Most of the individuals from the marginal Chibougamau populations and population DZ6 from the

discontinuous zone (Abitibi) were assigned to only one cluster. Similarly, almost all individuals from the Témiscamingue populations (CZ1 to CZ4) were assigned to only one cluster.

The results of the NJT that were based on Nei's (Nei 1987) standard genetic distance (Ds) were partially consistent with the geographic origins of the populations (Fig. 2.4). Four clusters can be identified at increased confidence levels (bootstrap values $\geqslant$ 50). Two of these clusters were also identified using STRUCTURE. MZ1, MZ2, MZ3, MZ4 and DZ6 were assigned to one cluster, while CZ1, CZ3, CZ2, and CZ4 were assigned to another cluster.

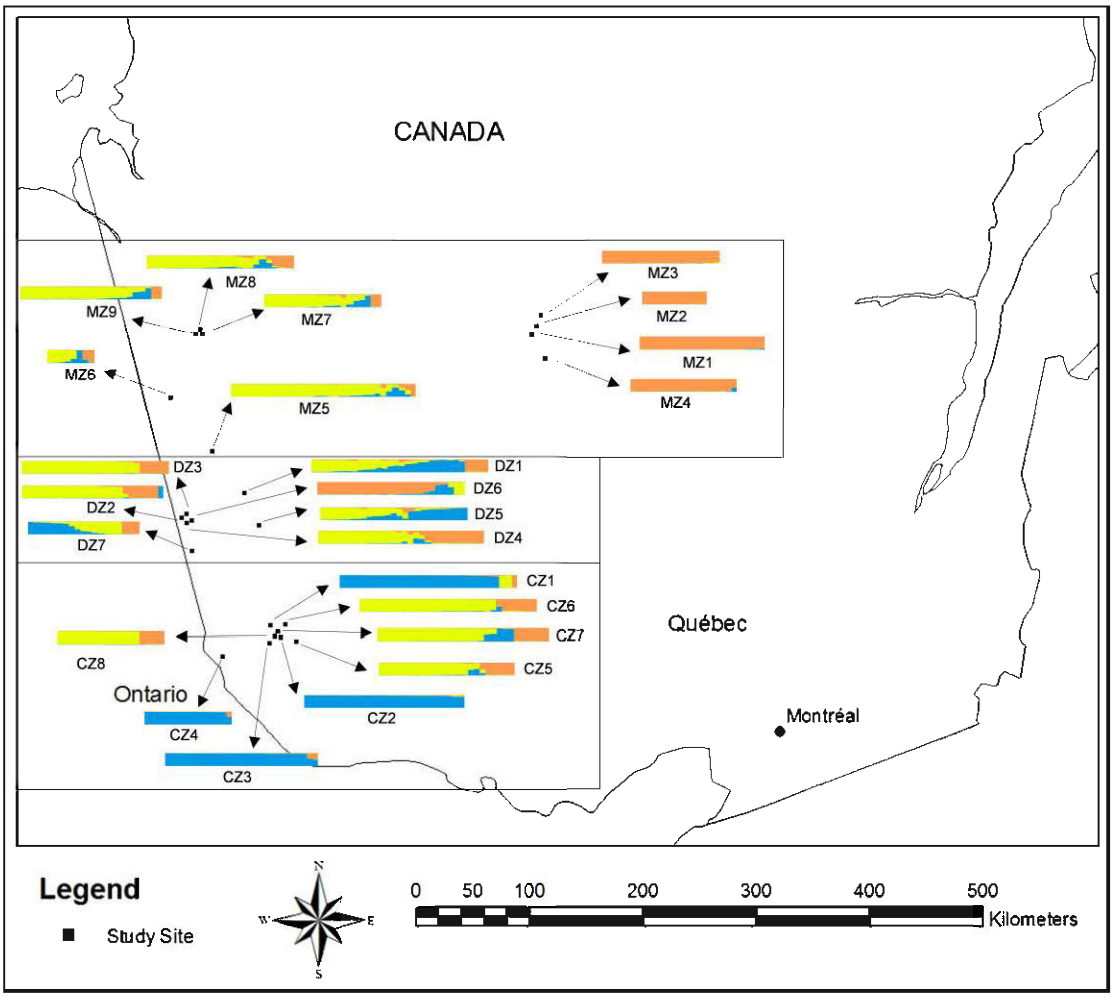


Figure 2.3 Study site, geographic origin, and genetic structure of *Thuja occidentalis* populations deduced by STRUCTURE at K = 3 (Orange cluster: MZ1, MZ2, MZ3, MZ4, and DZ6; yellow: MZ5, MZ6, MZ7, MZ8, MZ9, DZ1, DZ2, DZ3, DZ4, DZ5, CZ5, CZ6, CZ7, and CZ8; blue: CZ1, CZ2, CZ3, CZ4, and DZ7)

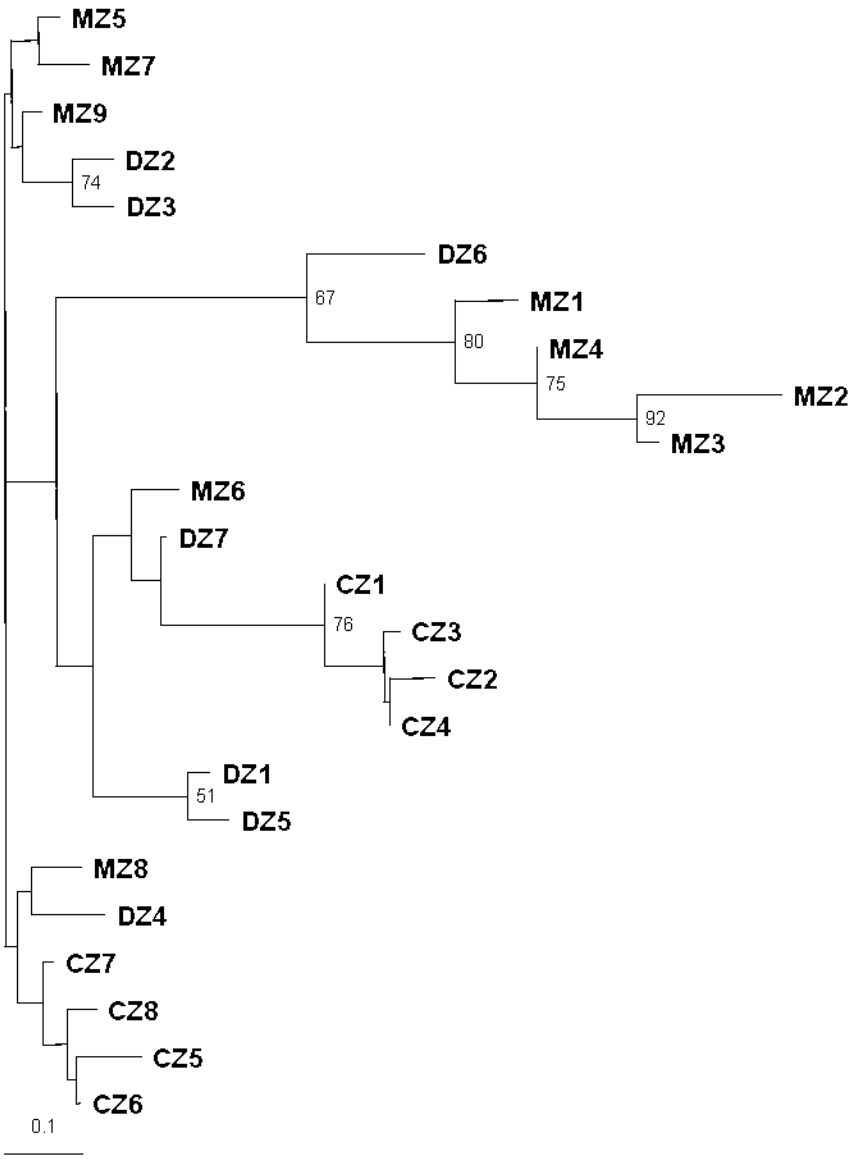


Figure 2.4 Neighbor-joining tree of *Thuja occidentalis* populations based on Nei's standard genetic distance, Ds (Nei 1987). The numbers indicate the bootstrap values; ouly values $\geqslant$ 50% are presented.

2.5.4 Genetic Variation Partitioning

AMOVA revealed a significant level of differentiation among the EWC populations, with 7.7% of the variation found among populations and 92.3% within populations (Table 2.4).

Table 2.4 Analysis of molecular variance for 24 populations, for populations pooled by zones (continuous, discontinuous, and marginal), for populations pooled in groups identified by STRUCTURE, and for populations at the level of each zone.

Source of variation	Sum of squares	Variance components	Percentage variation	*P* value
All populations				
Among populations	176.034	0.12952	7.69904	0.00000
Within populations	904.304	0.12401	92.30096	0.00000
Pooled by zones				
Among zones	33.492	0.02610	1.51048	0.00059
Among populations within zones	142.543	0.11134	6.61151	0.00000
Within populations	904.304	0.12401	91.87801	0.00000
Groups by clusters (3) identified by STRUCTURE				
Among groups	84.830	0.12628	7.12498	0.00000
Among populations within groups	91.204	0.05736	3.39239	0.00000
Within populations	904.204	0.12401	89.48263	0.00000
Populations at the level of each zone Marginal				
Among populations	48.625	0.10511	6.00038	0.00000
Among individuals	334.361	0.33193	18.94962	0.00000
Discontinuous				
Among populations	46.001	0.11803	6.98706	0.00000
Among individuals	295.332	0.17632	10.43782	0.00000
Continuous				
Among populations	47.917	0.10980	6.60170	0.00000
Among individuals	274.611	-0.10815	-6.50205	1.00000

When the populations are pooled based on their distribution zones (marginal, discontinuous, continuous), 1.5% of the variability occurred among zones and 6.6% occurred among populations within a zone. When the populations are pooled according to the results obtained with STRUCTURE (MZ1, MZ2, MZ3, MZ4, and DZ6 (orange); MZ5, MZ6, MZ7, MZ8, MZ9, DZ1, DZ2, DZ3, DZ4, DZ5, CZ5, CZ6, CZ7, and CZ8 (yellow); CZ1, CZ2, CZ3, CZ4, and DZ7 (blue)), the variation among groups was estimated to be 7.1% and 3.4% among populations within groups. The level of variation among populations within zones was

generally similar (6.0, 7.0, and 6.6%; Table 2.4). The variance explained by individuals within populations from the continuous zone is negative (-6.5%) and can be interpreted as being zero, which indicates an absence of genetic structure.

The correlation between genetic and geographic distances was positive and significant when all 24 populations were included in the analysis (Mantel test: $r = 0.645$, $P = 0.001$). However, this correlation became non-significant when the populations from Chibougamau (which are geographically distant from all other sampled populations, > 300 km from the populations of James Bay) were excluded from the analysis ($r = - 0.0002$, $P = 0.571$) (Fig. 2.3). Moreover, no significant correlation between geographic and genetic distances was detected when the IBD (isolation by distance) was tested at the level of each zone (data not shown).

Table 2.5 Results of MSVAR analysis of population expansion or decline

Parameter	N_0	SE	Lower Bound	Upper Bound	N_1	SE	Lower Bound	Upper Bound	T	SE	Lower Bound	Upper Bound	r-ratio
Marginal Zone(MZ)													
MZ1	4.35	0.0066	3.19	5.42	4.15	0.0115	1.80	6.75	4.43	0.0148	1.29	7.91	1.05
MZ2	4.37	0.0065	3.23	5.54	4.21	0.0128	1.43	7.42	4.39	0.0150	1.45	8.07	1.04
MZ3	4.43	0.0082	2.85	6.27	4.75	0.0117	2.68	7.61	4.67	0.0180	1.08	8.67	0.93
MZ4	4.33	0.0060	3.30	5.39	4.50	0.0117	2.33	7.27	4.29	0.0154	1.32	8.24	0.96
MZ5	4.43	0.0077	3.11	6.19	4.66	0.0127	2.31	7.60	4.42	0.0175	1.21	8.61	0.95
MZ6	4.47	0.0054	3.46	5.45	3.82	0.0108	1.55	6.42	4.48	0.0120	1.74	7.25	1.17
MZ7	4.45	0.0057	3.46	5.42	3.87	0.0107	1.17	5.79	4.31	0.0133	1.71	7.72	1.15
MZ8	4.36	0.0083	2.73	6.33	5.02	0.0114	3.27	7.81	4.70	0.0191	1.12	8.91	0.87
MZ9	4.47	0.0047	3.61	5.27	3.35	0.0099	1.09	4.65	3.99	0.0115	1.57	6.79	1.33
Discontinous Zone (DZ)													
DZ1	4.66	0.0030	3.98	5.30	2.41	0.0062	1.01	3.69	3.78	0.0060	2.42	5.04	1.94
DZ2	4.55	0.0043	3.79	5.33	3.25	0.0086	1.62	4.37	4.09	0.0091	2.17	5.94	1.40
DZ3	4.47	0.0053	3.47	5.45	3.56	0.0095	1.41	4.67	4.44	0.0100	2.38	6.52	1.26
DZ4	4.56	0.0033	3.85	5.25	3.29	0.0047	2.24	4.33	4.40	0.0063	2.99	5.72	1.39
DZ5	4.51	0.0032	3.81	5.20	2.62	0.0061	1.31	3.98	3.70	0.0067	2.09	5.03	1.72
DZ6	4.47	0.0029	3.82	5.07	3.05	0.0077	1.55	4.30	3.74	0.0067	2.34	5.09	1.46
DZ7	4.67	0.0030	4.01	5.35	2.57	0.0047	1.58	3.59	4.11	0.0047	3.10	5.15	1.82
Continuous Zone (CZ)													
CZ1	4.48	0.0037	3.76	5.13	2.72	0.0092	0.85	4.23	3.75	0.0092	1.53	5.22	1.65
CZ2	4.18	0.0057	3.18	5.10	3.46	0.0103	0.95	4.95	4.11	0.0117	1.58	6.81	1.21
CZ3	4.28	0.0053	3.33	5.23	3.17	0.0091	0.99	4.43	4.24	0.0114	1.83	6.81	1.35
CZ4	4.30	0.0071	3.00	5.65	3.56	0.0111	0.97	5.56	4.64	0.0141	1.63	7.86	1.21
CZ5	4.28	0.0073	2.93	5.67	4.19	0.0117	1.85	6.97	4.53	0.0155	1.16	7.89	1.02
CZ6	4.54	0.0073	3.20	6.05	4.35	0.0098	2.58	6.63	4.94	0.0159	1.54	8.46	1.04
CZ7	4.72	0.0122	2.51	7.62	4.50	0.0064	3.31	5.64	4.30	0.0166	1.04	8.32	1.05
CZ8	4.74	0.0107	2.98	7.37	4.41	0.0078	3.06	6.06	4.94	0.0151	1.77	8.33	1.07

N_0, current effective population size; N_1, ancestral effective population size; T, time in generations since population size changes; Lower and upper bound are presented as 90% Highest Probability Density intervals.

2.5.5 Population Genetic Bottleneck

A genetic bottleneck was detected by heterozygosity excess test in only one marginal population (MZ4) under both TPM and SMM models. However, population MZ4 had a normal L-shaped allelic distribution, indicating that the bottleneck was not recent or that the population is not completely isolated. Bayesian MSVAR detected a population decline in marginal population MZ8 (r = 0.87). Several populations (MZ3, MZ4, and MZ5) had r-ratios slightly below 1, which indicated a slight decline in population size (Table 2.5). The remaining populations showed a signal of recent expansion (r > 1) (Table 2.5).

2.6 Discussion

Microsatellite markers revealed a significant effect of habitat fragmentation on the genetic structure in EWC populations. Populations from the marginal and discontinuous distribution ranges showed an excess of homozygotes, whereas populations from the continuous range were in HW equilibrium. Therefore, the impact of population fragmentation on the EWC genetic structure is the existence of a positive inbreeding coefficient, which was, on average, nearly two times in marginal zone of that of populations from the discontinuous zone (Table 2.3). This pattern could also partially reflect historical events (e.g., effects of postglacial migration and colonization) as the farthest north population experienced population decline (Hoban *et al.* 2010; Dudaniec *et al.* 2012). This result indicated the presence of a higher occurrence of selfing within fragmented EWC populations that was coupled with a higher degree of gene exchange among near-neighbour relatives, leading to significant inbreeding. In their review, Aguilar *et al.* (2008) reported a trend of increased inbreeding due to habitat fragmentation; however, they reported a non-significant overall effect on F_{is}, possibly because the fragmentation was too recent. In many published studies, the sampled adults were established before fragmentation occurred (Young *et al.* 1996; Lowe *et al.* 2005; Kettle *et al.* 2007). Indeed, the effect of population fragmentation on inbreeding coefficients can be detectable only after the first generation of progeny has been established.

The presence of a high level of self-fertilisation in EWC has been reported in previous studies (Perry & Knowles 1990; Lamy *et al.* 1999). Lamy *et al.* (1999) showed that mating patterns

are biased towards higher selfing in recently fragmented, small EWC populations. This life-history characteristic contrasts with most coniferous species, which are generally much more affected by inbreeding (Mitton 1983; Plessas & Strauss 1986; Gauthier *et al.* 1992; Beaulieu & Simon 1995; Ledig *et al.* 2000; Gamache *et al.* 2003; Gapare *et al.* 2005). A high level of inbreeding, maintained over several generations, is expected to lead to progressive genetic erosion, higher between-population differentiation and an overall decrease in genetic diversity. This pattern was not observed in the present study. Genetic variation among populations was similar in the marginal, discontinuous and continuous populations (6.0%, 7.0% and 6.6%, respectively), as were the levels of genetic diversity (Hs, Table 3), except that only populations from continuous zones had private alleles. This is probably because the fragmentation has not progressed long enough to have detectable effects on progressive genetic erosion. Long-lived trees may be buffered against genetic erosion for centuries (Templeton & Levin 1979; Cabin 1996; Piotti 2009).

The global level of differentiation among EWC populations was relatively high and similar to that reported by Lamy *et al.* (1999) (7.7% vs. 7.3%) in populations sampled over a much smaller geographical area (180 km^2). It was also higher than those values that were reported in EWC populations by Matthes-Sears *et al.* (1991) (1.9%) and Perry *et al.* (1990) (1.6%). Most alleles were distributed in populations throughout the three zones. Populations from the continuous distribution zone harboured the highest proportion of rare alleles (frequency < 1%), with a decreasing trend towards the northern range margins. Yet, no significant differences were observed in allelic richness among populations from the three bioclimatic zones, indicating that populations residing in the discontinuous or marginal distribution ranges have not experienced a great decrease in population size or, if so, have overcome previous bottlenecks (Nei *et al.* 1975; Leberg 2002). The evidence of population decline was detected in marginal populations (MZ3, MZ4, MZ5, and MZ8). However, the detection power of our bottleneck analysis was weak due to the limited number of polymorphic microsatellite loci available for the EWC. Our results were still comparable to other studies that detected significant bottlenecks based on four polymorphic loci (Aizawa *et al.* 2009; Heuertz *et al.* 2010). Genetic bottleneck effects could also be obscured by immigration events.

The majority of studies that have examined geographic variation in genetic diversity have used a 'categorical approach' in which only groups of peripheral and central populations were sampled (Eckert *et al.* 2008). Yet, the 'categorical approach' has also been blamed for confounding geographical position with region compared to 'continuous sampling approach'. Our study relaxed this confounding by sampling along a latitudinal transect that encompasses central, intermediate, and peripheral populations. The geographic distribution of EWC along the latitudinal gradient was estimated from the analysis of a large inventory database (a total of 5476 sample plots) and found to decrease from 55% to 9% to 3% from the continuous to the discontinuous to the marginal zones, respectively (Paul 2011). This pattern conforms to the 'abundant centre model', which predicts an increase in the spatial isolation of populations from the range centre towards the range limits (Sagarin & Gaines 2002; Eckert *et al.* 2008). This increase in population isolation was apparently not correlated with a detectable effect on genetic diversity. One plausible explanation involves the life-history characteristics of EWC. Selfing species naturally retain most of their genetic diversity within populations, and their level of population genetic diversity is less affected by restricted gene flow. Moreover, the ability of EWC to reproduce vegetatively, via layering, may buffer the genetic effects of fragmentation by delaying the time between generations (Honnay & Bossuyt 2005). A parallel study conducted at the same sites showed higher levels of layering in populations in the north (marginal and discontinuous zones) than in the south (continuous zone), with equivalent seed production along the gradient (Paul 2011). Finally, the effect of inbreeding on genetic erosion may also be buffered by selection against homozygotes in young EWC individuals, which will eliminate a higher proportion of these individuals before they become adults.

Population structure

Both Bayesian and NJT analyses detected a certain level of genetic structure among the 24 EWC populations. Interestingly, the four marginal populations (MZ1, MZ2, MZ3, and MZ4) from Chibougamau and one population (DZ6) from Abitibi were assigned to one cluster, even though more than 400 km separated DZ6 from the Chibougamau marginal populations. One explanation may be that these populations followed the same post-glacial migration route. Apparently, the four populations from Témiscamingue (CZ1, CZ2, CZ3, and CZ4)

belonged to the same cluster, indicating that gene flow (via seed or pollen dispersal) was high among them. Some sub-branches of the NJT were significant (bootstrapped values $\geqslant$ 50), such as the sub-branch clustering of DZ1 and DZ5 or that of DZ2 and DZ3. These populations that clustered together are genetically closer and may have followed similar post-glacial migration routes. Fourteen populations (marginal: MZ5, MZ6, MZ7, MZ8, and MZ9; discontinuous: DZ1, DZ2, DZ3, DZ4, and DZ5; and continuous: CZ5, CZ6, CZ7, and CZ8) were assigned into a single (yellow) cluster.

2.7 Conservation Implications

Our results converged to demonstrate that spatial isolation of marginal EWC populations is not associated with low genetic diversity. Therefore, increased inbreeding does not lead to a loss of genetic variation in northern EWC populations and, therefore, they have the potential to respond and adapt to environmental changes. The actual distribution and expansion of white cedar at the northern edge of its range has been limited by climate in association with fires (Paul 2011). This limitation illustrates the complexity of the species' population dynamics and the difficulty of predicting future EWC distributions in a changing environment. If climate favours improved regeneration of this species and its northward migration, peripheral populations could play a major role as seed sources and in the further movement of the geographic range in response to climate changes. In contrast, if global warming triggers an increase in fire frequency (Bergeron *et al.* 2010), the EWC distribution could be negatively affected and reduced to lower latitudes. In such a context of uncertainty, the precautionary principle should apply, and marginal populations should be protected to allow continuity of natural evolutionary processes.

2.8 Acknowledgements

This work forms part of the PhD thesis of H. Xu. We would like to acknowledge Marie-Hélène Longpré and Danielle Charron for logistical support and Dr. Marc Mazerolle for statistical advice. Special thanks go to industrial supervisors from the company Tembec: Louis Dumas, Sonia Légaré, and Genevieve Labrecque. We thank two anonymous referees for their comments, which helped improve this manuscript. We also thank Dr. W. F. J.

Parsons for correcting the references to match the journal. This research was supported by an Industrial Innovation Doctoral Scholarship from NSERC (Natural Sciences and Engineering Research Council of Canada), a BMP innovation doctoral scholarship from FRQ-NT the (Fonds de recherché du Québec- Natures et technologies) and Tembec, and a scholarship from FERLD (Forêt d'enseignement et de recherche du lac Duparquet) to H. Xu. This research was also funded by a NSERC strategic grant (STPGP 336871) to F. Tremblay.

CHAPTER III

SEXUAL OR ASEXUAL REGENERATION, WHICH IS MORE IMPORTANT FOR A LATE SUCCESSIONAL SPECIES, EASTERN WHITE CEDAR (*THUJA OCCIDENTALIS* L.)?

Article submitted to *Annals of Botany*

3.1 Résumé

•**Contexte et objectif** Comprendre le mode de colonisation et d'invasion d'une espèce de fin de succession suite à une perturbation naturelle (feu).

• **Méthode** Nous avons sélectionné trois peuplements forestiers ayant brûlé pour la dernière fois en 1916, 1823 et 1760, représentant une chronoséquence après feu d'une longueur de 250 ans. Un total de 535 individus (régénération: 281, et arbres adultes: 254) de *Thuja occidentalis* (CBE) ont été génotypés à l'aide de 16 marqueurs microsatellites.

•**Résultats** Nos résultats montrent une influence du stade de succession après feu sur la dynamique de régénération du CBE. Le pourcentage de régénération asexuée augmente légèrement avec le temps écoulé depuis le dernier feu (1916 : 21.8%, 1823 : 27.0%, 1760 : 30.9%), et s'accompagne d'une diminution de la diversité génotypique (1916 : 0.779, 1823 : 0.727, 1760 : 0.688). La plus grande proportion des échanges génétiques avaient lieu à l'intérieur de la population (pollen : 66.1%, graines : 62.5%). Le site le plus jeune (1916) a reçu une grande proportion du flux de gènes (82.4 %) des deux autres sites d'études (1823, 1760). L'analyse de la structure génétique spatiale (SGS) a montré des patrons contrastés entre les adultes et les gaules (hauteur moyenne: 60± cm). L'autocorrélation spatiale était élevée et significative pour les gaules et faible pour les arbres adultes. La propagation végétative des gaules entraine une augmentation de la structure génétique spatiale sur de courtes distances et en accroit l'Intensité (*Sp*).

• **Conclusions** Notre étude a permis d'accroître nos connaissances de la dynamique et le mode de recrutement du CBE dans la succession après feu en forêt boréale. Elle nous renseigne sur la dynamique spatio-temporelle de la propagation végétative et sur la structure génétique spatiale fine d'une espèce de fin de succession.

3.2 Abstract

•**Background and Aims** To understand the mode of colonization and invasion by a late successional tree species into boreal forest stands following primary natural disturbance (forest fire).

•**Methods** We selected three sites that were last burnt in 1916, 1823, 1760, respectively, representing a 250-year-long post-fire successional gradient. A total of 535 individuals (281 saplings, 254 adult trees) of the eastern white cedar (EWC; *Thuja occidentalis* L.), were successfully genotyped using 16 polymorphic microsatellite loci.

•**Key Results** Our results revealed the influence of a post-fire succession on regeneration pattern of EWC. The percentage of asexual regeneration slightly increased with stand development (1916, 21.8%; 1823, 27.0%; 1760, 30.9%), while genotypic diversity decreased (1916, 0.779; 1823, 0.727; 1760, 0.688). Most gene dispersal was realised within site (pollen, 66.1%; seed, 62.5%). The youngest site (1916) received a great portion (82.4%) of genes from older sites (1823, 1760). Fine-scale spatial genetic structure (SGS) analysis showed contrasting patterns of SGS between saplings (mean height: 60± cm) and adult trees. SGS was high and significant in saplings, and weaker in adult trees. Clonal growth increased SGS in saplings over short distances, together with SGS intensity (*Sp*).

•**Conclusions** Our study clarified the dynamics and the mode of natural population regeneration along succession for EWC, in the boreal forest. It has shed insights into spatio-temporal dynamics of clonal growth and fine-scale spatial genetic structure (SGS) in late-successional species.

3.3 Introduction

Most plants can propagate both vegetatively and sexually, and the balance between the two modes of reproduction varies widely among and within species (Prati & Schmid 2000; Rasmussen & Kollmann 2004; Burczyk *et al.* 2006). Many factors affect this balance. Sexual reproduction can be limited by both biotic and abiotic factors, such as climate and soil conditions, stressful local environments, as well as biased sex ratios (Dorken & Eckert 2001; Frenne *et al.* 2012). Extensive asexual growth that results in large clone sizes (i.e. over $100/m^2$) could limit sexual reproduction by making the habitat unsuitable for seedling establishment (Pigott 1992; Rasmussen & Kollmann 2004; Karst *et al.* 2008). Vegetative reproduction could be limited by microsite conditions (Oddou-Muratorio *et al.* 2004).

Trees that regenerate vegetatively tend to clump together. This results in an uneven spatial distribution (Oddou-Muratorio *et al.* 2004), which generates spatial genetic structure (SGS) over short distances (Reusch *et al.* 1999; Ohsako 2010). Moreover, temporal variation in genotypic diversity reflects the relative contributions of sexual and asexual recruitment during stand development. Thus, genotypic diversity is predicted to decrease in the absence of frequent sexual recruitment (Balloux *et al.* 2003). Consequently, highly clonal populations are expected to exhibit a decrease in genotypic diversity with time (Koppitz & Kühl 2000; Hock *et al.* 2008).

Sexual reproduction usually initiates re-colonisation following stand-replacing disturbances, i.e. forest fires. In Canada, wildfire has been recognised as the principal agent of disturbance in the boreal forest and is characterised by high-intensity crown fires that initiate secondary succession processes in the burned areas (Heinselman 1981; Van Wagner 1983; Bergeron 2000). Many studies have reported that late successional species could invade burnt areas immediately following fire (Bergeron & Charron 1994). However, further recruitment of late-successional species could be delayed until suitable substrates become more abundant (Simard *et al.* 1998; Simard *et al.* 2003).

One of these species, eastern white cedar (EWC, *Thuja occidentalis* L.), invades stands slowly and becomes dominant in old-growth forests (Bergeron 2000). Sexual and asexual

propagation interact throughout the succession. Relatively small groups of mature trees may disperse their seeds by wind to colonise a nearby area, but this process is strongly influenced by the distance from seed sources (Asselin *et al.* 2001). Subsequent expansion of EWC, or other late-successional species, might be guaranteed through vegetative propagation. The balance between sexual and vegetative modes of propagation in EWC that contribute to these forest dynamics is unknown and probably depends upon multiple factors, including disturbance regimes.

In this study, we combined molecular genetics and forest ecology to understand the mode of colonisation and invasion by late successional tree species into boreal forest stands following natural disturbances (i.e. forest fire), using EWC as a model species. Specifically, our objectives were: 1. to examine the relative contributions of sexual and asexual regeneration in EWC colonisation along the boreal mixed wood succession and to shed insight into the temporal dynamics of EWC clonal growth; 2. To understand the effects of the balance between the two forms of regeneration on the fine-scale spatial genetic structure (SGS) of EWC; 3. To estimate gene dispersal distances that are mediated by pollen and seeds. These three aims will help to understand the process of EWC invasion into forest stands. We hypothesised that: 1. the ratio of vegetative recruitment through layering versus sexual reproduction increases along the succession, and genotypic diversity decreases concomitantly; 2. Vegetative reproduction increases fine-scale spatial genetic structure (SGS) over short distances in EWC; and 3. Seed trees within the stand contribute to the local gene pool of the regeneration along the succession.

3.4 Material and Methods

3.4.1 The Species

Eastern white cedar (*Thuja occidentalis* L.) is a monoecious and wind-pollinated tree species (Johnston 1990). It is a shade-tolerant, slow-growing conifer (Godman 1958), which can grow to between 15 and 38 metres in height (Chambers 1993). Flowering usually commences in late April or May, with cone formation starting in late June, and maturation in August. These cones may remain attached to the tree (Johnston 1990). The seed of EWC has two flat

lateral wings with two cotyledons, with seed production usually occurring at an early age but peaking after age 75, with abundant seed crops being produced every 3 to 5 years (Johnston 1990). Seed is mainly disseminated by wind, usually beginning in September and ending in November, but it may continue to fall throughout the winter. Seed germination normally begins in late May or June (Johnston 1990). EWC reproduces vegetatively by layering, and more rarely by coppicing and root suckering. It tends to occur in pure and mixed stands, over a wide range of organic and mineral soils in the boreal forest (Johnston 1990). In the southern Canadian boreal forest, EWC is rarely encountered in stands at early post-fire successional stages, while it is found more frequently in intermediate stages in association with balsam fir (*Abies balsamea* [L.] Mill.), white (*Picea glauca* [Moench] Voss) and black spruce (*Picea mariana* [Mill.] BSP), trembling or quaking aspen (*Populus tremuloides* Michx.), and paper or white birch (*Betula papyrifera* Marsh). EWC generally dominates old-aged stands in association with species such as balsam fir and white birch (Bergeron 2000; Chen & Popadiouk 2002; Bergeron *et al.* in press).

3.4.2 Study Area

The study was located in the Lake Duparquet Research and Teaching Forest (FERLD, Figure 3.1) (79°10'W, 48°30'N), which is in northwestern Quebec, Canada. This forest is a part of a larger region that had been covered by lacustrine deposits from the maximum post-Wisconsinian extension of the postglacial lakes Barlow and Ojibway (Vincent & Hardy 1977). Vegetation types vary, depending on successional stage and soil deposits (Bergeron & Dubuc 1989). This forest has been barely touched by human disturbance, and extensive research has been conducted regarding its disturbance history and successional dynamics (Bergeron & Dubuc 1989; Bergeron 1991; Dansereau & Bergeron 1993; Kneeshaw & Bergeron 1998). Young successional stages (< 100-years-old) are dominated by trembling aspen, while intermediate stages (100- to 200-years-old) are dominated by balsam fir, trembling aspen, and white spruce. Old stages (> 200-years-old) are dominated by balsam fir and eastern white cedar (Bergeron 2000).

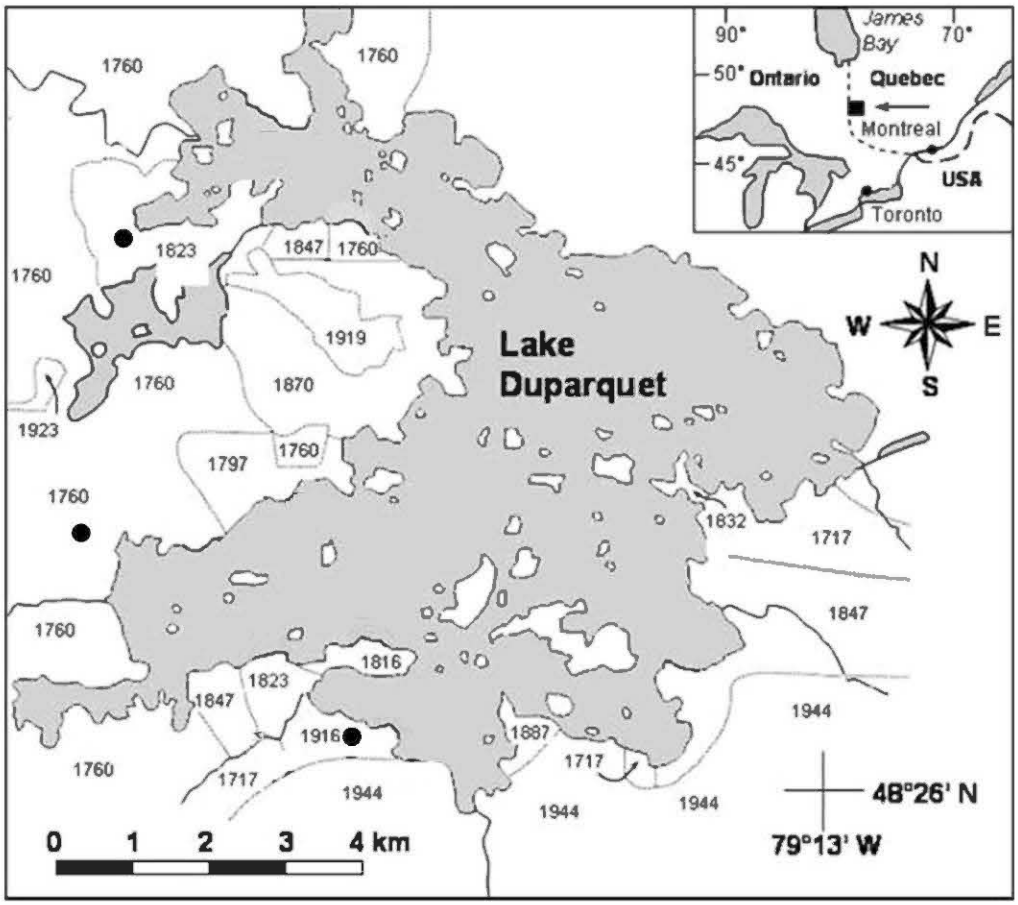


Figure 3.1 Study area and fire years and in FERLD forest in Canada and the three sampling sites were dotted

3.4.3 Sampling

Three fire-initiated sites (1916 AD, 1823 AD, and 1760 AD) were selected that covered a 250-year-long post-fire succession. One 1-hectare plot (100 x 100 m) was established in each site, in which every tree (DBH > 5 cm) was tagged and mapped in two dimensions (Table 3.1, Figure 3.1). Mean distances between the sampling plots were 7.2 km (1916, 1823), 4.39 km (1916, 1760), and 4.33 km (1760, 1823), respectively. A total of 14 EWC trees (DBH > 5 cm) were counted in plot 1916, 1020 in 1823, and 489 in 1760 (Table 3.1).

Table 3.1 Fire years in FERLD forest in Quebec (Canada) and the composition and characteristics of the three study sites

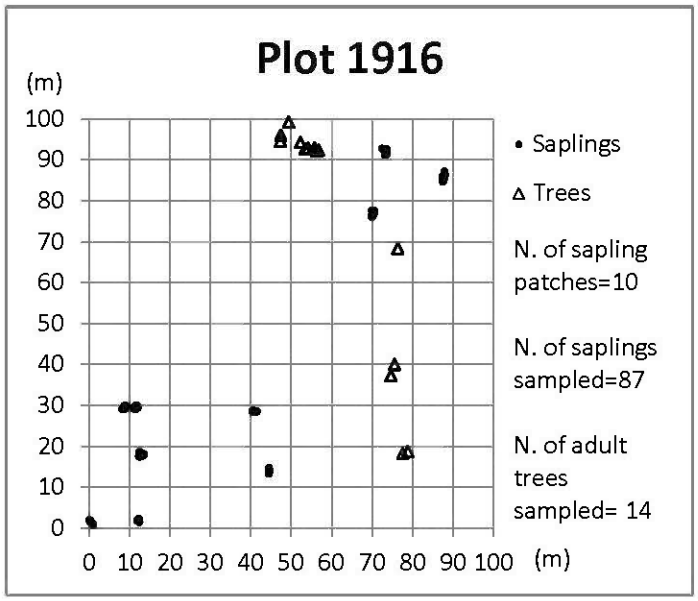
Fire years	1 hectare sampling plot						
	Tree Species (DBH>5cm)	Mean DBH (cm)		Density (stems/ha)		Basal area (m^2/ha)	
		live	dead	live	dead	live	dead
1916	Balsam fir	10.3	11.7	597	96	5.69	1.20
	Alder	13.3	6.3	2	1	0.04	0.01
	White birch	13.7	10.7	352	202	5.87	2.06
	White spruce	12.8	15.8	218	8	4.04	0.26
	Trembling aspen	36.7	24.7	195	41	22.17	2.38
	Willows	7.5	-	2	-	0.01	-
	Eastern white cedar	16.8	-	14	-	0.39	-
	Total			1380	348	38.22	5.91
1823	Balsam fir	9.1	8.1	450	64	3.31	0.37
	Alder	6.1	6.7	17	3	0.05	0.01
	White birch	18.5	21.7	44	17	1.40	0.72
	White spruce	23.4	10.4	64	3	3.35	0.04
	Black spruce	16.4	5.9	21	2	0.56	0.01
	Trembling aspen	22.1	16.8	243	30	12.17	1.04
	Willows	5.4	-	1	-	0.01	-
	Eastern white cedar	13.7	12.1	1020	24	19.37	0.33
	Total			1860	143	40.22	2.52
1760	Balsam fir	10.7	17.5	23	3	0.24	0.08
	White birch	29.2	25.6	52	19	4.17	1.09
	White spruce	31.8	-	7	-	0.64	-
	Trembling aspen	37.5	25.2	35	2	4.26	0.12
	Eastern white cedar	27.0	19.3	489	12	33.57	0.47
	Total			606	36	42.88	1.76

Note: Balsam fir: *Abies balsamea* [L.] Mill.; Alder: *Alnus rugosa*; White birch: *Betula papyrifera* Marsh.; White spruce: *Picea glauca* [Moench] Voss; Black spruce: *Picea mariana* [Mill.] BSP; Trembling aspen: *Populus tremuloides* Michx.; Willows: *Salix* spp. Eastern white cedar: *Thuja occidentalis* L. Coordinates for the 1 hectare sampling plots: 1916, 48°26'27.1"N, 79°17'44.1"W; 1823, 48° 30' 12.4"N, 79° 19' 15.6"W; 1760, 48°27'59.6"N, 79°20'31.4"W. Only the dead trees that were still standing were included in calculations. FERLD, Forêt d'enseignement et de recherche du lac Duparquet.

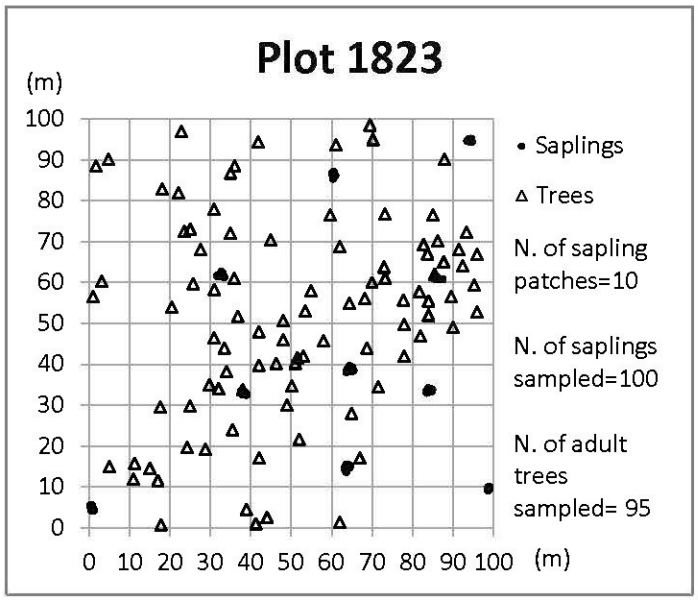
Each 1-ha plot was divided into one hundred 10 x 10 m subplots (Figure 3.2 a, b, c). We first surveyed each hectare to spot regeneration patches. Regeneration patches (< 100cm in height) were very scarce and small in the 1916 plot, while they were common and large in the other two plots, especially in 1823. Ten subplots (10 x 10 m) were then selected as representative

of the regeneration of EWC saplings. In the 1916 plot, patches were selected on the basis of their size rather than their location. One regeneration patch was selected in each 100 m^2 subplot, where up to ten EWC saplings were sampled (Figure 3.2 a, b, c). Tree diameter at forest floor level (cm) and height (m) of each sampled sapling was measured and foliage was collected for genetic analysis (Table S3.1).

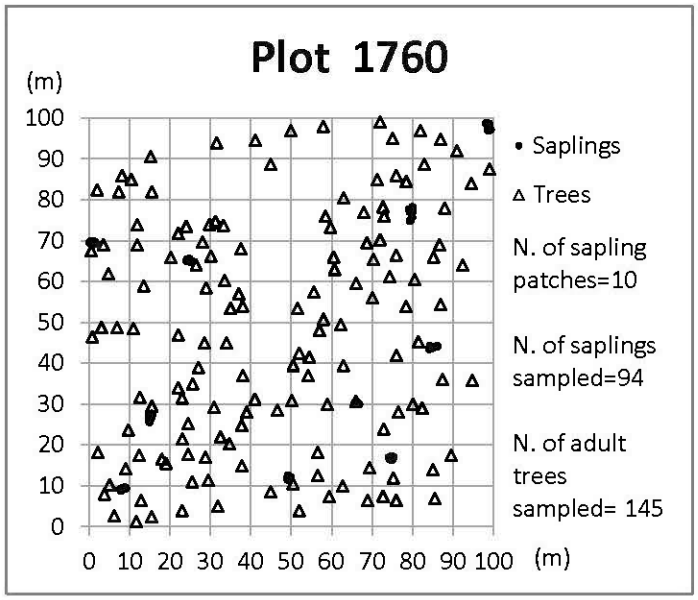
For mature trees, only EWC trees with DBH $\geqslant$ 35cm in 1760 (145 trees), DBH $\geqslant$ 25cm in 1823 (95 trees), and all trees in 1916 (14 trees, mean DBH = 16.8 cm) were sampled for genetic analysis (Figure 3.2). Foliage or cambial tissue was collected for the genetic analyses. All samples were stored frozen (-20 °C).



(a) Plot1916



(b) Plot 1823



(c) Plot 1760

Figure 3.2 Spatial distributions of adult trees of *Thuja occidentalis* and saplings that were sampled in three plots (a, plot1916; b, plot 1823; c, plot 1760)

3.4.4 DNA Extraction and Microsatellite Genotyping

Frozen samples were ground in liquid nitrogen and genomic DNA was extracted using the DNeasy Plant Mini kit (QIAGEN, Hilden, Germany). DNA amplification was conducted at 16 polymorphic microsatellite loci following the protocol that had been developed by (Xu *et al.* 2013). Capillary electrophoresis of amplified DNA was performed on an ABI Prism 310 Genetic Analyzer (Applied Biosystems, Carlsbad, California, USA), and then analysed and genotyped using Genemapper v3.7 Software (Applied Biosystems).

3.4.5 Genetic Data Analysis

The probability of finding individuals with identical multi-locus genotypes (MLGs) that derive from sexual reproductive events (P_{sex}) and its analogue (P_{sex} *(f)*) when considering possible departures from Hardy-Weinberg equilibrium (Tibayrenc *et al.* 1990; Parks & Werth 1993; Arnaud-Haond & Belkhir 2007) were estimated. Genotypic richness was measured by using the ratio (R, R= (G-1)/(N-1), where G is the number of observed genets, and N is the number of individuals that were genotyped (Dorken & Eckert 2001). This ratio tends to zero when the number of genotypes is very low, and reaches a maximum of 1 when each individual has a unique multi-locus genotype. The number of single-ramet genets and its percentage (*SR*), and the number of multi-ramet genets and its percentage (*MR*) were calculated. All of the above analyses were conducted in GenClone v. 2.0 (Arnaud-Haond & Belkhir 2007).

In the ‘‘genet’’ data set (N = 458), samples with identical MLGs that originated from same clone were only counted once. The ‘‘ramet’’ data set (N = 535) included genotypic data from all of the samples that were collected. The "genet" data set was used in the following analysis. The presence of null alleles and large allele dropouts at each locus were tested using Micro-Checker software (Van Oosterhout *et al.* 2004). The frequencies of putative null alleles (r) and global genetic differentiation (F_{st}) (Weir 1996) were estimated using the FreeNA program, with and without the inclusion of null alleles at each locus (Chapuis & Estoup 2007). Since the 16 loci were in linkage equilibrium (Xu *et al.* 2013), we tested Hardy-Weinberg equilibrium, and calculated the inbreeding coefficient (F_{is}) and allelic

richness (*AR*) using rarefaction in FSTAT v. 2.9.3 (Goudet 1995; Goudet 2001). Bonferroni corrections (Rice 1989) were applied to test Hardy-Weinberg equilibrium.

3.4.6 Fine-scale Spatial Genetic Structure (SGS)

We characterised fine-scale spatial genetic structure of sampled individuals within each plot for regeneration and adult trees using SPAGeDi v. 1.3 (Hardy & Vekemans 2002). The "genet" data set was used for these computations. The statistic of Nason's kinship coefficient (F_{ij}) (Loiselle *et al.* 1995) was computed, as it does not assume Hardy-Weinberg equilibrium as a prerequisite. The maximum distance between paired samples was 133.4 m, 130.8 m, and 122.7 m for plots 1760, 1823, and 1916, respectively. Ten distance intervals were allocated by the program to maintain a similar number of paired samples within each distance interval (Table S3.2). For each distance interval, *% partic* >50% (*% partic*, proportion of all individuals represented in each distance interval), and *CV partic* <= 1 (*CV partic*, the coefficient of variation for the number of times each individual is represented) were maintained following the recommendation of the authors (Hardy & Vekemans 2002). The adult trees in plot 1916 were excluded from the analysis due to their low number (N = 14). Random permutations (10 000) were conducted to test the significance of SGS. Jackknifing over loci was used for the calculations. SGS intensity in each plot was quantified by the statistic Sp following the formula $Sp = -b_{log}/(1- F_1)$ (Vekemans & Hardy 2004), where b_{log} is the regression slope of the kinship coefficient on the logarithm of geographical distance, while F_1 is kinship coefficient for adjacent individuals in the first distance interval. Ten-thousand random permutations were applied to test the significance of b_{log}. To test the effects of vegetative reproduction on SGS, we also ran a second SGS analysis which included all clonal saplings.

3.4.7 Parentage Analysis

CERVUS v. 3.0 was used to assign parentage to saplings ("genet" data set) (Kalinowski *et al.* 2007). CERVUS excludes unlikely parents when a mismatch occurs during the comparison of genotypes of candidate parents against those of the offspring, and statistically distinguishes the non-excluded candidate parents using likelihood ratios that are calculated from simulation stages (Marshall *et al.* 1998; Slate *et al.* 2000; Kalinowski *et al.* 2007). CERVUS tolerates

moderate deviations from Hardy-Weinberg equilibrium and low levels of null alleles, as well as weak linkages among loci (Kalinowski *et al.* 2007). We assigned parentage within plots, and between candidate parents from one plot and saplings from the other to test possible pollen and seed dispersals over longer distances. The program settings followed the authors' recommendations, except for genotyping error (0.05) which took into account the presence of null alleles and the proportion of parents sampled (5%) assuming most parents were not sampled. In our study, we report parentage assignments under strict confidence (95%) (Hoban *et al.* 2012). We calculated the distance of realised pollen dispersal by averaging the distances between all parent pairs for each offspring, and the distance of realised seed dispersal by averaging the distances between each offspring and its closer parent, which we assumed was the maternal parent.

3.5 Results

3.5.1 Clonal Structure, Genetic and Genotypic Diversity

We detected a very low level of null alleles and no evidence for large allele dropout using MICRO-CHECKER. Similar levels of global genetic differentiation (F_{st}) were obtained when either including or excluding the null alleles (Table S3.3). The probabilities (P_{sex}) of finding individuals with identical MLGs that derive from sexual reproductive events were very low ($<10^{-6}$), as was the case for the analogue $P_{sex}(f)$ ($<10^{-5}$) when considering possible departures from Hardy-Weinberg equilibrium. A total of 458 genets out of 535 ramets were detected (Table 3.2). Among saplings, 68, 73, 65 genets were identified for plots 1916, 1823, and 1760, respectively. Allelic richness (AR) decreased from 5.4 (1760) to 4.7 (1823), and to 4.6 (1916) for saplings (Table 3.2). A significant positive inbreeding coefficient (F_{is}) was found for saplings in 1760 (0.477) and 1916 (0.223), indicating heterozygote deficiency (Table 3.2). A low negative inbreeding coefficient was found for adult trees in 1760 (-0.051) and 1823 (-0.088), indicating slight heterozygote excess. When all samples were pooled, it revealed a slight heterozygote deficiency (F_{is} = 0.100, data not shown). Genotypic richness (R) was high, with a slight decrease from 0.779 (1916) to 0.727 (1823), and to 0.688 (1760). The percentage of saplings that regenerated vegetatively varied from 21.8% (1916) to 27.0% (1823), and to 30.9% (1760). Most clones were small, with clone sizes (CS) ranging from 2

to 5 (Table 3.2). The largest clone (*CS* = 5) was found in the 1760 plot. There were 15 clones among saplings in 1916, 19 in 1823, and 18 in 1760. Clonal saplings were within the same sapling patches across plots. Only one small clone (*CS* = 2) was found among adult trees in both 1916 and 1823, while there were no clones among adult trees in 1760. The number of single-ramet genets (*SR*) was 53, 54, and 47 among saplings in 1916, 1823, and 1760, respectively. The number of multi-ramet genets (*MR*) among saplings ranged from 15 (22.1%) in 1916 to 19 in 1823 (26.0%), and 18 in 1760 (27.7%).

Table 3.2 Clonal structure, genetic and genotypic diversity of *Thuja occidentalis*

Sample type		*N*	*G*	*AR*	*Fis*	*R*	*V (%)*	*CS*	*SR (%)*	*MR (%)*
Regeneration (saplings)	1916	87	68	4.6	0.223*	0.779	21.8%	2(11);3(4)	53(77.9)	15(22.1)
	1823	100	73	4.7	-0.093*	0.727	27.0%	2(13);3(4); 4(2)	54(74.0)	19(26.0)
	1760	94	65	5.4	0.477*	0.688	30.9%	2(11);3(4); 4(2);5(1)	47(72.3)	18(27.7)
Adult trees	1916	14	13	3.8	0.186*	0.923	-	2(1)	12(92.3)	1(7.7)
	1823	95	94	5.2	-0.088*	0.989	-	2(1)	93(98.9)	1(1.1)
	1760	145	145	5.1	-0.051*	1.000	-	-	145(100.0)	-

N, number of individuals genotyped or ramets; *G*, number of individuals with unique multi-locus genotypes (MLGs) or genets; *AR*, allelic richness; F_{is}, inbreeding coefficient; *, $P < 0.05$; *R*, $R = (G-1)/(N-1)$, proportion of unique MLGs; *V*(%), $V = (N-G)/N$, percentage of saplings regenerated vegetatively; *CS*, Clone size and its number in parentheses; *SR*(%), number of single-ramet genets and percentage in parentheses; *MR*(%), number of multi-ramet genets and percentage in parentheses.

3.5.2 Fine-scale Spatial Genetic Structure (SGS)

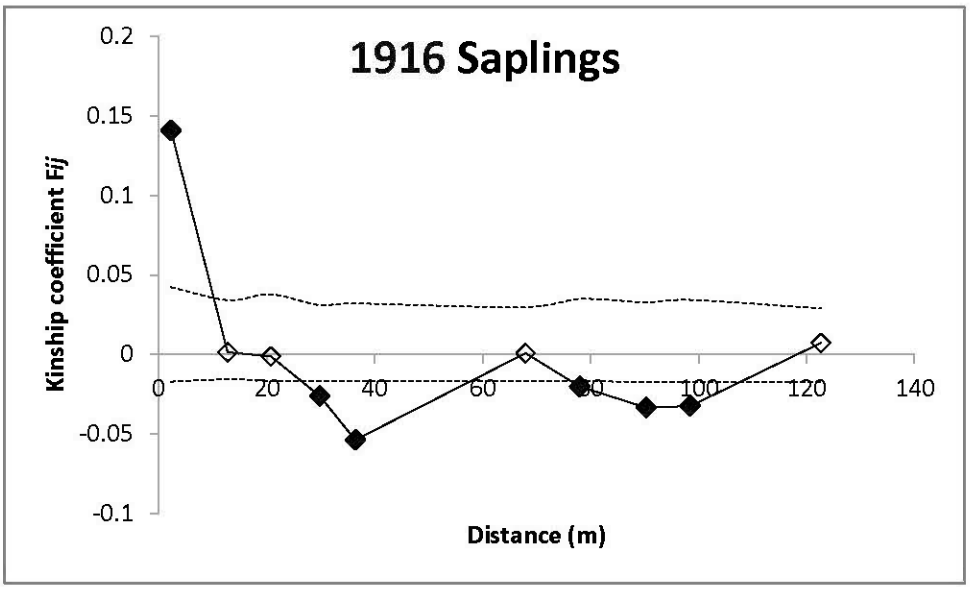
A significant negative regression slope was detected for the relationship between the kinship coefficient and distance, indicating significant SGS (Table 3.3). SGS was weaker for adult trees (1823, b_{log} = -0.0025; 1760, b_{log} = -0.0045) than for saplings. Kinship coefficient was significant for the first distance interval (F_1) for saplings in all three plots, as well as for adult trees in 1760 (Table 3.3, Figure 3.3 a, b, c, d, e).

Table 3.3 Spatial Genetic Structure (SGS) statistics of *Thuja occidentalis*

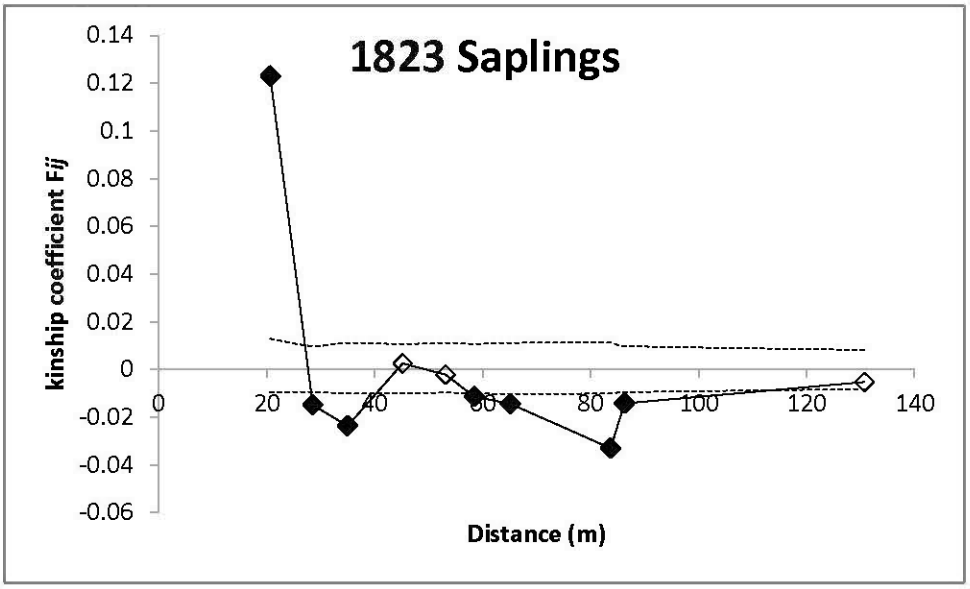
Sample type		F_1	F_L	b_{log}	Sp
Saplings (clones included)	1916	0.1674***	-0.0012	-0.0362391***	0.04352522
	1823	0.1429***	0.0006	-0.0357056***	0.04165862
	1760	0.151***	-0.0025	-0.0400867***	0.04721637
Saplings (clones excluded)	1916	0.1409***	-0.0017	-0.0283321***	0.03297882
	1823	0.1230***	0.0006	-0.0305266***	0.03480798
	1760	0.0895***	-0.0037	-0.0259973***	0.02855277
Adult trees	1823	0.0043ns	0.0005	-0.0024706ns	0.00248122
	1760	0.0076**	0.0002	-0.0044975***	0.00453195

F_1, kinship coefficient for adjacent individuals in the first distance interval; F_L, average kinship coefficient (Loiselle *et al.* 1995) between individuals over all distance intervals; Sp, SGS intensity, $Sp=(-b_{log}/(1-F_1))$,where b_{log} is the regression slope of kinship coefficient on the logarithm of geographical distance. ns, $P > 0.05$, not significant; **, $0.01 \geq P > 0.001$; ***, $P \leq 0.001$.

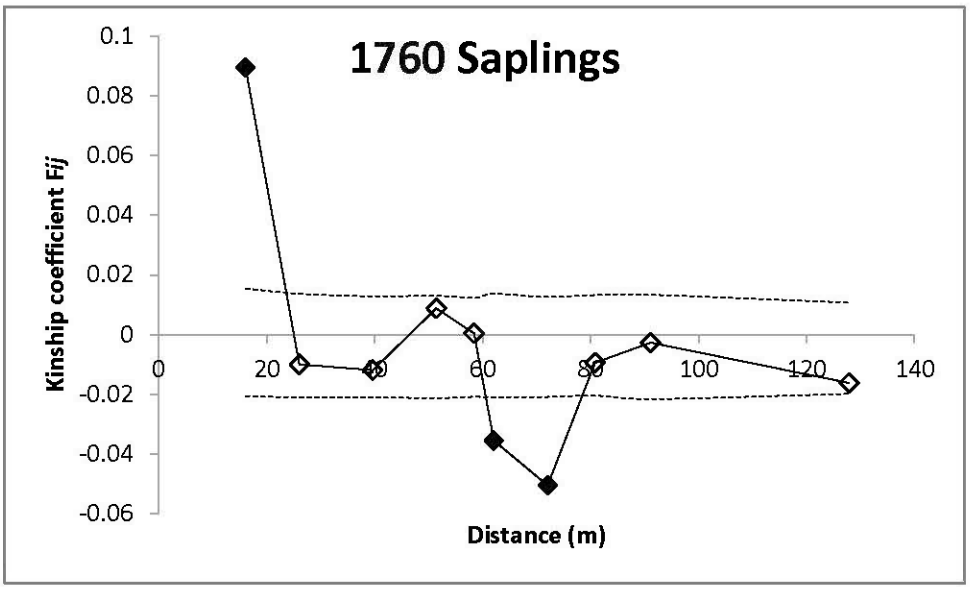
The kinship coefficient was significant at 49.5 m (adult trees) in 1760 (Figure 3.3 e). The negative kinship coefficient found between some individuals meant that those individuals are less related compared to random occurrence, as would occur normally (Hardy & Vekemans 2002). Overall SGS (*Sp* statistic) was similar for saplings in 1916 and 1823 (0.0330, 0.0348), and slightly lower in 1760 (0.0286). *Sp* was much lower for adult trees in 1823 and 1760 (0.0025, 0.0045). When clonal saplings were included into the analysis, SGS increased at the first distance interval in all three plots, as did overall SGS (*Sp*) (Table 3.3).



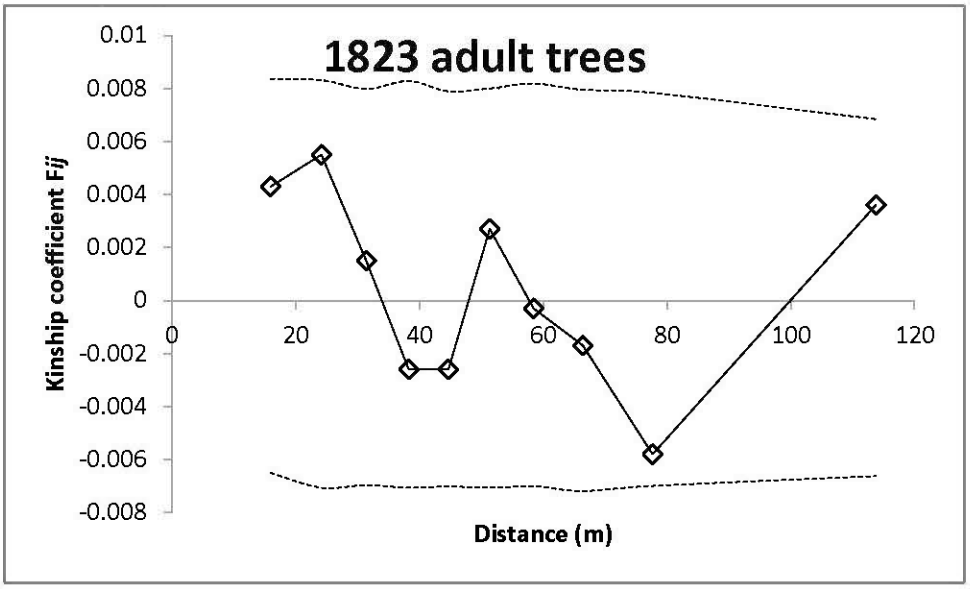
(a) 1916 saplings



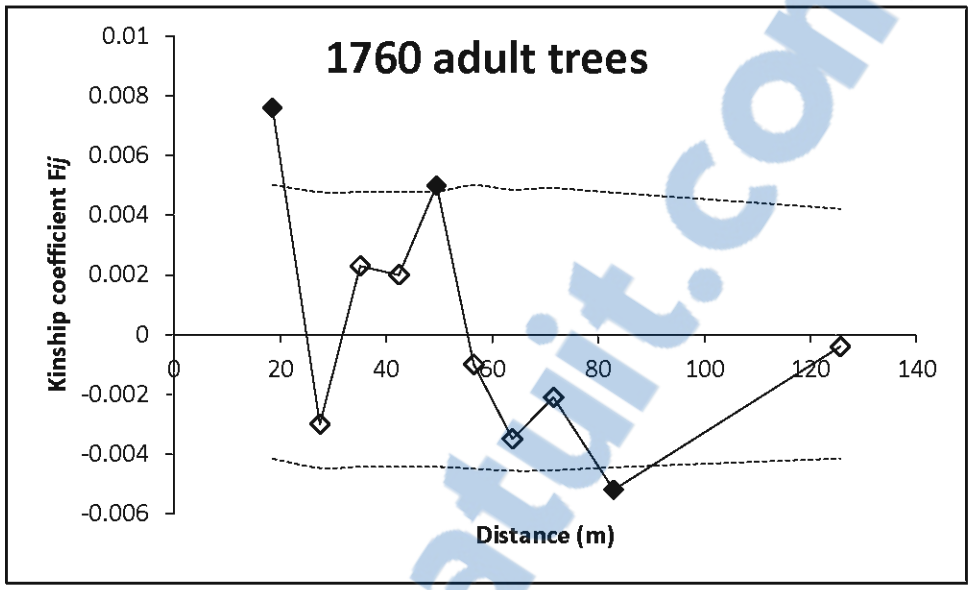
(b) 1823 saplings



(c) 1760 saplings



(d) 1823 adult trees



(e) 1760 adult trees

10 distance intervals were allocated based on similar number of paired samples in each distance class. Solid diamond represents significant autocorrelation. Double dashed lines indicate 95% confidence limits based upon null hypothesis that no autocorrelation exists.

Figure 3.3 Spatial genetic structures (SGS) of *Thuja occidentalis* in three plots based on autocorrelation between kinship coefficient F_{ij} and geographic distances ("genet" data set; a, 1916 saplings; b, 1823 saplings; c, 1760 saplings; d, 1823 adult trees; e, 1760 adult trees)

3.5.3 Parentage Analysis, Family Structure and Gene Dispersal Distance

In plot 1916, 36.8% (25) of the saplings that were analysed had two identified parents, 35.3% (24) had one identified parent, and 27.9% (19) had no identified parents (Table 3.4). In plot 1823, 24.7% (18), 69.9% (51), and 5.5% (4) of the saplings had two parents, one parent, and no parent, respectively. In 1760, these proportions of parentage were 20.0% (13), 26.2% (17), and 53.8% (35), respectively. In 1916, 17.6% of the parentage was assigned within the hectare, while 32.4% and 50.0% of the parents originated from 1823 and 1760, respectively. Most parents were found inside the hectare in 1823 (90.8%), as well as in1760 (79.1%).

Table 3.4 Parentage analysis for saplings of *Thuja occidentalis* ("genet" data set) in three plots since latest fires

Plot	Parent pair	One parent	No parent	Parentage distribution
1916	36.8% (25)	35.3% (24)	27.9% (19)	17.6% (1916); 32.4% (1823); 50.0% (1760)
1823	24.7% (18)	69.9% (51)	5.5% (4)	90.8% (1823); 9.2% (1760)
1760	20.0% (13)	26.2% (17)	53.8% (35)	79.1% (1760); 20.9% (1823)

In 1916, 7 of 13 adult trees were identified as parents, while 53 of 94 trees were identified as parents in 1823. In 1760, 40 of 145 trees were identified as parents (Table 3.5). Overall, among these parent trees (100), 47 were assigned as parents for a single offspring, while 53 contributed gametes to more than one offspring. Fifty-one of those trees contributed gametes to at least one offspring that was outside of their own plot, 23 came from 1823, and 28 from 1760. For example, 22 adult trees in 1760 were assigned as parents (either maternal or paternal) and 4 pairs as parent pairs (maternal and paternal) to saplings that were sampled in the 1916 plot. No selfing was detected.

Table 3.5 Number of adult trees of *Thuja occidentalis* identified as parents and their family structure, and gene dispersal distance

Plot	*N*	*P*	*P (1)*	*P (2+)*	*P& PP* (outside plot)	Shared Parents	Family structure	Gene dispersal
1916	13	7	4	3	0	0	Half-sib: 53, size: 3.0 (2 to 10) ; Full-sib: 1, size: 2.	Pollen: 1519.8 (7.5 to 4578.8), 66.1%; Seed: 1780.6(5.4 to7399.7), 62.5%.
1823	94	53	22	31	23(P:16& PP:2(1916); P:8&PP:1(1760))	13(1:1916,1760; 9:1916,1823; 3:1823,1760)		
1760	145	40	21	19	28(P:22& PP:4(1916); P:7& PP: 2(1823))	11(10:1916,1760; 1:1916,1823,1760)		
Total	252	100	47	53	51	24		

N: number of adult trees tested; *P*: number of adult trees assigned as parents within plot; *P (1):* number of adult trees assigned as parents within plot with a single offspring; *P (2+):* number of adult trees assigned as parents within plot with more than one offspring; P& PP (outside plot): number of the adult trees assigned as Parent & Parent pair to offspring outside their own plot; Shared Parents: number of adult trees as shared parents by offspring from different plots; Family structure: number of half-sib family and its average size, number of full-sib family and its average size; Gene dispersal: the average distance (m) of gene dispersal distance and its range mediated by pollen and seed calculated on the basis of the offspring that had their parent pairs assigned, and the percentage of gene dispersal realised within plot.

Twenty-four adult trees (1823, 13; 1760, 11) that were assigned as parents were shared by multiple saplings from different plots, among which 1 from 1760 contributed gametes to saplings within all three plots (Table 3.5). In total, 53 half-sib families (offspring sharing one parent) were found with numbers averaging 3.0 offspring per family. Among these, 1 full-sib family (offspring sharing the parent pairs) was found, with a size of 2 (Table 3.5). Mean dispersal distance was estimated to be 1519.8 m for pollen (ranging from 7.5 m to 4578.8 m), and 1780.6 m for seeds (5.4 m to 7399.7 m) based on the offspring that had their parent pairs assigned (Table 3.5). Most of the detected dispersals (66.1% for pollen and 62.5% for seeds) were realised within the plots over short distances.

3.6 Discussion

3.6.1 Regeneration Dynamics along a Post-fire Succession

The relative proportion of vegetative propagation to EWC recruitment slightly increases with time-since-fire. As the stand ages, the percentage of asexual contributions to regeneration increased from 21.8% in 1916 to 27.0% in 1823, and to 30.9% in 1760. At the same time, genotypic diversity decreased along the successional gradient.

Layering by buried branches and the formation of vertical stems from fallen or leaning trees are the two most common types of vegetative reproduction observed for EWC. Adventitious roots can be produced from buried branches or stems when moisture conditions are favourable (Godman 1958). Young seedlings (5-years-old) can produce layers (Johnston 1990). Stems originating from vegetative recruitment are generally more tolerant to shade and drought (Curtis 1946). Further, the development of roots on a stem while it is still attached to its parent is common (Newell 2005).

Contrary to our expectations, there were a limited proportion of stems that originated from layers versus seedlings in older stands. Much higher levels of vegetative propagation have been observed for EWC. For example, Nelson (1951) reported that more than 60% of EWC regeneration originated from layers in lowland areas. This is probably because the site conditions favoured vegetative propagation. The abundance of broadleaf trees in old stands (1823, 1760; see Table 3.1) does not produce a sufficiently thick humus layer that would allow for efficient layering. Moreover, broadleaf litters decompose faster than do conifer litters (Simard *et al.* 2003). In contrast, leaf litter of broadleaf species does not provide a good seedbed for seeds and smothers EWC seed germination (Simard *et al.* 1998; Simard *et al.* 2003).

The number of ramets per genet ranged from 2 to 5, and EWC clones were small. It is low relative to other clonal species, but still comparable to tree species that reproduce asexually through layering. For example, the number of ramets per genet has been commonly observed to range from 2 to 6 for black spruce (Viktora *et al.* 2011). Clones remain small because the mode of propagation (through layering) is not very prolific. We also found very small clones

in adult EWC (only 2 trees from the same clone) (Table 3.2). This is probably because intra-specific competition among clonal individuals, together with inter-specific competition, is very high, given that strong competition at the juvenile stage is common in forest trees (Kremer *et al.* 2012).

3.6.2 Gene Dispersal Distance Affects Family Structure

Previous studies have estimated seed dispersal distances to be 45 m - 70 m for EWC under normal conditions (Johnston 1990). Nevertheless, longer distance dispersal could be aided by animals (Rooney *et al.* 2002; Newell 2005). With parentage analyses, we successfully located the two parents of offspring and revealed the pattern of seed-mediated gene dispersal of EWC. The maximum realised seed dispersal distance was estimated to be 7399.7 m (mean = 1780.6 m) in this study. This contrasts remarkably with many other wind-disseminated tree species, which have exhibited limited seed dispersal (with mean distances of tens of metres), based on microsatellite parentage analyses that were reviewed by Ashley (2010) and Kremer *et al.* (2012). For example, maximum effective seed dispersal distance was around 100 m in temperate *Pinus densiflora* Siebold & Zucc (Iwaizumi *et al.* 2010), 1.4 km in temperate *Fraxinus excelsior* L. (Bacles *et al.* 2006), and 237 m in the subboreal conifer *Abies sachalinensis* F.Schmidt (Lian *et al.* 2008). Realised gene flow in our study was detected at the stage of naturally regenerated saplings (mean height: 56.6 to 63 cm, see Table S3.1), which differs from other studies that observed effective gene flow at the seed stage. In addition to differences in the life history traits among species and other ecological and environmental factors, the scale of the study can contribute to variation in the estimation of effective gene dispersal distance. The distance between our sampling plots ranged from 4.3 km to 7.2 km, which is greater than the scale used in previous studies. Using microsatellite parentage assignment, the longest pollination distance that was previously detected in wind-pollinated tree species ranged from 629 m (mean = 47 m) in temperate conifer *Pinus sylvestris* L. (Robledo-Arnuncio & Gil 2005) to over 10 km in conifer *Pinus ponderosa* (Lesser and Jackson 2013 EL) and in temperate *Populus trichocarpa* Torrey & Gray (mean = 7.6 km) (Slavov *et al.* 2009). The longest realised pollen movement distance was estimated to be 4.5788 km (mean = 1519.8 m) in EWC for an area covering 80 km^2 compared to > 300 km^2 for *Populus trichocarpa*. Indeed, the maximum possible pollination distance within the

study area reflects the maximum effective pollen movement distance between trees (Ashley 2010).

Longer distances of realised seed dispersal (maximum: 7399.7 m) were detected between sites 1760 and 1916, indicating that two-winged light seed of EWC travels up to several kilometres on air currents or by gliding on ice, which demonstrates its good colonising abilities. This explains that the initial invasion of the youngest site (1916) by EWC was achieved by seed. Established trees also received external pollen and started to produce offspring, as revealed by parentage assignment analysis. This colonisation process was intensified by vegetative reproduction through layering, a process that was neither space- nor parent-limited, as we saw substantial numbers of trees being assigned as parents across plots. In particular, a great number of adult trees in the older sites were shared as parents by saplings. The family was not very large in size (2 to 10), but spanned the three sites over distances of several kilometres ($\approx$ 7 km). Long-distance parent-offspring relationships are common in wind-pollinated species (Robledo-Arnuncio & Gil 2005; Slavov *et al.* 2009).

The detected seed dispersal distance in our study was greater than that of pollen based on the offspring that had their parent pairs assigned. Generally, pollen dispersal distance is greater than that of seed in wind-disseminated systems (Kremer *et al.* 2012). Very few examples can be found where seed flow exceeds pollen flow, except for the temperate tree species *Fraxinus excelsior* (Bacles *et al.* 2006) and the tropical palm *Iriartea deltoidea* Rui & Pav (Sezen *et al.* 2007). However, the interactions between biological and physical processes that influence dispersal distance are not well understood (Kremer *et al.* 2012). It has been reported that the seed of black spruce can be dispersed over snowpack by strong winds, with subsequent seed redistribution being aided by melting snow in the following spring (Gamache *et al.* 2003). In addition to weather conditions (particularly wind), features of the lacustrine landscape in our study area contributed to the longer seed dispersal distance, including seeds gliding over ice (Xu *et al.* submitted-a). In addition, pollen competition most likely affects pollination distance. The 1760 stand is located between the 1823 and 1916 stands. The pollen cloud that originates from the 1760 stand has much greater chance of saturating EWC trees in the 1916 stand than in the 1823 stand, given that the former is located at half of the distance. This competition does not exist for seeds.

Saplings ("genet" data set) from the same patch often shared the same maternal parent. Parents having many offspring are more likely to have offspring that were located across sites (longer distance) than parents having a single offspring. In older sites, most of parentage was found inside the plot (1823, 88.8%; 1760, 95.7%), indicating that mature trees located within the plot produced most of the offspring. Only 2.7% of the offspring in plot 1823 had no parents assigned. This suggests that trees with DBH $\geqslant$ 25 cm contributed the most to descendants. We did not find selfing among saplings (i.e., a tree that could be assigned as both a maternal and a paternal parent). Yet, high self-fertilisation has been frequently reported in EWC seeds (Perry & Knowles 1990). This indicates that self-fertilised seed of EWC did not successfully germinate, or that germinated seeds were not sufficiently competitive to survive.

3.6.3 Clonal Growth Increases Fine-scale SGS and the Pattern of SGS along a Post-fire Succession

Our results presented contrasting patterns of fine-scale SGS between saplings and adult trees. The significant fine-scale SGS observed in saplings is likely the result of patterns of seedling recruitment. Maternally related seedlings are clumped together (phalanx propagation) and generate a significant fine-scale SGS at the first distance interval for all three plots. There was a clear pattern of decreasing SGS at first distance interval in saplings along the succession. When identical MLG (clonal) saplings were included in the analysis, SGS at the first distance interval increased in all the plots, as did overall SGS (*Sp*). The overall SGS (*Sp*) in the oldest plot (1760) was subsequently the highest. Indeed, the relative contributions of vegetative versus sexual reproduction have an impact on fine-scale SGS (Epperson 2000). Clonal reproduction increased fine-scale SGS within EWC in this study. This result contrasts with another study that reported no contribution of vegetative propagation to fine-scale SGS in EWC (Pandey & Rajora 2012b).

There were low levels of fine-scale SGS in adult trees (Sp, 0.0025~ 0.0045), while localised seeds induced a strong SGS in saplings ("genet" data set) (*Sp*, 0.0286 ~ 0.0348). This finding over life stages is consistent with most studies (Chung *et al.* 2003; Oddou-Muratorio *et al.* 2004). The *Sp* value found in adult trees was comparable to the mean value of *Sp* (0.0068) in

trees with both wind-dispersed pollen and seed (Vekemans & Hardy 2004). However, it is lower than the value of *Sp* (0.014 ~ 0.023) previously reported for adult EWC trees (Pandey & Rajora 2012b). This is probably the consequence of restricted gene flow between fragmented populations in that study. The significant fine-scale SGS in adult trees in the 1760 plot was not observed in the 1823 plot. It was likely caused by a decrease in the population density of adult trees between 1823 and 1760 plots, which is the major determinant of the level of fine-scale SGS (Vekemans & Hardy 2004). Negative F_{is} values in old trees at late successional stages indicated heterozygote excess. This pattern of increasing heterozygosity can be explained as selection against inbred genotypes, which has been commonly reported in many tree species (Bush & Smouse 1992; Oddou-Muratorio *et al.* 2004; Hoebee *et al.* 2006).

3.7 Conclusions

This study revealed the importance of sexual and asexual regeneration along a post-fire succession for a late-successional tree species in the boreal forest. Contrary to our expectations, higher asexual recruitment does not explain the increased abundance of EWC in older stands. It is likely due to a higher abundance of seed trees and to better conditions for seed germination and establishment. The distance from seed sources and the time needed for post-fire regeneration to produce seed-bearing trees are among the factors that controlled EWC abundance, together with the abundance of suitable microsites. Our results provide important insights into the dynamics of natural population regeneration along a post-fire succession. Our study showed empirical evidence for extensive pollen- and seed-mediated gene flow, which is associated with post-fire colonisation. It also demonstrated a great ability of EWC in dispersing seeds over relatively long distances in a lacustrine landscape. Vegetative recruitment accompanied sexual reproduction during EWC stand development. It increased SGS at fine-scales in saplings, especially in older stands in which vegetative recruitment became more common.

3.8 Acknowledgements

This work forms part of the doctoral dissertation of H. X. We thank Marie-Hélène Longpré and Danielle Charron for logistical support, Véronique Paul, Albanie Leduc, Geneviève Trudeau, Rébecca Libert, and Sandrine Chauchard for help with the field work, and Yiqin Zhou for assistance in the laboratory. Special thanks are owed to the industrial supervisors from the company Tembec: Louis Dumas, Sonia Légaré, and Geneviève Labrecque. We are grateful to Dr. William F.J. Parsons for careful English proofreading. This research was supported by an Industrial Innovation Doctoral Scholarship from NSERC (Natural Sciences and Engineering Research Council of Canada), a BMP innovation doctoral scholarship from FRQNT (Fonds de recherche du Québec - Natures et technologies) and Tembec, and a scholarship from FERLD (Forêt d'enseignement et de recherche du lac Duparquet) to H. X. This research was also funded by a NSERC strategic grant [STPGP 336871 to F. T].

CHAPTER IV

IMPORTANCE OF LARGE OLD-GROWTH FOREST PATCHES IN MAINTAINING THE GENETIC DIVERSITY OF "*ARBOR VITAE*", THE EASTERN WHITE CEDAR (*THUJA OCCIDENTALIS* L.), IN BOREAL FIRE-DOMINATED LANDSCAPES

Article submitted to *Conservation Biology*

4.1 Résumé

Les perturbations à grande échelle fréquentes, comme les feux de forêts, créent une mosaïque complexe dans le paysage de la forêt boréale, représentée par de grandes surfaces brûlées et des peuplements résiduels non brûlés. Les forêts du cèdre blanc de l'Est résiduelles non brûlées sont essentielles dans la mesure où elles constituent des habitats pour les espèces sensibles aux perturbations et des refuges biologiques potentiels. Nous avons utilisé des marqueurs moléculaires (18 microsatellites) pour étudier la diversité génétique du cèdre blanc de l'Est (CBE, *Thuja occidentalis* L.). Pour mieux comprendre les effets des patrons du paysage sur la structure génétique du CBE, nous avons comparé des peuplements représentés par de petits îlots de CBE préservés et fragmentés dans un matrice forestière dominée par les espèces de début de succession, un massif de vielles forêts et des peuplements de CBE situés sur des îles du lac Duparquet. Nos résultats montrent une dynamique source-réservoir associée à un taux élevé d'échanges génétiques à l'échelle du paysage. Nous avons mis en évidence des différences significatives de la richesse allélique et de la différentiation des populations avec les petits îlots non brûlés de CBE renfermant la plus faible richesse allélique (5.06), la plus forte différentiation des populations (0.052) et le nombre d'allèles privés le plus bas (peuplements résiduels non brûlés, 5; îles, 8; vieilles forêts, 15). Notre étude suggère que le vent et le paysage façonnent la structure génétique du CBE. Nous avons montré que les massifs de vieilles forêts ont une valeur de conservation plus forte que les petits peuplements non brûlés pour la préservation de la diversité génétique du CBE. Ainsi, des mesures appropriées de protection devraient être prises pour éviter la fragmentation et la diminution de la taille des massifs de vieilles forêts afin de limiter la perte de diversité allélique de populations de CBE.

4.2 Abstract

Frequent large-scale natural disturbances by wildfires create complex forest mosaics in the boreal landscape comprising large burned areas in combination with small remaining unburned stands. The unburned forests remnants of the eastern white cedar, or fire residuals, are essential to provide crucial habitats for disturbance-sensitive species and function as fire-free refugia. In this study, we used molecular markers (18 polymorphic microsatellite loci) to investigate the genetic diversity of the eastern white cedar (EWC, *Thuja occidentalis* L.), in forest remnants that successfully escaped consecutive wildfires. We took advantage of the co-existence of three types of landscape, including naturally fragmented EWC islands, small-fragmented EWC fire skips, and non-fragmented old-growth EWC mainland forests in the terrestrial landscape, to better understand the effects of landscape features on genetic structure. Our results revealed a source-sink dynamic associated with a high level of gene flow in the landscape. There were significant differences in allelic richness and population differentiation among the three landscape types, with small fire skips having the lowest allelic richness (5.06), the highest population differentiation (0.052), and the fewest private alleles (fire skips, 5; islands, 8; mainland, 15). Our study suggests that both wind and landscape features shape the genetic structure of the EWC in this landscape. We provided genetic evidence that the large, mainland old-growth EWC patches within a successional forest matrix have a higher conservation value than small fire skips embedded within single fires. Therefore, appropriate protection measures should be taken to avoid fragmentation that may lead to a reduction in allelic diversity and to ensure that remnant stands can continue to function as biological refugia for fire-sensitive species.

4.3 Introduction

The discipline of "landscape genetics" combines population genetics and landscape ecology. It aims to study how landscape features interact with micro-evolutionary processes and their effects on the genetic structure (Manel *et al.* 2003; Storfer *et al.* 2007; Manel & Segelbacher 2009; Manel *et al.* 2010). Much of the past research efforts in this field have focused on animals, and plants have received less attention (Holderegger *et al.* 2010; Storfer *et al.* 2010). Most studies on plants have documented the impact of landscape fragmentation on genetic processes and predicted an increase in population differentiation and a reduction in genetic diversity in a fragmented landscape (Sork & Waits 2010). Wind-pollinated tree species may have different responses, i.e., maintaining a high genetic diversity (Bacles *et al.* 2006) and being resilient to the negative genetic consequences of fragmentation (Craft & Ashley 2007).

Wildfire, insect, and windthrow are the most common natural disturbances in the boreal forest (Kneeshaw *et al.* 2011). Wildfire, in particular, is the one that controls boreal forest dynamics (Heinselman 1981; Payette 1992; Morin *et al.* 1993). Large-scale disturbances by wildfires create complex shifting forest mosaics in the boreal landscape that include large burned areas in combination with a matrix of patches of mature forest of different ages, including many patches that have reached the old growth stage. Old growth forests can also be observed as much smaller patches of unburned stands (small fire skips) embedded in larger burned areas. These forest remnants are usually late-successional forest and offer excellent refuges for fire-sensitive species (Segerström 1997; Fenton & Bergeron 2008). They also serve as a seed source for forest reestablishment (Asselin *et al.* 2001). Hence, natural fire residuals gain a special status in conservation and have been considered to be biological refugia and hotspots of biodiversity in the boreal landscape (Hörnberg *et al.* 1998; Gandhi *et al.* 2001; Mosseler *et al.* 2003).

In the eastern Canadian boreal landscape, old growth forests observed in the matrix and small forest remnants that escaped the forest fires are often dominated by eastern white cedar (EWC, *Thuja occidentalis* L.). It is a late successional, long-lived conifer tree species (Archambault & Bergeron 1992) that becomes dominant 200-250 years after disturbance (Bergeron 2000). The EWC natural forest remnants function as refugia in a landscape

affected by recurrent large-scale wildfires (Ouarmim *et al.* submitted). However, under certain ecological conditions, recurrent wildfires can wipe out forest remnants or turn them into even smaller-sized forest patches (small fire skips). Fire skips are often found within fragmented landscape and embedded in a matrix of young pioneer tree species (e.g., paper birch, trembling aspen). These fire skips can eventually lose their function as refugia.

In this study, our main objective was to investigate the genetic conservation value of EWC in fire residuals that play a key role in providing habitat reserves for fire-sensitive species. We took advantage of the co-existence of three types of landscapes, including naturally fragmented EWC islands in a lacustrine landscape, small-fragmented EWC fire skips, and non-fragmented mainland EWC old forests in a terrestrial landscape, as a model to better understand the effects of landscape features on the genetic structure in EWC. Our hypotheses were the following: 1) mainland EWC old forests are the main source of gene flow; 2) small EWC fire skips are more genetically differentiated due to the barrier to gene flow created by the presence of an early successional forest matrix; 3) small EWC fire skips have the lowest genetic diversity.

4.4 Materials and Methods

4.4.1 Study Area

The study area is located in northwestern Quebec, in the Lake Duparquet Research and Teaching Forest (79°10'W-48°30'N), covering an approximately 80 km^2 boreal landscape and more than 170 islands (Bergeron & Dubuc 1989). The forest is dominated by balsam fir (*Abies balsamea* [L.] Mill.), EWC, white spruce (*Picea glauca* [Moench] Voss), and black spruce (*Picea mariana* [Mill.] BSP). Pioneer species, such as paper birch (*Betula papyrifera* Marsh), trembling aspen (*Populus tremuloides* Michx.), and jack pine (*Pinus banksiana* Lamb.), occupy large areas following a disturbance (Bergeron & Bouchard 1984; Bergeron & Dubuc 1989). The periodic occurrence of fires since the end of the last glaciation period (cal. 8000 BP) has been well documented (Carcaillet *et al.* 2001). Eight major fires burnt in this forest between 1717 and 1944 (Bergeron 2000). The study forest was selected because it has been minimally affected by human intervention (Bergeron & Dubuc 1989; Bergeron 2000).

4.4.2 The Species

Eastern white cedar (EWC; *Thuja occidentalis* L.), also known as northern white cedar, has a large area of natural distribution in North America. It is listed as endangered in Indiana, Massachusetts, and New Jersey and as threatened in Connecticut, Illinois, Kentucky, and Maryland (National Plant Data Center, USDA). It is a monoecious and wind-pollinated conifer species (Johnston 1990). Its flowers usually develop in late April or May; cone formation starts in late June, and the cones mature in August but may remain attached to the tree until the following spring (Johnston 1990). The seeds have two flat lateral wings with two cotyledons, and their dispersal usually begins in September and ends in November, being mainly disseminated by the wind, but may continue throughout the winter (Johnston 1990).

4.4.3 Sampling

Nine sites (Figure 4.1) covering three different types of landscape were selected. Three small EWC fire skips located in an area that last burned in 1944 were selected to represent sites fragmented in a forest landscape dominated by surrounding young successional broadleaved species. All the three sites (f441, f442, f443) have escaped forest fires since 1717 (Bergeron 1991; Ouarmim *et al.* submitted). The number of EWC adult trees ranged from 20 (f442) and 25 (f441) to <100 (f443). Three naturally fragmented islands (> 100 EWC adult trees) – is39, is42, and is134 – last burned in 1889, 1825, and after 1825 (precise year unknown), respectively, representing the lacustrine landscape that was selected.

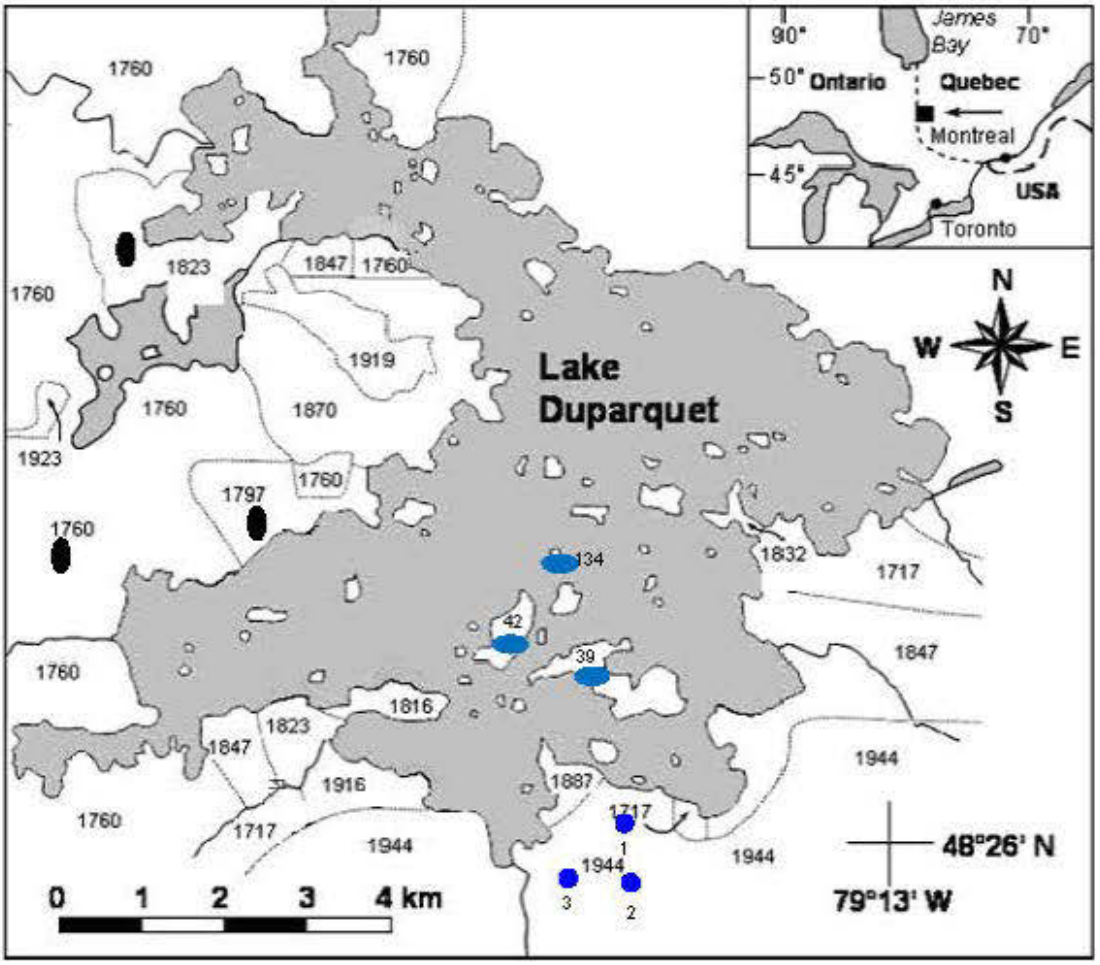


Figure 4.1 Nine study sites of Eastern white cedar covering three landscape types in Quebec, Canada

Three mainland old sites (f60, f97, f23) that had 500 to more than 1000 EWC adult trees per hectare, which were last burned in 1760, 1797, and 1823, respectively, were selected to represent non-fragmented sites with different times since a fire on the mainland. The distance ranged from 300 to 900 m between small fire skips, 600 m to 2.1 km between islands, and 1.7 to 4.4 km between mainland sites. It varied from 1.6 to 8.8 km across the three different landscapes.

The EWC adult trees were randomly selected for sampling. The distance between each selected tree was greater than 20 m whenever possible to reduce the probability of selecting ramets or closely related individuals. Foliage or cambial root tissues were collected for genetic analysis. All the sampled materials were kept in the freezer (-20°C) before the genetic analysis.

4.4.4 DNA Extraction and Genotyping

The leaves were ground in liquid nitrogen, and genomic DNA was extracted and amplified over 18 loci (TO791, TO605, TO328, TO53, TO925, TO727, TO659, TO29, TO737, TO587, TO512, TO503, TO715, TO521, TO418, TO20, TP10, and TP11) using methods previously described (O'Connell & Ritland 2000; Xu *et al.* 2013). The amplified DNA was sent to Genome Quebec (Montreal, Canada) for genotype analysis on an ABI 3730 genetic analyzer (Applied Biosystems, Carlsbad, California, USA). The results were analyzed using the GeneMapper 3.7 software (Applied Biosystems).

4.4.5 Data Analysis

The presence of null alleles and scoring errors were tested in Micro-Checker (Van Oosterhout *et al.* 2004). The basic genetic diversity estimates within stands, including the number of alleles per locus (N_a), observed heterozygosity (H_o), expected heterozygosity (H_e) and private alleles (PA), were calculated in GenAlex 6.4 (Peakall & Smouse 2006). Allelic richness (AR) was estimated in FSTAT 2.9.3.2 (Goudet 2001). We tested for Hardy-Weinberg equilibrium (HWE) using samples from each stand over each locus (162 tests: 9 x 18) and for all pooled samples as one group (18 tests: 1 x 18) and for linkage disequilibrium (LD) over 18 loci (153 pairs) for all pooled samples in FSTAT 2.9.3.2 (Goudet 2001). The Bonferroni correction (Rice 1989) was applied to test for the significance of the equilibrium.

The samples were grouped for each landscape type, and genetic diversity estimates including AR, H_o, H_e, inbreeding coefficient (F_{is}) (Weir & Cockerham 1984), population differentiation (F_{st}) (Weir & Cockerham 1984), and population relatedness (Rel) (Hamilton 1971) among the groups (1000 permutations) were compared. The pairwise genetic differentiation among the

stands and the associated significance were tested with a G-test and 36000 permutations (Goudet *et al.* 1996). All the analyses were performed in FSTAT 2.9.3.2 (Goudet 2001).

The mode-based Bayesian clustering approach implemented in STRUCTURE (version 2.3.2) was used to explore the genetic differentiation among the stands without using their locations (Pritchard *et al.* 2000). The model assumed that within assumptive populations (K), the loci are in Hardy-Weinberg equilibrium and in linkage equilibrium. This version (2.3.2) can handle weakly linked markers. The number of clusters (K, ranging from 1 to 9) was estimated with an admixture model and the option of correlated allele frequency. This model is the most appropriate for inferring structure among populations that likely share a common ancestry or that experienced migrations (Falush *et al.* 2003). All parameters were set following the user's manual. To choose an appropriate run length, we performed a pilot run that showed that burn-in and MCMC (Markov chain Monte Carlo) lengths of 300,000 each were sufficient to obtain consistent data. Increasing the burn-in or MCMC lengths did not improve the results significantly. Ten replicate runs for each value of K were performed. The most likely number of cluster (value of K) was chosen by following the methods of Evanno *et al.* (2005). The results were then graphically displayed using DISTRUCT (Rosenberg 2004). A principal coordinates analysis (PCoA) was also performed using the covariance-standardized method based on the population genotypic genetic distance in GenAlex 6.4 (Peakall & Smouse 2006).

The correlation between the geographic and genetic distances (F_{st}) (Weir & Cockerham 1984) among the populations was tested with a Mantel test (999 permutations) (Mantel 1967) in GenAlex 6.4 (Peakall & Smouse 2006). A Bayesian multilocus genotyping procedure implemented in BAYESASS v. 1.3 using MCMC methods (Wilson & Rannala 2003) was used to estimate the direction and rate of recent immigration. It was estimated to be lower than 5 generations (or < 150 years) given the estimated generation time of 20-30 years (Johnston 1990). The number of iterations for the chain was 3,000,000, with 999,999 being burn-in, and the sampling frequency was 2000. An increase in the iterations did not affect the output. We repeated the analysis three times to verify its stability.

The "detection of first generation migrants," implemented in the GeneClass2 software, was chosen to detect the migrants among the populations (Piry *et al.* 2004). The ratio of the likelihood (L=L_home/L_max) was used for migrant detection (Paetkau *et al.* 2004), in which L_home represents the likelihood of drawing a given individual's genotype from the population where this individual was sampled and L_max represents the greatest likelihood value among all population samples including the L_home population. The Bayesian method of computing genotypes (Rannala & Mountain 1997) was selected as the criterion for likelihood computations, which appears to be a better method than the distance-based one (Cornuet *et al.* 1999). A Monte Carlo re-sampling algorithm (Paetkau *et al.* 2004) was chosen to compute the probability.

4.5 Results

4.5.1 Genetic Diversity, Private Allele and Test of Equilibrium

We did not detect any null alleles or scoring errors. The N_a ranged from 4.9 (f442) to 6.4 (f23), with an average of 5.8. The *AR* averaged 5.4, ranging from 4.9 (f442) to 5.8 (f23). The mean *Ho* was 0.718, ranging from 0.674 (f443) to 0.774 (is39), and the mean *He* was 0.619, ranging from 0.594 (f442) to 0.638 (f23) (Table 4.1). All sites contained private alleles. The mainland had the highest number of private alleles (15), followed by the islands (8), and the small fire skips had the lowest number (5).

Table 4.1 Genetic diversity parameters in nine Eastern white cedar stands from three landscape types in Quebec, Canada

Landscape type	Pop	*N*	N_a	*AR*	*PA*	H_o	H_e
Small fire skips	f441	25	5.3	5.1	1	0.744	0.605
	f442	20	4.9	4.9	2	0.722	0.594
	f443	30	5.6	5.1	2	0.674	0.595
Islands	is39	30	6.2	5.6	5	0.774	0.637
	is42	30	5.8	5.3	1	0.681	0.598
	is134	30	5.7	5.3	2	0.717	0.637
Mainland	f97	30	6.1	5.5	4	0.693	0.636
	f23	30	6.4	5.8	4	0.731	0.638
	f60	30	6.0	5.5	7	0.726	0.629
	Mean	28	5.8	5.4	-	0.718	0.619

N: number of individuals genotyped per population; N_a: average number of alleles per locus; *AR*: allelic richness; *PA*: private alleles; H_o: observed heterozygosity; H_e: expected heterozygosity.

The samples were generally in HWE. Among the 18 loci, 13 had an insignificant excess of heterozygotes ($P > 0.00278$, adjusted 5% nominal level: 0.00278), and five had an insignificant deficit of heterozygotes ($P > 0.00278$ adjusted 5%: 0.00278) (data not shown). Among 162 tests of the samples from each stand over each loci, 15 (9.3%) showed a slight excess of heterozygotes ($P = 0.00030$, adjusted 5%: 0.00031), and the remainder did not deviate from HWE ($P > 0.009$, adjusted 5%: 0.00031) (data not shown). Among 153 loci pairs, 145 were not in linkage disequilibrium ($P > 0.0007$, adjusted 5%: 0.000327), and only eight were weakly linked ($P = 0.000330$, adjusted 5%: 0.000327) and could be considered in linkage equilibrium (data not shown).

4.5.2 Population Structure and Differentiation in the Landscape

There were significant differences in allele richness and population differentiation among the three different landscape groups (Table 4.2). Small fire skips had the lowest *AR* (5.061; $P =$

0.02) and the highest F_{st} (0.052; P = 0.05). No significant differences in the H_o, H_e, F_{is}, and *Rel* were observed.

Table 4.2 Comparison of genetic diversity parameters among nine Eastern white cedar stands from three landscape types in Quebec, Canada

Group	*AR*	H_o	H_e	F_{is}	F_{st}	R_{el}
Small fire skips (f441, f442, f443)	5.061	0.710	0.608	-0.168	0.052	0.117
Islands (is39, is42, is134)	5.419	0.724	0.633	-0.144	0.015	0.034
Mainland (f97, f23, f60)	5.588	0.717	0.644	-0.113	0.023	0.050
P-values	0.02	0.94	0.12	0.47	0.05	0.06
	*	ns	ns	ns	*	ns

AR: allelic richness; H_o: observed heterozygosity; H_e: expected heterozygosity; F_{is}: inbreeding coefficient; F_{st}: population differentiation; R_{el}: population relatedness; ns, not significant; *: $P \leq 0.05$.

Most of the pair-wise F_{st} values (34 out of 36) were significantly different from zero, indicating a substantial genetic population differentiation (Table 4.3). Within each group, the pair-wise F_{st} values were the highest between small fire skips (0.0473-0.0588), followed by the mainland old stands (0.0188-0.0287) and the islands (0.009- 0.0242).

Table 4.3 Population differentiation (pair-wise F_{st}) among nine Eastern white cedar stands from three landscape types in Quebec, Canada

	f441	f442	f443	is39	is42	is134	f97	f23	f60
f441		***	***	**	***	***	***	***	***
f442	0.0588		***	***	***	***	***	***	***
f443	0.0473	0.0530		***	***	***	***	***	***
is39	0.0205	0.0489	0.0491		***	**	***	***	***
is42	0.0384	0.0563	0.0423	0.0242		ns	***	***	***
is134	0.0320	0.0574	0.0458	0.0117	0.0090		ns	***	***
f97	0.0400	0.0498	0.0435	0.0157	0.0165	0.0082		***	***
f23	0.0314	0.0594	0.0566	0.0239	0.0195	0.0145	0.0188		***
f60	0.0518	0.0586	0.0336	0.0402	0.0307	0.0211	0.0287	0.0204	

F_{st}: below diagonal; Significance: above diagonal; ns, not significant; **, $0.001 \leq P < 0.01$; ***$P < 0.001$.

The PCA analysis revealed a clustering of populations with the first and second axes explaining 33.86% and 24.15%, respectively, of the variation (Figure 4.2). For example, the three fire skips (f441, f442, f443) were far from each other, indicating a higher differentiation among stands. In contrast, stands from the mainland and from the islands were clumped together.

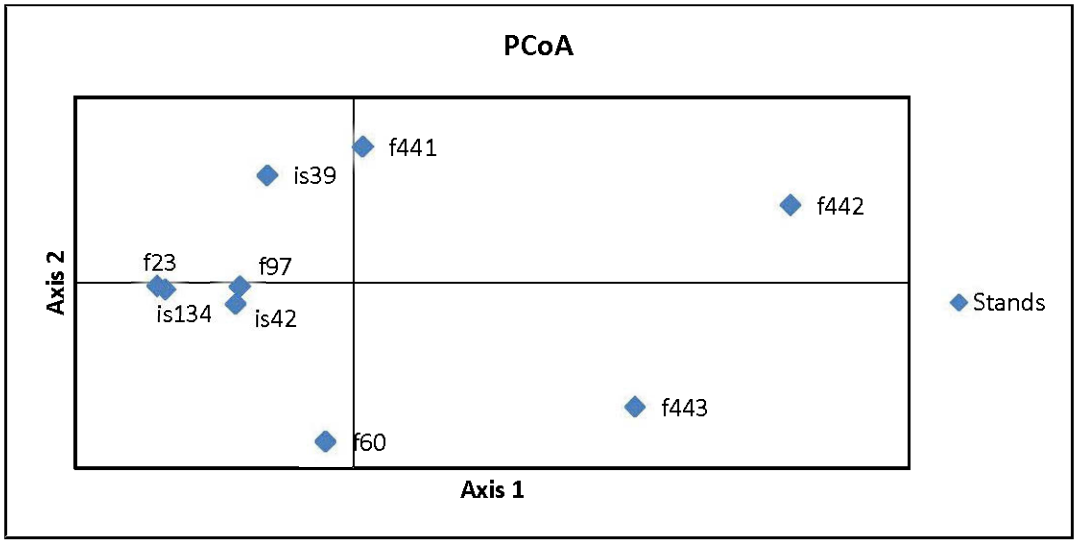


Figure 4.2 Principal coordinates analysis (PCoA) of genetic distances among nine populations of Eastern white cedar from three landscape types. Axis 1 explained 33.86% of the variation, and axis 2 explained 24.15%; Lacustrine landscape (islands): is39, is42, is134; Non-fragmented forest landscape (mainland old stands): f60, f97, f23; Fragmented forest landscape (small fire skips): f441, f442, f443.

All loci were used in the Bayesian analysis. The most likely number of clusters (K=4) was inferred by plotting L (K) (Figure S4.1), and confirmed by plotting ΔK (Figure S4.2). The four clusters were displayed in blue, yellow, orange, and green (Figure 4.3). The patterns were consistent with the results of the PCA and pair-wise F_{st} analysis. For example, there was a quite different pattern of the membership of clusters (Q) among the three small fire skips (i.e., >50% of Q: f441, orange 58.9%; f442, orange 48.3% and yellow 30.4%; f443, green 63.5%), indicating that they were genetically different from each other. Two of the islands

had a similar pattern of Q (is42, blue 50.7% and yellow 21.2%; is134, blue 42.3% and yellow 26.8%).

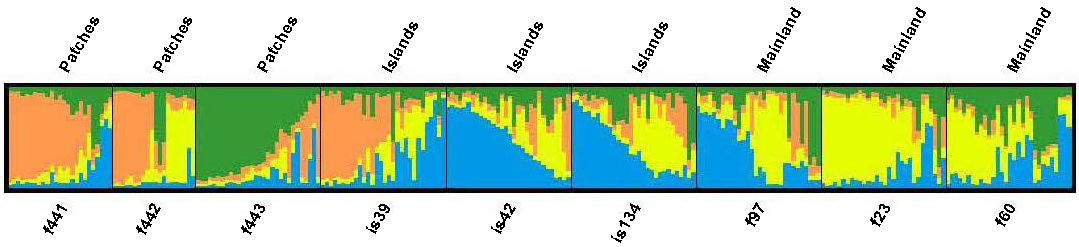


Figure 4.3 Assignment plotting of posterior probability for 255 Eastern white cedar trees collected in nine stands from three landscape types in Quebec, Canada. The trees were grouped into four genetic clusters (K=4) based on Bayesian analysis implemented in STRUCTURE 2.3.2 at 18 microsatellite loci. The four clusters are displayed in blue, yellow, orange, and green.

4.5.3 Gene Flow in the Landscape

No significant correlation between geographical and genetic distances (F_{st}) was detected (R = 0.123, P = 0.252), indicating that the populations were not isolated by distance. The recent immigration analysis showed that two mainland populations (f97, f23) served as the main source of migrants (Table S4.1). The gene flow pattern is shown in Figure 4.4 (migration rate $\geqslant$ 0.136). The receiver populations had a large proportion of migrants from f23 for f441 (0.225) and f442 (0.271); from f97 for f443 (0.273), is42 (0.291), is134 (0.291), and f60 (0.294); and from both f97 and f23 for is39 (f97: 0.136; f23: 0.155). The source population f23 received a significant proportion of migrants (0.290) from f97. The population f97 had the highest proportion of non-migrants (0.988). The detection of first generation migrants showed that only 6 individuals out of 255 (3 in small fire skips and 3 in islands) had a probability value < 0.01 of originating from the population where they were sampled and were thus considered immigrants. No immigrant was detected in the mainland.

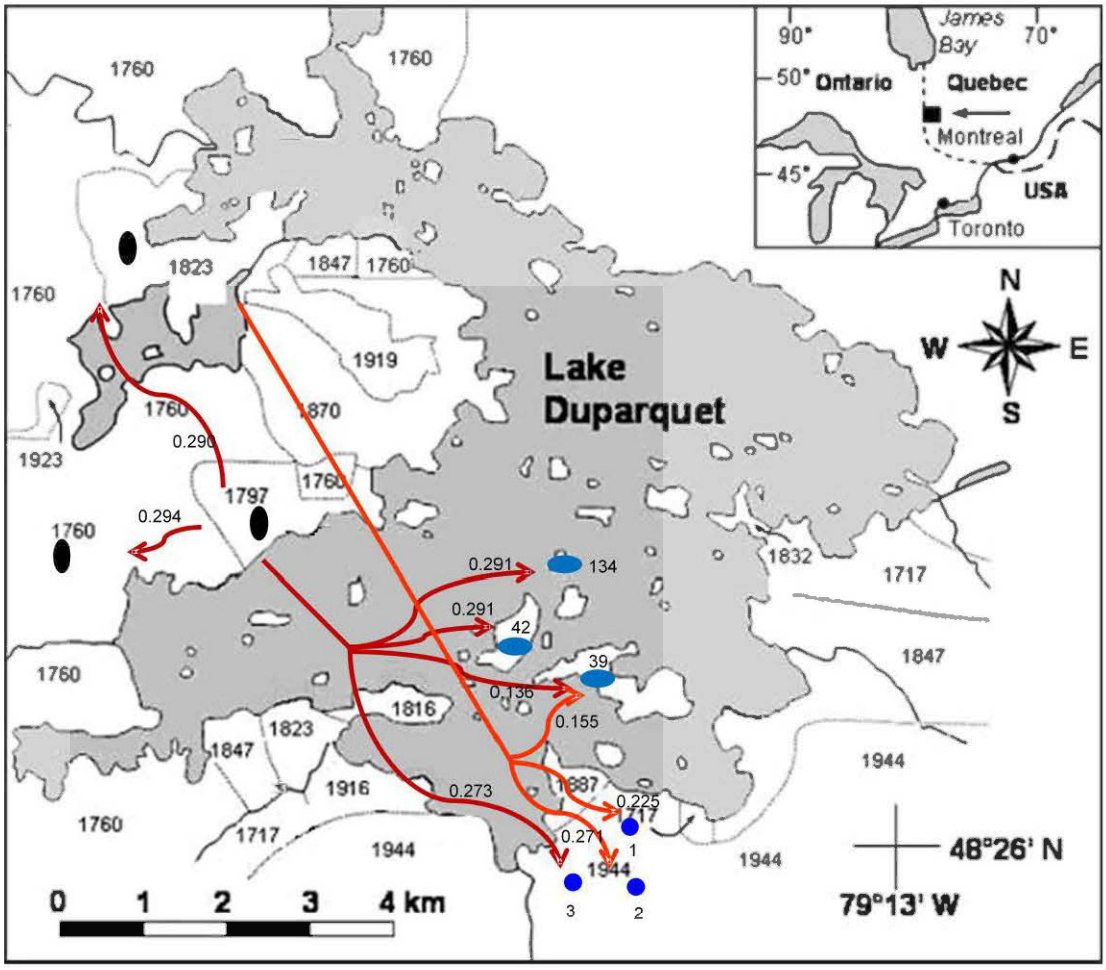


Figure 4.4 Gene flow pattern among nine Eastern white cedar stands from three landscape types in Quebec, Canada

4.6 Discussion

4.6.1 The Differences of Genetic Diversity among the Three Landscape Types

Landscape features may act as barriers to isolate populations or may enhance dispersal and gene flow (Manel *et al.* 2003; Antolin *et al.* 2006). Landscape futures significantly influenced the level of genetic differentiation between the EWC populations. Population genetic differentiation increased from the lacustrine landscape (islands, F_{st} = 0.015) to the non-fragmented forest landscape (mainland old stands, F_{st} = 0.023) and the fragmented forest landscape (small fire skips, F_{st} = 0.052). Increased levels of genetic differentiation (F_{st}) across the landscape types suggested that the forest matrix acted as a slight barrier to gene flow. Our study also revealed a significant decrease in allelic richness from the non-fragmented forest landscape (mainland old stands, AR = 5.59) to the lacustrine landscape (islands, AR = 5.42) and the fragmented forest landscape (small fire skips, AR = 5.06). This result also indicated that allelic richness was more sensitive to population decline than other estimates of genetic diversity (e.g., heterozygosity), and variation in this parameter was easier to detect. Small fire skips were fragmented in the landscape. This fragmentation increased the distances between EWC stands and created barriers to pollen and seed dispersal as the surrounding forest matrix hindered the gene movement between stands. Our results confirmed the hypotheses that small EWC fire skips were more genetically differentiated due to the barrier to gene flow created by the early successional forest matrix.

4.6.2 Gene Flow among Stands in the Landscape

The absence of obstacles in the lacustrine landscape increased the effectiveness of seed and pollen dispersal (Gauthier *et al.* 1992). In winter, conifer seeds or cones can glide more easily on the ice of the lake than through densely forested areas (Critchfield 1985). Indeed, tree populations within open landscapes exchange a substantial number of migrants (McRae 2006). The gene dispersal of EWC is not spatially limited in this area (Xu *et al.* submitted-b), and EWC is usually found to be dominant in old growth stands that escaped from fires, which makes them strong recipients of genes from outside the stand.

Although small EWC fire skips are fragmented, there is still an exchange of genes occurring among the EWC stands. The absence of significant isolation by distance (IBD) between the genetic (F_{st}) and geographic distances suggested a high level of gene flow in the landscape. The high amount of gene flow among the EWC stands in the landscape buffers the effects of fragmentation on the genetic characteristics of the small EWC fire skips. The relative long dispersal distance of EWC seed and pollen (Xu *et al.* submitted-b) facilitates the inter-stand gene exchange among naturally fragmented EWC island populations in lacustrine landscape and among non-fragmented mainland old EWC stands and small-fragmented EWC fire skips in the terrestrial forest landscape. In *Prunus mahaleb*, landscape features markedly affected the gene flow through seed dispersal (estimated seed dispersal distance <400 m) (Godoy & Jordano 2001).

F_{st} measured historical population connectivity, and Bayesian assignment tests complemented it by revealing contemporary connectivity (over the last several generations) (Holderegger *et al.* 2010). This analysis revealed a source (f97, f23) – sink (other populations) dynamic that has rarely been identified in landscape genetic studies (Storfer *et al.* 2010). It confirmed the hypothesis that mainland EWC old forests were the main source of gene flow. This process was observed because the northwest winds are dominant during the time of pollen and seed dispersal (Denneler *et al.* 1999; Denneler *et al.* 2008). Indeed, variation in the wind direction and surface wind speed affect the genetic structure (Wright *et al.* 2008; Nathan *et al.* 2011). In addition, EWC has relatively long seed dispersal distance (Xu *et al.* submitted-b) compared with many other wind-pollinated tree species (Bacles *et al.* 2006; Kremer *et al.* 2012). This indicates that colonisation achieved through seed dispersal could affect migration rates among populations.

The level of genetic differentiation (F_{st}) among populations detected in this study was generally very low (0.0082- 0.0594), although some were significant. This low F_{st} was because extensive gene flow may not always fully prevent the divergence in tree species, which agrees with the observations reviewed by Savolainen *et al.* (2007) and Kremer *et al.* (2012). Genetic differentiation among populations occurs when populations are separated either by distance or by barriers (Balkenhol *et al.* 2009). Coniferous species do not show evidence of strong population genetic differentiation (Hamrigk & Godt 1996). The

populations in this study exhibited low levels of genetic differentiation, which agrees with previous findings in EWC (Xu *et al.* 2012). This result indicates a historical inter-stand high connectivity maintained by gene flow, which is generally a homogenizing force (Wright 1965; Frankham 1995). The Bayesian assignment analysis showed highly probable first generation immigrants, both in the islands and the small fire skips. The initial EWC populations could be established from quite a few individuals.

4.6.3 Conservation Value of Fire Residuals

Habitat patchiness has been recognized as one of the salient features of landscapes by population geneticists, ecologists and conservationists. Human activities, e.g., logging, superimpose upon natural disturbances, largely contributed to population fragmentation at the landscape scale (Addicott *et al.* 1987; Boucher *et al.* 2009; Bélisle *et al.* 2011). Forest remnants were found to function as "life boats" for organisms in the Scandinavian forest (Matveinen-Huju *et al.* 2006) and to harbor high species richness in the Canadian boreal forest (Fenton & Bergeron 2008). The high level of genetic diversity of a single species does not only manifest its own persistence and resilience to disturbance but also demonstrates its capacity in playing a key role in the habitat ecosystem that it inhabits; thus, its conservation must be an urgent priority (Reusch *et al.* 2005; Hughes *et al.* 2008; Bailey *et al.* 2009). The mainland old-growth EWC stands that escaped forest fires maintained high genetic diversity and function as transient fire-free refugia in boreal landscapes. Gene flow from mainland old-growth EWC stands ensures the maintenance of genetic variability in the small EWC fire skips. Thus, the conservation of large mainland old-growth EWC patches should be given careful attention, and they should be protected effectively.

4.7 Conclusion

Our data revealed a substantial level of gene flow in the EWC among different landscapes in boreal mixed-wood forests. The lacustrine landscape facilitated gene flow, whereas the fragmented terrestrial forest landscape increased population differentiation in EWC. Both wind and landscape features shaped the genetic structure of EWC in this landscape. Our study provided evidence that the mainland EWC old stands within a successional forest

matrix possessed a higher conservation value than the small EWC fire skips. Therefore, appropriate protection measures should be taken before relatively large old-growth forests massively turn into small patches, such as EWC fire skips, which are characterized by a reduction in allelic diversity and elevated genetic differentiation.

4.8 Acknowledgements

This work is part of the PhD thesis of H. X. We would like to acknowledge M.H. Longpré and D. Charron for logistical support; A. Leduc, G. Trudeau, R. Libert, and S. Chauchard for help during the field work; and Y. Zhou for assistance in the lab. Special thanks go to the industrial supervisors from the company Tembec: L. Dumas, S. Légaré, and G. Labrecque. This research was supported by an Industrial Innovation Doctoral Scholarship from the Natural Sciences and Engineering Research Council of Canada (NSERC), a BMP innovation doctoral scholarship from the Fonds de recherché du Québec-Natures et technologies (FRQNT) and Tembec, and a scholarship from the Forêt d'enseignement et de recherche du lac Duparquet (FERLD) to H. X. This research was also funded by a NSERC strategic grant (STPGP 336871) to F. T.

CHAPTER V

GENERAL CONCLUSION

5.1 Conclusions

The impact of population fragmentation upon genetic structure of marginal EWC was the presence of a positive inbreeding coefficient. However, increased population isolation apparently did not correlate with a detectable effect on genetic diversity. Overall, the fragmented populations of EWC appear to be well-buffered against effects of inbreeding on genetic erosion. Therefore, increased inbreeding does not lead to a loss of genetic variation in northern EWC populations, which have the potential to respond and to adapt to environmental changes. The marginal populations would be able to play a major role as seed sources in further northward range expansion under favourable climate changes.

We revealed the importance of sexual and asexual regeneration for EWC along a post-fire succession in the boreal forest. We have shown empirical evidence for extensive pollen- and seed-mediated gene flow in association with post-fire colonisation, demonstrating a considerable ability of EWC to disperse seeds over relatively long distances in a lacustrine landscape. We have revealed the effect of disturbance history on the development of genetic structure in EWC along a 250-year-long post-fire succession as well as a dynamic pattern of fine-scale genetic diversity and of structure over time since last-disturbance in association with clonal reproduction.

The contemporary levels of genetic diversity were established early in population development. The relative proportion of vegetative propagation to EWC recruitment slightly increased with time since fire. Contrasting patterns of fine-scale spatial genetic structure (SGS) were observed between saplings and adult trees. We have detected significant SGS

over time only amongst saplings. Vegetative recruitment increased SGS at fine-scales in saplings, especially in older stands in which vegetative recruitment became more common.

We have revealed extensive gene flow in EWC in landscapes comprising lacustrine landscape (islands), mainland old forests, and small fire skips in the boreal mixed-wood forest. Both wind and landscape features shaped the genetic structure of EWC in this landscape. There were significant differences in allelic diversity and in genetic differentiation among the three landscape types with small EWC fire skips having lowest allelic diversity and highest population differentiation. This effect of fragmentation on EWC genetic characteristics as detected in small fire skips indicates that dense forest matrix acts as an efficient barrier to gene flow.

We have provided evidence that the mainland EWC old stands within a successional forest matrix possess a higher conservation value than the small EWC fire skips. Therefore, one should take appropriate protection measures before relatively large old-growth forests turn into small patches, such as EWC fire skips, which a reduction in allelic diversity and an elevated genetic differentiation characterise.

5.2 Implications for Sustainable Forest Management

Our perception of the forest has changed. This ecosystem is no longer viewed only as a source of timber supply. This calls for new sustainable management practices that aim to reconcile forest exploitation with biodiversity preservation. For example, as public environmental awareness gradually increases, the FSC (Forest Council Stewardship) label has become one of the salient indicators of the forest industry's social and environmental responsibilities. The forest industry must meet FSC standards in how they implement the forest management plan to sell FSC labeled products (FSC 2004). The maintenance of species at the limits of their range, the use of natural regeneration, and the conservation of rare forest types (i.e. old-growth forest) are commonly-used approaches to maintain genetic diversity since 1999, when the FSC introduced the concept of High Conservation Value Forests (HCVFs) (FSC 2004).

We have provided genetic information on EWC populations at the limit of EWC range, as well as about old-growth EWC forests. These findings will help to direct the development of sustainable forest management strategies and the preservation of the adaptive ability of local and regional EWC populations. For instance, marginal EWC populations (MZ1 to MZ9) and old-growth mainland stands (f1797, f1823) have exhibited high levels of genetic diversity. Old-growth mainland stands also serve as the source of gene flow and they play a fundamental role in maintaining connectivity among fragmented EWC stands. We identify them (*i.e.*, MZ1 to MZ9, f1797, f1823) as High Conservation Value Forest. Therefore, these HCVFs EWC stands should be persevered to fulfil the requirements for sustainable forest management practice. The dimension of old-growth EWC forest to be preserved should be at least on a scale of tens of hectares (e.g. large old-growth patch) whenever possible.

5.3 Future Research Recommendations

5.3.1 Conducting a Range-wide Genetic survey, Testing Glacial Refugia, and Revealing Post-glacial History

Ice covered most of the North American continent during the Pleistocene. During that period most species thus retreated into small areas, then re-expanded northward during interglacial periods (Davis 1983). Boreal tree species experienced range contraction into refugia during a glacial-interglacial cycle. Then, post-glacially, they expanded into their present range (Payette *et al.* 2002; de Lafontaine *et al.* 2010). The present range of boreal tree species is the result of postglacial migration from the glacial refugia, and of species adaptations to abiotic (i.e. climate) and biotic interactions, and to historical disturbances (*e.g.*, fire) (Larsen 1980; Lomolino *et al.* 2006; Banks *et al.* 2010; Anderson *et al.* 2011).

However, only fossil and pollen records provide evidence before the emergence both of allozyme markers and of polymorphic molecular markers, even though these aforementioned records have limitations in taxonomic and morphological resolutions (Bryant Jr & Holloway 1985; Ritchie & Macdonald 1986; Cruzan & Templeton 2000). Macrofossil records and geographic distribution of genetic variation have become two complementary approaches to reconstruct species history (Davis 1983; Hewitt 1996; Jackson *et al.* 1997; Petit &

Vendramin 2005). Molecular markers thus have gained attention and have wide usage in, for example, phylogeographical studies using these markers to address questions of the localisation of glacial refugia, of postglacial migration routes reconstruction, and of colonisation dynamics in many species (Hewitt 1996; Comes & Kadereit 1998; Burg *et al.* 2005; Takahashi *et al.* 2005).

Pollen records suggest that the ice-free areas of eastern Beringia (Alaska and adjacent Canada) served as glacial refugia for boreal tree species and facilitated their post glacial re-colonisation (Brubaker *et al.* 2005; Anderson *et al.* 2006b; Zazula *et al.* 2006). Hotspots of genetic diversity have been found to be associated with glacial refugia (Taberlet *et al.* 1998; Petit *et al.* 2003; Gómez & Lunt 2007). Population genetic diversity is generally higher in eastern regions for transcontinental tree species which range from the Pacific to the Atlantic Oceans (Brant & Ortí 2003; Jaramillo-Correa *et al.* 2004; Lee-Yaw *et al.* 2008; Jaramillo-Correa *et al.* 2009). These data suggest the presence of distinct glacial and postglacial history in North America.

The EWC has a large range, and that fact provides a good basis for glacial and postglacial history studies. The large variety of habitats which the EWC occupy may indicate that its survival at glaciations' time occurred at various sites. Relevant questions regarding phylogeographic history of EWC include but are not limited to: 1) what is the historic dynamics of EWC? 2) Were there many glacial refugia? If yes, 3) are they in accordance with the clusters identified in our study? If not, 4) where are they? What are their sizes? How many of them are they? Lastly, 5) is there any historic barrier to the movement of pollen and seeds?

5.3.2 Using Organelle DNA Sequences, and Developing More Molecular Markers

With technological advances, range-wide surveys of genetic diversity have applied molecular markers in thousands of species (Petit *et al.* 1998). Conservation plans thus have taken into account into such information (Riggs 1990; Millar & Libby 1991). Nuclear DNA (ncDNA) of conifers is biparentally inherited, while the cytoplasmic DNA has contrasting modes of inheritance. Pollen and seeds disperse chloroplast DNA (cpDNA), and it is paternally

inherited; while mitochondrial DNA (mtDNA) is maternally inherited and dispersed by seeds alone (Dong & Wagner 1993; Ouborg *et al.* 1999; Provan *et al.* 2001; Lian *et al.* 2008) .

Studying both cpDNA and mtDNA can disentangle the evolutionary forces acting upon gene flow (Slatkin 1985). However, these markers are less polymorphic compared to nuclear makers. Seed-bore mtDNA may be better than other DNA markers to assess the impact of historical factors on genetic structure, due to its maternal mode of inheritance and its much shorter dispersal distances than those of airborne pollen (Gamache *et al.* 2003). The comparisons of genetic parameters deriving from cpDNA, mtDNA, and ncDNA will permit a better understanding of the relative contribution of post-glacial migration and of gene flow to current genetic structure of tree species. For instance, population structure of black spruce derived from mtDNA is distinct from the structure which derives from ncDNA (Jaramillo-Correa *et al.* 2004). However, Anderson *et al.* (2006a) found that broad-scale phylogeographic patterns of white spruce populations are consistent with cpDNA and ncDNA.

5.3.3 Scanning Genome to Detect Adaptive Genes

Adaptation has been the center of many ecological genetic studies, which provide increased evidences of global warming (Prunier *et al.* 2011). Information observed by scanning adaptive genes provides insights for forest genetic resource management, *e.g.,* improving the selection of seed source during afforestation, and tree breeding for wood traits (Aitken *et al.* 2008). Understanding the genetic basis of adaptation is also important to evolutionary biology (Prunier *et al.* 2011), since tree species tend to reveal phenotypic adaptations to environmental changes (Morgenstern 1996).

Some studies have established the relationship between climate-related traits and functional genes in forest trees (Rehfeldt 1989; St Clair *et al.* 2005; St. Clair 2006). Many such studies in tree species have identified the genes which pertain to adaption , for example, those related to cold-hardiness in *Pseudotsuga menziesii* var. *menziesii* (Eckert *et al.* 2009) and *Picea sitchensis* (Holliday *et al.* 2010), to radial growth in *Picea glauca* (Beaulieu *et al.* 2011), and to bud set in *Picea* (Kayal *et al.* 2011).

Next-generation sequencing has greatly reduced the cost and time of developing new molecular markers, even for non-model organisms (Selkoe & Toonen 2006; Frankham 2010; Xu *et al.* 2013), and it has accelerated the transition from study on neutral genetic diversity to such focus on adaptive genes (Gilad *et al.* 2009). The application of this technology to the EWC would reveal interesting stories. For example, detecting genes that are in correlation with environmental and ecological parameters will greatly assist us in accounting for the EWC's potential for adaptation.

APPENDIX I

DEVELOPMENT AND MULTIPLEXED AMPLIFICATION OF SSR MARKERS FOR *THUJA OCCIDENTALIS* (CUPRESSACEAE) USING SHOTGUN PYROSEQUENCING

Article published in 2013

In *Applications in Plant Sciences*, 1(5), 1200427

(Formerly *American Journal of Botany - Primer Notes and Protocols in the Plant Sciences*)

A1.2 Abstract

• *Premise of the study:* Sixteen novel, polymorphic, multiplexed microsatellite loci were developed for eastern white cedar (*Thuja occidentalis*) using simple sequence repeat (SSR)-enriched shotgun pyrosequencing.

• *Methods and Results:* Sixteen loci were tested on a panel of 24 individuals from different populations. The number of observed alleles ranged from four to 22. Four sets of multiplex PCR for the 16 loci were then carried out on 60 individuals of two populations from islands of FERLD Duparquet Forest, Canada. Mean number of alleles, observed heterozygosity and expected heterozygosity were respectively 5.75, 0.594, 0.574 for Island 58, and 5.50, 0.704, and 0.624 for Island 134.

• *Conclusions:* Four sets of multiplex microsatellite loci can be used for future genetic studies, which includes investigating genetic diversity and structure, and fragmentation and regeneration studies.

A1.3 Introduction

Eastern white cedar (*Thuja occidentalis*) is a native, wind-pollinated conifer with a broad distribution across North America (Fowells 1965). The species' range extends from Gulf of St. Lawrence in the East to southeastern Manitoba in the West, and from James Bay in the North to Tennessee and North Carolina in the South (Fowells 1965). A member of the Cupressaceae, it is also commonly called eastern arborvitae, American arborvitae, northern white cedar, Atlantic red cedar, and swamp cedar in English (USDA 2013), and thuya occidental, cèdre, balai, cèdre blanc, thuier cèdre, and arborvitae in French (Brouillet *et al.* 2010). Eastern white cedar (EWC) is listed as endangered in Indiana, Massachusetts, and New Jersey, as a threatened species in Connecticut, Illinois, Kentucky, and Maryland, and of special concern in Tennessee (USDA 2013). Genetic analyses previously conducted on EWC have been mainly based on allozyme markers (Hofmeyer *et al.* 2007), while highly polymorphic markers such as microsatellites have not been developed for EWC. We report on the development and characterization of microsatellite markers for EWC using shotgun pyrosequencing on a simple sequence repeat (SSR)-enriched library (Malausa *et al.* 2011).

A1.4 Methods and Results

Foliage of EWC individuals from 14 sites across northern Quebec (Table SA1.1) was collected and maintained at -20°C before genetic analysis. Genomic DNA was extracted using the DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany). DNA extracts of 14 individuals were combined and sent on dry ice to Genoscreen (Lille, France) for microsatellite enriched GS-FLX library construction following the methodology developed by Malausa *et al.* (2011). Briefly, main steps included: 1) digestion of genomic DNA with *Rsa*I (Fermentas, International Inc., Burlington, Ontario, Canada); 2) enrichment of microsatellite sequences in fragmented DNA with eight types of probes (TG, TC, AAC, AAG, AGG, ACG, ACAT, ACTC), which was accomplished by using Dynabeads (Invitrogen, Carlsbad, California, USA); and 3) PCR amplification of enriched DNA with primers specific to adaptor sequences (Malausa *et al.* 2011). In total, 11,393 raw sequences with average read length of 400 bp were obtained, with 2175 sequences containing microsatellite motifs. One hundred seventeen of the sequences successfully had primers

designed for them using QDD software (Meglécz *et al.* 2010) with default parameters except optimal primer length of 22bp (range 18-27bp) and 50% GC content (range 40%-60%).

To minimize screening costs, we initially selected 48 of 117 pairs following criteria detailed in Lepais and Bacles (2011). In brief, we restricted our selection to loci with hexa-, tetra-, and dinucleotide motif types. Dinucleotide motif was limited to AC, CA, TG, GT, AG, GA, CT, and TC types, because AT and TA types are notoriously hard to amplify (Temnykh *et al.* 2001). Each of the selected 48 loci was initially tested for amplification with unlabeled primers (Invitrogen) on a screening panel that included seven EWC trees collected across northern Quebec (one tree per site) (Table SA1.1) and a negative control. Amplifications were carried out in a total volume of 10 µL using four 96-Well Mastercycler Pro S PCR Systems (Eppendorf Gmbh, Wesselling-Berzdorf, Germany). Each reaction mixture contained 1 µl of DNA extract, 5 µl of 2X QIAGEN Multiplex PCR Master Mix (QIAGEN), and a final concentration of 0.2 µM for each forward and reverse primer. The PCR program consisted of an initial heat-activation step at 95°C for 15 min, 36 cycles of 3-step cycling: denaturation at 94°C for 30 s, annealing at 54°C for 90 s, extension at 72°C for 60 s, and a final extension at 60°C for 30 min. A total of 2.5 µl of PCR products were visualized on 3% agarose gel (Promega Corporation, Madison, Wisconsin, USA), with electrophoretic migration performed at 100V for 20 min on the Bio-Rad Imaging System (Bio-Rad, Montreal, Canada).

The initial test on 48 loci using agarose gel showed that 38 had amplified products, and four of 38 had three or more products in one PCR reaction (non-specific). Among the remaining 34 loci, 16 had one or two amplification products, variable in size. Thus, we used all of them to verify polymorphisms and further design Multiplex PCR. Each was tested with fluorescent dye-labeled primers (Applied Biosystems, Carlsbad, California, USA) on a panel that included 24 EWC individuals collected from 24 sites across northern Quebec (one individual per site) (Table SA1.1) plus one negative control. PCR cycles were the same as those mentioned previously, except for increased annealing temperatures to achieve specific amplifications (Table A1.1). A total of 2 µl of 1:100 diluted PCR products labeled with four different dyes (6-FAM, VIC, NED, and PET; Applied Biosystems) were mixed with 8.35 µl of HI_DI formamide (Applied Biosystems) and 0.15 µl of GeneScan™ 500 LIZ™ Size

Standard (Applied Biosystems), and sent to GenoQuebec (Montreal, Canada) for genotype reading on an ABI 3730 genetic analyzer. Results were analyzed with GeneMapper 3.7 software (Applied Biosystems).

All 16 loci showed interpretable, repeatable, and polymorphic patterns (Table A1.1). We used Multiplex Manager (Holleley & Geerts 2009) to design and optimize multiplex PCRs to find annealing temperatures for each multiplex group to ensure specific amplifications and avoid complementary sequences among primers. Primer pairs were multiplexed to reduce amplification costs (Table A1.1). Co-amplifications of all multiplexed primers were tested on two populations (30 trees per population) sampled from islands in Lake Duparquet, northwestern Quebec (Table A1.2). PCR cycles were the same as those mentioned previously, except for multiplexed annealing temperatures (M1 at 57 °C; M2 at 56°C; M3 at 55°C; M4 at 54°C). PCR products were genotyped as previously detailed.

Table A1.1 Characteristics of 16 microsatellites validated for *Thuja occidentalis*

Locus	Primer Sequence (5′ → 3′)	GenBank ID	Dye	Repeat Motif	Size (bp)[a]	T_a (°C)	MG	Observed		
								A	Min	Max
TO791	F: AAGAGATTTATTTGCCCTCCG	JX475983	VIC	$(ca)_{12}$	141	57	1	16	133	167
	R: ATGGTTGATGGACTCCTTGG									
TO605	F: GAATAACTTCTTCTGGGAAAGATACA	JX475984	PET	$(ac)_8$	190	59	1	10	174	196
	R: GAGGTGGAAAGAAGTGGATAAAA									
TO328	F: CCCGCAACACCTACTTGTCT	JX475985	FAM	$(taca)_7$	215	57	1	4	203	215
	R: TGCTCCATGTTTGAAGTTGC									
TO53	F: AAATGGCCCATAAGCACAAA	JX475986	NED	$(ca)_5$	184	58	1	6	174	184
	R: GGATGTTTCCAGTTGACGGT									
TO925	F: TGTGTTTGTGGTGGCTGACT	JX475987	FAM	$(tg)_{20}$	151	58	2	22	129	217
	R: CATTCATACATTTCCCATCCA									
TO727	F: GAGATTCCTTTAAAATATTGGCAT	JX475988	VIC	$(ga)_{11}$	241	57	2	14	233	325
	R: CCCTCCCATTCCTCTTAATG									
TO659	F: TGATGCACCAATTTTCTTTGG	JX475989	PET	$(ct)_9$	191	56	2	7	181	195
	R: TGATGCACTTTAAGGTGTAGGG									
TO29	F: TGCAGTGTTAGTGGAGCAACTT	JX475990	NED	$(ca)_5$	162	57	2	14	148	186
	R: TCATTGTTTATTCCCTAAGATGGA									
TO737	F: GAGCAAGAAGGAGAGTGGGA	JX475991	PET	$(agat)_{11}$	124	63	3	6	102	130
	R: CCTAGGTTGCCTTGTTGTCC									
TO587	F: GTGCCAAACTTTTCAAGGTAAGA	JX475992	NED	$(ct)_8$	167	62	3	13	139	211
	R: GCAAGAGCACAAATGATCACA									
TO512	F: TGCATAACAACTCTTCTTAAATCAGC	JX475993	FAM	$(ct)_8$	194	63	3	11	146	212
	R: AGGTCCTATCTAGGTCTTAGACAACTT									
TO503	F: CTTGTCCGTCTGACATGTGTTT	JX475994	VIC	$(ga)_8$	190	55	3	12	138	202
	R: CACATAGGTTAAGGGTAGTTTCCT									
TO715	F: CATCTACATGGTCGATGATTTAAC	JX475995	VIC	$(ag)_{10}$	106	60	4	6	100	110
	R: TATCCCAAACCAGCAAAACC									
TO521	F: CAAATATGGCACCAATGCCT	JX475996	PET	$(ct)_8$	121	54	4	17	113	239
	R: CAATTTCCTCAGGTTTGGGA									
TO418	F: ATGCTTTTCTAACCCTTTTGGA	JX475997	NED	$(ac)_7$	253	61	4	8	163	255
	R: TGATCAGTTGGATTTCTAGATTGC									
TO20	F: TTTGGCTTGTAGGTGGTTTT	JX475998	FAM	$(tg)_5$	192	57	4	15	168	204
	R: CTCCATTTTGGAGTGTTGGT									

Note: A = number of alleles observed; Max = maximum allele size observed during screening; MG = multiplex group; Min = minimum allele size observed during screening; T_a = annealing temperature.
[a] Product size from shotgun pyrosequencing.

The number of different alleles per locus (A), observed heterozygosity (H_o), and expected heterozygosity (H_e) were calculated in GenAlex6.2 (Peakall & Smouse 2006). Inbreeding coefficient (F_{is}) and Hardy-Weinberg equilibrium (HWE) tests were done in FSTAT2.9.3 (Goudet 2001). Null allele presence was checked in Micro-Checker (Van Oosterhout *et al.* 2004). Mean values for A, H_o, and H_e were respectively 5.75, 0.594, and 0.574 on Island 58, and 5.50, 0.704, and 0.624 on Island 134 (Table A1.2). F_{is} ranged from -0.706 to 0.665 on Island 58, and from -0.357 to 0.194 on Island 134 (Table A1.2).

A1.5 Conclusions

Shotgun pyrosequencing has proved to be effective for isolating microsatellite markers in EWC. The four sets of multiplex microsatellite loci that were developed here for the first time will facilitate future studies of population genetics in EWC, including investigating phylogeographic patterns of post-glacial expansion in North America, and studying the impacts of habitat fragmentation on population genetic structure and gene flow. They will also help resolve questions regarding regeneration patterns in this species along post-fire successions (Bergeron 2000).

Table A1.2 Results of initial primer screening in *Thuja occidentalis* samples from Lake Duparquet, Lake Duparquet Research & Teaching Forest, Quebec, Canada

	Island 58 (N=30)[a]					Island 134 (N=30)[a]				
Locus	*A*	H_o	H_e	F_{is}	Null	*A*	H_o	H_e	F_{is}	Null
TO53	5.00	0.567	0.705	0.212	no	4.00	0.900	0.652	-0.366	no
TO328	4.00	0.567	0.585	0.048	no	3.00	0.667	0.491	-0.343	no
TO605	3.00	0.200	0.580	0.665*	yes(0.24)	2.00	0.367	0.433	0.169	no
TO791	12.00	0.667	0.794	0.177	no	11.00	0.800	0.839	0.063	no
TO29	7.00	0.500	0.543	0.096	no	7.00	0.700	0.712	0.034	no
TO659	4.00	0.400	0.581	0.326	yes(0.11)	6.00	0.500	0.608	0.194	no
TO727	2.00	0.833	0.486	-0.706*	no	5.00	0.900	0.686	-0.296	no
TO925	18.00	0.800	0.847	0.072	no	10.00	0.667	0.770	0.151	no
TO503	6.00	0.967	0.644	-0.487*	no	4.00	0.633	0.521	-0.199	no
TO512	5.00	0.367	0.322	-0.121	no	7.00	0.733	0.556	-0.305	no
TO587	5.00	1.000	0.712	-0.391*	no	7.00	0.867	0.710	-0.204	no
TO737	5.00	0.733	0.634	-0.139	no	4.00	0.867	0.624	-0.375	no
TO20	2.00	0.400	0.320	-0.234	no	4.00	0.933	0.676	-0.366	no
TO418	3.00	0.067	0.065	-0.009	no	3.00	0.433	0.443	0.038	no
TO521	8.00	0.933	0.777	-0.185	no	8.00	0.800	0.738	-0.067	no
TO715	3.00	0.500	0.583	0.159	no	3.00	0.500	0.529	0.072	no
Mean	5.75	0.594	0.574	-	-	5.50	0.704	0.624	-	-
SE	1.031	0.068	0.050	-	-	0.658	0.045	0.030	-	-

Note: *A* = number of alleles; F_{IS} = inbreeding coefficient; H_e = expected heterozygosity; H_o = observed heterozygosity; Null, null alleles present (frequency). * $P \leq 5\%$; Bonferroni correction was applied, and indicative adjusted *P* value for 5% nominal level was 0.0031.[a] Geographical coordinates: Island 58 (48°26'41.4"N, 79°15'51.9"W), Island 134 (48°27'52.5"N,79°16'19.6"W).

APPENDIX II

SUPPORTING INFORMATION

A 2.1 Supporting information for chapter II

Table S2.1 Primer characteristics, frequencies of null alleles (*r*), and *F*-statistics over all populations at each locus. F_{st} was estimated by harbouring and excluding the null allele (INA).

Locus	Primer sequences (5'–3')	Repeats of cloned allele	Size range (bp)	Allele (n)	*r*	F_{st}	F_{st} INA	F_{is}
TP9	TCTCCTTGTCTTGGATTTGG CGGAAAGTAGTCTCATTATCAC	(AC)> 20	123-433	12	0.131	0.124	0.109	0.146
TP10	TAGTTGTGTCCATTCAGGCAT GCTCTTATCTTCTTTTAGGGC	(GT)4GC (GT)12	136-156	10	0	0.031	0.038	-0.211
TP11	GCTCTTATCTTCTTTTAGGGC CCTGATCCGCTTTGATGGGT	(CT)12 (CA)16	140-256	15	0.131	0.061	0.064	0.22
TP12	GATAAGAGGCATCACTCGAG CCGGATCATTAAGGGCTCTA	(CA)29	130-258	17	0.104	0.083	0.083	0.143
All	-	-	-	-	-	0.076	0.073	-

Table S2.2 Allele frequencies and the number of alleles (N) by group

Locus	Allele	Marginal	Discontinuous	Continuous
TP9	N	9	11	12
	123	-	0.023	0.01
	139	0.081	0.099	0.066
	141	-	0.06	0.02
	143	0.073	0.054	0.038
	147	-	0.02	0.003
	159	0.048	0.014	0.01
	211	0.132	0.045	0.01
	215	0.126	-	0.013
	257	0.135	0.21	0.25
	259	0.039	0.014	0.073
	431	0.174	0.213	0.303
	433	0.191	0.247	0.205
TP10	N	9	8	10
	136	0.051	0.043	0.053
	138	0.056	0.031	0.033
	140	0.056	-	0.008
	142	0.062	0.023	0.003
	144	0.149	0.205	0.285
	146	0.202	0.284	0.152
	148	0.292	0.301	0.348
	150	0.126	0.085	0.101
	152	0.006	0.028	0.01
	156	-	-	0.008
TP11	N	14	13	15
	140	0.098	0.188	0.028
	142	0.025	0.017	0.013
	158	0.045	0.04	0.008
	160	0.062	-	0.005
	166	0.157	0.122	0.258
	180	0.059	0.034	0.109
	182	0.034	0.006	0.071
	200	0.065	0.009	0.043
	202	0.053	0.048	0.04

	212	0.09	0.139	0.093
	214	0.045	0.031	0.015
	220	-	0.003	0.035
	222	0.183	0.233	0.255
	224	0.039	-	0.008
	256	0.045	0.131	0.02
TP12	N	14	15	17
	130	0.045	0.043	0.015
	138	0.081	0.023	0.03
	140	0.143	0.119	0.086
	142	0.216	0.25	0.096
	144	0.174	0.102	0.061
	146	0.025	0.006	0.028
	148	0.062	0.057	0.063
	150	0.056	0.037	0.03
	156	0.039	0.102	0.061
	158	0.062	0.02	0.018
	160	-	0.009	0.015
	202	0.037	0.014	0.018
	214	0.048	0.102	0.005
	240	-	0.037	0.146
	242	-	-	0.058
	256	0.006	0.08	0.184
	258	0.006	-	0.086
N (total allele)		46	47	54
The Rate of Rare Alleles (Frequency < 1%)		0.065	0.106	0.148

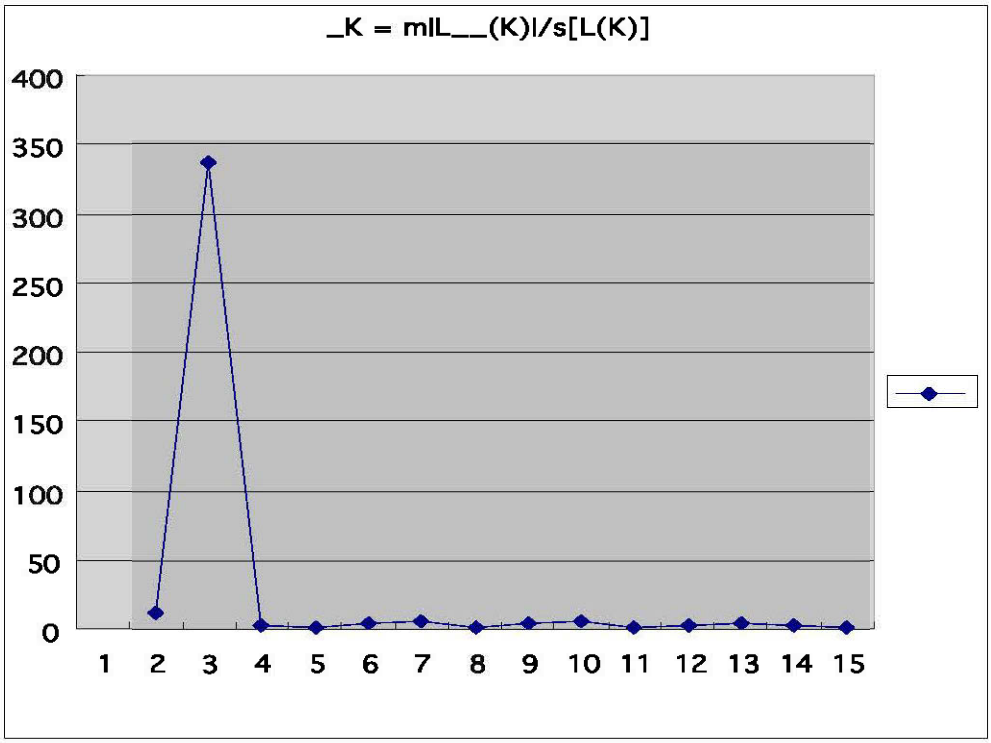


Figure S2.1 Detection of the number of clusters, K, using STRUCTURE for eastern white cedar (*Thuja occidentalis* L.) populations according to Evanno et al. (2005)

A 2.2 Supporting information for chapter III

Table S3.1 Average diameter at forest floor and height of saplings of *Thuja occidentalis* sampled from 10 regeneration patches in 10 subplots (10x10 m) in three 1-ha sampling plots

Plot		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Mean
1916	N_s	9	6	7	10	8	10	10	10	9	8	8.7
	D	1.32	0.50	0.90	1.27	0.77	0.77	1.22	0.88	0.94	1.44	1.00
		(0.61)	(0.30)	(0.17)	(0.52)	(0.28)	(0.20)	(0.66)	(0.32)	(0.55)	(0.44)	(0.30)
	H	81.78	33.50	67.14	79.70	43.88	49.30	72.10	50.60	63.89	87.88	62.97
		(25.07)	(27.79)	(14.72)	(31.23)	(10.63)	(11.07)	(36.28)	(12.93)	(25.98)	(29.42)	(18.04)
1823	N_s	10	10	10	10	10	10	10	10	10	10	10
	D	1.42	0.92	0.46	1.66	0.93	0.46	0.74	0.83	1.22	0.59	0.92
		(0.58)	(0.25)	(0.20)	(0.76)	(0.33)	(0.16)	(0.23)	(0.20)	(.0.41)	(0.23)	(0.40)
	H	94.40	62.60	42.60	78.10	68.50	36.00	63.50	69.60	10.98	48.90	57.52
		(26.18)	(14.11)	(19.27)	(27.61)	(23.10)	(10.47)	(16.33)	(18.16)	(35.63)	(13.76)	(23.66)
1760	N_s	10	9	10	10	10	8	8	10	9	10	9.4
	D	1.14	0.81	0.60	1.01	0.43	1.34	0.69	0.92	0.89	1.55	0.94
		(0.44)	(0.27)	(0.14)	(0.68)	(0.20)	(0.74)	(0.40)	(0.32)	(0.42)	(0.63)	(0.34)
	H	64.40	60.78	44.20	64.60	40.20	87.88	55.75	71.11	66.67	10.28	56.59
		(16.95)	(12.11)	(8.01)	(37.54)	(13.24)	(33.68)	(13.69)	(22.66)	(21.32)	(35.97)	(21.09)

N_s, Number of saplings sampled in each patch; D, Diameter (cm) at forest floor, mean (SD); H, Height (cm), mean (SD); P1, the patch sampled in subplot 1; P2, the patch sampled in subplot 2; and so forth.

Table S3.2 The list of 10 distance intervals (m) used for SGS analysis in each plot of *Thuja occidentalis*

Distance classes		1	2	3	4	5	6	7	8	9	10
Saplings (clones included)	1916	2.52	15.14	27.25	30.04	36.93	57.29	77.48	89.75	100.68	122.67
	1823	20.22	28.34	34.84	44.78	52.53	58.00	64.58	82.85	87.34	130.75
	1760	15.87	25.45	39.22	50.73	58.41	61.75	72.46	80.68	91.05	128.04
Saplings (clones excluded)	1916	2.30	12.79	20.80	29.83	36.45	67.98	77.99	90.30	98.47	122.67
	1823	20.68	28.52	34.97	45.18	53.14	58.50	65.12	83.70	86.28	130.75
	1760	16.11	26.07	39.66	51.45	58.40	62.09	72.04	80.83	91.14	127.88
Adult trees	1823	15.90	24.13	31.38	38.26	44.63	51.31	58.42	66.36	77.66	113.79
	1760	18.52	27.50	35.09	42.41	49.51	56.53	63.79	71.55	82.86	125.65

Table S3.3 Estimation of global F_{st} both using and without using the ENA correction

Locus	F_{st} not using ENA	F_{st} using ENA
TO53	0.052400	0.051407
TO328	0.081412	0.071728
TO605	0.073067	0.089496
TO791	0.045815	0.041879
TO29	0.045732	0.038129
TO659	0.021037	0.016004
TO727	0.074908	0.069833
TO925	0.078753	0.063574
TO503	0.129544	0.141431
TO512	0.031348	0.030279
TO587	0.048548	0.034095
TO737	0.020711	0.020645
TO20	0.122001	0.101828
TO418	0.115272	0.115493
TO521	0.027111	0.028057
TO715	0.025395	0.028130
All loci	0.061975	0.057774

ENA correction described in Chapuis and Estoup (2007)

A 2.3 Supporting information for chapter IV

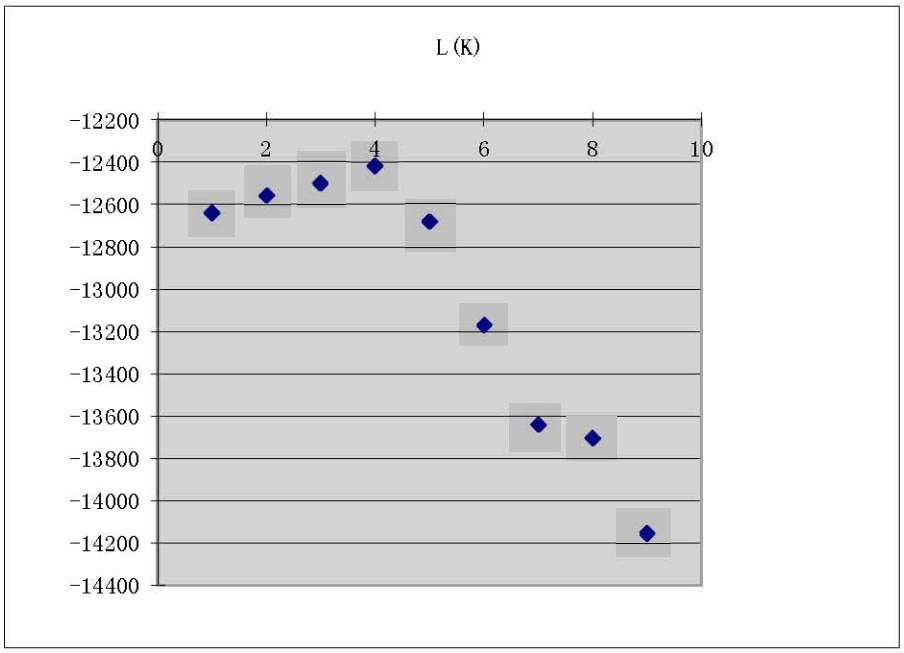


Figure S4.1 Detection of the number of clusters (K) for the STRUCTURE for Eastern white cedar (*Thuja occidentalis* L.) populations by plotting L (K) according to Evanno et al. (2005).

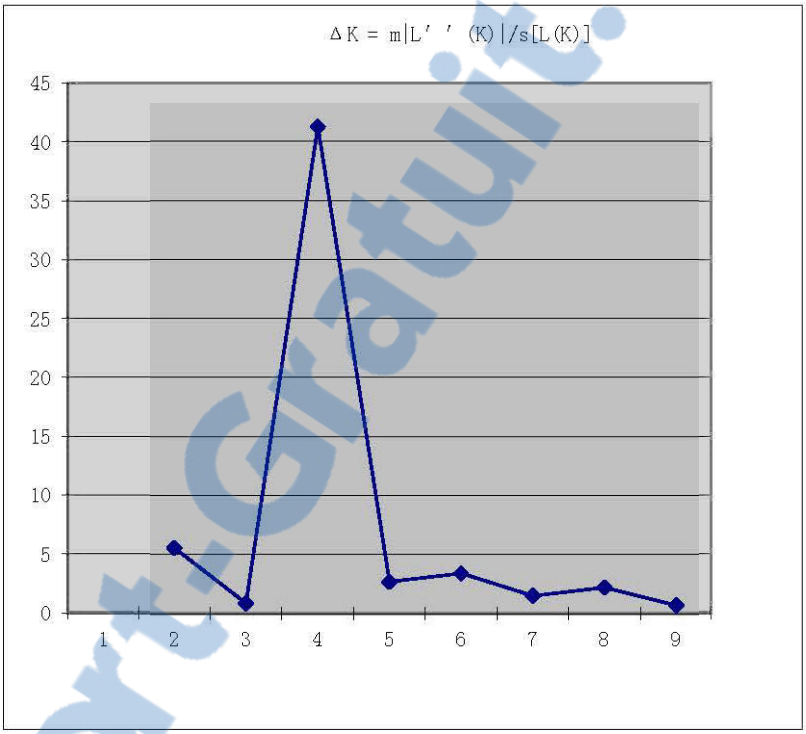


Figure S4.2 Detection of the number of clusters (K) for the STRUCTURE for Eastern white cedar (*Thuja occidentalis* L.) populations by plotting ΔK according to Evanno et al. (2005).

Table S4.1 Recent migration analysis implemented in BayesAss 1.3 (95% Confidence Interval) for nine Eastern white cedar stands form three landscape types in Quebec, Canada

	f441	f442	f443	is39	is42	is134	f97	f23	f60
From f441 into	0.6787 (0.6670, 0.7100)	0.0059 (4.9e-07, 0.0399)	0.0038 (4.7e-07, 0.0216)	0.0056 (1.5e-05, 0.0282)	0.0036 (2.1e-07, 0.0224)	0.0038 (3.5e-07, 0.0224)	0.0014 (2.6e-12, 0.0114)	0.0043 (6.0e-07, 0.0234)	0.0041 (2.6e-07, 0.0236)
From f442 into	0.0055 (8.6e-07, 0.0281)	0.6814 (0.6670, 0.7222)	0.0039 (6.0e-07, 0.0232)	0.0054 (2.1e-05, 0.0292)	0.0042 (2.5e-07, 0.0238)	0.0039 (1.1e-06, 0.0238)	0.0017 (2.3e-12, 0.0144)	0.0045 (3.8e-07, 0.0257)	0.0041 (9.2e-08, 0.0228)
From f443 into	0.0054 (3.7e-06, 0.0280)	0.0062 (6.3e-07, 0.0334)	0.6766 (0.6669, 0.7037)	0.0055 (1.7e-05, 0.0306)	0.0040 (3.2e-07, 0.0254)	0.0038 (3.1e-07, 0.0202)	0.0013 (2.9e-13, 0.0114)	0.0037 (1.7e-07, 0.0215)	0.0038 (9.2e-08, 0.0209)
From is39 into	0.0050 (7.6e-06, 0.0252)	0.0059 (7.9e-07, 0.0354)	0.0037 (5.6e-07, 0.0230)	0.6769 (0.6669, 0.7036)	0.0037 (2.0e-07, 0.0244)	0.0041 (7.3e-08, 0.0240)	0.0017 (2.6e-14, 0.0140)	0.0037 (1.3e-07, 0.0220)	0.0041 (2.7e-07, 0.0229)
From is42 into	0.0050 (1.9e-06, 0.0299)	0.0057 (4.3e-07, 0.0363)	0.0038 (1.4e-07, 0.0229)	0.0056 (2.1e-05, 0.0280)	0.6773 (0.6670, 0.7026)	0.0039 (4.4e-08, 0.0229)	0.0015 (1.8e-13, 0.0127)	0.0037 (2.7e-07, 0.0214)	0.0041 (2.4e-07, 0.0251
From is134 into	0.0052 (2.0e-06, 0.0311)	0.0055 (2.9e-07, 0.0357)	0.0034 (2.8e-07, 0.0195)	0.0052 (2.2e-05, 0.0250)	0.0040 (1.3e-07, 0.0228)	0.6775 (0.6670, 0.7058)	0.0013 (2.7e-13, 0.0116)	0.0037 (7.9e-08, 0.0227)	0.0041 (1.5e-07, 0.0232)
From f97 into	0.0641 (0.0150, 0.1293)	0.0116 (9.8e-07, 0.0649)	0.2730 (0.2055, 0.3194)	0.1357 (0.0459, 0.2197)	0.2909 (0.2397, 0.3264)	0.2914 (0.2412, 0.3248)	0.9876 (0.9559, 0.9995)	0.2895 (0.2370, 0.3245)	0.2943 (0.2442, 0.3269)
From f23 into	0.2254 (0.1499, 0.2944)	0.2712 (0.1959, 0.3244)	0.0280 (0.0005, 0.0851)	0.1548 (0.0760, 0.2434)	0.0083 (3.9e-07, 0.0430)	0.0080 (1.5e-06, 0.0401)	0.0019 (1.2e-12, 0.0169)	0.6828 (0.6671, 0.7274)	0.0046 (1.6e-07, 0.0276)
From f60 into	0.0056 (2.9e-06, 0.0285)	0.0065 (8.3e-07, 0.0370)	0.0039 (5.8e-07, 0.0225)	0.0053 (1.6e-05, 0.0237)	0.0038 (4.2e-07, 0.0233)	0.0037 (4.0e-07, 0.0207)	0.0015 (5.5e-13, 0.0128)	0.0040 (1.7e-07, 0.0230)	0.6767 (0.6670, 0.7023)

A 2.4 Supporting information for Appendix I

Table SA1.1 Voucher information for *Thuja occidentalis* samples. All samples were preserved at Institut de recherche sur les forêts, Université du Québec en Abitibi-Témiscamingue, Canada.

No.	Site	Latitude	Longitude	Location	Country	Year of collection
1	MZ1	49°52'31.44"N	74°23'34.224"W	Chibougamau	Canada	2007
2	MZ2	49°54'32.976"N	74°19'21.396"W	Chibougamau	Canada	2007
3	MZ3	49°57'12.636"N	74°13'44.688"W	Chibougamau	Canada	2007
4	MZ4	49°38'30.336"N	74°20'2.58"W	Chibougamau	Canada	2007
5	MZ5	48°55'39.792"N	78°53'8.808"W	James Bay	Canada	2007
6	MZ6	49°25'23.412"N	79°12'39.492"W	James Bay	Canada	2007
7	MZ7	49°51'30.708"N	78°36'25.956"W	James Bay	Canada	2007
8	MZ8	49°53'0.564"N	78°38'45.78"W	James Bay	Canada	2007
9	MZ9	49°51'21.924"N	78°38'41.496"W	James Bay	Canada	2007
10	DZ1	48°32'24.72"N	78°38'30.696"W	Abitibi	Canada	2007
11	DZ2	48°28'12.54"N	79°27'8.46"W	Abitibi	Canada	2007
12	DZ3	48°28'47.244"N	79°26'12.624"W	Abitibi	Canada	2007
13	DZ4	48°25'53.796"N	79°24'6.588"W	Abitibi	Canada	2007
14	DZ5	48°15'46.656"N	78°34'29.208"W	Abitibi	Canada	2007
15	DZ6	48°25'51.636"N	79°23'2.976"W	Abitibi	Canada	2007
16	DZ7	48°12'4.752"N	79°25'8.796"W	Abitibi	Canada	2007
17	CZ1	47°25'45.192"N	78°40'42.528"W	Témiscamingue	Canada	2007
18	CZ2	47°25'0.084"N	78°40'55.704"W	Témiscamingue	Canada	2007
19	CZ3	47°23'44.052"N	78°43'53.904"W	Témiscamingue	Canada	2007
20	CZ4	47°20'42.18"N	79°23'33.396"W	Témiscamingue	Canada	2007
21	CZ5	47°18'39.96"N	78°30'55.8"W	Témiscamingue	Canada	2007
22	CZ6	47°27'14.22"N	78°35'15.54"W	Témiscamingue	Canada	2007
23	CZ7	47°25'8.184"N	78°40'42.384"W	Témiscamingue	Canada	2007
24	CZ8	47°24'56.844"N	78°42'41.94"W	Témiscamingue	Canada	2007
25	IS58	48°26'41.4"N	79°15'51.9"W	Abitibi	Canada	2008
26	IS134	48°27'52.5"N	79°16'19.6"W	Abitibi	Canada	2008

REFERENCE:

Addicott JF, Aho JM, Antolin MF, Padilla DK, Richardson JS, Soluk DA (1987) Ecological neighborhoods: scaling environmental patterns. *Oikos,* **49**, 340-346.

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**, 5177-5188.

Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S (2008) Adaptation, migration or extirpation: Climate change outcomes for tree populations. *Evolutionary Applications,* **1**, 95-111.

Aizawa M, Yoshimaru H, Saito H *et al.* (2009) Range-wide genetic structure in a north-east Asian spruce (*Picea jezoensis*) determined using nuclear microsatellite markers. *Journal of Biogeography,* **36**, 996-1007.

Albani M, Andison DW, Kimmins JP (2005) Boreal mixedwood species composition in relationship to topography and white spruce seed dispersal constraint. *Forest Ecology and Management,* **209**, 167-180.

Ali AA, Carcaillet C, Bergeron Y (2009) Long-term fire frequency variability in the eastern Canadian boreal forest: The influences of climate vs. local factors. *Global Change Biology,* **15**, 1230-1241.

Anderson LL, Hu FS, Nelson DM, Petit RJ, Paige KN (2006a) Ice-age endurance: DNA evidence for a white spruce refugium in Alaska. *Proc Natl Acad Sci,* **103**, 2447-12450.

Anderson LL, Hu FS, Nelson DM, Petit RJ, Paige KN (2006b) Ice-age endurance: DNA evidence of a white spruce refugium in Alaska. *Proceedings of the National Academy of Sciences of the United States of America,* **103**, 12447-12450.

Anderson LL, Hu FS, Paige KN (2011) Phylogeographic history of white spruce during the last glacial maximum: Uncovering cryptic refugia. *Journal of Heredity,* **102**, 207-216.

Antolin MF, Savage LT, Eisen RJ (2006) Landscape features influence genetic structure of black-tailed prairie dogs (*Cynomys ludovicianus*). *Landscape Ecology,* **21**, 867-875.

Archambault S, Bergeron Y (1992) An 802-year tree-ring chronology from the Quebec boreal forest. *Canadian Journal of Forest Research,* **22**, 674-682.

Arnaud-Haond S, Belkhir K (2007) GENCLONE: A computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. *Molecular Ecology Notes,* **7**, 15-17.

Ashley MV (2010) Plant parentage, pollination, and dispersal: How DNA microsatellites have altered the landscape. *Critical Reviews in Plant Sciences,* **29**, 148-161.

Asselin H, Fortin MJ, Bergeron Y (2001) Spatial distribution of late-successional coniferous species regeneration following disturbance in southwestern Québec boreal forest. *Forest Ecology and Management,* **140**, 29-37.

Bacles CFE, Lowe AJ, Ennos RA (2006) Effective seed dispersal across a fragmented landscape. *Science,* **311**, 628.

Bailey JK, Schweitzer JA, Úbeda F *et al.* (2009) From genes to ecosystems: A synthesis of the effects of plant genetic factors across levels of organization. *Philosophical Transactions of the Royal Society B: Biological Sciences,* **364,** 1607-1616.

Bakuzis EV, Hansen HL (1959) A provisional assessment of species synecological requirements in Minnesota forests. *Minnesota Forestry Notes,* **84,** 1-2.

Balkenhol N, Gugerli F, Cushman SA *et al.* (2009) Identifying future research needs in landscape genetics: Where to from here? *Landscape Ecology,* **24,** 455-463.

Balloux F, Lehmann L, De Meeûs T (2003) The population genetics of clonal and partially clonal diploids. *Genetics,* **164,** 1635-1644.

Banks SC, Ling SD, Johnson CR, Piggott MP, Williamson JE, Beheregaray LB (2010) Genetic structure of a recent climate change-driven range extension. *Molecular Ecology,* **19,** 2011-2024.

Bartlett RM, Reader RJ, Larson DW (1991) Multiple controls of cliff-edge distribution patterns of Thuja occidentalis and Acer saccharum at the stage of seedling recruitment. *Journal of Ecology,* 79, 183-197.

Beaulieu J, Doerksen T, Boyle B *et al.* (2011) Association genetics of wood physical traits in the conifer white spruce and relationships with gene expression. *Genetics,* **188,** 197-214.

Beaulieu J, Simon JP (1995) Mating system in natural populations of eastern white pine in Quebec. *Canadian Journal of Forest Research,* **25,** 1697-1703.

Beaumont MA (1999) Detecting population expansion and decline using microsatellites. *Genetics,* **153,** 2013-2029.

Bélisle AC, Gauthier S, Cyr D, Bergeron Y, Morin H (2011) Fire regime and old-growth boreal forests in central quebec, Canada: An ecosystem management perspective. *Silva Fennica,* **45,** 889-908.

Bergeron Y (1991) The influence of island and mainland lakeshore landscapes on boreal forest fire regimes. *Ecology,* **72,** 1980-1992.

Bergeron Y (1998) Consequences of climate changes on fire frequency and forest composition in the southwestern boreal forest of Quebec. *Les consequences des changements climatiques sur la frequence des feux et la composition forestiere au sud-ouest de la foret boreale quebecoise,* **52,** 167-173.

Bergeron Y (2000) Species and stand dynamics in the mixed woods of Quebec's southern boreal forest. *Ecology,* **81,** 1500-1516.

Bergeron Y, Bouchard A (1984) Use of ecological groups in analysis and classification of plant communities in a section of western Quebec. *Vegetatio,* **56,** 45-63.

Bergeron Y, Charron D (1994) Postfire stand dynamics in a southern Boreal forest (Quebec): a dendroecological approach. *Ecoscience,* **1,** 173-184.

Bergeron Y, Chen HYH, Kenkel NC, Leduc AL, Macdonald SE (in press) Boreal mixedwoods stand dynamics: ecological processes underlying a diversity of expressions. *Forestry Chronicle.*

Bergeron Y, Cyr D, Girardin MP, Carcaillet C (2010) Will climate change drive 21st century burn rates in Canadian boreal forest outside of its natural variability: Collating global climate model experiments with sedimentary charcoal data. *International Journal of Wildland Fire,* **19,** 1127-1139.

Bergeron Y, Dubuc M (1989) Succession in the southern part of the Canadian boreal forest. *Vegetatio,* 79, 51-63.

Bergeron Y, Gauthier S, Flannigan M, Kafka V (2004) Fire regimes at the transition between mixedwood and coniferous boreal forest in northwestern Quebec. *Ecology,* **85,** 1916-1932.

Bouchard A, Domon G (1997) The transformations of the natural landscapes of the Haut-Saint-Laurent (Quebec) and their implications on future resource management. *Landscape and Urban Planning,* **37,** 99-107.

Boucher Y, Arseneault D, Sirois L, Blais L (2009) Logging pattern and landscape changes over the last century at the boreal and deciduous forest transition in Eastern Canada. *Landscape Ecology,* **24,** 171-184.

Brant SV, Ortí G (2003) Phylogeography of the Northern short-tailed shrew, *Blarina brevicauda* (Insectivora: Soricidae): Past fragmentation and postglacial recolonization. *Molecular Ecology,* **12,** 1435-1449.

Briand CH, Posluszny U, Larson DW (1992) Differential axis architecture in *Thuja occidentalis* (eastern white cedar). *Canadian Journal of Botany,* **70,** 340-348.

Briand CH, Posluszny U, Larson DW, Matthes-Sears U (1991) Patterns of architectural variation in *Thuja occidentalis* L. (eastern white cedar) from upland and lowland sites. *Botanical Gazette,* **152,** 494-499.

Brouillet L, F. Coursol, M. Favreau, M. Anions, P. Bélisle, Desmet. P (2010) VASCAN, the Database of Vascular Plants of Canada. http://data.canadensys.net/vascan/.

Brubaker LB, Anderson PM, Edwards ME, Lozhkin AV (2005) Beringia as a glacial refugium for boreal trees and shrubs: New perspectives from mapped pollen data. *Journal of Biogeography,* **32,** 833-848.

Bryant Jr VM, Holloway RG (1985) *Pollen records of late-Quaternary North American sediments*. American Association of Stratigraphic Palynologists Foundation.

Buckley BM, Wilson RJS, Kelly PE, Larson DW, Cook ER (2004) Inferred summer precipitation for southern Ontario back to AD 610, as reconstructed from ring widths of *Thuja occidentalis*. *Canadian Journal of Forest Research,* **34,** 2541-2553.

Burczyk J, Adams WT, Birkes DS, Chybicki IJ (2006) Using genetic markers to directly estimate gene flow and reproductive success parameters in plants on the basis of naturally regenerated seedlings. *Genetics,* **173,** 363-372.

Burg TM, Gaston AJ, Winker K, Friesen VL (2005) Rapid divergence and postglacial colonization in western North American Steller's jays (*Cyanocitta stelleri*). *Molecular Ecology,* **14,** 3745-3755.

Bush RM, Smouse PE (1992) Evidence for the adaptive significance of allozymes in forest trees. *New Forests,* **6,** 179-196.

Cabin RJ (1996) Genetic comparisons of seed bank and seedling populations of a perennial desert mustard, Lesquerella fendleri. *Evolution,* **50,** 1830-1841.

Carcaillet C, Bergeron Y, Richard PJH, Fréchette B, Gauthier S, Prairie YT (2001) Change of fire frequency in the eastern Canadian boreal forests during the Holocene: Does vegetation composition or climate trigger the fire regime? *Journal of Ecology,* **89,** 930-946.

Caulkins HL, Jr. (1967) The ecology and reproduction of northern white-cedar. M.S. Thesis, University of Michigan.

Chambers KL (1993) *Thuja*. *Flora of North America, north of Mexico. Vol. 2, Pteridophytes and Gymnosperms*, pp. 410-411. Oxford University Press, New York.

Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution,* **24,** 621-631.

Chen HYH, Popadiouk RV (2002) Dynamics of North American boreal mixedwoods. *Environmental Reviews,* **10,** 137-166.

Chimner RA, Hart JB (1996) Hydrology and microtopography effects on northern white-cedar regeneration in Michigan's Upper Peninsula. *Canadian Journal of Forest Research,* **26,** 389-393.

Chung MY, Epperson BK, Chung MG (2003) Genetic structure of age classes in *Camellia japonica* (Theaceae). *Evolution,* **57,** 62-73.

Collingwood GH (1940) Northern white cedar. *American Forests,* **12,** 558-559.

Comes HP, Kadereit JW (1998) The effect of quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science,* **3,** 432-438.

Cornett MW, Puettmann KJ, Frelich LE, Reich PB (2001) Comparing the importance of seedbed and canopy type in the restoration of upland *Thuja occidentalis* forests of Northeastern Minnesota. *Restoration Ecology,* 9, 386-396.

Cornett MW, Reich PB, Puettmann KJ (1997) Canopy feedbacks and microtopography regulate conifer seedling distribution in two Minnesota conifer-deciduous forests. *Ecoscience,* **4,** 353-364.

Cornett MW, Reich PB, Puettmann KJ, Frelich LE (2000) Seedbed and moisture availability determine safe sites for early *Thuja occidentalis* (Cupressaceae) regeneration. *American Journal of Botany,* **87,** 1807-1814.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics,* **144,** 2001-2014.

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics,* **153,** 1989-2000.

Craft KJ, Ashley MV (2007) Landscape genetic structure of bur oak (*Quercus macrocarpa*) savannas in Illinois. *Forest Ecology and Management,* **239,** 13-20.

Critchfield WB (1985) The late Quaternary history of lodgepole and jack pines. *Canadian Journal of Forest Research,* **15,** 749-772.

Cruzan MB, Templeton AR (2000) Paleoecology and coalescence: Phylogeographic analysis of hypotheses from the fossil record. *Trends in Ecology and Evolution,* **15,** 491-496.

Curtis JD (1944) Northern white cedar on upland soils in Maine. *Journal of Forestry,* **42,** 756-759.

Curtis JD (1946) Preliminary observations on northern white cedar in Maine. *Ecology,* **27,** 23-36.

D.m.page R (1996) TreeView: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences,* **12,** 357-358.

Dansereau PR, Bergeron Y (1993) Fire history in the southern boreal forest of northwestern Quebec. *Canadian Journal of Forest Research,* **23,** 25-32.

Davis A, Puettmann K, Perala D (1998) Site preparation treatments and browse protection affect establishment and growth of northern white-cedar. U.S. Department of Agriculture, Forest Service, North Central Experiment Station.

Davis MB (1983) Quaternary history of deciduous forests of eastern North America and Europe. *Annals of the Missouri Botanical Garden,* **70,** 550-563.

Davis MB, Shaw RG (2001) Range Shifts and Adaptive Responses to Quaternary Climate Change. *Science,* **292,** 673-679.

De Blois S, Bouchard A (1995) Dynamics of *Thuja occidentalis* in an agricultural landscape of southern Quebec. *Journal of Vegetation Science,* **6,** 531-542.

de Lafontaine G, Turgeon J, Payette S (2010) Phylogeography of white spruce (*Picea glauca*) in eastern North America reveals contrasting ecological trajectories. *Journal of Biogeography,* **37,** 741-751.

Denneler B, Asselin H, Bergeron Y, Bégin Y (2008) Decreased fire frequency and increased water levels affect riparian forest dynamics in southwestern boreal Quebec, Canada. *Canadian Journal of Forest Research,* **38,** 1083-1094.

Denneler B, Bergeron Y, Bégin Y (1999) An attempt to explain the distribution of the tree species composing the riparian forests of Lake Duparquet, southern boreal region of Quebec, Canada. *Canadian Journal of Botany,* 77, 1744-1755.

Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994) Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences of the United States of America,* **91,** 3166-3170.

Diniz-Filho JAF, Nabout JC, Bini LM *et al.* (2009) Niche modelling and landscape genetics of *Caryocar brasiliense* ("Pequi" tree: *Caryocaraceae*) in Brazilian Cerrado: An integrative approach for evaluating central-peripheral population patterns. *Tree Genetics and Genomes,* **5,** 617-627.

Dong J, Wagner DB (1993) Taxonomic and population differentiation of mitochondrial diversity in *Pinus banksiana* and *Pinus conforta*. *Theoretical and Applied Genetics,* **86,** 573-578.

Dorken ME, Eckert CG (2001) Severely reduced sexual reproduction in northern populations of a clonal plant, *Decodon verticillatus* (Lythraceae). *Journal of Ecology,* **89,** 339-350.

Dudaniec RY, Spear SF, Richardson JS, Storfer A (2012) Current and Historical Drivers of Landscape Genetic Structure Differ in Core and Peripheral Salamander Populations. *PLoS ONE,* **7,** e36769.

Eckert AJ, Bower AD, Wegrzyn JL *et al.* (2009) Association genetics of coastal Douglas fir (*Pseudotsuga menziesii* var. *menziesii, Pinaceae*). I. Cold-hardiness related traits. *Genetics,* **182,** 1289-1302.

Eckert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographical ranges: The central-marginal hypothesis and beyond. *Molecular Ecology,* **17,** 1170-1188.

Ellstrand NC, Elam DR (1993) Population Genetic Consequences of Small Population Size: Implications for Plant Conservation. *Annual Review of Ecology and Systematics,* **24,** 217-242.

Epperson BK (2000) Spatial genetic structure and non-equilibrium demographics within plant populations. *Plant Species Biology,* **15,** 269-279.

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.

Excoffier L, G. Laval, Schneider aS (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online,* **1,** 47-50.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics,* **131,** 479-491.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics,* **164,** 1567-1587.

Fenton NJ, Bergeron Y (2008) Does time or habitat make old-growth forests species rich? Bryophyte richness in boreal Picea mariana forests. *Biological Conservation,* **141,** 1389-1399.

Fowells HA (1965) Silvics of forest trees of the United States. *United States Department of Agriculture Handbook,* **271,** 1189-1209.

Frankham R (1995) Conservation genetics. *Annual Review of Genetics,* **29,** 305-327.

Frankham R (2010) Challenges and opportunities of genetic approaches to biological conservation. *Biological Conservation,* **143,** 1919-1927.

Frenne PD, Graae BJ, Brunet J *et al.* (2012) The response of forest plant regeneration to temperature variation along a latitudinal gradient. *Annals of Botany,* **109,** 1037-1046.

FSC (2004) National boreal standard. *Forest stewardship council Canada working group.* Available at https://ca.fsc.org/national-boreal-standard.203.htm [accessed 30 April 2013].

Gajewski K, Garralla S, Milot-Roy V (1996) Postglacial vegetation at the northern limit of lichen woodland in northwestern Québec. *Géographie physique et Quaternaire,* **50,** 341-350.

Gamache I, Jaramillo-Correa JP, Payette S, Bousquet J (2003) Diverging patterns of mitochondrial and nuclear DNA diversity in subarctic black spruce: Imprint of a founder effect associated with postglacial colonization. *Molecular Ecology,* **12,** 891-901.

Gandhi KJK, Spence JR, Langor DW, Morgantini LE (2001) Fire residuals as habitat reserves for epigaeic beetles (*Coleoptera*: *Carabidae* and *Staphylinidae*). *Biological Conservation,* **102,** 131-141.

Gapare WJ, Aitken SN, Ritland CE (2005) Genetic diversity of core and peripheral Sitka spruce (*Picea sitchensis* (Bong.) Carr) populations: Implications for conservation of widespread species. *Biological Conservation,* **123,** 113-123.

Garralla S, Gajewski K (1992) Holocene vegetation history of the boreal forest near Chibougamau, central Quebec. *Canadian Journal of Botany,* **70,** 1364-1368.

Gauthier S, Simon JP, Bergeron Y (1992) Genetic structure and variability in jack pine populations: effects of insularity. *Canadian Journal of Forest Research,* **22,** 1958-1965.

Gilad Y, Pritchard JK, Thornton K (2009) Characterizing natural variation using next-generation sequencing technologies. *Trends in Genetics,* **25,** 463-471.

Girod C, Vitalis R, Leblois R, Fréville H (2011) Inferring population decline and expansion from microsatellite data: A simulation-based evaluation of the msvar method. *Genetics,* **188,** 165-179.

Godman RM (1958) Silvical characteristics of northern white-cedar (*Thuja occidentalis*). pp. 1-17. United States Department of Agriculture, Forest Service, Lake States Forest Research Experiment Station.

Godoy JA, Jordano P (2001) Seed dispersal by animals: Exact identification of source trees with endocarp DNA microsatellites. *Molecular Ecology,* **10,** 2275-2283.

Gómez A, Lunt DH (2007) Refugia within refugia: Patterns of phylogeographic concordance in the Iberian Peninsula. *Phylogeography of Southern European Refugia* (eds Weiss S., Ferrand N.), pp. 155-188. Springer. Dordrecht.

Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity,* **86,** 485-486.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). http://www2.unil.ch/popgen/softwares/fstat.htm, Updated from Goudet (1995).

Goudet J, Raymond M, De Meeüs T, Rousset F (1996) Testing differentiation in diploid populations. *Genetics,* **144,** 1933-1940.

Habeck JR (1958) White Cedar Ecotypes in Wisconsin. *Ecology,* **39,** 457-463.

Hamilton WD (1971) Selection of selfish and altruistic behavior in some extreme models. *Man and Beast: Comparative Social Behavior* (ed. Dillon E.a.), pp. 57-91. Smithsonian Inst. Press, Wash, DC.

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: The rear edge matters. *Ecology Letters,* **8,** 461-467.

Hamrick JL, Godt MJW, Sherman-Broyles SL (1992) Factors influencing levels of genetic diversity in woody plant species. *New Forests,* **6,** 95-124.

Hamrigk JL, Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society B: Biological Sciences,* **351,** 1291-1298.

Hardy OJ, Vekemans X (2002) SPAGeDI: A versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes,* **2,** 618-620.

Heinselman ML (1981) Fire and succession in the conifer forests of North America. *Forest Succession: Concepts and Application* (eds West D.C., Shugart H.H., Botkin D.B.), pp. 374-406. Springer-Verlag, New York, USA.

Heitzman E, Pregitzer KS, Miller RO (1997) Origin and early development of northern white-cedar stands in northern Michigan. *Canadian Journal of Forest Research,* **27,** 1953-1961.

Hély C, Flannigan M, Bergeron Y, McRae D (2001) Role of vegetation and weather on fire behavior in the Canadian mixedwood boreal forest using two fire behavior prediction systems. *Canadian Journal of Forest Research,* **31,** 430-441.

Heuertz M, Teufel J, González-Martínez SC *et al.* (2010) Geography determines genetic relationships between species of mountain pine (*Pinus mugo* complex) in western Europe. *Journal of Biogeography,* **37,** 541-556.

Hewitt G (2000) The genetic legacy of the quaternary ice ages. *Nature,* **405,** 907-913.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society,* **58,** 247-276.

Hoban S, Schlarbaum S, Brosi S, Romero-Severson J (2012) A rare case of natural regeneration in butternut, a threatened forest tree, is parent and space limited. *Conservation Genetics,* **13,** 1447-1457.

Hoban SM, Borkowski DS, Brosi SL *et al.* (2010) Range-wide distribution of genetic diversity in the North American tree *Juglans cinerea*: A product of range shifts, not

ecological marginality or recent population decline. *Molecular Ecology,* **19,** 4876-4891.

Hock Z, Szövényi P, Schneller JJ, Urmi E, Tóth Z (2008) Are sexual or asexual events determining the genetic structure of populations in the liverwort *Mannia fragrans*? *Journal of Bryology,* **30,** 66-73.

Hoebee SE, Menn C, Rotach P, Finkeldey R, Holderegger R (2006) Spatial genetic structure of Sorbus torminalis: The extent of clonal reproduction in natural stands of a rare tree species with a scattered distribution. *Forest Ecology and Management,* **226,** 1-8.

Hofmeyer PV, Kenefic LS, Seymour RS (2007) Northern white-cedar: An annotated bibliography., pp. 30 p. Cooperative Forestry Research Unit Research Report.

Holcombe JW (1976) The bryophyte flora of *Thuja* seedbed logs in a northern white-cedar swamp. *The Michigan Botanist,* **15,** 173-181.

Holderegger R, Buehler D, Gugerli F, Manel S (2010) Landscape genetics of plants. *Trends in Plant Science,* **15,** 675-683.

Holleley CE, Geerts PG (2009) Multiplex Manager 1.0: A cross-platform computer program that plans and optimizes multiplex PCR. *BioTechniques,* **46,** 511-517.

Holliday JA, Ritland K, Aitken SN (2010) Widespread, ecologically relevant genetic markers developed from association mapping of climate-related traits in Sitka spruce (*Picea sitchensis*). *New Phytologist,* **188,** 501-514.

Honnay O, Bossuyt B (2005) Prolonged clonal growth: Escape route or route to extinction? *Oikos,* **108,** 427-432.

Hörnberg G, Zackrisson O, Segerström U, Svensson BW, Ohlson M, Bradshaw RHW (1998) Boreal swamp forests: Biodiversity 'hotspots' in an impoverished forest landscape. *BioScience,* **48,** 795-802.

Hughes AR, Inouye BD, Johnson MTJ, Underwood N, Vellend M (2008) Ecological consequences of genetic diversity. *Ecology Letters,* **11,** 609-623.

Hunter Jnr ML, Hutchinson A (1994) The virtues and shortcomings of parochialism: Conserving species that are locally rare, but globally common. *Conservation Biology,* **8,** 1163-1165.

IPCC (2007) Summary for Policymakers of the Synthesis Report of the IPCC Fourth Assessment Report. *Intergovernmental Panel on Climatic Change*. Available at http://www.ipcc.ch/publications_and_data/publications_ipcc_fourth_assessment_report_synthesis_report.htm [accessed 30 April 2013].

Iverson LR, Schwartz MW, Prasad AM (2004) How fast and far might tree species migrate in the eastern United States due to climate change? *Global Ecology and Biogeography,* **13,** 209-219.

Iwaizumi MG, Takahashi M, Watanabe A, Ubukata M (2010) Simultaneous evaluation of paternal and maternal immigrant gene flow and the implications for the overall genetic composition of *Pinus densiflora* dispersed seeds. *Journal of Heredity,* **101,** 144-153.

Jackson ST, Overpeck JT, Webb Iii T, Keattch SE, Anderson KH, Orndorff S (1997) Mapped plant-macrofossil and pollen records of late quaternary vegetation change in eastern North America. *Quaternary Science Reviews,* **16,** 1-70.

Jaramillo-Correa JP, Beaulieu J, Bousquet J (2004) Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (*Picea mariana*), a transcontinental North American conifer. *Molecular Ecology,* **13,** 2735-2747.

Jaramillo-Correa JP, Beaulieu J, Khasa DP, Bousquet J (2009) Inferring the past from the present phylogeographic structure of North American forest trees: Seeing the forest for the genes. *Canadian Journal of Forest Research,* **39,** 286-307.

Johnston WF (1990) *Thuja occidentalis. Silvics of North America. Volume 1. Conifers. Agriculture Handbook 654,* (eds Burns R.M., Honkala B.H.), pp. 580-589. USDA Forest Service, Washington, D.C.

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology,* **16,** 1099-1106.

Karst AL, Antos JA, Allen GA (2008) Sex ratio, flowering and fruit set in dioecious *Rubus chamaemorus* (Rosaceae) in Labrador. *Botany,* **86,** 204-212.

Kayal WE, Allen CCG, Ju CJT *et al.* (2011) Molecular events of apical bud formation in white spruce, *Picea glauca. Plant, Cell and Environment,* **34,** 480-500.

Kelly PE, Cook ER, Larson DW (1992) Constrained Growth, Cambial Mortality, and Dendrochronology of Ancient *Thuja occidentalis* on Cliffs of the Niagara Escarpment: An Eastern Version of Bristlecone Pine? *International Journal of Plant Sciences,* **153,** 117-127.

Kelly PE, Cook ER, Larson DW (1994) A 1397-year tree-ring chronology of *Thuja occidentalis* from cliff faces of the Niagara Escarpment, southern Ontario, Canada. *Canadian Journal of Forest Research,* **24,** 1049-1057.

Kettle CJ, Hollingsworth PM, Jaffré T, Moran B, Ennos RA (2007) Identifying the early genetic consequences of habitat degradation in a highly threatened tropical conifer, *Araucaria nemorosa* Laubenfels. *Molecular Ecology,* **16,** 3581-3591.

Kneeshaw D, Bergeron Y, Kuuluvainen T (2011) Forest ecosystem structure and disturbance dynamics across the circumboreal forest. *The SAGE Handbook of biogeography* (eds Millington A., Blumler M., Schickhoff U.), pp. 261-278. SAGE, London, U.K.

Kneeshaw DD, Bergeron Y (1998) Canopy gap characteristics and tree replacement in the southeastern boreal forest. *Ecology,* 79, 783-794.

Koppitz H, Kühl H (2000) To the importance of genetic diversity of Phragmites australis in the development of reed stands. *Wetlands Ecology and Management,* **8,** 403-414.

Kremer A, Ronce O, Robledo-Arnuncio JJ *et al.* (2012) Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecology Letters,* **15,** 378-392.

Lamy S, Bouchard A, Simon JP (1999) Genetic structure, variability, and mating system in eastern white cedar (*thuja occidentalis*) populations of recent origin in an agricultural landscape in southern quebec. *Canadian Journal of Forest Research,* **29,** 1383-1392.

Laroche C, Ruel J-C, Lussier JM, Kenefic LS (2006) Regeneration of *Thuja occidentalis* L. in mixedwood stands on mesic sites after partial cuts in Quebec. *Proceedings of the Eastern Canada – USA Forest Science Conference*, pp. 90-94. Quebec, QC.

Larsen JA (1980) *The Boreal Ecosystem.* Academic Press, New York, USA.

Larson DW (2001) The paradox of great longevity in a short-lived tree species. *Experimental Gerontology,* **36,** 651-673.

Larson DW, Kelly PE (1991) The extent of old-growth *Thuja occidentalis* on cliffs of the Niagara Escarpment. *Canadian Journal of Botany,* **69,** 1628-1636.

Larson DW, Matthes-Sears U, Kelly PE (1993) Cambial Dieback and Partial Shoot Mortality in Cliff-Face *Thuja occidentalis*: Evidence for Sectored Radial Architecture. *International Journal of Plant Sciences,* **154,** 496-505.

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

Ledig FT, Bermejo-Velazquez B, Hodgskiss PD, Johnson DR, Flores-Lopez C, Jacob-Cervantes V (2000) The mating system and genic diversity in Martinez spruce, an extremely rare endemic of Mexico's Sierra Madre Oriental: An example of facultative selfing and survival in interglacial refugia. *Canadian Journal of Forest Research,* **30,** 1156-1164.

Lee-Yaw JA, Irwin JT, Green DM (2008) Postglacial range expansion from northern refugia by the wood frog, *Rana sylvatica. Molecular Ecology,* **17,** 867-884.

Lepais O, Bacles CFE (2011) Comparison of random and SSR-enriched shotgun pyrosequencing for microsatellite discovery and single multiplex PCR optimization in *Acacia harpophylla* F. Muell. Ex Benth. *Molecular Ecology Resources,* **11,** 711-724.

Lesica P, Allendorf FW (1995) When are peripheral populations valuable for conservation? *Conservation Biology,* 9, 753-760.

Lian C, Goto S, Kubo T, Takahashi Y, Nakagawa M, Hogetsu T (2008) Nuclear and chloroplast microsatellite analysis of *Abies sachalinensis* regeneration on fallen logs in a subboreal forest in Hokkaido, Japan. *Molecular Ecology,* **17,** 2948-2962.

Little EL, Jr. (1971) Atlas of United States trees, volume 1, conifers and important hardwoods: U.S. Department of Agriculture Miscellaneous Publication 1146, 9 p., 200 maps.

Liu KB (1990) Holocene paleoecology of the boreal forest and Great Lakes-st. Lawrence forest in northern Ontario. *Ecological Monographs,* **60,** 179-212.

Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany,* **82,** 1420-1425.

Lomolino MV, Riddle BR, Brown JH (2006) *Biogeography,* 3rd edn., Sinauer Associates, Sunderland, MA, USA.

Lönn M, Prentice HC (2002) Gene diversity and demographic turnover in central and peripheral populations of the perennial herb Gypsophila fastigiata. *Oikos,* 99, 489-498.

Loveless MD, Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. *Annual review of ecology and systematics. Vol. 15,* 65-95.

Lowe AJ, Boshier D, Ward M, Bacles CFE, Navarro C (2005) Genetic resource impacts of habitat loss and degradation; reconciling empirical evidence and predicted theory for neotropical trees. *Heredity,* **95,** 255-273.

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity,* **89,** 238-247.

Malausa T, Gilles A, Meglécz E *et al.* (2011) High-throughput microsatellite isolation through 454 GS-FLX Titanium pyrosequencing of enriched DNA libraries. *Molecular Ecology Resources,* **11,** 638-644.

Manel S, Joost S, Epperson BK *et al.* (2010) Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology,* **19,** 3760-3772.

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

Manel S, Segelbacher G (2009) Perspectives and challenges in landscape genetics. *Molecular Ecology,* **18,** 1821-1822.

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

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology,* **7,** 639-655.

Matthes-Sears U, Larson DW (1990) Environmental controls of carbon uptake in two woody species with contrasting distributions at the edge of cliffs. *Canadian Journal of Botany,* **68,** 2371-2380.

Matthes-Sears U, Larson DW (1991) Growth and Physiology of *Thuja occidentalis* L. from Cliffs and Swamps: Is Variation Habitat or Site Specific? *Botanical Gazette,* **152,** 500-508.

Matthes-Sears U, Larson DW (1995) Rooting characteristics of trees in rock: A study of Thuja occidentalis on cliff faces. *International Journal of Plant Sciences,* **156,** 679-686.

Matthes-Sears U, Stewart SC, Larson DW (1991) Sources of allozymic variation in Thuja occidentalis in southern Ontario, Canada. *Silvae Genet.,* **40,** 100-105.

Matveinen-Huju K, Niemelä J, Rita H, O'Hara RB (2006) Retention-tree groups in clear-cuts: Do they constitute 'life-boats' for spiders and carabids? *Forest Ecology and Management,* **230,** 119-135.

McRae BH (2006) Isolation by resistance. *Evolution,* **60,** 1551-1561.

Meglécz E, Costedoat C, Dubut V *et al.* (2010) QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics (Oxford, England),* **26,** 403-404.

Messaoud Y, Bergeron Y, Asselin H (2007) Reproductive potential of balsam fir (*Abies balsamea*), white spruce (*Picea glauca*), and black spruce (*P. mariana*) at the ecotone between mixedwood and coniferous forests in the boreal zone of western Quebec. *American Journal of Botany,* **94,** 746-754.

Millar CI, Libby WJ (1991) Strategies for conserving clinal, ecotypic, and disjunct population diversity in widespread species. *Genetics and Conservation of Rare Plants* (eds Falk D.A., Holsinger K.E.), pp. 149-170. Oxford University Press, New York, USA.

Mimura M, Aitken SN (2010) Local adaptation at the range peripheries of Sitka spruce. *Journal of Evolutionary Biology,* **23,** 249-258.

Mitton JB (1983) Conifers. *Isozymes in Plant Genetics and Breeding* (ed. Orton S.D.T.a.T.J.), pp. 443-472. Elsevier, Amsterdam.

Morgenstern EK (1996) *Geographic variation in forest trees*. UBC Press, Vancouver, BC, Canada.

Morin H, Laprise D, Bergeron Y (1993) Chronology of spruce budworm outbreaks near Lake Duparquet, Abitibi region, Quebec. *Canadian Journal of Forest Research,* **23,** 1497-1506.

Mosseler A, Major JE, Rajora OP (2003) Old-growth red spruce forests as reservoirs of genetic diversity and reproductive fitness. *Theoretical and Applied Genetics,* **106,** 931-937.

Musselman RC, Lester DT, Adams MS (1975) Localized Ecotypes of *Thuja occidentalis* L. in Wisconsin. *Ecology,* **56,** 647-655.

Nakao Y, Iwata H, Matsumoto A, Tsumura Y, Tomaru N (2001) Highly polymorphic microsatellite markers in *Chamaecyparis obtusa*. *Canadian Journal of Forest Research,* **31,** 2248-2251.

Nathan R, Horvitz N, He Y, Kuparinen A, Schurr FM, Katul GG (2011) Spread of North American wind-dispersed trees in future environments. *Ecology Letters,* **14,** 211-219.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution,* **29,** 1-10.

Nelson TC (1951) A reproduction study of northern white cedar including results of investigations under federal aid in wildlife restoration project Michigan 49-R. pp. 100. Department of Conservation Game Division, Lansing,MI.

Newell RE (2005) Provincial (Nova Scotia) Status Report on Northern White Cedar *Thuja occidentalis*. E.C. Smith Herbarium, Irving Biodiversity Collection, K.C. Irving Environmental Science Centre, Acadia University, Wolfville, NS.

O'Connell LM, Ritland CE (2000) Characterization of microsatellite loci in western redcedar (*Thuja plicata*). *Molecular Ecology,* 9, 1920-1922.

Oddou-Muratorio S, Demesure-Musch B, Pélissier R, Gouyon PH (2004) Impacts of gene flow and logging history on the local genetic structure of a scattered tree species, *Sorbus torminalis* L. Crantz. *Molecular Ecology,* **13,** 3689-3702.

Ohsako T (2010) Clonal and spatial genetic structure within populations of a coastal plant, *Carex kobomugi* (Cyperaceae). *American Journal of Botany,* 97, 458-470.

Ouarmim S, Ali AA, Asselin H, Hély-Alleaume C, Bergeron Y (submitted) Permanent versus random post-fire residual stands in the eastern Canadian boreal mixedwood forest.

Ouborg NJ, Piquot Y, Van Groenendael JM (1999) Population genetics, molecular markers and the study of dispersal in plants. *Journal of Ecology,* **87,** 551-568.

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

Pamilo P (1984) Genotypic correlation and regression in social groups: Multiple alleles, multiple loci and subdivided populations. *Genetics,* **107,** 307-320.

Pamilo P (1985) Effect of inbreeding on genetic relatedness. *Hereditas,* **103,** 195-200.

Pandey M, Rajora OP (2012a) Genetic diversity and differentiation of core vs. peripheral populations of eastern white cedar, *Thuja occidentalis* (cupressaceae). *American Journal of Botany,* 99, 690-699.

Pandey M, Rajora OP (2012b) Higher fine-scale genetic structure in peripheral than in core populations of a long-lived and mixed-mating conifer - Eastern white cedar (*Thuja occidentalis* L.). *BMC Evolutionary Biology,* **12,** 48.

Parks JC, Werth CR (1993) A study of spatial features of clones in a population of bracken fern, *Pteridium aquilinum* (Dennstaedtiaceae). *American Journal of Botany,* **80,** 537-544.

Paul V (2011) Les facteurs écologiques limitant la répartition nordique du thuja de l'est (*thuja occidentalis* L.). Master in Biology Master's thesis, Université du Québec en Abitibi-Témiscamingue. .

Payette S (1992) Fire as controlling process in the North American boreal forest. *A systems analysis of the global boreal forest* (eds Shugart H.H., Leemans R., Bonan G.B.), pp. 145-169. Cambridge University Press, Cambridge.

Payette S, Eronen M, Jasinski JJP (2002) The circumboreal tundra-taiga interface: Late Pleistocene and Holocene changes. *Ambio Special Report,* **12,** 15-22.

Peakall R, Smouse PE (2006) GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes,* **6,** 288-295.

Perry DJ, Knowles P (1989) Inheritance and linkage relationships of allozymes of eastern white cedar (*Thuja occidentalis*) in northwestern Ontario. *Genome,* **32,** 245-250.

Perry DJ, Knowles P (1990) Evidence of high self-fertilization in natural populations of eastern white cedar (*Thuja occidentalis*). *Can. J. Bot.,* **68,** 663-668.

Perry DJ, Knowles P (1991) Short-note: Are inferred out-crossing rates affected by germination promptness? *Silvae Genet.,* **40,** 35-36.

Perry DJ, Knowles P, Yeh FC (1990) Allozyme variation of *Thuja occidentalis* L. in Northwestern Ontario. *Biochemical Systematics and Ecology,* **18,** 111-115.

Petit RJ, Aguinagalde I, De Beaulieu JL *et al.* (2003) Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science,* **300,** 1563-1565.

Petit RJ, Mousadik AEL, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology,* **12,** 844-855.

Petit RJ, Vendramin GG (2005) Phylogeography of organelle DNA in plants: An introduction. *Phylogeography of Southern European Refugia* (eds Weiss S., Ferrand N.). Kluwer, Amsterdam, The Netherlands.

Pigott CD (1992) Are the distribution of species determined by failure to set seed. *Fruit and Seed Production* (eds C. Marshall, Grace. J.), pp. 203-215. Cambridge University Press, Cambridge, UK.

Piotti A (2009) The genetic consequences of habitat fragmentation: The case of forests. *IForest,* **2,** 75-76.

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

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity,* **90,** 502-503.

Plessas ME, Strauss SH (1986) Allozyme differentiation among populations, stands, and cohorts in Monterey pine. *Can. J. For. Res.,* **16,** 1155-1164.

Prati D, Schmid B (2000) Genetic differentiation of life-history traits within populations of the clonal plant *Ranunculus reptans*. *Oikos,* **90,** 442-456.

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

Provan J, Maggs CA (2011) Unique genetic variation at a species' rear edge is under threat from global climate change. *Proceedings of the Royal Society B: Biological Sciences,* **279,** 39-47.

Provan J, Powell W, Hollingsworth PM (2001) Chloroplast microsatellites: New tools for studies in plant ecology and evolution. *Trends in Ecology and Evolution,* **16,** 142-147.

Prunier J, Laroche J, Beaulieu J, Bousquet J (2011) Scanning the genome for gene SNPs related to climate adaptation and estimating selection at the molecular level in boreal black spruce. *Molecular Ecology,* **20,** 1702-1716.

Queller DC, Goodnight KF (1989) Estimating Relatedness Using Genetic Markers. *Evolution,* **43,** 258-275.

Raftery AE, Lewis SM (1992) One long run with diagnostics: Implementation strategies for Markov chain Monte Carlo. *Statistical Science,* **7,** 493-497.

Raftery AE, Lewis SM (1995) The number of iterations, convergence diagnostics and generic Metropolis algorithms. *Practical Markov Chain Monte Carlo* (eds Gilks W.R., Spiegel-halter D.J., Richardson S.), pp. 115-130. Chapman and Hall, London, U.K.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America,* **94,** 9197-9201.

Rasmussen KK, Kollmann J (2004) Poor sexual reproduction on the distribution limit of the rare tree *Sorbus torminalis*. *Acta Oecologica,* **25,** 211-218.

Rehfeldt GE (1989) Ecological adaptations in Douglas-Fir (*Pseudotsuga menziesii* var. *glauca*): a Synthesis. *Forest Ecology and Management,* **28,** 203-215.

Reusch TBH, Ehlers A, Hämmerli A, Worm B (2005) Ecosystem recovery after climatic extremes enhanced by genotypic diversity. *Proceedings of the National Academy of Sciences of the United States of America,* **102,** 2826-2831.

Reusch TBH, Hukriede W, Stam WT, Olsen JL (1999) Differentiating between clonal growth and limited gene flow using spatial autocorrelation of microsatellites. *Heredity,* **83,** 120-126.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution,* **43,** 223-225.

Richard P (1980) Postglacial history of the vegetation, south of Lake Abitibi, Ontario and Quebec. *Géographie physique et Quaternaire,* **34,** 77-94.

Riggs LA (1990) Conserving genetic resources on-site in forest ecosystems. *Forest Ecology and Management,* **35,** 45-68.

Ringius GS (1979) Status of eastern white cedar, *Thuja occidentalis*, in western Nova Scotia. *Canadian Field Naturalist,* **93,** 326-328.

Ritchie JC, Macdonald GM (1986) The patterns of post-glacial spread of white spruce. *Journal of Biogeography,* **13,** 527-540.

Robledo-Arnuncio JJ, Gil L (2005) Patterns of pollen dispersal in a small population of *Pinus sylvestris* L. revealed by total-exclusion paternity analysis. *Heredity,* **94,** 13-22.

Rooney TP, Solheim SL, Waller DM (2002) Factors affecting the regeneration of northern white cedar in lowland forests of the Upper Great Lakes region, USA. *Forest Ecology and Management,* **163,** 119-130.

Rosenberg MS, and C.D. Anderson (2011) PASSaGE: Pattern Analysis, Spatial Statistics and Geographic Exegesis. Version 2. *Methods in Ecology and Evolution,* **2,** 229-232.

Rosenberg NA (2004) DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes,* **4,** 137-138.

Sagarin RD, Gaines SD (2002) The 'abundant centre' distribution: To what extent is it a biogeographical rule? *Ecology Letters,* **5,** 137-147.

Sagarin RD, Gaines SD, Gaylord B (2006) Moving beyond assumptions to understand abundance distributions across the ranges of species. *Trends in Ecology and Evolution,* **21,** 524-530.

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution,* **4,** 406-425.

Savolainen O, Pyhäjärvi T, Knürr T (2007) Gene flow and local adaptation in trees. *Annual Review of Ecology, Evolution, and Systematics,* **38,** 595-619.

Schopmeyer CS (1974) *Seeds of woody plants in the United States*. USDA Forest Service, Agriculture Handbook 450, Washington, DC, USA.

Scott ML, Murphy PG (1986) Origin and composition of an old-growth cedar-hardwood stand: the role of dune activity. *The Michigan Botanist,* **25,** 57-65.

Segerström U (1997) Long-term dynamics of vegetation and disturbance of a southern boreal spruce swamp forest. *Journal of Vegetation Science,* **8,** 295-306.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecology Letters,* 9, 615-629.

Sexton JP, McIntyre PJ, Angert AL, Rice KJ (2009) Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, and Systematics.,* **40,** 415-436.

Sezen UU, Chazdon RL, Holsinger KE (2007) Multigenerational genetic analysis of tropical secondary regeneration in a canopy palm. *Ecology,* **88,** 3065-3075.

Simard MJ, Bergeron Y, Sirois L (1998) Conifer seedling recruitment in a southeastern Canadian boreal forest: The importance of substrate. *Journal of Vegetation Science,* 9, 575-582.

Simard MJ, Bergeron Y, Sirois L (2003) Substrate and litterfall effects on conifer seedling survivorship in southern boreal stands of Canada. *Canadian Journal of Forest Research,* **33,** 672-681.

Sirois L (2000) Spatiotemporal variation in black spruce cone and seed crops along a boreal forest-tree line transect. *Canadian Journal of Forest Research,* **30,** 900-909.

Slate J, Marshall T, Pemberton J (2000) A retrospective assessment of the accuracy of the paternity inference program CERVUS. *Molecular Ecology,* 9, 801-808.

Slatkin M (1985) Gene flow in natural populations. *Annual review of ecology and systematics. Vol. 16,* 393-430.

Slavov GT, Leonardi S, Burczyk J, Adams WT, Strauss SH, Difazio SP (2009) Extensive pollen flow in two ecologically contrasting populations of *Populus trichocarpa*. *Molecular Ecology,* **18,** 357-373.

Sork VL, Waits L (2010) Contributions of landscape genetics - Approaches, insights, and future potential. *Molecular Ecology,* **19,** 3489-3495.

St Clair JB, Mandel NL, Vance-Borland KW (2005) Genecology of Douglas fir in Western Oregon and Washington. *Annals of Botany,* **96,** 1199-1214.

St. Clair JB (2006) Genetic variation in fall cold hardiness in coastal Douglas-fir in western Oregon and Washington. *Canadian Journal of Botany,* **84,** 1110-1121.

Storfer A, Murphy MA, Evans JS *et al.* (2007) Putting the 'landscape' in landscape genetics. *Heredity,* **98,** 128-142.

Storfer A, Murphy MA, Spear SF, Holderegger R, Waits LP (2010) Landscape genetics: Where are we now? *Molecular Ecology,* **19,** 3496-3514.

Storz JF, Beaumont MA (2002) Testing for genetic evidence of Population expansion and contraction: An empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model. *Evolution,* **56,** 154-166.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology,* **7,** 453-464.

Takahashi T, Tani N, Taira H, Tsumura Y (2005) Microsatellite markers reveal high allelic variation in natural populations of Cryptomeria japonica near refugial areas of the last glacial period. *Journal of Plant Research,* **118,** 83-90.

Tardif J, Bergeron Y (1997) Comparative dendroclimatological analysis of two black ash and two white cedar populations from contrasting sites in the Lake Duparquet region, northwestern Quebec. *Canadian Journal of Forest Research,* **27,** 108-116.

Tardif J, Stevenson D (2001) Radial growth-climate association of *Thuja occidentalis* L. at the northwestern limit of its distribution, Manitoba, Canada. *Dendrochronologia,* **19,** 179-187.

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S (2001) Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): Frequency, length variation, transposon associations, and genetic marker potential. *Genome Research,* **11,** 1441-1452.

Templeton AR, Levin DA (1979) Evolutionary consequences of seed pools. *Am. Nat.,* **114,** 232-249.

Thomas CD, Cameron A, Green RE *et al.* (2004) Extinction risk from climate change. *Nature,* **427,** 145-148.

Tibayrenc M, Kjellberg F, Ayala FJ (1990) A clonal theory of parasitic protozoa: The population structures of *Entamoeba*, *Giardia*, *Leishmania*, *Naegleria*, *Plasmodium*, *Trichomonas*, and *Trypanosoma* and their medical and taxonomical consequences. *Proceedings of the National Academy of Sciences of the United States of America,* **87,** 2414-2418.

Tollefsrud MM, Sønstebø JH, Brochmann C, Johnsen Ø, Skrøppa T, Vendramin GG (2009) Combined analysis of nuclear and mitochondrial markers provide new insight into the genetic structure of North European Picea abies. *Heredity,* **102,** 549-562.

Tremblay MF, Bergeron Y, Lalonde D, Mauffette Y (2002) The potential effects of sexual reproduction and seedling recruitment on the maintenance of red maple (*Acer rubrum* L.) populations at the northern limit of the species range. *Journal of Biogeography,* **29,** 365-373.

USDA (2013) The PLANTS database. Natural Resources Conservation Service, National Plant Data Team. , Greensboro, North Carolina, USA. Website http://plants.usda.gov [accessed 27 March 2013].

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes,* **4,** 535-538.

Van Wagner CE (1983) Fire behaviour in northern conifer forests. *The Role of Fire in Northern Circumpolar Ecosystems* (eds Wein R.W., MacLean D.A.), pp. 45-80. John Wiley & Sons, New York, USA.

Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology,* **13,** 921-935.

Viktora M, Savidge RA, Rajora OP (2011) Clonal and nonclonal genetic structure of subarctic black spruce (*Picea mariana*) populations in Yukon territory. *Botany,* **89,** 133-140.

Vincent JS, Hardy L (1977) L'évolution et l'extinction des lacs glaciaires Barlow et Ojibway en territoire québécois. *Géogr. Phys. Quat.,* **31,** 357-372.

Weir BS (1996) *Genetic Data Analysis II.* Sinauer Associates, Sunderland, MA.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution,* **38,** 1358-1370.

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics,* **163,** 1177-1191.

Woodward FI (1987) *Climate and Plant Distribution.* Cambridge University Press Cambridge, U.K.

Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution,* **19,** 395-420.

Wright SJ, Trakhtenbrot A, Bohrer G *et al.* (2008) Understanding strategies for seed dispersal by wind under contrasting atmospheric conditions. *Proceedings of the National Academy of Sciences of the United States of America,* **105,** 19084-19089.

Xu H, Tremblay F, Bergeron Y (2013) Development and Multiplexed Amplification of SSR Markers for *Thuja occidentalis* (Cupressaceae) Using Shotgun Pyrosequencing. *Applications in Plant Sciences,* **1,** 1200427.

Xu H, Tremblay F, Bergeron Y (submitted-a) Landscape features influence genetic diversity in "*arbor vitae*", eastern white cedar (*Thuja occidentalis* L.) and conservation value of its fire residuals.

Xu H, Tremblay F, Bergeron Y (submitted-b) Sexual or asexual reproduction, which is more important in regeneration for a model, late successional species, eastern white cedar (*Thuja occidentalis* L.)? .

Xu H, Tremblay F, Bergeron Y, Paul V, Chen C (2012) Genetic consequences of fragmentation in "*arbor vitae*," eastern white cedar (*Thuja occidentalis* L.), toward the northern limit of its distribution range. *Ecology and Evolution,* **2,** 2501-2515.

Yakimowski SB, Eckert CG (2008) Populations do not become less genetically diverse or more differentiated towards the northern limit of the geographical range in clonal *Vaccinium stamineum* (Ericaceae). *New Phytologist,* **180,** 534-544.

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

Zazula GD, Telka AM, Harington CR, Schweger CE, Mathewes RW (2006) New spruce (Picea spp.) macrofossils from Yukon Territory: Implications for late Pleistocene refugia in Eastern Beringia. *Arctic,* **59,** 391-400.