Genetic subpopulation structuring and its implications in a mature eastern white pine stand

Samuel E. Nijensohn, Paul G. Schaberg, Gary J. Hawley, and Donald H. DeHayes

Abstract: We examined patterns of genetic structuring within a mature eastern white pine (*Pinus strobus* L.) forest, using geographic information system (GIS)-based data and maps that combined genetic (isozyme analysis of 46 loci) and other tree-specific information (e.g., size, growth, age, and location) for 220 trees in Jericho, Vermont. Interconnections between genotypic information with other tree characteristics revealed several patterns of genetic structuring. Average observed heterozygosity generally increased with tree age-class, and trees with a high number of rare alleles were disproportionally represented in suppressed crown classes. Spatial structuring was also evident: trees within 5 m of one another were highly related, and levels of relatedness generally decreased with increasing distance between trees. In general, a 35-m-radius circle around any tree circumscribed its zone of genetic similarity. Hierarchical cluster analysis indicated the stand consisted of five family groups that exhibited greater genetic similarity within than among clusters. Temporal structuring (a generation gap) was also evident: trees of similar age showed significant positive relatedness, as did trees 30–40 years apart. Patterns of genetic structuring likely resulted from the combined influences of natural selection, isolation by distance, and functional generation times. Genetic structuring may also have biological and management implications. Computer-based simulated harvests suggested that the stand could experience genetic alteration when tree removal criteria disrupted existing structural patterns.

Résumé : Les auteurs ont étudié la structure génétique d'une forêt mature de pin blanc (*Pinus strobus* L.) à l'aide de données et de cartes associées à un SIG combinant des données de nature génétique (analyse de 46 loci d'isoenzymes) et des informations relatives aux arbres (p. ex., la taille, la croissance, l'âge et l'emplacement) pour 220 arbres à Jericho dans l'État du Vermont. L'étude des relations entre l'information génotypique et les autres caractéristiques des arbres a permis de mettre en évidence plusieurs patrons de structure génétique. L'hétérozygotie moyenne observée augmente généralement avec la classe d'âge, et les arbres possédant un nombre élevé d'allèles rares sont disproportionnellement plus représentés dans les classes de cime supprimées. Il y a aussi une structure spatiale : les arbres situés à 5 m et moins les uns des autres sont très apparentés et le degré de parenté diminue généralement avec l'augmentation de la distance entre les arbres. En général, un cercle de 35 m de rayon autour d'un arbre délimite sa zone de similitude génétique. L'analyse de regroupement hiérarchique indique que le peuplement est composé de cinq groupes familiaux qui démontrent une plus grande similitude génétique à l'intérieur des regroupements qu'entre ces derniers. Une structure temporelle (une coupure générationnelle) est également évidente : les arbres d'âge similaire sont plus apparentés, de même que les arbres ayant des différences d'âge de 30 à 40 ans. Les patrons de structure génétique sont probablement le résultat de l'influence combinée de la sélection naturelle, de l'isolement par la distance et de la durée de génération fonctionnelle. La structure génétique peut également avoir des implications de nature biologique et pour l'aménagement. Des coupes forestières simulées par ordinateur laissent croire que le peuplement pourrait subir une altération génétique si les critères de coupe perturbent les patrons actuels de la structure.

[Traduit par la Rédaction]

Introduction

Genetic diversity is an essential component of biodiversity. By providing the means for selective adaptation to environmental change, genetic diversity plays a critical role in maintaining forest ecosystem health. Trees in particular, because of their long life spans and limited mobility, depend heavily on genetic diversity within populations to allow for survival and adaptation in response to environmental change. In fact, trees (particularly conifers) are among the most genetically variable organisms measured in allozyme research (Bush and Smouse 1992).

Change, human-induced or otherwise, occurs continuously throughout forests across a range of temporal and spatial scales as a result of numerous factors, including wind, fire, insects, pathogens, drought, temperature extremes, storms, timber harvests, pollution, and climate change. Informed attempts at maintaining ecological sustainability must ensure

Received 1 December 2004. Accepted 29 December 2004. Published on the NRC Research Press Web site at http://cjfr.nrc.ca on 19 May 2005.

S.E. Nijensohn and P.G. Schaberg.[1] USDA Forest Service, Northeastern Research Station, Burlington, VT 05402, USA.
G.J. Hawley and D.H. DeHayes. The University of Vermont, Rubenstein School of the Environment and Natural Resources, Burlington, VT 05405, USA.

[1]Corresponding author (e-mail: pschaberg@fs.fed.us).

 doi: 10.1139/X05-029

that tree populations retain sufficient genetic diversity to successfully survive and adapt to such changes (Rajora 1999). Structures of genetic diversity (i.e., the nonrandom heterogeneous distribution of genes and genotypes) are dynamic and mutable. Directed management can guide genetic structuring towards various ends, including maximizing effective population and subpopulation sizes, selecting for specific adaptations, and increasing genetic diversity in reserves for the future (Namkoong 1991). On the other hand, management also has the potential to cause degradation of genetic structure within forests (Savolainen and Karkkainen 1992). Understanding the structuring of genetic diversity is therefore an integral part of managing ecosystem resilience and long-term health (Riggs 1990).

It is generally accepted that genotypes tend to be nonrandomly distributed within populations (Perry and Knowles 1991; Epperson 2000). However, empirical data that elucidates specific patterns of genetic structure within native, and especially managed forests remains incomplete (Ledig 1988; Namkoong 1991; Epperson and Chung 2001). This applies both to the structuring of genotypes with respect to phenotypes, and to the structuring of genotypes with respect to space and time.

Our research utilized a unique tool to examine genetic structuring: a GIS (geographic information systems) database of the physical, spatial, temporal, and genetic attributes of all mature individual trees in an eastern white pine (*Pinus strobus* L.) stand in Jericho, Vermont. This framework allowed us to categorically inventory and query baseline genotypic and physical data to examine patterns of genetic structuring with respect to variations in location, age, and crown class. This database also provided a unique platform for examining the potential influence of silvicultural manipulation on stand genetics by conducting a series of nondestructive simulated harvests with varying selection criteria. Eastern white pine was chosen as a study species because of its economic and historic importance, and because it exhibits high genetic variability with substantial local adaptation (Buchert 1994).

Methods

Site

The study site is an approximately 10-ha stand in Jericho, Vermont, USA, that is dominated by mature eastern white pine. This site is representative of Vermont's land use history; the oldest pines are "wolf" trees that were likely pioneers in a pasture deserted before the turn of the 20th century (Johnson 1980). Other tree species with a minor presence include red maple (*Acer rubrum* L.), red spruce (*Picea rubens* Sarg.), red oak (*Quercus rubra* L.), white ash (*Fraxinus americana* L.), black cherry (*Prunus serotina* Ehrh.), American beech (*Fagus grandifolia* Ehrh.), paper birch (*Betula papyrifera* Mash.), and striped maple (*Acer pensylvanicum* L.).

GIS database

A comprehensive database was constructed by first gathering inventory data, including size, age, growth, vigor, crown class, and spatial location for all 232 eastern white pine trees in the forest. These data were combined with genetic data, determined from starch gel electrophoresis, in a GIS framework.

Table 1. Enzyme systems, number of loci resolved, and buffer systems used in starch gel electropheresis isozyme analyses.

Enzyme system	No. of loci	Buffer system
Asparate aminotransferase	3	TC
Aconotase	1	LB
Alcohol dehydrogenase	4	8MC
Acid phosphotase	4	6MC
Adenylate kinase	3	6MC
Aldolase	2	8MC
Catalase	1	TC
Diaphorase	3	TC
Fumarase	1	8MC
Glutamate dehydrogenase	1	TC
Glucose-6-phosphate dehydrogenase	4	8MC
Isocitrate dehydrogenase	2	6MC
Leuicine aminopeptidase	2	LB
Malate dehydrogenase	4	6MC
Mannose-6-phosphate isomerase	1	TC
Phosphoglucoseisomerase	3	LB
Phosphoglucomutase	2	LB
6-Phosphogluconic dehydrogenase	3	6MC
Shikimate dehydrogenase	2	TC

Note: Buffer systems are the following: Tris citrate, pH 8.3 (TC); lithium borate, pH 8.8 (LB); morpholine citrate, pH 6.1 (6MC); and morpholine citrate, pH 8.1 (8MC) (Jech and Wheeler 1984; Beaulieu and Simon 1994; Chagala 1996; Buchert et al. 1997; Rajora et al. 1998).

Physical and age measurements

Stem diameter at 1.3 m (DBH) was recorded, and two increment cores were taken from each tree. After crossdating (Stokes and Smiley 1968), cores were used to calculate tree age using the methods of Norton and Ogden (1990). Crown class was visually determined following the Kraft crown classification system (Smith et al. 1997). Precise spatial coordinates for each individual were obtained using a Total Station (Pentax 2CS, Asahi Optical Corp. Ltd., Tokyo, Japan) surveying tool.

Genetic measurements

Cones were collected from individual trees in the fall of 2000, during an approximately 3-week period between seed maturity and cone opening. Almost all trees (220 of 232 trees) produced seed, allowing for a nearly complete assay of stand genetics. The remaining 12 trees were all suppressed and did not produce seed at any time during the 3 years of this study.

Cones were dried, and seed was manually extracted and dewinged. Following a 24-h soak in water, seed was germinated on moist filter paper in a growth chamber at 22 °C. Seed coats and embryos were removed, and haploid megagametophytes were extracted for electrophoretic assessments. A minimum of seven haploid megagametophytes from each tree were used in genetic analyses. Starch gel electrophoresis resolved 46 loci (26 of which exhibited polymorphisms) from 19 enzyme systems (Table 1). Previous studies have indicated that these loci show independent Mendelian inheritance (e.g., Buchert et al. 1997; Rajora et al. 1998).

Population genetics and statistical analyses

Stand level genetic assessments were determined using BIOSYS-1 (Swofford and Selander 1989). Statistics of interest included heterozygosity (observed and expected), percentage of loci exhibiting polymorphism, effective number of alleles per locus, observed number of alleles per locus, individual locus allele frequencies, and the number of rare ($p \leq 0.01$) and uncommon ($0.01 < p \leq 0.05$) alleles.

Analyses of variances and Tukey–Kramer honestly significant difference (HSD) tests were used to test for significant differences in average heterozygosity among age-classes, and differences in the number of rare alleles among crown classes. Paired t tests were used to test whether levels of expected and observed heterozygosities differed, and thereby assess whether cohorts were in Hardy–Weinberg equilibrium.

Family structuring was analyzed by comparing genetic relatedness values, spatial distances, and age differences for all possible pairs of trees (24 090 distinct pairs for 220 trees excluding self-comparisons). The relatedness, spatial distance, and age difference between two trees was referred to as the pairwise relatedness, pairwise distance, and pairwise age difference of these trees, respectively.

Individual tree multilocus genotypes were converted into a 220 × 220 matrix of pairwise relatedness values (i.e., a measure of the relatedness between pairs of trees) ranging between 1 and –1 using Relatedness 5.0.8 (Goodnight and Queller 1999). This program compares the genotypes of each pair of individuals relative to the entire population, preferentially weighting rarer alleles that possess greater value for determinations of heredity (Queller and Goodnight 1989). A relatedness value of +1 indicates that the individuals compared are genetically identical for the loci assessed, +0.5 indicates genetic similarities equivalent to full-siblings, +0.25 indicates genetic similarities equivalent to half-siblings, etc. A relatedness value of 0.0 indicates that compared individuals have the same degree of similarity to each other as they have on average to randomly selected members of the population. Conversely, negative relatedness values indicate that compared individuals are more dissimilar to each other than they are to the population at large. At the extreme, a value of –1 indicates that the compared individuals are dissimilar for all loci assessed. The following equation was used to calculate R, the relatedness value:

$$R = \frac{\sum_x \sum_k \sum_l (P_y - P^*)}{\sum_x \sum_k \sum_l (P_x - P^*)}$$

where

- x indexes individuals in the data set.
- k indexes loci.
- l indexes allelic position (i.e., $l = 1$ or 2 for a diploid individual, or only 1 for a haploid).
- P_x is the frequency within the current x individual of the allele found at x's locus k and allelic position l. In a diploid, this value must be either 0.5 or 1.0.
- P_y is the frequency of that same allele in the set of "partners" of x – the individual(s) to which you are assessing x's relatedness.
- P^* is the frequency of the allele in the population at large (Goodnight and Queller 1999).

Matrices describing pairwise distances and pairwise age differences between all pairs of trees were created in ArcView® GIS 3.2 using the extension "dmatrix_en.avx" (Maoh 2001).

Mean pairwise relatedness values within categorically derived subpopulations (i.e., groups of tree pairs whose memberships were determined by pairwise distance rankings or pairwise age difference rankings) were compared with random means created using Monte Carlo simulations (1000 resampling iterations) (Simon 2002). Monte Carlo simulations repeatedly select random samples of user-determined sample sizes (i.e., of the same sample size as a categorically derived subpopulation) from the complete data set. The means from these resampling groups create a random distribution used as a comparison to determine the statistical probability that a categorically derived subpopulation's mean could occur by chance.

Clustering analyses and mapping were performed with a hierarchical averaging technique using JMP software (version 4.0.2, SAS Institute Inc., Cary, N.C.). Relatedness values were used as clustering variables. Five cluster groups were identified based on the first significant distance gap between bridged clusters comprising the output dendrogram. Maps were created for cluster visualization. Clusters were compared using analyses of variances to determine if they differed in mean age, heterozygosity, and rare alleles.

Harvesting scenarios were developed using the following characteristics as silvicultural selection criteria: tree age, basal area of the most recent 10 years of growth, DBH, and crown class at various harvesting intensities (e.g., simulating the removal of different proportions of the original stand). An additional harvesting scenario was developed that used the on-site judgment of the regional Vermont State Forester to select trees for silvicultural removal.

Several genetic diversity parameters, including heterozygosity (observed and expected), and the mean number of uncommon alleles, were calculated for all residual stands following harvest simulations. Paired t tests were used to compare the genetics of stands following simulated harvests to the uncut stand and to determine whether residual stands were in Hardy–Weinberg equilibrium. Monte Carlo simulations were also used to compare the genetics of residual stands following harvest simulations to repeated random harvests of equal intensities (1000 resampling iterations) (Simon 2002). By comparing harvests that left the same number of trees in residual stands, the confounding effects of differential sample size on genetic diversity measures were avoided.

Results

Physical overview

A diverse group of 220 genotyped mature eastern white pine trees comprised the stand (Fig. 1). The largest tree's DBH was over 5 times greater than the smallest; DBHs ranged from 21.4 to 113.4 cm. Although all crown classes were present, most trees were codominant or dominant, while only 36 and 5 trees were intermediate or suppressed, respectively. Average tree age was about 77 years old at breast height, with individuals ranging from 60 to 118 years old. Sixteen of the oldest individuals were "wolf" trees with large diameter branches extending nearly to ground level (Table 2).

Fig. 1. One of the GIS layers for the eastern white pine study site in Jericho, Vermont. Various stand attributes (phenotypic or genotypic) can be queried and depicted. In this example, tree locations and DBHs are displayed.

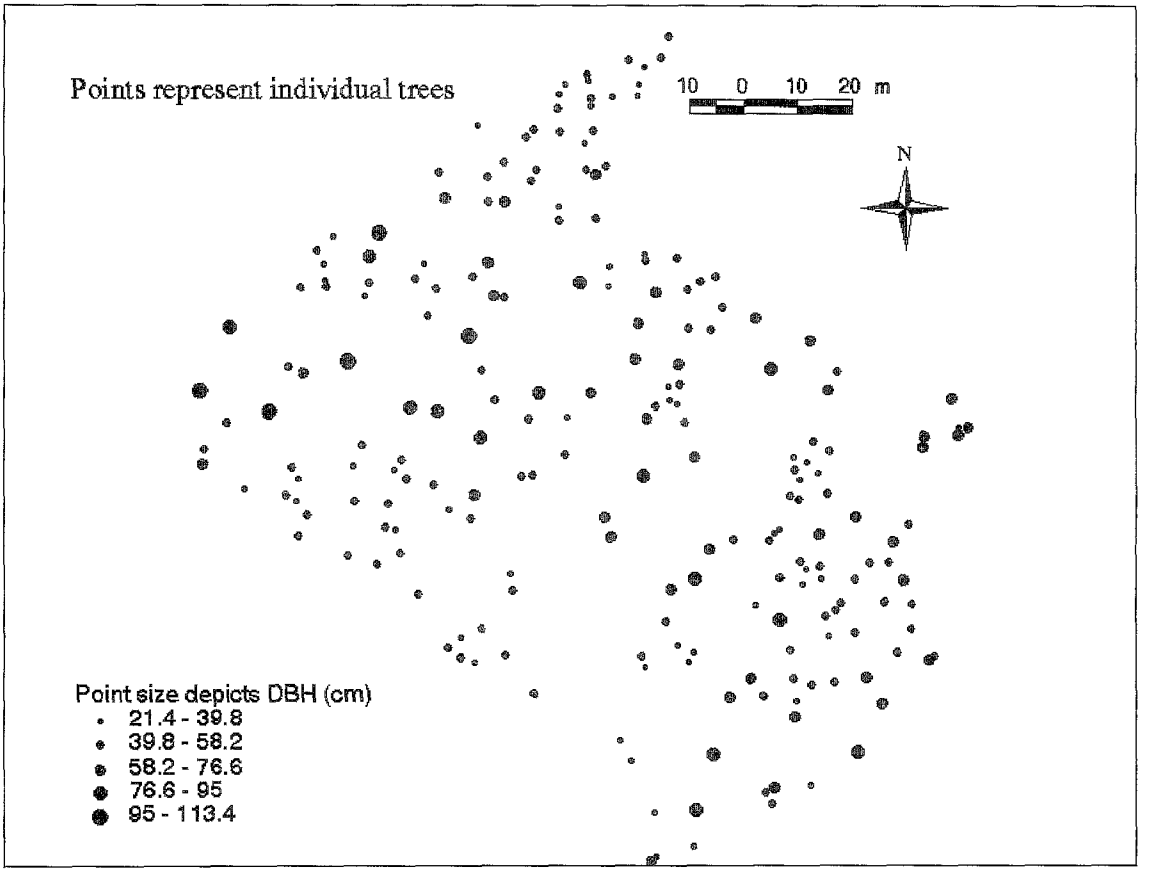


Table 2. Overview of the physical characteristics for the eastern white pine study site in Jericho, Vermont.

Physical characteristic	Value(s)
Trees assessed	220
Wolf trees	16
Crown class	
Dominant	62
Codominant	115
Intermediate	36
Suppressed	5
Mean DBH (cm) (max., min.)	52.2 (113.4, 21.4)
Mean age (years) (max., min.)	77.4 (118, 60)

Table 3. Overview of the genetic characteristics for the eastern white pine study site in Jericho, Vermont.

Genetic characteristic	Value(s)
Trees genotyped	220
Polymorphic loci	46
Uncommon alleles ($0.01 < p \leq 0.05$)	26
Rare alleles ($p \leq 0.01$)	27
Mean (SE) observed heterozygosity	0.113 (0.027)
Mean (SE) expected heterozygosity	0.114 (0.027)
Mean (SE) no. of alleles per locus	1.91 (0.15)

Genetic overview

Overall, the stand was in Hardy–Weinberg equilibrium. Mean observed heterozygosity was 0.113, and of the 46 loci resolved, 26 were polymorphic. The mean number of alleles per locus was 1.91. There were numerous rare and uncommon alleles (Table 3). These measures are similar to previous reports for eastern white pine (e.g., Rajora et al. 1998; Beaulieu and Simon 1994; Buchert at al. 1997).

Evidence of structuring by age and crown class

The interconnections between age and genotypic characteristics resulted in several patterns of genetic structuring. For example, when trees were ranked by age into five equal-sized age-classes, significant differences in mean observed heterozygosity were found between classes (ANOVA, $p = 0.02$). Heterozygosity generally increased for older age-classes, with age-class 4 (the second oldest class with trees ranging from 77 to 87 years old) exhibiting significantly higher heterozygosity than the age-class 1 (the youngest class with trees ranging from 60 to 68 years old) ($p \leq 0.05$, Tukey–Kramer HSD test) (Fig. 2). Also, trees from more dominant crown classes tended to have fewer rare ($p \leq 0.01$) alleles, although the low sample sizes for the intermediate ($n = 36$) and suppressed ($n = 5$) classes limited statistical resolution, so that

Fig. 2. Mean heterozygosities and standard errors of trees from the Jericho, Vermont, study site grouped into five equal-sized age groups. Age-classes that do not share the same letters are significantly different ($p \leq 0.05$) based on results from a Tukey–Kramer honestly significant difference test.

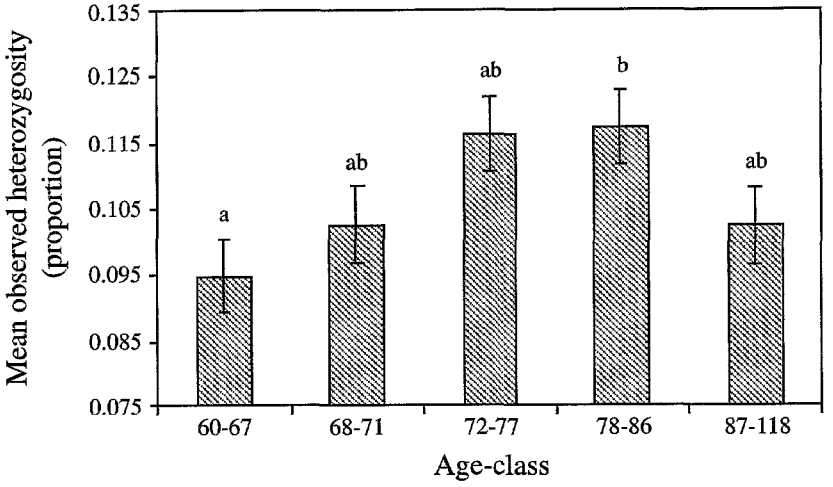
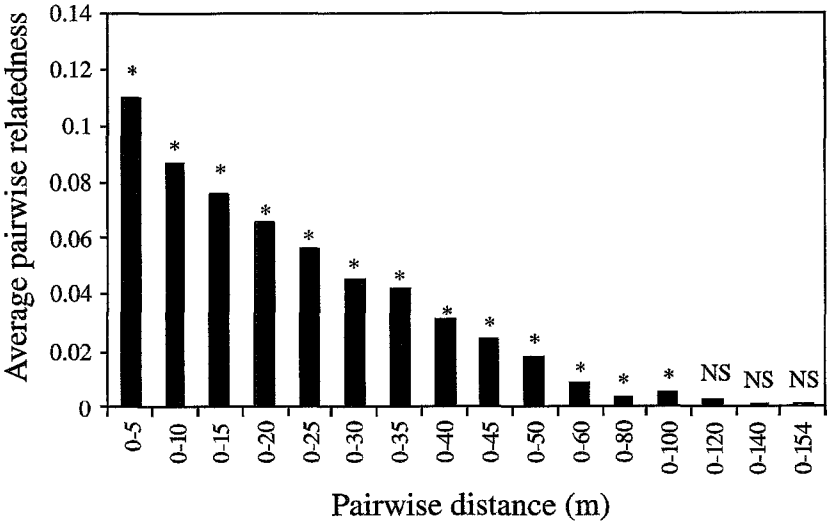


Fig. 3. Average pairwise relatedness among pairwise distance groups of increased cumulative size. Asterisks represent significant positive relatedness ($p \leq 0.05$), and NS designates relatednesses that are not significantly positive ($p > 0.05$) based on results from Monte Carlo simulations. Relatedness values can range from 1 (individuals are identical for all alleles assessed) to –1 (individuals have no alleles in common).

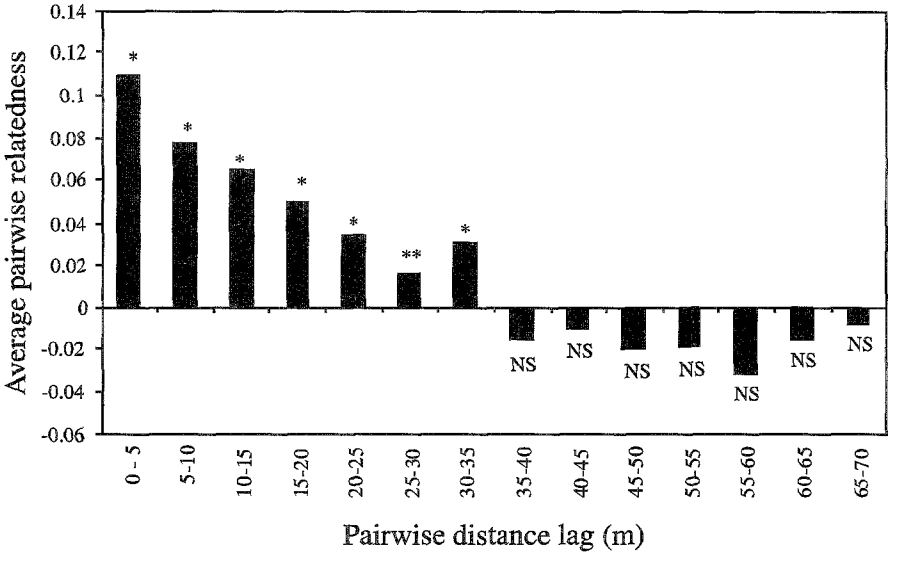


differences were not statistically significant (Tukey–Kramer HSD test).

Evidence of family structuring

More striking than patterns between heterozygosity and age, or rare alleles and crown class, was evidence of family structuring in space and time. Comparing every possible pair of trees in terms of genetic relatedness, location, and age provided a visual as well as quantitative assessment of distinct family structures. Overall, pairs of trees close to one another demonstrated levels of relatedness significantly higher than random pairings. A clear indication of this was that, on average, pairwise relatedness values (i.e., a measure of the relatedness between pairs of trees) decreased as observations were made at ever increasing distances radiating from each tree (Fig. 3). However, this measure evaluated average relatedness for trees over cumulative distances, which likely overestimated the extent of relatedness across the stand, because strong genetic similarities among very close trees were included even in more distant comparisons. A slightly different approach was used to estimate average family size (i.e., the distance at which average pairwise relatedness is no longer significantly positive). We divided pairwise distance, the spatial distance between pairs of trees, into 5-m lags (e.g., distance lag "30–35" included all pairs of trees that were between 30 and 35 m apart) for a detailed view of relatedness across distance. Distance lags above 70 m were not included in this analysis, because low sample sizes unduly weighted the influence of a few trees at the stand's perimeter. This analysis showed that, on average, an approximately 35-m-radius circle drawn around any tree circumscribed its zone of genetic similarity (Fig. 4). Within this zone, the center

Fig. 4. Average pairwise relatedness of pairwise lagged distance groups. Asterisks represent significant positive relatedness with $p \leq 0.05$, the double asterisk represents significant positive relatedness with $p \leq 0.10$, and NS designates relatednesses that are not significantly positive ($p > 0.10$) based on results from Monte Carlo simulations. Relatedness values can range from 1 (individuals are identical for all alleles assessed) to –1 (individuals have no alleles in common).



tree was significantly more related to other trees than to randomly selected trees from the overall stand.

This general relationship between spatial location and relatedness can be more specifically examined by actually mapping family group boundaries. We performed hierarchical cluster analysis on the pairwise relatedness matrix, clustering individuals into family groups according to their relatedness values. Five family groups in this stand exhibited significantly greater genetic similarity within clusters than among the clusters. When a stand map was labeled according to which group each tree belonged, the family groups (of varying size) exhibited distinct spatial structures (Fig. 5). Analyses of variance showed significant differences between clusters in mean observed heterozygosity (ranging from 0.09 to 0.18, $p \leq 0.01$), mean number of uncommon alleles (ranging from 0.58 to 3.71, $p \leq 0.01$), and mean age (ranging from 75.19 to 85.29, $p \leq 0.02$). Cluster 1, which has the most members (n = 126), is the only cluster that did not exhibit spatial aggregation.

Temporal patterns between relatedness and age (expressed as lags of age differences) also revealed family structuring (Fig. 6). As in distance lag analysis, extreme age difference lags were not included in this analysis, because low sample sizes exaggerate the influence of the few oldest trees. The significant positive relatedness exhibited by like-aged individuals ("0–5" year age difference lag) may indicate within cohort relatedness (e.g., sibling and cousin relationships). Whereas the significant positive relatedness at "30–35" and "35–40" year age difference lags likely reflects intergenerational relationships. Based on this, we estimated that the effective generation time in this stand (i.e., the minimum significant age difference between parents and offspring) was approximately 30 years. Although significantly positive, mean relatedness within like-aged and intergenerational groups was not high, averaging about 0.05 for the 0–5 lag and approximately 0.03 for the greater age lags. However, these groups included over 25% of individual pairings with relatedness levels ≥ 0.25 (genetic similarities equivalent to half-sibs) and about 5% of pairings with relatedness levels ≥ 0.50 (analogous to full-sibs).

Discussion

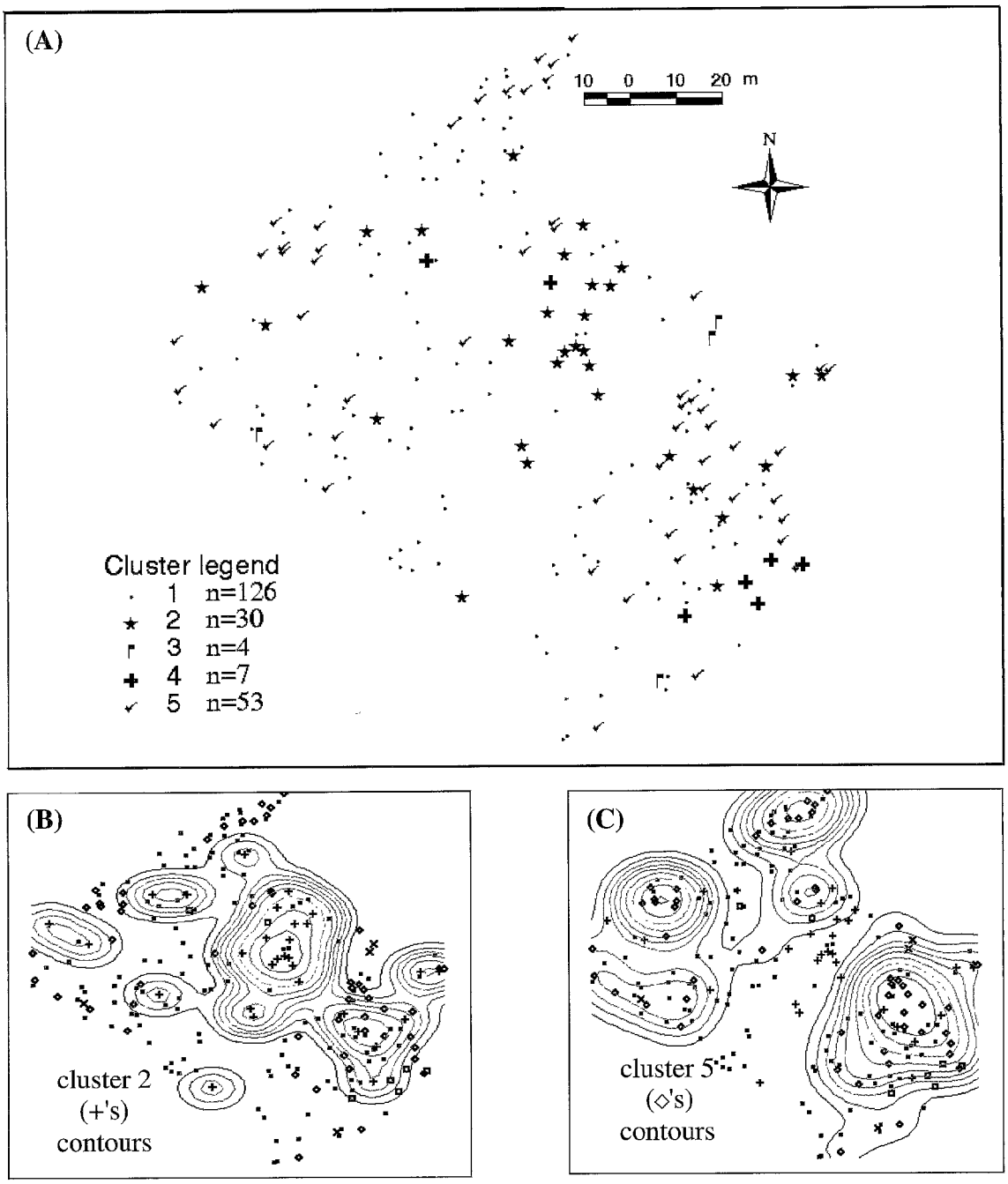
Patterns and processes

The patterns of genetic structuring that we observed likely resulted from the influence of several processes, including fitness-based selection on genotype-influenced phenotypes (natural selection), spatial family structuring determined by limited pollen and seed transport with distance (isolation by distance), and temporal family structuring resulting from a generation gap between offspring and parents determined by the age of effective reproductive maturity (generation time). Microhabitat selection could also create genetic structuring; however, this process is likely to affect only a few loci. By combining 46 loci in the analysis, we diluted the influence of this potentially confounding effect (Berg and Hamrick 1995).

Natural selection

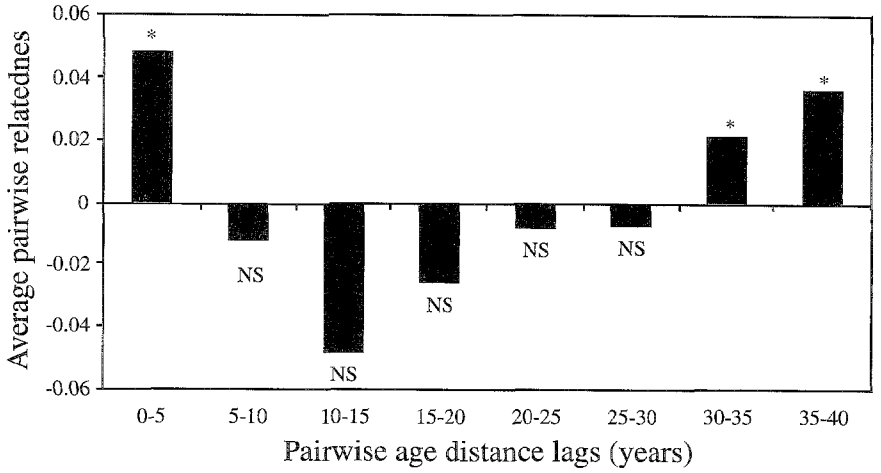
Heterozygosity has been shown to be related to fitness (e.g., Jelinski 1993; Li and Wu 1996; Arcade et al. 1996; Li et al. 1998). This phenomenon, referred to as heterosis, has been linked to two possible processes: "inbreeding depression or dominance", whereby deleterious recessive alleles are homozygous in inbred individuals, and "overdominance" whereby heterozygosity for multiple alleles may provide a survival advantage by allowing greater plasticity in physiological function (Knowles and Grant 1981; Ledig et al. 1983; Bush et al. 1987). Older populations or cohorts have been shown to have higher heterozygosity, presumably because they have been subjected to greater cumulative selective pres-

Fig. 5. (A) Family clusters based on relatedness shown in space. Symbols depict a tree's spatial location and distinct cluster group (e.g., each tree is a member of one of the five cluster groups). Note that clusters (except group 1) visually demonstrate spatial grouping. (B and C) Two examples of nonparametric density contours delineating a cluster's spatial boundaries (clusters 2 and 5 shown here). Density contours represent a smooth surface that describes how dense data points are at each location based on bivariate density estimation models within the Nonpar Density feature of JMP statistical software (2002; SAS Institute Inc., Gary, N.C.).



sure over time, which preferentially selected for individuals with a relatively high proportion of heterozygous loci and (or) against homozygosity (Neale and Adams 1985; Plessas and Strauss 1986; Hawley et al. 1988; Bush and Smouse 1992). This stand showed a similar pattern of increasing heterozygosity with age-class, although the oldest age-class diverged from the trend (Fig. 2). Slightly lower heterozygosities for the oldest trees may indicate that competitive pressures were low when this group, which included all the open-grown wolf trees, initially colonized the site. Low heterozygosity levels may also reflect a "founder effect", whereby diversity was restricted by the genetics of a limited number of initiating seed sources. The majority of wolf trees shared positive relatedness values (approximately 64% of wolf

Fig. 6. Average pairwise relatedness at age difference lags (e.g., the 0–5 lag includes all pairwise comparisons of trees that differ in age from 0 to 5 years). Asterisks represent significant positive relatedness ($p \leq 0.05$), and NS designates relatednesses that are not significantly positive ($p > 0.05$) based on results from Monte Carlo simulations. Relatedness values can range from 1 (individuals are identical for all alleles assessed) to –1 (individuals have no alleles in common).



tree pairs), and the overall mean relatedness of wolf trees was comparatively high (0.105), which would be expected if this group resulted from a small founder population.

Another interesting relationship between fitness and genotype involved rare alleles. Both in heterozygous and homozygous forms, rare alleles have been associated with weaker phenotypes (Cheliak et al. 1988; Bush and Smouse 1991; Adams et al. 1998; Hawley et al. 2005). Consistent with this, the mean number of rare alleles per individual exhibited an inverse relationship with crown class dominance. Mean number of rare alleles increased from 0.06 for dominant and 0.08 for co-dominant trees, to 0.11 for intermediate and 0.2 for suppressed trees, although crown class means were not significantly different from one another because of the low number of intermediate and suppressed individuals.

Because heterozygosity and rare allele levels both showed some association with fitness, it is not surprising that a group unique in one of these measures was also distinctive in the other. The second age-class (age 68–71 years old) was the only one deviating from Hardy–Weinberg equilibrium (because of an excess of heterozygotes; $p \leq 0.02$, paired t test). Interestingly, this age-class was also the only class to show a significantly low number of rare and uncommon alleles ($p \leq 0.03$ for uncommon alleles, $p \leq 0.02$ for rare alleles, with comparisons made against the distribution of means derived from Monte Carlo simulations). Selection, perhaps during the vulnerable period of seedling establishment and early growth, may have shaped the unique genetics of this group.

Isolation by distance

The GIS framework allowed for substantial advances in evaluating Wright's isolation by distance model (Wright 1969). Although it has long been supposed that spatial and temporal genetic structures exist as a result of familial relationships, tree population research has often failed to find family patterning in accordance with theoretical expectations (Smouse and Peakall 1999), leading some to suggest that high outcrossing rates in conifers prevent such structuring (Savolainen and Karkkainen 1992). Various researchers have proposed that typically used statistical procedures, such as the per locus autocorrelation assessments using Moran's I statistics, may fail to detect generalized evidence of spatial structuring (e.g., Epperson 1995*a*, 1995*b*; Berg and Hamrick 1995; Smouse and Peakall 1999). Indeed, by using a broader multilocus technique, we provide evidence of significant family structuring. Our method determined average pairwise relatedness values (Queller and Goodnight 1989) within distance and age difference classes, and then used Monte Carlo resampling statistics to determine significance levels.

Our results indicate that family structuring has occurred in space and time in this stand. Spatial structuring was evident from the strongly significant relationship between pairwise distance and pairwise relatedness among trees (Fig. 3). This pattern likely resulted from the exponential decrease of effective pollen and seed transport with distance (Adams 1992), and is consistent with patterns detected for eastern white pine from a variety of locations and growth stages (Epperson and Chung 2001; Rajora et al. 2002). In addition, by dividing pairwise distance into 5-m lag groups, we were able to perceive a family structure, averaging 35 m in radius, surrounding each tree (Fig. 4). Also, a cluster analysis performed on the relatedness matrix allowed us to group trees into clusters of highly related individuals (i.e., denoting families). This procedure further confirmed the existence of spatial structure, when family clusters, plotted on a map, visually exhibited discernable spatial patterning (Fig. 5). Density contours were drawn on the map to further delineate cluster boundaries (Figs. 5B and 5C).

Generation time

The primary indication of generation time (i.e., the minimum age differences between generations, which is driven by the age at which an individual becomes part of the effective breeding population) has been rather loosely associated with the age at which trees arrived at sexual maturity. For some monoecious species like eastern white pine, this age is

not specific; eastern white pine can produce female strobili at young ages (5–10 years old), but male strobili are mostly found on older trees. Furthermore, estimates of the age of substantial seed production in eastern white pine range from 20 to 30 years old, to over 50 years old (Lancaster and Leak 1978; Wendel and Smith 1990). In addition, the age of sexual maturity may be younger than the age at which an individual produces enough gametes to effectively contribute to the gene pool. Thus, just as the concept of an "effective population size" (Chambers 1983) accounts for the fact that not all members of a population equally contribute gametic material to the next generation, a concept of an "effective age of reproductive maturity" should account for the likelihood that an individual's contribution to the future gene pool varies with age. The results of comparisons between pairwise relatedness and pairwise age differences indicated that significant positive relatedness occurred at age differences greater than 30 years (Fig. 6). Thus, the effective age of reproductive maturity (i.e., the functional generation time) for this stand was approximately 30 years. This empirical estimate corresponds favorably with the generalized expected ages based on seed production criteria.

Management implications

By affecting processes that shape genetic structure (including natural selection, isolation by distance, and generation time), management could alter the genetics of forest stands.

For example, in theory the gene pools of forest species could be influenced by human selection and removal of trees (and genes) through timber harvesting (Smith et al. 1997). Through the differential removal of trees of certain age, size, or quality characteristics from the reproductive pool, harvests might alter the genetic structure of the residual population in ways distinct from natural selection.

As a test of the possibility that selective tree removals might alter the genetics of the residual stand, we performed simulated harvests on our GIS-based data set using various tree characteristics (e.g., DBH, age, crown class, and basal area of the last 10 years of growth) as selection criteria. By recalculating levels of heterozygosity and rare alleles in residual stands after specific selection harvests and comparing these to the distribution of genetic measures following 1000 random harvests of equal intensity (using Monte Carlo simulations, where the number of trees in all residual populations compared were held constant so there was no confounding of differential sample size), we evaluated the relative influence of selection on these genetic parameters. Most simulations, including a harvest prescription marked by a Vermont state county forester, left a residual stand that did not significantly alter genetic diversity. However, harvests that targeted parameters associated with genetic structuring were a noted exception to this trend (Table 4). For example, simulated harvests that removed the oldest individuals significantly reduced mean observed heterozygosity more than harvests that removed the same number of trees at random ($p \leq 0.02$ for harvests that removed 75% of the trees, retaining only the youngest 25% of trees). This was especially true when the wolf trees (which are not likely to be cut in a commercial harvest) were left ($p \leq 0.05$ for a 43% harvest intensity, $p \leq 0.01$ for a 68% harvest intensity). These same simulations (i.e., cutting the oldest, but leaving the wolf trees) also significantly reduced the mean number of uncommon alleles ($p \leq 0.03$ for a 43% harvest intensity, $p \leq 0.06$ for a 68% harvest intensity).

In general, simulated harvests that manipulated patterns of genetic structuring had the potential to either increase or decrease genetic diversity measures depending on the selection criteria employed (Table 4). Some of these alterations have theoretical implications for stand fitness. For example, it is generally assumed that rare alleles convey some reduction in fitness and are selected against (Rajora and Mosseler 2001). Considering this, the artificial amplification of the frequency of rare alleles could reduce stand fitness in the near term assuming stable environmental conditions (Hawley et al. 2005). However, because they can act as a reserve for altered gene expression that may be of adaptive benefit under new environmental and competitive conditions, rare alleles may also be important to long-term population and species survival (Bergmann et al. 1990; Buchert et al. 1997; Rajora et al. 2002). Thus, the loss of rare alleles could reduce options for adaptation to changing conditions, especially if the gene flow from outside the stand is limited.

Whatever the implications of our simulated harvests, it is important to note that published reports from field-based silvicultural studies provide some support for the possibility that selective harvesting can influence genetic reserves. Neale (1985) and Neale and Adams (1985) reported that shelterwood harvesting had little impact on the genetic structure of Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) stands. They attributed this resiliency to the high levels of genetic diversity inherent to Douglas-fir, and the fact that the species occurs in almost pure, high-density stands. In contrast, Cheliak et al. (1988) found that phenotypically selected white spruce (*Picea glauca* (Moench) Voss) trees had on average only 75% of the genes found among randomly selected trees from the same population. For coastal Douglas-fir, Adams et al. (1998) reported fewer alleles in a shelterwood treatment (that preferentially removed low quality trees) relative to an unmanaged control. The alleles absent from the shelterwood stand were rare in the control stand (Adams et al. 1998). Phenotypically based harvests have also been associated with shifts in allelic frequency for eastern hemlock (*Tsuga canadensis* (L.) Carrière) (Hawley et al. 2005). Here, when smaller trees with inferior phenotypes were preferentially removed, the number of rare alleles was lower than the control stand. In contrast, alleles that were rare in the control existed at a higher frequency in a diameter-limit cut that preferentially retained small and poor quality trees (Hawley et al. 2005).

Valuable for comparisons to our simulated harvests, significant alterations in genetic diversity following actual harvests have also been noted for old growth eastern white pine (Buchert et al. 1997). Here, a 75% reduction in tree density was accompanied by a 25% decrease in the number of alleles, a 33% reduction in the proportion of polymorphic loci, a 40% loss of low frequency alleles, and an 80% decrease in the number of rare alleles. It was estimated that the latent genetic potential of these stands was reduced by about 50%, and that this loss likely compromised the ability of these gene pools to adapt to changing environmental conditions.

Table 4. The effects of simulated harvests on genetic diversity parameters.

Trees removed in simulated harvest	*n* residual	ΔHo (*t* test)	ΔHe (MCS)	ΔHe (*t* test)	Δmean no. uncommon alleles (MCS)	Mean Ho per locus	Mean He per locus	Mean no. uncommon alleles/tree
Original population	220					0.113	0.114	0.973
Oldest 25%	165					0.113	0.113	0.988
Oldest 50%	110		↓	↓↓	↓	0.107	0.103	0.809
Oldest 75%	55	↓	↓↓	↓↓		0.100	0.097	0.800
Youngest 25%	165	↑		↑↑		0.117	0.118	1.102
Youngest 50%	110	↑		↑↑		0.110	0.111	0.955
Youngest 75%	55					0.112	0.115	0.927
Biggest DBH 25%	165	↑		↓		0.110	0.110	0.903
Biggest DBH 50%	110			↓↓		0.110	0.110	0.955
Biggest DBH 75%	55					0.114	0.111	0.909
Smallest DBH 25%	165					0.113	0.115	0.994
Smallest DBH 50%	110			↑↑		0.116	0.117	1.000
Smallest DBH 75%	55	↑		↑↑		0.112	0.123	1.180
Highest BA–10 25%	165					0.115	0.115	0.952
Highest BA–10 50%	110					0.114	0.115	0.945
Highest BA–10 75%	55			↑		0.114	0.115	0.945
Lowest BA–10 25%	165			↓↓		0.113	0.112	0.982
Lowest BA–10 50%	110					0.112	0.112	1.000
Lowest BA–10 75%	55			↓		0.108	0.108	1.036
Dominant (D)	158					0.113	0.113	1.032
Codominant (C)	105	↓				0.110	0.113	1.010
Intermediate (I)	184					0.114	0.115	0.935
Suppressed (S)	215					0.113	0.114	0.953
All but D	62					0.113	0.115	0.823
All but C	115					0.116	0.115	0.940
All but I	36			↓		0.108	0.107	1.167
All but S	5				↑	0.104	0.112	1.800
D and C	43	↓		↓		0.107	0.108	1.279
D, C, and I	7				↑	0.099	0.107	1.857
All but D and C	177	↑				0.115	0.115	0.898
Oldest but leave:								
18% wolfs	181			↓		0.112	0.114	0.956
43% wolfs	126	↓	↓↓	↓↓	↓↓	0.107	0.102	0.786
68% wolfs	71	↓↓	↓↓	↓↓	↓	0.101	0.114	0.761

Note: ↑ denotes a significant increase ($p \leq 0.10$), ↑↑ denotes significant increase ($p \leq 0.05$), ↓ denotes a significant decrease ($p \leq 0.10$), and ↓↓ denotes significant decreases ($p \leq 0.05$). Paired *t* tests or Monte Carlo simulations (MCS; 1000 iterations) were used to test significance levels. Mean Ho (homozygous loci) and He (heterozygous loci) per locus standard errors were 0.025–0.028.

Microsatellite DNA analysis by Rajora et al. (2000) on samples from the same study produced similar results as those reported by Buchert et al. (1997) using allozyme analysis.

Genetic structures are naturally dynamic (Namkoong 1991), and there is a need to study their fluctuations over time to better understand how they contribute to forest health and productivity. Managed forests should be monitored and genetically compared with existing baseline data from unmanaged stands to assess the biological impacts of human manipulations (Ledig 1992). The GIS-based approach that we describe here could provide a useful framework for this analysis. Evaluations of GIS-based genetics information can highlight the importance of life history characteristics and provide a method to quantify their influence on genetic structuring. An enhanced understanding of life-history characteristics and resulting structures could assist the development of management strategies that are ecologically sustainable (Namkoong 1983). Whether harvesting, collecting seed (for research, gene conservation, or reforestation nursery stock), or setting aside land as reserves, corridors, or harvest patch retentions, knowledge of genetic structuring may prove useful in the preservation of future genetic resources.

Acknowledgments

We extend our thanks to personnel at the Ethan Allen Firing Range, Michael Snyder (Chittenden County Forester, Vermont), and countless US Forest Service personnel for

their help with field and laboratory data collection. We also thank Drs. Bryan Epperson (Michigan State University), Alex Mosseler (Canadian Forest Service), David Barrington (The University of Vermont), and two anonymous reviewers for their thoughtful reviews of an earlier version of this manuscript. The research described in this manuscript was supported in part by USDA Cooperative State Research, Education, and Extension Service McIntire–Stennis Forest Research Program funds.

References

Adams, W.T. 1992. Gene dispersal within forest tree populations. New For. **6**: 217–240.

Adams, W.T., Zuo, J., Shimizu, J.Y., and Tappeiner, J.C. 1998. Impact of alternative regeneration methods on genetic diversity in coastal Douglas-fir. For. Sci. **44**: 390–396.

Arcade, A., Faivre-Rampant, P., Le Guerroué, B., Pâques, L.E., and Prat, D. 1996. Heterozygosity and hybrid performance in larch. Theor. Appl. Genet. **93**: 1274–1281.

Beaulieu, J., and Simon, J.P. 1994. Genetic structure and variability in *Pinus strobus* in Quebec. Can. J. For. Res. **24**: 1726–1733.

Berg, E.E., and Hamrick, J.L. 1995. Fine-scale genetic structure of a turkey oak forest. Evolution, **49**: 110–120.

Bergmann, F., Gregorius, H.-R., and Larsen, J.B. 1990. Levels of genetic variation in European silver fir (*Abies alba*) — are they related to the species' decline? Genetica, **82**: 1–10.

Buchert, G.P. 1994. Genetics of white pine and implications for management and conservation. For. Chron. **70**: 427–434.

Buchert, G.P., Rajora, O.P., Hood, J.V., and Dancik, B.P. 1997. Effects of harvesting on genetic diversity in old-growth eastern white pine in Ontario, Canada. Conserv. Biol. **11**: 747–758.

Bush, R.M., and Smouse, P.E. 1991. The impact of electrophoretic genotype on life history traits in *Pinus taeda*. Evolution, **45**: 481–498.

Bush, R.M., and Smouse, P.E. 1992. Evidence for the adaptive significance of allozymes in forest trees. New For. **6**: 179–196.

Bush, R.M., Smouse, P.E., and Ledig, F.T. 1987. The fitness consequences of multiple-locus heterozygosity: the relationship between heterozygosity and growth rate in pitch pine (*Pinus rigida* Mill.). Evolution, **41**: 787–798.

Chagala, E.M. 1996. Inheritance and linkage of allozymes in *Pinus strobus* L. Silvae Genet. **45**: 181–187.

Chambers, S.M. 1983. Genetic principles for managers. *In* Genetics and conservation: a reference for managing wild animal and plant populations. *Edited by* C. Schonewald-Cox, S. Chambers, B. MacBryde, and W. Thomas. The Benjamin/Cummings Publishing Company, Inc., Menlo Park, Calif. pp. 15–50.

Cheliak, W.M., Murray, G., and Pitel, J.A. 1988. Genetic effects of phenotypic selection in white spruce. For. Ecol. Manage. **24**: 139–149.

Epperson, B.K. 1995*a*. Spatial distributions of genotypes under isolation by distance. Genetics, **140**: 1431–1440.

Epperson, B.K. 1995*b*. Fine-scale spatial structure: correlations for individual genotypes differ from those for local gene frequencies. Evolution, **49**: 1022–1026.

Epperson, B.K. 2000. Spatial genetic structure and non-equilibrium demographics within plant populations. Plant Species Biol. **15**: 269–279.

Epperson, B.K., and Chung, M.G. 2001. Spatial genetic structure of allozyme polymophisms within populations of *Pinus strobus* (Pinacea). Am. J. Bot. **88**: 1006–1010.

Goodnight, K.F., and Queller, D.C. 1999. Relatedness: a program for calculating genetic relatedness using codominant genetic markers. Available from http://www.gsoftnet.us/GSoft.html [accessed 4 March 2002].

Hawley, G.J., DeHayes, D.H., and Gage, S.F. 1988. The relationship between genetic diversity and stand viability: a case study with jack pine. *In* Proceedings of the 31st Forest Tree Improvement Conference and the Sixth North Central Tree Improvement Association, 7–8 July 1988, University Park, Pa. *Edited by* Maurice E. Demeritt, Jr. Pennsylvania State University, University Park, Pa. pp. 80–91.

Hawley, G.J., Schaberg, P.G., DeHayes, D.H., and Brissette, J.C. 2005. Silviculture alters the genetic structure of an eastern hemlock forest in Maine, USA. Can. J. For. Res. **35**: 143–150.

Jech, K.S., and Wheeler, N.C. 1984. Laboratory manual for horizontal starch gel electrophoresis. Weyerhaeuser Company, Federal Way, Wash. Weyerhaeuser Res. Rep. 050-3210.

Jelinski, D.E. 1993. Associations between environmental heterogeneity, heterozygosity, and growth rates of *Populus tremuloides* in a cordilleran landscape. Arct. Alp. Res. **25**: 183–188.

Johnson, C.W. 1980. The nature of Vermont. University Press of New England, Hanover, N.H.

Knowles, P., and Grant, M.C. 1981. Genetic patterns associated with growth variability in ponderosa pine. Am. J. Bot. **68**: 942–946.

Lancaster, K.F., and Leak, W.B. 1978. A silvicultural guide for white pine in the northeast. USDA For. Serv. Gen. Tech. Rep. NE-41.

Ledig, F.T. 1988. The conservation of diversity in forest trees: why and how should genes be conserved? BioScience, **38**: 471–479.

Ledig, F.T. 1992. Human impacts on genetic diversity in forest ecosystems. Oikos, **63**: 87–108.

Ledig, F.T., Guries, R.P., and Bonefeld, B.A. 1983. The relation of growth to heterozygosity in pitch pine. Evolution, **37**: 1227–1238.

Li, B., and Wu, R. 1996. Genetic cause of heterosis in juvenile aspen: a quantitative comparison across intra- and inter-specific hybrids. Theor. Appl. Genet. **93**: 380–391.

Li, B., Howe, G.T., and Wu, R. 1998. Developmental factors responsible for heterosis in aspen hybrids (*Polulus tremuloides* × *P. tremula*). Tree Physiol. **18**: 29–36.

Maoh, H. 2001. dmatrix_en.avx. Arcscript available from http://arcscripts.esri.com/details.asp?dbid = 10408 [accessed 1 April 2002].

Namkoong, G. 1983. Preserving natural diversity. *In* Genetics and conservation: a reference for managing wild animal and plant populations. *Edited by* C. Schonewald-Cox, S. Chambers, B. MacBryde, and W. Thomas. The Benjamin/Cummings Publishing Co. Inc., Menlo Park, Calif. pp. 317–333.

Namkoong, G. 1991. Biodiversity — issues in genetics, forestry, and ethics. For. Chron. **68**: 438–444.

Neale, D.B. 1985. Genetic implications of shelterwood regeneration of Douglas-fir in Southwest Oregon. For. Sci. **31**: 995–1005.

Neale, D.B., and Adams, W.T. 1985. Allozyme and mating-system variation in balsam fir (*Abies balsamea*) across a continuous elevational transect. Can. J. Bot. **63**: 2448–2453.

Norton, D.A., and Ogden, J. 1990. Problems in the use of tree rings in the study of forest population dynamics. *In* Methods of dendrochronology: applications in the environmental sciences. Kluwer Academic Publishers, Dordrecht, Netherlands. pp. 284–288.

Perry, D.J., and Knowles, P. 1991. Spatial genetic structure within three sugar maple (*Acer saccharum* Marsh.) stands. Heredity, **66**: 137–142.

Plessas, M.E., and Strauss, S.H. 1986. Allozyme differentiation among populations, stands, and cohorts in Monterey pine. Can. J. For. Res. **16**: 1155–1164.

Queller, D.C., and Goodnight, K.F. 1989. Estimating relatedness using genetic markers. Evolution, **43**: 258–275.

Rajora, O.P. 1999. Genetic biodiversity impacts of silvicultural practices and phenotypic selection in white spruce. Theor. Appl. Genet. **99**: 954–961.

Rajora, O.P., and Mosseler, A. 2001. Challenges and opportunities for conservation of forest genetic resources. Euphytica, **118**: 197–212.

Rajora, O.P., DeVerno, L., Mosseler, A., and Innes, D.J., 1998. Genetic diversity and population structure of disjunct Newfoundland and Central Ontario populations of eastern white pine (*Pinus strobus*). Can. J. Bot. **76**: 500–508.

Rajora, O.P., Rahman, M.H., Buchert, G.P., and Dancik, B.P. 2000. Microsatellite DNA analysis of genetic effects of harvesting in old-growth eastern white pine (*Pinus strobus*) in Ontario, Canada. Mol. Ecol. **9**: 339–348.

Rajora, O.P., Mosseler, A., and Major, J.E. 2002. Mating system and reproductive fitness traits of eastern white pine (*Pinus strobus*) in large, central versus small, isolated, marginal populations. Can. J. Bot. **80**: 1173–1184.

Riggs, L.A. 1990. Conserving genetic resources on-site in forest ecosystems. For. Ecol. Manage. **35**: 45–68.

Savolainen, O., and Karkkainen, K. 1992. Effects of forest management on gene pools. New For. **6**: 329–345.

Simon, J.L. 2002. Resampling stats (Software). Resampling Stats, Inc., Arlington, Va.

Smith, D.M., Larson, B.C., Kelty, M.J., and Ashton, P.M.S. 1997. The practice of silviculture. 9th ed. John Wiley & Sons, New York.

Smouse, P.E., and Peakall, R. 1999. Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. Heredity, **82**: 561–573.

Stokes, M.A., and Smiley, T.L. 1968. An introduction to tree ring dating. University of Chicago Press, Chicago, Ill.

Swafford, D.L., and Selander, R.B. 1989. BIOSYS-1: a computer program for the analysis of allelic variation in population genetics and biochemical systematics. Release 1.7. Illinois Natural History Survey, Champaign, Ill.

Wendel, G.W., and Smith, H.C. 1990. Silvics of North America: 1. Conifers. *Technical coordinators* R.M. Burns and B.H. Honkala. USDA For. Serv. Agric. Handb. 654. Available at http://www.na.fs.fed.us/spfo/pubs/silvics_manual/Volume_1/pinus/strobus.htm [accessed 22 May 2002].

Wright, S. 1969. Evolution and the genetics of populations: a treatise in three volumes. Vol. 2. The theory of gene frequencies. The University of Chicago Press, Chicago, Ill. pp. 290–344.