Use of Molecular Genetic Markers in Forest Management

Craig S. Echt[1]

Abstract.—When managing forests for biodiversity or sustainability, attention must be given to how silvicultural practices affect genetic diversity. A new generation of DNA-based markers affords a greater detail of genetic analysis than previously possible. These new markers, SSRs or microsatellites, have been used to demonstrate genetic diversity and infer evolutionary history of red pine, something that has not been possible with other markers. SSR markers developed by the Forest Service Research Biotechnology Unit are also being used to monitor how methods of sustainable timber management affect genetic diversity and breeding patterns within white pine stands on the Menominee Indian reservation.

INTRODUCTION

Often the goal of silvicultural prescriptions is nothing less than the management, or manipulation, of ecosystems. The importance of considering the effects on the genetic diversity of forest species when making such management decisions has been expertly reviewed (Conkle 1992; DeWald and Mahalovich 1997; El-Kassaby 1992; Ledig 1988; Li et al. 1992; Millar and Westfall 1992; Namkoong 1991; Savolainen and Kärkkäinen 1992; Yang and Yeh 1992). Maintenance of genetic diversity is also part of the biodiversity standards that have recently been proposed for forest plantations (Spellerberg and Sawyer 1996). The most efficient way to monitor natural or managed changes in genetic diversity is with protein or DNA-based (molecular) markers that are neutral with respect to natural selection. (Definitions of various terms, such as 'genetic marker', are provided in the final section of this paper.) A recently discovered class of DNA-based markers promises efficient and thorough assessment of alterations in genetic diversity for select forest species. In this presentation, progress in the application of these markers to address issues in forest management will be described, as will the approaches being used to communicate the role of forest research to those involved or interested in resource management.

SSR MARKERS IN FOREST GENETICS

Microsatellites, or simple sequence DNA repeats (SSRs), came to prominence in the field of genetics only during this past decade, due in large part to their medical applications in human disease research and DNA fingerprinting. But SSRs are highly abundant and variable in most organisms, not just humans, and thus serve as a universal source of highly informative genetic markers. In pine genomes, for example, there are several hundred thousand SSR sites (Echt and May-Marquardt 1997). It is expensive and time-consuming to develop the molecular genetic information needed to use SSRs as genetic markers, but several forest research laboratories are involved in the process for pines (Echt et al. 1996; Fisher et al. 1996; Kostia et al. 1995; Smith and Devey 1994), spruces (Pfeiffer et al. 1997; VandeVen and McNicole 1996), and oaks (Dow et al. 1995). Preliminary results from research in the author's laboratory indicate that SSR markers developed in one species, can be used in closely related species, thus leveraging marker development investments. The major advantages of using SSR markers over other types of markers, such as isozymes, RAPDs, or RFLPs, is that they generally have a large number of alleles at a locus, allele identification is unambiguous, heterozygosity is easily determined, they can be used among all members of a species, and they are quickly and efficiently analyzed from very small amounts of plant tissue. Given the increased resolution of genetic discrimination possible with SSR markers, they can be considered the "Hubble telescope" of genetics research.

In addition to their presence in the nuclear genome, SSRs are also found in the DNA of chloroplasts, and can serve as highly informative organellar markers. A number of chloroplast SSR (cpSSR) markers are available for use in conifers (Cato and Richardson 1996; Powell et al. 1995; Vendramin et al. 1996). Since in many species chloroplast DNA is uniparentally inherited (in conifers it is transmitted through pollen), cpSSR markers can provide information about evolutionary lines of decent among populations.

Fundamental information about the SSR markers and their uses can be found on-line through the Dendrome Forest Genetics World Wide Web server. This information is provided and updated by the author as a service to the forest research community. The URL address for white pine SSR markers is: http://s27w007.pswfs.gov/Data/echt_ssr_primers.html, that for hard pine SSR markers is: http://s27w007.pswfs.gov/Data/chloroplast.html, and information about cpSSR markers can be found at: http://s27w007.pswfs.gov/Data/ hardssr.html.

RED PINE POPULATION DIVERSITY AND EVOLUTIONARY HISTORY

Background

Genetic diversity of red pine, *Pinus resinosa* Ait., is extremely low, perhaps dangerously low, throughout its range. Red pine is an important timber species in the northcentral and northeastern United States, as well as in Canada. In Minnesota, Wisconsin and Michigan alone there are almost 1.8 million acres of red pine, valued at over $3 billion, and this area is only a fraction of the North American range of the species (Leatherberry et al. 1996; Miles et al. 1995; Spencer et al. 1988). Knowledge of the amount and distribution of genetic diversity of red pine is critical to management of this important natural resource.

[1]Research Plant Molecular Geneticist, Biotechnology Unit, U.S. Department of Agriculture, Forest Service, North Central Experiment Station, Rhinelander, WI

Table 1.—Locations of red pine populations

code	provenance	seedlot#	lat./long.	Elevation.(m)
A	Nova Scotia, Beaver Lake	7010280	44.14/65.20	140
B	Ontario, Eldridge Twnshp	7030250	47.00/79.30	310
D	Ontario, Sioux Lookout	6830060	50.04/91.57	370
E	Quebec, Norway Bay	7023040	45.32/76.26	80
F	Michigan, Delta Co.	5780350	46.00/87.00	NA
G	New Brunswick, Tracy	7010310	45.43/66.42	60
H	Ontario, Macdiarmid	7030260	49.18/88.05	370

The high level of homozygosity observed in red pine is thought to have resulted from one, or a series of, population bottlenecks. The most recent drastic decrease in population size (a population bottleneck) to have affected the species as a whole is thought to have occurred during the last Pleistocene glaciation 20,000 years ago, when red pine was restricted to refugial populations in the Appalachian highlands of present day West Virginia (Fowler and Morris 1977). The disjunct, dispersed populations found throughout the species' current range promote inbreeding which further increases homozygosity (Fowler and Lester 1970; Mosseler 1992).

Morphological and phenological uniformity are characteristic of the species (Fowler 1964; Fowler and Lester 1970). While some variation among provenances has been reported, it is much less than what is observed for other northern pines and the heritability of the variation has not been established (Fowler 1965; Wright et al. 1972; Ager et al. 1983). A narrow genetic base puts red pine at risk of extinction from exotic pests, disease epidemics, or rapid climate change, and can be further eroded by mismanagement. Thus there is a need to efficiently identify sources of genetic diversity so that divergent germplasm may be preserved, both in the forests and in seed orchards where it could be utilized in tree improvement programs.

Results

Since no other marker system had previously revealed genetic variation in red pine, cpSSR markers were used to survey 159 individuals among seven populations distributed across the natural range of the species (Table 1). As expected, the higher variability of cpSSRs allowed population differences to become evident. Measures of within population diversity (Table 2) indicated that a population in Tracy, New Brunswick harbored more chloroplast haplotype variability than any other population, and that no two populations were genetically identical. Examination of the genetic relationships among populations, and of the distribution of haplotype differences among individuals within populations, indicated that red pine, as a species, recovered rapidly from a population bottleneck. These results lend strong support to the previously formulated population bottleneck theory, but what was not so evident from previous observations, and what the cpSSR data clearly indicated, was that individual populations of red pine arose at different evolutionary times and possibly from different lines of decent.

Table 2.—Measures of cpSSR haplotype variation within populations

Population	n_e	f_a	H_e
A	1.430	0.833	0.314
B	2.153	0.667	0.559
D	2.998	0.545	0.698
E	2.924	0.429	0.691
F	1.754	0.750	0.449
G	8.397	0.217	0.920
H	1.497	0.810	0.348

n_e, Effective number of haplotypes in each population
f_a, Frequency of the most common haplotype
H_e, measure of unbiased genetic diversity

Red pine is thus not a genetically homogeneous specie, at least where the evolutionary origins of individual populations are concerned. More extensive cpSSR surveys should identify populations throughout the range of red pine that have the highest levels of diversity, and thus help set guidelines for genetic resource conservation programs of red pine. Genotyping of populations with nuclear SSR markers will be needed to assess the degree of population genetic differentiation of genes in the nuclear genome. Since the chloroplast microsatellite approach revealed population genetic differences in a species characterized by no detectable allozyme variation, it should also be considered for studying population structures of other forest species that have low genetic diversity, such as Torrey pine, *Pinus torreyana* Parry ex Carr, (Ledig and Conkle 1982) or western red cedar, *Thuja plicata* Donn ex E. Don., (Copes 1981).

GENETIC DIVERSITY OF NATURAL AND MANAGED WHITE PINE STANDS

Background

To achieve sustainable and ecologically sound white pine management there is a need to know whether certain practices maintain native levels of genetic diversity, or whether they narrow that diversity and foster greater inbreeding within managed stands. When artificial reforestation is used, it is also of benefit to know whether there are genetic subdivisions, or subpopulations, of trees within a management area. Such knowledge is useful for establishing seed transfer guidelines.

A previous study of eastern white pine populations in Quebec found that there are high levels of isozyme diversity, high gene flow among populations and no measurable population genetic differentiation across the range (Beaulieu and Simon 1994). In an isozyme study of eastern white pine mating systems, essentially no excess of inbreeding was found in two natural Quebec populations that differed in their stand densities of from 800 to 100 trees/acre (Beaulieu and Simon 1995).

Table 3.—Populations samples from white pine stands and site characteristics

stand	overstory (density)	cone seeds collected	regen., natural	regen., artificial
SP	~240 yr (2/ac)	-	-	-
PP	~160 yr (40/ac)	-	-	-
MC	~160 yr (3/ac)	√	40 yr	-
C1	~160 yr (40/ac)	√	3 yr	-
OL	~160 yr (20/ac)	-	9 yr	-
EL	none	-	-	8 yr

Ongoing research in the Biotechnology Unit at Rhinelander involves assessing the genetic diversity of white pine, *Pinus strobus* L. A study is in progress with white pines managed by Menominee Tribal Enterprises in Menominee County, Wisconsin, to determine how certain silvicultural practices affect genetic diversity across the 234,000-acre reservation. Instead of isozymes, SSR markers developed by Echt et al.(1996) and cpSSR markers developed by Vendramin et al. (1996) are being used.

Here are some of the questions that the Menominee foresters hope to have answered by this research: Is gene diversity distributed equally among populations across the county? Does genetic variation and gene flow change under different management strategies? How does overstory density affect genetic diversity and degree of inbreeding among the regenerated progeny? Is off-site seed genetically equivalent to local seed sources?

Besides assisting forest managers in long-term planning of tribal resources, the Menominee study will supply additional basic information on white pine diversity in the North Central Lake States region. It is expected that regional white pine information eventually will be combined with data from eastern U.S. populations, and from similar studies occurring in Canada, to construct a continental database for white pine population genetic diversity and structure. Such information could be used in formulating area-wide management and planning decisions.

The Menominee Study Plan

Six sites, involving nine populations of individuals, were selected for study by Dan Pubanz, a Menominee Tribal Enterprises forester (Figure 1). Five of the site are actively managed, and the sixth (School Pines - SP) is representative of a remnant, natural population that has not been thinned. Of the five managed stands, one (East Line Plantation - EL) was artificially regenerated from off-reservation seedlings, three sites (Potato Patch - PP, Camp One - C1, Oconto Line - OL) are under shelterwood management, and one (Minnow Creek - MC) is a pine release management site. Age classes of the overstory and regenerated populations are provided in Table 3. Fifty individuals from both the overstory and regenerated generations at each site will be sampled. From the same 50 mature trees at the MC and C1 sites, which differ ten-fold in their overstory densities, seed will also be collected to estimate pollen flow and current levels of inbreeding and outcrossing that are occurring within the sparse overstory. GPS coordinates were obtained for all the trees sampled, so spatial genetic analyses will be possible to look for patterns in genetic differences within and among populations. Each tree, seedling, and seed will be genotyped for 10 nuclear SSR loci and haplotyped for 10 cpSSR loci. Statistical analyses will be performed to quantify genetic diversity and levels of inbreeding, both within and among individual populations. The nuclear marker data from trees and seedlings will provide information on how certain silvicultural practices affect levels of heterozygosity, inbreeding and gene diversity, while the nuclear SSR marker and cpSSR marker data from seeds and seedlings will provide information on patterns of pollen flow. Data from cpSSR haplotypes of trees will also be used to look at historic patterns of dispersal of white pine. Once the Menominee study is complete, the tribal foresters hope to use the genetic diversity information as part of the biodiversity component of their timber recertification process with Scientific Certification Systems.

Results

To date, DNA has been isolated from 450 trees and genotyping has been done for two SSR loci. Forty seeds will be collected from each of 100 trees in the fall of 1997. Although no firm conclusions can be drawn from just two loci, the preliminary results demonstrate the type of information that genetic surveying with SSR markers can provide silviculturalists.

The two SSR loci each had 14 alleles, which resulted in total gene diversity being high across the reservation ($H_t = 0.73$). There was very little genetic differentiation among stands ($G_{st} = 2\%$), meaning that diversity across the reservation was not highly structured. Even so, there were small, but significant genetic differences separating each stand such that genetic distance trees (phenograms) could be constructed to represent the genetic relationships of populations to each other. Naturally regenerated progenies, as expected, were genetically most similar to their parents, with the interesting exception of the Minnow Creek pine release management site. At that site the regenerated saplings had a distinctly different genetic makeup from the overstory trees that were left after thinning. The reason for this difference is not known and additional study is needed. Off-site seedlings (EL) were a bit less diverse than, but were genetically quite similar to, on-site seedlings.

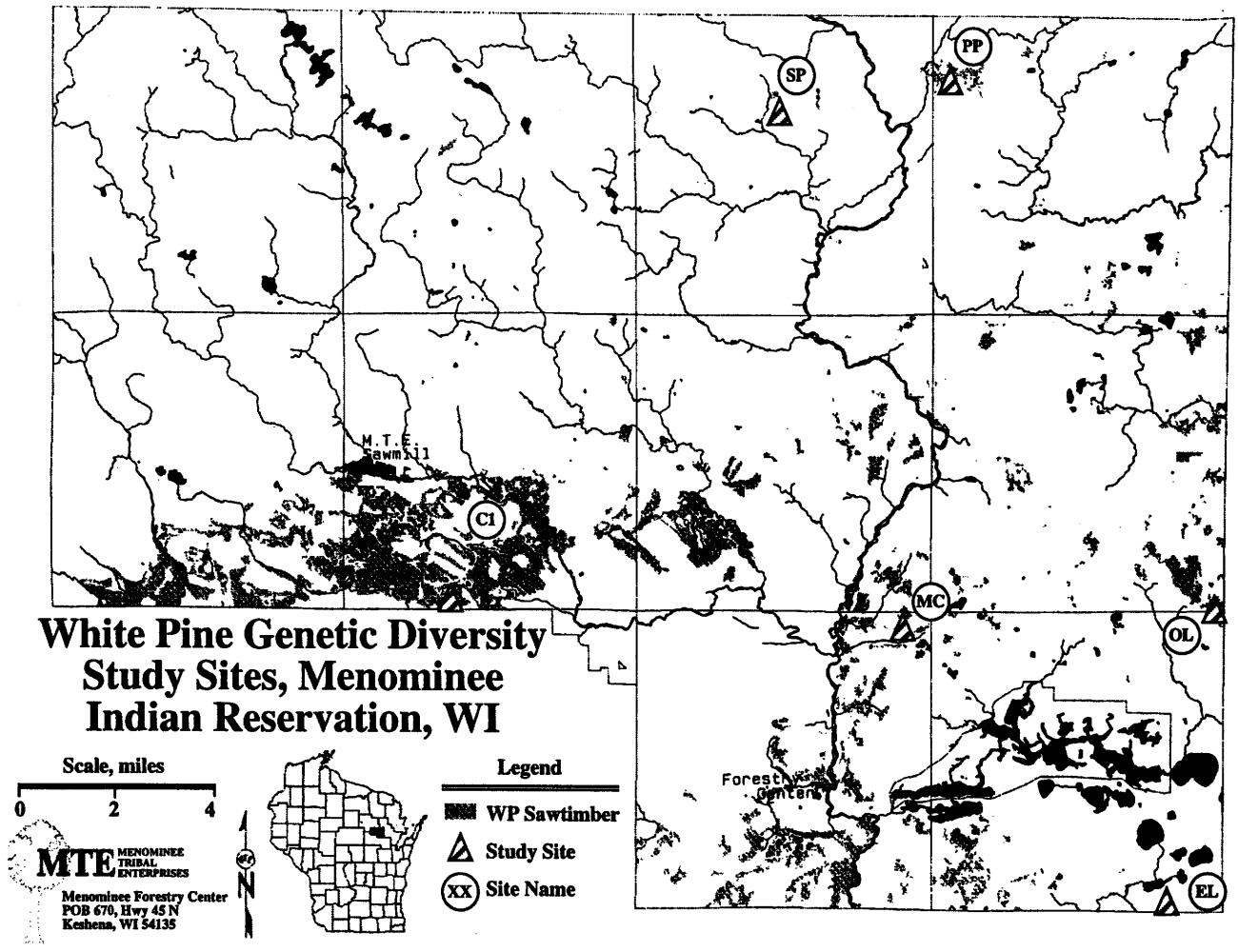

Figure 1.—Site map of study areas, and sawtimber white pine range, in Menominee County, Wisconsin.

Again based on data from just two loci, there was a general, but slight, deficiency of heterozygotes across the reservation (F_{it} = 0.20), suggesting that there has been more mating among relatives (inbreeding) than would have been expected in purely randomly mating populations. Inbreeding was always more pronounced among the parental generations. That is, there appeared to have been more inbreeding occuring 160 to 240 years ago than is occuring now in the managed stands. Since the regenerated seedling populations were less inbreed (more heterozygous) than their parents, Menominee silvicultural practices appear to be preserving, or even increasing, genetic diversity within managed white pine stands.

These preliminary results stand in contrast to a study of white pines in Ontario that used 18 isozyme loci, in which there were an average of only 2 alleles per locus, and gene diversity was about one sixth of that in the Menominee forests (Beaulieu and Simon 1994). While that study found a similar distribution of gene diversity within and among populations, the Ontario populations did not demonstrate an appreciable degree of inbreeding. Whether these differences result from differences between the information provided by isozyme and SSR markers, or from true biological differences between the populations, remains to be seen.

DEFINITIONS OF TERMS

Genome

The total genetic component of an organelle, individual or species. If only the word 'genome' is used, then it refers to all the DNA present in the nucleus. 'Chloroplast genome' and 'mitochondrial genome' refer to the total DNA present in those particular organelles.

Genotype, Haplotype and Phenotype

A genotype is an abstract, symbolic expression of the genetic factors (genes or loci) responsible for a phenotype. Genotyping is the process of determining the specific alleles that are present in an individual or population. While genotypes are expressions of the genetic constitution of diploids, where both sets of chromosomes are considered, haplotypes are expressions of the genetic constitution of haploids, where only one set of chromosomes (either the mother's or the father's) are considered. Scoring alleles in leaves or roots will give you genotypes, while scoring alleles in conifer megagametophyte or in chloroplast DNA, will give you haplotypes. A phenotype is an observable, heritable character, and is the physical aspect of the underlying

genetic factors. Once the genetic factors are identified, every phenotype can be symbolically represented by a genotype. A band on a gel, a cellular metabolite, or the angle of a branch, can all be considered phenotypes, as long as they are heritable characters.

Markers, Alleles and Loci

A marker is a quantifiable character that distinguishes, or marks, underlying genetic differences between individuals. It is often encountered as a particular enzyme or DNA fragment having a defined position on an electrophoretic gel. Any given 'band on a gel' does not gain marker status, however, unless two or more forms (alleles) of it exist at a single chromosomal location (locus). The plural of locus is loci. At their most fundamental level alleles are simply the DNA sequence variants found at a locus.

Oligonucleotide, Primer

DNA is a polymer of nucleotides, or a polynucleotide. A short piece of DNA is an oligonucleotide. An oligonucleotides usually refers to a piece of single stranded DNA, although it can be double-stranded, and is generally 10-100 nucleotides in length. An oligonucleotide primer, or simply, primer, is a single stranded oligonucleotide that anneals to a complementary sequence on single stranded DNA (the template DNA) and directs, or primes, the synthetic action of a class of enzymes called DNA polymerases. When two primers are used to direct the amplification of a locus by PCR they are called PCR primer pairs.

PCR (Polymerase Chain Reaction)

A Nobel Prize-winning technique for replicating and amplifying (cloning) specific DNA fragments in a test tube. The key component of PCR is a heat stable DNA polymerase called Taq polymerase, after the hot springs bacteria *Thermus aquaticus* from which it is purified. The extent of DNA that is amplified is determined by the two PCR primers which anneal to opposite DNA strands at the ends of the target DNA segment. In theory, millions of copies of target DNA can be amplified from a single DNA template molecule (a single cell!), but in common practice the starting point is several hundred copies of the DNA template. Thus, from just several hundred cells, a specific locus can be genotyped with a PCR-based marker. In pines, a typical PCR marker of 200 base pairs in length represents 1/100 millionth of the total length of DNA present in the nucleus. It is this discriminatory power of PCR that makes it the most powerful tool available for genetic analysis.

SSR (simple sequence repeat), or microsatellite DNA

A class of tandemly repeated DNA sequences that are highly variable, and which are used extensively as genetic markers. Examples of SSRs are $(AG)_{10}$ - the sequence of nucleotides deoxyAdenosine and deoxyGuanosine repeated 10 times - or $(ACT)_8$ - the sequence deoxyAdenosine, deoxyCytidine and deoxyThymidine repeated 8 times. SSR repeats are found in abundance throughout the nuclear DNA of most organisms. Each repeat that serves as a SSR marker is surrounded by unique, non-repetitive DNA. It is this unique DNA that allows identification of individual SSR loci along the chromosomes, while the embedded repeat DNA provides the informative variation. SSR loci can be analyzed by PCR once the DNA sequence of the repeat and its surrounding DNA is determined. These DNA sequences allow the unique PCR primer pairs to be designed and chemically synthesized.

ACKNOWLEDGMENT

This work would not have been possible without the contributions from my collaborators on the red pine study, L. L. DeVerno of the Canadian Forest Service, and G. G. Vendramin and M. Anzidei of the National Research Council of Italy, or from my collaborators on the Menominee white pine study, Dan Pubanz of Menominee Tribal Enterprises and Dave Congas of the Bureau of Indian Affairs. The expert technical assistance of Paula Marquardt and Cheryl Zobel of the Rhinelander Forestry Sciences Lab has been indispensable.

LITERATURE CITED

Ager, A., Guries, R., Lee, C. H. 1983. **Genetic gains from red pine seedling seed orchards.** Proceedings of the 28th Northeastern Forest Tree Improvement Conference, pp175-194. Institute of Natural and Environmental Resources, Univ. of New Hampshire, Durham, NH. July 7-9.

Beaulieu, J., and Simon, J. P. 1994. **Genetic structure and variability in *Pinus strobus* in Quebec.** Canadian Journal of Forest Research, 24:1726-1733.

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

Cato, S. A., and Richardson, T. E. 1996. **Inter- and intraspecific polymorphism at chloroplast SSR loci and the inheritance of plastids in *Pinus radiata*, D. Don.** Theoretical and Applied Genetics, 93:587-592.

Copes, D. L. 1981. **Isoenzyme uniformity in western red cedar seedlings from Oregon and Washington.** Canadian Journal of Forest Research, 11:451-453.

DeWald, L. E., and Mahalovich, M. F. 1997. **The role of forest genetics in managing ecosystems.** The Forestry Journal, 95:12-16.

Dow, B. D., Ashley, M. V., and Howe, H. F. 1995. **Characterization of highly variable (GA/CT)n microsatellites in the bur oak, *Quercus macrocarpa*.** Theoretical and Applied Genetics, 91:137-141.

Echt, C. S., and May-Marquardt, P. 1997. **Survey of microsatellite DNA in pine.** Genome, 40:9-17.

Echt, C. S., May-Marquardt, P., Hseih, M., and Zahorchak, R. 1996. **Characterization of microsatellite markers in eastern white pine.** Genome, 39:1102-1108.

El-Kassaby, Y. A. 1992. **Domestication and genetic diversity - should we be concerned?** Forestry Chronicle, 68: 687-700.

Fowler, D. P. 1964. **Effects of inbreeding in red pine, *Pinus resinosa* Ait. I. Natural variation.** Silvae Genetica, 13, 170-177.

Fowler, D. P. 1965. **Effects of inbreeding in red pine, *Pinus resinosa* Ait. IV. Comparison with other northeastern *Pinus* species.** Silvae Genetica, 14, 76-81.

Fowler, D. P. And Lester, D. T. 1970. **The genetics of red pine.** USDA Forest Service Research Paper W-8, pp 13.

Fowler, D. P. and Morris, R. W. 1977. **Genetic diversity in red pine: evidence for low genic heterozygosity.** Canadian Journal of Forest Research, 7:343-347.

Fisher, P. J., Gardner, R. C., and Richardson, T. E. 1996. **Single locus microsatellites isolated using 5' anchored PCR.** Nucleic Acids Research, 24:4369-4372.

Kostia, S., Varvio, S., Vakkari, P., and Pulkkinen, P. 1995. **Microsatellite sequences in a conifer, *Pinus sylvestris*.** Genome, 38:1244-1248.

Leatherberry, E. C. and Spencer, J. S., Jr. 1996. **Michigan Forest Statistics, 1993.** USDA Forest Service. Resource Bulletin NC-170.

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. and Conckle, M. T. 1983. **Gene diversity and genetic structure in a narrow endemic, Torrey pine (*Pinus torreyana* Parry ex. Carr.).** Evolution, 37:79-85.

Li, P., MacKay, J., and Bousquet, J. 1992. **Genetic diversity in Canadian hardwoods: implications for conservation.** Forestry Chronicle, 68: 709-719.

Miles, P. C. Chen, C. M., and Leatherberry, E. C. 1995. **Minnesota Forest Statistics, 1990, Revised.** USDA Forest Service, Resource Bulletin NC-158.

Millar, C. I., and Westfall, R. D. 1992. **Allozyme markers in forest genetic conservation.** New Forests, 6:347-371.

Mosseler, A. 1992. **Life history and genetic diversity in red pine: implications for gene conservation in forestry.** The Forestry Chronicle, 68:701-708.

Namkoong, G. 1991. **Biodiversity - issues in genetics, forestry and ethics.** The Forestry Chronicle, 68:438-443.

Savolainen, O., and Kärkkäinen, K. 1992. **Effect of forest management on gene pools.** New Forests, 6:329-345.

Smith, D. N., and Devey, M. E. 1994. **Occurrence and inheritance of microsatellites in *Pinus radiata*.** Genome, 37:977-983.

Spellerberg, I. F., and Sawyer, J. W. D. 1996. **Standards for biodiversity:A proposal based on biodiversity standards for forest plantations.** Biodiversity and Conservation, 5:447-459.

Spencer, J. S., Jr., Smith, W. B., Hahn, J. T., and Raile, G. K. 1988. **Wisconsin's Fourth Forest Inventory 1983.** USDA Forest Service, Resource Bulletin NC-107.

Pfeiffer, A., Olivieri, A., and Morgante, M. 1997. **Identification and characterization of microsatellites in Norway spruce (*Picea abies* K.) Genome, 40:** in press.

Powell, W., Morgante, M., McDevitt, R., Vendramin, G., and Rafalski, J.A. 1995. **Polymorphic simple sequence repeat regions in chloroplast genomes: Applications to the population genetics of pines.** Proc. Natl. Acad. Sci. USA, 92:7759-7763.

Wright, J. W., Read, R. A., Lester, D. T., Merritt, C., and Mohn, C. 1972. **Geographic variation in red pine: 11-year data from the North Central states.** Silvae Genetica, 21:205-222.

VandeVen, W. G. and McNicole, R. J. 1996. **Microsatellites as DNA markers in Sitka spruce.** Theoretical and Applied Genetics, 93:613-617.

Vendramin, G. G., Lelli, L., Rossi, P., and Morgante, M. 1996. **A set of primers for the amplification of 20 chloroplast microsatellites in *Pinaceae*.** Molecular Ecology, 5:595-598.

Yang, R., and Yeh, F.C. 1992. **Genetic consequences of *in situ* and *ex situ* conservation of forest trees.** Forestry Chronicle, 68:720-729.