Whole-tree silvic identifications and the microsatellite genetic structure of a red oak species complex in an Indiana old-growth forest

Preston R. Aldrich, George R. Parker, Charles H. Michler, and Jeanne Romero-Severson

Abstract: The red oaks (*Quercus* section *Lobatae*) include important timber species, but we know little about their gene pools. Red oak species can be difficult to identify, possibly because of extensive interspecific hybridization, although most evidence of this is morphological. We used 15 microsatellite loci to examine the genetic composition of a red oak community in 20.6 ha of an Indiana old-growth forest. The community included northern red oak (*Quercus rubra* L.), Shumard oak (*Quercus shumardii* Buckl.), and pin oak (*Quercus palustris* Muenchh.). Species were identified using whole-tree silvic characters, the approach most often implemented by foresters. We found high genetic diversity within species but limited genetic differences between species. Phenetic clustering showed that *Q. rubra* and *Q. shumardii* were more genetically similar than either was to *Q. palustris*, but a neighbor-joining tree revealed that individuals of the different species did not resolve into single-species clusters. We identified four mixed-species subpopulations using Structure, a computer program based on Monte Carlo simulation. The three largest groups are consistent with the following biological interpretations: (*i*) pure *Q. rubra*, (*ii*) *Q. rubra*, *Q. shumardii*, and their hybrids, and (*iii*) *Q. rubra*, *Q. shumardii*, *Q. palustris*, and their hybrids. We discuss the implications of these findings for the whole-tree silvic approach to selection and for management of the red oak gene pool.

Résumé : Même si les chênes rouges (*Quercus* section *Lobatae*) comprennent des espèces importantes pour la production de bois, peu d'informations existent à propos de leur diversité génétique. Les espèces de chêne rouge peuvent être difficiles à reconnaître, possiblement en raison d'une hybridation interspécifique importante. Cependant, ce phénomène n'est surtout étayé que par des données morphologiques. Les auteurs ont utilisé 15 marqueurs de loci microsatellites afin d'étudier la composition génétique d'une association de chênes rouges couvrant une superficie 20,6 ha au sein d'une forêt ancienne en Indiana. L'association comprenait le chêne rouge d'Amérique (*Quercus rubra* L.), le chêne de Shumard (*Quercus shumardii* Buckl.) et le chêne des marais (*Quercus palustris* Muenchh.). Les espèces ont été reconnues à partir d'un ensemble de caractères morphologiques et sylvicoles distinctifs. Cette approche est la plus fréquemment utilisée par les forestiers. Une diversité génétique élevée a été remarquée à l'intérieur des espèces, mais les différences génétiques étaient limitées entre les espèces. L'analyse de regroupement phénétique a démontré que *Q. rubra* et *Q. shumardii* étaient plus similaires génétiquement que l'un ou l'autre avec *Q. palustris*. Cependant, le dendrogramme découlant de la méthode de « neighbor-joining » a démontré que les individus des différentes espèces ne formaient pas de groupes cohérents représentatifs de chacune des espèces. En utilisant la procédure *structure*, un programme d'ordinateur basé sur la simulation Monte Carlo, les auteurs ont identifié quatre sous-populations rassemblant des individus des différentes espèces. Les trois plus grands groupes correspondaient à la classification biologique suivante : (*i*) *Q. rubra* pur, (*ii*) *Q. rubra*, *Q. shumardii* et leurs hybrides, et (*iii*) *Q. rubra*, *Q. shumardii*, *Q. palustris* et leurs hybrides. Les auteurs discutent des implications de leurs observations sur l'approche d'identification des chênes à l'aide des caractères morphologiques et sylvicoles et ce, dans le contexte de la sélection ainsi que pour la gestion des ressources génétiques chez ce complexe d'espèces.

[Traduit par la Rédaction]

Received 20 January 2003. Accepted 3 July 2003. Published on the NRC Research Press Web site at http://cjfr.nrc.ca on 12 November 2003.

P.R. Aldrich[1] and C.H. Michler. USDA Forest Service, North Central Research Station, Hardwood Tree Improvement and Regeneration Center, Purdue University Department of Forestry and Natural Resources, West Lafayette, IN 47907-2033, U.S.A.
G.R. Parker. Purdue University Department of Forestry and Natural Resources, West Lafayette, IN 47907-2033, U.S.A.
J. Romero-Severson. Hardwood Tree Improvement and Regeneration Center, Purdue University Department of Forestry and Natural Resources, West Lafayette, IN 47907-2033, U.S.A.

[1]Corresponding author (e-mail: preston@fnr.purdue.edu).

 doi: 10.1139/X03-160 © 2003 NRC Canada

Introduction

Management of oak gene pools is encumbered by unclear species boundaries. Generally, there exists a weak correlation between any one morphological trait and the ability to discriminate species of oaks, as many species exhibit morphological tendencies or affinities rather than fixed differences (Palmer 1942; Burger 1975; Jensen 1988). It is easy to find extreme forms of an oak species that exhibit the full suite of traits as described in field manuals, but a review of numerous specimens frequently reveals many intermediate forms.

Such taxonomic uncertainties present a challenge to foresters working with large numbers of trees that must be identified rapidly in the field. Foresters usually rely on whole-tree silvic characters to select trees for seed crop and nursery stock (Tomlinson et al. 2000) and timber as well. Here, an overall phenotype is evaluated based on dendrological characters that tend to distinguish the species (some of economic interest) such as bole and crown shape, bark, branchiness, and leaf shape (e.g., Harlow et al. 1996). Although the logs of some oak species are pooled at the mill with little consequence, knowledge of gene pool structure would benefit breeding and germplasm management programs. Knowing whether the whole-tree silvic approach identifies genetically discrete species of oaks would help guide these management and improvement programs.

Molecular genetic studies have revealed several instances in which the gene pools of oak species were poorly differentiated (e.g., Manos and Fairbrothers 1987; Guttman and Weigt 1989). This null hypothesis of no genetic difference can arise from low sample sizes, low marker variation, recent speciation, or introgressive hybridization. The former two factors are readily addressed, especially with the advent of hypervariable microsatellite markers (Schlotterer 2001). The latter two factors are more difficult to disentangle and in some instances are interdependent. The shared retention of ancestral polymorphisms in recently formed species can prevent the resolution of what are otherwise independent lineages (Pamilo and Nei 1988; Edwards and Beerli 2000). But recently diverged genomes also are likely to be more tolerant of continued hybridization. In areas of sympatry, the populations of hybridizing species may come to share more alleles at a site than they do with geographically distant conspecific populations (Whittemore and Schaal 1991; Petit et al. 1997).

There is little evidence of hybrids forming between the two major groups of oaks, the red oaks (*Quercus* section *Lobatae*) and white oaks (section *Quercus*), but there is ample evidence of hybridization within each group, especially the red oaks. Guttman and Weigt (1989) surveyed 10 red and eight white oak species using 18 isozyme loci and found that eight loci clearly distinguished the red from the white oaks, but there was less genetic distance among red oak species than among white oaks. Extensive research on the morphology of red oaks (Jensen 1977, 1988, 1994; Jensen and Eshhbaugh 1976; Jensen et al. 1984, 1993) supports this model and implies that many of the species may be fairly young and that hybridization within the group is widespread. Patterns of morphological gradation between red oak species can vary greatly by site, however, and should depend heavily on the local species composition if hybridization is occurring.

One of the best-studied red oak hybrid systems involves northern red oak (*Quercus rubra* L.), a highly prized timber species, and northern pin oak (*Quercus ellipsoidalis* E.J. Hill), whose branches persist and diminish the value of its wood. Jensen et al. (1993) revealed a cline of morphometric variation between the species in populations extending from the Wisconsin mainland into the Apostle Islands. These same populations showed little isozymic differentiation (Hokanson et al. 1993). The authors concluded that introgressive hybridization, mainly from *Q. ellipsoidalis* into *Q. rubra*, was the most plausible interpretation for their data. In a separate study, Tomlinson et al. (2000) evaluated the utility of selecting seed sources in this complex using silvic characters of the maternal tree. They categorized 30 mother trees as *Q. rubra*, *Q. ellipsoidalis*, or a putative hybrid using whole-tree silvics. Subsequent isozyme analysis showed little genetic differentiation, and morphometric analyses of leaves and acorns revealed a continuum consistent with the presence of hybrids. Multivariate analysis resolved two groups (*Q. ellipsoidalis* and *Q. rubra* plus putative hybrids) rather than three, and the authors concluded that *Q. ellipsoidalis* could be culled from publically collected seed by whole-tree silvics but that hybrid germplasm may persist.

Here, we examine the microsatellite genetic structure of *Q. rubra*, pin oak (*Quercus palustris* Muenchh.), and Shumard oak (*Quercus shumardii* Buckl.) in an old-growth forest in Indiana. Morsink and Pratt (1984) described a hybrid swarm involving these three species in Ontario. At our site in Indiana, individuals were identified to species using whole-tree silvics (Prentice 1927; Parker et al. 1985), although individual trees exhibit a variety of pure species and intermediate forms that would suggest a history of hybridization in the community. We applied 15 microsatellite markers to address these specific questions: (*i*) does the whole-tree silvic approach identify species that are genetically differentiated, (*ii*) which species is the most genetically distinct, and (*iii*) what are the implications for management given the genetic structure of this red oak community?

Materials and methods

Site and species

The Davis-Purdue Research Forest (DPRF) is located in east-central Indiana (site descriptions in Parker and Leopold 1983; Parker et al. 1985). The largest stand (20.6 ha) is among the oldest of the remnant deciduous forests in the region. Exogenous disturbance has been suppressed since acquisition of the property by Purdue University in 1917. Censuses (Prentice 1927; Parker and Leopold 1983; Parker et al. 1985) have registered seven oak species in the stand with a combined importance value of 41.2% out of 25 tree species. Three red oak species were present including *Q. rubra*, *Q. shumardii*, and *Q. palustris*. Judging from gradation and overlap of morphological traits in some individuals, hybrids among these species may also have been present, namely *Quercus* ×*riparia* Laughlin (*Q. rubra* × *Q. shumardii*), *Quercus* ×*columnaris* Laughlin (*Q. rubra* × *Q. palustris*), and *Quercus* ×*mutabilis* Palmer & Steyerm. (*Q. shumardii* × *Q. palustris*). We show characteristic leaves (Fig. 1) and acorns (Fig. 2) along with some of putative hybrid origin.

© 2003 NRC Canada

Fig. 1. Representative morphological diversity of leaf shapes in the red oaks at the Davis-Purdue Research Forest. Group memberships of the individuals from which the leaves were taken reflect the whole-tree silvic and genetic (program Structure) designations as follows: A, *Quercus rubra* (R); B1–2, *Quercus shumardii* (R × S × P); B3–6, *Q. shumardii* (R × S); C, *Quercus palustris* (R × S × P); D, various species: indeterminate genetic clustering at 90% confidence level toward the center of Fig. 5.

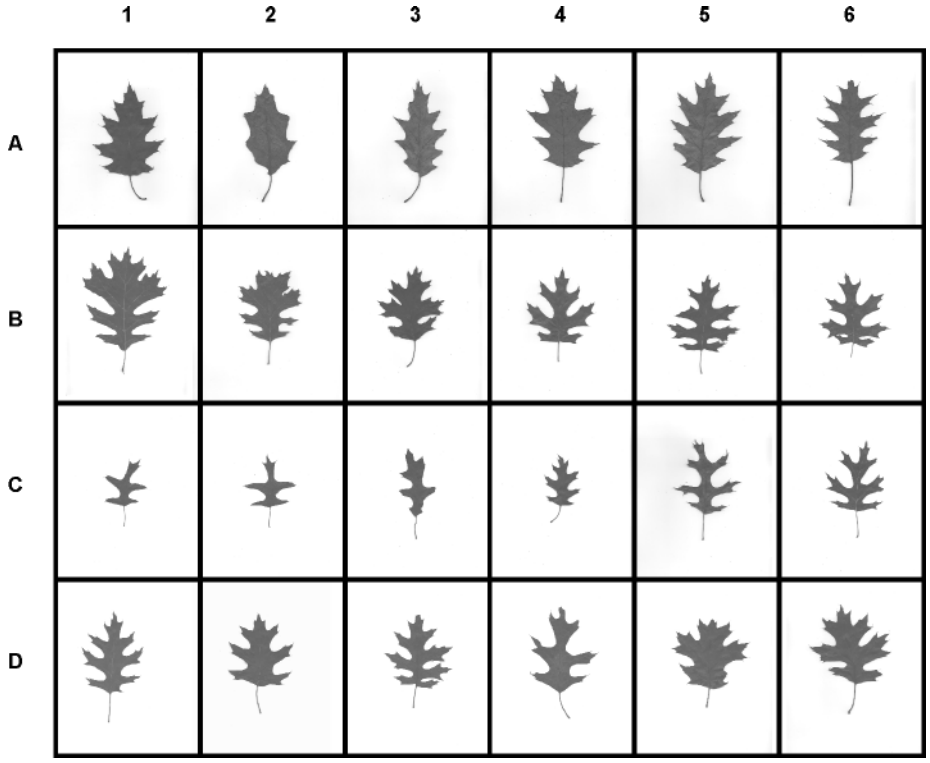
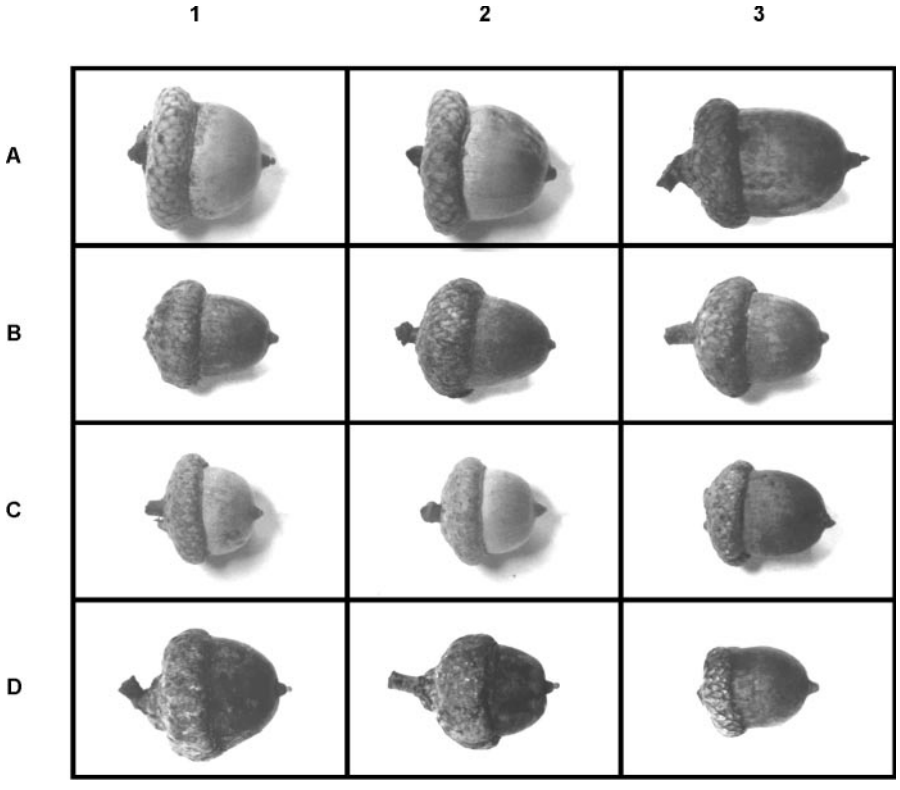


Fig. 2. Representative morphological diversity of acorn shapes in the red oaks at the Davis-Purdue Research Forest. Group memberships are as follows: A, *Quercus rubra* (R); B1, *Quercus shumardii* (R × S × P); B2–3, *Q. shumardii* (R × S); C, *Quercus palustris* (R × S × P); D, various species: indeterminate genetic clustering.



© 2003 NRC Canada

Table 1. Whole-tree silvic characters and habitat used to delineate species in the red oak community at the Davis-Purdue Research Forest.

Trait	*Quercus rubra*	*Quercus shumardii*	*Quercus palustris*
Acorn cup	Thick flat saucer	Thick flat saucer	Thin saucer
Acorn length (cm)	1.5–3	1.5–3	<1.5
Acorn shape	Long, narrow or broad	Tapers	Long as wide
Bark	Furrowed	Furrowed	Smooth
Branch habit	Lower branches prune	Lower branches prune	Lower branches persist
Crown	Rounded, bushy	Open and spreading	Pyramidal, lower drooping
Leaf buds	Large (1/4 in.), red, silky tipped	Large (1/4 in.), grey to straw colored	Small (1/8 in.), red
Leaf length (cm)	15–20	15–20	5–15
Leaf lobes	7–11, narrow	7–9, broad	5 (7), narrow
Leaf lobe tips	4 or fewer, bristle-tipped teeth	5–7 bristle-tipped, teeth	4 or fewer, bristle-tipped teeth
Leaf sinuses	Shallow, wide, 1/2–1/3 to midrib	Deep, narrow, 3/4 to midrib	Deep, wide, nearly to midrib
Soil	Mesic to xeric, well drained	Moist depressions and bottomlands	Wet, poorly drained

Trees were identified using the whole-tree silvic method (see Tomlinson et al. 2000), which is analogous to the standard selection method used by foresters in the field. Rather than judging species status based on one or two diagnostic traits, or on numerous quantitative traits that are later analyzed using multivariate analyses and computer programs (e.g., Perron and Bousquet 1997), foresters most often judge species status based on the overall fit of an individual to a suite of key traits that typically are drawn from field manuals (e.g., Fowells 1965; Little 1980; Harlow et al. 1996) and dendrology courses. In the case of the DPRF, all stems ≥10 cm diameter at breast height (DBH) were identified to species in the field during prior censuses (Prentice 1927; Parker and Leopold 1983; Parker et al. 1985) using whole-tree silvic characters (see Table 1 for red oak character states). Individuals that failed to exhibit all of the character states of a single species were assigned to the species that they most resembled. We did not explicitly consider putative hybrid classes based strictly on morphology but instead sought to evaluate whether these whole-tree silvic classes comprised mixtures and hybrids of the three species.

Sampling

We sampled 118 adults 37.4–149.0 cm DBH. According to the 1998 census, *Q. rubra* was numerically dominant in the canopy (N = 603 adults ≥10 cm DBH, 74.3% of the red oak community) followed by *Q. shumardii* (N = 108, 13.3%) and *Q. palustris* (N = 101, 12.4%). Consequently, we sampled more *Q. rubra* (N = 66, 55.9% of total sample) than *Q. shumardii* (N = 36, 30.5%) or *Q. palustris* (N = 16, 13.6%). For genotyping, we used a chisel and hammer to harvest bark cambium and sapwood (~2 cm × 6 cm, 1–2 cm thick) from the base of the trunk. Samples remained on ice until they were brought to the laboratory within 2 days.

Molecular methods

We ground sample tissue from bark cambium and sapwood in liquid nitrogen using a wood pulverizer (CertiPrep 6750 freezer/mill; Spex, Inc., Metuchen, N.J.) followed by further breakup of tissue with a metal bead in an agitating rotor (FastPrep Bio101; Savant Instruments, Inc., Holbrook, N.Y.). DNA extraction and microsatellite primer development are described in Aldrich et al. (2002). We selected 15 microsatellite loci for the study (Table 2): eight from Aldrich et al. (2002), one developed from durmast oak (*Quercus petraea* (Mattuscka) Liebl.) (*ssrQpZAG9*; Steinkellner et al. 1997), and five additional loci developed from *Q. rubra* (Aldrich et al. 2003). Template concentrations were adjusted to 5–10 ng/μL. Polymerase chain reaction (PCR) amplifications (25-μL volume) included 1× *Ex Taq* buffer (Invitrogen, Carlsbad, Calif.; proprietary except 2.0 mmol/L $MgCl_2$), 100 μmol/L dNTP each, 72 nmol/L of each upper and lower primer, 0.01 U of Takara *Ex Taq* polymerase/μL (Invitrogen), and 0.2–0.4 ng DNA/μL. PCR profiles were 94 °C for 1 min; 40–50 cycles of 94 °C for 30 s, T_a for 45 s, and 72 °C for 1.5 min; and 72 °C for 10 min. We analyzed fragment sizes on an eight-capillary CEQ2000XL genotyper (Beckman Coulter, Inc., Fullerton, Calif.) after initial screens on 1.5% agarose gels (1× TAE). Nonoverlapping loci were pooled and ethanol-precipitated. The forward primer was 5′ end labeled with phosphoramadite dyes (D2-PA, D3-PA or D4-PA; Invitrogen, Carlsbad, Calif.) allowing detection of ±1 bp size differences judged relative to a 22-band standard (60–420 bp). We called +A or –A alleles consistently within a locus, and in cases where loci exhibited stutter bands, we selected the tallest peak as the allele.

Analyses

We analyzed genetic structure using the program Genetic Data Analysis (Lewis and Zaykin 2001). We report single-locus diversity indices for the *Q. rubra* – *Q. shumardii* – *Q. palustris* complex and multilocus estimates for individual species. We tested each locus for Hardy–Weinberg (HW) disequilibrium using Fisher's exact test (3200 permutations). Inbreeding coefficients measure the extent to which the species within the complex have differentiated over time. The F (equivalent to F_{IT}; Wright 1951) measures inbreeding of individuals relative to the species complex, f (or F_{IS}) is the inbreeding in individuals relative to their species, and θ_p (or F_{ST}) represents inbreeding within species relative to the complex (Weir and Cockerham 1984; Weir 1996, chap. 5). We report locus-specific inbreeding coefficients, means, and 95% confidence intervals based on 1000 bootstrap replicates across loci. We also produced an unweighted pair group method arithmetic average (UPGMA) phenogram of the three species using Nei's (1972) genetic distance.

We employed two methods of assigning individuals to groups based on their multilocus genotypes to further judge

© 2003 NRC Canada

Table 2. Genetic diversity at 15 microsatellite loci for the entire red oak species complex (*Quercus rubra*, *Quercus shumardii*, and *Quercus palustris*).

Locus	*N*	*A*	H_e	H_o	*f*	*F*	θ_p
quru-GA-0A01	108	9	0.569	0.481	0.117*	0.177	0.068
quru-GA-0A03	109	16	0.895	0.908	–0.020	–0.011	0.009
quru-GA-0C19	109	18	0.917	0.807	0.111*	0.127	0.018
quru-GA-0I21	106	14	0.893	0.934	–0.049	–0.043	0.006
quru-GA-0M07	99	13	0.836	0.707	0.145*	0.153	0.009
quru-GA-1C06	110	26	0.880	0.818	0.069	0.071	0.003
quru-GA-1C08	99	15	0.685	0.596	0.119*	0.129	0.011
quru-GA-1D09	104	15	0.851	0.798	0.06	0.063	0.003
quru-GA-1F02	113	18	0.855	0.912	–0.068	–0.064	0.003
quru-GA-1F07	99	23	0.921	0.919	0.004	0	–0.004
quru-GA-1G13	107	14	0.753	0.729	0.033	0.032	–0.001
quru-GA-1H14	105	26	0.916	0.848	0.068	0.078	0.011
quru-GA-2G07	112	12	0.488	0.393	0.195	0.192	–0.003
quru-GA-2N03	95	17	0.820	0.832	–0.016	–0.011	0.005
ssrQpZAG9	106	20	0.847	0.877	–0.053	–0.024	0.027
Mean	105.4	17.1	0.808	0.771	0.04	0.05	0.01
Upper bound					0.078	0.088	0.019
Lower bound					0.002	0.011	0.005

Note: *N*, number of trees; *A*, number of alleles per locus; H_e, expected heterozygosity; H_o, observed heterozygosity; *f*, *F*, and θ_p, inbreeding coefficients (F_{IS}, F_{IT}, and F_{ST}, respectively). Upper and lower bounds on inbreeding coefficients are 95% confidence intervals based on 1000 bootstrap replicates.

*Rejection of HW equilibrium based on Fisher's exact test ($P < 0.05$).

the correspondence between morphologically based species determinations and genetic composition in the red oak community. First, we produced a neighbor-joining tree of individuals using the Neighbor subroutine of the program PHYLIP (Felsenstein 1989). This approach used a pairwise allele-sharing distance matrix (Mountain and Cavalli-Sforza 1997) calculated as follows. For a pair of individuals genotyped at *L* loci, add 1/*L* for every locus in which they shared no alleles, 1/(2*L*) for every locus in which they shared one allele (i.e., AA:AB or AB:AC), and 0 for every locus for which two alleles are shared (i.e., AA:AA or AB:AB). The result is an unrooted tree of individuals, labeled by species, and clustered according to genetic affinities.

We also used the program Structure (Pritchard et al. 2000), a Markov chain Monte Carlo (MCMC) clustering approach that assigns individuals to *K* subpopulations, or clusters, based on their multilocus genotypes. Individuals are assigned in a fashion that minimizes the amount of HW or gametic disequilibrium present within populations. Population allele frequencies are estimated simultaneously. Null alleles might have adverse effects on the clustering (Pritchard 2000), so we based our analysis on the 11 loci conforming to HW expectations. All runs involved 10^6 MCMC repetitions after an initial burn-in period ($N = 10^4$).

Results

Genetic structure

These 15 loci revealed fairly high levels of diversity in the red oak complex but limited genetic differentiation among the three species (Table 2). The number of alleles per locus varied from 9 to 26 (mean 17.1). There were numerous private alleles ($N = 76$), i.e., those occurring only in one species, but these were all low frequency (*Q. rubra*: $N = 45$, mean frequency = 0.0187; *Q. shumardii*: $N = 27$, mean frequency = 0.021; *Q. palustris*: $N = 3$, mean frequency = 0.049). Six loci exhibited the same most common allele in each species. The other nine loci differed among species as follows: *Q. rubra* and *Q. shumardii* carried different most common alleles at four loci, *Q. rubra* and *Q. palustris* at five loci, and *Q. shumardii* and *Q. palustris* at six loci.

Table 3. Genetic diversity for individual species based on 15 microsatellite loci.

Species	*N*	*A*	H_e	H_o	*f*
Q. rubra	58.6	14.9	0.809	0.774	0.043
Q. shumardii	32.1	12.9	0.812	0.781	0.039
Q. palustris	14.7	8.4	0.762	0.737	0.034
Mean	35.1	12.1	0.794	0.764	0.039

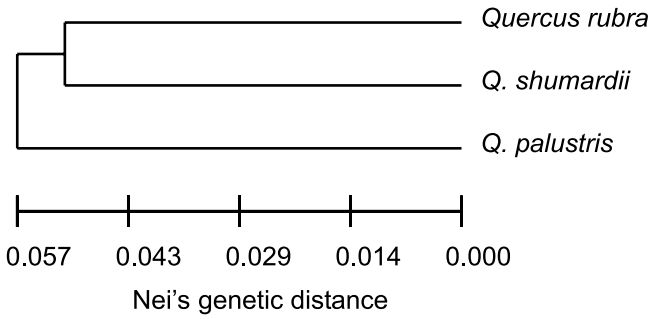
Note: See Table 2 for definitions of variables.

Observed heterozygosities were high (mean $H_o = 0.771$) and overall very similar to the expected values (mean $H_e = 0.808$) (Table 2). The levels of genetic diversity detected in each species were very similar (Table 3). The number of alleles per locus (A) is probably most sensitive to sample size, and the estimates of *A* increased with *N*, as expected. Observed and expected heterozygosities were very similar between species, and there was only a minor deficit of heterozygotes present in each species considering all 15 loci.

Over all 15 loci, inbreeding within species (mean $f = 0.040$) and within the species complex (mean $F = 0.050$) was small but significant (Table 2). Similarly, genetic differentiation among species was limited (mean $\theta_p = 0.010$) but significant. Most loci conformed to HW expectations except for four (*quru-GA-0A01*, *quru-GA-0C19*, *quru-GA-0M07*,

© 2003 NRC Canada

Fig. 3. Unweighted pair group method arithmetic average phenogram based on Nei's (1972) genetic distance showing relationships between the three red oak species according to evidence from 15 microsatellite loci.



and *quru-GA-1C08*) that registered a significant heterozygote deficit by Fisher's exact test ($P < 0.05$) (Table 2). The deficit at *quru-GA-2G07* was comparable in magnitude but did not register as significant.

Clustering methods

Phenetic clustering of species (Fig. 3) showed that *Q. rubra* was more similar genetically to *Q. shumardii* than either was to *Q. palustris*. The level of resolution was very limited, however, with the deepest branch involving a genetic distance (Nei 1972) of only 0.057.

Individuals did not resolve into distinct monospecific clusters when genetic data were analyzed using the neighbor-joining method (Fig. 4). Some portions of the unrooted tree comprised predominantly *Q. rubra* individuals and the *Q. palustris* samples clustered mainly on the right side of the tree, but otherwise, it was difficult to discern clear species-based clustering.

The program Structure resolved four subpopulations in the species complex (Table 4). Using Bayes' Rule (formulae in Pritchard 2000), we found that $P(K = 4 \text{ subpopulations}) = 0.999$, $P(K = 5) = 0.001$, and $P(K = 1\text{–}3, 6) < 0.001$ in the DPRF red oaks. There was not a clear correspondence between the inferred clusters and the silvic species designation of the individuals. Most individuals were assigned to either one of two clusters (2 and 3), and all clusters were a mixture of the three species. Only three individuals were assigned to cluster 4 using 90% probability intervals, and these may represent migrants from another stand. We visualized population substructuring in three-dimensional space (Fig. 5) by collapsing the rare fourth dimension and representing individual positions in triangular space as standardized q_1, q_2, and q_3 estimates, wherein q_k is the probability of membership in each of the inferred k subpopulations and is the vertical distance from the opposite side of the triangle.

Discussion

Genetic structure and species relationships

We found significant genetic differentiation among species within the DPRF red oak community, but the magnitude of the differentiation ($\theta_p = 0.010$) was small and commensurate with values reported elsewhere for differences among populations within *Q. rubra*. These isozyme studies of *Q. rubra* range in scale from comparisons across adjacent

Fig. 4. Unrooted neighbor-joining tree (Felsenstein 1989) showing genetic similarities among individuals of three red oak species. Genetic distances among individuals were calculated using a pairwise allele-sharing matrix (Mountain and Cavalli-Sforza 1997) determined from microsatellite genotypes.

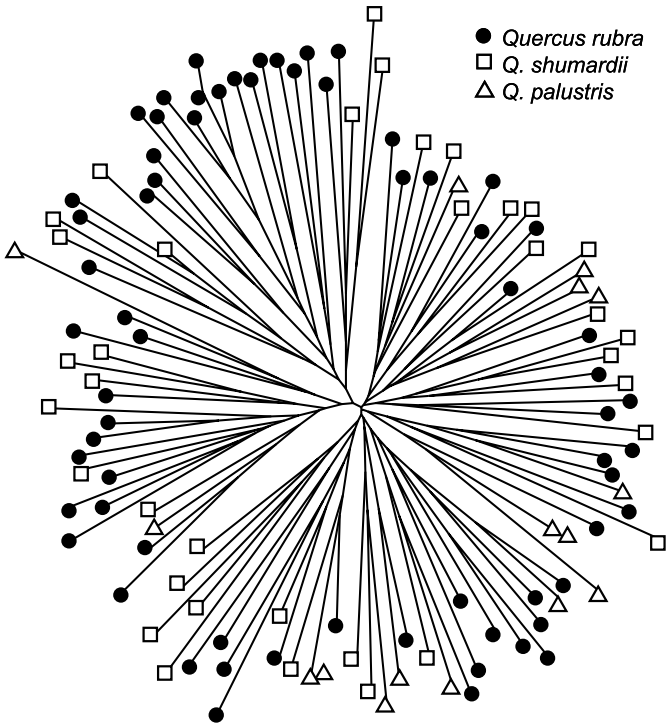


Table 4. Proportion of membership of each species in each of four subpopulation clusters inferred by the program Structure (Pritchard et al. 2000).

	Cluster			
Species	1	2	3	4
Q. rubra	0.272	0.294	0.405	0.030
Q. shumardii	0.025	0.253	0.708	0.014
Q. palustris	0.053	0.343	0.525	0.080

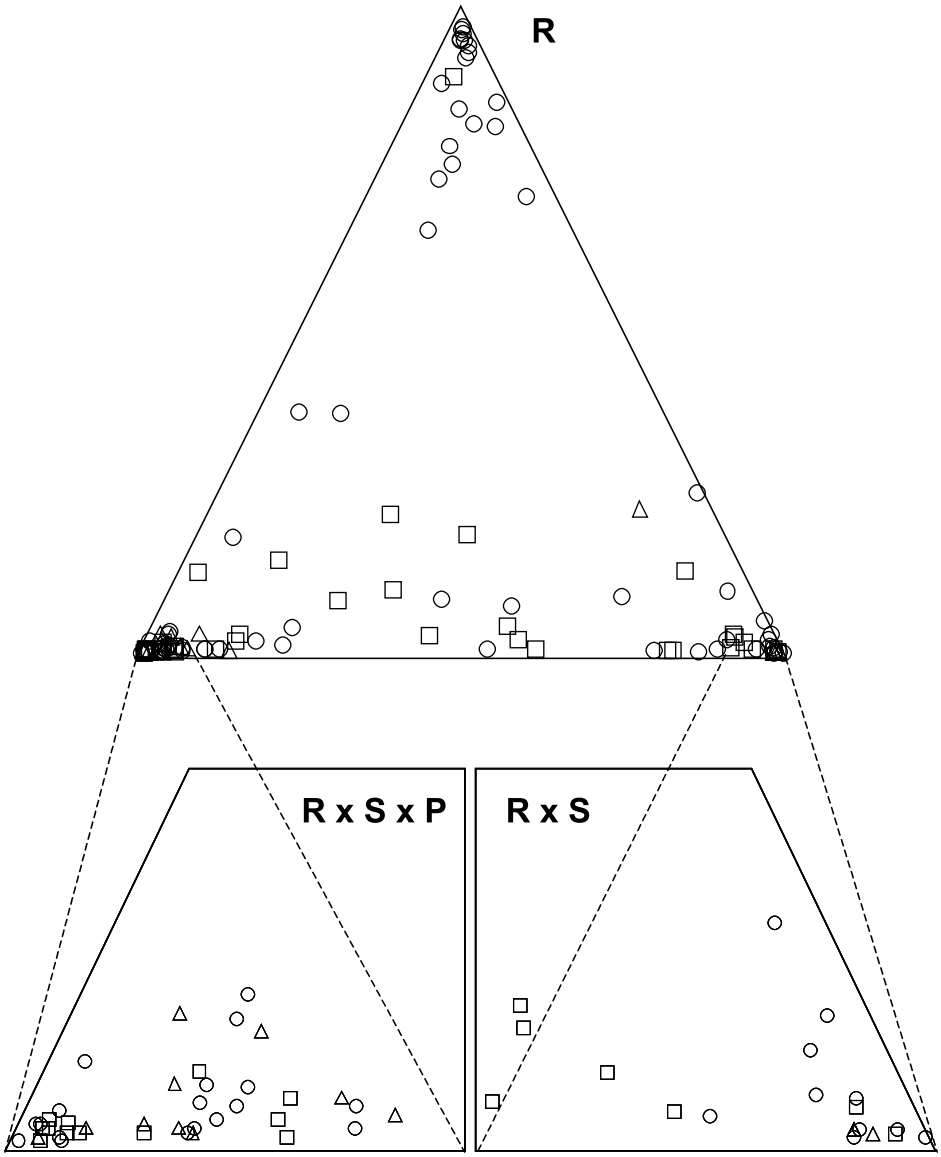
subpopulations ($F_{ST} = 0.011$, Sork et al. 1993) to eight populations from Pennsylvania ($G_{ST} = 0.009$; Schwarzmann and Gerhold 1991) and 10 populations from several states in the Midwest ($F_{ST} = 0.092$; Sork et al. 1993). Guttman and Weigt (1989) compared the three species that we studied and found that they were poorly differentiated at isozyme loci; *Q. shumardii* and *Q. rubra* shared the same most common allele at 17 of 18 loci and *Q. palustris* and *Q. rubra* at 16 of 18 loci. *Quercus palustris* was most genetically distinct in our study, which agrees with conclusions drawn from isozyme research (Manos and Fairbrothers 1987; Guttman and Weigt 1989).

Whole-tree silvics

The whole-tree silvic approach failed to resolve distinct red oak species at the DPRF according to the neighbor-joining tree and the Monte Carlo clustering algorithm. This indicates a fairly weak perceived association at this site be-

© 2003 NRC Canada

Fig. 5. Resolution of substructuring in the gene pool of the red oak complex at the Davis-Purdue Research Forest (DPRF) based on analysis using Structure, a Markov chain Monte Carlo program (Pritchard et al. 2000). Symbols represent individual trees identified to species (circles, *Quercus rubra*; squares, *Quercus shumardii*; triangles, *Quercus palustris*) using whole-tree silvics. The position of each symbol in triangular space denotes the respective probability of membership (q_k) of that individual in each of the three most common subpopulations inferred by this analysis at the DPRF. Symbols near the uppermost apex reside in a cluster (R) that is mostly *Q. rubra* (q_1 is the distance from the bottom edge), symbols near the rightmost apex reside in a cluster (R × S) that contains mostly individuals identified as either *Q. rubra* or *Q. shumardii*, and the leftmost apex (R × S × P) contains a mixture of all three species. Symbols that failed to cluster near any of the apices may be advanced generation hybrids.



tween phenotypic traits commonly used to identify these species and the underlying patterns of genetic variation as measured by these 15 loci. Potential causative factors include (*i*) inadequacies of the whole-tree silvic approach and (*ii*) phenotypic and genetic overlap between species as a result of recent speciation or repeated hybridization and introgression.

A preponderance of errors in the whole-tree silvic assignments because of phenotypic or developmental plasticity, for instance, could mean that the three groups resolved in Fig. 5 represent the pure species *Q. rubra* (R, cluster 1), *Q. shumardii* (R × S, cluster 2), and *Q. palustris* (R × S × P, cluster 3). It is possible, then, that rigorous morphological measurements followed by computer analyses would resolve the assignments into pure species groups, although wide application of such a time-intensive approach is unlikely in applied forestry. It also is conceivable that other morphological characters that we did not consider may have provided better resolution of these three species, although the traits that we used are frequently cited as most useful (e.g., Fowells 1965; Little 1980; Harlow et al. 1996). Ultimately, it may be that there are few truly diagnostic morphological traits distinguishing the red oak species (Palmer 1942; Burger 1975; Jensen 1988). Regardless of causation, however, it is clear

© 2003 NRC Canada

that the whole-tree silvic approach yielded species groups that were poorly differentiated genetically.

Our results corroborate other findings that morphological markers can be poor indicators of underlying genetic structure in some groups of tree species. Perron and Bousquet (1997) reported that species-specific markers for black spruce (*Picea mariana* (Mill.) BSP) and red spruce (*Picea rubens* Sarg.) were more effective at identifying naturally occurring hybrids than were morphological traits, even after analysis with a discriminant function. Tomlinson et al. (2000) showed that whole-tree silvic identifications of *Q. rubra*, *Q. ellipsoidalis*, and putative hybrids revealed only limited genetic differences based on isozyme markers. They also used multivariate morphometric analysis but could not distinguish *Q. rubra* from the putative hybrids, although they could distinguish *Q. ellipsoidalis*. Similarly, we were most able to distinguish *Q. palustris* as a distinct entity by use of an MCMC genetic clustering algorithm, whereas *Q. rubra* was the least distinct of the three species. If the phenotypic overlap seen in this study and in Tomlinson et al. (2000) arose from fairly recent speciation and (or) ongoing hybridization, the failure of the whole-tree silvic method to separate these species simply reflects the underlying biology of the group. Indeed, Kashani and Dodd (2002) used 311 polymorphic amplified fragment length polymorphism AFLP band classes and still resolved few differences between two California red oak species (*Quercus parvula* var. *shreveii* and *Qurcus wislizeni*) suspected of hybridization and (or) recent divergence.

Recent speciation and hybridization

We cannot determine if recent speciation or introgression has played a larger role in limiting genetic differences between these species, and indeed the two mechanisms may act synergistically to limit the effectiveness of whole-tree silvic identifications. Most red oak species appear to have sufficient phenotypic cohesion to be recognizable over large expanses of territory (Harlow et al. 1996), like the European white oaks that retain phenotypic and genetic distinctions despite extensive hybridization (Muir et al. 2000). Failure of the neighbor-joining and MCMC approaches to resolve distinct lineages in our study may owe to several factors relating to speciation and hybridization dynamics. (*i*) Resolution of relationships in a three-species group from a single site, such as ours, may pose a significantly more complex problem compared with a two-species group collected from numerous sites, as with the European white oaks (Muir et al. 2000). In the latter case, subtle differences between the species may become evident if they are retained across several geographic locations. (*ii*) Stands can differ greatly in the degree to which morphological boundaries blur between red oak species (Jensen 1988). Populations of oak species in old-growth forests would have a long history of sympatry with ample opportunity for the accumulation of advanced generation crosses. The result may be a complex patchwork of genomic segments shared among the species, something that might be less noticeable in stands with a higher rate of turnover and shorter bouts of sympatry between population pairs of different oak species. (*iii*) Many red oak species may have evolved more recently than the white oak species (e.g., Guttman and Weigt 1989), in which case, red oak lineages could be fundamentally more difficult to resolve, as less time has elapsed for the fixation of distinct alleles in the process of lineage sorting. (*iv*) It also is possible that differentiation of oak species depends on changes at fairly few places in the genome so that a sample of relatively few loci might result in numerous loci showing little differentiation across species and only a few, if any, of the sampled loci revealing substantial differences. The latter might explain the heterogeneous values of θ reported here (Table 2).

If one accepts a history of hybridization among red oaks at the DPRF, then the interpretation of the MCMC clustering results is fairly straightforward. The red oak community comprises three primary subpopulations or demes: relatively pure *Q. rubra* (R, cluster 1), *Q. rubra* × *Q. shumardii* (R × S, cluster 2), and *Q. rubra* × *Q. shumardii* × *Q. palustris* (R × S × P, cluster 3). This is not to say that the R × S subpopulation consists entirely of *Q.* ×*riparia*, for example, but the genetic and morphological evidence is consistent with the presence of this hybrid along with the parental forms. The extent to which advanced generation hybrids are present is unclear and would require a more thorough examination of hypotheses regarding relationships among the multilocus genotypes. It is reasonable to assert, however, that individuals that fail to cluster near the vertices of the triangle (see also Figs. 1D and 2D) may represent later stages of introgression that have accumulated over a long history of sympatry in the old-growth stand.

The case for hybridization and introgression between *Q. rubra*, *Q. shumardii*, and *Q. palustris* is compelling based on published nongenetic reports. Naturally occurring hybrids have been described formally for all pairwise crosses (Palmer and Steyermark 1935; Laughlin 1963, 1964). Jensen (1994) characterized a number of trees as putative *Q.* ×*riparia* in a northern Indiana woodlot that also contained the pure parental species *Q. rubra* and *Q. shumardii*. Putative hybrids displayed a mixture of parental traits in leaves and fruits, and morphometric analysis showed that hybrids occupied an intermediate position between parent species. The author concluded that there had been considerable local introgression at this site. Similarly, *Q.* ×*columnaris* (*Q. rubra* × *Q. palustris*) was first described by Laughlin (1964) who claimed to have identified both backcross and F_1 individuals in an Illinois woodlot. He noted that the leaf characters looked more like *Q. palustris* but the acorns were more like *Q. rubra*.

Morsink and Pratt (1984) described a hybrid swarm between *Q. rubra*, *Q. shumardii*, and *Q. palustris* in Ontario, Canada, that resembles the DPRF red oak community. Both sites involved soil with poor internal drainage and standing surface water during the winter and spring, with *Q. rubra* in the better-drained areas. They described the Ontario community as a Shumard oak hybrid complex, *Q. shumardii* hybridizing primarily with *Q. palustris* (i.e., *Q.* ×*mutabilis*) and secondarily with *Q. rubra* (i.e., *Q.* ×*riparia*), judging from leaf and acorn morphology, bark, and branching habit. Similarly, at the DPRF, each of the subpopulations that we identified by genetic means contained the most common species (*Q. rubra*) either as a fairly pure entity (R) (Fig. 5) or as part of a mixed-species group or putative hybrid complex (R × S, R × S × P). It makes sense that the most abundant species (*Q. rubra*, DPRF; *Q. shumardii*, Ontario) would have the

© 2003 NRC Canada

greatest opportunity to occur as a pure entity and to cross with all other species in the stand because of its frequency advantage. By contrast, gene pools of the less common species (*Q. palustris*, DPRF) should be more subject to dilution by introgression (R × S × P), since the rare species would cross most often to the common species. Fowells (1965) noted that their flowering periods overlap. After several generations, most individuals of the rare species would carry at least a few introgressed traits of the common species. Such morphological mosaics were evident both in Ontario and at the DPRF.

Practical implications for germplasm management

It is reasonable to expect that the whole-tree silvic approach will fail to identify genetically distinct red oak species under many circumstances. This may arise from errors in the methodology that generate misclassifications, or there simply may be no clear boundaries between some species. Consequently, the selection of seed crop and nursery stock from red oak species complexes may draw from a broader gene pool than is intended. This may be desirable in a *Q. rubra* – *Q. shumardii* complex, since both species produce good-quality timber (Harlow et al. 1996), but it is more problematic if lower-quality species are present. Tomlinson et al. (2000) determined that pure *Q. ellipsoidalis* could be removed from seed collections using silvic methods but that hybrid germplasm would be unavoidably pooled with that of pure *Q. rubra*. We have found a similar pattern at the DPRF. We could use whole-tree silvic identifications to avoid collecting from what appears to be the purest portion of the *Q. palustris* gene pool (R × S × P) (Fig. 5). However, several individuals in this cluster were identified morphologically as *Q. rubra* and *Q. shumardii* despite their overall genetic affinities with *Q. palustris*. These individuals may carry some of the key genes that typify *Q. rubra* and *Q. shumardii* embedded in a *Q. palustris* genomic background.

Acknowledgements

We thank D. Jacobs, R. Jensen, D. Weigel, and three anonymous reviewers for valuable comments on the manuscript and J.K. Pritchard for advice on application of the program Structure. Y. Feng and S. Meyer helped with field collections, M. Jagtap provided assistance with PCR, and W. Sun provided logistical support. Funding was provided by the USDA Forest Service and the Purdue University Department of Forestry and Natural Resources.

References

Aldrich, P.R., Michler, C.H., Sun, W., and Romero-Severson, J. 2002. Microsatellite markers for northern red oak (Fagaceae: *Quercus rubra*). Mol. Ecol. Notes, **2**: 472–474.

Aldrich, P.R., Jagtap, M., Michler, C.H., and Romero-Severson, J.A. 2003. Amplification of North American red oak microsatellite markers in European white oaks and Chinese chesnut. Silvae Genet. In press.

Burger, W.C. 1975. The species concept in *Quercus*. Taxon, **24**: 45–50.

Edwards, S.V., and Beerli, P. 2000. Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. Evolution, **54**: 1839–1854.

Felsenstein, J. 1989. PHYLIP — Phylogeny Inference Package (version 3.2). Cladistics, **5**: 164–166.

Fowells, H.A. 1965. Silvics of forest trees of the United States. U.S. Dep. Agric. Agric. Handb. 271.

Guttman, S.I., and Weigt, L.A. 1989. Electrophoretic evidence of relationships among *Quercus* (oaks) of eastern North America. Can. J. Bot. **67**: 339–351.

Harlow, W.M., Haraar, E.S., Hardin, J.W., and White, F.M. 1996. Textbook of dendrology. 8th ed. McGraw-Hill, Inc., New York.

Hokanson, S.C., Isebrands, J.G., Jensen, R.J., and Hancock, J.F. 1993. Isozyme variation in oaks of the Apostle Islands in Wisconsin: genetic structure and levels of inbreeding in *Quercus rubra* and *Q. ellipsoidalis* (Fagaceae). Am. J. Bot. **80**: 1349–1357.

Jensen, R.J. 1977. A preliminary numerical analysis of the red oak complex in Michigan and Wisconsin. Taxon, **26**: 399–407.

Jensen, R.J. 1988. Assessing patterns of morphological variation of *Quercus* spp. in mixed-oak communities. Am. Midl. Nat. **120**: 120–135.

Jensen, R.J. 1994. *Quercus rubra* dans le sous-genre *Erythrobalanus*: systematique, origine et hybridation. *In* Le chêne rouge d'Amerique. *Edited by* J. Timbal, A. Kremer, N. le Gof, and G. Nepveu. Institut National de la Recherche Agronomique, Paris. pp. 61–74.

Jensen, R.J., and Eshbaugh, W.H. 1976. Numerical taxonomic studies of hybridization of *Quercus*. I. Populations of restricted areal distribution and low taxonomic diversity. Syst. Bot. **1**: 1–10.

Jensen, R.J., DePiero, R., and Smith, B.K. 1984. Vegetative characters, population variation and the hybrid origin of *Quercus ellipsoidalis*. Am. Midl. Nat. **111**: 364–370.

Jensen, R.J., Hokanson, S.C., Isebrands, J.G., and Hancock, J.F. 1993. Morphometric variation in oaks of the Apostle Islands in Wisconsin: evidence of hybridization between *Quercus rubra* and *Q. ellipsoidalis* (Fagaceae). Am. J. Bot. **80**: 1358–1366.

Kashani, N., and Dodd, R.S. 2002. Genetic differentiation of two California red oak species, *Quercus parvula* var. *shreveii* and *Q. wislizeni*, based on AFLP genetic markers. U.S. For. Serv. Gen. Tech. Rep. PSW-GTR-184. pp. 417–426.

Laughlin, K. 1963. *Quercus* × *riparia* Laughlin Kansas City oak. Phytologia, **9**: 101–107.

Laughlin, K. 1964. *Quercus* × *columnaris* Laughlin Caldwell oak. Phytologia, **9**: 488–495.

Lewis, P.O., and Zaykin, D. 2001. Genetic Data Analysis: computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors over the internet. Available from http://lewis.eeb.uconn.edu/lewishome/software.html [cited 2003].

Little, E.L. 1980. The Audubon Society field guide to North American trees: eastern region. Alfred A. Knopf, New York.

Manos, P.S., and Fairbrothers, D.E. 1987. Allozyme variation in populations of six northeastern American red oaks (Fagaceae: *Quercus* subg. *Erythrobalanus*). Syst. Bot. **12**: 365–373.

Morsink, W.A.G., and Pratt, P.D. 1984. Shumard Oak, *Quercus shumardii*, in Essex County, Ontario. Can. Field-Nat. **98**: 470–478.

Mountain, J.L., and Cavalli-Sforza, L.L. 1997. Multilocus genotypes, a tree of individuals, and human evolutionary history. Am. J. Hum. Genet. **61**: 705–718.

Muir, G., Fleming, C.C., and Schlotterer, C. 2000. Species status of hybridizing oaks. Nature (London), **405**: 1016.

Nei, M. 1972. Genetic distance between populations. Am. Nat. **106**: 283–292.

Palmer, E.J. 1942. The red oak complex in the United States. Am. Midl. Nat. **27**: 732–740.

© 2003 NRC Canada

Palmer, E.J., and Steyermark, J.A. 1935. *Quercus* × *mutabilis* Palmer & Steyerm. Ann. Mo. Bot. Gard. **22**: 521.

Pamilo, P., and Nei, M. 1988. Relationships between gene trees and species trees. Mol. Biol. Evol. **5**: 568–583.

Parker, G.R., and Leopold, D.J. 1983. Replacement of *Ulmus americana* L. in a mature east-central Indiana woods. Bull. Torrey Bot. Club, **110**: 482–488.

Parker, G.R., Leopold, D.J., and Eichenberger, J.K. 1985. Tree dynamics in an old-growth, deciduous forest. For. Ecol. Manage. **11**: 31–57.

Perron, M., and Bousquet, J. 1997. Natural hybridization between black spruce and red spruce. Mol. Ecol. **6**: 725–734.

Petit, R.J., Pineau, E., Demesure, B., Bacilieri, R., Ducousso, A., and Kremer, A. 1997. Chloroplast DNA footprints of postglacial recolonization by oaks. Proc. Natl. Acad. Sci. U.S.A. **94**: 9 996 – 10 001.

Prentice, B. 1927. Forest survey No. 1. Herbert Davis, Forestry Farm. Unpublished report to the Department of Forestry and Conservation, Purdue University, West Lafayette, Ind.

Pritchard, J.K. 2000. Documentation for *structure* software. University of Oxford, Oxford, U.K.

Pritchard, J.K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics, **155**: 945–959.

Schlotterer, C. 2001. Genealogical inference of closely related species based on microsatellites. Genet. Res. Camb. **78**: 209–212.

Schwarzmann, J.F., and Gerhold, H.D. 1991. Genetic structure and mating system of northern red oak (*Quercus rubra* L.) in Pennsylvania. For. Sci. **37**: 1376–1389.

Sork, V.L., Huang, S., and Wiener, E. 1993. Macrogeographic and fine-scale genetic structure in a North American oak species, *Quercus rubra* L. *In* Genetics of oaks. *Edited by* A. Kremer, P.S. Savill, and K.C. Steiner. Elsevier, Paris. pp. 261–270.

Steinkellner, H., Fluch, S., Turetschek, E., Lexer, C., Streiff, R., Kremer, A., Burg, K., and Glossl, J. 1997. Identification and characterization of $(GA/CT)_n$ — microsatellite loci from *Quercus petraea*. Plant Mol. Biol. **33**: 1093–1096.

Tomlinson, P.T., Jensen, R.J., and Hancock, J.F. 2000. Do whole tree silvic characters indicate hybridization in red oak (*Quercus* section *Lobatae*)? Am. Midl. Nat. **143**: 154–168.

Weir, B.S. 1996. Genetic data analysis II. Sinauer Associates, Inc., Sunderland, Mass.

Weir, B.S., and Cockerham, C.C. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution, **38**: 1358–1370.

Whittemore, A.T., and Schaal, B.A. 1991. Interspecific gene flow in sympatric oaks. Proc. Natl. Acad. Sci. U.S.A. **88**: 2540–2544.

Wright, S. 1951. The genetical structure of populations. Ann. Eugen. **15**: 323–354.

© 2003 NRC Canada