*Ecology*, 82(7), 2001, pp. 2012–2021


# ASSOCIATION BETWEEN ENVIRONMENTAL AND GENETIC HETEROGENEITY IN FOREST TREE POPULATIONS

WENDY K. GRAM[1] AND VICTORIA L. SORK[2]

*University of Missouri, Department of Biology, St. Louis, Missouri 63121 USA*

*Abstract.* Within a plant species, environmental heterogeneity has the potential to influence the distribution of genetic variation among populations through several evolutionary processes, including natural selection, differential gene exchange, and chance associations caused by genetic drift or founder effects. We evaluated the relationship between environmental characteristics and the distribution of genotypes for three common woody species (*Quercus alba, Carya tomentosa,* and *Sassafras albidum*) in Missouri Ozark forests. We measured the correlation between multivariate genotypes and forest structure variables, and assessed the influence of physical landscape on multivariate genotypes. For all three species, we found significant relationships between genetic variation and environmental heterogeneity. We discovered that populations in local habitat patches with different forest structures also differed in combinations of multivariate genotypes. In contrast, we did not detect significant differences in multivariate genotypes among soil-type/aspect classes in any of the three study species, suggesting that genotypic differentiation is operating on a finer scale than soil or aspect differences. We conclude that natural selection, possibly interacting with founder events, has influenced the population differentiation of these three long-lived plant species. Such microgeographic variation in response to environmental heterogeneity is expected for a broad range of species, even when extensive gene flow is present.

*Key words:* *canonical correlation;* Carya tomentosa; *environmental heterogeneity; founder effects; Missouri Ozarks; multilocus genetic structure; multivariate genotypes; natural selection; population differentiation;* Quercus alba; Sassafras albidum.

## INTRODUCTION

Environmental heterogeneity across a landscape produces microenvironments that vary in plant species composition. Within a plant species, environmental heterogeneity has the potential to influence the distribution of genetic variation among populations (Antonovics 1971, Linhart and Grant 1996, Mitton 1997). Environmental heterogeneity can create genetic heterogeneity through several evolutionary processes (Linhart and Grant 1996). First, natural selection can cause populations to adapt to their local environment, resulting in fine-scale microgeographical variation. This process has been demonstrated for a variety of traits (see reviews by Antonovics and Bradshaw 1970, Linhart and Grant 1996, Mitton 1997), including single or few locus traits (e.g., pathogen tolerance; Dirzo and Harper 1982), quantitative traits (e.g., Dudley 1996*a, b*), single locus allozymes (e.g., allozymes and stomatal size; Mitton et al. 1998), and multilocus allozymes (e.g., soil moisture and temperature; Allard et al. 1972, 1993). Second, genetic heterogeneity may be the result of differential gene exchange, influenced by variation in flowering phenology among local habitats. Several studies have demonstrated ecological barriers that prevent gene exchange among populations along an elevational gradient (Mitton et al. 1980). Within a region, phenological differences in flowering time associated with local habitat (e.g., Jackson 1966) may create genetic barriers that, in turn, exacerbate local population differences. Third, chance associations caused by genetic drift or founder effects could also create genetic heterogeneity, as founders colonize different sites but gene flow is not sufficient to homogenize differences (e.g., see Husband and Barrett 1996, Antonovics et al. 1997). Thus, microevolutionary processes have the potential to create genetic patchiness across a heterogeneous landscape.

Theoretically, gene flow is expected to homogenize the distribution of genetic variation unless selection is quite strong (Endler 1973, Slatkin 1985). Thus, we cannot assume that environmental heterogeneity will lead to genetic heterogeneity for all traits or for all plant species. Even when selection is strong enough within the range of gene flow to create population differentiation (e.g., selection for tolerance to toxic soils; Antonovics and Bradshaw 1970), other aspects of the genome may remain undifferentiated. It may be for this reason that $F$ statistics based on single, putatively neutral loci, often reveal low levels of population differ-

Manuscript received December 13 1999; revised 15 June 2000; accepted 5 July 2000; final version received 7 August 2000.

[1] Present address: Sam Noble Oklahoma Museum of Natural History, University of Oklahoma, 2401 Chautauqua Avenue, Norman, Oklahoma 73072 USA.

[2] Corresponding author. E-mail: sork@umsl.edu

entiation (Latta 1998). However, $F$ statistics, which assume migration/drift equilibrium (Neigel 1997), may not be a sensitive measure of population differentiation (Westfall and Conkle 1992, Kremer and Zanetto 1997; Sork et al., *in press*). To detect genetic differentiation, it is preferable to examine multiple loci simultaneously (Taylor and Mitton 1974, Westfall and Conkle 1992, Kremer and Zanetto 1997). These measures will be able to detect small differences in allelic state that accumulate across loci (Smouse et al. 1982), and they can detect the effects of linkage disequilibrium caused by selection (Yang and Yeh 1993). While it is not clear that multivariate analyses will effectively detect differentiation associated with quantitative trait loci (Latta 1998), it is true that such analyses improve discrimination among populations (Westfall and Conkle 1992, Latta 1998).

An effective way to test whether multilocus genetic differentiation is associated with environmental heterogeneity is to summarize multivariate genotypes and environments with multivariate statistics, such as principal component analysis, canonical correlation analysis, or canonical discriminant analysis (Westfall and Conkle 1992). Multivariate analyses are more efficient than univariate analyses because only one analysis is required for all variables. Even if one variable (e.g., a single locus or environmental measure) is independent of the others and is responsible for generating a pattern, this relationship will be clearly shown in the structure of a multivariate analysis. Furthermore, single loci are not likely to reveal patterns of differentiation, even if they are significant (Thorpe 1985, Westfall and Conkle 1992). Several studies have used multivariate genotypes to describe macrogeographical patterns of genetic differentiation within species by correlating genetic variables with elevational or latitudinal gradients (Yeh et al. 1986, Westfall and Conkle 1992, Aradhya et al. 1993). Other studies have focused on smaller geographic scales (Allard et al. 1972, 1993, Hamrick and Allard 1972) using multiallelic associations, which are another way to look at multiple loci simultaneously. These studies found nonrandom associations with patchily distributed environmental variation.

Multilocus genotypes are appropriate genetic variables to test with environmental heterogeneity, but choosing spatially and temporally relevant environmental variables is more complicated. Evolutionary processes work across many generations of a population, so ideal environmental variables are measures of other organisms or physical attributes that have experienced the same conditions through time and space as the target organism. In this study, we were interested in evaluating the effects of environmental variation on the genetic differentiation of forest tree populations. Forest structure measures, based on densities of trees of different size classes, basal area, and canopy cover, characterize the local environment of each sample tree population. These measures may not be directly responsible for genetic differentiation because forest structure clearly changes through time, but forest structure variables are a good surrogate for unmeasured environmental factors, and thus reflect environmental conditions through time. To complement these quantitative local environmental measures, we also classified populations based on broad, abiotic categories of soil type and slope aspect. Other studies have demonstrated ecological and genetic differences between north and east-facing slope (NE) and south and west-facing slope (SW) aspect classes (Kabrick et al. 1997, Sork et al. 1997). Environmentally, the forest-structure measures enabled us to evaluate associations with the local environment whereas the soil-type/aspect classes permitted assessment of the influence of the physical landscape on genotypic distribution.

In this paper, we evaluated the relationship between environmental characteristics and the distribution of genotypes for three common woody species (*Quercus alba, Carya tomentosa,* and *Sassafras albidum*) in Missouri Ozark forests. These three species have significant amounts of multilocus genetic differentiation among populations (Koop 1996; Sork et al., *in press*) in spite of nonsignificant single locus genetic structure as measured by $F$ statistics (Koop 1996, Sork et al. 1997). Given the topographical and local habitat diversity of the Ozark landscape (Brookshire and Shifley 1997), we tested whether extensive genetic differentiation was associated with environmental heterogeneity. Our first objective was to assess the correlation between forest structure variables and individual genotypes for each species using composite variables generated from a canonical correlation analysis. Our second objective was to test the association between the physical landscape and multivariate genotypes. Our populations were sampled from one of two aspect classes (south and west-facing slopes, and north and east-facing slopes) that were found in one of two soil types (Alfisol or Ultisol; Brookshire and Shifley 1997). We tested whether genotypes differ among four soil-type/aspect classes with multivariate analysis of variance (MANOVA), treating individual genotypes as dependent variables.

## Methods

### *Study area*

The study area was part of the Missouri Ozark Forest Ecosystem Project (MOFEP), a multi-investigator ecosystem project administered by the Missouri Department of Conservation, in southeastern Missouri (Brookshire and Shifley 1997) (see Plate 1). In this forest, *Quercus alba* shared the canopy with other species of oaks, including *Q. stellata, Q. velutina, Q. coccinea,* as well as with *Pinus echinata* and *Carya tomentosa* (Kurzejeski et al. 1993). The study area was divided into nine sites (Fig. 1), ranging in size from 260 to 527 ha. The sites were contiguous tracts of forest

PLATE 1. Oak–hickory forest in Missouri Ozarks, Missouri, USA. Left: View of the canopy illustrating topographic heterogeneity. Right: Interior view of the forest. Photographs by Victoria Sork.

with minimal edge in close proximity to each other and largely free from manipulation for at least 40 yr (Brookshire and Shifley 1997). The MOFEP study area includes 13 ecological land types (i.e., classes based on landform, aspect, soil, and vegetation; Miller 1981, Meinert et al. 1997); south and west-facing slope (SW), north and east-facing slope (NE), and ridge-top ecological land types account for 90% of the total study area (Meinert et al. 1997). Each site was further divided into ~5.0-ha forestry stands within an ecological land-type. Sampled populations ranged in elevation from 182 to 275 m, and were between latitude 37°00′ N and 37°15′ N, and between longitude 91°07′ W and 91°00′ W in the Current River and Peck Ranch Conservation Areas. The minimum interpopulation distance was 200 m, while the two most extreme populations at each end of the study area were 24 km apart.

### *Study species*

We selected two canopy trees and one understory shrub to be the focal study species: *Carya tomentosa* Nuttell (Juglandaceae) (common name, mockernut hickory); *Quercus alba* L. (Fagaceae) (common name, white oak); and *Sassafras albidum* (Nuttell) Nees (Lauraceae). Hereafter, these species will be referred to as *Carya, Quercus,* and *Sassafras,* respectively. *Quercus* and *Sassafras* are diploid, whereas *Carya* is a tetraploid. These three species were chosen because they are widely distributed among MOFEP study sites (Brookshire and Shifley 1997). These species differ in pollen vector, seed dispersal vector, successional status, canopy status, and abundance in the MOFEP sites (Brookshire and Shifley 1997). *Carya* and *Quercus* are wind pollinated and vertebrate dispersed, while *Sassafras* is insect pollinated and bird dispersed. In the Ozarks, *Sassafras* is early successional and is usually an understory tree (V. L. Sork, *personal observation*). These three species differ in abundance in the MOFEP sites. Mean *Quercus* density per site ranges from 185 to 337 individuals/ha across the study area, whereas mean *Carya* density per site ranges from 83 to 140 individuals/ha, and mean *Sassafras* density per site ranges from 27 to 83 individuals/ha (Missouri Department of Conservation, *unpublished data*).

In earlier studies of the genetic structures of these three study species using single-locus models, we found that *Carya* and *Sassafras* had significant genetic differentiation among populations ($F_{ST}$ = 0.034, 0.139, respectively; Sork et al., *in press*; $F_{ST}$ represents a genetic measure of differentiation), but *Quercus* did not ($F_{ST}$ = 0.009; Koop 1996). In contrast, the multivariate genetic structures based on MANOVA revealed highly significant population differentiation for all three species. The percentage of genetic variation attributable to populations was 30% for *Carya* and *Sassafras* (Sork et al., *in press*), and 16% for *Quercus* (A. L. Koop and V. L. Sork, *unpublished manuscript*). In fact, for all three species, the differentiation among the nine MOFEP sites was much less than differentiation among populations within a site. Thus, these species showed a great deal of genetic heterogeneity across the Ozark landscape.

### *Genetic sampling*

The data for this study were collected as part of a genetic survey of tree populations in the MOFEP study area (for full details of sampling and electrophoretic protocols, see Sork et al. 1997). We sampled 48 individuals within an area of 1.0 ha from four stands (i.e., populations) in each of the nine sites of the MOFEP study area (Fig. 1), resulting in a total sample size of 36 populations per species. Stands sampled were located on either south and west-facing slope (SW) or north and east-facing slope (NE) ecological land types. Leaves were collected and placed on ice until they could be stored in an ultracold freezer (−70°C).

All leaf samples were analyzed using standard horizontal starch gel electrophoresis procedures (see Kephart 1990). For each species, we surveyed 20 enzymes on several combinations of five gel–electrode buffer systems (see Koop 1996, Sork et al. 1997). As a tetraploid species, *Carya* showed different banding pat-

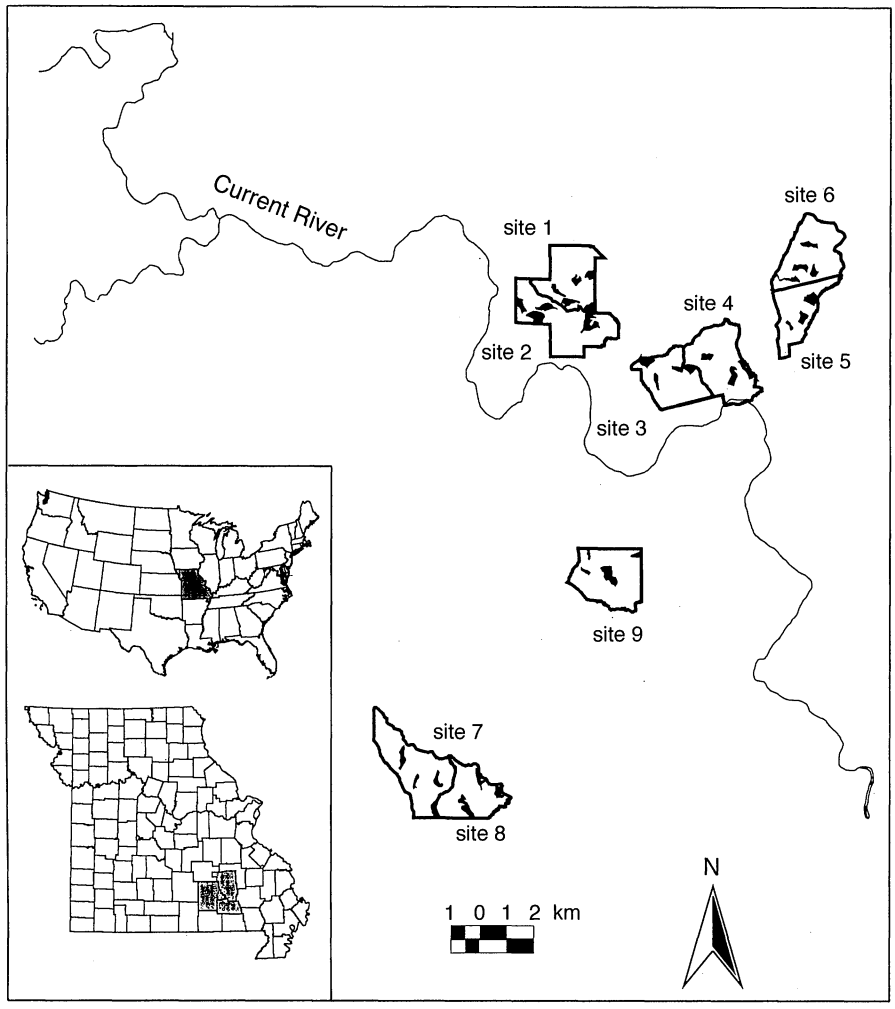


FIG. 1. Map of the Missouri Ozark Forest Ecosystem Project study area with study sites 1–9. Forest stands where populations of *Carya tomentosa, Sassafras albidum,* and *Quercus alba* were sampled are shaded in black.

terns from the other two study species, requiring special interpretation. For homozygotes, we assigned four alleles to the same locus when only one band was observed. For heterozygotes, we assigned the number of alleles according to the intensity of the banding pattern (for description of banding interpretation of tetraploid genotypes, see Murawski et al. 1994). For analyses used in this study, we used only polymorphic loci. We were unable to obtain complete sample sizes for all three species due to a freezer breakdown that allowed degradation of some of the *Carya* samples, and difficulties in obtaining clear enzyme expression across all *Sassafras* individuals, a problem typical of species in the Lauraceae (V. L. Sork, *personal observation*). For multivariate statistics, if the genotype of one locus is lost, the entire multilocus genotype must be removed from the analysis. Fortunately, we had sufficient replication within populations that this loss of observations did not eliminate any populations from the *Sassafras* analysis. The final sample sizes for statistical analysis were 1717 multivariate genotypes (i.e., individuals) from 36 populations across seven isozymes for *Quercus,* 1077 multivariate genotypes from 31 populations across five isozymes for *Carya,* and 1094 multivariate genotypes from 36 populations across five isozymes for *Sassafras.*

### *Data analysis*

The Missouri Department of Conservation provided data from randomly assigned 0.2-ha vegetation plots on density and basal area of all trees, percent canopy cover, soil type, and mean stand aspect for each population sampled (Brookshire and Shifley 1997). We calculated eight separate measures to represent forest structure of each study population, including total basal area, total tree density, small tree density (diameter at breast height [dbh] $\geq$ 4 cm and dbh $<$ 11 cm), medium tree density (dbh $\geq$ 11 cm and dbh $<$ 37 cm), large

TABLE 1. Means, standard deviations, and ranges (Min., Max.) of the eight forest structure variables from the 36 sample stands. All tree densities are reported in units of trees per hectare.

| Forest structure variable | Mean | 1 SD | Min. | Max. |
|---|---|---|---|---|
| Total basal area ($m^2$/ha) | 176 | 19.6 | 137 | 244 |
| Total density (trees/ha) | 609 | 134.1 | 392 | 1007 |
| Small tree density | 264 | 107.1 | 45 | 620 |
| Medium tree density | 137 | 41.9 | 48 | 226 |
| Large tree density | 17 | 9.3 | 0 | 40 |
| Large pine density | 10 | 12.3 | 0 | 50 |
| Small pine density | 6 | 20.2 | 0 | 90 |
| Canopy cover (%) | 84 | 7.3 | 63 | 92 |

tree density (dbh $\geq$ 37 cm), small pine density (dbh $\geq$ 4 cm and dbh $<$ 11 cm), large pine density (dbh $\geq$ 11 cm), and percent canopy cover. We used densities of small and large pine as separate variables, because pines differ from angiosperm wood plants in tree architecture, and they are known to be associated with acidic soils. All forest structure variables showed considerable variation among sample stands (Table 1).

To generate a multivariate genotype of each individual in a population, we transformed all diploid and tetraploid genotypes into linear combinations of traits, by scoring each polyploid genotype into a score for each allele minus one. This method was outlined by Smouse and colleagues (Smouse and Williams 1982, Smouse et al. 1982) and described in detail by Westfall and Conkle (1992). For example, an allele at a locus was assigned a value of 0.5 each time it was present and 0.0 when it was absent. Thus, the individual score for an allele at a homozygous locus would be 1.0 (i.e., 0.5 + 0.5), while the score for the same allele at a heterozygous locus would be 0.5 (i.e., 0.5 + 0.0). The score of a tetraploid heterozygote could be either 0.25, 0.5, or 0.75 depending on whether the allele was represented one, two, or three times in the heterozygote. This calculation was repeated for each allele at a locus minus one allele. For a population, the score of an allele was its total frequency (i.e., the sum of the individual allele scores) for the population. Linear combinations of these scores are usually normally distributed (Smouse et al. 1982, Westfall and Conkle 1992), and we evaluated our data for normality by examining the residuals from the canonical model.

To test for associations between multivariate genotypes and environmental variation, we used canonical correlation analyses (CCA) for each study species, generating new composite variables with maximal correlation between individual genotypes and forest structure. We kept the model that best fit the data as determined by the overall canonical correlation, canonical structure, and normality of the residuals (Table 2). The best fit CCA model produced a set of univariate canonical variables (i.e., a genotypic vector and forest structure vector, in this case) with maximal correlations that were standardized to a mean of zero and standard deviation of one. Thus, each individual tree sampled had a standardized genotypic canonical score and a standardized forest structure canonical score that could be used in further univariate analyses. Because we were ultimately interested in population-level genotypic associations, we correlated mean genotypic canonical scores per population with forest structure canonical scores for each species (which only exist at the population level; Gram and Sork 1999), and we report these Pearson correlations.

To examine whether broad physical characteristics of the site were associated with multivariate genotypic composition, we tested whether mean population genotypes differed among soil-type/aspect classes. Our populations were found in four different soil-type/aspect classes, including Ultisol soil/NE aspect, Alfisol soil/NE aspect, Ultisol soil/SW aspect, and Alfisol soil/SW aspect. We used a nested multivariate analysis of variance (MANOVA) model, with individual genotypes as dependent variables. The MANOVA model was of the form

$$Y_{ijk} = C_i + P_{ij} + e_{ijk}$$

where $C_i$ is soil/aspect class, $P_{ij}$ is population within soil/aspect class, and $e_{ijk}$ is the residual effect.

We used population nested within soil-type/aspect class ($P_{ij}$) as the error term in the soil-type/aspect class significance test, and the residual as the error term for the population (soil/aspect class) significance test. We report $F$ values from Hotelling-Lawley statistics with associated degrees of freedom and significance levels (SAS 1996).

To test for the presence of spatial autocorrelation in genotypic and forest structure data, we compared geographical distances between populations with genotypic and forest structure differences between populations, with a series of Mantel tests (Mantel 1967), using a permutation test with 999 iterations to test for significance. We calculated Mahalanobis genetic distances (Mahalanobis 1948) between populations to construct the genotypic distance matrices, and Euclidean distances between populations based on forest structure variables to construct the forest structure distance matrices.

## RESULTS

Genotypic composition was associated with environmental variation. Using the first canonical vectors, we found significant correlations between the mean population genotypic vector and the forest structure vector for *Quercus* ($r = 0.67$, $n = 36$, $P < 0.0001$), *Carya* ($r = 0.85$, $n = 31$, $P < 0.0001$), and *Sassafras* ($r = 0.57$, $n = 36$, $P = 0.0003$; Fig. 2). These canonical vectors accounted for most of the variation in the canonical correlation analysis model for all three species (*Quercus,* 68%; *Carya,* 33%; *Sassafras,* 46%), and exhibited significant canonical correlations (Table 2). Al-

TABLE 2. Canonical structure from the canonical correlation analyses (CCA) between genotypic and forest structure variables for *Quercus, Carya,* and *Sassafras.*

| Allozymes | Canonical structure (correlation with genotypic vector) | Forest structure variables | Canonical structure (correlation with forest structure vector) |
|---|---|---|---|
| *Quercus alba*† ($R$ = 0.17, $P$ = 0.0024) | | | |
| Pgi-2[3] (Phosphoglucoisomerase) | −0.7703 | Large tree density‡ | 0.9056 |
| Pgi-2[2] | 0.6202 | Medium tree density§ | −0.5133 |
| Pgi-2[4] | 0.5268 | Small pine density‖ | 0.3234 |
| Ces-1[4] (Colormetric esterase) | −0.3168 | Total basal area ($m^2$/ha) | 0.2439 |
| Fes-1[3] (Fluorescent esterase) | 0.2443 | | |
| Fes-1[2] | −0.2388 | | |
| Per-1[6] (Peroxidase) | 0.2055 | | |
| Fes-4[4] | 0.1898 | | |
| Per-1[2] | −0.1752 | | |
| Ces-1[3] | 0.1390 | | |
| *Carya tomentosa* ($R$ = 0.30, $P$ = 0.0001) | | | |
| Skdh[3] (Shikimate dehydrogenase) | −0.5919 | Total basal area ($m^2$/ha) | 0.9237 |
| Skdh[4] | 0.3959 | Medium tree density§ | −0.8563 |
| Skdh[2] | 0.3672 | Total density | 0.5657 |
| Pgi-2[3] | 0.3446 | Large pine density‡ | −0.4319 |
| Dia[5] (Diaphorase) | 0.2727 | Small tree density‖ | −0.2984 |
| Pgi-2[2] | −0.2418 | Small pine density‖ | −0.2014 |
| Dia[7] | −0.2285 | Large tree density‡ | −0.1422 |
| Skdh[1] | −0.1812 | Canopy cover (%) | 0.0896 |
| Aat[7] (Aspartate aminotransferase) | −0.1802 | | |
| Pgi-2[7] | 0.1653 | | |
| Dia[9] | −0.1652 | | |
| Aat[5] | 0.1629 | | |
| Pgi-2[5] | −0.1299 | | |
| Pgi-2[8] | −0.0809 | | |
| Pgi-2[6] | 0.0662 | | |
| Mnr[6] (Menadione reductase) | 0.0646 | | |
| Dia[3] | −0.0296 | | |
| Pgi-2[4] | −0.0286 | | |
| Mnr[7] | −0.0264 | | |
| Mnr[5] | −0.0108 | | |
| *Sassafras albidum* ($R$ = 0.36, $P$ = 0.0001) | | | |
| Mnr[7] | −0.6636 | Canopy cover (%) | 0.8046 |
| Dia-2[3] | 0.5444 | Large pine density‡ | −0.4991 |
| Dia-2[5] | −0.4775 | Total basal area ($m^2$/ha) | −0.4345 |
| Aat-2[5] | −0.4169 | Total density | 0.1729 |
| Aat-2[3] | 0.4165 | Medium tree density§ | 0.1638 |
| Dia-1[1] | −0.3988 | Small tree density‖ | 0.0578 |
| Dia-1[2] | 0.2842 | Large tree density‡ | 0.0089 |
| Aat-3[7] | 0.1860 | | |
| Aat-3[5] | −0.1641 | | |
| Dia-1[4] | −0.1228 | | |
| Aat-3[3] | 0.0773 | | |
| Dia-1[3] | 0.0537 | | |

*Notes:* The correlation between each input variable included in the final CCA model (e.g., see *Methods*) and its canonical vector are listed for the first genotypic and forest structure vectors of each species. Canonical correlation ($R$) and significance for each model are presented after species name. Allozymes are listed with an abbreviation to indicate the enzyme (with full name given at first usage), and locus number. The superscript represents allele number for that allozyme. All tree densities are in trees per hectare.

† Per-3 (Peroxidase-3) and Mnr-1 (Menadione reductase) alleles were part of the analysis but were not included in the best-fit model.

‡ Density of trees with diameter at breast height (dbh) ≥ 37 cm, except for large pines which includes trees with dbh >11 cm.

§ Density of trees with dbh ≥ 11 cm and dbh < 37 cm.

‖ Density of trees with dbh ≥ 4 cm and dbh < 11 cm.

lele patterns among the three species were not similar, but all genotypic canonical vectors were multilocus as indicated by relatively high correlations with several alleles per species. The environmental canonical vector was represented by four to eight forest structure variables, depending on species (Table 2).

Population genotypes did not differ among broad physical site characteristics, but genotypic variation

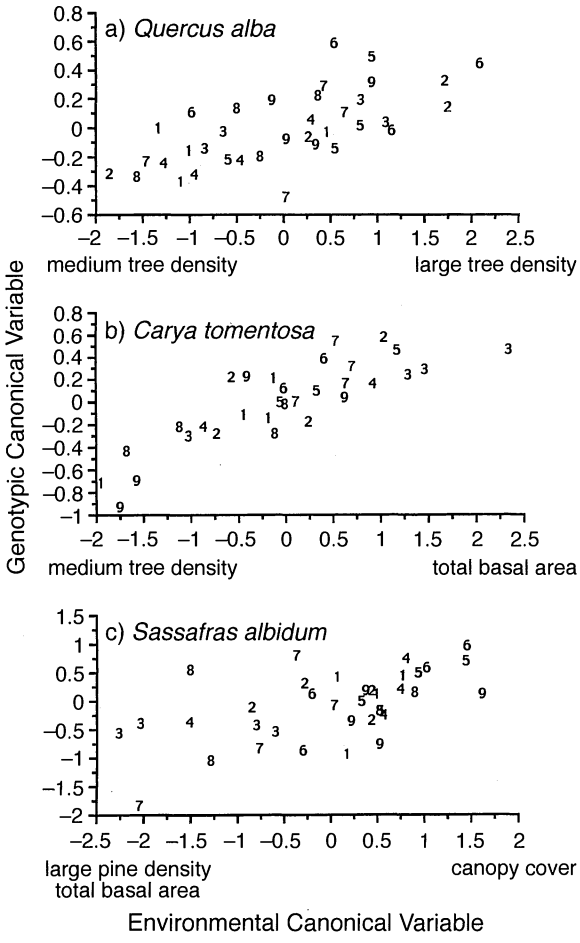


FIG. 2. Plots of the mean genotypic canonical variable vs. the environmental canonical variable for: (a) *Quercus alba,* (b) *Carya tomentosa,* and (c) *Sassafras albidum.* Canonical variables are composite variables produced with canonical correlation analysis and are expressed as standardized values with a mean = 0 and SD = 1. Forest structure variables that are most highly correlated (positively and negatively) with the environmental canonical variable are identified on the *x*-axis for each species (see Table 2). Populations ($n$ = 36 for *Sassafras* and *Quercus,* $n$ = 31 for *Carya*) are represented by their study site numbers (1–9).

among populations within a soil-type/aspect class was substantial. We did not find significant differences among different soil-type/aspect classes for *Quercus* ($F_{42,53} = 1.38$, $P = 0.14$), *Carya* ($F_{60,23} = 0.94$, $P = 0.60$), and *Sassafras* ($F_{36,59} = 0.88$, $P = 0.65$). We did, however, find significant differentiation among populations within classes for *Quercus* ($F_{448,21478} = 1.40$, $P < 0.0001$), *Carya* ($F_{560,20482} = 2.47$, $P < 0.0001$), and *Sassafras* ($F_{384,12542} = 5.63$, $P < 0.0001$).

The Mantel tests for spatial autocorrelation did not show significant correlations for forest structure measures for *Quercus* ($r = 0.07$, $n = 36$, $P = 0.22$), *Carya* ($r = 0.07$, $n = 31$, $P = 0.12$), or *Sassafras* ($r = 0.10$, $n = 36$, $P = 0.51$). *Sassafras* population genotypes exhibited a trend towards spatial autocorrelation ($r = 0.12$, $n = 36$, $P = 0.06$), but *Quercus* ($r = -0.09$, $n = 36$, $P = 0.19$) and *Carya* ($r = -0.06$, $n = 31$, $P = 0.22$) were not spatially autocorrelated.

## DISCUSSION

Our results demonstrate significant genetic variation in association with environmental heterogeneity for all three species. The observed genetic heterogeneity occurred in spite of supposedly extensive gene flow that should have homogenized the populations. Based on the genetic structure estimates for wind-pollinated species (Loveless and Hamrick 1984), we expect the wind-pollination systems of *Carya* and *Quercus* to promote extensive gene flow. In fact, with $F$ statistic analyses, we find that the estimates of $F_{ST}$ are sufficiently small to conclude that historical gene flow in the wind-pollinated species has been extensive ($F_{ST} = 0.034$, 0.009, and 0.139 for *Carya, Quercus,* and *Sassafras,* respectively; Sork et al., *in press*; A. L. Koop and V. L. Sork, *unpublished manuscript*). However, the association with environmental heterogeneity implies that this gene flow has not been sufficient to prevent the occurrence of differentiation among populations.

Many biologists would propose that natural selection is the most likely evolutionary process responsible for genetic heterogeneity (Antonovics and Bradshaw 1970, Linhart and Grant 1996, Mitton 1997). The genetic mosaic observed in this heterogeneous Ozark landscape is consistent with a local selection hypothesis. While we recognize that correlation does not demonstrate cause and effect (Linhart and Grant 1996), identifying a pattern is the first step in demonstrating the role of natural selection (Endler 1986). In our study, *Carya, Quercus,* and *Sassafras* each exhibited an association with a different set of forest-structure components. *Quercus* and *Carya* genotypes varied with large tree density or total basal area, respectively (Table 2, Fig. 2). Populations of *Quercus* in stands with many large trees may have been established earlier than those with lower densities of large trees. Older sites may include less shade for a longer period of time, which, in turn, may favor different genotypes. Similarly, *Carya* genotypes may vary with successional status or local habitat characteristics associated with basal area. *Sassafras* populations are known to be among the first species to invade an abandoned agricultural land, which may be related to our finding that *Sassafras* multivariate genotypes varied mostly with canopy cover. It is not surprising to find species-specific responses to ecological factors, yet we cannot determine whether these forest-structure variables are responsible for genotypic differentiation, or that they are correlated with environmental factors, such as microclimate, successional status, or disturbance history, that influence the study species.

The patterns that we observed indicate that populations found in local habitat patches with different forest structure characteristics also differ in combina-

tions of multivariate genotypes. Although our multivariate analyses produce gradients of genotypic and environmental scores, the populations along these gradients are not contiguous. In fact, we do not find any spatial autocorrelation in our genetic or environmental axes. Linhart and Grant (1996) proposed that genetic differentiation produced by biotic agents such as parasites or herbivores are likely to create such a pattern. Allard et al. (1993), who studied patterns of diploid and tetraploid multiallelic associations in *Avena hirtula* and *Avena barbata* in California and Spain, found a mosaic of multivariate genotypes that mapped precisely onto environmental patterns of moisture and temperature variation. The Allard et al. study has an implied causal mechanism for the genetic mosaic. In our case, we are unable to identify which biotic or abiotic environmental factors are likely causal agents, but we conclude that the forest structure variables, as bioassays of the environment, provide good evidence that genetic heterogeneity is influenced by local selection pressures.

Differential flowering times among microenvironments could also cause genetic heterogeneity. Jackson (1966) elegantly demonstrated that herbaceous plants differ in their flowering phenology in response to local microclimate. Overlap in flowering time is directly related to gene exchange; little overlap leads to a greater potential for genetic differentiation among populations. In *Carya, Quercus,* and *Sassafras,* we did not observe the degree of temporal isolation in flowering that would be required for genetic isolation. Furthermore, it is unlikely that phenological differences would be sufficient to create the observed genetic patterns for all three species, even though differential gene exchange and natural selection might work together to create genetic mosaics (Linhart and Grant 1996).

We also point out that genetic drift via founder effects may contribute to genetic heterogeneity. In the late 1880s, a large portion of Ozark forests was harvested extensively for timber, and then managed for pasture through grazing and fire (Cunningham and Hauser 1989). Since the 1920s, the Missouri Department of Conservation has purchased large tracts of land, suppressed fires, and managed the region for forest, allowing tree populations to become re-established by stump sprouts, scattered saplings, and newly established seedlings. Such colonization by different founders could increase genetic differentiation among populations, but it is difficult to explain how random founder events would result in a correlation between genetic and environmental variables. It is possible that many of the founders of current populations were stump sprouts from cut trees and saplings that had escaped removal (R. Jensen and D. Dey, *personal communication*). If so, then these founders would actually reflect the impact of selection on forest populations before clear-cutting. To the extent that natural selection has contributed to the heterogeneity that we observe, we cannot conclude whether it was natural selection on the founders of current populations in predisturbance forest, or natural selection on the individuals who have established postdisturbance. Nonetheless, we conclude that founder effects have probably interacted synergistically with natural selection to create genetic heterogeneity.

We suggest that gene flow should homogenize genetic variation rather than cause genetic structure, at least in *Carya* and *Quercus,* because wind-pollinated species are usually associated with high gene flow (Loveless and Hamrick 1984, Hamrick et al. 1995). However, we should point out that these forest stands were established after the early 1900s, from trees that were stump sprouts or escaped cutting. Thus, these tree populations are not at genetic equilibrium and may not have had a chance to establish a genetic structure that reflects the outcome of all evolutionary forces. Under these circumstances, founder effects, especially in combination with natural selection, have tremendous potential to influence the genetic mosaic without any mitigating effects of extensive gene flow. Only time will demonstrate whether selection will override gene flow, if it is indeed extensive.

Finally, our analyses did not detect significant differences in multivariate genotypes among soil-type/aspect classes in any of the three study species. Populations within each soil-type/aspect class differed significantly, which demonstrates local genetic heterogeneity, but soil type and aspect alone were not the primary factors associated with genotypic differentiation. These broad, abiotic characteristics probably interact with other local factors to create microenvironments, which, in turn, may influence genetic composition. These findings are consistent with multivariate analyses of genetic structure of these three species (Koop 1996, Sork et al. 1999). They found significant population differentiation within aspect class and site (17–30% of total variation) but little of the total variation was attributable to differences between the aspect classes (<3%). Thus, genotypic differentiation appears to be operating on a finer scale than soil or aspect differences.

We conclude that natural selection, possibly interacting with founder events, has influenced the population differentiation of these three long-lived plant species. Regardless of the evolutionary mechanisms responsible for genetic differentiation, environmental heterogeneity is likely to create genetic structure. Ecologists need to be aware of the extent to which plant populations can be genetically different even for populations within gene flow distance of each other. Just as community ecologists recognize variation in species composition across complex landscapes, we will benefit from the notion that populations are a community of genes with alleles that vary across heterogeneous environments.

ACKNOWLEDGMENTS

This project would not have been possible without the cooperative efforts of the people associated with the Missouri Ozark Forest Ecosystem Project, a collaborative research project of the Missouri Department of Conservation (MDC). We thank T. Treiman (MDC) for assistance with the forestry data, and R. Westfall for statistical advice. A. L. Koop, M. A. de la Fuente, and P. Foster conducted laboratory and field work. A. L. Koop generated raw genetic data on *Quercus alba* as part of his M.S. thesis at the University of Missouri–St. Louis. MDC provided awards to V. L. Sork for field and laboratory assistance and for support of the postdoctoral associate, W. K. Gram. V. L. Sork was supported by a sabbatical fellowship at the National Center for Ecological Analysis and Synthesis at the University of California–Santa Barbara (a Center funded by NSF grant DEB-92-21535), the University of California–Santa Barbara, and the State of California. Finally, we thank V. Apsit, K. Bailey, R. Dyer, J. Fernandez, C. Kelly, J. Wall, R. Westfall, E. Wiener, K. Clay, and two anonymous reviewers for comments on this manuscript.

LITERATURE CITED

Allard, R. W., G. R. Babbel, M. T. Clegg, and A. L. Kahler. 1972. Evidence for coadaptation in *Avena barbata.* Proceedings of the National Academy of Sciences, USA **69**: 3043–3048.

Allard, R. W., P. Garcia, L. E. Saenz-de-Miera, and M. Peres de la Vega. 1993. Evolution of multilocus genetic structure in *Avena hirtula* and *Avena barbata.* Genetics **135**:1125–1139.

Antonovics, J. 1971. The effects of a heterogeneous environment on the genetics of natural populations. American Scientist **59**:593–599.

Antonovics, J., and A. D. Bradshaw. 1970. Evolution in closely adjacent plant populations. 8. Clinal patterns at a mine boundary. Heredity **25**:349–362.

Antonovics, J., P. H. Thrall, and A. M. Jarosz. 1997. Genetics and the spatial ecology of species interactions: the Silene–Ustilago System. Pages 158–180 *in* D. Tilman and P. Kareiva, editors. Spatial ecology. The role of space in population dynamics and interspecific interactions. Princeton University Press, Princeton, New Jersey, USA.

Aradhya, K. M., D. D. Mueller, and T. A. Ranker. 1993. Genetic structure and differentiation in *Metrosideros-Polymorpha* Myrtaceae along altitudinal gradients in Maui Hawaii. Source Genetical Research **61**:159–170.

Brookshire, B. L., and S. R. Shifley, editors. 1997. Proceedings of the Missouri Ozark Forest Ecosystem Project symposium: an experimental approach to landscape research. General Technical Report NC-193. U.S. Forest Service, St. Paul, Minnesota, USA.

Cunningham, R. J., and C. Hauser. 1989. The decline of the Missouri Ozark forest between 1880 and 1920. Pages 34–37 *in* T.A. Waldrup, editor. Proceedings of pine–hardwood mixtures: a symposium on management and ecology of the type. General Technical Report 58. U.S. Forest Service, Atlanta, Georgia, USA.

Dirzo, R., and J. L. Harper. 1982. Experimental studies on slug–plant interactions. IV. The performance of cyanogenic and acyanogenic morphs of *Trifolium repens* in the field. Journal of Ecology **70**:119–138.

Dudley, S. A. 1996*a*. Differing selection on plant physiological traits in response to environmental water availability: a test of adaptive hypotheses. Evolution **50**:92–102.

Dudley, S. A. 1996*b*. The response to differing selection on plant physiological traits: evidence for local adaptation. Evolution **50**:103–110.

Endler, J. A. 1973. Gene flow and population differentiation. Science **179**:243–250.

Endler, J. A. 1986. Natural selection in the wild. Princeton University Press, Princeton, New Jersey, USA.

Gram, W. K., and V. L. Sork. 1999. Population density as a predictor of genetic variation for woody plant species. Conservation Biology **13**:1079–1087.

Hamrick, J. L., and R. W. Allard. 1972. Microgeographical variation in allozyme frequencies in *Avena barbata.* Proceedings of the National Academy of Sciences USA **69**: 2100–2104.

Hamrick, J. L., M. J. W. Godt, and S. L. Sherman-Broyles. 1995. Gene flow among plant populations: evidence from genetic markers. Pages 215–232 *in* P. C. Hoch and A. G. Stephenson, editors. Experimental and molecular approaches to plant biosystematics. Missouri Botanical Garden, St. Louis, Missouri, USA.

Husband, B. C., and S. C. H. Barrett. 1996. A metapopulation perspective in plant population biology. Journal of Ecology **84**:461–469.

Jackson, M. T. 1966. Effects of microclimate on spring flowering phenology. Ecology **47**:407–415.

Kabrick, J. M., D. R. Larsen, and S. R. Shifley. 1997. Analysis of pre-treatment MOFEP woody vegetation and environmental data: evaluation for compartment, block, and treatment trends in species counts, density, diameter and basal area. Pages 150–168 *in* B. L. Brookshire and S. R. Shifley, editors. Proceedings of Missouri Ozark Forest Ecosystem Project symposium: an experimental approach to landscape research. General Technical Report NC-193. U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station, St. Paul, Minnesota, USA.

Kephart, S. R. 1990. Starch gel electrophoresis of plant isozymes: a comparative analysis of techniques. American Journal of Botany **77**:693–712.

Koop, A. L. 1996. Genetic variation and structure in *Quercus alba* L. in a Missouri Ozark landscape. Thesis. University of Missouri, St. Louis, Missouri, USA.

Kremer, A., and A. Zanetto. 1997. Geographical structure of gene diversity in *Quercus petraea* (Matt.) Liebl. II: multilocus patterns of variation. Heredity **78**:476–489.

Kurzejeski, E. W., R. L. Clawson, R. B. Renken, S. L. Sheriff, L. D. Vangilder, C. Hauser, and J. Faaborg. 1993. Experimental evaluation of forest management: the Missouri Ozark Forest Ecosystem Project. Pages 599–609 *in* Transactions of the 58th North American Wildlife and Natural Resource Conference. Washington D.C., USA.

Latta, R. G. 1998. Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. American Naturalist **151**:283–292.

Linhart, Y. B., and M. C. Grant. 1996. Evolutionary significance of local genetic differentiation in plants. Annual Review of Ecology and Systematics **27**:237–277.

Loveless, M. D., and J. L. Hamrick. 1984. Ecological determinants of genetic structure in plant populations. Annual Review of Ecological Systematics **15**:65–95.

Mahalanobis, P. C. 1948. Historic note on the $D^2$ statistic. Sankhya **9**:237.

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

Meinert, D., T. Nigh, and J. Kabrick. 1997. Landforms, geology and soils of the MOFEP study sites. Pages 56–68 *in* B. L. Brookshire and S. R. Shifley, editors. Proceedings of Missouri Ozark Forest Ecosystem Project symposium: an experimental approach to landscape research. General Technical Report NC-193. U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station, St. Paul, Minnesota, USA.

Miller, M. R. 1981. Ecological Land Classification Terrestrial Subsystem—A basic inventory system for planning and management of the Mark Twain National Forest. U.S. De-

partment of Agriculture, Forest Service, Rolla, Missouri, USA.

Mitton, J. B. 1997. Selection in natural populations. Oxford University Press, Oxford, UK.

Mitton, J. B., M. C. Grant, and A. M. Yoshino. 1998. Variation in allozymes and stomatal size in Pinyon (*Pinus edulis,* Pinaceae), associated with soil moisture. American Journal of Botany **85**:1262–1265.

Mitton, J. B., K. B. Sturgeon, and M. L. Davis. 1980. Genetic differentiation in ponderosa pine along a steep elevational transect. Silvae Genetica **29**:100–103.

Murawski, D. A., T. H. Fleming, K. Ritland, and J. L. Hamrick. 1994. Mating system of *Pachycereus pringlei*: an autotetraploid cactus. Heredity **72**:86–94.

Neigel, J. E. 1997. A comparison of alternative strategies for estimating gene flow from genetic markers. Annual Review of Ecology and Systematics **28**:105–128.

SAS Institute. 1996. Version 6.12. SAS Institute, Cary, North Carolina, USA.

Slatkin, M. 1985. Gene flow in natural populations. Annual Review of Ecology and Systematics **16**:393–430.

Smouse, P. E., R. S. Spielman, and M. H. Park. 1982. Multiple-locus allocation of individuals to groups as a function of the genetic variation within and differences among human populations. American Naturalist **119**:445–463.

Smouse, P. E., and R. C. Williams. 1982. Multivariate analysis of HLA-Disease associations. Biometrics **38**:757–768.

Sork, V. L., A. Koop, M. A. de la Fuente, P. Foster, and J. Raveill. 1997. Patterns of genetic variation in woody plant species in the Missouri Ozark Forest Ecosystem Project (MOFEP). Pages 231–247 *in* B. L. Brookshire and S. R. Shifley, editors. Proceedings of Missouri Ozark Forest Ecosystem Project symposium: an experimental approach to landscape research. Gen. Tech. Rep. NC-193. U.S. Department of Agriculture, Forest Service, St. Paul, Minnesota, USA.

Sork, V. L., J. Nason, and D. Campbell. 1999. Landscape approaches to the study of gene flow in plants. Trends in Ecology and Evolution **14**:219–224.

Sork, V. L., A. R. Templeton, R. D. Westfall, M. A. de la Fuente, P. Foster, and A. L. Koop. *In press.* Single and multi-locus genetic structure of tree populations in a heterogeneous landscape. Evolution.

Taylor, C. E., and J. B. Mitton. 1974. Multivariate analysis of genetic variation. Genetics **76**:575–585.

Thorpe, R. S. 1985. Clines: character number and the multivariate analysis of simple patterns of geographic variation. Biological Journal of Linnean Society **26**:201–214.

Westfall, R. D., and M. T. Conkle. 1992. Allozyme markers in breeding zone designation. New Forests **6**:279–309.

Yang, R. C., and F. C. Yeh. 1993. Multilocus structure in *Pinus contorta* Dougl. Theoretical and Applied Genetics **87**:568–576.

Yeh, F. C., M. A. Khalil, Y. El-Kassaby, and D. C. Trust. 1986. Allozyme variation in *Picea mariana* from Newfoundland: genetic diversity, population structure and analysis of differentiation. Canadian Journal of Forest Research **16**:713–720.