ANALYSIS OF POPULATION STRUCTURE FROM ALLOZYME FREQUENCIES

R. P. GURIES and F. T. LEDIG[1/]

Abstract -- Selection strategies for tree improvement depend on population structure which will probably vary among species and even among regions within a species' range. Analysis of population structure using allozyme information can help in determining improvement strategy. Population structure can be described in terms of inbreeding, the grouping of genotypes within stands, and genetic distances among stands.

There was no evidence of inbreeding in stands of pitch pine based on analysis of 15 allozyme loci, nor was there evidence of familial groups within stands. However, there were differences among stands over distances of several kilometers. The implication for tree improvement strategy in pitch pine is to use comparison-tree methods of selection, and for provenance testing, to expend little effort on sampling different stands within regions.

INTRODUCTION

An important aspect of population structure is the subdivision of a species into smaller breeding units. In forest trees, adjacent individuals may have greater opportunities for mating than those separated by distance. Also, the closer two trees, the greater the likelihood they are related because of limited seed dispersal. Such situations could produce stands characterized by moderate levels of inbreeding and broken into family groups or neighborhoods. Alternatively, pollen and seed dispersal may be so great that, genotypes are distributed at random within stands. Strategy and tactics of tree improvement should be chosen to take advantage of actual population structure, and in fact, tactics inappropriate to the population structure may limit improvement.

Allozymes are allelic variants of genes which code for enzymes and other proteins. Allozyme frequencies have been used by several authors to study population structure in forest trees (Feret, 1974; Sakai and Park, 1971; Sakai, Miyazaki and Matsuura, 1972; Rudin et al., 1974; Tigerstedt, 1973). In general, these workers found populations to be in close agreement with Hardy-Weinberg expectations, indicating that inbreeding was unlikely to be important. Nevertheless, differentiation between sub-populations was observed over distances as short as several hundred meters. Selection and genetic drift were suggested as factors responsible for the observed heterogeneity among populations.

As part of a continuing study of the genecology of pitch pine (Pinus rigida Mill.), we have used 15 allozyme loci to examine some aspects of population structure in natural stands. The analysis includes only populations from the central portion of the pitch pine range, and conclusions concerning the amount and distribution of variation may not apply to other ecological situations.

METHODS

Seed were collected from four populations of pitch pine in New Jersey, two

---

[1/] Research Associate and Associate Professor, School of Forestry and Environmental Studies, Yale University, New Haven, CT 06520.

from the dwarf Pine Plains (East Plains, N=153; West Plains, N=118) and two from the Pine Barrens (Lebanon State Forest, N=69; Helmetta, N=61). The Pine Plains and Lebanon populations are within 10 km of one another; the Helmetta population is located about 69 km to the north. All trees within a circle of approximately 50 m diameter and which were bearing cones (80 - 95% of all trees) were mapped and included in the sample. Handling of seed and conditions for horizontal starch gel electrophoresis are described in Guries and Ledig (1977). The loci analyzed included isocitrate dehydrogenase (IDH), fumarase (FUM), leucine aminopeptidase (LAP-1), acid phosphatase (ACP), aconitase (ACO), and two each of glutamate oxalate transaminase (GOT-1, GOT-2), malate dehydrogenase (MDH-1, MDH-2), phosphoglucomutase (PGM-1, PGM-2), glucose phosphate isomerase (GPI-1, GPI-2), and 6-phosphogluconate dehydrogenase (6-PGD-1, 6-PGD-2).

RESULTS

Within population comparisons

The frequency of the most common allele varied from 0.598 to 1.000 depending on gene locus and population (Table 1). Only three loci, GOT-2 and ACO in the West Plains population and 10H-1 in the Helmetta population, showed a significant deviation from the Hardy-Weinberg frequencies expected under random mating. Three deviations out of sixty cases could be expected by chance. However, it has often been noted that the Chi-square goodness-of-fit test lacks the statistical power to detect deviations due to inbreeding unless the sample sizes are very large or the inbreeding is pronounced (Ward and Sing, 1970; Smith, 1970). Inbreeding coefficients averaged over all loci were virtually zero for each stand (a maximum of 0.016).

Table 1.-- Frequency of the most common allele for 15 allozyme loci.

	Population			
Locus	East Plains	West Plains	Lebanon	Helmetta
MDH-1	.970	.997	.964	.959
MDH-2	.741	.765	.775	.721
IDH	.924	.850	.913	.959
FUM	.997	.996	.986	.984
PGM-1	.977	.962	.978	.992
PGM-2	.974	.987	.971	.934
GPI-1	.997	.991	1.000	1.000
GPI-2	.951	.974	.986	.967
6-PGD-1	.810	.716	.768	.623
6-PGD-2	.670	.754	.703	.598
LAP-1	.847	.912	.899	.951
GOT-1	.951	.924	.899	.893
GOT-2	.967	.934	.920	.951
ACP	.971	.962	.942	.926
ACO	.680	.725	.680	.598

Examination of the pattern of genotype distribution in the stands also provided no indication of inbreeding, i.e. clustering of genotypes. The distribution of MDH-2 genotypes in the Helmetta population (Fig. 1) is typical of the patterns observed for other loci in these stands. To statistically examine relationships among individuals with respect to their positions within the stand, we developed a genetic similarity value. Genetic similarity between individuals was calculated by averaging the number of alleles held in common over all loci. Each allele was weighted by its frequency. The measure is a genetic analogy to the "disagreement count" of Sakai and Miyazaki (1972) for scoring phenotypic banding patterns in peroxidase isozymes. Genetic similarity did not decrease with distance between individuals (Fig. 2).We suggest that the distribution of genotypes is essentially at random over areas as large as 2000 m^2 (0.2 hectares).

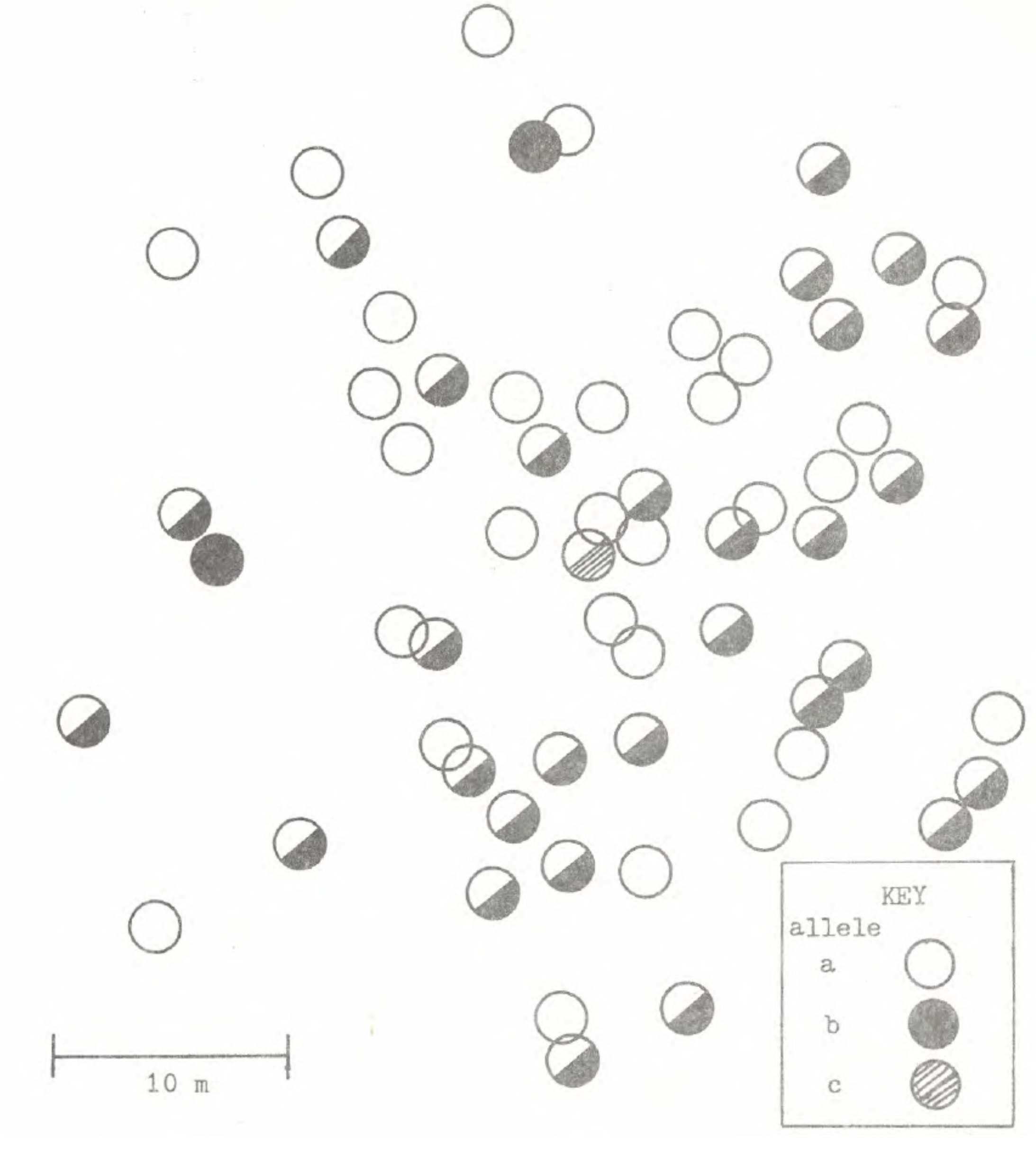

Figure 1. Spatial distribution of MDH-2 genotypes in the Helmetta population. Solid circles are homozygotes; split circles are heterozygotes.

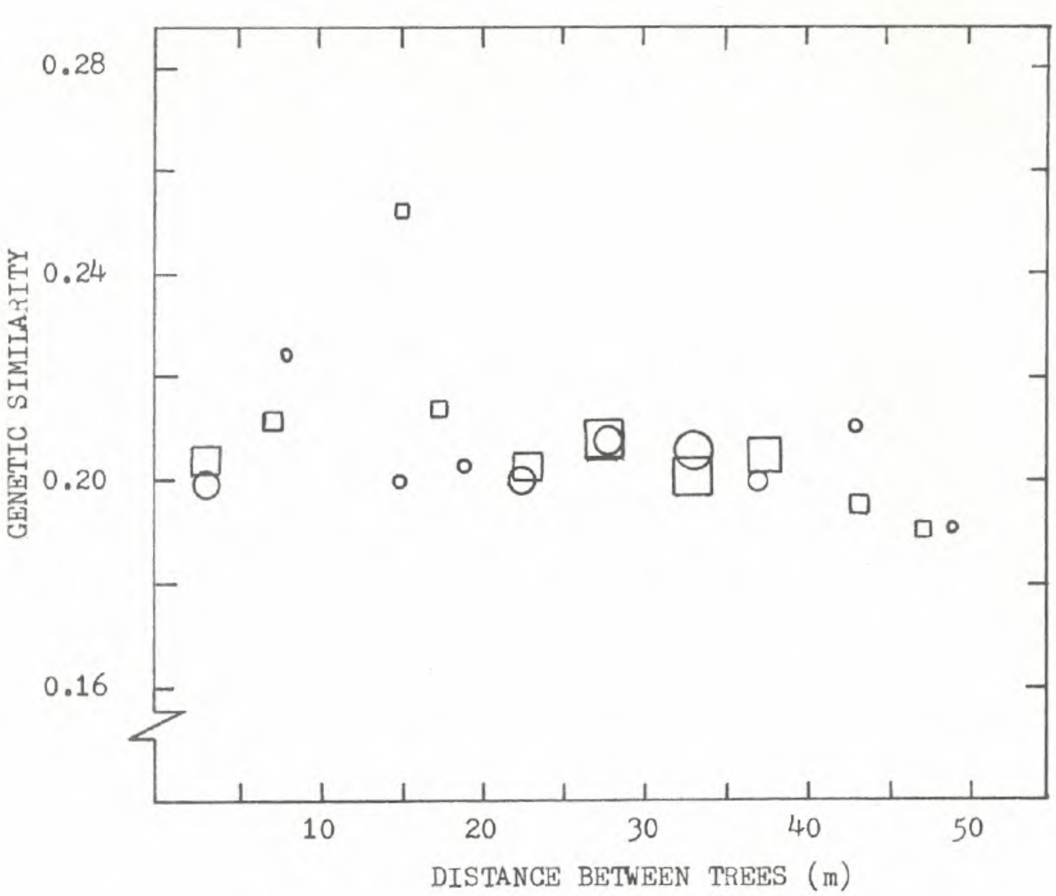

Figure 2. - Genetic similarity for trees 346 (○) and 190 (□) compared to 67 neighboring trees in relation to the distance between 346 or 190 and their neighbors. Points are means by 5 m intervals; size of symbol indicates the number of values included, from 1 to 15. Results for this pair of trees is typical for the Lebanon population, and regression for all 69 trees was approximately zero.

between population analysis of gene frequency Heterogeneity

Population subdivision is expected to result in a heterogeneous distribution of allele frequencies between subpopulations. Variations in allele frequencies can be tested by Chi-square contingency test (Workman and Niswander, 1970). There was significant heterogeneity of allele frequencies at most loci (Table 2), suggesting that the four populations of pitch pine were genetically differentiated. The variation in Chi-square values suggests that different loci are probably responding independently to the factors causing the heterogeneity.

Genetic distance

An alternative method of analyzing population structure is that of genetic distance. This technique measures the accumulated number of gene differences per locus among populations and expresses them in a simple index (Nei, 1972). If population differentiation is largely the result of isolation by distance, then spatial distance and genetic distance are expected to be positively related; i.e. as physical distance increases, genetic distance should also increase.

Table 2.-- Chi-square analysis of gene frequency heterogeneity among four pitch pine populations.

Locus	$X^2(3)$	p
MDH-1	14.90	<.005
MDH-2	4.36	.10<p<.25
IDH	14.84	<.005
FUM	7.51	.05<p<.10
PGM-1	3.19	.25<p<.50
PGM-2	24.90	<.005
GPI-1	5.62	.10<p<.25
GPI-2	8.76	.01<p<.05
6-PGD-1	38.01	<.005
6-PGD-2	13.04	<.005
LAP-1	19.54	<.005
GOT-1	15.06	<.005
GOT-2	5.70	.10<p<.25
ACP	5.20	.10<p<.25
ACO	22.26	<.005

Under an isolation by distance model, the Helmetta population, located some 69 km from the other three populations, would be expected to have the greatest genetic distance. However, the values for the Helmetta-Pine Plains comparisons were lower than those for the Lebanon-Pine Plains comparisons (Table 3), even though Lebanon is separated from the West Plains by a mere 5 km. The fact that large differences were not observed between Helmetta and the other populations suggests that isolation by distance may not be the major factor in population differentiation of pitch pine. While it is tempting to extrapolate to other wind-pollinated conifers, this conclusion must be accepted with caution awing to the small number of populations compared here.

Table 3.__ Estimates of genetic distance among four pitch pine populations. Values above diagonal are estimates of genetic distance, D; values below diagonal are estimates of the normalized identity of genes, I.

	Populations			
	East Plains	West Plains	Lebanon	Helmetta
East Plains		.0034	.0123	.0055
West Plains	.9966		.0102	.0065
Lebanon	.9877	.9899		.0147
Helmetta	.9945	.9935	$9854	

DISCUSSION

Pitch pine of the Pine Plains' and Pine Barrens' populations appear to be composed of individuals distributed randomly with respect to genotype. Vagile pollen and seed dispersal are the factors most likely to produce such a distribution over areas the size of those studied. The heterogeneity of gene frequencies among populations indicates that differentiation has occurred over distances of several kilometers. Isolation by distance may be partially responsible for this differentiation, but additional factors such as environmental heterogeneity and genetic drift could also be involved. It is interesting to note that the dwarf populations of the Pine Plains are not radically different from the Pine Barrens' populations in terms of allozyme frequencies. These dwarf populations have originated and been maintained without extensive genic differentiation of the soluble enzymes investigated here.

Population structure should determine the pattern of sampling for provenance selection. If stands within regions are widely separated relative to pollen and seed dispersal distance, then adjacent stands may differentiate, and several stands will be required to adequately sample a region. Alternatively, large population size with wide outcrossing leads to predictable patterns of gene frequencies, and sampling need not be as intense.

Distribution of genotypes within stands also affects selection tactics and resulting gain. If adjacent individuals tend to be related, then rigidly applied comparison-tree methods of selection will provide less improvement than base-line selection (Ledig, 1970. Alternatively, if genotypes are randomly distributed without subdivision into local neighborhoods, then comparison-tree selection is favored because it provides a correction for environmental differences among select-tree candidates.

For pitch pine, inbreeding within clusters as small as 0.2 hectares is essentially non-existent, and genotypes are randomly distributed. Such a situation favors comparison-tree selection. While true for pitch pine, other species may differ in population structure as a result of ecological and genetic factors. For example, in white spruce (Picea glauca (Moench.) Voss.), and in Japanese arbor-vitae (Thujopsis dolabrata Sieb. & Zucc.) there was indication of substantial relationship among neighboring trees, which decreased as distance among trees increased (Coles and Fowler, 1976; Sakai and Miyazaki, 1972). In black spruce (Picea mariana (Mill.) BSP) the relationship among neighboring trees was slight except in small, isolated stands at the southern extreme of the range (Morgenstern, 1972). In Norway spruce (Picea abies (L.) Karst.) there was a completely random distribution of genotypes within stands (Tigerstedt, 1973). Indirect evidence suggests relationships among adjacent trees in some yellow pines (Snyder, 1969; pers. comm., 1973; Sittman and Tyson, 1971).

The relative amounts of genetic variation among individuals within stands and among stands or regions should also influence selection method. For allozymes there was little variation among stands within the same region, further suggesting the utility of the comparison-tree method f selection. The same conclusion was reached by comparison of provenance growth in nursery trials (Ledig et al., 1976). Genetic distances for allozyme loci did not increase appreciably with geographic separation, although admittedly results were based on limited comparisons. Differences among provenances in growth did increase clinally with respect to either latitude or climatic variables (Ledig et al., 1976). Thus, there must be some point at which variation among populations exceeds variation within, favoring some emphasis on provenance selection. Therefore, it seems reasonable to conclude that selection within regions should be sufficient to sample almost all the genetic variability available, although provenance selection might be a more efficient means of achieving some objectives. The implications for provenance versus within-provenance selection are not as clear as those for within-stand selection and further study is required.

The technique used here could be applied to any conifer, and would yield results in a short time. The benefit in planning tree improvement operations could be substantial, and studies using additional conifer species are currently under way.

ACKNOWLEDGEMENTS

This research was supported by National Science Foundation Grant DEB74-11794.

LITERATURE CITED

Coles, J. F., and D. P. Fowler. 1976. Inbreeding in neighboring trees in two white spruce populations. Silvae Genet. 25:29-35.

Feret, P. P. 1974. Genetic differences among three small stands of *Pinus pungens.* Theor. Appl. Genet. 44:173-177.

Guries, R. P., and F. T. Ledig. 1977. Inheritance of some polymorphic isoenzymes in pitch pine (*Pinus rigida* Mill.). Heredity: in press.

Ledig, F. T. 1974. An analysis of methods for the selection of trees from wild stands. For. Sci. 20:2-16.

Ledig, F. T., C. C. Lambeth, and D. I. H. Linzer. 1976. Nursery evaluation of a pitch pine provenance trial. In Proc. Twenty-third Northeast.For. Tree Improv. Conf., New Brunswick, NJ, pp. 93-108.

Morgenstern, E. K. 1972. Preliminary estimates of inbreeding in natural populations of black spruce, *Picea mariana.* Can. J. Genet. & Cytol. 14:443-446.

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

Rudin, D., G. Eriksson, I. Ekberg, and M. Rasmuson. 1974. Studies of allele frequencies and inbreeding in Scots pine populations by the aid of the isozyme technique. Silvae Genet. 23:10-13.

Sakai, K.-I., and Y. Miyazaki. 1972. Genetic studies in natural populations of forest trees. II. Family analysis: a new method for quantitative genetics studies. Silvae Genet. 21:149-154.

Sakai, K.-I., and Y. Park. 1971. Genetic studies in natural populations of forest trees. III. Genetic differentiation within a forest of *Cryptomeria japonica.* Theor. Appl. Genet. 41:13-17.

Sakai, K.-I., Y. Miyazaki, and T. Matsuura. 1972. Genetic studies in natural populations of forest trees. I. Genetic variability on the enzymatic level in natural forests of *Thujopsis dolabrata.* Silvae Genet, 20:168-173.

Sittman, K., and H. Tyson. 1971. Estimates of inbreeding in *Pinus banksiana.* Can. J. Bot. 49:1241-1245.

Smith, C. A. B. 1970. A note on testing the Hardy-Weinberg law. Ann. Human Genet. 33:377-383.

Snyder, E. B. 1969. Parental selection versus half-sib family selection of longleaf pine. In Proc. Tenth So. Conf. For. Tree Improv., Houston, TX, pp. 84-88.

Tigerstedt, P. M. A. 1973. Studies on isozyme variation in marginal and central populations of *Picea abies* L. Hereditas 75:47-60.

Ward, R. H., and C. F. Sing. 1970. A consideration of the power of the X2 test to detect inbreeding effects in natural populations. Am. Nat. 104: 355-363.

Workman, P. L., and J. D. Niswander, 1970. Population studies on southwestern Indian tribes. II. Local genetic differentiation in the Papago. Am. J. Human Genet. 22:24-49.