GENETIC VARIATION IN JACK PINE: CONCORDANCE BETWEEN ELECTROPHORETIC AND METRIC DATA FOR LAKE STATES' POPULATIONS[1]

David M. O'Malley and R. P. Guries[2]

Abstract. Concordance between patterns of allozyme variation and metric variation was assessed for populations of jack pine in the Lake States. An ordination of principal components based on allozyme data revealed two broad groupings of populations, a southern group, and a central and northern group. These groups correspond to those identified by Jeffers and Jenson (1980) based on height and diameter growth after 20 years. Patterns of variation for both allozyme and metric data could be related, to growing degree days. However, each data base provides a somewhat different perspective on variation among jack pine populations in the Lake States.

Electrophoretic techniques offer a convenient method for detecting and quantifying genetic variation at a number of allozyme loci (Brown 1979). Studies of allozyme variation can reveal patterns of genetic differentiation but it has proven difficult to relate such patterns to information breeders consider important (eg. adaptation of seed sources to particular locations). In part, this is due to uncertainties concerning the nature of allozyme variation (Gottlieb 1981); the issue of the selective neutrality of allozyme polymorphisms remains controversial. It is also partly due to to the lack of reliable and extensive studies of seed source adaptation.

Variation in growth performance of 26 Lake States jack pine (Pinus banksiana Lamb.) seed sources at 14 planting sites throughout the Lake States was shown to be continuous and related to climatic variables, although some sources expressed adaptation to local conditions (Jeffers and Jenson 1980). The availability of detailed information on variation in growth among Lake States jack pine seed sources provides an opportunity to assess, in a general way, the concordance between the patterns of allozyme variation as revealed in the current study, and patterns of variation in metric traits.

Materials and Methods

Seed from 18 populations of jack pine were collected for paired stands from 9 locations in Minnesota, Wisconsin, and Michigan. Populations in each pair were separated by 6-23 km (Figure 1, Table 1). Allozyme variation was analyzed using starch gel electrophoresis (O'Malley et al. 1980) to obtain estimates of gene frequencies. One megagametophyte per tree for 60 trees per

[1] Research supported in part by McIntire-Stennis Project No. 2385.

[2] Research Assistant and Associate Professor, respectively; Department of Forestry, School of Natural Resources, University of Wisconsin, Madison, WI 53706.

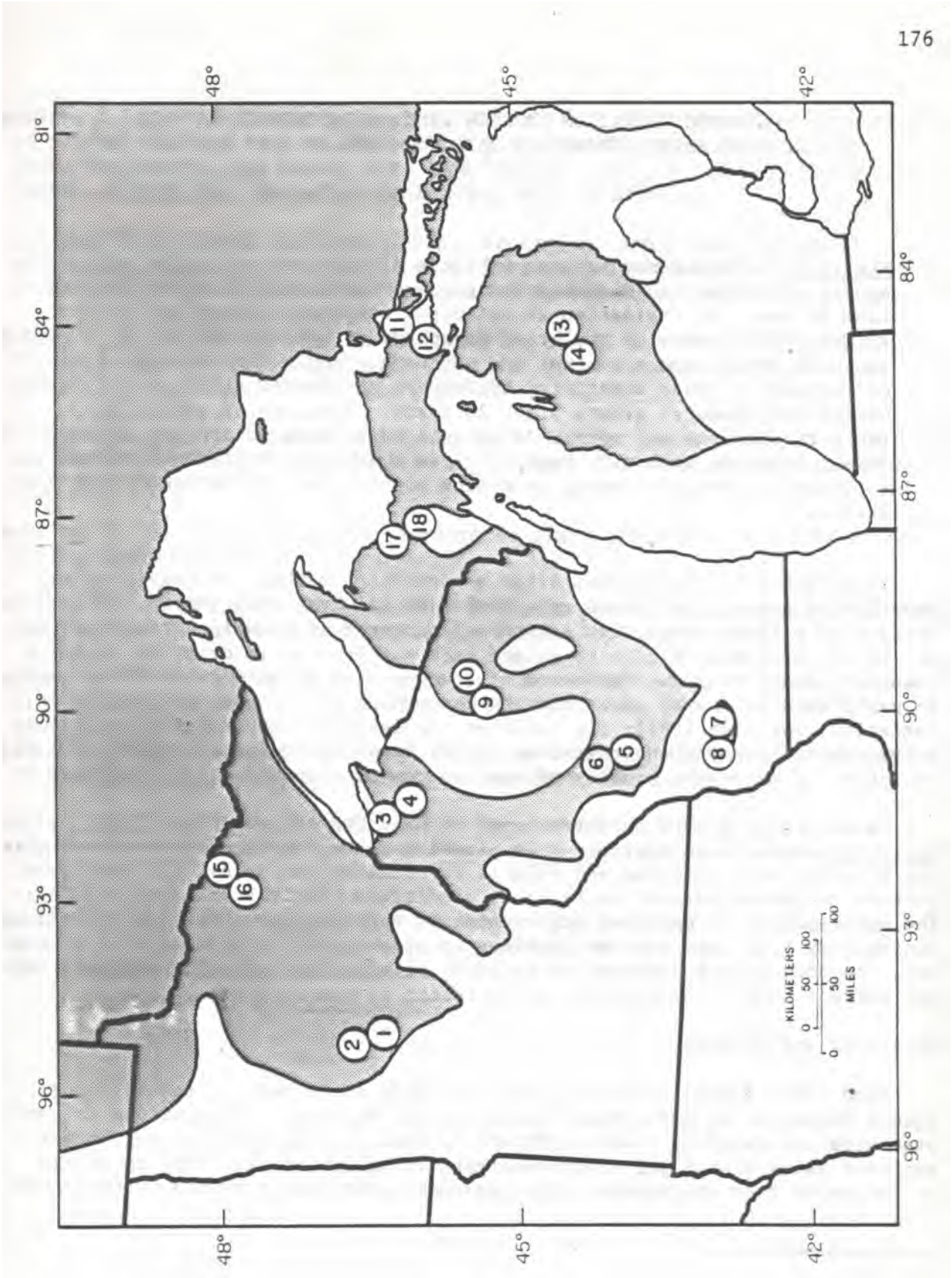

Figure 1. Locations of eighteen populations of jack pine included in study.

Table 1. Locations and distances between pairs of jack pine populations in Minnesota, Wisconsin, and Michigan. N is the number of trees sampled.

Location	Stand	Accn. Name	N	Lat.	Long.	Dist.
A	1	Germ	60	46*33	94*47'	23 km
	2	Shell	60	46*45'	94*55'	
B	3	Hug	61	46*48'	91*55'	10 km
	4	Dru	76	46*23'	91*26'	
C	5	Mil	60	44*10'	90*38'	11 km
	6	Ko	60	44*21	90*45'	
D	7	LR	60	43*12'	90*16'	16 km
	8	Mus	38	43*12'	90*30'	
E	9	Nok	60	45*37'	89*45'	6 km
	10	Har	60	45*39'	89*41'	
F	11	Pick	60	46*13'	84*26'	6 km
	12	Soo	60	46*10'	84*28'	
G	13	Ros	60	44*36'	84*37'	16 km
	14	Pig	60	44*37'	84*50'	
H	15	Et	60	48*10'	92*28'	23 km
	16	Hlr	60	48*00'	92*44'	
I	17	Sand	60	48*23'	87*25'	8 km
	18	Lit	38	46*18'	87*21'	

stand was scored for allelic variation. Assuming independence, this sample size is large enough to be certain of detecting at least one megagametophyte per stand carrying an allele present with a frequency of 0.05; genotyping was unnecessary for our objective. The frequency of the most common allele at 6 highly polymorphic loci was transformed (arcsin) and standardized before performing a principal components analysis based on the correlation matrix. The loci reported here encode the enzymes aconitase (ACO), acid phosphatase (ACP2), glyceric-2 dehydrogenase G2D, 6-phosphogluconic dehydrogenase (6PG1 and 6PG2), and uridine diphosphoglucosepyrophosphorylase (UGPP1) (Table 2).

Results

The frequency of the common allele at six allozyme loci varied substantially among the 18 paired populations of jack pine (Table 2). Variation among loci was not independent as shown by the correlations of gene frequencies at different loci (Table 2). When such a covariance structure exists, the variation in the data can often be summarized by a smaller number of variables using the multivariate technique of principal components (Seal

Table 2. Frequency of the most common allele at six polymorphic loci in jack pine, and the number of gametes sampled (n) in eighteen populations (see Table 1 for accession place name).

		GERM	SHEL	HUG	DRU	MIL	KO	LR	MUS	NOK	HAR	PIC	SOO	ROS	PIG	ET	HLR	SAND	LIT
Aco		.828	.793	.700	.808	.741	.678	.448	.635	.833	.723	.850	.678	.654	.763	.780	.817	.833	.817
	n	58	58	60	52	58	59	58	52	60	47	60	59	52	59	59	60	60	60
Acp		.828	.793	.833	.833	.877	.746	.793	.808	.800	.840	.700	.593	.855	.797	.783	.814	.833	.817
	n	58	58	60	60	57	59	58	50	60	50	60	59	55	59	60	59	60	60
G2dh		.586	.655	.586	.431	.474	.482	.632	.574	.667	.745	.407	.483	.667	.672	.644	.475	.500	.717
	n	58	58	58	58	57	56	57	54	60	55	59	58	54	58	59	59	60	60
6pgd1		.358	.362	.633	.533	.569	.695	.750	.731	.667	.392	.350	.534	.638	.776	.542	.559	.500	.500
	n	53	58	60	60	58	59	56	52	60	51	60	58	50	58	59	59	60	60
6Pgd2		.897	.879	.800	.967	.793	.881	.946	.885	.733	.765	.683	.948	.698	.862	.864	.814	.800	.783
	n	58	58	60	60	58	59	56	52	60	51	60	58	53	58	59	59	60	60
Ugpp1		.845	.867	.750	.911	.638	.627	.679	.788	.883	.804	.800	.897	.769	.831	.881	.949	.767	.867
	n	58	60	60	56	58	59	53	52	60	51	60	58	52	59	59	59	60	60

1964). Principal components were constructed to summarize independent trends in the original data. The first principal component of our gene frequency data summarizes 34% of the total variability and is highly correlated with ACO, 6PG1, and UGPP1. Similarly, the second principle component summarizes 24% of the total variability and is highly correlated with ACP2, 6PG2, and G2D. The first four principal components account for 89% of the variability in gene frequencies at these six loci.

Jeffers and Jenson (1980) divided Lake States seed sources into three groups; northern, central and southern based upon correlations among seed sources at different planting sites. To compare their metric data with our allozyme data, we reanalyzed their height growth data at twelve planting locations using principal components analysis. Their original divisions are retained and easily recognized from an ordination of the first two principal components based upon height growth variation for 26 seed sources at 12 planting sites (Figure 2a). The first principal component is correlated with seed source performance at nearly all planting sites and summarizes 56.5% of the total variance. The second principal component summarizes 16.4% of the total variance and is correlated with means at two plantations, Superior National Forest and Argonne Experimental Forest, the two coldest sites included.

We initially assigned our populations to the three groups recognized by Jeffers and Jenson (1980). However, our allozyme data was unable to separate two populations from northeastern Minnesota (St. Louis Co.) from those in central Minnesota, Upper Peninsula Michigan and northern Wisconsin (Figure 2b). Regression of the first principal component of the gene frequency data on a northern plus central vs southern classification yields an R of 65%. A higher R can be obtained by regressing this variable on growing degree days (70%) following Rudoph (1956). The other principal components were not readily interpreted.

Discussion

Gene frequency variation at these six loci allozyme appears to be continuous among Lake States jack pine populations. The southern Wisconsin populations (LR and MUS) differed substantially from the others for several loci. However, they do not appear to be isolated outliers as neighboring locations show similar trends, though not as extreme as the frequencies found in these populations. These southern Wisconsin sites are very sandy; associated species include prickly pear cactus. The area has average July temperatures considerably warmer than northern Wisconsin, but it is in a large cold air drainage with average January minimum temperatures more typical of locations more than 100 miles to the north. The trees are non-serotinous and have seed weights much greater than average for jack pine. These southern Wisconsin populations account for a substantial portion of the gene frequency differentiation found for Lake States jack pine.

Jeffers and Jenson (1980) differentiated northern, central, and southern Lake States jack pine seed sources on the basis of correlations among sources for height, diameter, and volume at 20 years. Their groupings are strongly correlated with growing degree days at the seed source origins. This pattern of geographic variation in growth can be readily recognized in a plot of the first and second principal component which summarize 73% of the variation in

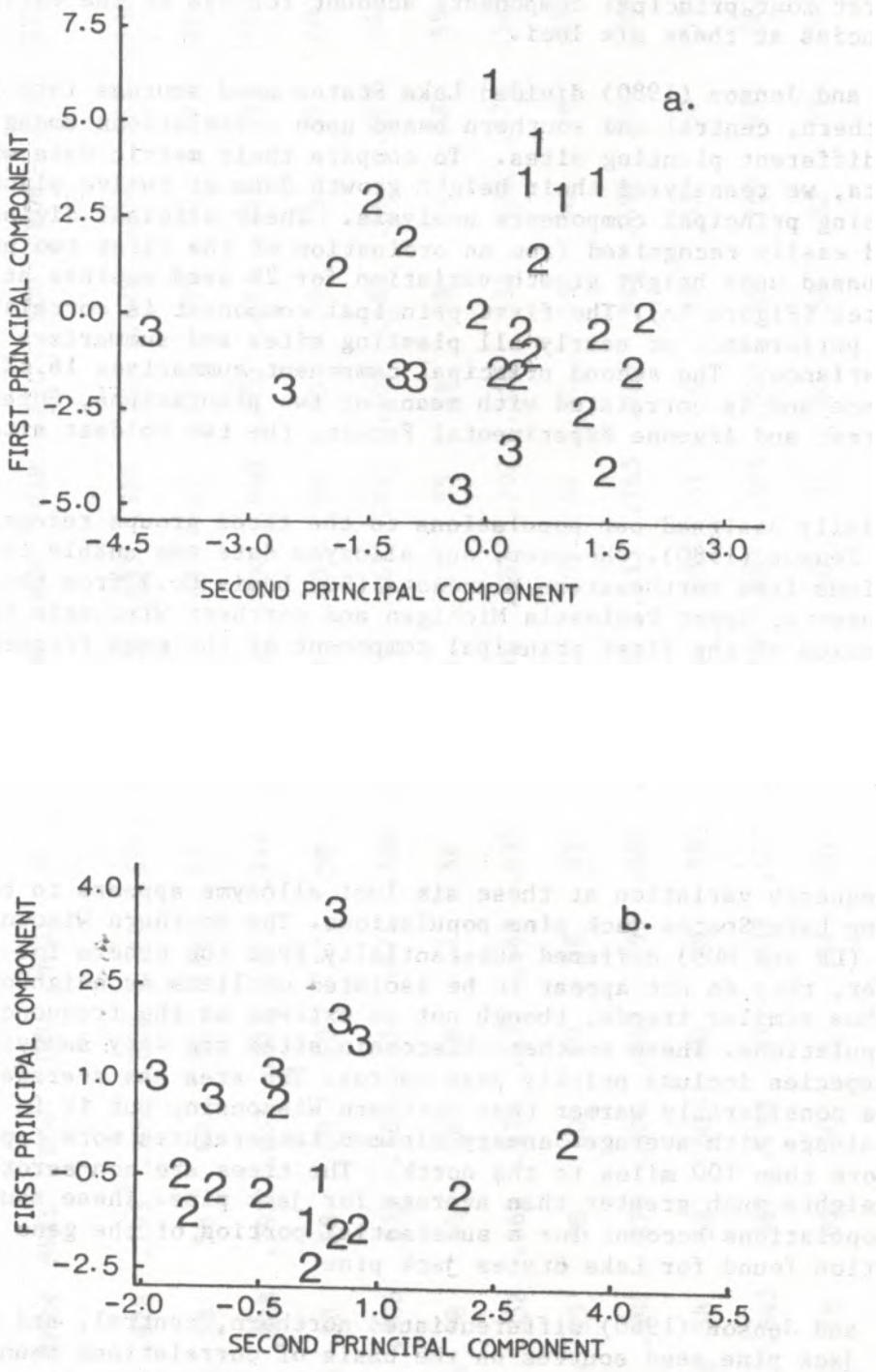

Figure 2. Ordination of principal components; a) metric data of Jeffers and Jenson (1980), and b) allozyme data of present study. See text for additional details.

height growth performance (Figure 2a). No large-scale geographic pattern is apparent for the values of the second principal component, but it may be related to adaptation of seed sources to local conditions.

Growing degree days account for a large portion of the variation in the first principal component of allozyme gene frequencies. This variable summarizes a trend in variation involving primarily three loci: ACO, 6PG1, and UGPP1. Thus, it appears that a portion of the gene frequency variation of these six allozyme loci follows the same trend noted by Jeffers and Jenson (1980) to account for the pattern of variation in growth rates among seed sources. It is not unreasonable to expect that different loci could be 'tracking' the same environmental gradients or similar identical patterns of migration. This is the sort of trend which can be detected through multivariate analysis. Regressions of the gene frequencies at ACO, 6PG1, and UGPP1 individually on the north/central vs south pattern of variation are each significant (40% < R < 55%), but the correlation of the first principal component with the classification is considerably stronger, hence the higher R^2 (65%).

Allozyme variation would be considered more useful in forestry if it could be related directly to traits of economic importance. Population studies of some selfing annuals such as Hordeum have detected allozyme variation patterns which are strongly correlated with edaphic and other environmental factors (Nevo et al. 1979). Studies in corn have demonstrated that selection for quantitative traits results in changes in allozyme gene frequencies, and reciprocally, selection on allozymes can yield changes in quantitative traits (Brown and Allard 1971, Stuber and Moll 1982). It is not entirely clear whether these effects are direct or indirect; however, they can not yet be predicted in advance and require validation studies to establish the relationships between allozymes and other traits and factors.

It has been notoriously difficult to establish the role of natural selection in maintaining allozyme polymorphisms. Allendorf and Phelps (1981) accepted neutrality as a working model for allozyme variation reasoning that if selection is an important factor in explaining allozyme variation in natural populations, people would not still be debating the issue. Population studies in forest trees generally have not been directed toward establishing the effects of natural selection, but rather consist of gene frequency surveys in which some loci often show weak correlations with an environmental variable. However, the range of variation in gene freqencies is usually small with respect to the standard error of the estimates, thus resolution should be low. In addition, genetic studies are biased towards defining a population in very local terms, such as a collection of neighboring individuals. If local structure is important, this kind of sample may not be adequate to describe the gene frequencies of the region the population is purported to represent. To assess more directly the question of neutrality vs selection, other studies have 'replicated' nearby pairs of populations contrasting in environments and have detected consistent gene frequency differences more supportive of selection than drift (Mitton et al. 1977).

It is as difficult to avoid concluding that selection may sometimes be important as it is to prove selection is really responsible for observed frequency differences. With respect to application of the knowledge based on patterns of gene frequency differentiation, the question is moot; if a useful

relationship between seed source adaptation and allozyme gene frequencies exists, it can be exploited. Two factors will limit the application of allozyme information in this context. First, relationships between allozymes and growth must be established empirically, but few provenance studies have been undertaken of the scope of the Lake States jack pine trial. Second, the sampling design of allozyme studies would have to revised. The density of sampling must be increased and perhaps a random sampling effort undertaken for a specified group of loci. Considerably greater effort would be required to provide sufficiently detailed information to be of practical importance.

Allozyme studies clearly will never replace provenance trials. However, it is important to consider the shortcomings of the provenance approach. 'Common garden' experiments provide little insight into the basic genetic structure of populations (Brown and Munday 1982). The extent of genetic differentiation between populations is a function of the number of genes involved. At the level of the gene, allelic differences contribute equally to differentiation, but the relationship between gene and phenotype is difficult to assess for quantitative traits. The phenotypic expression of allelic differences in provenence trials, and hence perceived differences, depends upon the test environment. Slight differentiation at the level of the gene could result in radical differences in performance in growth trials.

Allozyme studies provide genetic information at a level close to the genes themselves. This information can be useful for 1) identifying genetically differentiated groups of populations, and 2) developing programs to collect and deploy genotypes in breeding programs. The differentiation of the southern Wisconsin jack pine populations should be taken into account in breeding efforts either as a source of variation or as a potential source of unusual performance.

LITERATURE CITED

Allendorf, F. and S. Phelps. 1981. Use of allelic frequencies to describe population structure. Can. J. Fish. Aquat. Sci. 38: 1507-1514.

Brown, A. H. D. 1978. Enzyme polymorphism in plant population structure. Theor. Pop. Biol. 15: 1-42.

Brown, A. H. D. and R. W. Allard. 1971. Effect of reciprocal recurrent selection for yield on isozyme polymorphisms in maize (Zea mays). Crop Sci. 11: 888-893.

Brown, A. H. D. and B. Munday. 1982. Population-genetic structure and optimal sampling of land races of barley from Iran. Genetica 58: 85-96.

Gottlieb, L. D. 1981. Electrophoretic evidence and plant populations. Prog. Phytochem. 7: 1-46.

Jeffers, R. and R. Jenson. 1980. Twenty-year results of the Lake States jack pine seed source study. USDA-Forest Service, North Central Forest Expt. Stn. Res. Paper NC-181, 20 p.

Mitton, J. B., Y. Linhart, J. Hamrick, and J. Beckman. 1977. Observations on the genetic structure and mating system of ponderosa pine in the Colorado front range. Theor. Appl. Genet. 51: 5-13.

Nevo, E., D. Zohary, A. H. D. Brown, and M. Haber. 1979. Genetic diversity and environmental associations of wild barley, Hordeum spontaneum, in Israel. Evolution 33: 815-833.

O'Malley, D. M., N. C. Wheeler, and R. P. Guries. 1980. A manual for starch gel electrophoresis. UW-Madison, Dept. of Forestry Staff Paper #11, 16 p.

Rudolf, P. 1956. A basis for forest tree seed collection zones in the Lake States. Minn. Acad. Sci. 24: 21-28.

Seal, H. 1964. Multivariate Statistical Analysis for Biologists. Wiley, New York. 207 p.

Stuber, C. W., M. M. Goodman, and R. H. Moll. 1982. Improvement of yield and ear number resulting from selection at allozyme loci in a maize population. Crop. Sci. 22: 737-740.