

United States
Department of
Agriculture

Forest Service

**Pacific Southwest
Forest and Range
Experiment Station**

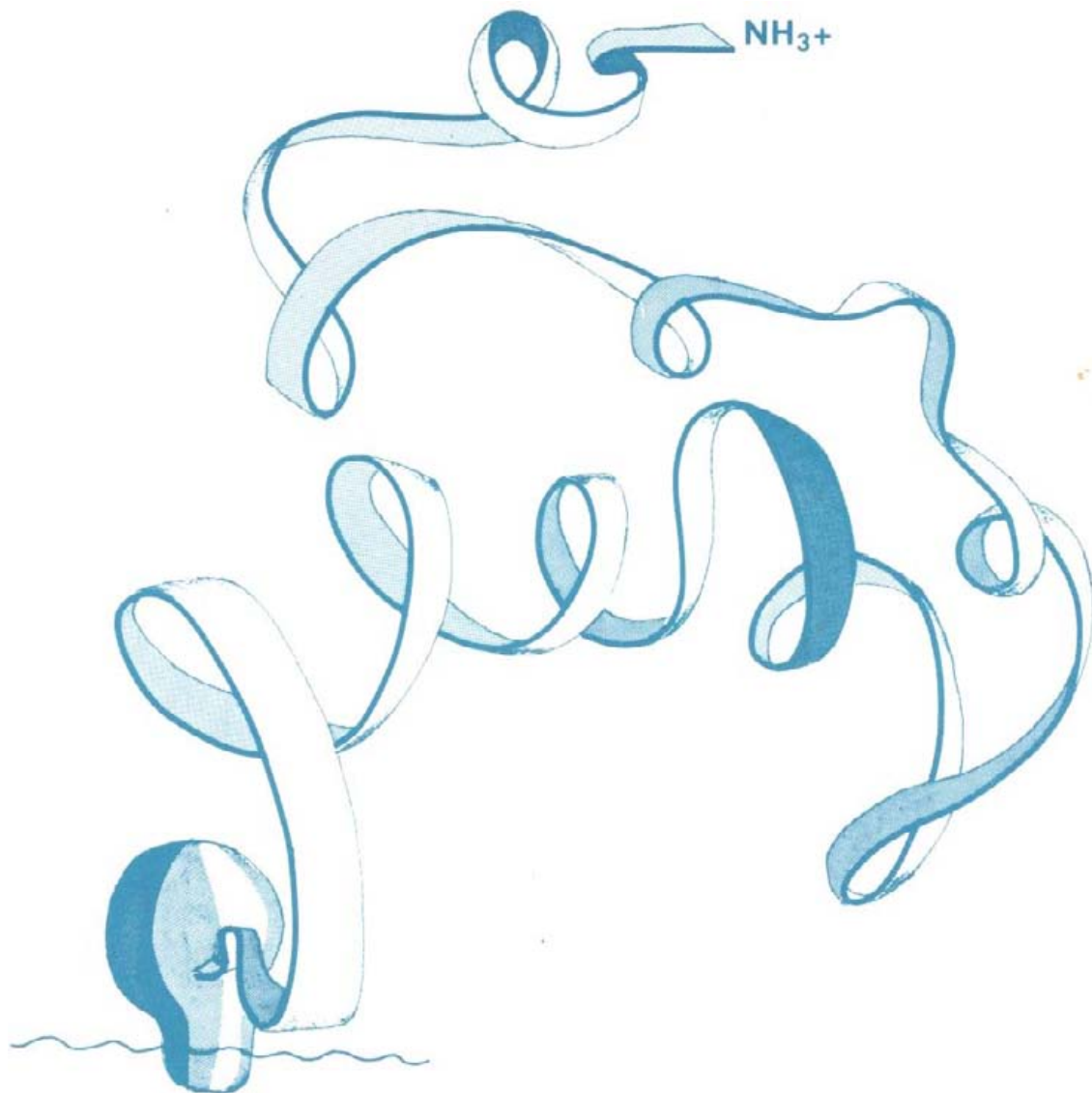
P.O. Box 245
Berkeley California
94701

General Technical
Report PSW-48

Proceedings of the Symposium on

Isozymes of North American Forest Trees and Forest Insects

July 27, 1979, Berkeley, California



Technical Coordinator:

M. THOMPSON CONKLE, a geneticist, works in the Genetics of Western Forest Trees Research Unit, with headquarters in Berkeley, California, and field facilities at the Institute of Forest Genetics, in Placerville, California. He earned a bachelor's degree (1959) in forestry at Michigan State University, East Lansing, a master's degree (1962) in forest genetics at North Carolina State University, Raleigh, and a doctorate (1976) in forestry and genetics at the University of California, Berkeley. He joined the Station's research staff in 1965.

Cover: The polypeptide strand forming will become an enzyme. A ribosome translates genetic information as it moves along a messenger-RNA molecule.

Publisher:

**Pacific Southwest Forest and Range Experiment Station
P.O. Box 245, Berkeley, California 94701**

April 1981

Proceedings of the Symposium on

Isozymes of North American Forest Trees and Forest Insects

July 27, 1979, Berkeley, California

M. Thompson Conkle
Technical Coordinator

CONTENTS

Isozymes and the Genetic Resources of Forest Trees	1
<i>A. H. D. Brown and G. F. Moran</i>	
Isozyme Variation and Linkage in Six Conifer Species	11
<i>M. Thompson Conkle</i>	
Analysis of Mating System Parameters and Population Structure in Douglas-fir Using Single-locus and Multilocus Methods	18
<i>D. V. Shaw and R. W. Allard</i>	
Isozyme Studies of Forest Insect Populations	23
<i>Molly W. Stock</i>	
Associations Between Heterozygosity and Growth Rate Variables in Three Western Forest Trees	27
<i>Jeffry B. Milton, Peggy Knowles, Kareen B. Sturgeon, Yan B. Linhart, and Martha Davis</i>	
Levels of Genetic Variation in Trees: Influence of life history characteristics	35
<i>J. L. Hamrick; J. B. Milton, and Y. B. Linhart</i>	
Genetic Structure of Populations and Differentiation in Forest Trees	42
<i>Raymond P. Guries and F. Thomas Ledig</i>	
Analyses of Gene Diversity in Some Species of Conifers	48
<i>Francis C. Yeh</i>	
An Analysis of Genetic Architecture in Populations of Ponderosa Pine	53
<i>Yan B. Linhart, Jeffry B. Mitton, Kareen B. Sturgeon, and Martha L Davis</i>	
Applying Isozyme Analyses in Tree-Breeding Programs	60
<i>W. T. Adams</i>	

PREFACE

Isozyme techniques are currently the best means for evaluating the genetic structure of forest populations. Data from isozyme studies can help researchers evaluate, sample, and conserve the genetic variation of forest tree species. Our perception of genetic variation in forest populations has been significantly enlarged by new findings from studies detecting enzyme alleles of numerous genes of forest trees and insects.

The 1979 meeting of the Western Forest Genetics Association in Berkeley, California, provided an opportunity for a symposium, on July 27, to present some of these findings. It was appropriate to hold the symposium in conjunction with the Western Forest Genetics Association meeting, since most researchers of isozymes of North American forest organisms work in Western North America. The symposium was sponsored by the Pacific Southwest Forest and Range Experiment Station's Institute of Forest Genetics, which maintains a center for isozyme research in population genetics of forest trees.

The gathering of researchers and their presentations attested to the high value and multiple uses of isozyme technology for resolving genetic variation in forest organisms. Participants reported the distribution of genetic variation between and within several species, and proposed strategies for maintaining the variation through future generations. The population structure of species was evaluated at several levels—from broad regional comparisons to groups of trees on mountain slopes.

For many years, growth tests coupled with quantitative analyses of forest trees have estimated the proportion of total phenotypic variation that is attributable to genetic causes. But almost all the observed variation in isozymes stems from genetic differences: species and populations can be compared on the basis of the amount of genetic variation each possesses. The amount of genetic variation, whether considering an entire species or a small group of trees, is measured by the average heterozygosity of numerous loci. Presentations in the proceedings give information that relates average heterozygosities to growth rates and to life history characteristics of trees.

The relationship between a host tree and its pests adds another dimension to the study of forest genetics. Both the

trees and the pests can differ in the genetic structure of their respective populations and in the amount of genetic variation each possesses. One presentation in this symposium points to the importance of the genetic variation of insects in designing effective pest management strategies.

Forest trees generally are difficult to work with in genetic studies because of their large sizes and slow developmental rates. One feature of conifers, however, elevates their value in isozyme studies. The female gametophyte tissue—the endosperm of conifer seeds—and its egg cells develop from a single haploid nucleus; mature seeds contain a haploid gametophyte and a diploid embryo. The genes segregating among the gametophytes of a tree permit direct genetic tests of allelism. One presentation in the proceedings uses segregation from multiple loci to determine linkage relationships. Others analyze the gametophytes and accompanying embryos from individual seeds to determine the seed tree and pollen parent genotypes. The proportion of outcrossing and selfing are evaluated for open-pollinated seed. Knowledge of the seed tree genotype and of pollen parentage are shown to be valuable for a variety of uses in tree breeding programs.

In all, an impressive diversity of information is represented in the papers of this symposium. It is obvious that a technology that yields precise genetic information for numerous genes, with great efficiency of time and space and at relatively low cost, will be a great aid in tree improvement and gene conservation.

We, at the Pacific Southwest Forest and Range Experiment Station in Berkeley, were delighted with the enthusiasm of our friends and colleagues who attended the symposium and prepared papers for these proceedings. To each participant, we extend our warmest appreciation for current, advanced information of the highest quality. Their successful efforts in resolving genetic variation in forest organisms will undoubtedly stimulate new work in forest genetics and greatly assist in the wise management of genetic resources.

We extend appreciation to the Western Forest Genetics Association for publicizing the symposium and providing local arrangements. Special thanks are due Ray Steinhoff, 1979 Chairman of the Western Forest Genetics Association, and William B. Critchfield and William J. Libby, co-chairmen of the 1979 meeting of the Association.

M. Thompson Conkle
Forest Service, U.S. Department of Agriculture
Technical Coordinator

Isozymes and the Genetic Resources of Forest Trees¹

A. H. D. Brown² and G. F. Moran³

Abstract: Genetic data are an essential prerequisite for analysing the genetic structure of tree populations. The isozyme technique is the best currently available method for obtaining such data. Despite several shortcomings, isozyme data directly evaluate the genetic resources of forest trees, and can thus be used to monitor and manipulate these resources. For example, preliminary isozyme data indicate that domestication of cultivated species has generally reduced variation within populations. Geographic differences among natural populations are less evident in trees than in herbaceous plants, but need to be considered in sampling and conservation strategies. In dealing with remnant population size (N_r) and distribution, management programs should recognize that the half-life of heterozygosity is about $1.4 N_r$ generations, whereas the number of trees required to retain half the current alleles after a bottleneck is about the square root of the original population size.

The growing awareness of two major theses provide the occasion of this opening contribution to the symposium. The first is that the genetic resources of economic plant species, and in particular of forest trees species, are finite. Yet, these resources have been managed without concern for their conservation or renewal. It is now clear that our remaining forest resources must be used prudently to guarantee a long future for managed species. The second thesis is that the isozyme technique is well suited for monitoring and assisting the manipulation of our genetic heritage, at least in part (Brown 1978). The last decade has witnessed the growth and impact of this technique in population biology. The ensuing contributions to this symposium will amply demonstrate the ways in which electrophoretic data can contribute to our genetic knowledge of populations of forest trees.

This paper first considers the properties of isozyme data and then discusses how the data can be used to monitor and conserve the genetic resources of forest trees.

ISOZYMES: BONANZA OR BLIND ALLEY IN FOREST GENETICS?

Phenotypic as Opposed to Genetic Evaluation

Two radically different approaches can be taken when measuring differences among individuals, populations or species. The first we define as the "phenotypic" approach. It is motivated by the argument that it deals directly with characters of economic or biological importance. The ideal of this approach is to know the phenotypic expression of

yield, quality, and pest and disease resistance in all possible environments. Clearly, this ideal is unattainable, because the need for decisions limits, in space and time, the number of environments that may be tested. Quantitative genetics attempts to overcome these limitations by using a predictive model. The model assumes that unspecified polygenes can simulate the genetic and developmental complexities that underly phenotypic measurements and their variation. Various classes of effects on phenotypic variation—additive, dominance, epistasis, genotype $\times$ environment interaction—are defined (Libby and others 1969). The predictive power, however, is severely limited to the genotypes tested, the environments experienced, and the mode of measurement. Of course phenotypic (agronomic or silvicultural) assessment and phenotypic selection are essential in any plant breeding program. It is less certain whether such data provide valid measures of genetic variation and genetic similarity among populations.

The second approach to evaluation, the "genetic" approach, is entirely different in emphasis. Its aim is to detect genetic differences as close as possible to the DNA (deoxyribonucleic acid) level. The ideal of this approach is to know all the DNA sequences. Such basic data would form direct measures of how genetically variable is one population, and how similar it is to another, entirely free of environmental effects. Data from different species would be directly comparable. Again, this ideal is far from being a reality. Perhaps it may be attainable for very restricted classes of DNA (for example, organellar DNA, highly repeated DNA) or in very simple organisms (Smith 1979). But the need for "genetic," as distinct from "phenotypic," evaluation is apparent.

Because we cannot presently measure variation directly at the DNA level, we must lay down criteria for the choice of techniques that are currently feasible. *Table 1* summarizes four criteria first suggested by Hubby and Lewontin (1966; see also Lewontin 1974). These criteria require that the effect of an allelic substitution at a locus be detectable (C), and distinguishable from the substitution of any other allele at the same locus (A), or other loci (B).

¹ Presented at the Symposium on Isozymes of North American Forest Trees and Forest Insects, July 27, 1979, Berkeley, Calif.

²Principal Research Scientist, Commonwealth Scientific and Industrial Research Organization (CSIRO), Division of Plant Industry, Canberra, Australia.

³Experimental Officer, CSIRO, Division of Forest Research, Canberra, Australia.

Table 1—*The Hubby-Lewontin criteria for genetic markers, and the properties and pitfalls of the isozyme technique with respect to each criterion*

Criteria to be met	Properties of isozymes	Problems occurring
A. Alleles are distinct in individuals	Codominance; free of epistatic or environmental effects	Genetic or environmental mobility; mobility modification
B. Effects are locus-specific	Enzyme specificity	Nonspecific assays; duplication or polyploid
C. All substitutions are detectable	Charge-state separation, irrespective of function or variation	Detectable class of base substitutions is restricted and biased in variation
D. Loci are sampled randomly	Assay availability is independent of variation	Loci assayed are a limited class and differ in variation

Further, the sample of loci studied should be random, irrespective of function, or likely level of polymorphism (D).

The Isozyme Method

As perceived by Hubby and Lewontin (1966), the isozyme technique meets these criteria more closely than any preexisting method. Its general properties relevant to the criteria are also listed (*table 1*). In short, one can study discrete Mendelian genes. This applies even in long-lived organisms such as trees, because individual seeds or seedlings can be assayed as members of progeny arrays.

An additional and important advantage of the method is the facility with which an array of enzymatic loci can be assayed using material in minute quantities, and with minimal preparation. The isozyme technique, therefore, represents the best currently available method for measuring genetic variation, close to the DNA level, relatively free of environmental effects, and on a manageable number of samples.

Pitfalls and Shortcomings of Isozyme Data

Some of the problems of isozymes associated with each of the Hubby-Lewontin criteria are listed in *table 1*. First mention is of some known cases where simple Mendelian codominant inheritance does not apply. Post-translational modification, either genetic (Law 1967) or environmental (Cullis 1977) in origin, is a possibility of which isozyme geneticists are becoming increasingly aware. The best way to avoid this pitfall is to verify by a Mendelian analysis that a set of isozyme variants are *allozymes*, that is, they are coded by alleles at a locus. This is especially the case for enzymes which do not form heteromultimers, or which are assayed nonspecifically (for example, esterases), or which

are known to be affected by environment (for example, peroxidases).

Enzyme assays vary in their specificity. Diploid barley has at least 10 genetically defined esterase loci (Hvid and Nielsen 1977). With such nonspecific enzyme assays, a Mendelian analysis is essential. Duplication of genetic material, as in polyploidy, can often lead to the inability to ascribe a particular variant to a single locus. Special chromosomal stocks, such as nullisomics (Hart and Langston 1977) or haploids, if available, provide a definitive solution to this problem. This problem is rare in forest species, however, as diploidy is the common condition (Libby and others 1969).

With respect to the third criterion, two problems arise. The first is that electrophoresis does not detect all single base substitutions. About 23 percent of substitutions are synonymous and do not alter the amino acid because of redundancy in the code. About 67 percent of amino acid replacements do not alter net charge (Marshall and Brown 1975b). Some fraction of these replacements may be detectable by more refined electrophoretic techniques, or by altered heat sensitivity of the enzyme. Also, apparently long sequences of DNA inserted within the coding sequence of many eukaryotic genes may or may not be removed in the production of mRNA (messenger ribonucleic acid) (Crick 1979). These sequences are not translated, and are thus beyond the reach of electrophoretic surveys. Their function is not known, but preliminary evidence suggests that they can be quite variable. Crick (1979) describes the discovery of inserted sequences and split genes as "a mini-revolution in molecular genetics".

The other problem with respect to the third criterion is that, given access to a restricted class of base substitutions at a locus, can we assume that this class is unbiased with respect to the level of overall variability? Unfortunately, mounting evidence indicates that we cannot. For example, careful hybridization studies of variation in the DNA of sea urchins reveal a remarkably high level—4 percent—of base pair mismatching among individuals (Britten and others 1978). This mismatching is computed to be an order of magnitude greater than that predicted from electrophoretic heterozygosity, even after the latter has been corrected for synonymy and electrophoretically silent amino acid replacements. Isozymes are not so "close to the gene level" after all.

The fourth criterion raises the same two questions, but with respect to the loci sampled. Electrophoresis is used to study variation in the structural region of nuclear genes for certain restricted classes of proteins. Several studies of variation at these loci, either within functional classes of enzymes (Johnson 1974), or in relation to their physical parameters (Koehn and Eanes 1978), indicate that such loci differ in their tendency to vary genetically. This point must be taken into account in comparative studies of data on different species from different laboratories that may not have covered comparable sets of loci. Leigh Brown and Langley (1979) have suggested that the enzymes typically studied may overestimate genetic variation. They found

markedly less variation (4 percent heterozygosity in about 54 of the abundant proteins of adult *Drosophila*) than that predicted from allozyme studies (14 percent). Also, the inability of comparative electrophoretic divergence to measure evolutionary morphological divergence has been stressed by Wilson and others (1977). The lack of information on eukaryotic regulation systems and developmental sequences, and of data on sequence variation in regulatory loci or variation in their chromosomal location are severe shortcomings. Fortunately, the future application of other techniques of molecular biology to population studies does offer the hope of complementing the isozyme picture. These techniques include restriction nuclease analysis of organellar DNA's (Scowcroft⁴), DNA hybridization (Hall and others 1976), and ultimately cloning of specific sequences and DNA sequencing, and characterization of highly repeated classes of DNA for their base sequences and chromosomal location.

A major shortcoming—perhaps the greatest of all—is that the adaptive or selective significance of the bulk of allozyme variation remains obscure (Lewontin 1974). Is this variation adaptive or is it "evolutionary noise"? A less esoteric way of framing this question is to ask whether the genes studied by isozyme techniques have any relation at all to the characters of interest to the plant breeder. Is the study of allozyme variation a blind alley for forest tree breeding? Preliminary studies have shown it is possible to examine this question with a combination of population, physiological, and biochemical approaches. For example, alcohol dehydrogenase polymorphism in certain plant species shows population differentiation, selection responses, adaptive differences in tolerance to waterlogging, and in germination rates, and parallel variation in enzyme activity (Brown 1979). Much more evidence must be assembled, however, before a general answer is possible.

FOREST TREES AND THEIR GENETIC RESOURCES

The genetic resources of forest tree species consist of both primary and secondary gene pools (Frankel 1977). The primary reservoirs are those of undisturbed natural forests; the secondary sources include the remaining trees after ecosystem disturbance or harvesting, and the productive populations planted from local or exotic seed. Many authors have reviewed the genetic research on forest trees from the standpoint either of utilization or of genetic conservation (Libby and others 1969, Richardson 1970, Stern and Roche 1974). We now outline some issues in the status, handling, and conservation of genetic resources, which arise from the distinctive features of forest trees, in contrast to crop plants.

Size and Individuality

As large organisms, forest trees can readily be the target of individual phenotypic selection for growth of timber quality characteristics. This facility raises the question of the role of individual selection when collecting reproductive material. Should seed be taken only from a limited number of outstanding or "plus" trees? To answer this question, we first consider natural stands and ask whether the sampling is for conservation or breeding. Is the goal to obtain the largest amount of potentially useful genetic variation from that population; or is it to select a limited number of trees as parents in a seed orchard or breeding program? If the goal is *conservation*, Marshall and Brown⁵ (1975a) have argued against sampling biased by phenotype. The sampling pattern at a site should be random with respect to phenotype, but representative with respect to ecological variation within the site. *Where* a plant grows is, perhaps, as reliable a guide as *what* it looks like when we consider its likely contribution to the genetic resources of the species. If the sampling aim is restricted to *breeding*, then the collection of seed from "plus" trees is debatable. It is difficult, however, to avoid unwanted biases from, for example, ephemeral selection criteria, diverse age structure of the population, and physical or biotic microenvironmental effects. If the "plus" trees are to be propagated clonally, one may have more justification for selection. In plantations, with a uniform age structure, and more random distribution of genotypes in microenvironments, it is usually easier to designate "plus" trees. Still, they could amount to a random sample. Most important, however, is the *number* of trees sampled. This number should not be restricted solely on the grounds that few of the available trees are outstanding. Genetic data that show the effect of biased as opposed to random sampling in tree populations are obviously needed.

A second practical consequence of tree size contrasts with the situation in most crop and forage species. The culture of a single plant uses considerable resources, and severely limits the population numbers in seed orchards, and *ex situ* conserved stands (Bouvarel 1970).

Longevity

As long-lived plants, the phenotypes of mature forest trees sum up a long history of environmental variation in space and time. This fact raises the issue of genic variation and fitness. Is genic variation maintained by heterozygote advantage in most environments? This may or may not mean that heterozygotes show less phenotypic plasticity for certain morphological traits (Bradshaw 1965). Or, are the individual alleles at polymorphic loci adapted to cope with different environments? Testing for isozyme-envi-

⁴Scowcroft, W. R. Nucleotide polymorphism in chloroplast DNA of *Nicotiana debneyi*. Theor. Appl. Genet. (In press.)

⁵Marshall, D. R. and A. H. D. Brown. Theory of forage plant collection. In Genetic Resources in Forage Plants. R.A. Bray and J.G. McIver, eds. CSIRO, Melbourne (In press).

ment associations must take account of this complex of possibilities and of conflicting predictions from different modes of selection.

Longevity and the delay in maturity in forest species, compared with annual crops, raise a practical issue in conservation by seed storage (Wang 1975). The viability of stored seed can approach the duration of the life cycle. If so, it must be regenerated as often as it takes to grow one crop of seed. Stored seed has an important role in the distribution of genetic resources without repeated collecting. What, then, is its role in genetic conservation of trees? Stern and Roche (1974) view *in situ* conservation as an interim measure until conservation is possible by "other means." In contrast, it can be argued that seed storage is essentially a provisional measure until *ex situ* stands are established, or a temporary insurance against the accidental destruction of *in situ* stands or plantations. The role of seed storage is severely limited in the case of tropical species with "recalcitrant" seeds.

Fecundity

The number of seeds produced by a single tree, even in one crop, may be enormous. This high fecundity enables tree species to tolerate high genetic loads and therefore, high levels of heterozygosity. From the standpoint of seed collection, this is one reason why the number of seeds sampled per mature individual can and should be much greater in trees than in crop or pasture species. Another reason is that the subsequent multiplication in plantations of any seed sample from trees is a slow and costly process. More specific recommendations on sample size are given below.

Outcrossing

Almost all the forest trees in which the breeding system has been studied are outbreeders. From this we can infer that heterozygosity is beneficial. A corollary is that inbreeding depression is likely to be evident in progeny if selfing rates are high. Further, the evidence suggests that outbreeders have less scope for the evolution of coadapted complexes of genes (at least in the sense of genes in linkage disequilibrium) than inbreeders (Brown 1979). Because of maternal heterozygosity and outcrossing, the seed crop on an individual tree is genetically heterogeneous (in contrast to inbred crops).

Harvesting

The link between economic utilization and the threat to genetic resources in forest species is more direct and final than in most other economic species. Logging destroys the individual as a genetic resource, and renewal requires reforestation based on sound genecological principles.

Stage of Domestication

Forest trees are wild plants at the earliest stages of domestication (Libby 1973). The time scale of the domestication process will be long. Unless the use of vegetative propagation becomes widespread, breeding and agronomic advances in themselves will not be the major causes of genetic erosion, as they are in agricultural crop species. The chief threat to agricultural genetic resources is the replacement of the older, more variable strains with modern advanced cultivars. The principal task in their conservation is the collection and preservation in gene resource centers, of the dwindling stocks of land races (Frankel 1974). The major threats to the integrity or, indeed, the survival of the genetic resources of forest trees are (a) ecosystem destruction and changing patterns of land use, (b) widespread planting with a few fast-growing species, (c) selective harvest of trees, and (d) plantation or reforestation practices based on limited or maladapted germplasm. Several case histories of tropical pines, teak, eucalypts, and others were reviewed at the Third World Consultation on Forest Tree Breeding, Canberra, 1977. The desirability and need for *in situ* conservation was frequently stressed.

With these general features of forest trees as a background, we now return to considering the importance of population genetic research based on the isozyme technique.

ISOZYME STUDIES AND GENE RESOURCE MANAGEMENT IN FOREST TREES

What are the roles that genetic evaluation based upon the isozyme technique might play in the management of the genetic resources of forest species (Ferret and Bergmann 1976)? We would like to stress two interrelated functions—monitoring and decisionmaking.

Isozymes are ideal for monitoring:

- (a) the current level of genetic variability in populations of economic species. This would furnish a base-line for evaluating the genetic consequences of any practice, as well as furnishing data for comparing levels of variability in different species (Hamrick and others 1979);
- (b) increased genetic vulnerability in plantations as indicated by a decline in variability because of selection;
- (c) loss of variation resulting from ecosystem destruction;
- (d) the extent of divergence and variation in heterozygosity between natural populations of a species;
- (e) the differentiation within a species between natural and planted populations, especially in other countries where exotic species (for example *Pinus radiata* D. Don, *Eucalyptus grandis* W. Hill ex Maiden) may have undergone adaptive "naturalization" (Frankel 1977);
- (f) the extent of effective cross-pollination in seed orchards;
- (g) pollen contamination of seed orchards;

- (h) pollen contamination of *in situ* reserves by plantations;
- (i) the inbreeding load in severely bottlenecked populations, such as those that regenerate from decimated stands following clear felling;
- (j) seed and clone identification.

Several of these topics will be discussed in this symposium. As an example, we quote some of our own and other results to illustrate the impact of domestication on levels of genetic diversity. Table 2 gives estimates of genetic diversity (H) (Nei 1975), and alleles per locus (A) in four cases. The statistic H is the mean expected panmictic heterozygosity ($H = 1 - \sum p_i^2$ Where p_i is the frequency of the i^{th} allele). In barley (*Hordeum*), the sequence wild progenitor → primitive varieties or land races → advanced cultivars shows a fall in genetic variation both within lines and in the total sample. In contrast, domestication in *Phlox* is accompanied by increased total diversity. This increase probably arose from the deliberate collection and continual use of interesting variants. Darwin described an analogous consequence of artificial selection in the domestication of pigeons. In forest trees, the preliminary data indicate important changes in diversity levels with management, especially in *P. radiata*.

Recently, Hamrick and others (1979) have concluded from an analysis of published data that trees are genetically more variable on the average than herbaceous plants, because they are widespread, long-lived, large, fecund, and outcrossing species. This conclusion differs from that of animal data (Powell 1975, Nevo 1978): many vertebrates have less than one-half the heterozygosity of invertebrates. Is longevity or size of an organism (the attribute that Frankel and Soulé⁶ use as an index of conservation strategies) a conflicting indicator of genetic variability in plants as

compared with animals? Perhaps the plant literature is limited and biased by studies of inbreeding, colonizing, exotic annuals on the one hand, and the genecological studies of forest trees on the other. Several of these latter studies, including ours on mating systems, deliberately focused *only* on variable loci.

Table 3 summarizes some estimates of population genic diversity per variable locus ($\bar{H}$) and population differentiation ($D_{ST}/\bar{H}$, where $D_{ST} = H_T - \bar{H}$ [Nei 1975]) in tree species. These estimates are compared with the same statistics in inbreeding and outbreeding herbaceous plants. As a group, the wind-pollinated conifers show the least degree of population differentiation. Trees pollinated by insects or birds show about three times as much, and are comparable with herbaceous outbreeders. *Eucalyptus caesia* Benth., a bird-pollinated, and geographically restricted species shows the highest differentiation of these trees, approaching the level for inbreeding herbs. Clearly, a wide range of population structures is found in trees.

DECISIONS

We will now outline some issues in which management decisions could be made with more precision if extensive isozyme data were available.

Sampling Strategies

How large should samples be and how should they be disposed among trees and among populations? Answers to this question depend on the purpose of sampling. This may be (a) to assemble the genetic resources in a target area for direct use, breeding or conservation; (b) to measure population genetic structure at individual loci, or obtain a sample that faithfully reflects gene frequencies; (c) to

⁶Frankel, O. H. and M. Soulé. Conservation and evolution. (Manuscript in preparation.)

Table 2—Comparative genetic diversity (H) and alleles per locus (A) of domesticated lines and related natural populations

Species	Lines	Allozyme loci ¹	Within lines		Total		Source ²
			A	H	A	H _T	
<i>Hordeum vulgare</i>	32 Scandinavian varieties	11 ¹	1.02	0.007	1.55	0.127	a
	Composite cross XXI, F ₁₇	17 ¹	—	—	1.53	.094	b
	11 Iran land races	17	1.21	.072	1.94	.180	c
<i>H. spontaneum</i>	28 Israel populations	17 ¹	1.56	.109	4.12	.216	b
<i>Phlox drummondii</i>	16 Cultivars	17	1.12	.041	1.41	.133	d
	10 Populations	17	1.25	.063	1.29	.089	d
<i>Pinus sylvestris</i>	2 Seed orchards	3	3.5	.36	3.7	.37	e
	3 Sweden populations	3 ¹	4.3	.33	5.3	.38	d
<i>Pinus radiata</i>	1 Seed orchard (30 clones)	18 ¹	—	—	1.67	.118	f
	20 Breeding clones	19	—	—	1.74	.145	f
	5 California populations	19	1.84	.139	2.63	.161	f

¹Loci: More loci are in the original source, but only comparable loci are included here

²Sources: (a) Almgard and Landegren (1974), (b) 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. (In press.) (c) Brown (unpublished). (d) See Brown (1979) for references, (e) Rudin and Lindgren (1977), (t) Moran (unpublished).

Table 3—Population genic diversity per variable locus ($\bar{H}$) and differentiation ($D_{ST}/\bar{H}$) in trees

Species	Variable loci	Populations	$\bar{H}$	DST/ $\bar{H}$ Percent	Source ¹
<i>Abies lasiocarpa</i>	1	3	0.40	3.0	(a)
<i>Picea abies</i>	4	8	.42	4.5	(a)
<i>P. abies</i>	6	10	.40	4.7	(a)
<i>P. abies</i>	4	8	.35	2.4	(a)
<i>P. abies</i>	10	4	.41	2.9	(b)
<i>P. engelmannii</i>	1	3	.47	12	(a)
<i>Pinus longaeva</i>	11	5	.47	3.6	(c)
<i>P. nigra clusiana</i>	3	7	.48	13	(d)
<i>laricio</i>	4	12	.30	6.9	(d)
<i>nigricans</i>	3	11	.52	12	(d)
<i>pallasiana</i>	4	10	.30	7.0	(d)
<i>P. radiata</i>	19	5	.14	15	(e)
<i>P. rigida</i>	15	4	.17	1.1	(b)
<i>P. sylvestris</i>	3	3	.33	14	(a)
<i>P. virginiana</i>	2	4	.29	2.6	(g)
<i>Pseudotsuga menziesii</i>	2	6	.55	3.0	(h)
<i>P. menziesii</i>	4	9	.59	2.9	(i)
<i>Eucalyptus caesia</i>	8	10	0.11	78	(j)
<i>E. cloeziana</i>	5	17	.25	17	(k)
<i>E. delagatensis</i>	4	8	.34	24	(a)
<i>E. obliqua</i>	3	7	.39	21	(a)
<i>E. pauciflora</i>	7	3	.28	2.2	(a)
<i>Ficus carica</i>	2	4	.53	2.3	(a)
Summary		Average		Weighted ²	Average
		$\bar{H}$	DST/ $\bar{H}$ ³	$\bar{H}$	DST/ $\bar{H}$
Outbreeders	Wind-pollinated	0.372	7.4	0.298	6.8
	Animal-pollinated	.317	16	.265	20
	Herbaceous	(b) .325	22	.282	18
Inbreeders		(b) .126	118	.133	102

¹Source: (a) see Brown (1979) for references, (b) Lundkvist (1979), (c) Hiebert, see Hamrick and others (1979), (d) Bonnet-Masimbert and Bikay-Bikay (1978), (e) Moran (unpublished) (f) Guries and Ledig (1977), (g) Witter and Feret (1978), (h) Mejnartowicz (1976), (i) Yang and others (1977), (j) Moran and Hopper (unpublished), (k) Turnbull (unpublished).

²Weighted by the number of variable loci in each study.

³Computed by dividing the sum of D_{ST} estimates by the sum of $\bar{H}$ estimates.

measure multilocus associations; or (d) to estimate parameters of the mating system.

Sampling for the first purpose has been discussed extensively by Marshall and Brown⁵ (1975a). The key to this sampling is to maximize the number of *locally common* alleles in a collection whose total size is strictly limited. Locally common alleles are defined here as alleles that assume a frequency greater than 0.10 in only one locality. Alleles with two such occurrences are termed sporadically common. The existence of locally common alleles in plant populations is discussed elsewhere (Brown 1978). The limited data from forest species (*table 4*) indicate that of all alleles detected, locally common alleles make up a lesser (but significant) fraction, compared with herbaceous plants. Our rule of thumb has been to sample about 50 plants at random and visit as many contrasting sites as possible. If this is difficult or time consuming, then samples of fewer plants from more sites should be taken (Marshall and Brown⁵).

In forest trees, this rule applies directly to the sampling of vegetative material. For seed samples, however, it requires variation because of the high fecundity, the outbreeding system, and the difficulty of sampling each additional tree. Samples from fewer trees (n), but more seed per tree (k), such that $nk > 100$ are in order. In practice, however, much larger samples than this are taken in forestry to allow for seed requests, and to avoid early resampling or regeneration.

Samples for research, or for the representation of site allele frequencies, must be based on a reasonable number of trees. A minimum value for n follows from the limiting formula for the variance of allele frequencies in the seed generation (Brown and others 1975) with large k and assuming $p = q = 0.5$

$$\text{Lim } \sigma_p^2 = (1 + 3F)^2 / 32n(1 + F)$$

$$k \rightarrow \infty$$

Table 4—The number of alleles with various kinds of distribution

Species	Populations	Loci	Kinds of enzymes	Common			Rare alleles	
				Wide-spread	Sporadic	Localized	Wide-spread	Localized
<i>Picea abies</i>	11	4	4	10	1	1	1	1
<i>P. abies</i>	10	6	4	12	—	2	3	1
<i>P. abies</i>	4	11	6	19	3	3	16	13
<i>Pinus longaeva</i>	5	14	7	27	3	—	3	—
<i>P. nigra clusiana</i>	7	4	1	9	3	1	1	—
<i>laricio</i>	12	4	1	8	1	—	1	1
<i>nigricans</i>	11	4	1	10	—	—	2	1
<i>pallasiana</i>	10	4	1	7	—	1	1	3
<i>P. radiata</i>	5	19	13	26	1	3	10	10
<i>P. rigida</i>	4	15	10	19	2	1	8	—
<i>Pseudotsuga menziesii</i>	6	2	1	7	—	—	4	1
<i>P. menziesii</i>	9	4	3	12	—	—	—	—
<i>Eucalyptus caesia</i>	10	8	6	12	1	6	—	2
<i>E. cloeziana</i>	17 ¹	5	5	9	3	1	1	—
<i>E. delagatensis</i>	8	4	4	8	1	—	1	—
<i>E. obliqua</i>	7	3	3	6	—	2	4	2
Summary				Percentage of variants				
Wind-pollinated trees		91	—	41	8	7	27	17
Animal-pollinated trees		20	—	39	13	23	75	10
Wild herbaceous (8 entries ²)		14	—	31	18	28	7	16
Cultivars (4 entries ²)		25	—	45	14	23	4	14

¹ These populations were grouped into seven geographic regions.² From Brown (1978). See table 3 for sources of remaining data.

where F is the inbreeding coefficient.

The minimum number of trees required to ensure that the difference (d), between the sample allele frequency and its true population frequency, is exceeded on less than 5 percent of occasions is

$$n = (1 + 3F)^2 / 8(1 + F)d^2$$

$$\simeq 0.2 d^{-2}, \text{ assuming } F < 0.10$$

Ford = 0.1, we would require large samples from at least n = 20 trees. A sample of 10 trees from each of two populations can detect differences in gene frequency between them of 0.2 at the 5 percent level of significance.

Much larger samples are needed to detect weak multi-locus associations in the adult generation (Brown 1975). Finally, sampling for the study of mating systems unfortunately is ideally served by one contrary principle: the sampling of trees from within the same microenvironment and in close proximity (Brown and others 1975). Otherwise, microgeographic differentiation can lead to inflated estimates of self-fertilization. In general, more information on allelic frequency patterns could be used to assess optimal sampling strategies within and among sites.

Harvesting in Naturally Regenerating Systems

What and where to log and what to leave? Selective harvesting followed by natural or assisted reforestation is a

common practice. There are important genetical issues for which we have little genetic data to inform us. If size or growth rate is related to heterozygosity, will selective logging amount to selection against heterozygotes and, therefore, lead to destabilized genetic equilibria (Charlesworth 1972) and the loss of certain polymorphic alleles? Dysgenic selection could also arise in an additive model, simply by removing the preferred alleles. Perhaps it would be a better strategy, genetically, to clear fell most of the area and leave certain areas untouched as a seed source. A major variable here would be the pattern and scale of microdifferentiation. For example, the study of Mitton and others (1977) of *Pinus ponderosa* Dougl. ex Laws., indicates that reserved areas should be located on both north- and south-facing slopes, and that of Bergmann (1978) of *Picea abies* K indicates that elevational gradients are important.

Critical Population Sizes for Conservation

How large a population should be set aside and managed as a long-term source of genetic diversity (conservation *in situ*)?

As the need for conservation becomes more pressing, ecologists and geneticists are being asked to recommend the minimum size of a viable population. As yet there is no single answer to this question. Frankel and Soulé⁶ have suggested 50 for the basic irreducible number for

conservation genetics. They regard this number as the absolute minimum below which inbreeding ($\Delta F = 1 - (1/2N)$) would be unacceptably greater than 1 percent per generation. Of course, we refer to *effective* population number—actual numbers being somewhat larger. This reasoning, however, does not apply to inbreeding or to apomictic plants. Furthermore, populations of this size have little scope for adaptive evolution, a prerequisite for long-term survival. Franklin (see Frankel and Soulé⁶) has proposed the population size of 500, so that the effects of drift would be negligible when compared with those of selection of measurable intensity.

Latter and Frankel (in Frankel 1970) have suggested using the equilibrium formula of Crow and Kimura (1970) which is

$$H_e = 4Nu / (1 + 4Nu) = \Theta / (1 + \Theta)$$

in which

H_e = equilibrium heterozygosity

N = effective population number

u = mutation rate

$\Theta = 4Nu$

If we assume $H_e = 0.3$ and $u = 2 \times 10^{-6}$ (Mukai and Cockerham 1977), both figures for electrophoretically detectable mutation, we obtain the number, $N \simeq 50,000$. This is the effective number required to prevent neutral heterozygosity declining at all from its present level. Thus, we have a range of answers (from 50 to 50,000) based on a range of rationales. The answer will vary for different species, so that no rationale is universally applicable.

Another way to look at this issue is to introduce the time element into the calculations, and ask what is the *half-life of heterozygosity* in generations, given a particular effective size? The loss of one-half the heterozygosity is biologically equivalent to one round of complete selfing.

Let t_j denote the time in generations required for the current heterozygosity (H_0) to fall to a value $j H_0$ ($0 < j < 1$) in a restricted population of effective size N_e and neutral mutation rate u . If $H_e = 4N_e u / (1 + 4N_e u)$, the new equilibrium heterozygosity, and assuming [$j H_0 > H_e$], the number of generations is

$$t_j = \frac{\log_j \{ [j H_0 - H_e] / [H_0 - H_e] \}}{\log_j \{ 1 - 2u / (1 + 4N_e u) \}}$$

In particular, the half-life of heterozygosity is approximately

$$t_{0.5} \simeq 1.4 N_e$$

If we assume that the "time scale of concern" (Frankel 1974) for forest species is of the order of 100 generations or 1000 years (Frankel and Soulé⁶), and that we wish to

preserve at least one-half the heterozygosity by that date, the effective number required is $N_e = 70$.

The loss of heterozygosity is but one dimension to the problem of population size. A second and potentially more important problem is the numbers of alleles lost in a bottleneck. We may again ask an analogous question.

If the original population size is N_0 , what size of relic population, N_r , is required to retain half the alleles currently present? We can use the sampling theory for neutral alleles (Ewens 1972) to attempt an answer to this question. We will assume two values for $\Theta = 4N_0 u$, namely $\Theta = 0.5$ or 1.0 . In this section, u denotes the *total* mutation rate.

In the population, the actual number of alleles is

$$n_a = \theta \int_{1/2N_0}^1 (1-x)^{\theta-1} x^{-1} dx$$

For $\Theta = 1$,

$$n_a | \Theta=1 = \log_e(2N_0)$$

and for $\Theta = 0.5$,

$$n_a | \Theta=0.5 \simeq [\log_e(8N_0)] / 2$$

From Ewens (1972), the expected number of alleles retained in the sample (the relic population) is

$$n_r = \sum_{i=1}^{2N_r} \frac{\theta}{\theta + i - 1}$$

Using an integral approximation, this summation can be evaluated as

$$n_r \simeq \Theta \{ \log_e [(\Theta + 2N_r - 1) / \Theta] \} + 0.5$$

We wish to find the values of N_r which will yield $n_r/n_a = 0.5$, to retain half the alleles.

For $\Theta = 1$; $N_r = 0.43(N_0)^{1/2}$, and for $\Theta = 0.5$;

$$N_r = 0.26(N_0)^{1/2}$$

Unfortunately, rather large standard errors are associated with the number of alleles per locus retained in a relic population. Applying the same approximation procedure to Ewens' (1972) variance formula, we obtain for the variance,

$$\text{var}(n_r) \simeq \Theta \{ \log_e [(\Theta + 2N_r - 1) / \Theta] - 1 \}$$

However, if one considers that a large number of independent loci ($v > 100$) are polymorphic in the population, it is possible to ask what is the number required to ensure survival of one-half of all the alleles with 95 percent certainty. For $\Theta = 1$, the effective population size (N_e) required for this is

$$N_e \simeq 0.43N_0^{1/2} \times \exp [2 \{ \log_e(0.1N_0) \} / v]^{1/2}$$

Original population size (N_0)	Number of polymorphic loci (v)		
	100	1000	∞
10,000	62	48	43
100,000	209	158	136

These figures indicate that, as a rough guide, the number to retain to ensure one-half the alleles survive with 95 percent certainty is about the square root of the original population size. More precise calculations would be possible when more estimates of θ , u and v are available.

Population Distribution

Should only one large population be set aside, or is it sounder genetically to divide the requirements into several partially isolated subpopulations?

If s is the number of subpopulations, each of effective size N_s , and m is the migration rate between them in the island model (Nei 1975), the total heterozygosity maintained in the whole ensemble (H_T) is

$$H_T = 1 - [1 + u \left\{ s/m + 4sN_s \right\}]^{-1}$$

From this formula, it can be seen that H_T will increase under subdivision as $s/m + 4sN_s$ increases. The island model of migration is the most conservative from this viewpoint. The quantity sN_s nominally equals N_r , the total number retained from the original population. This quantity, however, is not likely to remain constant with increasing s ; rather, it will decline because of the need for buffer zones around each subpopulation. More and more subpopulations of smaller size may render each one steadily more vulnerable, so that the total (N_r) declines as s increases. The requirement that the quantity (s/m) increase with increasing s means that, as the number of subpopulations and hence their proximity increases, a less than commensurate proportional increase in migration will occur between them. The optimum value of s to maximize H_T could be defined if we had information on buffer-zone requirements and on gene flow with increasing s .

CONCLUSIONS

The isozyme technique is a quantum jump forward in research on the genetics of forest trees. It does more than just measure another set of characteristics, because it enables a direct genetic evaluation of the status of genetic resources. Despite the real limitations of isozyme data, which may in the near future be partly met by the application of newer molecular techniques in population studies, it is possible to make decisions on the manipulation and conservation of these resources, aided

by such data. An isozyme laboratory may be considered demanding of personnel and financial resources. Yet these costs should be compared with those of establishing and maintaining large-scale field trials at several sites for several years. It is our view that genetic evaluation is as essential a component as phenotypic evaluation in research programs aimed at supporting decisions on the management of the genetic resources of forest trees.

Preliminary isozyme studies in several cultivated plants (for example, barley, tomatoes, *Phlox*, radiata pine), have quantified the impact of domestication on genetic variation. The erosion of genetic resources and, hence, the greater genetic vulnerability of planted populations, is evident. Comparative studies of variation patterns in trees and herbaceous plants emphasize the diversity of population genetic structures in plants. The degree of geographic differentiation among populations is more limited in trees than in herbs, but nonetheless significantly affects sampling and conservation strategies. Estimates of present day heterozygosity, mutation rates, gene flow, and differentiation permit an objective approach to questions of remnant population size and distribution. Several recommendations of minimum populations size may be compared within practical and species-specific constraints. Effective population sizes of about 70 allow one-half the current heterozygosity to remain after 100 generations, whereas the number of trees required to ensure with 95 percent certainty that one-half the current alleles survive the bottleneck is approximately the square root of the current population size. As further isozyme data accrue, more precise estimates and strategies will be possible.

Acknowledgments

We are grateful to our colleagues A.G. Brown, K.G. Eldridge, O.H. Frankel, A.C. Matheson, R.N. Oram and J. W. Turnbull for their critical reading of the manuscript.

LITERATURE CITED

- Almgard, G., and U. Landegren.
1974. Isoenzymatic variation used for the identification of barley cultivars. *Zeitschrift für Pflanzenzüchtung* 72:63-73.
- Bergmann, F.
1978. The allelic distribution at an acid phosphatase locus in Norway Spruce (*Picea abies*) along similar climatic gradients. *Theor. Appl. Genet.* 52:51-64.
- Bonnett-Masimbert, M., and V. Bikay-Bikay.
1978. Variabilité intraspécifique des isozymes de la glutamate-oxaloacetate transaminase chez *Pinus nigra* Arnold intérêt pour la taxonomie des sous espèces. *Silvae Genetica* 27:71-79.
- Bouvarel, P.
1970. The conservation of gene resources of forest trees. In *Genetic resources in plants*. O. H. Frankel and E. Bennett, eds. p. 523-530. Blackwell, Oxford.
- Bradshaw, A. D.
1965. Evolutionary significance of phenotypic plasticity in plants. *Adv. Genet.* 13:115-155.
- Britten, R. J., A. Cetta, and E. H. Davidson.
1978. The single-copy DNA sequence polymorphism of the sea urchin *Strongylocentrotus purpuratus*. *Cell* 15:1175-1186.
- Brown, A. H. D.
1975. Sample sizes required to detect linkage disequilibrium between two and three loci. *Theor. Popul. Biol.* 8:184-201.

- Brown, A. H. D.
1978. **Isozymes, plant population genetic structure and genetic conservation.** Theor. Appl. Genet. 52:145-157.
- Brown, A. H. D.
1979. **Enzyme polymorphism in plant populations.** Theor. Popul. Biol. 15:1-42.
- Brown, A. H. D., A. C. Matheson, and K. G. Eldridge.
1975. **Estimation of the mating system of *Eucalyptus obliqua* L'Hérit. using allozyme polymorphisms.** Aust. J. Biol. Sci. 23:931-949.
- Charlesworth, Brian.
1972. **Selection in populations with overlapping generations. III. Conditions for genetic equilibrium.** Theor. Popul. Biol. 3:377-395.
- Crick, Francis.
1979. **Split genes and RNA splicing.** Science 204:264-271.
- Crow, James F., and Motoo Kimura.
1970. **An introduction to population genetics theory.** p. 322. Harper and Row, New York.
- Cullis, C. A.
1977. **Molecular aspects of the environmental induction of heritable changes in flax.** Heredity 38:129-154.
- Ewens, W. J.
1972. **The sampling theory of selectively neutral alleles.** Theor. Popul. Biol. 3:87-112.
- Feret, P. P., and F. Bergmann.
1976. **Gel electrophoresis of proteins and enzymes.** In Modern methods in forest genetics. J. P. Miksche, ed. p. 49-78. Springer-Verlag, Berlin.
- Frankel, O. H.
1970. **Variation—the essence of life.** Proceedings of Linnean Society of N.S.W. 95:158-169.
- Frankel, O. H.
1974. **Genetic conservation: our evolutionary responsibility.** Genetics 78:53-65.
- Frankel, O. H.
1977. **Philosophy and strategy of genetic conservation in plants.** 3d World Consult. For. Tree Breeding. Canberra 1:1-11.
- Guries, R. P., and F. T. Ledig.
1977. **Analysis of population structure from allozyme frequencies.** Proc. South. For. Tree Improv. Conf. 14:246-253.
- Hall, R. B., J. P. Miksche, and K. M. Hansen.
1976. **Nucleic acid extraction, purification, reannealing and hybridization methods.** In Modern methods in forest genetics. J. P. Miksche, ed. p. 19-48. Springer-Verlag, Berlin.
- Hamrick, J. L., Y. B. Linhart, and J. B. Mitton.
1979. **Relationships between life history characteristics and electrophoretically-detectable genetic variation in plants.** Ann. Rev. Ecol. Syst. 10:173-200.
- Hart, G. E., and P. J. Langston.
1977. **Chromosomal location and evolution of isozyme structural genes in hexaploid wheat.** Heredity 39:263-277.
- Hubby, J. L., and R. C. Lewontin.
1966. **A molecular approach to the study of genetic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*.** Genetics 54:577-594.
- Hvid, Søren, and Gunnar Nielsen.
1977. **Esterase isoenzyme variants in barley.** Hereditas 87:155-162.
- Johnson, George B.
1974. **Enzyme polymorphism and metabolism.** Science 184:28-37.
- Koehn, Richard K., and Walter F. Eanes.
1978. **Molecular structure and protein variation within and among populations.** Evol. Biol. 11:39-100.
- Law, G. R. J.
1967. **Alkaline phosphatase and leucine aminopeptidase association in plasma of the chicken.** Science 156:1106-1107.
- Leigh Brown, A. J., and C. H. Langley.
1979. **Reevaluation of level of genic heterozygosity in natural population of *Drosophila melanogaster* by two dimensional electrophoresis.** Proc. Natl. Acad. Sci. 76:2381-2384.
- Lewontin, R. C.
1974. **The genetic basis of evolutionary change.** 346 p. Columbia Univ. Press, New York.
- Libby, W. J.
1973. **Domestication strategies for forest trees.** Can. J. Forest Res. 3:265-276.
- Libby, W. J., R. F. Stettler, and F. W. Seitz.
1969. **Forest genetics and forest tree breeding.** Ann. Rev. Genet. 3:469-494.
- Lundkvist, Kenneth.
1979. **Allozyme frequency distributions in four Swedish populations of Norway spruce (*Picea abies* K.). I. Estimates of genetic variation within and among populations, genetic linkage and among mating system parameters.** Hereditas 90:127-143.
- Marshall, D. R., and A. H. D. Brown.
1975a. **Optimum sampling strategies in genetic conservation.** In Genetic resources for today and tomorrow. O. H. Frankel and J. G. Hawkes, eds., p. 53-80. Cambridge Univ. Press, Cambridge.
- Marshall, D. R., and A. H. D. Brown.
1975b. **The charge state model of protein polymorphisms in natural populations.** J. Mol. Evol. 6:149-63.
- Mejnartowicz, L.
1976. **Genetic investigations on Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco) populations.** Arboretum Kornickie 21:126-187.
- Mitton, J. B., Y. B. Linhart, J. L. Hamrick, and J. S. Beckman.
1977. **Observations on the genetic structure and mating system of ponderosa pine in the Colorado front range.** Theor. Appl. Genet. 51:5-13.
- Muskai, Terumi, and C. Clark Cockerham.
1977. **Spontaneous mutation rates at enzyme loci in *Drosophila melanogaster*.** Proc. Natl. Acad. Sci. 74:2514-2517.
- Nei, Masatoshi.
1975. **Molecular population genetics and evolution.** 288 p. Elsevier, New York.
- Nevo, Eviatar.
1978. **Genetic variation in natural populations: patterns and theory.** Theor. Popul. Biol. 13:121-177.
- Powell, Jeffry R.
1975. **Protein variation in natural populations of animals.** Evol. Biol. 8:79-119.
- Richardson, S. D.
1970. **Gene pools in forestry.** In Genetic resources in plants. O. H. Frankel and E. Bennett, eds. p. 353-365. Blackwell, Oxford.
- Rudin, Dag, and Dag Lindgren.
1977. **Isozyme studies in seed orchards.** Studia Forestalia Suecica 139:1-23.
- Smith, M.
1979. **The first complete nucleotide sequencing of an organism's DNA.** Amer. Sci. 67:57-67.
- Stern, Klaus, and Laurence Roche.
1974. **Genetics of forest ecosystems.** 330 p. Springer-Verlag, Berlin.
- Wang, B. S. P.
1975. **Tree seed and pollen storage for genetic conservation: possibilities and limitations.** In Methodology of conservation of forest genetic resources. p. 93-103. FAO, Rome.
- Wilson, Allan C., Steven C. Carlson, and Thomas J. White.
1977. **Biochemical evolution.** Ann. Rev. Biochem. 46:573-639.
- Witter, M. S., and P. P. Feret.
1978. **Inheritance of glutamate oxaloacetate transaminase isozymes in Virginia pine megagametophytes.** Silvae Genetica 27:128-134.
- Yang, J.-Ch., T. M. Ching, and K. K. Ching.
1977. **Isoenzyme variation of coastal Douglas-fir. 1. A study of geographic variation in three enzyme systems.** Silvae Genetica 26:10-18.

Isozyme Variation and Linkage in Six Conifer Species¹

M. Thompson Conkle²

Abstract: Isozymes of female gametophyte tissue were analyzed for allelic variation in knobcone, lodgepole, loblolly, Jeffrey, and sugar pines and in Douglas-fir. Linkage was studied in the five pines. The average number of alleles and average heterozygosity per enzyme locus were estimated. Knobcone pine ranked lowest among the six species in number of alleles and average heterozygosity; loblolly pine and Douglas-fir ranked highest. Numerous linkages were identified, and one chromosome segment contained between one-fourth and one-third of all loci examined. Linked loci had a consistent gene order in the different pine species. Linkage relationships support the concept that similar isozymes represent the same loci in each pine species. The consistency with which loci map in the same order and with similar map distances in species from different subsections of the genus, further suggests that evolution of the pines is not associated with major chromosome rearrangements.

The analysis of genetic variation in forest trees entered a new phase with the application of electrophoresis to detect allelic differences at isozyme loci. Gametophytes from seed of single trees provide excellent material for assessing isozyme variation. Enzymes are analyzed at the just-germinated stage of development, and diploid genotypes of parent trees are inferred from the segregation of allele phenotypes of the haploid female gametophytes. Seed from some of the heterozygous parent trees are analyzed in large numbers to determine if enzyme variants segregate according to simple Mendelian ratios. A one-to-one segregation ratio of phenotypic variants is evidence for allelism. Homozygous loci within a single tree are identified without error and the samples per tree are sufficient to minimize the misclassification of heterozygous loci.

This paper reports a two-stage analysis. In the first stage, the amount of genetic variation in several species of two conifer genera was estimated. Gametophytes from seed of individual trees were analyzed for segregation and data were obtained on the number of alleles and the average heterozygosity per locus. The six species compared were Douglas-fir and five North American pines. In the second stage, linkage from segregation in female gametophytes of heterozygous trees was estimated. Linkage data are provided for the five pine species.

ISOZYME VARIATION

The six conifers in this study were knobcone pine (*Pinus attenuata* Lemm.), lodgepole pine (*P. contorta* Dougl. ex Loud.), loblolly pine (*P. taeda* L.), Jeffrey pine (*P. jeffreyi* Grev. & Balf.), sugar pine (*P. lambertiana* Dougl.), and

Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco). The geographic origins and numbers per species are listed in *table 1*. Sugar pine is classed within the subsection *Strobi* in the subgenus *Strobus* (the white pines). Within the subgenus *Pinus* (the hard pines), loblolly pine is within the subsection *Australes*, Jeffrey pine is in the subsection *Ponderosae*, lodgepole pine is within the subsection *Contortae*, and knobcone pine is in the subsection *Oocarpae*.

The minimum sample size was six gametophytes per tree and heterozygous loci were only misclassified when all six gametophytes carried one allele or its alternate, a probability of $2X (1/2)^6$, or 0.03. Standard laboratory techniques for starch gel electrophoresis provided information for 21 enzyme systems (*table 2*) and as many as 43 loci per gametophyte.

While analyzing the loci of these species, some general observations concerning gene phenotypes became obvious:

- The number of zones that developed bands within a particular enzyme system was consistent for different species. That is, if an enzyme system resolved only one zone of activity for a particular species, it resolved a single zone of activity for the other species.
- The relative mobilities of bands within enzyme systems were similar for different species. For example, if an enzyme system of one species had two zones of activity, two zones were also found in other species and the bands had identical or nearly identical mobilities.
- The characteristic isozyme banding phenotypes for different loci were consistent among different species. For example, band phenotypes for a particular locus were very dark staining and wide, and a second locus had band phenotypes that were light staining and narrow. These two loci had similar appearances in different species and genera. The number, mobility, and appearances of primary bands, and the characteristics of nearby secondary bands—whether leading, trailing, or in multimeric combinations—were constant for the conifer species I analyzed.

¹Presented at the Symposium on Isozymes of North American Forest Trees and Forest Insects, July 27, 1979, Berkeley, Calif.

²Geneticist, Pacific Southwest Forest and Range Experiment Station, Forest Service, U.S. Department of Agriculture, Berkeley, Calif.

Table 1—Species, geographic origin, and number of trees in the enzyme analyses

Species	Stands	Geographic origin	Trees
Knobcone pine (<i>Pinus attenuata</i> Lemm.)	10	Range wide	49
Lodgepole pine (<i>P. contorta</i> Dougl. ex Loud.)	1	Sierra Nevada, California	40
Loblolly pine (<i>P. taeda</i> L.)			
Natural stand	1	Schenck Forest, Raleigh, North Carolina	146
Superior trees	(Sample)	Throughout Southeastern U.S. sample of trees from the North Carolina State Tree Improvement Program	90
Jeffrey pine (<i>P. jeffreyi</i> Grev. & Balf.)	4	Central Sierra Nevada, Sierra Nevada, California	75
Sugar pine (<i>P. lambertiana</i> Dougl.)	(Sample)	Rangewide in California	58
Douglas-fir (<i>Pseudotsuga menziesii</i> [Mirb]Franco)	1	Cascades, west-central Oregon	152

Table 2—Enzyme systems analyzed from conifer gametophytes

Isozyme	Abbrev.	Commission Enzyme Code	Reference
Acid phosphatase	ACPH	3.1.3.2	Scandalios 1969
Aconitase	ACO	4.2.1.3	Yeh and O'Malley ¹
Alcohol dehydrogenase	ADH	1.1.1.1	Scandalios 1969
Aldolase	ALD	4.1.2.13	Yeh and O'Malley ¹
Catalase	CAT	1.11.1.6	Scandalios 1969
Diaphorase	DIA	1.6.4.3	Yeh and O'Malley ¹
Esterase (colorimetric)	EST	3.1.1.2	Scandalios 1969
Esterase (florescent)	FLEST	3.1.1.2	Mitton and others 1979
Glutamate dehydrogenase	GDH	1.4.1.2	Shaw and Koehn 1968
Glutamate oxalacetic transaminase	GOT	2.6.1.1	Brewbaker and others 1968
Glucose-6-phosphate dehydrogenase	G6PD	1.1.1.49	Shaw and Prasad 1970
Isocitrate dehydrogenase	IDH	1.1.1.42	Nichols and Ruddle 1973
Malic dehydrogenase	MDH	1.1.1.37	Nichols and Ruddle 1973
Peptidases			
Leucine-amino peptidase	LAP	3.4.11.1	Scandalios 1969
Alanine-amino peptidase	ALAP	3.4.11.1	Ott and Scandalios 1978
Peptidase	PEP	3.4.13.11	Nichols and Ruddle 1973
Peroxidase	PER	1.11.1.7	Shaw and Prasad 1970
6-Phospho-gluconate dehydrogenase	6PGD	1.1.1.44	Brewer 1970
Phosphoglucose isomerase	PGI	5.3.1.9	Brewer 1970
Phosphoglucomutase	PGM	2.7.5.1	Brewer 1970
Superoxide dismutase	SOD	1.15.1.1	

¹ Yeh, F. C., and D. M. O'Malley. 1980. Enzyme variation in natural populations of Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco) from British Columbia. I. Genetic variation patterns in coastal populations. (In press).

It seemed reasonable to assume, therefore, that the same genes were being expressed in different species.

Two measures of isozyme variation were evaluated. The number of alleles per locus (A), was a count of the different allele phenotypes; heterozygosity per locus (H), was computed from allele frequencies. The H values, derived from Hardy-Weinberg expectations, were computed by subtracting the sum of the squared allele frequencies for each locus from 1.00. This measure was chosen instead of the direct count of heterozygotes because it is the value most widely reported in the literature. Admittedly, it is a first approximation that is accurate only when the trees within a species are randomly interbreeding, and when geographic differentiation within a species is lacking. Although supporting data will not be provided here, the observed proportions of heterozygotes per gene per species closely approximated the computed values.

The estimates of isozyme variation for the six species studied resulted in few general statements regarding variability (*table 3*). The values represented contrasts; for example, the number of alleles per locus (A) for ACPH-2 were low (2 or 3) in knobcone pine, sugar pine, and Douglas-fir, intermediate (5) in lodgepole and Jeffrey pines, and high (8) in loblolly pine. Heterozygosity (H) of ACPH-2 was low for knobcone pine and sugar pine, 0.06 and 0.02, but high for Douglas-fir (0.49). ADH-2 had values of A ranging from 2 to 6 with H values from 0.02 to 0.68. A and H are sometimes widely divergent: Jeffrey pine had 3 alleles for ADH-2 and an H value of 0.60; the superior trees in the loblolly sample had 5 alleles and a low H estimate (0.06). GDH was homozygous in knobcone, Jeffrey, and sugar pines, but was highly variable within lodgepole and loblolly pines. GOT-1 was nearly homozygous in Douglas-fir but variable in the five pine species. Characteristically, EST-1 had numerous alleles per locus and high values for H, and LAP-1 and LAP-2 were variable in all six species.

Values for A and H averaged for all loci show differences for the six species in the study. In general, the A and H values have similar trends. The average number of alleles per locus grouped into three general classes, with values near 2, 3, and 4 (*table 3*). Knobcone pine had the lowest value for A with an average of 2. Lodgepole, Jeffrey, and sugar pines had values approaching 3 alleles; loblolly pine and Douglas-fir had high average values approaching 4 alleles per locus. Average heterozygosity values ranged from a low of 0.13 for knobcone pine to a high of 0.36 for the natural stand of loblolly pine; Jeffrey and sugar pines approached intermediate values; and, loblolly pine and Douglas-fir averaged high H values.

I believe the two values for loblolly pine may be similar. Slightly different H values for these resulted from inclusion of the additional loci in the superior tree sample. The additional loci did not modify the A mean values, but usually one allele was in high frequency. The result was the addition of several loci with low values for H. The data from the two sources of loblolly pine should not be used to argue that a higher average heterozygosity exists in natural

stands than in a broad geographic sample of superior trees. When comparisons were restricted to the nine loci analyzed within both the natural stand and the select trees, the superior trees averaged more alleles per locus than the natural stand (4.78 compared with 3.89) and H values were similar for both samples.

Estimates of A and H for numerous species were compared by Brown and Moran (1981) and Hamrick and others (1981), but two points need mention. The lowest estimate of A and H for conifers has been found in red pine (Fowler and Morris 1977). In a sample of nine isozyme loci, no variation could be attributed to allelic differences; therefore, the estimates of genetic variation are $A = 1.00$ and $H = 0.00$. This finding is remarkable when compared with other conifers. It is, however, consistent with findings that red pine is highly self-fertile, has low variability in growth traits, and does not exchange genes with any other species. Some researchers would argue that significant genetic variance in phenotypic traits has yet to be demonstrated in red pine.

The estimates of nearly four alleles per locus and heterozygosities near one-third for loblolly pine and Douglas-fir are among the highest values obtained for any species of higher organism thus far reported (Hamrick and others 1981). As data become available for other conifers, a continuum of values for genetic variation in isozymes is expected. The six species reported here probably will rank from below average to among most variable.

LINKAGE

Isozyme analyses of female gametophytes tested deviations from independent assortment between gene loci. Data on genetic linkage were obtained on the five pines, but not on Douglas-fir. Highly heterozygous trees, identified in the species surveys for allele frequencies, were efficient items for estimating linkages. The number of pair-wise gene combinations increased exponentially with increasing numbers of heterozygous loci.

Pair-wise linkage of loci was established when the sum of the recombinant classes had depressed frequencies and the parental classes had inflated frequencies in comparison with random expectations. Species recombination values were computed from pooled single-tree estimates and all recombination values were adjusted to provide estimates of map distances in centiMorgans (cM) (Kasambi 1944). Complete data on sample size and segregation ratios will be published elsewhere; however, each distance was estimated with seed from a minimum of two different trees within each species. The minimum number of gametophytes used to estimate distances was 144 and the maximum exceeded 1000. Gene order was established by comparing distances in multiple heterozygotes. Because data were limited by the availability of appropriate multiple heterozygotes, not all gene combinations could be examined in each species.

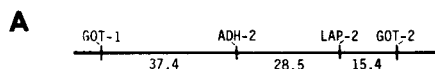
Linked genes were distributed among several linkage blocks in the five pine species. Linkage maps were

Table 3—Number of alleles per locus (A) and heterozygosity per locus (H) for five pine species and Douglas-fir

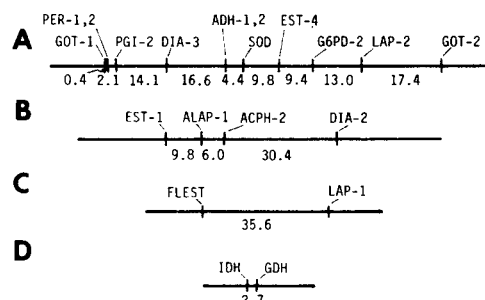
Locus	Knobcone pine		Lodgepole pine		Loblolly pine (natural stand)		Loblolly pine (superior trees)		Jeffrey pine		Sugar pine		Douglas-fir	
	A	H	A	H	A	H	A	H	A	H	A	H	A	H
ACPH 1	2	0.06	2	0.04			4	0.33	4	0.27				
ACPH 2	2	0.06	5	0.41	7	0.32	8	0.20	5	0.40	2	0.02	3	0.49
ACPH 3			1	0.00										
ACPH 4	2	0.06	2	0.04			2	0.46						
ACO			2	0.21					3	0.40				
ADH 2	3	0.45	5	0.53	2	0.02	5	0.06	3	0.60	6	0.68		
ALD 1			2	0.04					2	0.33				
ALD 2			2	0.04					1	0.00				
CAT 1			2	0.04					3	0.35				
DIA 1			3	0.08					5	0.55				
DIA 2			3	0.08					2	0.25				
DIA 3									2	0.01				
DIA 4			3	0.26					2	0.32				
EST 1	6	0.46	7	0.52	6	0.72	9	0.51	6	0.63	3	0.34	8	0.69
EST 2			3	0.65					6	0.61				
EST 4			3	0.37	4	0.47	4	0.53	3	0.25				
FLEST			2	0.08					3	0.43				
GDH	1	0.00	3	0.30			3	0.37	1	0.00	1	0.00		
GOT 1	2	0.23	4	0.32	3	0.37	3	0.41	4	0.41	4	0.37	2	0.01
GOT 2	2	0.04			3	0.54			3	0.08	2	0.04	4	0.21
GOT 3	2	0.02	2	0.02	2	0.33	3	0.27	2	0.13	5	0.30	3	0.16
G6PD 1	2	0.04	2	0.08			3	0.22	3	0.16	3	0.08	4	0.54
G6DP 2	1	0.00	2	0.08			3	0.29	2	0.64	2	0.02	2	0.40
IDH	2	0.06	2	0.02			2	0.04	2	0.01	1	0.00	5	0.39
MDH 1	1	0.00	2	0.08			4	0.10	2	0.01	2	0.48	4	0.05
MDH 3	2	0.08	5	0.76			4	0.52	3	0.24	2	0.45	2	0.32
MDH 4	1	0.00	3	0.18			3	0.04	2	0.51	6	0.70	5	0.44
LAP 1	3	0.46	4	0.32	4	0.24	4	0.24	6	0.55	3	0.20	8	0.72
LAP 2	2	0.20	5	0.20	4	0.51	4	0.52	4	0.12	6	0.68	4	0.18
ALAP 1			2	0.25			3	0.08	3	0.31				
ALAP 2			2	0.10			3	0.13	4	0.32				
PEP 1			1	0.00					1	0.00				
PEP 2									1	0.00				
PER 1			2	0.21			4	0.51	3	0.45				
PER 2			7	0.65			7	0.40	5	0.64				
6PGD 1	2	0.43	2	0.04			8	0.65	3	0.28	2	0.48	3	0.14
6PGD 2	2	0.02	2	0.08			2	0.08	3	0.07			3	0.44
PGI 1			2	0.04					2	0.01				
PGI 2			2	0.04					3	0.24				
PGM	2	0.08	1	0.00	3	0.10	3	0.10	2	0.20	2	0.38	3	0.31
SOD 1	1	0.00	1	0.00			1	0.00	2	0.03	1	0.00	4	0.14
SOD 2	1	0.00	2	0.04					1	0.00	1	0.00		
SOD 3									1	0.00				
SOD 4									2	0.03				
SOD 5									3	0.39				
TOTAL A	22		39		10		25		43		19		17	
Mean	2.00		2.74		3.80		3.96		2.86		2.84		3.94	
S.E.	±0.22		±0.23		±0.49		±0.39		±0.21		±0.39		±0.42	
H														
Mean	0.125		0.185		0.362		0.282		0.261		0.275		0.331	
S.E.	±0.035		±0.032		±0.063		±0.038		±0.032		±0.057		±0.049	

¹Construction of the table implies that a particular gene is common to all six species. No importance should be attached to the blanks within the table; species with the greater numbers of loci reflect laboratory advances over earlier analyses.

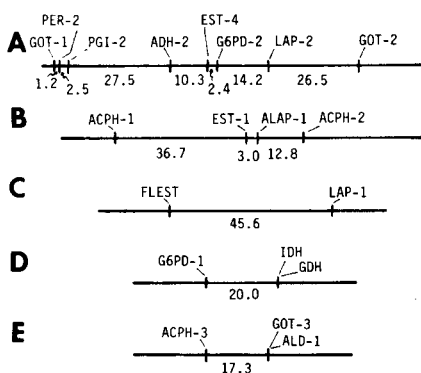
Knobcone Pine



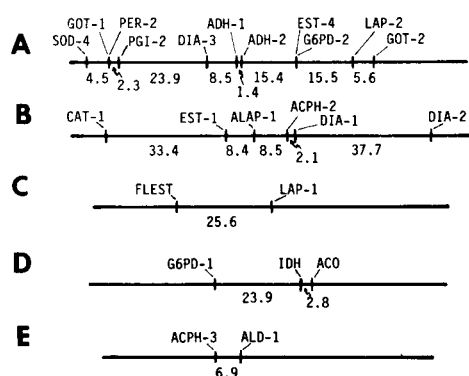
Lodgepole Pine



Loblolly Pine



Jeffrey Pine



Sugar Pine

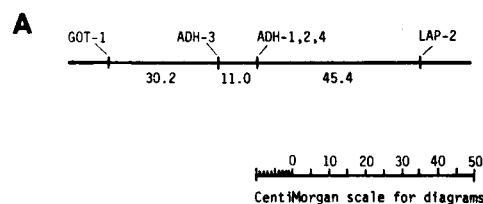


Figure 1—Linked genes detected in five species of pine with map distances in centiMorgan units.

constructed with a common base using cM distances and linkage groups were indicated by capital letters (*fig. 1*).

A high proportion of the genes sampled were linked. Some were closely associated, others were separated by distances that required large sample sizes for detection.

The presence and arrangement of particular genes within chromosome segments for different species were not random; linkage blocks had common relationships for all species. Segment A, for example, was the longest linkage group found. For lodgepole, loblolly, and Jeffrey pines, the number of genes in linkage block A was large; for example, 12 for lodgepole pine. And the gene order within the block was almost identical for all five species. Three closely associated genes, GOT-1, PER-2, and PGI-2 anchored the left-hand portion of the segment, ADH-2 was 30 to 35 map units to the right, and EST-4 and G6PD-2 were to the right of ADH-2. Because LAP-2 was a substantial distance from ADH-2, the estimates were subject to large sampling errors. Nevertheless, ADH-2 and LAP-2 linkage estimates were statistically significant and map distances ranged from 25.4 in Jeffrey pine to 45.4 in sugar pine. In most species GOT-2 was difficult to resolve but some families provided estimates that located it to the right of LAP-2. Weak evidence suggests that linkage group A was linked with group B. If so, this means that a larger proportion of the genes in the total sample of loci were on a single chromosome. The evidence for linkage between blocks A and B needs further analysis; it is premature to consider them linked at this time.

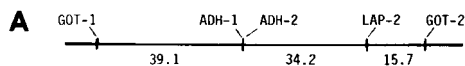
Linkage group B was characterized by a block of three genes: EST-1, ALAP-1, and ACPH-2. They were found in the same order in lodgepole, loblolly, and Jeffrey pines with ALAP-1 in the center.

Linkage block C represented a long segment including FLEST and LAP-1. The main features of the D linkage block, investigated in lodgepole, loblolly, and Jeffrey pines, was a close linkage between IDH and GDH, with G6PD-1 located about 20 map units from IDH and GDH. In loblolly pine, GOT-3 was linked with ALD-1 with no recombination, and both genes were linked to ACPH-3 (linkage block E). One final linkage, block F, was observed between MDH-2 and TO-5 in families of Jeffrey pine.

Two reports of linkage groups in pines compare favorably with these findings. Rudin and Ekberg (1978) report a linkage block in Scots pine (*P. sylvestris* L.) consisting of GOT-1, ADH-1 and 2, LAP-2, and GOT-2 (*fig. 2*). Map distances in Scots pine are comparable to those reported for the species in this paper, and the arrangement is identical. Similar linkage in Scots pine is important because it samples the subsection *Sylvestres* which includes most pines of the Eastern Hemisphere. Guries and others (1978) found recombination frequencies of 0.03 and 0.04 between GOT-1 and PGI-2 (their GPI-2) for two trees of pitch pine (*P. rigida* Mill.) and these genes are tightly linked in lodgepole, loblolly, and Jeffrey pines.

Isozyme phenotypes led me to assume that many genes were common to different species and the linkage data supports this assumption. Genes, with common pheno-

Scots Pine
(Rudin and Ekberg, 1978)



Pitch Pine
(Guries and others, 1978)

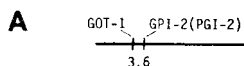


Figure 2—Linkage in two pine species, previously reported, compare well with the findings of this study.

types, map in the same order and with about the same distances in the different species.

Gene arrangement is highly conservative in the pines studied. These data indicate that evolution within the genus has not required large structural rearrangements within or between the gene blocks thus far examined. The linkage results invite comparison with other work on the arrangement of the genetic material in conifers. Specifically, the cytological studies of Sax (1960) and Saylor (1972) on pine chromosomes indicate that the karyotypes of conifers are uniform and almost indistinguishable. Major structural differentiation in the form of inversions and translocations is lacking, according to evidence from interspecific hybrids.

The interpretation of isozyme studies substantiate the conclusions of cytologists. Heterozygous inversions would suppress recombination, and inversion homozygotes would produce recombinants. If the trees differed for inversions, then linkage estimates could be expected to show wide divergence from tree to tree. Although more data are desirable, variances in linkage estimates from different trees in this study were within limits of random error and did not suggest the presence of inversion polymorphisms.

Saylor and Smith (1966) found no evidence of heterozygous translocations and the isozyme data support this observation. Translocations would disturb linkage relationships but linkage block A thus far demonstrates a consistent relationship between major genes in the block, not only within but also between the pine species thus far studied.

Data showed that about one-fourth to one-third of the sampled loci were associated with one linkage block; pines have 12 chromosomes, all but one or two with median centromeres (Saylor 1972). Twelve or 24 linkage groups might be expected, because whether considering the number of chromosomes or the number of chromosome arms, a random distribution throughout the chromosome complement of the 10 to 43 genes analyzed would lead one to expect a moderate number of linkages with low numbers of genes per linkage block. The likelihood of various numbers of genes being linked by chance on pine chromosomes, or on arms of chromosomes, can be

approximated by the binomial distribution function. Linkage analyses evaluated about 35 genes and the expected distributions were computed on this basis (fig. 3). For 12 chromosomes, about 5 percent would not receive a gene, 15 percent would get one, 23 percent would have 2, and so on. This study indicated that as many as 12 genes were linked in the A block. The probability of a linkage block this large or larger on a chromosome is 1.5×10^{-5} . When chromosome arms are considered, the proportions of blocks with numerous genes are reduced. The probability of a linkage block with 10 or more genes on a chromosome arm is 1.0×10^{-7} .

If we are to maintain that the large number of genes in the A linkage block results from some causal association of loci, the obvious observation is that the genes are all active during a specific developmental stage, namely, the growth of the developing embryo. The function of ADH, which is in the center of this segment, is the conversion of the end products of anaerobic respiration. Esterases are involved in the metabolism of fats and peroxidases are known to serve as IAA oxidases. These functions would implicate them in processes of embryo growth.

Close linkages suggest the possibility that selection operates upon specific allelic combinations. Fitness values for tightly linked loci depend on the multilocus combination of alleles rather than the fitness of individual loci (Lewontin 1974). Although pine species vary in the number of alleles per locus for the three tightly linked loci on the A block, GOT-1, PER-2, and PGI-2, three alleles per locus would produce 27 gametic classes and a total of 378 zygotic classes. The problem of estimating fitness values for such a complex system is overwhelming.

Lewontin (1974) cited Fisher's proposal that natural selection favors closer linkage of genes influencing the same polymorphism. Tight linkage could slow the process

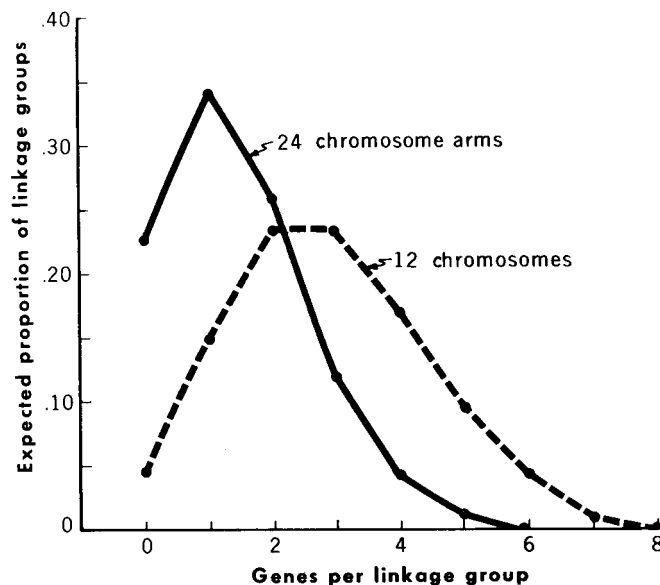


Figure 3—Expected proportion of linkage groups with different numbers of genes per group when 35 genes are randomly allotted to chromosomes and chromosome arms.

of allele fixation at a locus by operating upon complexes of loci and the effect could extend to include genes at greater map distances (Lewontin 1974). The analysis of pine loci points to a significantly large block of linked genes. One could question if this linkage block represents genes that are related in function but sufficiently separated to accommodate recombination.

Conifers stand, almost unchallenged, as higher organisms that readily yield information on isozyme linkage. Data in pines strongly suggest nonrandom association of numerous loci and provide new evidence on conifer evolution. Allelic variation within these linkage blocks may provide for precise genetic evaluation of species relationships. Linkages may also serve as the basis for mapping and manipulating genetically controlled phenotypic traits. These possibilities bring closer the goal of relating genetic variation in enzymes to variation in phenotypic responses; and conifers, once thought to be poor organisms for precise genetic analyses, offer advantages over many, if not most, other higher organisms.

LITERATURE CITED

- Brewbaker, James L., Mahesh D. Upadhyaya, Yrjo Makinen, and Timothy MacDonald.
1968. **Isozyme polymorphism in flowering plants. III. Gel electrophoretic methods and applications.** *Physiologia Plantarum* 21:930-940.
- Brewer, George J.
1970. **An introduction to isozyme techniques.** 186 p. Academic Press, New York and London.
- Brown, A.H.D., and G.F. Moran.
1981. **Isozymes and the genetic resources of forest trees.** In Proceedings of symposium on isozymes of North American forest trees and forest insects, July 27, 1979, Berkeley, Calif. Gen. Tech. Rep. PSW-48, p. 1-10. Pacific Southwest Forest and Range Exp. Stn., Forest Serv., U.S. Dep. Agric., Berkeley, Calif.
- Fowler, D.P., and R.W. Morris.
1977. **Genetic diversity in red pine: evidence for low genic heterozygosity.** *Can. J. For. Res.* 7:343-347.
- Guries, R.P., S.T. Friedman, and F.T. Ledig.
1978. **A megagametophyte analysis of genetic linkage in pitch pine (*Pinus rigida* Mill.).** *Heredity* 40:309-314.
- Hamrick, J.L., J.B. Mitton, and Y.B. Linhart.
1981. **Levels of genetic variation in trees: influence of life history characteristics.** In Proceedings of symposium on isozymes of North American forest trees and forest insects, July 27, 1979, Berkeley, Calif. Gen. Tech. Rep. PSW-48, p. 35-41. Pacific Southwest Forest and Range Exp. Stn., Forest Serv., U.S. Dep. Agric., Berkeley, Calif.
- Kosambi, D.D.
1944. **The estimation of map distance from recombination values.** *Ann. Eugen.*, London 12:172-175.
- Lewontin, R.C.
1974. **The genetic basis of evolutionary change.** 346 p. Columbia Univ. Press, New York and London.
- Lundkvist, Kenneth.
1979. **Allozyme frequency distributions in four Swedish populations of Norway spruce (*Picea abies* K.).** *Hereditas* 90:127-143.
- Nichols, Elizabeth A., and Frank H. Ruddle.
1973. **A review of enzyme polymorphism, linkage and electrophoretic conditions for mouse and somatic cell hybrids in starch gels.** *J. Histochem. and Cytochem.* 21:1066-1081.
- Mitton, J. B., Y. B. Linhart, K. B. Sturgeon, and J. L. Hamrick.
1979. **Allozyme polymorphisms detected in mature needle tissue of ponderosa pine.** *J. Heredity* 70:86-89.
- Ott, Lila, and John G. Scandalios.
1978. **Genetic control and linkage relationships among aminopeptidases in maize.** *Genetics* 89:137-146.
- Rudin, D., and I. Ekberg.
1978. **Linkage studies in *Pinus sylvestris* L. —using macro gametophyte allozymes.** *Silvae Genetica* 27:1-12.
- Sax, Karl.
1960. **Meiosis in interspecific pine hybrids.** *Forest Sci.* 6:135-138.
- Saylor, L.C.
1972. **Karyotype analysis of the genus *Pinus*—subgenus *Pinus*.** *Silvae Genetica* 21:155-163.
- Saylor, LeRoy C., and Ben W. Smith.
1966. **Meiotic irregularity in species and interspecific hybrids of *Pinus*.** *Amer. J. Bot.* 53:453-468.
- Scandalios, John G.
1969. **Genetic control of multiple molecular forms of enzymes in plants: a review.** *Biochem. Genetics* 3:37-39.
- Shaw, Charles R., and Ann L. Keen.
1968. **Starch gel zone electrophoresis of enzymes.** In *Chromatographic and electrophoretic techniques. II. Zone electrophoresis*, Charles R. Smith, 2d ed., p. 325-364. Interscience Publishers, New York.

Analysis of Mating System Parameters and Population Structure in Douglas-fir Using Single-locus and Multilocus Methods¹

D. V. Shaw² and R. W. Allard³

Abstract: Two methods of estimating the proportion of self-fertilization as opposed to outcrossing in plant populations are described. The first method makes use of marker loci one at a time; the second method makes use of multiple marker loci simultaneously. Comparisons of the estimates of proportions of selfing and outcrossing obtained using the two methods are shown to yield additional information about inbreeding and the genetic structure of populations. Single-locus estimates of the proportion of selfing in Douglas-fir were heterogeneous over marker loci and lower than estimates obtained with the multilocus estimator. Family structure and/or microhabitat selection were discussed as likely causes of the biased estimates of the proportion of selfing obtained from the single-locus estimation procedure.

An important goal of research in forest genetics is the development of methods that allow for the breeding of superior stock and also provide for the long-term maintenance of adequate germplasm resources. In realizing this dual goal, knowledge of the forces that are responsible for the distribution of genotypes in natural stands of forest trees, including mating system and gene flow, will be particularly useful. Several methods are available for the quantification of parameters which specify different aspects of mating systems, family structure, and gene flow. These methods have been used to estimate inbreeding coefficients (F) and the proportions of selfing (s) and outcrossing (t) (Sorensen 1973, Rudin 1977), as well as the effects of inbreeding depression on quantitative characters (Franklin 1970).

Electrophoretically detectable loci have been particularly useful in the study of mating systems and gene flow in plant populations (Brown and Allard 1970, Clegg and others 1978) for two main reasons. First, electrophoresis often provides an abundance of markers; second, enzyme loci are frequently codominant, allowing all genotypic classes to be identified directly and thus increasing the efficiency of the estimation procedures.

This paper presents estimates of the proportions of selfing and outcrossing in two stands of Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco) based on two different methods of estimation. We show that comparisons of the estimates from the two methods yield additional information about the mating system and about gene flow within and between stands of this species.

MATERIALS AND METHODS

The data we use to illustrate the estimation procedures are from a study of two stands of Douglas-fir. The two stands occupy ridge-top sites which are located about 40 kilometers east of Springfield, Oregon, and about 20 kilometers from each other. The two sites, Springfield 3 (S3) and Springfield 5 (S5), which are both at about 500-m elevation, are occupied by young stands of naturally regenerated trees. Trees on S3 are about 30 years of age and S5 contains trees of 40 to 45 years of age. The stands are composed almost entirely of Douglas-fir, with an occasional western hemlock (*Tsuga heterophylla* [Raf.] Sarg.) interspersed. *Figure 1* gives the spatial distribution of the maternal trees sampled on the two sites. Sampling followed a modification of the Nelder Design (Nelder 1962).

Seeds were germinated and both diploid embryo tissue and haploid gametophytic tissue were analyzed by starch gel electrophoresis.⁴ A minimum of 7 haploid-diploid pairs were analyzed per maternal tree sampled. The following 11 loci were scored for each haploid-diploid pair (numbers in parentheses represent the number of alleles scored for each locus): Got-I (3), Got-III (3), G6pd (2), Gdh (2), To (2), Est (4), Lap-I (3), Lap-II (3), Pgi (3), Pgm-I (3), Pgm-II (3).

STATISTICAL TECHNIQUES

Single-locus Estimation

Fyfe and Bailey (1951) gave the theory of estimation of the proportion of selfing (s) and outcrossing ($t = 1 - s$) for the mixed mating model, using recessive markers. They used

¹Presented at the symposium on Isozymes of North American Forest Trees and Forest Insects, July 27, 1979, Berkeley, Calif.

²Graduate Student, Department of Genetics, University of California, Davis, Calif.

³Professor of Genetics and Agronomy, University of California, Davis, Calif.

⁴Starch gel electrophoretic techniques were developed by M. Thompson Conkle. (Personal commun. 1978.)

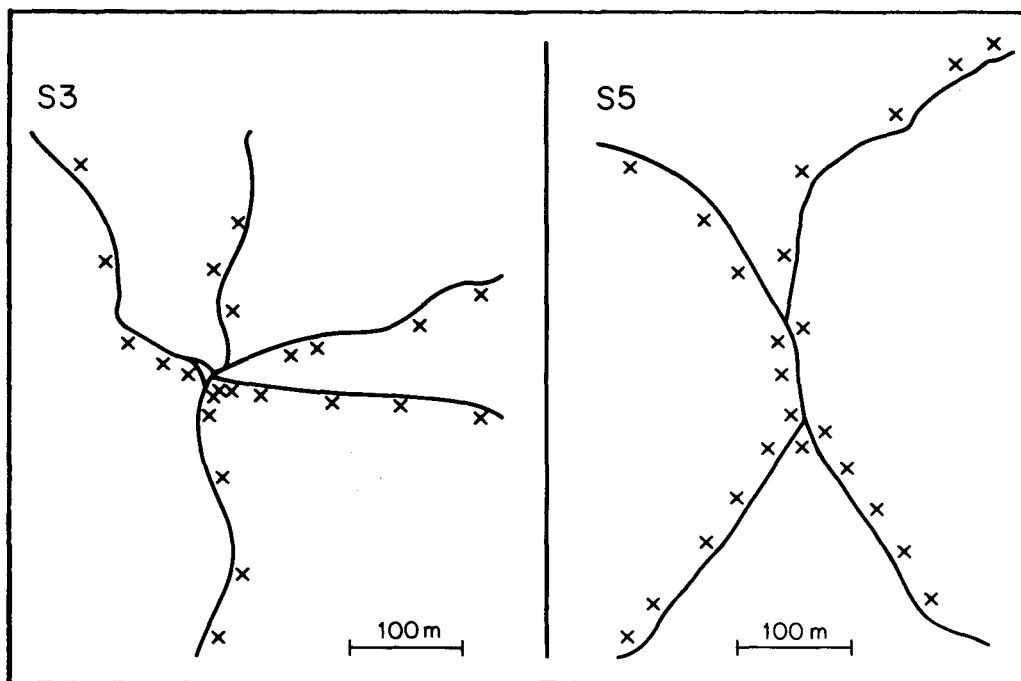


Figure 1—Spatial distribution of sample trees (S3) Springfield 3 site, (S5) Springfield 5 site.

the transformation $f = (1-t)/(1+t)$, where f is the inbreeding coefficient, and formulated the joint maximum likelihood estimators and their variances for f and p (allelic frequencies). Brown and Allard (1970) and Clegg and others (1978) developed maximum likelihood methods for estimating outcrossing rates from the genotypic arrays of half-sib families in which the maternal parent is of unknown genotype. In these estimators the genotype of the maternal parent is inferred from the array of progeny produced by each maternal individual. This inference is affected by the maternal genotype, by allelic frequencies in the pollen pool, and by the mating system parameters themselves. When data from gametophytes are also available, as in the present case, the genotype of each maternal individual can be inferred from the gametic array of each maternal individual and an estimate of maternal genotypic frequencies can be obtained which is statistically independent of the other parameters. The estimator used in the present study was modified from that of Clegg and others (1978) to take advantage of the information provided by the gametophytic analysis. Fyfe and Bailey (1951) pointed out that an important assumption of the mixed mating model is that allelic frequencies are homogeneous in the pollen pool over the entire area sampled. A variety of causes (for example, family structure, microhabitat selection) can lead to heterogeneity of the pollen pool in natural stands and the resulting Wahlund effect leads to overestimation of the amount of selfing (Wahlund 1928). Consider the hypothetical sampling situation given by figure 2. Two contiguous stands, P_1 and P_2 , are reproductively isolated from each other but mating occurs at random within each stand. For a diallelic locus the frequency of allele A_1 is 0.9 in P_1 and 0.1 in P_2 . If an investigator were aware of the existence of the two stands, and sampled the two areas separately, the

single-locus estimator would yield correct estimates of the amount of selfing ($\hat{s} = 0$) and outcrossing ($\hat{t} = 1$). If, however, the investigator were unaware that there were two stands, and sampled as if there were only one stand, the estimate of s would be 0.89 and t would be 0.11. This example demonstrates the potential of pollen heterogeneity for introduction of bias into single-locus estimates, and it points out the need for careful sampling technique. Two important additional assumptions in the mixed mating model are: (a) that the probability of an outcross will not be affected by the maternal genotype and, (b) that selection does not intervene between mating and the determination of progeny genotype distributions. The fact that significant between-locus differences in s and t are often obtained from a single sample of progeny indicates that these assumptions, like the assumption of homogeneity of the pollen pool, are also often not valid.

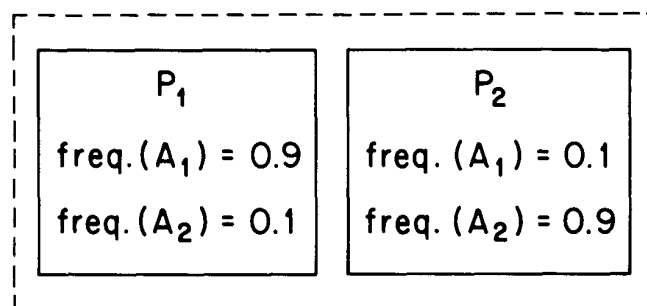


Figure 2—Hypothetical sampling situation showing actual stand boundaries (—) and sampling boundaries (---).

Multilocus Estimation

An alternate method for estimating proportions of outcrossing and selfing, which avoids many of the problems of single-locus estimation, has been developed.⁵ The large number of marker loci provided by electrophoretic analysis makes application of this alternative method feasible in experimental situations. Briefly this method involves the identification of outcross progeny by the comparison of multilocus progeny genotypes with maternal genotypes and using the information gained from single-locus data for statistical compensation for outcrosses which are not identifiable. As more loci are considered the identification of true outcross progeny becomes nearly complete and estimates of s and t become increasingly dependent on observation and less dependent on statistical compensation. One advantage of this estimator is its insensitivity to failure of assumptions that can seriously affect single-locus estimation.

Another advantage is that comparisons of estimates derived from single-locus and multilocus techniques provide a means of separating the effects of selfing from other causes of nonrandom pollen dispersal, thus providing additional information concerning the breeding structure of populations.

RESULTS

Single-locus estimates of the proportions of outcrossing, t , for three loci in the Springfield 3 and Springfield 5 sites are presented in *table 1*. The estimates oft range of 0.66 to 0.97. This heterogeneity indicates that the estimates are affected by factors other than the mating system. Likelihood ratio tests indicate that 7 of the 9 single-locus estimates differ significantly from those expected under random mating ($t = 1$) and an eighth value approached significance ($P = .06$) (Rao 1973). Note also that the estimates obtained by combining the samples from the two sites are always lower than the mean of the two independent samples. This situation parallels the hypothetical case given in *figure 1* and demonstrates that bias is in fact introduced by combining heterogeneous samples.

Estimates of the proportion of outcrossing($\hat{t}_m$) obtained using the multilocus estimator are given below.

Stand:	$\hat{t}_m$ (S.E.)
Springfield 3	0.91 (0.03)
Springfield 5	.93 (.02)
Combined	.93 (.02)

These estimates are based on 11 loci. The estimates do not differ significantly between the two sites and they are not affected by the combination of the samples from the two

⁵Shaw, D.V., A.L. Kahler, and R.W. Allard. A multilocus method for estimating mating system parameters in plant populations. (Manuscript in preparation.)

Table 1—Single-locus estimates of outcrossing rates, $\hat{t}$ for two stands of Douglas-fir near Springfield, Oregon. Standard errors are in parentheses

Stand	PGM	Locus LAP	EST
Springfield 3	0.91* (.05)	0.90† (.05)	0.83**(.05)
Springfield 5	.93* (.05)	.98 (.05)	.66**(.06)
Combined	.91**(.02)	.79**(.03)	.71**(.03)

†, *, **The likelihood ratio test $H_0: \hat{t} = 1$, is significant at probability levels, 0.06, 0.05 and 0.01, respectively.

sites. These estimates probably reflect the actual proportions of self and outcross progeny which germinated and survived to the time of electrophoretic assay.

Comparisons of single-locus and multilocus estimates of t are presented in *table 2*. Likelihood ratio tests were performed against the null hypothesis; $H_0: \hat{t} = \hat{t}_m$ (Rao 1973). The three loci each demonstrate a unique situation and they will be discussed separately.

1) Pgm-I: None of the three single-locus estimates of outcrossing based on this locus deviated significantly from $\hat{t}_m$. Combining the samples, therefore, did not generate any additional heterogeneity above and beyond the departure from randomness generated by the approximately 9 percent of selfing which occurred in both stands.

2) Lap-I: The single-locus estimate based on the combined samples differed significantly from $\hat{t}_m$. Thus the combining of data for the two stands introduced significant bias for this locus, indicating that allelic frequencies were not the same in the pollen pools of the two stands.

3) Est: All estimates based on this locus differed significantly from $\hat{t}_m$. This indicates that allelic frequencies in the pollen pool were not the same in the two stands and it indicates further that single-locus assumptions, probably including the assumption that the pollen pool is homogeneous within stands, do not hold.

DISCUSSION

A general expectation is that the pollen which falls on a given maternal tree is more likely to come from near neighbors than from distant trees. Also, the genotypes of neighboring trees may be more similar than those of

Table 2—Comparison of multilocus with single-locus estimates of outcrossing rates for two stands of Douglas-fir near Springfield, Oregon

Stand	$\hat{t}_m$	PGM	$\hat{t}$ LAP	EST
Springfield 3	0.91	0.93	0.90	0.83*
Springfield 5	.93	.93	.97	.69*
Combined	.93	.91	.79*	.71*

*Significant differences between $\hat{t}_m$ and $\hat{t}$ by likelihood ratio tests, at the .05 probability level.

random trees in the population due to family structure or microhabitat selection. If such nonrandom distribution of pollen occurs over maternal trees, resulting estimates of selfing, s , made using the mixed mating model, will be larger than the true amount of selfing occurring in the population. In this study we compared estimates of selfing and outcrossing made with single-locus techniques, which are affected by heterogeneity of pollen allelic frequencies, with multilocus estimates made using a multilocus technique which is insensitive to such heterogeneity. With one locus (Pgm-I) we observed no differences in single- and multilocus estimates of selfing. However, with a second locus (Lap-I) the results indicated heterogeneity of pollen gene frequencies between distant sites. With a third locus (Est) heterogeneity was indicated both between distant sites and within small areas. Comparisons of single-locus estimates with multilocus estimates indicated within-stand heterogeneity of pollen gene frequency for one locus (and thus nonrandom pollen dispersal), whereas this effect was not detected at other loci. Because different loci can give different results, it is desirable to include several loci in studies of mating systems.

Microhabitat selection, which causes individuals of similar genotype to be clustered nonrandomly within populations, is one possible cause of the heterogeneity observed in Douglas-fir. Another possible cause is the tendency of seeds to fall to the ground and grow near their maternal parent. The result of this restricted seed dispersal is that relatives, whose genotypes are more similar than those of random members of the populations, become concentrated in family clusters. Such nonrandom distribution of genotypes might lead to heterogeneity of the pollen pool. It can also lead to inbreeding: near neighbors are more likely to be relatives and they are also more likely to mate than are random members of the population.

The demonstration that nearly 10 percent of assayable embryos in Douglas-fir resulted from self-fertilization, and

the indications that additional types of assortative matings may also occur, provide evidence that substantial inbreeding occurs in this species. Douglas-fir, however, shows moderate to severe inbreeding depression which raises the question of the equilibrium level of inbreeding that develops in response to the effects of these opposing forces. Table 3, which gives Est genotypic frequencies of adult maternal individuals and also genotypic frequencies of the progeny of these same individuals, provides information on this point. In the progeny generation homozygotes are in excess and heterozygotes are in deficiency. This is expected when mating is known to be assortative and when seedlings are assayed at a sufficiently early stage such that little selection is likely to have occurred between mating and time of assay. However, genotypic frequencies in the adult parents give a good fit to expectations based on the trinomial square rule. This pattern was repeated for the Lap-I and Pgm-I loci, using combined samples from the two stands (combined samples were used to obtain adequate sample sizes). Thus, even though significant inbreeding occurs in Douglas-fir, leading to excesses of homozygotes over expectations based on the assumption of random mating, it appears that selection removes the excess of inbred individuals. The adult populations fit Hardy-Weinberg expectations but only because the effects of nonrandom mating and selection balance and cancel each other.

In this paper we have illustrated methods of studying the mating system and breeding structure of populations of forest trees. Analysis of preliminary data from an experiment with Douglas-fir indicates that significant numbers of self-fertilized embryos survive through embryonic stages in this species, and that other phenomena, such as family structure and habitat selection, may also contribute to nonrandomness in family arrays derived from open pollination.

Knowledge of the breeding structure in natural stands of forest trees may influence the interpretation of results from open-pollinated genetic tests, and may be useful in planning plus-tree selections and germplasm conservation programs. The methods described appear to provide information useful to the planning and development of tree breeding programs.

Acknowledgments

We are grateful to Dr. Rex McCullough and the Weyerhaeuser Company for providing financial support and for advice concerning the field aspects of the problem. We also thank Dr. M. Thompson Conkle for advice concerning electrophoretic techniques and Michael Carlson for assistance in the collection of seeds.

LITERATURE CITED

Brown, A. H. D., and R. W. Allard.
1970. Estimation of mating systems in open-pollinated maize populations using isozyme polymorphisms. *Genetics* 66:133-145.
Clegg, M. T., A. L. Kahler, and R. W. Allard.
1978. Estimation of life cycle components of selection in an experimental plant population. *Genetics* 89:765-792.

Table 3—Test for goodness of fit for Esterase genotypes to expectations assuming random mating for the parents and the progeny of the combined Springfield samples

Class	Genotype					
	11	22	33	12	13	23
Progeny						
Observed	45	150	68	104	50	90
Expected	29.2	121.7	37.0	117.2	65.7	134
Chi-square ₍₃₎	= 61.24***					
Parents						
Observed	1	14	3	9	8	14
Expected	1.76	13.2	4.1	9.6	5.8	14
Chi-square ₍₃₎	= 1.88 n.s.					

***The chi-square goodness of fit test, with 3 degrees of freedom is significant at the .001 probability level; n.s. indicates nonsignificant deviation.

- Franklin, E. C.
 1970. **Survey of mutant forms and inbreeding depression in species of the family Pinaceae.** USDA Forest Service Res. Paper SE-61, 21 p. Southeastern Forest Exp. Stn., Asheville, N.C.
- Fyfe, J. L., and N. T. J. Bailey.
 1951. **Plant breeding studies in leguminous forage crops. I. Natural cross-breeding in winter beans.** J. Agric. Sci. 41:371-378.
- Nelder, J. A.
 1962. **New kinds of systematic designs for spacing experiments.** Biometrics 18(3):283-307.
- Rao, C. R.
 1973. **Linear Statistical Inference and Its Applications.** (2d ed.) p. 417-418. John Wiley and Sons, New York.
- Rudin, D., G. Eriksson, and M. Rasmunson.
 1977. **Inbreeding in a seed tree stand of *Pinus sylvestris* L., in Northern Sweden. A study by the aid of the isozyme technique.** Research notes, Dep. Forest Genetics, Royal College of Forestry, Sweden.
- Sorensen, F.
 1969. **Embryonic genetic load in coastal Douglas fir, *Pseudotsuga menziesii* var. *menziesii*.** Amer. Nat. 103(932):389-398.
- Sorensen, F.
 1974. **Self-pollination effects on Douglas-fir and ponderosa pine seed and seedlings.** Silvae Genetica 23:135-165.
- Wahlund, S.
 1928. **Zusammensetzung von Populationen und Korrelationserscheinungen von Standpunkt der Vererbungslehre aus befrachtet.** Hereditas 11:65-106.

Isozyme Studies of Forest Insect Populations¹

Molly W. Stock²

Abstract: Data from isozyme analyses are being used to help answer many basic biological questions about forest insect pests and to provide information for a variety of other purposes as well. This paper summarizes the uses of isozymes in quality control of laboratory insect colonies, in studies of insecticide response, as markers of insect parasitoids, and in investigations of species relationships and population differentiation, including patterns of dispersal.

Until recently, genetic studies in forest entomology were handicapped by two rather general assumptions (Wellington 1977). One was that individuals and populations within species are homogeneous units that will behave similarly when treated in a certain way or when subjected to certain environmental conditions. This is an extension of the “if you've seen one, you've seen them all” attitude mentioned earlier in this symposium. If, for example, a particular silvicultural treatment affected an insect population in a certain way in one area, it was often assumed that all other populations of that insect species would respond similarly. A second assumption that hindered initiation of applied biochemical genetics research is a corollary of the first. That is, numerical changes in insect populations (outbreaks, population crashes, and other events) could be explained adequately by understanding external factors affecting the populations.

The life system concept of Clark and others (1967) exemplifies a more realistic approach. According to this scheme, an organism's phenotype and, ultimately, its success in terms of population numbers and persistence, results from the interaction of the genetic makeup of its component individuals with the effective environment.

Since 1974, the importance of intrinsic genetic variation in insect population dynamics has been taken into account in the development of research programs aimed at managing forest insect pests. Data from isozyme analyses are being used to answer basic biological questions and are contributing greatly to our understanding of insects that attack forest trees.

To illustrate the range of isozyme uses in forest entomology, I will summarize some of these research areas.

QUALITY CONTROL OF LABORATORY INSECT COLONIES

Many of the problems associated with conserving genetic resources of forest trees are similar to those dealt with in establishing and maintaining laboratory colonies of insects.

Genetic variation is high in natural populations of insects. This variation is reduced when a limited sample from one area of an insect's range is taken to propagate a laboratory colony. If the number of individuals used to start the colony is small, genetic variation is further reduced in subsequent generations by inbreeding. Colony insects become closely adapted to the laboratory environment and, unless precautions are taken, they may lose the adaptive flexibility characteristic of the wild population. In addition, they may come to differ in important ways from the wild population, making information derived from them of limited value (Bush 1977). After many generations in the laboratory, for example, cultured Lepidoptera, such as budworms and tussock moths, often prefer artificial media to foliage, and groups that diapause (undergo a dormant phase) in the field may become nondiapausing in the laboratory. Although these qualities make the insects easier to rear, their responses in experiments may vary from those of their wild counterparts.

Because responses of colony insects under controlled conditions in the laboratory are often used to estimate behavior of natural populations, it is essential that the degree of overall similarity between the laboratory colony and that of the wild population be known. It is also essential that the colony be monitored over time to detect any further divergence in genetic diversity from wild populations. Quality control—precise, detailed information on differences between laboratory colonies and wild populations, obtained using isozyme analyses—promises to be an integral part of the maintenance of laboratory insect colonies in the future.

¹Presented at the Symposium on Isozymes of North American Forest Trees and Forest Insects, July 27, 1979, Berkeley, Calif.

²Assistant Professor of Forest Resources, University of Idaho, Moscow, Idaho.

GENETICS OF INSECTICIDE RESPONSE

Because they provide a means of identifying insecticide-resistant genotypes, isozyme analyses have been of great value in relating the presence or absence of specific enzymes to insecticide response. Several studies have shown that changes in esterase sensitivity to inhibition by organophosphorus and carbamate insecticides can confer high levels of resistance (Plapp 1976, Georgiou and Pasteur 1978). Marked differences may occur between esterase zymograms of susceptible and resistant insect groups. For example, the organophosphorus insecticide, dimethoate, acted as a selective agent on esterases of the olive fruit fly; insects that were heterozygous at one of two polymorphic esterase-producing gene loci had a better chance of surviving (Tsakas and Krimbas 1970). Pasteur and Sinegre (1975) found a high correlation between certain gene frequencies at an esterase locus in a mosquito and sensitivity to the toxicant, Dursban.

Among forest insects, inter- and intra-specific differences in response to chemical control agents have been observed repeatedly in the field and, more recently, have been demonstrated in controlled laboratory tests. Robertson and others (1978a) compared the response of six budworm species to selected insecticides and found no single species consistently most or least susceptible to these chemicals. It has been common practice, however, when an insecticide has not been tested on a particular species, to predict its response from results of tests on closely related species. Differential response to insecticides has also been shown in laboratory tests on Douglas-fir tussock moth populations (Robertson and others 1978b).

Effective integrated control of forest pests requires that variations within and between species in insecticide response be more fully assessed and considered. Toward this end, isozyme analyses are now being used on these insects. For example, recent work has shown a relationship between gene frequencies at an esterase locus and insecticide resistance in the tussock moth (Stock and Robertson 1979).

Although, like other areas of isozyme application, this work is in its early phases, we may look ahead to a time when levels of insecticide tolerance can be estimated by simple electrophoretic analyses. From a practical standpoint, adoption of routine genetic assays as part of insect population surveys before implementation of control operations could provide estimates of population response to various insecticides. Ultimately, such information could help in the decisionmaking processes involved in integrated pest management.

MARKERS OF INSECT PARASITIDS

Parasitoids—parasites that kill the host—are important in natural control of pest insect populations. Techniques to rapidly and accurately assess levels of parasitism are essential in the development of integrated control

programs. Parasitism levels are presently assessed through host dissection, which is tedious and often inaccurate, or rearing hosts until parasitoids emerge, a procedure that delays estimates of percent parasitism and requires laboratory space, time, and personnel during the field season. Electrophoresis is emerging as a promising practical alternative to these methods. Because host insects and their parasitoids belong to different taxonomic groups, some isozymes are characteristic enough to be readily associated with one or the other species. On this basis, the presence and identification of parasitoids can be determined (May and others 1977a, Wool and others 1978).

SPECIES RELATIONSHIPS

The value of isozyme data in elucidating systematic relationships is well known (Avisé 1974, Ayala 1975). On the basis of electrophoretic differences, for example May and others (1977b) separated the human biting blackfly in Maine from several isomorphic species that do not feed upon humans.

Isozyme information is helping us unravel some taxonomic problems with forest insects, such as the spruce budworm, a species complex described by Freeman in 1967. Most of these recent budworm divisions are based upon traditional morphological comparisons. The large amount of variation present among populations, however, has resulted in considerable uncertainty as to the species status and degree of isolation of many of these groups. In all species, both polymorphic and continuously variable characteristics appear in size and color of larvae, pupae and adults (including wing markings), host plant preferences, cold hardiness, diapause characteristics, and seasonal timing of developmental stages. In addition, varying degrees of reproductive isolation between sympatric species and between allopatric populations within some species occur. To clarify relationships in this group, an integrated approach incorporating pheromone analyses, isozyme analyses, and traditional morphological studies is being used. The electrophoretic method is thus providing a valuable addition to classical approaches to taxonomic studies of forest insect pests.

POPULATION DIFFERENTIATION

Effective control strategies for forest insects are based on a thorough understanding of the population dynamics of each pest. As we learn more about how insect populations are adapted to specific regional host types, we come closer to a management scheme that minimizes disruption of the ecological system of which the pest insect is part. Toward this end, isozyme studies are proving of considerable value. We have begun to answer questions about population differentiation of pest insects as it relates to adaptation to sympatric host plant varieties or species (Bush 1975). The

mountain pine beetle, for example, attacks a diversity of coniferous forest trees over its range, with host species preference varying from place to place. In a mixed stand, the beetles may attack lodgepole pine and apparently ignore other tree species. In a nearby stand, another tree species might be attacked while lodgepole pine goes relatively unfested. These and other differences revealed in the field and laboratory experiments suggest that stratification of sympatric mountain pine beetle populations may be occurring in many areas. Preliminary work in our laboratory suggests that beetles attacking white pine and lodgepole pine in northern Idaho have different characteristic acid phosphatase frequencies. Similar types of differentiation have been observed in mountain pine beetle populations in Colorado.³

Larger differences have been noted between insects in different host tree types when geographic barriers restrict gene flow between them. For example, considerable genetic divergence has occurred between Douglas-fir beetles in Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco var. *menziesii*) in coastal Oregon and Douglas-fir var. *glauca* [Beissn.] Franco in northern Idaho (Stock and others 1979). This work supports evidence of behavioral and physiological differences between the two groups of beetles and underlines the necessity for using caution in extrapolating information obtained on one population to the other when control measures are being planned.

Regional population differentiation is also being examined in the southern pine beetle (Namkoong and others 1979), the forest tent caterpillar,⁴ the eastern spruce budworm,⁵ and several other species of spruce budworm. In the western spruce budworm, we are applying isozyme studies toward clarification of patterns of gene flow among Pacific Northwest populations. It has not been clear, for example whether new budworm outbreaks originate by dispersal from older outbreaks or whether they derive from low-density populations already present in an area. To help answer this question, we have examined genetic characteristics of 20 outbreak populations of western spruce bud worm in Idaho and Montana (Willhite 1979). In this work, one technique that has proven useful is the study of shared rare alleles among the populations (Pashley and Rush 1979). It appears that in some areas of Montana, dispersal plays a more important role in the spread of budworm outbreaks than was previously believed. We have also been able to delineate major patterns of gene flow over the area to help aid prediction of future trends in the spread of budworm outbreaks.

In summary, the union of electrophoretic methodology with applied research in forest entomology is certainly a promising one and we welcome opportunities, such as we have had here today, to exchange ideas and work toward development of productive integrated research programs in forestry.

LITERATURE CITED

- Avise, J. C.
1974. **Systematic value of electrophoretic data.** Syst. Zool. 23:465-481.
- Ayala, F. J.
1975. **Genetic differentiation during the speciation process.** In Evolutionary biology. T. Dobzhansky, M. K. Hecht, and W. C. Steere, eds. Vol. 8, p. 1-78. Plenum Press, New York.
- Bush, G. L.
1975. **Sympatric speciation in phytophagous insects.** In Evolutionary strategies of parasitic insects and mites. P. W. Price, ed. p. 187-206. Plenum Press, New York.
- Bush, G. L.
1977. **The use of gel electrophoresis to monitor genetic variation and maintain quality in mass reared fruit flies.** In Quality control: an idea book for fruit fly workers. E. F. Boller and D. L. Chambers, eds. p. 143-145. **SROP/ WPRS** Bulletin 1977/ 5. Published by the International Organization for Biological Control.
- Clark, L. R., P. W. Geier, R. D. Hughes, and R. F. Morris.
1967. **The ecology of insect populations in theory and practice.** 232 p. Methuen and Co., Ltd., London.
- Freeman, T. N.
1967. **On coniferophagous species of *Choristoneura* (Lepidoptera: Tortricidae) in North America. I. Some new forms of *Choristoneura* allied to *C. fumiferana*.** Can. Entomol. 99:449-455.
- Georghiou, G. P., and N. Pasteur.
1978. **Electrophoretic esterase patterns in insecticide resistant and susceptible mosquitoes.** J. Econ. Entomol. 71:201-205.
- May, B., D. E. Leonard, and R. L. Vadas.
1977a. **Electrophoretic variation, average heterozygosity, and sex-linkage in the spruce budworm.** J. Hered. 68:355-359.
- May, B., L. S. Bauer, R. L. Vadas, and J. Granett.
1977b. **Biochemical genetic variation in the family Simuliidae: electrophoretic identification of the human biter in the isomorphic *Simulium jenningsi* group.** Ann. Entomol. Soc. Amer. 70:637-640.
- Namkoong, G., J. H. Roberds, L. B. Nunnally, and H. A. Thomas.
1979. **Isozyme variations in populations of southern pine beetles.** Forest Sci. 25:197-203.
- Pasteur, N., and G. Sinegre.
1975. **Esterase polymorphism and sensitivity to Dursban organophosphorus insecticide in *Culex pipiens pipiens* populations.** Biochem. Genet. 13:789-803.
- Pashley, D. P., and G. L. Bush.
1979. **The use of allozymes in studying insect movement.** In Movement of highly mobile insects: concepts and methodology in research. R. L. Rabb and G. G. Kennedy, eds. p. 333-341. North Carolina State University, Raleigh.
- Plapp, F. W.
1976. **Biochemical genetics of insecticide resistance.** Ann. Rev. Entomol. 21:179-196.
- Robertson, J. L., N. L. Gillette, B. A. Lucas, N. E. Savin, and R. M. Russell.
1978a. **Comparative toxicity of insecticides to *Choristoneura* species (Lepidoptera: Tortricidae).** Can. Entomol. 110:399-406.
- Robertson, J. L., L. M. Boelter, R. M. Russell, and N. E. Savin.
1978b. **Variation in response to insecticides by Douglas-fir tussock moth, *Orgyia pseudotsugata* (Lepidoptera: Lymantriidae), populations.** Can. Entomol. 110:325-328.

³Personal communication from K. Sturgeon, graduate student, Dep. EPO Biology, University of Colorado, Boulder, Colo. June 1979.

⁴Personal communication from N. Lorimer, Principal Insect Geneticist, North Central Forest Experiment Station, Folwell Avenue, St. Paul, Minn., Sept. 1978.

⁵Personal communication from G. Harvey, Research Scientist, Great Lakes Forest Research Centre, Box 490, Sault Ste. Marie, Ontario, Canada, Nov. 1979.

- Stock, M. W., and J. L. Robertson.
1979. **Differential response of Douglas-fir tussock moth, *Orgyia pseudotsugata* (Lepidoptera: Lymantriidae), populations and sibling groups to acephate and carbaryl: toxicological and genetic analyses.** Can. Entomol. 111:1231-1239.
- Stock, M. W., G. B. Pitman, and J. D. Guenther.
1979. **Genetic differences between Douglas-fir beetles (*Dendroctonus ponderosae*) from Idaho and coastal Oregon.** Ann. Entomol. Soc. Amer. 72:394-397.
- Tsakas, S., and C. B. Krimbas.
1970. **The genetic of *Dacus oleae*. IV. Relation between adult esterase genotypes and survival to organophosphate insecticides.** Evolution 24:807-815.
- Wellington, W. G.
1977. **Returning the insect to insect ecology: some consequences for pest management.** Environ. Entomol. 6:1-8.
- Willhite, E. A.
1979. **Genetics of outbreaking western spruce budworm, *Choristoneura occidentals* Freeman (Lepidoptera: Tortricidae), populations in Idaho and Montana.** Master's thesis, University of Idaho, Moscow, 106 p.
- Wool, D., H. F. van Emden, and S. W. Bunting.
1978. **Electrophoretic detection of the internal parasite, *Aphidus matricariae*, in *Myzus persicae*.** Ann. Appl. Biol. 90:21-26.

Associations Between Heterozygosity and Growth Rate Variables In Three Western Forest Trees¹

Jeffrey B. Mitton,² Peggy Knowles,³ Kareen B. Sturgeon,³ Yan B. Linhart,² and Martha Davis³

Abstract: For each of three species, quaking aspen, ponderosa pine, and lodgepole pine, we determined the relationships between a ranking of heterozygosity of individuals and measures of growth rate. Genetic variation was assayed by starch gel electrophoresis of enzymes. Growth rates were characterized by the mean, standard deviation, logarithm of the variance, and coefficient of variation of annual ring widths measured from cores. In aspen, heterozygosity was associated with high mean growth rate, but in ponderosa and lodgepole pines heterozygosity was not associated with mean growth rate. High levels of heterozygosity were associated with high growth variability in aspen and ponderosa pine, and with low growth variability in lodgepole pine.

Several frequently observed phenomena suggest strong genetic influence on growth characteristics of forest trees. When trees from different origins are grown in common gardens they show marked differences in growth traits (Squillace and Silen 1962, Weidman 1939). These interpopulation differences are thought to be primarily the result of natural selection forming adaptations to local environmental conditions. Differences in growth characteristics are seen also between trees within a single population. Genetic differences within a population arise from segregation and a variety of mating types; some individuals are produced by selfing, some by outcrossing between related individuals, and some by outcrossing between unrelated individuals. These matings produce progeny with relatively low, intermediate, or high levels of heterozygosity, respectively. One generation of selfing, for example, reduces heterozygosity by one-half relative to outcrossing. In selfed progeny, an increase in homozygosity of deleterious recessive alleles may result in a lack of vigor (Fowler 1965, Sorensen and Miles 1974). In outbred progeny, however, increased vigor may result from heterosis because of a summation of heterozygote superiority over many loci. Heterosis for growth and viability is commonly observed in crosses between strains of domesticated plants and animals.

This paper considers the relationships between heterozygosity and growth rate variables within populations of three forest tree species. In this study, we used starch gel electrophoresis to determine heterozygosity of several protein genotypes and estimated mean growth rate from measurements of annual radial growth rings taken from increment cores.

THEORETICAL BACKGROUND

Considerable theoretical and empirical evidence suggest that heterozygosity is associated with fitness and morphological variation.

Heterozygosity Associated With Fitness

Heterosis with respect to growth or viability has been reported extensively in crosses between strains or stocks of domesticated plants and animals (Lerner 1954). Offspring from these crosses have levels of heterozygosity higher than those maintained in either parental stock. Experiments have been designed to detect heterosis in individuals that differ in either heterozygosity of a single pair of chromosomes or a few allozyme loci. For example, chromosomes may be made heterozygous or homozygous in several species of *Drosophila*. When heterozygotes were compared to homozygotes, the homozygotes had fitnesses of only 10 to 20 percent that of the heterozygotes (Sved and Ayala 1970, Tracey and Ayala 1974). In three species of marine pelecypods—the blue mussel, *Mytilus edulis*, the ribbed mussel, *Modiolus demissus*, and the California mussel, *Mytilus californianus*—the proportion, of heterozygotes at one or more loci increased with size (Koehn and others 1973, Koehn and others 1976, Tracey and others 1975). Size is highly correlated with age in these animals, so this relationship probably results from higher viability of heterozygotes. In individuals of the common killifish *Fundulus heteroclitus*, heterozygosity of twelve polymorphic loci was monitored for several years in two localities. Fish could be aged accurately by size, so the same cohort of fish was sampled in successive years. In both localities, level of heterozygosity increased with age, and the data were interpreted as evidence for higher viability in individuals with higher heterozygosity (Mitton and Koehn 1975). In the slender wild oat *Avena barbata*, viability differentials favoring heterozygotes were found at each of three esterase loci (Clegg and Allard 1973). In populations

¹Presented at the Symposium on Isozymes of North American Forest Trees and Forest Insects, July 27, 1979, Berkeley, Calif.

²Associate Professor, Department of Environmental, Population, and Organismic Biology, University of Colorado, Boulder, Colo.

³Graduate Student, Department of Environmental, Population, and Organismic Biology, University of Colorado, Boulder, Colo.

of *Colias* butterflies, *Colias philodice eriphyle* (Watt 1977) and the sagebrush lizard *Sceloporus graciosus* (Tinkle and Selander 1973), heterozygosity increased steadily throughout the life cycle.

Heterozygosity confers advantages for other aspects of fitness. Growth rates recorded for the American oyster *Crassostrea virginica*, were higher for heterozygotes for several loci than for homozygotes (Singh and Zouros 1978). In the tiger salamander *Ambystoma tigrinum*, growth rates in the field and in the laboratory were associated with heterozygosity of a few allozyme loci. Higher growth rates in highly heterozygous individuals could be detected within the progeny of a single pair cross raised in the laboratory.⁴ A study of the demography of the perennial composite *Liatris cylindracea* revealed that highly heterozygous individuals attained greater size and produced more seeds than highly homozygous individuals (Schaal and Levin 1976). Heterozygosity at a single, highly polymorphic locus in the dark-eyed junco, *Junco hyemalis*, was found to be related to both dominance rank and viability; heterozygotes generally enjoyed higher positions in the dominance order than homozygotes, and also survived stress tests better than homozygotes (Baker and Fox 1978).

The papers cited in this section represent careful studies that gave positive results. Equally careful and comprehensive studies have led some investigators to conclude that allozyme variation is essentially unrelated to either fertility or viability (Christiansen and others 1973, Mukai 1977, Yamazaki 1972).

Heterozygosity Associated With Morphological Variation

Associations between heterozygosity and morphological variation are interesting primarily because variation in morphology may be associated with variation in fitness. To date, the majority of studies of associations between heterozygosity and morphological variation have been conducted utilizing data from several or many populations. In studies comparing a series of island populations, for example, the proportion of polymorphic loci and the level of heterozygosity for a population have been related to the degree of morphological variation in those populations. In both the side-blotched lizard *Uta stansburiana* (Soulé and others 1973) and in macaques *Macaca* sp. (Morris and Kerr 1974), populations with more polymorphic loci and higher levels of heterozygosity showed higher levels or greater ranges of morphological variation. Soulé and others (1973) pointed out that these observations suggest allozyme variation is a reasonable indicator for genetic variation throughout the genome. The

studies just cited agree in that populations with greater numbers of genotypes show greater morphological variation.

We are concerned here with heterozygosity, not as it affects variation between populations but, rather, within populations. Lerner (1954) presented a large body of evidence, primarily from domesticated animals and plants, to support his hypothesis that high levels of heterozygosity enhanced developmental homeostasis among individuals within populations. He proposed that highly heterozygous individuals have lower levels of phenotypic variation than do predominantly homozygous individuals. Lewontin (1956) stressed that homeostasis is most evident for traits directly involved in the determination of fitness. An extension of Lerner's hypothesis is that comparisons of individuals with different levels of heterozygosity, but sampled from the same randomly mating populations, have differences in the levels of developmental homeostasis. The effects of heterozygosity upon developmental homeostasis can be measured when only single chromosomes or a few allozyme loci are used to construct groups with low or high heterozygosity. When viability and fertility in *Drosophila* homozygous or heterozygous for specific chromosomes were compared, heterozygous individuals had lower variance in both traits (Dobzhansky and Wallace 1953). Similarly, when Robertson and Reeve (1952) compared *Drosophila* heterozygous or homozygous for specific chromosomes, they found that heterozygotes had lower variance in wing morphology. By using the level of bilateral asymmetry of side-blotched lizards to estimate their degree of developmental homeostasis, Soulé (1979) found that developmental homeostasis increased with heterozygosity as measured by allozyme loci. When populations were divided into homozygous and heterozygous classes on the basis of single allozyme markers, homozygous classes had higher phenotypic variation for both the common killifish *Fundulus heteroclitus* (Mitton 1978) and the monarch butterfly *Danaus plexippus* (Eanes 1978). These data support the extension of Lerner's hypothesis mentioned earlier: developmental homeostasis can vary with levels of heterozygosity within populations, and the effect of heterozygosity may be detected when only single chromosomes or a few allozymes are used to sort individuals into different levels of heterozygosity.

Relationships between heterozygosity and some aspects of fitness have been found primarily in studies of animals. Long-lived forest trees, however, are subjected to much temporal and spatial heterogeneity during their life cycle and recent reviews indicate that forest trees harbor more genetic variation than other groups of plants or animals (Hamrick 1979, Hamrick and others 1979, Hamrick and others 1981).

This paper discusses the results of recent investigations of forest trees to rank individual trees for levels of heterozygosity and to assess how these ranks relate to mean growth rate and variability of growth rate. Examples are taken from three species native to the Colorado Rocky Mountains: quaking aspen *Populus tremuloides* Michx.,

⁴Pierce, B. A., and J. B. Mitton. *Growth and heterozygosity in the tiger salamander, Ambystoma tigrinum*. (Manuscript submitted for publication.)

ponderosa pine *Pinus ponderosa*, Dougl. ex Laws., and lodgepole pine *Pinus contorta* Dougl. ex Loud.

MATERIALS AND METHODS

Mean growth rate and variability of growth rate were estimated from annual radial growth ring measurements obtained from one core extracted from each tree. A binocular stereoscope fitted with an ocular micrometer was used to measure ring widths. For each tree, three measures of variability in growth were calculated: the variance, the standard deviation, and the coefficient of variation.

Vegetative or gametophytic tissue was subjected to starch gel electrophoresis and heterozygosity scores were obtained directly from the banding patterns on gels. The number of loci heterozygous were summed to obtain an individual heterozygosity score for each clone or tree. If four loci were being scored for each individual, for example, the possible heterozygosity score ranged from 0 to 4.

The analytical techniques of multiple regression and analysis of covariance compared growth characteristics among individuals of different heterozygosity levels. These techniques offer the advantage of removing the confounding effects of independent variables that are strongly correlated.

Aspen

In the Western United States, aspen reproduces predominantly asexually by sending up ramets from lateral roots. The unit of study in this species is, therefore, the clone, and cores from the largest ramets within a proximity of about 10 m were used to estimate growth rate variables. Aspen is dioecious, and the sex of each clone was determined in the spring by examining catkins. Age, sex, elevation, and growth rate variables were recorded for each of 106 clones at elevations from 1700 to 3100 m west of Boulder, Colorado. In addition, genotypes at three loci—peroxidase, phosphohexose isomerase, and glutamate dehydrogenase—were obtained from leaf tissue for each clone (Mitton and Grant 1980).

Ponderosa Pine

Pollen tissue from 14 trees on a ridgetop and a south-facing slope at an elevation of 2590 meters near Glacier Lake was analyzed for four loci—peroxidase, aminopeptidase, phosphohexose isomerase, and phosphoglucumutase (Knowles 1978).

Lodge pole Pine

Needle tissue was sampled from 152 lodgepole pines at the University of Colorado's Mountain Research Station,

Table 1—Distributions of loci heterozygous per individual in quaking aspen, ponderosa pine, and lodgepole pine

Sample	Sample size	Loci studied	Heterozygous loci				
			0	1	2	3	4
Quaking aspen	100	3	40	41	15	4	¹ -
Ponderosa pine	117	4	23	57	30	7	0
Lodgepole pine	152	4	32	69	42	9	0

¹ Dash indicates that this number of loci was not considered.

Ward, Colorado, at an elevation of 2800 m. Needle tissue was prepared for electrophoresis (Mitton and others 1979) to obtain genotypes for four loci—peroxidase, fluorescent esterase, phosphohexose isomerase, and alcohol dehydrogenase.

RESULTS

Results of starch gel electrophoresis were summarized for each population sample as a distribution of the number of loci that were heterozygous per individual (table 1).

Aspen

Correlations among variables complicated the search for an association between heterozygosity and growth rate. Growth rate varied with age, sex, and elevation, and distribution of the sexes was not random with respect to elevation. At low elevations, females outnumbered males slightly, and at high elevations, males outnumbered females by 2:1, or more. Growth rate of females was superior to that of males at all elevations, but female growth rate dropped more sharply with increasing elevation (Grant and Mitton 1979). Radial growth decreases regularly with increasing age of a stem (Fritts 1976). Data were analyzed by multiple regression, with mean growth rate of a clone as the dependent variable, and age of the oldest standing ramet, sex, elevation, and heterozygosity for each clone as independent variables (table 2). The multiple R indicated that the suite of independent variables accounts for about 28 percent ($p < 0.001$) of the variance in mean growth rate among the clones. As expected, growth rate decreased, with age and elevation, but was not affected by sex in this analysis. With

Table 2—Multiple regression analysis of standard deviation of annual ring width in quaking aspen

	Coefficients	Probability
Multiple R	0.530	$P < 0.001$
Independent variables		
Age	-0.064	$P < 0.001$
Heterozygosity	.811	$P < 0.05$
Sex	.593	n.s.
Elevation	-.008	$P < 0.01$

the effects of other variables held constant, the relationship between heterozygosity and growth rate was significant. The range of heterozygosity in this sample was 0 to 3 and mean growth rate increased with the number of heterozygous loci ($p < 0.05$).

Within this sample of clones, mean growth rate and variance of growth rate were highly correlated ($r = 0.92$, $p < 0.001$). This suggests that the average growth rate and the variability of the growth rate over time are intimately associated in some way. A multiple regression performed with the same independent variables but with either variance or standard deviation of growth rate as dependent variables gave similar results; increases in the number of heterozygous loci were associated with increases in growth variability.

Ponderosa Pine

Ponderosa pine sampled near Glacier Lake were distributed over a ridgetop and a south-facing slope. Age structures, growth rates, and gene frequencies for some of the four loci differed between these geographically adjacent but ecologically distinguishable sites (Knowles 1978). Mean growth rate and age were negatively correlated ($r = -0.057$, $p < 0.001$). To eliminate the confounding effects of age, slope aspect, and mean growth, an analysis of covariance was employed. A growth rate variable was the dependent variable, a heterozygosity class was the independent variable, and age, slope aspect, and mean growth rate were covariates. The effect of heterozygosity on mean growth, therefore, could be tested with the effects of age and slope held constant; or, the effect of heterozygosity on variance of growth rate could be tested with the effects of age, slope, and mean growth rate held constant. For this analysis, heterozygosity scores were pooled to create two classes—those predominantly homozygous (0 to 1 loci heterozygous) and those predominantly heterozygous (2, 3, or 4 loci heterozygous).

An analysis of covariance, with mean growth rate as the dependent variable, revealed no differences between trees predominantly heterozygous and those predominantly

homozygous (*table 3*). Analyses of covariance were also done with the standard deviation, the logarithm of the variance, and the coefficient of variation used to characterize growth variability (*table 4*). Results of these analyses were all consistent, and were either statistically significant or at the borderline of significance. Although mean growth rates did not differ between heterozygosity classes, variability of growth rates over time were different for heterozygotes and homozygotes. Growth of heterozygotes was more erratic from year to year than was growth of homozygotes.

Lodge pole Pine

A relationship between genotype and growth rate was tested in lodgepole pine. The pines sampled for this analysis were all from a single relatively homogeneous site. Within this sample of 152 trees, age and mean growth were positively correlated ($r = 0.19$, $p < 0.05$); although within any tree radial increments tended to decrease with age. To correct for the effect of age, analyses of covariance were done with mean growth rate or coefficient of variation as the dependent variable, heterozygosity class as the independent variable, and age as the covariate. The mode for the distribution of individual heterozygosity was at 1, and none of the trees in this sample were heterozygous for all four of the loci (*table 1*). Once again, heterozygosity scores were pooled to create two classes, predominantly homozygous (0 to 1 loci heterozygous) and predominantly heterozygous (2 or 3 loci heterozygous). In this analysis, the contribution of the covariate (age) was not significant, so that similar results could be obtained with a single classification analysis of variance, or with a t-test between means. Mean growth rates did not differ between heterozygosity classes but the classes differed in variability of growth rates. The distribution of coefficients of variation of growth rates (*fig. 1*) shows that, although there was considerable overlap in growth variability between heterozygosity classes, the predominantly heterozygous individuals had a mean of 70.3, in contrast to the predominantly homozygous individuals, which had a

Table 3—Comparisons in ponderosa pine and lodgepole pine of mean growth increment between predominantly homozygous and predominantly heterozygous individuals

Locality	Heterozygosity level	Trees	Unadjusted mean	Adjusted mean	F ¹	Probability
Glacier Lake (Ponderosa pine)	Low	78	14.55	15.21	0.15	P > 0.50
	High	36	16.62	15.11		
Ward (Lodgepole pine)	Low	101	5.30	1.10	P > 0.50	
	High	51	5.58			

¹F statistics for ponderosa pine analyses are from analyses of covariance, with age and slope as covariates. Adjusted values incorporate adjustments for covariates. F statistic for lodgepole pine is based on analysis of variance.

Table 4—Comparisons at two localities in ponderosa pine of annual growth increment variabilities between trees of different heterozygosity labels

Locality	Heterozygosity Level	Trees	Standard Dev.		Probability	Log of Variance		Probability	Coef. of Variation		Probability
			unadj.	adj.		unadj.	adj.		unadj.	adj.	
Glacier Lake	Low	78	9.36	9.68	<0.05	1.87	1.88	<0.05	59.33	58.45	0.10>p 0.05
Glacier Lake	High	36	11.18	11.07		1.01	1.01		63.35	65.26	

Note: Significance levels come from analyses of covariance. Age, slope, and mean growth rate are covariates for standard deviation and logarithm of the variance, and slope and age are covariates for the coefficient of variation. Each probability corresponds to the contrast immediately to its left in this table.

mean of 79.2. The differences between the means were statistically significant. Once again, level of heterozygosity was associated with growth variability, but the direction of the relationship found in lodgepole pine was opposite to those found in ponderosa pine and aspen; that is, in lodgepole pine, high heterozygosity was associated with low variability in growth rate.

DISCUSSION

The data presented here from quaking aspen, ponderosa pine, and lodgepole pine are consistent with data from previous studies in that they find a relationship between heterozygosity and some measure of growth rate. Although we do not find an association between mean growth rate and heterozygosity in either ponderosa pine or lodgepole pine, mean growth rate of highly heterozygous clones of aspen are higher than those of predominantly

homozygous clones, just as the mean growth rates of highly heterozygous American oysters (Singh and Zouros 1978) and tiger salamanders⁵ exceed those of more homozygous individuals. These observations may be considered extensions of Lerner's (1954) hypothesis to smaller and smaller portions of the genome. For example, Lerner established that high heterozygosity in the progeny of crosses between domesticated strains or stocks often bestows vigor in the progeny. The same phenomenon, on a much finer genetic scale, is apparent in our results; aspen clones that are heterozygous for a few loci have growth rates that exceed those of homozygotes. Also, variability in ring width in all species is related to level of heterozygosity, and in this respect, these results are similar to those of Mitton (1978), Eanes (1978) and Soulé (1979).

The results of tests for associations between heterozygosity level and growth variability are consistent in the three species studied in that associations were significant in each species. They are inconsistent, however, in their directions. Predominantly heterozygous individuals show higher growth variability in aspen (table

⁵Pierce, B. A., and J. B. Mitton. *Growth and heterozygosity in the tiger salamander, *Ambystoma tigrinum**. (Manuscript submitted for publication.)

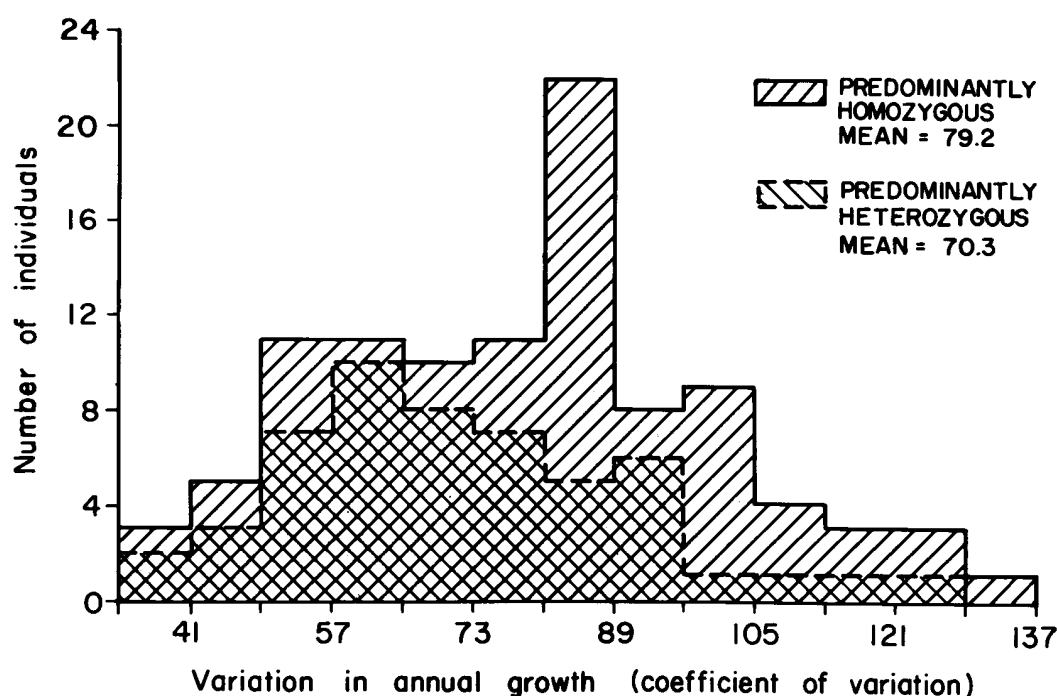


Figure 1—Distribution of the coefficients of variation in annual growth of the predominantly homozygous and predominantly heterozygous lodgepole pine trees ($t = 2.52$, $P > 0.05$).

2) and ponderosa pine (tables 3 and 4), but lower growth variability in lodgepole pine (fig. 1). The sample sizes in each of the studies are substantial, so that we have confidence that the results are real. Although we have no simple, compelling hypothesis to explain the inconsistency, we offer the following possibility. As Lewontin (1956) cautioned, the inverse association between heterozygosity and phenotypic variation would only exist for those phenotypic characters that relate directly to the fitness of the organism. Is it possible that growth characteristics directly contribute to the fitness of lodgepole pine but not to that of ponderosa pine or aspen? Such a situation is doubtful. Rather, we would speculate that growth characteristics are one step removed from a more adaptively significant character, for example, reproductive output. We expect, all else being equal, that annual ring width and seed production are negatively correlated (Fritts 1976). Our measure of growth variability, therefore, may also contain considerable information about variation in seed production. Both quaking aspen and ponderosa pine are known to be highly variable in seed production from year to year (Linhart and others 1979, Linhart and others 1981, Man 1961, Schubert 1974). Years of high seed production are scattered among more common years of low seed production. In contrast, the lodgepole pine in this study has predominantly serotinous cones, and yearly cone production is rather regular. Ponderosa pine and aspen may decrease their growth rates as a result of shunting energy into reproduction during years of heavy seed production. Lodgepole pine, however, may exhibit regular annual growth due to regular reproduction. If high heterozygosity is associated with the extreme of the frequency of reproduction typical for a given species, this could produce the observed association between heterozygosity and growth variability. This conjecture is offered tentatively with the intent of encouraging replication of these studies and initiation of new studies.

Another possible explanation for the discrepancy in direction of relationships between heterozygosity and growth variability is our use of different loci in the different species. If these loci play specific roles in controlling growth rate or growth variability, congruence may occur only when homologous loci are examined in each study.

The mechanisms that produce an association between level of heterozygosity and growth rate variables are of interest both theoretically and practically. Perhaps the first question that must be addressed in consideration of these mechanisms is: When individuals in a population are ranked or placed into groups on the basis of their heterozygosity at a few allozyme loci, just what is being measured? Heterozygosity at a few allozyme loci may estimate heterozygosity (a) of the whole genome, (b) of sections of the chromosomes tightly linked to the allozyme loci, (c) within one or several metabolic pathways, or (d) of only a few allozyme loci. A few loci can be used to reasonably reflect heterozygosity at one or a few dozen loci but it is unlikely that a few loci could accurately rank the members of a population by heterozygosity of the whole

genome.⁶ Perhaps the heterozygosity rankings used here are actually measuring heterozygosity in one of two major metabolic pathways (for example, pentose shunt, or dicarboxylic acid cycle). Heterozygosity levels within a pathway may provide more economic use of energy (Berger 1976), greater degrees of regulation (Johnson 1974), or efficiency for a greater range of temperature or of pH values (Fincham 1972, Gillespie and Langley 1974, Hochachko and Somero 1968). Different levels of heterozygosity within a major pathway could be reflected in growth variability measured from tree cores. Finally, the relationship between heterozygosity and growth rate variables could result solely from the specific loci being monitored. This possibility, however, is incompatible with the hypothesis that protein polymorphisms have little or no effect upon fitness (Kimura and Ohta 1971). A substantial number of kinetic studies of protein polymorphisms now indicate significant differences in *in vitro* performance of different genotypes at a locus (Beckman 1977; Day and others 1974a, 1974b; Merritt 1972; Miller and others 1975; Watt 1979) and these differences may reflect the performance of these genotypes in different environments.

Although the mechanisms for establishing relationships between heterozygosity and growth rate variables are unknown, much is to be gained from this type of study. If these relationships are general and reliable, they can be incorporated into tree improvement programs, because growth rate is likely to affect both quality and quantity of wood being produced. Finally, recent reviews have indicated that forest trees have higher levels of genetic variation than most other groups of animals and plants (Hamrick 1978, Hamrick and others 1979, Hamrick and others 1981). If the associations between heterozygosity and growth rate variables are general, and if differential growth rate and growth variability are adaptive, these associations may help us to understand why so much variation is maintained in forest trees. The influence of allozyme loci upon growth rate could be a fertile field for the joint efforts of geneticists and physiologists.

LITERATURE CITED

- Baker, M. C., and S. F. Fox.
1978. Dominance, survival, and enzyme polymorphism in dark-eyed juncos, *Junco hyemalis*. *Evolution* 32:697-711.
- Beckman, J. S.
1977. Adaptive peroxidase allozyme variation between colonizing and established populations of *Pintos ponderosa* laws on the Shanahan Mesa, Boulder, Colorado. Master's thesis, Univ. Colorado, Boulder, 74 p.
- Berger, E.
1976 Heterosis and the maintenance of enzyme polymorphism. *Amer. Nat.* 110:823-839.

⁶Mitton, J. B., and B. A. Pierce. *The distribution of individual heterozygosity in natural populations*. (Manuscript submitted for publication.)

- Christiansen, F. B., O. Frydenberg, and V. Simonsen.
1973. **Genetics of zoarces populations. IV. Selection component analysis of an esterase polymorphism using population samples including mother-offspring combinations.** *Hereditas* 73:291-304.
- Clegg, M. T., and R. W. Allard.
1973. **Viability versus fecundity selection in the slender wild oat, *Avena barbata*.** *L. Science* 181:667-668.
- Day, T. H., P. C. Hillier, and B. Clarke.
1974a. **Properties of genetically polymorphic isozymes of alcohol dehydrogenase in *Drosophila melanogaster*.** *Biochem. Genet.* 11:141-153.
- Day, T. H., P. C. Hillier, and B. Clarke.
1974b. **The relative quantities and catalytic activities of enzymes produced by alleles at the alcohol dehydrogenase locus in *Drosophila melanogaster*.** *Biochem. Genet.* 11:155-165.
- Dobzhansky, T. H., and B. Wallace.
1953. **The genetics of homeostasis in *Drosophila*.** *Proc. Natl. Acad. Sci.* 39:162-171.
- Eanes, W. F.
1978. **Morphological variance and enzyme heterozygosity in the Monarch butterfly.** *Nature* 276:263-264.
- Fincham, F. R. S.
1972. **Heterozygous advantage as a likely general basis for enzyme polymorphisms.** *Heredity* 28:387-391.
- Fowler, D. P.
1965. **Effects of inbreeding in red pine, *Pinus resinosa* Ait. III. Factors affecting natural selfing.** *Silvae Genetica* 14:37-46.
- Fritts, H. C.
1976. **Tree rings and climate.** p. 567. Academic Press, New York.
- Gillespie, J. H., and C. H. Langley.
1974. **A general model to account for enzyme variation in natural populations.** *Genetics* 76:837-848.
- Grant, M. C., and J. B. Mitton.
1979. **Elevational gradients in adult sex ratios and sexual differentiation in vegetative growth rates of *Populus tremuloides* Michx.** *Evolution* 33:914-918.
- Hamrick, J. L.
1978. **Genetic variation and longevity.** In *Topics in plant population biology.* O. Solbrig, S. Jain, G. Johnson, and P. Raven, eds. p. 84-113. Columbia Univ. Press, New York.
- Hamrick, J. L., Y. B. Linhart, and J. B. Mitton.
1979. **Relationships between life history characteristics and electrophoretically detectable genetic variation in plants.** *Ann. Rev. Ecol. Syst.* 10:173-200.
- Hamrick, J. L., J. B. Mitton, and Y. B. Linhart.
1981. **Levels of genetic variation in trees: influence of life history characteristics.** In *Proceedings of symposium on isozymes of North American forest trees and forest insects, July 27, 1979, Berkeley, Calif. Gen. Tech. Rep. PSW-48, p. 35-41. Pacific Southwest Forest and Range Exp. Stn., Forest Serv., U.S. Dep. Agric., Berkeley, Calif.*
- Hochachka, P. W., and G. N. Somero.
1968. **The adaptation of enzymes to temperature.** *Comp. Biochem. Physiol.* 27:659-668.
- Johnson, G. B.
1974. **Enzyme polymorphisms and metabolism.** *Science* 184:28-37.
- Kimura, M., and T. Ohta.
1971. **Theoretical aspects of population genetics.** 219 p. Princeton University Press, Princeton.
- Knowles, M. H.
1978. **Genetic characteristics associated with growth in ponderosa pine.** Master's thesis, Univ. Colorado, Boulder.
- Koehn, R. K.
1978. **Physiology and biochemistry of enzyme variation: the interface of ecology and population genetics.** In *Ecological genetics: the interface.* P. Brussard, ed. p. 51-72. Springer-Verlag, New York.
- Koehn, R. K., R. Milkman, and J. B. Mitton.
1976. **Population genetics of marine pelecypods. IV. Selection, migration, genetic differentiation in the blue mussel, *Mytilus edulis*.** *Evolution* 30:2-32.
- Koehn, R. K., F. J. Turano, and J. B. Mitton.
1973. **Population genetics of marine pelecypods. II. Genetic differences in microhabitats of *Modiolus demissus*.** *Evolution* 27:100-105.
- Lerner, I. M.
1954. **Genetic homeostasis.** 134 p. Oliver and Boyd, Edinburgh and London.
- Lewontin, R. C.
1956. **Studies on homeostasis and heterozygosity. I. General considerations. Abdominal bristle number in second chromosome homozygotes of *Drosophila melanogaster*.** *Amer. Nat.* 90:237-255.
- Linhart, Y. B., J. B. Mitton, D. M. Bowman, K. B. Sturgeon, and J. L. Hamrick.
1979. **Genetic aspects of fertility differentials in ponderosa pine.** *Genet. Res. Camb.* 33:237-242.
- Linhart, Yan B., Jeffery B. Mitton, Kareen B. Sturgeon, and Martha L. Davis.
1981. **An analysis of genetic architecture in populations of ponderosa pine.** In *Proceedings of symposium on isozymes of North American forest trees and forest insects, July 27, 1979, Berkeley, Calif. Gen. Tech. Rep. PSW-48, p. 53-59. Pacific Southwest Forest and Range Exp. Stn., Forest Serv., U.S. Dep. Agric. Berkeley, Calif.*
- Marr, J.
1961. **Ecosystems of the east slope of the Front Range in Colorado.** Univ. Colo. Studies Series in Biol. 8. Univ. Colorado Press, Boulder.
- Merritt, R. B.
1972. **Geographic distribution and enzymatic properties of lactate dehydrogenase allozymes in the Fathead minnow, *Pimephales promelas*.** *Amer. Nat.* 196:173-184.
- Miller, S. R., W. Pearcy, and E. Berger.
1975. **Polymorphism at the L-glycerophosphate dehydrogenase locus in *Drosophila melanogaster*. I. Properties of adult allozymes.** *Biochem. Genet.* 13:175-188.
- Mitton, J. B.
1978. **Relationship between heterozygosity for enzyme loci and variation of morphological characters in natural populations.** *Nature* 273:661-918.
- Mitton, J. B., and M. C. Grant.
1980. **Observations on the ecology and evolution of quaking aspen, *Populus tremuloides*, in the Colorado Front Range.** *Amer. J. Bot.* 67(2):202-209.
- Mitton, J. B., and R. K. Koehn.
1975. **Genetic organization and adaptive response of allozymes to ecological variables in *Fundulus heteroclitus*.** *Genetics* 79:97-111.
- Mitton, J. B., Y. B. Linhart, J. L. Hamrick, and J. S. Beckman.
1977. **Observations on the genetic structure and mating system of ponderosa pine in the Colorado Front Range.** *Theor. Appl. Genet.* 51:5-13.
- Mitton, J. B., Y. B. Linhart, K. B. Sturgeon, and J. L. Hamrick.
1979. **Allozyme polymorphisms detected in mature needle tissue of ponderosa pine.** *J. Heredity* 70:86-89.
- Morris, L. N., and B. A. Kerr.
1974. **Genetic variability in Macaques: correlations between biochemical and dermatoglyphic systems.** *J. Human Evol.* 3:223-235.
- Mukai, T.
1977. **Lack of experimental evidence supporting selection for the maintenance of isozyme polymorphism in *Drosophila melanogaster*.** In *Proceedings of the second Taniguchi international symposium on biophysics: molecular evolution and polymorphism.* p. 103-126.
- Robertson, F. W., and E. C. R. Reeve.
1952. **Heterozygosity, environmental variation and heterosis.** *Nature* 179:286-287.
- Schaal, B. A., and D. A. Levin.
1976. **The demographic genetics of *Liatris cylindracea* Michx.** *Amer. Nat.* 110:191-206.
- Schubert, G. H.
1974. **Silviculture of southwestern ponderosa pine: the status of our knowledge.** USDA Forest Serv. Res. Paper RM-123. 71 p. Rocky Mountain Forest and Range Exp. Stn., Fort Collins, Colo.

- Singh, S. M., and E. Zouros.
1978. **Genetic variation associated with growth rate in the American oyster (*Crassostrea virginica*)**. *Evolution* 32:342-345.
- Sorensen, F. C., and R. S. Miles.
1974. **Self-pollination effects of Douglas-fir and ponderosa pine seeds and seedlings**. *Silvae Genetica* 23:135-138.
- Soulé, M. E., S. Y. Yang, M. C. W. Weiler, and G. C. Gorman.
1973. **Island lizards: the genetic-phenetic variation correlation**. *Nature* 242:191-193.
- Soulé, M.
1979. **Heterozygosity and developmental stability: another look**. *Evolution* 33:396-401.
- Squillace, A. E., and R. R. Silen.
1962. **Racial variation in ponderosa pine**. *Forest Sci. Monogr.* 2:1-27.
- Sved, J. A., and F. J. Ayala.
1970. **A population cage test for heterosis in *Drosophila pseudoobscura***. *Genetics* 66:97-113.
- Tinkle, D. W., and R. K. Selander.
1973. **Age dependent allozyme variation in a natural population of lizards**. *Biochem. Genet.* 8:231-237.
- Tracey, M. L., and F. J. Ayala.
1974. **Genetic load in natural populations: Is it compatible with the hypothesis that many polymorphisms are maintained by natural selection?** *Genetics* 77:569-589.
- Tracey, M. L., N. F. Bellett, and C. B. Graven.
1975. **Excess of allozyme homozygosity and breeding population structure in the mussel *Mytilus californianus***. *Mar. Biol.* 32:303-311.
- Watt, W. B.
1979. **Adaptation of specific loci. I. Natural selection on phosphoglucose isomerase of *Colias* butterflies: biochemical and population aspects**. *Genetics* 87:177-194.
- Weidman, R. H.
1939. **Evidences of racial influence in a 25-year test of ponderosa pine**. *J. Agric. Res.* 59:855-887.
- Yamazaki, T.
1972. **Detection of single gene effect by in-breeding**. *Nature: New Biol.* 240:53-54.

Levels of Genetic Variation in Trees: Influence of life history characteristics¹

J. L. Hamrick,² J. B. Mitton,³ and Y.B. Linhart³

Abstract: In a previous study, levels of genetic variation, as measured by isozyme analyses, were compared for 113 taxa of vascular plants. Each species was classified for 12 life history and ecological traits and three measures of genetic variation were calculated. Plants with large ranges, high fecundities, an outcrossing mode of reproduction, wind pollination, a long generation time, and from habitats representing later stages of succession tended to have more isozyme variation than species with other combinations of characteristics. This paper discusses the results of the previous study and examines the available isozyme data for similar trends in forest trees. Special consideration was given to differences in genetic variation among 20 conifer species that have many of their life history characteristics in common. Successional stage, habitat type, cone type and historical events were associated with differences in genetic variation among the conifer species. These results are discussed in terms of expectations from current population genetics theory.

Forest trees have been the subject of many quantitative genetic investigations (Libby and others 1969, Stern and Roche 1974). These studies have concentrated on morphometric and physiological characteristics such as survival, growth initiation, height and diameter growth, hardiness to environmental stress, and various leaf, stem, fruit, and wood characteristics. As a result, the distribution of quantitative genetic variation is better understood in certain tree species than in most other naturally occurring plants.

Studies that use biochemical techniques to measure genetic variation in forest trees generally have lagged behind those that use quantitative traits. A few workers have used secondary plant compounds, such as terpenes, in studies of species (Zavarin and Snajberk 1969, Zavarin and others 1969), racial (Smith and others 1969) and population (Adams 1975a, 1975b) differentiation. But, because the genetic basis of variation in these compounds is poorly understood, they are of limited use in population genetic studies. Since the early 1970's, electrophoretic techniques have been used in genetic studies of forest tree populations. These techniques offer a number of advantages over other biochemical or quantitative approaches: (a) genetic inheritance of electrophoretically-detectable traits can be easily demonstrated; (b) most isozyme loci are codominant and gene frequencies can be calculated without the necessity of genetic crosses; (c) estimates of genetic variation can be compared directly between populations or between species. The relatively few isozyme studies of plant populations have been the subject of a number of recent reviews (Gottlieb 1977, Brown 1979,

Hamrick 1979, Hamrick and others 1979). Generally, these reviews have concluded that plants contain somewhat more variation than invertebrate animals and considerably more variation than most vertebrates. Furthermore, different plant species contain varying amounts of genetic variation. Trees, for example, have been found to contain significantly more variation than herbaceous plants (Hamrick 1979, Hamrick and others 1979).

This paper reviews the results of a previous study (Hamrick and others 1979), which compared the isozyme variation of plant species with different combinations of life history traits and examines the available isozyme data for similar trends in forest trees. Specifically, we address the following questions:

- Do plant species with certain combinations of life history traits contain higher levels of genetic variation?
- Does being a woody plant directly affect levels of isozyme variation?
- Can differences in genetic variation among tree species be explained by their life history and ecological characteristics?

PROCEDURES

In a previous study (Hamrick and others 1979), we examined the relationship between 12 life history and ecological traits and the levels of genetic variation maintained within populations of 113 taxa of plants. Each species was classified for taxonomic status, geographic range, generation length, mode of reproduction, mating system, pollination mechanism, fecundity, seed dispersal mechanism, chromosome number, successional stage, habitat type, and cultivation status. Where the data allowed, three measures of intrapopulation genetic variation were calculated for each species: the percent of polymorphic loci per population (*P*), the mean number of alleles per locus (*A*), and a polymorphic index (*PI*). The *PI*

¹Presented at the Symposium on Isozymes of North American Forest Trees and Forest Insects, July 27, 1979, Berkeley, Calif.

²Associate Professor of Biology, University of Kansas, Lawrence, Kans.

³Associate Professor of Biology, University of Colorado, Boulder, Colo.

Table 1—Levels of variability among categories of 12 life history and ecological traits. Weighted means and standard errors are given for each measure of variability. Differences in PI between categories are tested by ANOVA. Statistical significance levels are given in parentheses (from Hamrick and others 1979)

Variable	Species	Loci	Polymorphic loci (<i>P</i>)		Alleles per locus (<i>A</i>)		Polymorphic Index (<i>PI</i>)	
			$\bar{x}$	S.E.	$\bar{x}$	S.E.	$\bar{x}$	S.E.
			<i>Percent</i>					
Taxonomic status							<i>(P</i> < 0.01)	
Gymnospermae	11	9.2	67.01	7.99	2.12	0.20	0.270	0.041
Dicotyledoneae	74	11.4	31.28	3.31	1.46	.06	.113	.014
Monocotyledoneae	28	11.6	39.70	6.02	2.11	.19	.165	.026
Geographic range							<i>(P</i> < 0.05)	
Endemic	17	15.1	23.52	5.06	1.43	0.11	0.086	0.019
Narrow	22	11.4	36.73	6.01	1.60	.14	.158	.030
Regional	39	8.3	55.96	5.13	1.85	.10	.185	.025
Widespread	35	12.5	30.36	5.03	1.58	.15	.120	.021
Generation length							<i>(P</i> < 0.001)	
Annual	42	11.2	39.47	4.32	1.72	0.11	0.132	0.017
Biennial	13	17.2	15.78	5.12	1.26	.09	.060	.020
Short-lived perennial	31	12.0	28.09	5.06	1.46	.09	.123	.023
Long-lived perennial	27	7.6	65.77	5.08	2.07	0.13	0.267	0.027
Mode of reproduction							N.S.	
Asexual	1	8.0	50.00	0.00	1.91	0.00	0.139	0.000
Sexual	95	11.7	35.64	3.03	1.63	.07	.135	.012
Both	17	8.9	41.71	8.12	1.67	.14	.185	.034
Mating system							<i>(P</i> < 0.01)	
Primarily selfed	33	14.2	17.92	3.21	1.27	0.06	0.058	0.014
Mixed	42	8.6	14.16	4.89	1.76	.10	.181	.022
Primarily outcrossed	36	11.3	51.07	4.95	1.85	.12	.185	.022
Pollination mechanism							<i>(P</i> < 0.001)	
Selfed	33	14.2	18.99	3.51	1.31	0.07	0.058	0.028
Animal	55	9.5	38.83	3.94	1.55	.07	.130	.015
Wind	23	10.7	57.45	6.29	2.27	.17	.264	.028
Fecundity							<i>(P</i> < 0.001)	
<10 ²	21	12.0	40.06	6.45	1.72	0.16	0.127	0.026
10 ² to 10 ³	27	12.4	26.35	3.46	1.44	.07	.096	.013
10 ³ to 10 ⁴	22	11.9	36.98	6.37	1.64	.10	.199	.034
>10 ⁴	40	9.8	67.99	5.99	2.27	.17	.286	.033
Seed dispersal mechanism							N.S.	
Large	27	11.4	37.42	3.73	1.76	0.156	0.0230	.16
Animal-attached	16	11.1	28.79	5.55	1.55	.08	.092	.020
Small	26	12.4	32.98	5.10	1.51	.09	.118	.018
Winged or plumose	21	12.2	44.91	7.27	1.86	.13	.188	.029
Animal ingested	20	7.0	32.98	8.25	1.43	.10	.132	.036
Chromosome number							<i>(P</i> < 0.01)	
10 to 20	50	13.1	35.52	3.61	1.55	0.08	0.111	0.014
22 to 30	44	10.0	37.41	5.35	1.73	.10	.175	.023
>30	16	8.9	41.65	7.20	2.10	.14	.224	.030
Stage of succession							<i>(P</i> < 0.01)	
Weedy and early	54	12.5	29.67	3.82	1.60	0.08	0.116	0.015
Middle	49	9.7	37.90	4.43	1.56	.08	.137	.019
Late	10	12.0	62.76	5.28	2.14	.19	.271	.038
Habitat type							N.S.	
Xeric	4	8.8	15.39	8.20	1.11	0.09	0.048	0.040
Submesic	19	10.5	43.68	4.86	1.66	.08	.140	.020
Mesic	82	11.4	36.01	3.61	1.65	.07	.146	.016
Hydric	8	13.0	27.71	10.33	1.59	.22	.145	.050
Cultivation status							N.S.	
Cultivated	21	6.8	38.99	7.15	1.61	0.12	0.172	0.032
Noncultivated	89	12.6	36.10	3.12	1.63	.07	.136	.013
Both	3	2.7	50.00	—	1.75	.21	.209	.035

is equivalent to the Hardy-Weinberg heterozygosity (Hamrick and Allard 1972).

For each category of the 12 life history variables weighted (by the number of loci), means were calculated for P , A , and PI , and mean PI values were compared for heterogeneity by single classification ANOVA. Multivariate statistics were used to determine correlations among traits and to test whether combinations of life history variables influence genetic variation. The data were analyzed first with a principal components analysis to describe the major patterns of covariation. Associations noted in this analysis were explored further with a stepwise multiple regression.

Differences in genetic variation among tree species was the subject of a second analysis. This analysis included studies not available earlier but was limited to studies that used a wide variety of enzyme systems. Values of P , A , and PI were calculated for each species and weighted means were calculated for each category of six variables. Multivariate analyses were not used since most of the tree species were conifers and, therefore, similar for many of their life history traits.

RESULTS

Isozyme Variation in Plants

The mean values of P , A , and PI obtained by pooling the 113 taxa in our original study were: P = 36.8 percent, A = 1.69, and PI = 0.141. When compared with animals, plants tend to have levels of genetic variation that are roughly equivalent to the invertebrates (P = 46.9 percent, PI = 0.135 [Selander 1976]; P = 39.7 percent, PI = 0.112 [Nevo 1978]) but are considerably higher than those of vertebrate species (P = 24.7 percent, PI = 0.061 [Selander 1976], P = 17.3 percent, PI = 0.036 [Nevo 1978]).

Genetic Variation and Life History Traits

Statistically significant differences ($P < 0.05$) were found among categories of eight life history or ecological traits (table 1). Gymnosperm species tended to have more variation than angiosperms while regionally distributed species had more variation than those with other geographic ranges. We had expected species with the widest ranges to contain the most variation. Many species in this category, however, are weedy or early successional species that tend to have less variation. Long-lived perennials and plants with mixed or primarily outcrossed mating systems contained at least twice the variation of species in other categories. Within the outcrossed species, wind-pollinated species had more variation than those with animal pollination. Species with high fecundities and high chromosome numbers also had more variation as did species of the later stages of succession.

Multivariate techniques were used to identify groups of variables varying in concert, and to assess the covariation of genetic variation with ecological and life history variables. Approximately one-half of the correlations among the ecological, life history, and genetic variables were statistically significant. Pollination mechanism, mating system, and fecundity had the highest correlations with the genetic variables. Additional multivariate analyses demonstrated that only the first two principal components contributed to our understanding of the genetic variation among these species. The first principal component, which had high loadings from the three genetic variables, generation length, mating system, pollination mechanism, fecundity, seed dispersal and successional status explained 30 percent of the variation. The second principal component, which had high loadings from the genetic variables, taxonomic status, seed dispersal mechanisms, and successional stage explained 16 percent of the variation. These results were consistent with the univariate analyses and indicated that species with large ranges, high fecundities, an outcrossing mode of reproduction, wind pollination, a long generation time, and from habitats representing later stages of succession had more genetic variation than did species with other combinations of traits. It is worth noting that forest trees in general and conifers in particular combine many of the characteristics that are associated with high levels of genetic variation.

To determine which ecological and life history traits were most closely associated with genetic variation we employed a stepwise multiple regression analysis in which PI was the dependent variable and the 12 ecological and life history traits were independent variables. Only two variables were significant—pollination mechanism and fecundity. Species that are wind pollinated and highly fecund were the most genetically variable.

Variation Among Tree Species

The separate analyses of the tree data produced some noteworthy results. First, the mean level of variation within populations of coniferous trees was lower than that reported previously (Hamrick and others 1979). In the earlier review (Hamrick and others 1979), the inclusion of studies based solely on polymorphic loci tended to increase estimates of the mean levels of variation. Also, many of the earlier tree studies used enzyme systems such as esterases and peroxidases that are often highly variable. The data in the present analyses are relatively free of these biases and should be more representative of the actual levels of genetic variation in tree populations. It is significant, therefore, that the mean levels of variation in tree populations continue to exceed that of herbaceous species by more than 60 percent.

A second important result is the continued observation of interspecific differences in genetic variation. The PI varies from 0.000 for red pine (*Pinus resinosa* Ait.) to 0.364

Table 2—Summary of electrophoretic studies of genetic variation in trees. *P* = percent of polymorphic loci, *A* = alleles per locus, *PI* = polymorphic index

Species ¹	Location	Loci	Populations	<i>P</i>	<i>A</i>	<i>PI</i>	Source
A. Gymnosperms							
<i>Abies balsamea</i>	New Hampshire	14	1	64.0	1.86	0.150	Neale (1978)
<i>Picea abies</i>	Sweden	12	4	91.7	3.54	.341	Lundkvist (1979)
<i>P. sitchensis</i>	Rangewide	24	10	—	1.90	.150	Yeh, F. (unpubl.)
<i>Pinus aristata</i>	Rangewide	22	5	46.4	1.55	.139	Hiebert, R.D. and Hamrick, J.L. (unpubl.)
<i>P. attenuata</i>	Rangewide	22	10	73.0	2.09	.140	Conkle, M.T. (unpubl.)
<i>P. balfouriana</i>	Rangewide	23	4	57.6	1.61	.208	Hiebert, R.D. and Hamrick, J.L. (Unpubl.)
<i>P. banksiana</i>	Michigan	21	1	28.6	—	.083	Snyder, T. (unpubl.)
<i>P. contorta</i>	British Columbia	27	17	—	1.90	.150	Yeh, F. (unpubl.)
<i>P. contorta</i>	California	37	1	89.2	2.78	.190	Conkle, M.T. (unpubl.)
<i>P. contorta</i>	Colorado	26	1	42.0	1.42	.160	Hamrick, J.L. (unpubl.)
<i>P. lambertiana</i>	Rangewide	20	—	80.0	2.85	.260	Conkle, M.T. (unpubl.)
<i>P. longaeva</i>	Rangewide	14	5	78.6	2.35	.364	Hiebert (1977)
<i>P. jeffreyi</i>	California	44	1	—	2.90	.260	Conkle, M.T. (unpubl.)
<i>P. muricata</i>	Northern California	17	1	65.0	1.53	.090	Millar, C. (unpubl.)
<i>P. ponderosa</i>	Eastern Colorado	22	7	68.4	2.00	.226	Hamrick, J.L. (unpubl.)
<i>P. pungens</i>	Rangewide	15	3	40.0	1.33	.144	Feret (1974)
<i>P. resinosa</i>	Rangewide	9	5	0.0	1.00	.000	Fowler and Morris (1977)
<i>P. rigida</i>	Rangewide	21	11	78.8	2.19	.144	Guries, R. and Ledig, T. (unpubl.)
<i>P. strobus</i>	Seed Orchard,	17	—	52.9	2.06	.330	Eckert and others (1980)
	New Hampshire						
<i>P. taeda</i>	North Carolina	11	1	100.0	3.73	.340	Conkle, M.T. (unpubl.)
<i>P. taeda</i>	Superior Trees	30	—	93.3	3.87	.260	Conkle, M.T. (unpubl.)
<i>P. taeda</i>	Seed Orchard,	15	—	80.0	2.93	.200	Adams and Joly ²
	South Carolina						
<i>Pseudotsuga menziesii</i>	British Columbia,	21	11	—	2.23	.180	Yeh, F. (unpubl.)
	Interior						
<i>P. menziesii</i>	Coastal	21	11	—	2.23	.150	Yeh, F. (unpubl.)
<i>P. menziesii</i>	California Coastal	11	9	74.2	3.17	.332	Morris, R. (unpubl.)
<i>P. menziesii</i>	California Interior	17	1	100.0	1.78	.330	Conkle, M.T. (unpubl.)
<i>P. menziesii</i>	Eastern Colorado	22	5	64.0	1.86	.264	Hamrick, J.L. (unpubl.)
<i>Sequoiadendron giganteum</i>	Rangewide	8	34	50.0	2.63	.155	Fins, L. (unpubl.)
	Mean	20.1		67.7 ±4.9	2.29 ±0.14	.207 ±.017	
B. Angiosperms							
<i>Persea americana</i>		10		80.0	1.90	.195	Torres and others (1978)
cultivars							
<i>Phoenix dactylifera</i>		7		100.0	2.00	.332	Torres, A.M. (unpubl.)
cultivars							
	Mean	8.5		88.23 ±10.0	1.94 ±0.05	.251 ±.028	

¹Only those studies that used a variety of enzyme systems are included.

²Adams, W.T., and R.J. Joly. Allozyme studies in loblolly pine seed orchards: clonal variation and frequency of progeny due to self-fertilization (Manuscript in preparation.)

for Great Basin bristlecone pine (*P. longaeva* D.K. Bailey). Since the 20 conifer species have many of their life history characteristics in common, only six traits were available to explain the differences in genetic variation observed among species—taxonomic status, geographic range, stage of succession, habitat type, cone type, and U.S. distribution.

The present data (table 2) are inadequate to determine whether differences between angiosperm and gymnosperm trees were significant. The two studies of angiosperm trees that met our criteria involved commercial cultivars of fruit trees. Although angiosperm trees appear to have as much

genetic variation as conifers, final conclusions must await studies of naturally occurring angiosperm trees.

The geographic range of each conifer was classified as endemic, narrow, or regional (table 3). Most species had a regional distribution with only foxtail pine (*Pinus balfouriana* Grev. & Balf.) and giant sequoia (*Sequoiadendron giganteum* [Lindl.] Buchholz) classified as endemics. The differences in *PI* between these categories were small and were not significant ($P < 0.50$).

The ability of a species to successfully reproduce in its own shade and in the absence of environmental disturbance was used to classify the species into one of

Table 3—Levels of variability among categories of five life history and ecological variables for 28 conifer studies. Weighted means and standard errors are given. Differences in PI between categories are tested by ANOVA. Significance levels are indicated in parentheses

Variable	Species ¹	Studies	Mean loci	Polymorphic Index (PI)	
				Mean	S.E.
Geographic range				(P < 0.50)	
Endemic	2	2	15.5	0.194	0.052
Narrow	8	8	22.2	.197	.032
Regional	10	18	19.7	.207	.023
Stage of succession				(P < 0.10)	
Early	7	10	19.9	0.161	0.030
Middle	7	11	19.8	.241	.022
Late	7	7	20.8	.204	.035
Habitat type				(P < 0.10)	
Xeric	3	3	19.3	0.194	0.075
Submesic	11	15	21.7	.186	.021
Mesic	6	10	17.9	.238	.026
Cone type				(P < 0.01)	
Closed cone	6	7	21.3	0.132	0.012
Open cone	15	21	20.4	.226	.020
United States distribution				(P < 0.30)	
Northeastern	5	5	16.4	0.152	0.054
Southern	2	4	17.8	.235	.042
Western	12	18	22.1	.204	.018

¹ Some species may be represented in more than one category.

three stages of succession (table 3). More variation ($P < 0.10$) was observed in populations of middle and late successional species and less in populations of pioneering species. Only the presence of loblolly pine (*P. taeda* L.) in the early category prevented the results from being statistically significant.

Differences among the habitat categories approached statistical significance ($P < 0.10$), but no definite trend can be seen (table 3). Differences between the submesic and mesic categories suggest a pattern but with only two species in the xeric category no definite conclusions can be drawn. Studies of genetic variation in conifer species that are adapted to drought conditions, such as the piñon pines and junipers, should yield valuable information on this question.

Species with closed cones or ecotypes of species that have closed cones were grouped together. These species had significantly less variation ($P < 0.01$) than open-coned species (table 3).

Species which inhabit different geographic regions within the United States may also have different levels of genetic variation. The northeastern species (*Abies balsamea* [L.] Mill., *Pinus banksiana* Lamb., *P. resinosa* Ait., *P. rigida* Mill., and *P. strobus* L.) have somewhat less variation ($P < 0.30$) while the southern (*Pinus taeda* L. and *P. pungens* Lamb.) and the western species have higher values.

To summarize the results of the analyses on trees, species of later successional stages, mesic habitat types, with open

cones and a southern or western distribution have more genetic variation than species with alternate combinations of characteristics.

DISCUSSION

A significant proportion of the differences in genetic variation among plant species is accounted for by variation in their life history and ecological characteristics. Species with large ranges, high fecundities, an outcrossing mode of reproduction, wind pollination, a long generation time and from habitats representing later stages of succession have higher amounts of genetic variation. This result is generally consistent with that predicted by population genetics theory. Plants, in general, might be expected to maintain high levels of genetic variation within their populations since their sessile nature often leads to the evolution of locally adapted ecotypes (Antonovics 1971, Bradshaw 1972, Jain and Bradshaw 1966). Also, long-lived plant species with large ranges and high fecundities typically have large, stable populations. Such populations are resistant to chance fluctuations in gene or genotype frequencies and should maintain more variation than populations that experience large fluctuations in size. Longevity also ensures the representation of many cohorts within a population. If different alleles or genotypes are favored during the establishment phase of each cohort, individuals that survive to maturity will maintain a genetic

"record" of these evolutionary events. Their continued survival would retard the loss of genetic variation. Greater longevity could be especially effective in maintaining genetic variation in later stages of succession because individuals are continuously becoming established in a highly complex biotic environment. It should be noted, however, that the seed carryover abilities of many annuals and short-lived perennials could produce a similar genetic record. Finally, high rates of fecundity, outcrossing and wind pollination ensure large neighborhood sizes and the production of a variety of genotypes through recombination. Natural selection could act to maintain this variation through the evolution of locally adapted ecotypes or through various types of balancing selection.

The existence of high levels of genetic variation in forest tree populations can be explained primarily by their life history and ecological characteristics. The question remains, however, of whether woodiness has a direct effect on the levels of genetic variation that are maintained. Current evidence, although scanty, indicates that herbaceous species with life history or ecological traits similar to those of trees also have high levels of genetic variation. This would argue against woodiness itself as affecting the maintenance of genetic variation. An alternate argument can be made on the basis of differences in life forms (Raunkiaer 1934) between woody and herbaceous plants (Hamrick 1979). Woody plants are phanerophytes and maintain their apical meristems above ground. Their apical meristems, therefore, are exposed to environmental fluctuations throughout the year. Herbaceous perennials are either chamaephytes, hemicryptophytes or cryptophytes whose apical meristems are located at or below the soil surface during periods of severe environmental stress. The moderating effects of snow, soil, or litter cover may, in essence, reduce selection pressures for the internal buffering thought to be produced by increased heterozygosity (Lerner 1954). If this argument is valid we would expect to find differences in genetic variation between populations of herbaceous and woody plants that occur in the same habitats and have similar life history traits. Also, we would expect to find an increase in heterozygosity in older age classes and in temporally fluctuating environments. An increase in heterozygosity in older age classes was demonstrated in a population of *Liatris cylindracea* Michx. (Schaal and Levin 1976). A further test of this argument would be provided by surveying a wider variety of woody trees and shrubs.

Differences in genetic variation among the conifer species surveyed are somewhat more difficult to explain. This is because the 20 conifer species have many of their life history characteristics in common. Although only one of the six traits — cone type — was statistically significant, geographic range was the only trait that failed to have large differences among the mean *PI* values. This result is not surprising since endemic and narrowly distributed tree species often have rather large population sizes where they occur. Also both of the endemic species, *Pinus balfouriana* Grew. & Balf. and *Sequoiadendron giganteum* (Lindl.)

Buchholz, are long-lived and could maintain genetic variation by this mechanism.

Trends observed for stage of succession, habitat type, and cone type in the conifer data are generally consistent with previous results (Hamrick and others 1979); conifers from mid- and late-successional stages, mesic habitats, and with open cones tend to have higher levels of genetic variation. Trees of early successional stages must adapt to relatively homogeneous environmental conditions that are influenced primarily by physical factors. Coupled with their colonizing habit, such uniform selection pressures might lead to a reduction in genetic variation. As succession proceeds, the biotic environment becomes increasingly important and, as a result, habitats become more complex and heterogeneous. Such complex environments may select for the maintenance of higher levels of genetic variation. A similar explanation may apply to the higher levels of variation observed in the more mesic adapted trees.

The lower levels of variation found in the closed-cone pines could result from a combination of factors. First, the closed-cone species are adapted to habitats which experience periodic catastrophic wildfires and are, therefore, almost always members of early successional communities. Such species germinate and grow under relatively uniform environmental conditions. Also, the closed cone habit and short life expectancy ensures even-aged stands. The genetic record of the long-lived, uneven aged species is not, therefore, an important factor in the maintenance of genetic variation in closed-cone species.

The geographic region to which conifer species are native also seems to affect genetic variation. Four of the five northeastern conifers maintain less variation than their southern and western counterparts. Two factors may help to explain this observation. First, it could be argued that western habitats are environmentally more heterogeneous than those in the northeast. Support for this argument comes from the greater diversity of community types in the West (Küchler 1964). The high variation in the two southern species contradicts this argument, however. A second factor to consider is the evolutionary history of each species. The northeastern species have been exposed to numerous continental glacial events. Glaciation may have greatly reduced the ranges of the northeastern species and reduced population sizes and genetic variation. Although the western species have also been exposed to changes in climatic factors and shifts in their ranges, the more varied topography of the West may have provided a greater variety of refugia. In fact, some of the western species expanded their ranges during glacial periods (for example, *Pinus longaeva* D. K. Bailey⁴).

In conclusion, a thorough knowledge of a species' life history and ecological characteristics is essential if its genetic structure is to be understood. Much of the heterogeneity in genetic variation seen among species can

⁴Personal communication from P.V. Wells,

be explained by such considerations. A significant amount of unexplained heterogeneity remains, however. Past historical events and characteristics which have not been considered may partially explain the differences that remain. As additional plant species are studied we should gain a clearer understanding of the role that life history characteristics, or combinations of characteristics, play in shaping the genetic structure of plant populations, including those of forest trees.

Acknowledgments:

We wish to acknowledge the technical assistance provided by Diane Bowman, Kareen Sturgeon, and Karen Hamrick throughout the preparation of this paper. We especially wish to thank those workers who have shared their unpublished data with us: Thomas Adams, M. Thompson Conkle, Robert Eckert, Lauren Fins, Ray Guries, F. Thomas Ledig, Constance Millar, Richard Morris, Thomas Snyder, Andrew Torres, and Francis Yeh. Portions of this research were supported by NSF Grants DEB 76-01295 and BMS 75-14050.

LITERATURE CITED

- Adams, R.P.
1975a. **Gene flow versus selection pressure and ancestral differentiation in the composition of species: analysis of populational variation of *Juniperus ashei* Buch. using terpenoid data.** *J. Mol. Evol.* 5:177-185.
- Adams, R.P.
1975b. **Numerical-chemosystematic studies of infraspecific variation in *Juniperus pinchotii*.** *Biochem. Syst. Ecol.* 3:71-74.
- Antonovics, J.
1971. **The effects of a heterogeneous environment on the genetics of natural populations.** *Amer. Sci.* 59:593-599.
- Bradshaw, A.D.
1972. **Some of the evolutionary consequences of being a plant.** *Evol. Biol.* 5:25-46.
- Brown, A. H. D.
1979. **Enzyme polymorphism in plant populations.** *Theor. Popul. Biol.* 15:1-42.
- Eckert, R.T., R.J. Joly, and D.B. Neale.
1980. **Genetics of isozyme variants and linkage relationships among allozyme loci in eastern white pine (*Pinus strobus*).** *Can. J. For. Res.* [In press.]
- Feret, P.P.
1974. **Genetic differences among three small stands of *Pinus pungens*.** *Theor. Appl. Genet.* 44:173-177.
- Fowler, D.P., and R.W. Morris.
1977. **Genetic diversity in red pine: evidence for low heterozygosity.** *Can. J. For. Res.* 7:341-347.
- Gottlieb, L.D.
1977. **Electrophoretic evidence and plant systematics.** *Ann. Missouri Bot. Gard.* 65:161-180.
- Hamrick, J.L.
1979. **Genetic variation and longevity.** In *Topics in plant population biology*. O. Solbrig, S. Jain, G. Johnson, and P. Raven, eds. p. 84-113. Columbia Univ. Press, New York.
- Hamrick, J.L., and R.W. Allard.
1972. **Microgeographical variation in allozyme frequencies in *Avena barbata*.** *Proc. Nail. Acad. Sci.* 65:2100-2104.
- Hamrick, J.L., Y.B. Linhart, and J.B. Milton.
1979. **Relationships between life history characteristics and electrophoretically-detectable genetic variation in plants.** *Ann. Rev. Ecol. Syst.* 10:173-200.
- Hiebert, R.D.
1977. **The population biology of bristlecone pine (*Pinus longaeva*) in the eastern Great Basin.** Ph.D. dissertation. Univ. of Kansas, Lawrence, 82 p.
- Jain, S.K., and A.D. Bradshaw.
1966. **Evolutionary divergence among adjacent plant populations. I. The evidence and its theoretical analysis.** *Heredity* 21:407-441.
- Küchler, A.W.
1964. **Potential natural vegetation of the conterminous United States.** *Amer. Geogr. Soc. Spec. Publ.* 36.
- Lerner, I.M.
1954. **Genetic homeostasis.** 134 p. Oliver and Boyd, London.
- Libby, W.J., R.F. Stettler, and F.W. Seta.
1969. **Forest genetics and forest tree breeding.** *Ann. Rev. Genet.* 3:469-494.
- Lundkvist, K.
1979. **Allozyme frequency distributions in four Swedish populations of Norway spruce (*Picea abies* K.). 1. Estimates of genetic variation within and among populations, genetic linkage and a mating system parameter.** *Hereditas* 90:127-143.
- Neale, D.B.
1978. **Allozyme studies in balsam fir.** M.S. Thesis, University of New Hampshire, Durham, N.H. 83 p.
- Nevo, E.
1978. **Genetic variation in natural populations: patterns and theory.** *Theor. Popul. Biol.* 13:121-177.
- Raunkiaer, C.
1934. **The life forms of plants and statistical plant geography; being the collected papers of C. Raunkiaer.** 632 p. Clarendon Press, Oxford.
- Schaal, B.A., and D.A. Levin.
1976. **The demographic genetics of *Liatris cylindracea* Michx. (Compositae).** *Amer. Nat.* 110:191-206.
- Selander, R.K.
1976. **Genic variation in natural populations.** In *Molecular evolution*. F.J. Ayala, ed. p. 21-45. Sinauer, Sunderland, Massachusetts.
- Smith, R.H., R.L. Peloquin, and P.C. Passof.
1969. **Local and regional variation in the monoterpenes of ponderosa pine wood oleoresin.** USDA Forest Serv. Res. Paper PSW-56, 10 p. Pacific Southwest Forest and Range Exp. Stn., Berkeley, Calif.
- Stern, K., and L. Roche.
1974. **Genetics of forest ecosystems.** 330 p. Springer-Verlag, New York.
- Torres, A.M., U. Diedenhofen, B.O. Bergh, and R.L. Knight.
1978. **Enzyme polymorphisms as genetic markers in the avocado.** *Amer. J. Bot.* 65:134-139.
- Zavarin, E., and K. Snajberk.
1965. **Chemotaxonomy of the genus *Abies*, I. Survey of the terpenes present in the *Abies* balsams.** *Phytochemistry* 4:141-148.
- Zavarin, E., W.B. Critchfield, and K. Snajberk.
1969. **Turpentine composition of *Pinus contorta* X *Pinus banksiana* hybrids and hybrid derivatives.** *Can. J. Bot.* 47:1443-1453.

Genetic Structure of Populations and Differentiation in Forest Trees¹

Raymond P. Guries² and F. Thomas Ledig³

Abstract: Electrophoretic techniques permit population biologists to analyze genetic structure of natural populations by using large numbers of allozyme loci. Several methods of analysis have been applied to allozyme data, including chi-square contingency tests, F-statistics, and genetic distance. This paper compares such statistics for pitch pine (*Pinus rigida* Mill.) with those gathered for other plants and animals. On the basis of these comparisons, we conclude that pitch pine shows significant differentiation across its range, but appears to be less differentiated than many other organisms. Data for other forest trees indicate that they, in general, conform to the pitch pine model. An open breeding system and a long life cycle probably are responsible for the limited differentiation observed in forest trees. These conclusions pertain only to variation at allozyme loci, a class that may be predominantly neutral with respect to adaptation, although gene frequencies for some loci were correlated with climatic variables.

During the last 40 years, a considerable body of information has been assembled on genetic variability in forest tree species. Measurements of growth rate, cold-hardiness, phenology, and related traits have provided tree breeders with much useful knowledge, and have indicated in a general way how tree species have adapted to a spatially variable habitat (Wright 1976). Most information has come from studies designed to assess relative differences in metric traits among provenances representing a wide geographic range. When the identification of suitable seed sources for use in reforestation is a primary objective, provenance tests (that is, common garden studies) may be indispensable. However, such studies are expensive to conduct, require large areas of land, yield useful data only after many years, and use traits of unknown inheritance.

The development of electrophoretic techniques during the last two decades provided an alternative to common garden techniques for estimating levels of genetic variation in natural populations (Lewontin 1974). Electrophoretic techniques are now widely used in studies of plant and animal populations and have produced a large body of data on the levels and patterns of genic variability characteristic of many species. Although the difficulties inherent in the large size and late reproductive maturity of trees somewhat delayed their study, there is in recent years a rapid accumulation of useful data from the application of electrophoretic techniques to several conifers.

POPULATION STRUCTURE

A common notion among foresters is that most tree species are divided into relatively small breeding populations because of limited pollen or seed dispersal. Most matings are considered to be among near neighbors, leading in time to the division of a large population into numerous, small "neighborhoods." Other factors such as physical isolation, breeding system, population demography, natural selection, and random genetic drift accentuate such subdivision. The sum total of the ecological and genetic relationships among individuals and the populations they comprise is termed "population structure" (Jain 1975).

Of special interest to tree breeders are the genetic consequences of population structure. Several measures of the genetic structure of populations have been proposed, including assessments of (a) gene diversity in the average population, (b) levels of diversity in different populations, and (c) degree of differentiation among populations (Brown 1978). All these measures are amenable to allozyme analysis, and recent explorations have begun to probe the genetic structure of forest trees.

Our approach in this paper is to compare estimates of various parameters of population structure obtained from pitch pine (*Pinus rigida* Mill.) with those of other species. Rigorous comparisons are seldom possible because of variation in methods of data collection and analysis; therefore, we use them here only for purposes of general description.

CHI-SQUARE ANALYSIS OF HETEROGENEITY

The subdivision of populations may lead to a heterogeneity of gene frequencies among subpopulations as a result of selection or drift. Variation in genic proportions

¹Presented at the Symposium on Isozymes of North American Forest Trees and Forest Insects, July 27, 1979, Berkeley, Calif.

²Assistant Professor of Forestry, University of Wisconsin, Madison, Wisc.

³Project Leader, Institute of Forest Genetics, Pacific Southwest Forest and Range Experiment Station, Forest Service, U.S. Department of Agriculture, Berkeley, Calif.; and Professor of Forest Genetics, Yale University, New Haven, Conn.

can be subjected to a contingency chi-square analysis to determine the existence or extent of any heterogeneity (Workman and Niswander 1970). A major advantage of the technique is ease of calculation; however, the analysis does not identify the forces responsible for any observed heterogeneity. For pitch pine, chi-square was calculated using the technique of Snedecor and Irwin (1933):

$$\chi^2 = 2N \sum_i \frac{\sigma_i^2}{\bar{p}_i}$$

in which N is the total number of individuals, σ_i^2 is the variance in allele frequency, and $\bar{p}_i$ is the weighted mean frequency.

A significant amount of heterogeneity is apparent at most of 21 loci in 11 populations of pitch pine (table 1). Considering the wide range over which pitch pine samples were collected (from Quebec to North Carolina), differential selection in varying environments is one possible explanation. Distribution of allele frequencies was observed to be heterogeneous over relatively short distances in ponderosa pine (*Pinus ponderosa* Dougl. ex Laws.), perhaps because of differential selection on slopes of varying aspect (Mitton and others 1977). However, Neale (1978) found no evidence of allelic heterogeneity among populations of balsam fir (*Abies balsamea* [L.] Mill.) occurring along an elevational gradient in New Hampshire, although differences in growth and physiological processes were observed and related to environmental variation along the same transect (Fryer and Ledig 1972).

Genetic drift could also produce heterogeneity. In pitch pine, certain populations (for example, one in St. Chryso-

stome, Quebec) are isolated from the main body of the species range. If such populations were founded by a small number of individuals, genetic drift could account for much of the observed heterogeneity. Genetic drift was a possible explanation for observed heterogeneity of gene frequencies among three small populations of Table Mountain pine (*Pinus pungens* Lamb.) in Virginia (Feret 1974). In any event, the large range in chi-square values among loci in pitch pine suggests that different loci are probably responding independently to whatever factors are responsible for the heterogeneity.

F-STATISTIC ANALYSIS OF POPULATION STRUCTURE

The structure of a subdivided population can also be analyzed by F-statistics (Wright 1951, 1965, 1969; Kirby 1975; Nei 1977). F-statistics were originally devised to examine structuring in hierarchical populations by using the correlation between uniting gametes within and among subpopulations and for the population as a whole. Wright (1965) advanced three parameters "in terms of a total population (T), subdivisions (S), and individuals (I). F_{IT} is the correlation between gametes that unite to produce the individuals, relative to the gametes of the total population. F_{IS} is the average overall subdivisions of the correlation between uniting gametes relative to those of their subdivision. F_{ST} is the correlation between random gametes within subdivisions, relative to gametes of the total population. The list can be extended if there are further subdivisions. The above three F-statistics are not independent."

These statistics have as a common focal point the fixation index, F, which represents the total deviation from Hardy-Weinberg proportions because of the joint effects of finite population size, selection, inbreeding, and other factors. The estimate of F_{ISi} for the i^{th} subpopulation was calculated as:

$$F_{ISi} = 1 - H_i / \frac{1}{(1 + 2N_i - 1)} 2p_i q_i$$

in which H_i is the observed number of heterozygotes in the i^{th} subpopulation and the denominator is the expected number corrected for finite population size (Kirby 1975).

F_{IS} represents the average deviation of the population's genotypic proportions from Hardy-Weinberg equilibrium for a locus and is calculated as the weighted mean of the F values for all populations;

$$\hat{F}_{IS} = \sum_i \frac{N_i}{N} p_i q_i F_{ISi} / \sum_i \frac{N_i}{N} p_i q_i$$

A negative F_{IS} value represents an excess of heterozygotes. It is worth noting that F_{IS} not equivalent to the coefficient of inbreeding except in the unlikely event that inbreeding

Table 1—Chi-square contingency analysis of gene frequency heterogeneity by locus for 11 pitch pine populations

Locus	X ²	df	P<
MDH-1	45.76	10	0.005
MDH-2	107.77	10	.005
IDH	95.43	10	.005
FUM	18.22	10	.1
PGM-1	9.39	10	.5
PGM-2	42.22	10	.005
GPI-1	16.60	10	.1
GPI-2	28.07	10	.005
6-PGD-1	166.36	10	.005
6-PGD-2	52.58	10	.005
G-6-P	39.66	10	.005
LAM	51.91	10	.005
LAP-2	37.97	10	.005
GOT-1	44.39	10	.005
GOT-2	34.59	10	.005
ACP	54.92	10	.005
ACO	131.04	10	.005
GDH	31.72	10	.005
ADH	3.64	10	.95
ALD-1	30.00	10	.005
ALD-2	148.57	10	.005

alone is responsible for departures from Hardy-Weinberg equilibrium. F_{IS} can be averaged across all loci.

The extent of differentiation among subpopulations is measured by F_{ST} , the correlation between random gametes within subdivisions, and was calculated as Nei's (1975) G_{ST} :

$$G_{ST} = D_{ST} / H_T,$$

in which H_T is the gene diversity in the total population, D_{ST} is the average gene diversity among subpopulations.

$$H_T = 1 - \sum_k \left(\sum_i p_{ik} / S \right)^2, \text{ and}$$

$$D_{ST} = \sum_i \left(\sum_k p_{ik}^2 \right) - \sum_k \left(\sum_i p_{ik} / S \right)^2.$$

in which p_{ik} is the frequency of the k^{th} allele in the i^{th} subpopulation and S is the number of subpopulations.

For gene loci with only two alleles, F_{ST} was also calculated as:

$$F_{ST} = \sigma_p^2 / \bar{p}\bar{q},$$

in which σ_p^2 is the weighted sum of squared deviations of the individual subpopulation gene frequencies from the mean gene frequency, divided by the number of subpopulations, and P and q represent weighted mean frequencies.

F_{ST} was extended to multiple allele systems by Nei (1965):

$$F_{ST} = -\sigma_{jk} / \bar{p}_j \bar{p}_k,$$

in which

$$\sigma_{jk} = \sum_i \frac{N_i}{N} (p_{ji} - \bar{p}_j)(p_{ki} - \bar{p}_k),$$

$$\bar{p}_j = \sum_i \frac{N_i}{N} p_{ji}, \text{ and}$$

$$\bar{p}_k = \sum_i \frac{N_i}{N} p_{ki},$$

Table 2— F -Statistics from plant, animal, and human populations

Groups of plant, animal, and human populations	Loci	F_{IS}	F_{ST}	F_{IT}	Sampling unit	Source
Japanese	1	0.002	0.001	0.003	All Japan	Nei and Imaizumi 1966
Monarch butterfly	6	.008	.009	.017	Eastern United States	Eanes and Koehn 1978
Balsam fir	7	.001	.012	.013	New Hampshire mountainside	Neale 1978
Papago Indians	7	-.005	.023	.009	10 reservation districts	Workman and Niswander 1970
Pitch pine	21	.009	.024	.034	Eastern United States	Guries and Ledig (in preparation)
Bluegill	3	.012	.029	.024	Within reservoirs	Awise and Felley 1979
Yanomama Indians	8	.022	.064	.045	37 villages in Venezuela	Neel and Ward 1972
Cylindric blazing-star	15	.407	.069	.426	1/8 acre	Schaal 1975
Phlox (3 species)	4 to 6	.503	.200	.643	Central Texas	Levin 1978

where p_{ji} and p_{ki} are the j^{th} and k^{th} alleles in the i^{th} population. F_{ST} for multiple alleles must be calculated for each combination of alleles. For pitch pine, differences between Nei's G_{ST} and Wright's F_{ST} were negligible, and Nei's calculation was preferred.

The overall fixation index, F_{IT} , represents the correlation between uniting gametes relative to the gametes of the total population and was calculated from F_{IS} and F_{ST} as:

$$F_{IT} = F_{IS} + (1 - F_{IS}) F_{ST}.$$

Like F_{IS} , F_{IT} may be positive or negative, with a negative value indicative of excess heterozygotes. If all populations are in Hardy-Weinberg equilibrium, $F_{IS} = 0$ and $F_{IT} = F_{ST}$. However, even if Hardy-Weinberg proportions are obtained, differentiation because of differing allele frequencies in subpopulations can lead to significant F_{ST} and F_{IT} values.

F -statistics are arranged by increasing F_{ST} values for several species of plants and animals including human populations in table 2. Two species of trees, pitch pine and balsam fir, for which estimates are included in table 2, were both characterized by relatively small F_{IS} values, an indication that populations are at or near Hardy-Weinberg equilibrium. This may or may not indicate that inbreeding is insignificant, because a tendency toward increasing F_{IS} values by way of inbreeding could be offset by such factors as migration or differential fertility. In fact, opposing forces probably counter-balance each other in many instances, with the net result that F (and, therefore, F_{IS}) is near zero (Workman 1969).

For both tree species, F_{ST} values were also small, relative to other organisms. Balsam fir, which did not show allelic heterogeneity in contingency chi-square tests (Neale 1978), has a value of F_{ST} about one-half as large as pitch pine. However, the range sampled in pitch pine was much larger, more than 1000 km in both north-south and east-west directions, while balsam fir was sampled on a single mountainside. F_{ST} values have been calculated for a variety of other organisms not given in table 2, but including the house mouse ($F_{ST} = 0.024$, one barn; Selander 1970), *Drosophila robusta* ($F_{ST} = 0.055$, Eastern United States; Prakash 1973), and the brown snail ($F_{ST} = 0.116$, a city

block; Selander and Kaufman 1975). Geographic differentiation for pitch pine across its range is roughly no larger than that observed among bluegill populations within a single lake or Indian settlements on a single Indian reservation. The estimates for pitch pine are considerably smaller than those obtained for brown snails within the narrow confines of a city block or the herbaceous plant, cylindric blazing-star (*Liatris cylindracea* Michx.), growing on a plot of only 1/8-acre. Although population differentiation has occurred in pitch pine as indicated by the chi-square contingency test, it has not developed to the degree characteristic of many other organisms.

As a result of a low fixation index within populations (low F_{IS}) and a lack of extensive differentiation among populations (low F_{ST}), the total fixation index, F_{IT} , in pitch pine is also low. The two estimates for herbaceous plants included in table 2 have F_{IT} values from 10 to 20 times greater than pitch pine, a feature probably indicative of a high degree of inbreeding in the herbs.

ANALYSIS OF GENETIC DISTANCE

Perhaps the most widely applied index of genetic differentiation is that of genetic distance (Nei 1972). This method was developed to use isozyme data as a measure of the accumulated number of gene substitutions per locus and is defined by:

$$D = -\log_e I$$

in which I represents the normalized identity of genes. If two populations have the same alleles in the same frequency, $I = 1$; when two populations have no alleles in common, $I = 0$.

Estimates of genetic distance among populations within a species (or among species) are available for a number of plants and animals (see reviews of Avise 1974, Gottlieb 1977). Estimates of the mean genetic distance among conspecific populations for a number of species indicate that divergence among populations of tree species has been rather limited relative to that found in other organisms over much smaller areas (table 3). Results are consistent

with F_{ST} values presented earlier. It should be noted that all tree species for which such estimates are available are outbreeding and anemophilous, and gene flow may be substantially greater than for trees that use animals as pollen vectors.

Comparisons of genetic distance with geographic distance have frequently been useful in determining whether differentiation can be explained on the basis of isolation by distance. For example, genetic distance in cylindric blazing-star (table 2) was strongly correlated with geographic distance and "neighborhoods" became well-differentiated over distances of several meters (Schaal 1974).

To test whether a significant relationship exists between genetic and geographic distance for populations of pitch pine, a product-moment correlation was calculated and proved to be nonsignificant ($r = 0.263$; $0.1 > P > 0.05$). By contrast, a significant correlation was noted between genetic and geographic distance for Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco) populations in coastal British Columbia.⁴ However, it appears that much of the differentiation observed in Douglas-fir was the result of differences between island and mainland populations.⁵

RELATIONSHIP TO ENVIRONMENTAL GRADIENTS

Geographic distance *per se* may not be the most useful parameter by which to measure isolation, especially if the loci in question are under selection pressure. Therefore, product-moment correlations between the frequency of the most common allele and several climatic variables were calculated for the 11 most polymorphic loci in pitch pine (table 4).

Significant correlations were noted for six of the 11 loci. Although such correlations are not evidence for the operation of natural selection, they parallel clinal patterns

⁴Yeh, F.C., and D. M. O'Malley. Enzyme variations in natural populations of Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco) from British Columbia. Ms. submitted to *Silvae Genetica*.

⁵Personal communication from D. M. O'Malley, University of Wisconsin, Madison, July 1979.

Table 3—Mean genetic distance estimates, D , for a number of species and sampling units with estimates arranged in order of increasing $\bar{D}$

Species	Loci	$\bar{D}$	Sampling unit	Source
Douglas-fir	21	0.004	Coastal British Columbia	Yeh and O'Malley (in preparation)
Pitch pine	21	.007	Eastern United States	Guries and Ledig (in preparation)
Silver spot butterfly	3	.013	Northern California	Brittnacher and others 1978
Norway spruce	4	.014	Sweden	Lundkvist and Rudin 1977
House mouse	41	.014	Single barn	Selander, Hunt and Young 1969 (see Nei 1972)
Phlox	20	.019	Central Texas	Levin 1978
Mosquitos	18	.042	Coastal Kenya	Tabachnick and others 1979
American Indian tribes	6	.053	Northern South America	Ward and others 1975
Topminnow	25	.119	Northwestern Mexico	Vrijenhoek and others 1977
Pocket gophers	23	.144	Southwestern United States	Patton and Yang 1977
Newts	35 to 40	.292	River drainage in California	Hedgecock 1978

Table 4—Product-moment correlations between frequency of most common allele for 11 loci in pitch pine and selected climatic variables

Climatic variable ¹	Locus ²										
	MDH-1	MDH-2	IDH	6-PG-1	6-PG-2	G-6-P	LAP-1	LAP-2	GOT-1	GOT-2	ACO
Min. Jan. temp.	-0.06	0.406	0.655*	-0.718**	-0.367	-0.743**	-0.12	0.275	-0.776**	0.131	-0.502
Max. July temp.	.514	-.19	.432	-.342	.175	-.002	.139	.675	.036	-.257	.228
Heating ° days	-.18	-.577*	-.664*	.693*	.272	.734**	.095	-.477	.693*	.333	.333
Annual precipitation	-.042	.557	.453	-.401	-.463	-.772**	-.743**	-.181	-.514	-.344	-.441
Total snowfall	-.150	-.588*	.440	.687*	.338	.779**	.191	-.279	.721**	-.048	.292

¹ Source: 30-year climatic data summaries from NOAA Environmental Data Service (1979).

² *Statistically significant at 5 percent level.

**Statistically significant at 1 percent level.

of phenotypic variation noted in wood specific gravity (Ledig and others 1975) and cone serotiny (Ledig and Fryer 1972), and genetic variation in height growth revealed in a common garden study (Ledig and others 1976). In all of these studies patterns of variation were well correlated with ecological variables considered important to growth or survival. In the present study, for four of six loci, significant correlations between allele frequencies and climatic variables involved winter temperature or snowfall, which perhaps, reflect environmental stress; conditions which tend to reduce survival and limit reproductive rate in pitch pine.

DISCUSSION

The overall picture of pitch pine emerging from these analyses is one of a weakly differentiated series of populations. Although isolation by distance appears to be an important aspect of differentiation in herbaceous plant species, it is only mildly so in pitch pine. The lack of significant barriers to gene flow may be one important reason for the relatively undifferentiated status of pitch pine with regard to allozyme loci. Although the range of pitch pine includes disjunct populations, no major geographic barriers to gene flow exist, and the time since pitch pine migrated from its glacial refugium to occupy its present range has been brief, less than 100 to 300 generations. Furthermore, pitch pine, like most north-temperate tree species, is anemophilous, and wind can be an effective pollen vector for long distance transport (Koski 1970). In addition, pitch pine shows a dramatic depression when subjected to inbreeding (Wright 1962); therefore, selection reduces the impact of inbreeding and favors a strong tendency to outbreeding. For genetic distance, the estimates for pitch pine and other tree species (table 3) are similar and smaller than similar comparisons with inbreeding herbaceous species. The characteristic long generations and open breeding systems shared by tree species may act to distinguish them from herbaceous species.

Nevertheless, the data are insufficient to determine whether the results observed for pitch pine are really typical of other forest trees. Our current opinion is that

they will prove to be reasonably typical of many other eastern conifers. Similarities in breeding system, demography, range, time since post-Pleistocene recolonization, and other factors all suggest that patterns paralleling that of pitch pine might be expected. Conifers such as eastern hemlock (*Tsuga canadensis* [L.] Carr.), which is a late successional species, red pine (*Pinus resinosa* Ait.), which has limited genetic variability (Fowler and Morris 1977), or jack pine (*Pinus banksiana* Lamb.) with an extensive range, may differ from the pitch pine model. Conifers of the Western United States may be relatively more differentiated than pitch pine, especially those that have extensive and fractionated populations in mountainous areas where opportunities for differential selection are great. In mountainous areas population structure in its spatial aspects conforms to an island model, and is conducive to differentiation among populations. Some hardwoods, those that are insect-pollinated, may also deviate considerably from the pitch pine model, but most anemophilous hardwoods should parallel pitch pine. Electrophoretic studies of these and other species will undoubtedly increase our knowledge of the genetic structure and differentiation characteristic of forest trees. Results from such studies should be of great interest and value to population biologists and tree breeders alike.

Acknowledgments

This study was supported by Grant DEB74-11794, from the U.S. National Science Foundation and Regional Research Project 142-1826 from the Science and Education Administration, U.S. Department of Agriculture, and by the College of Agriculture and Life Sciences, University of Wisconsin, Madison.

LITERATURE CITED

- Avise, J. C.
1974. Systematic value of electrophoretic data. *Syst. Zool.* 23:465-481.
Avise, J. C., and J. Felley.
1979. Population structure of freshwater fishes. I. Genetic variation of bluegill (*Lepomis macrochirus*) populations in man-made reservoirs. *Evolution* 33:15-26.
Brittnacher, J. G., S. R. Sims, and F. J. Ayala.
1978. Genetic differentiation between species of the genus *Speyeria* (Lepidoptera: Nymphalidae). *Evolution* 32:199-210.

- Brown, A. H. D.
1978. **Isozymes, plant population genetic structure and genetic conservation.** Theor. Appl. Genet. 52:145-157.
- Eanes, W. F., and R. K. Koehn.
1978. **An analysis of genetic structure in the Monarch butterfly, *Danaus plexippus*.** Evolution 32:784-797.
- Feret, P. P.
1974. **Genetic differences among three small stands of *Pinus pungens*.** Theor. Appl. Genet. 44:173-177.
- Fowler, D. P., and R. W. Morris.
1977. **Genetic diversity in red pine: evidence for low genic heterozygosity.** Can. J. For. Res. 7:343-347.
- Fryer, J. H., and F. T. Ledig.
1972. **Microevolution of the photosynthetic temperature optimum in relation to the elevation] complex gradient.** Can. J. Bot. 50: 1231-1235.
- Gottlieb, L. D.
1977. **Electrophoretic evidence and plant systematics.** Ann. Missouri Bot. Gard. 64:161-180.
- Hedgecock, D.
1978. **Population subdivision and genetic divergence in red-bellied newt, *Taricha rivularis*.** Evolution 32:271-286.
- Jain, S. K.
1975. **Population structure and the effects of breeding system.** In Crop genetic resources for today and tomorrow. O. H. Frankel and J. G. Hawkes, eds., p. 15-36. Cambridge Univ. Press, Cambridge.
- Kirby, G. C.
1975. **Heterozygote frequencies in small subpopulations.** Theor. Popul. Biol. 8:31-48.
- Koski, V.
1970. **A study of pollen dispersal as a mechanism of gene flow in conifers.** Communicationes Instituti Forestalis Fenniae 70.4. Helsinki, Finland.
- Ledig, F. T., and J. H. Fryer.
1972. **A pocket of variability in *Pinus rigida* Mill.** Evolution 26: 259-266.
- Ledig, F. T., C. C. Lambeth, and D. I. H. Linzer.
1976. **Nursery evaluation of a pitch pine provenance trial.** In Proc. 23d Northeast. Forest Tree Improv. Conf., New Brunswick, N. J. p. 93-108.
- Ledig, F. T., B. J. Zobel, and M. F. Matthias.
1975. **Geoclimatic patterns in specific gravity and tracheid length in wood of pitch pine.** Can. J. For. Res. 5:318-329.
- Levin, D. A.
1978. **Genetic variation in annual phlox: self-compatible versus self-incompatible species.** Evolution 32:245-263.
- Lewontin, R. C.
1974. **The genetic basis of evolutionary change.** 346 p. Columbia Univ. Press, N. Y.
- Lundkvist, R., and D. Rudin.
1977. **Genetic variation in eleven populations of *Picea abies* as determined by isozyme analysis.** Hereditas 85:67-74.
- Mitton, J. B., Y. B. Linhart, J. L. Hamrick, and J. S. Beckman.
1977. **Observations on the genetic structure and mating system of ponderosa pine in the Colorado front range.** Theor. Appl. Genet. 51:5-13.
- Neale, D.
1978. **Allozyme studies in balsam fir.** M.S. thesis. Univ. New Hampshire, Durham.
- Neel, J. V., and R. J. Ward.
1972. **The genetic structure of a tribal population, the Yanomama Indians. VI. Analysis of F-statistics (including a comparison with the Makiritare and Xavante).** Genetics 72:639-666.
- Nei, M.
1965. **Variation and covariation of gene frequencies in subdivided populations.** Evolution 19:256-258.
- Nei, M.
1972. **Genetic distance between populations.** Amer. Nat. 106:283-292.
- Nei, M.
1975. **Molecular population genetics and evolution.** 288 p. Amer. Elsevier Publ., N. Y.
- Nei, M.
1977. **F-statistics and analysis of gene diversity in subdivided populations.** Ann. Human Genet. 41:225-233.
- Nei, M., and Y. Imaizumi.
1966. **Genetic structure of human populations. I. Local differentiation of blood group gene frequencies in Japan.** Heredity 21:9-35.
- NOAA-Environmental Data Service
1978. **Climatological data. National summary.** V. 27 (13) annual summary. National Climate Center, Asheville, N.C. 128 p.
- Patton, J. L., and S. Y. Yang.
1977. **Genetic variation in *Thomomys bottae* pocket gophers: macrogeographic patterns.** Evolution 31:697-720.
- Prakash, S.
1973. **Patterns of gene variation in central and marginal populations of *Drosophila robusta*.** Genetics 75:347-369.
- Schaal, B.
1974. **Isolation by distance in *Liatris cylindracea*.** Nature 252:703.
- Schaal, B.
1975. **Population structure and local differentiation in *Liatris cylindracea*.** Amer. Nat. 109:511-528.
- Selander, R. K.
1970. **Behavior and genetic variation in natural populations.** Amer. Zool. 10:53-66.
- Selander, R. K., and D. W. Kaufman.
1975. **Genetic structure of populations of the brown snail (*Helix aspersa*). I. Microgeographic variation.** Evolution 29:385-401.
- Selander, R. K., W. G. Hunt, and S. Y. Yang.
1969. **Protein polymorphism and genetic heterozygosity in two European subspecies of the house mouse.** Evolution 23:279-290.
- Snedecor, G., and M. R. Irwin.
1933. **On the chi-square test for homogeneity.** Iowa State Coll. J. Sci. 8:75-81.
- Tabachnick, W. J., L. E. Munstermann, and J. R. Powell.
1979. **Genetic distinctness of sympatric forms of *Aedes aegypti* in East Africa.** Evolution 33:287-295.
- Vrijenhoek, R. C., R. A. Angus, and R. J. Schultz.
1977. **Variation and heterozygosity in sexually versus clonally reproducing populations of *Poecidopsis*.** Evolution 31:767-781.
- Ward, R. H., H. Gershowitz, M. Layrisse, and J. V. Neel.
1975. **The genetic structure of a tribal population, the Yanomama Indians. XI. Gene frequencies for 10 blood groups and the ABH-Le secretor traits in the Yanomama and their neighbors; the uniqueness of the tribe.** Amer. J. Human Genet. 27:1-30.
- Workman, P. L.
1969. **The analysis of simple genetic polymorphisms.** Human Biol. 41: 97-114.
- Workman, P. L. and J. D. Niswander.
1970. **Population studies on southwestern Indian tribes. II. Local genetic differentiation in the Papago.** Amer. J. Human Genet. 22: 24-49.
- Wright, J. W.
1962. **Genetics of forest tree improvement.** 399 p. FAO, Rome.
- Wright, J. W.
1976. **An introduction to forest genetics.** 463 p. Academic Press, N.Y.
- Wright, S.
1951. **The genetical structure of populations.** Ann. Eugen. 15:323-354.
- Wright, S.
1965. **The interpretation of population structure by F-statistics with special regard to systems of mating.** Evolution 19:395-420.
- Wright, S.
1969. **Evolution and the genetics of populations. V. II. The theory of gene frequencies.** 511 p. Univ. Chicago Press, Chicago.

Analyses of Gene Diversity in Some Species of Conifers¹

Francis C. Yeh²

Abstract: Genetic variation at 21 to 25 loci in extracts of individual megagametophytes was surveyed in Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco), Sitka spruce (*Picea sitchensis* [Bong.] Carr.) and lodgepole pine (*Pinus contorta* ssp. *latifolia* [Engelm.] Critchfield). The overall mean proportion of polymorphic loci was 61.19 percent, and the overall mean heterozygosity per individual was 15.81 percent. Sitka spruce was on the low side and interior Douglas-fir on the high side of overall mean genetic variation. Distribution of loci relative to frequency of heterozygotes was rather even for heterozygosities between 0.05 and 0.60; however, between 38 to 56 percent of the loci had heterozygosities lower than 0.05. More than 90 percent of the total gene diversity resided within populations. Although subpopulations were differentiated by only 2.6 to 7.9 percent, level of population subdivision was considered significant for the species tested. The overall pattern of genetic differentiation agreed with the expected on the basis of the neutral-mutation theory. Some loci, however, demonstrated conspicuous clinal variation patterns that are not readily compatible with this stochastic model.

The recent use of gel electrophoresis in isozyme studies of genic variability in natural populations of conifers (Rudin 1976) has permitted researchers to investigate many basic questions of evolutionary biology. These questions concern levels of heterozygosity within populations, distribution of genic variation within and between local populations, and relative amounts of genetic variation in central, as opposed to, marginal populations. Until recently, most investigators addressing these questions on conifers were restricted to a small sample of loci. A striking feature of the more recent and extensive isozyme surveys, however, has been the demonstration of a high degree of interlocus variation in heterozygosity within populations (O'Malley and others 1979, Yeh and El-Kassaby 1979, Yeh and Layton 1979).³ It is imperative, therefore, that a large sample of loci be surveyed when assessing isozyme variation in conifers.

For the past 3 years, the major thrust of our research in British Columbia has concerned several conifers of commercial importance. Up to 30 loci were identified and used as genetic markers in population surveys to quantify the amount and organization of genetic variation in coastal and interior Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco), Sitka spruce (*Picea sitchensis* [Bong.] Carr.), and lodgepole pine (*Pinus contorta* ssp. *latifolia* [Engelm.] Critchfield). Our research has not been restricted to the theoretical issues of evolutionary relationships between populations, but has focused on problems that comple-

ment our applied tree-breeding program. This latter aspect includes defining subpopulations, delineating seed transfer rules, and investigating associations between allozyme frequencies and quantitative traits for indirect selection.

This paper summarizes results of our study on the amount and organization of genetic variation in Douglas-fir, Sitka spruce, and lodgepole pine.

GENETIC VARIATION IN SEVERAL SPECIES OF CONIFERS

Nineteen enzymes were surveyed by one of five buffer systems (*table 1*). Data were collected in our laboratory on the basis of electrophoretic surveys of protein extracts from individual megagametophytes (*table 2*). These data show the proportion of loci polymorphic, defined as the proportion of loci in which the most common allele does not exceed a frequency of 0.99, and the proportion of loci at which an individual can be expected to be heterozygous (Nei 1973). Because the criterion defining polymorphic loci is somewhat arbitrary³ and has a high variance as a result of the relatively small number of loci surveyed, the heterozygosity per individual is the more informative figure.

For the species shown (*table 2*), from 51 to 59 percent of the loci were segregating within a population, and the heterozygosity per individual fell in a relatively narrow range for all species, between 14.67 and 17.47 percent. The overall mean proportion of polymorphic loci was 61.19 percent and the overall mean heterozygosity per individual was 15.81 percent. Values for Sitka spruce were on the low side, and those for interior Douglas-fir were on the high side of the overall mean genetic variation. It is not incorrect, therefore, to characterize these conifers as being polymorphic for 60 percent of their genes, and individuals

¹Presented at the Symposium on Isozymes of North American Forest Trees and Forest Insects, July 27, 1979, Berkeley, Calif.

²Technical Advisor in Genetics, Research Branch, British Columbia Ministry of Forests, Victoria, B.C., Canada.

³Yeh, F. C., and D. M. O'Malley. Enzyme variations in natural populations of Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco.) from British Columbia. I. Genetic variation patterns in coastal populations. *Silvae Genetica* (in press).

within the species as being heterozygous for about 16 percent of their loci. It must be emphasized that these surveys were concerned only with alleles that are detected by using conventional starch gel techniques. Only a proportion—probably no more than 33 percent of all possible mutant alleles that produce structural changes in enzyme proteins—are detected in conventional electrophoretic surveys. The true level of gene diversity in these conifers, therefore, is likely to be greater than these estimates suggest. Probably many isozyme differences occur that are not detected by the procedures currently used in our laboratory (Coyne 1976). The rich gene pool in conifers is not surprising because they are exceedingly variable in morphology, both across their native range and from tree-to-tree within stands.

This high level of gene diversity probably results from a number of variables. Most conifer species grow in large continuous stands over wide geographic ranges. Divergent selection for macrogeographical adaptation (Allard and others 1972), balancing selection for microgeographical differentiation (Hamrick and Allard 1972, Milton and others 1977), combined with an open breeding system that facilitates gene flow within and between subpopulations, tend to maintain a rich gene pool. Heterosis would promote further the maintenance of genetic variation. The notable exception to this pattern is red pine (*Pinus resinosa* Ait.) (Fowler and Morris 1977). The lack of genetic variation within this species, however, has been hypothesized as being the result of a severe bottleneck, probably during the Pleistocene, when red pine was reduced to a small refugial population.

Estimates of mean heterozygosity (table 2) are comparable to that of 10 percent obtained by outbreeding organisms (Nei 1975) but are considerably lower than those reported previously for coastal Douglas-fir (Yang and others 1977) and Norway spruce (Lundkvist and Rudin 1977). This difference results from the bias introduced by the small sample of loci analysed in the above two studies.

The standard errors of average heterozygosity emphasize the importance of examining a large number of loci

(table 2). These standard errors are calculated from the variance among loci after the heterozygosities are averaged for populations within a species. They average about 25 percent as large as the mean heterozygosities. Such large

Table 1 —Enzymes assayed and procedures used in the survey of genic heterozygosity in some species of conifers

Enzyme	Buffer system	Loci scored		
		Douglas-fir	Sitka spruce	Lodgepole pine
Acid phosphatase (APH)	A	(*)	(*)	1
Aconitase (ACO)	A	1	1	1
Adenylate kinase (AK)	E	(*)	(*)	2
Alcohol dehydrogenase (ADH)	C	(*)	(*)	1
Aldolase (ALD)	A	1	1	(*)
Aspartate aminotransferase (AAT)	B	2	2	2
Diaphorase (DIA)	C	1	3	2
Esterase (EST)	C	1	1	(*)
Glucose-6-phosphate dehydrogenase (G6P)	A	1	1	1
Glutamate dehydrogenase (GDH)	B	1	1	1
β -Glucosidase	A	(*)	(*)	1
Isocitrate dehydrogenase (IDH)	C	1	1	1
Malate dehydrogenase (MDH)	E	4	3	4
Malic enzyme (ME)	A	2	1	2
Peptidase (PEP)	B	2	3	1
Phosphoglucose isomerase (PGI)	B	1	2	1
Phosphoglucomutase (PGM)	C	1	2	1
Superoxide dismutase (SOD)	B	1	(*)	(*)
6-phosphogluconic dehydrogenase (6PG)	C	1	2	2

¹ A = Morpholine-citrate pH 6.1; B = Tris-citrate:Li-borate pH 8.5; C = Tris-citrate pH 7.0; D = Histidine-citrate pH 7.0; E = Phosphate-citrate pH 7.0.

(*)Not assayed.

Table 2—Survey of genic heterozygosity in some species of conifers

Species	Populations	Loci	Proportion of loci polymorphic per population ¹	Heterozygosity per locus and standard error	Reference
<i>Pseudotsuga menziesii</i> (Mirb.) Franco					
coastal variety	11	21	0.6883	0.1546 $\pm$ 0.0370	Yeh and O'Malley (1979)
interior variety	11	21	.6800	.1747 $\pm$.0422	Yeh (In preparation) ²
<i>Picea sitchensis</i> (Bong.) Carr.	10	24	.5130	.1467 $\pm$.0400	Yeh and El-Kassaby (1979)
<i>Pinus contorta</i> ssp. <i>latifolia</i> (Engelm.)	10	25	.5914	.1544 $\pm$.0365	Yeh (In preparation) ³
Critchfield	9	25	.5867	.1601 $\pm$.0380	Yeh and Layton (1979)

¹The frequency of the most common allele is ≤ 0.99 .

²Yeh, F. C. Enzyme variations in natural populations of Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco) from British Columbia. II. Genetic variation patterns in interior populations. Manuscript in preparation at the Research Branch, British Columbia Ministry of Forests.

³Yeh, F. C. Altitudinal genetic differentiation in lodgepole pine (*Pinus contorta* ssp. *latifolia* [Engelm.] Critchfield). Manuscript in preparation at the Research Branch, British Columbia Ministry of Forests.

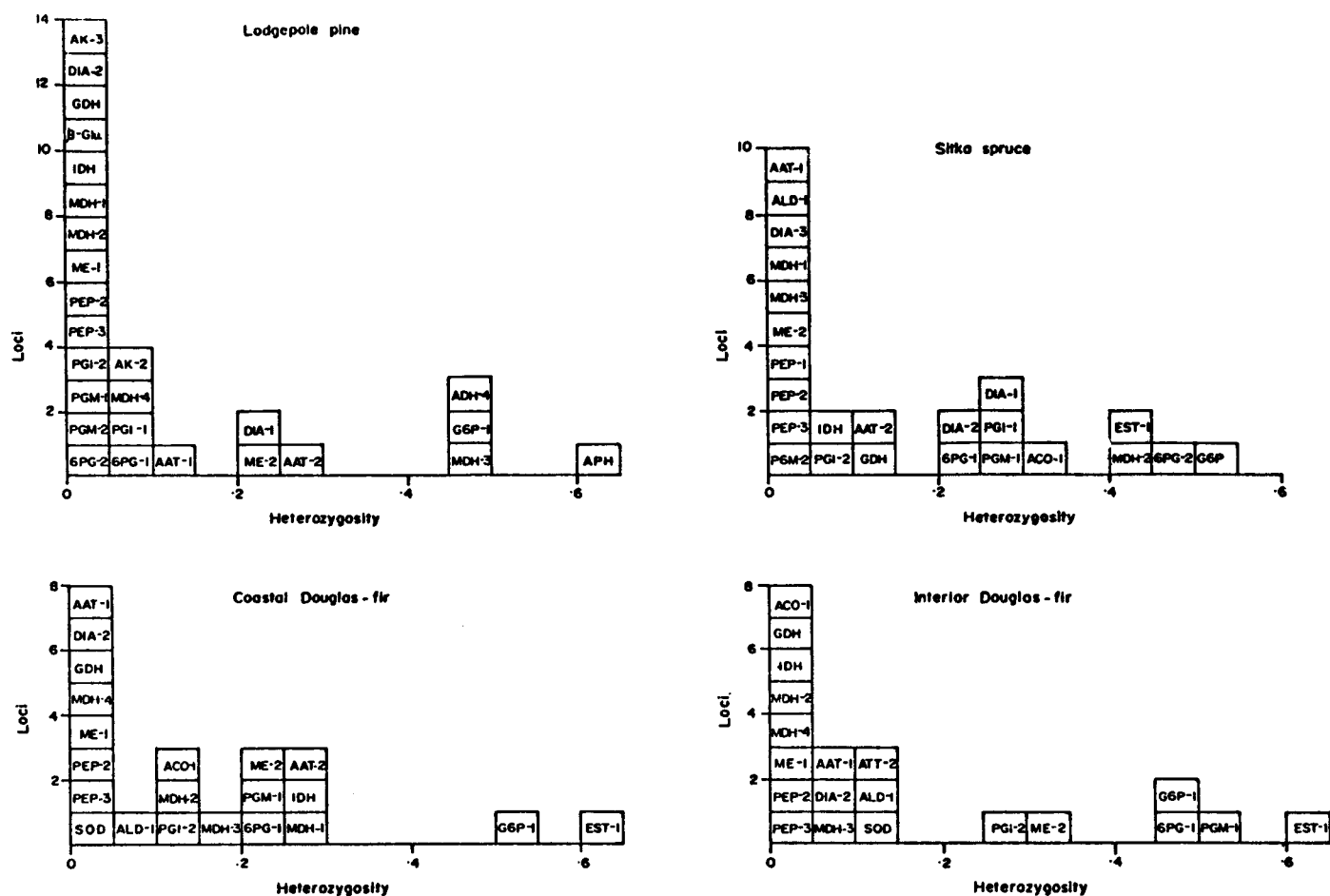


Figure 1—Distribution of heterozygosities for alleles determining electrophoretic variants in some species of conifers.

standard errors arise when loci differ markedly in their levels of variation.

Distribution of heterozygosities for alleles determining electrophoretic variants showed much interlocus variation in heterozygosity (fig. 1). Although no variation was apparent at some loci, at other loci more than 50 percent of the individuals were heterozygous. Distribution of the loci relative to the frequency of heterozygotes was rather even for heterozygosities between 0.05 and 0.60; however, between 38 to 56 percent of the loci analysed had heterozygosities lower than 0.05. The loci surveyed, therefore, do not seem to be equivalent in their contribution to the overall mean heterozygosity of the species. This broad range in heterozygosity, with a mode approaching zero, suggests that many isozyme loci should be surveyed to estimate reliably the amount of genetic variation in conifers. Because of such heterogeneity, interpretations on the basis of differences in detected heterozygosity between species should be considered tentative until a large number of loci are studied. Lewontin (1974) estimates that as many as 100 loci may be needed to adequately estimate heterozygosity and, at the present time, only studies of protein variation in human populations have approached this level of sampling.

ANALYSIS OF GENE DIVERSITY

Existing electrophoretic data on gene diversity was summarized in terms of its hierarchical organization in several species of conifers (table 3). This analysis (Nei 1973) enables genetic variation to be partitioned among different hierarchical levels of population structure—within, as opposed to between subpopulations. The technique is a modification of Wright's F-statistics (Wright 1965) expanded to multiple alleles, and is not dependent on the detection of genotypic frequencies. H_T estimates the total genetic variation sampled for all populations; it is a function of the mean allelic frequencies of the species. H_S is an estimate of the average amount of genetic variation maintained within any one subpopulation of a species. H_S can be interpreted as the proportion of loci at which an individual can be expected to be heterozygous. If all subpopulations are members of a single large panmictic unit and no gene differentiation exists among them, then all alleles will be equally distributed over the entire range and H_T will equal H_S . In nature, however, this is not true. Natural populations tend to differentiate over time into subpopulations because of the processes of mutation, selection, random drift, and restricted gene flow.

Therefore, H_S will be a subset of H_T . The extent of subdivision of a species can be described by partitioning the total gene diversity (H_T) into its components, the gene diversity within subpopulations (H_S) and between subpopulations (D_{ST}). The relative measure of genetic differentiation between subpopulations (G_{ST}) is defined by $G_{ST} = D_{ST}/H_T$ and its sampling variance $[V(G_{ST})]$ can be used to study the significance of the effect of population subdivision (Chakraborty 1974).

For the species shown (table 3), subpopulations were differentiated by 2.6 to 7.9 percent of the electrophoretically determined variation. Nevertheless, the standard errors of G_{ST} average about 11 percent of estimates of G_{ST} . This indicates that the level of population subdivision is significant for the different species. The apportionment of total gene diversity in these conifers is similar to that in man (Nei and Roychoudhury 1972) and in the horseshoe crab (Selander and others 1970) where more than 90 percent of the total gene diversity resides within local populations. That the majority of genic variation in conifers is maintained within populations is, perhaps, a reflection of their ecological amplitude, their breeding system, and the lack of effective barriers to gene flow between subpopulations.

MAINTENANCE OF GENETIC VARIATION

Genetic variation in natural populations detected by electrophoretic techniques is always confronted with the problem of determining the nature and relative roles of selection and neutral mutations in maintaining the observed variability. Only in a few examples have these been determined unequivocally. To encourage the possibility of making this distinction, I attempt here to infer the mechanism responsible for the maintenance of genetic variation in those conifers surveyed in this study.

One test of the neutral-mutation theory is to compare the theoretical variance of population heterozygosity, $Var(H)$,

with the observed variance. The theoretical variance is given by Stewart (1976) as

$$Var(H) = 2\theta / ((\theta + 1)^2(\theta + 2)(\theta + 3))$$

in which

$$\theta = 4N\mu$$

N = the effective population size

μ = the mutation rate per locus per generation

The value of θ may be estimated by $\hat{H}/(1 - \hat{H})$, in which $\hat{H}$ is the estimate of average heterozygosity, as the expectation of $\hat{H}$ is $\theta/(1 + \theta)$.

In the data for coastal Douglas-fir,³ Sitka spruce (Yeh and El-Kassaby 1979), and lodgepole pine (Yeh and Layton 1979), the expected and observed variances of population heterozygosity agree with each other surprisingly well. This agreement suggests that the observed patterns of isozyme variation are not primarily a response to selection along macroenvironmental gradients. Certain kinds of selection and varying mutation rates per locus, of course, may produce the same effect (Li 1978). Furthermore, our sampling strategies are not sensitive to microgeographical variation. Considering that the vast majority of a species' genic variation is maintained within local populations, selection may have much significance in operating on microsite differences.

An exhaustive analysis and discussion on micro- and macro-geographical genetic variation is beyond the scope of this report. It will suffice here, however, to note that patterns of variation at many loci studied in conifers are consistent generally with the expectations of a stochastic model. Several loci demonstrate conspicuous clinal variation patterns, however, that are not readily compatible with this model. Whether these nonrandom patterns of variation are the result of drift (Karlín and Richter-Dyn 1976), or selection on the enzyme themselves, or on the coadapted complexes that they mark, is problematical.

Table 3—Analysis of gene diversity in some species of conifers¹

Species	Populations	Loci	H_T	H_S	G_{ST}	Reference
<i>Pseudotsuga menziesii</i> (Mirb.) Franco						
coastal	11	21	0.1594	0.1546	0.0260 ± 0.0025	Yeh and O'Malley (1979)
interior	11	21	.1825	.1747	.0428 ± .0063	Yeh (In preparation) ²
<i>Picea sitchensis</i> (Bong.) Carr.	10	24	.1593	.1467	.0790 ± .0112	Yeh and El-Kassaby (1979)
<i>Pinus contorta</i> ssp. <i>latifolia</i> (Engelm.) Critchfield	9	25	.1670	.1601	.0411 ± .0055	Yeh and Layton (1979)

¹ H_T estimates the total genetic variation sampled for all populations; H_S estimates the average amount of genetic variation maintained within any one subpopulation of a species; G_{ST} is the relative measure of genetic differentiation between subpopulations defined by $(H_T - H_S)/H_T$.

² Yeh, F. C. Enzyme variations in natural populations of Douglas-fir in British Columbia. II. Genetic variation patterns in interior populations. Manuscript in preparation at the Research Branch, British Columbia Ministry of Forests.

LITERATURE CITED

- Allard, R. W., G. R. Babbel, M. T. Clegg, and A. L. Kahler.
1972. **Evidence of co-adaptation in *Avena barbata*.** Proc. Natl. Acad. Sci. 69:3043-3048.
- Chakraborty, R.
1974. **A note on Nei's measure of gene diversity in a substructured population.** Humangenetik 21:85-88.
- Coyne, J.
1976. **Lack of genic similarity between two sibling species of *Drosophila* as revealed by various techniques.** Genetics 84:593-607.
- Fowler, D. P., and R. W. Morris.
1977. **Genetic diversity in red pine: evidence of low genic heterozygosity.** Can. J. For. Res. 7:343-347.
- Hamrick, J. L., and R. W. Allard.
1972. **Microgeographical variation in allozyme frequencies in *Avena barbata*.** Proc. Natl. Acad. Sci. 69:2100-2104.
- Karlin, S., and N. Richter-Dyn.
1976. **Some theoretical analyses of migration selection interaction in a tine: a generalized two range environment.** In Population genetics and ecology. S. Karlin and E. Nevo, eds. p. 659-706. Academic Press, New York.
- Lewontin, R. C.
1974. **The genetic basis of evolutionary change.** 346 p. Columbia Univ. Press, New York and London.
- Li, W-H.
1978. **Maintenance of genetic variability under the joint effect of mutation, selection and random drift.** Genetics 90:349-382.
- Lundkvist, K., and D. Rudin.
1977. **Genetic variation in eleven populations of *Picea abies* as determined by isozyme analysis.** Hereditas 85:67-74.
- Mitton, J. B., Y. B. Linhart, J. L. Hamrick, and J. S. Beckman.
1977. **Observations on the genetic structure and mating system of ponderosa pine in the Colorado Front Range.** Theor. Appl. Genet. 51:5-13.
- Nei, M.
1973. **Analysis of gene diversity in subdivided populations.** Proc. Natl. Acad. Sci. 70:3321-3323.
- Nei, M.
1975. **Molecular population genetics and evolution.** 288 p. North Holland Publ. Co., Amsterdam.
- Nei, M., and A. K. Roychoudhury.
1972. **Sampling variances and heterozygosity and genetic distance.** Genetics 76:379-390.
- O'Malley, D. M., F. W. Allendorf, and G. M. Blake.
1979. **Inheritance of isozyme variation and heterozygosity in *Pinus ponderosa*.** Biochem. Genet. 17:233-250.
- Rudin, D.
1976. **Biochemical genetics and selection, application of isozymes in tree breeding.** In Proc. IUFRO joint meeting on advanced generation breeding [June 14-18, 1976, Bordeaux, France], p. 145-164.
- Selander, R. K., S. Y. Yang, R. C. Lewontin, and W. E. Johnson.
1970. **Genetic variation in the horseshoe crab (*Limulus polyphemus*) a phylogenetic relic.** Evolution 24:402-414.
- Stewart, F.M.
1976. **Variability in the amount of heterozygosity maintained by neutral mutations.** Theor. Popul. Biol. 9:188-201.
- Wright, S.
1965. **The interpretation of population structure by F-statistics with special regards to systems of mating.** Evolution 19:395-420.
- Yang, J. Ch., T. M. Ching, and K. K. Ching.
1977. **Isoenzyme variation of Coastal Douglas-fir. I. A study of geographic variation in three enzyme systems.** Silvae Genetica 26:10-18.
- Yeh, F. C., and Y. A. El-Kassaby.
1979. **Enzyme variations in natural populations of Sitka spruce (*Picea sitchensis* [Bong.] Carr.). I. Genetic variation patterns in ten IUFRO provenances.** In Proc., 52d Northwest Sci. Assoc. Meet. [Bellingham, Wash., March 29-31, 1979]. p. 25.
- Yeh, F. C., and C. Layton.
1979. **The organization of genetic variability in central and marginal populations of lodgepole pine (*Pinus contorta* ssp. *latifolia*).** Can. J. Genet. Cytol. 21:487-503.

An Analysis of Genetic Architecture in Populations of Ponderosa Pine¹

Yan B. Linhart,² Jeffry B. Mitton,² Kareen B. Sturgeon,³ and Martha L. Davis³

Abstract: Patterns of genetic variation were studied in three populations of ponderosa pine in Colorado by using electrophoretically variable protein loci. Significant genetic differences were found between separate clusters of trees and between age classes within populations. In addition, data indicate that differential cone production and differential animal damage have genetic components. These results suggest that many diverse phenomena can affect allele frequencies within populations and may contribute to the high levels of genetic variability detected within populations of this species.

Assessments of genetic variation in forest trees have been aided greatly by the use of electrophoretic analyses of protein polymorphisms. Many such studies have concentrated on the distribution of this variation over a broad geographic range (for example, Guries and Ledig 1981, Lundqvist and Rudin 1977, and Yeh 1981). Results document that substantial variation exists in trees and that this variation is distributed in patterns determined primarily by the heterogeneity of the physical environments. These findings agree with data on patterns of geographic variation collected by earlier forest geneticists (Langlet 1971, Libby and others 1969). Small-scale patterns—over a distance of 1 km or less—are less well documented but increasing evidence suggests that differentiation can occur in forest trees over distances as small as a few hundred meters (Barber 1965, Benson and others 1967, Herman and Lavender 1967, Mitton and others 1977, Grant and Mitton 1977). Again, this differentiation is associated with significant environmental heterogeneity produced by elevation, slope, or exposure.

The results mentioned above describe the organization of genetic variation between populations but say little about variation within populations. Within-population variation is particularly worthy of interest at this time, since several studies have suggested that in some forest trees, including ponderosa pine, such variation is significantly greater than variation between populations (Lines and Mitchell 1966, Madsen and Blake 1977, O'Malley and others 1979).

Many variables can affect the patterning of genetic variation within populations, including differential predation by animals and restricted pollen and seed dispersal. Also, trees are long-lived and environmental conditions change with time. As a result, different genotypes may reproduce in different years and natural

selection operating on seedlings may involve different forces in different years. Consequently, generations of different genetic constitutions may exist within multiaged stands.

This paper reports a study dealing specifically with patterns of intrapopulation variation. In 1976, we began studying populations that occupy relatively small areas, from 2 to 5 hectares, within which the heterogeneity of the physical environment is limited. We investigated three populations of ponderosa pine (*Pinus ponderosa* var. *scopulorum* Engelm.) in Colorado for (a) spatial distribution of genotypes within a given population and (b) genetic differences between age groups within a given population. We also investigated whether (c) genetic differences exist between trees that reproduce heavily and those that do not and, finally, (d) whether animal damage in the form of deer browsing or woolly aphid infestation is related to genetic constitution of the trees.

MATERIALS AND METHODS

All populations in our study grow in the Front Range of the Rocky Mountains, near Boulder, Colorado. None of them have been heavily disturbed by human activities, such as heavy logging or tree planting nearby. One population (Boulder Canyon) has been analyzed in detail; comparative data for some of the phenomena being studied are available from the Glacier Lake and Shanahan Mesa populations.

Genetic Methods

All genetic analyses are made on the basis of enzyme polymorphisms detectable in extracts of mature needle tissues. Because we are studying patterns of variation, only polymorphic loci are investigated.

The enzyme polymorphisms being studied include one peroxidase (PER-2, 3 alleles), one phosphohexose isomerase (PHI, 3 alleles), two phosphoglucumutase loci (PGM-1, 2 alleles and PGM-2, 3 alleles), one glutamate dehydrogenase (GDH, 2 alleles), one colorimetric esterase (C.E., 3 alleles), and one fluorescent esterase (F.E., 3

¹Presented at the symposium on Isozymes of North American Forest Trees and Forest Insects, July 27, 1979, Berkeley, Calif.

²Associate Professor, Department of Environmental, Population, and Organismic Biology, University of Colorado, Boulder, Colo.

³Graduate student, Department of Environmental, Population, and Organismic Biology, University of Colorado, Boulder, Colo.

alleles). Patterns of Mendelian inheritance are reported elsewhere (Mitton and others 1977, 1979).

Methods of preparation of needle extracts, resolution and staining are reported in the following publications: Mitton and others (1977) for PER-2; Mitton and others (1979) for all other loci.

Heterogeneities of allele frequencies were tested with the chi-square test of Workman and Niswander (1970).

Boulder Canyon Population

This population, within a pure ponderosa pine stand on the south-facing slope of Boulder Canyon at 1738 m, consists of all individuals—with the exceptions of a few seedlings—within an area of about 2 hectares (fig. 1). The population is located in the ponderosa pine belt typical of the lower elevations of the Colorado Front Range.

We numbered and mapped all pines in the population. We recorded age, diameter, pollen and cone output, presence or absence of infestations by woolly aphids (*Pineus coloradensis* Hopk.), browsing by deer (*Odocoileus demionus* Raf.) and the genetic constitution at all seven polymorphic protein loci.⁴

We compared the genetic constitutions of the six clusters of trees in the population. In 1977, we compared trees producing at least some pollen with those that did not, and we also counted all mature female cones. We compared prolific cone producers with poor cone producers, correcting for the effects of age and size on cone production (Linhart and others 1979a).

We analyzed browsing by mule deer, and we compared the genetic constitutions of heavily browsed trees to those of lightly or unbrowsed trees. We also compared trees that had woolly aphids on their foliage in 1979 with those that did not.

Glacier Lake Population

This population is also on a south-facing slope, at 2590 m. Here the forest is composed primarily of ponderosa pine, with Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco), limber pine (*Pinus flexilis* James), and juniper (*Juniperus scopulorum* Sarg.) intermixed. We tagged and mapped approximately one-third of the ponderosa pines within a 4-ha area. In contrast to the Boulder Canyon population, some disturbance occurred here in about 1900 when scattered logging was done in the area. Furthermore, during our study, an outbreak of bark beetles (*Dendroctonus ponderosae* Hopk.) killed more than one-third of the trees. As a result, the spatial structure of this population is not as stable as it is in the Boulder Canyon population. In this population we investigated primarily the spatial

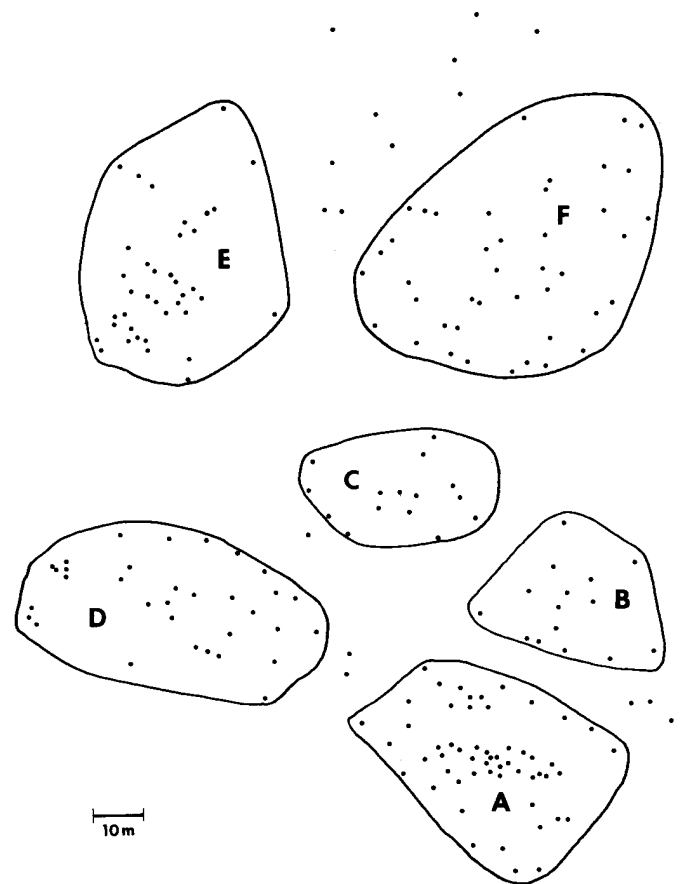


Figure 1—Approximate distribution of trees in a stand of ponderosa pine in Boulder Canyon, Colorado. Six clusters have been defined and are labeled A through E. In the text, these are called the Boulder Canyon Population. Age structures of all clusters are also shown.

distribution of genetic variation at three loci: PER, C.E. and R.E.

Shanahan Mesa Population

This population, at an elevation of 1700 to 2000 m, is on an east-facing mesa at the base of the Front Range. The mesa was grass-covered until about 120 years ago when ponderosa pine started to colonize the area from seed dispersed from stands growing on the steep slopes above the mesa. In one area, a second generation of saplings has become established under the closed canopy of the mature trees (Beckman 1977).

We compared variation at the PER locus in the original colonists, which grew up in the open grassland, with variation of their offspring, which grew up under their canopy.

RESULTS AND DISCUSSION

Spatial Clustering and Genetic Differentiation

Both the Boulder Canyon and the Glacier Lake populations have clusters of trees that are easily identified.

⁴Linhart, Y.B., J.B. Mitton, and K.B. Sturgeon. Some genetic features of clustering in a ponderosa pine population. (Manuscript in preparation.)

Open spaces are found between clusters. In each population, some trees, located either between clusters or at the edge of the mapped area, are not clearly assignable to a given cluster and, therefore, are omitted from the comparisons. In both populations we found significant differences in allele frequencies among the various clusters (table 1). Genetic heterogeneity is particularly striking in the Boulder Canyon population where nearby clusters can vary by as much as 20 percent at GDH, 21 percent at C.E., 22 percent at PER-2, and 29 percent at F. E. We found little intrapopulation variation at the other three loci tested—PGM-1, PGM-2, and PHI have their most common alleles at frequencies of 88 to 100 percent in all clusters. Similar results are found in the Glacier Lake population (table 1).

Similar genetic differentiation between carefully mapped, nearby clusters of plants has also been documented for *Linanthus parryae* (Gray) Greene (Epling and Dobzhansky 1942), *Liatris* (Schaal 1975), *Pinus sylvestris* L. (Rudin and others 1977), and *Picea abies* (L.) Karst. (Tigerstedt 1973). Tigerstedt believed that evidence for such heterogeneity was lacking within the populations he studied; however, an analysis of his figure 4 shows that, if one divides that population into two segments by an E-W line running through the center of the mapped area, the fast allele at the LAP locus is at a frequency of 0.67 in one area composed of 24 trees, and at a frequency of 0.39 in the other area composed of 23 trees. This difference is significant ($X^2 = 16.49$, 1 d.f., $P < 0.005$). Some evidence also exists for clustering and genetic heterogeneity in local populations of other trees. These include *Pseudotsuga menziesii* (Mirb.) Franco (Rehfeldt 1978, Shaw and Allard 1981), *Cupressus macrocarpa* Hartw. (Kafton 1977), *Shorea leprosula* Miq. (Dipterocarpaceae), and *Xerospermum intermedium* Radlk. (Sapindaceae) (Gan and others 1977) and *Fagus* (Stern and Roche 1974). No

evidence was found, however, for such clustering in populations of *Pinus rigida* Mill. (Guries and Ledig 1977) or *Pinus taeda* L.⁵

The genetic differentiation observed between nearby clusters probably results in part from the genetic relatedness of trees within a cluster. Such relatedness may occur when most seeds from a given parent tree are dispersed within a small area. Even if the pollen that fertilized the parent tree is a completely random sample of the pollen available in the stand, the group of trees arising from the seed rain of a single parent is a group of half-sibs. Also, those wind-dispersed seeds that are least likely to be transported any distance are those produced in the lower portion of a crown; those seeds also happen to be the ones that are most likely to have been produced by self-pollination (Fowler 1965). This fact may contribute to the genetic relatedness among individuals within a group.

Inbreeding, in the absence of selection, would be expected to produce deficiencies of heterozygotes relative to Hardy-Weinberg expectations. The distribution of genotype frequencies within given clusters compared with frequencies expected under Hardy-Weinberg equilibrium showed few significant departures from expectations. In the Glacier Lake population, no departures were observed. In Boulder Canyon, at the PER-2 locus, group B had 13 heterozygotes in a total of 15 trees ($X^2 = 7.13$, 2 d.f., $P < 0.05$), and group D had 24 heterozygotes in a total of 33 trees ($X^2 = 13.75$, 2 d.f., $P < 0.005$). Both of these departures result from excesses of heterozygotes. There are only two significant deviations in the 24 tests (6 clusters x 4 highly variable loci) that we performed, and these observations may be fortuitous. They may also reflect

⁵Personal communications from F. T. Ledig, and from M. Thompson Conkle, geneticists, Forest Service, U.S. Dep. Agric., July 26, 1979.

Table 1—Allele frequencies $\pm$ standard errors for the most common allele at several protein loci among clusters within two populations of ponderosa pine

<i>ponderosa pine</i>								
Locus (allele)	Cluster						X ²	P
	A	B	C	D	E	F		
	Boulder Canyon Population							
	PER-2 (2)	0.787± 0.04	0.567± 0.08	0.750± 0.08	0.712 ± 0.06	0.737 ± 0.05	0.721± 0.05	11.39(10 d.f.) <0.20
	GLU (1)	.723± .04	.633± .08	.532± .09	.697 ± .06	.711 ± .05	.733± .05	5.34(5 d.f.) < .40
	F.E. (2)	.630± .05	.667± .08	.750± .08	.742 ± .05	.461 ± .06	.547± .05	21.72(10 d.f.) < .025
	C.E. (1)	.722± .04	.767± .08	.563± .09	.773 ± .05	.697 ± .05	.767± .05	25.01(10 d.f.) < .01
	Total trees	54	15	16	33	38	43	
	Glacier Lake Population							
PER-2 (2)	0.709±0.05	0.783±0.06	0.725±0.06	0.833 ±0.07	0.727 ±0.05	0.833±0.07	1.28(4 d.f.) <0.90	
F.E. (2)	.535± .05	.522± .08	.500± .08	.714 ± .07	.583 ± .06	.667± .08	17.30(8 d.f.) < .05	
C.E. (1)	.628± .05	.696± .06	.550± .08	.785 ± .07	.693 ± .06	.667± .08	12.42(8 d.f.) < .20	
Total tree ¹	43	23	20	14 to 15	31 to 44	18		

¹ For clusters D and E, this number is variable because some trees died before their genotype at a given locus could be ascertained.

different selection pressures: other populations of ponderosa pine at comparable elevations also show excesses of heterozygotes at the PER-2 locus (Mitton and others 1977).

The extent of differentiation between clusters reported here is high: there may be greater genetic differences between some clusters *within* these pine populations than are found between populations separated by many kilometers in other species. One way to compare levels of differentiation is by estimating genetic distance, D (Nei 1972). Such comparisons must be done with caution because values of D are affected by the relative proportion of monomorphic loci used in the calculations. Because we used only polymorphic loci in our calculations of D , our results can be compared to those of Lundqvist and Rudin (1977) who used four polymorphic loci in studying *Picea abies* (L.) Karst. populations throughout Sweden, and those of Levin (1977) who used seven polymorphic loci in studies of *Phlox drummondii* Hook. subspecies throughout Texas. Lundqvist and Rudin reported D values of 0.002 to 0.039 between native populations. Levin reported a mean distance, $\bar{D}$ of 0.02 between all subspecies.

The Boulder Canyon clusters had a $\bar{D}$ of 0.015, with a range of 0.004 to 0.035; Glacier Lake had a $\bar{D}$ of 0.022, with a range of 0.006 to 0.055.

Genetic Differentiation Between Age Classes

The Shanahan Mesa population provides detailed evidence of how selection pressures within a population can change during a few years as a result of changes in the environment. Such changes in selection pressures produced significant genetic differentiation between age classes. The analysis was with the PER-2 locus, which has three alleles. Other studies have shown that allele 2 is found most frequently in the moister, cooler conditions prevalent at higher elevations (above 2400 m) and north-facing slopes. Alleles 1 and 3 are found most frequently in the warmer, more xeric conditions prevalent at lower elevations (below 1800 m) and on south-facing slopes (Mitton and others 1977). Also, 23 heterozygotes occur in frequencies significantly higher than those predicted by Hardy-Weinberg expectations in all populations growing in relatively warm, xeric conditions. Beckman (1977) studied the kinetics of the allozymes and found that the biochemical evidence was consistent with results from field observations: a 22 genotype produces an enzyme that functions best at cool temperatures; 13 and 33 genotypes produce enzymes that function best at warm temperatures; and a 23 genotype functions well over a broader range of temperatures than do the other genotypes.

At Shanahan Mesa, two well-defined size and age classes exist. In certain areas dominant trees that form the canopy are more than 40 years old, have colonized the open grasslands, and have had to survive in the relatively xeric conditions of these grasslands. Beneath them grow their offspring—trees that developed in the much moister,

cooler conditions provided by a closed canopy. Under this canopy, snow accumulates and can remain for several weeks. In contrast, the snowfalls on the adjacent prairie are blown away by frequent winds or melted by the sun within days. Significant differences exist between these two groups at the PER-2 locus. The old colonizers have significantly higher frequencies of alleles 1 and 3 (0.200) than do the younger trees (0.060); this difference is significant ($X^2 = 6.40$, $P < 0.025$). In addition, there are significant excesses of heterozygotes in the old colonizers, whereas genotypic frequencies in the young trees conform to Hardy-Weinberg expectations (Beckman 1977). The associations between allele 2 and cool conditions, and alleles 2 and 3 and hot, xeric conditions are so consistent, that we believe the results at Shanahan Mesa truly reflect genetic differentiation in response to environmental conditions that changed during the span of 120 years. In the Boulder Canyon population, the genetic make-up of the major age classes was compared at seven loci. There were no striking genetic differences between them.

Genetic Aspects of Differential Reproduction

Natural selection results in part from differential reproduction of genotypes (Darwin 1859), but only a few published reports document differential reproduction of genotypes in natural populations (Eanes and others 1977, Clegg and others 1978). We wish to know if reproduction in individual trees is related to their allozyme genotypes. Because it is difficult to estimate accurately total pollen production for a population, we compared the genetic constitution of trees that produced pollen with those that did not. Three loci—PER-2, F.E. and C.E.—were analyzed in the Boulder Canyon and Glacier Lake populations, and PER-2 only was analyzed in a third population near Glacier Lake. Results were the same for all comparisons: no genetic differences were detected between trees producing pollen and those that did not in 1977 (Linhart and others 1979b).

Female cone production can be measured accurately in the trees we study because the stands are open, and the trees are no taller than 15 m. Cone production is affected by both tree size and age in conifers (Schubert 1974, Dorman 1976). To describe cone production as a function of age and size, we performed a multiple regression analysis of cone production, using age and diameter of trees as independent variables. To get an accurate estimate of differential cone production in the Boulder Canyon population, we compared the trees whose cone production deviated most strikingly from that expected for their diameter and age. We compared the 50 trees that most exceeded their expected cone production to the 50 trees that were farthest below expectations. Results show that in 1977, a mast year, significant differences existed at several loci between high cone producers and low cone producers (table 2).

That differential cone production in trees has a significant genetic component is not unexpected. Much

agricultural research on selection for higher yields in plants and animals has underscored the importance of this genetic component (Allard 1960). Recent studies using allozyme markers have usually (Schall and Levin 1976, Eanes and others 1977, Clegg and others 1978), though not always (Christiansen and Frydenberg 1976), found genetic differences between reproducing and nonreproducing individuals. However, this genetic component needs more recognition in forestry research: a recent exhaustive review of research on reproduction in southern pines (Dorman 1976) shows that little of that research has focused on the genetic and evolutionary implications of differential reproduction.

Animal Damage and Its Genetic Implications

Patterns of deer browsing in 1977 and the presence of woolly aphids in 1979—the only years for which we have analyzed data so far—have distributions that are affected by the genetic constitutions of trees in the Boulder Canyon population. A comparison of the genetic constitution of heavily browsed trees with that of unbrowsed or lightly browsed trees at three loci showed that, although at PER-2 and F.E. we found no statistically significant differences, at C. E. the frequencies of alleles 1 and 2 combined were 0.791 in the heavily browsed and 0.672 in the lightly browsed trees ($X^2 = 4.635$, $P < 0.05$). In addition, a preference by deer for particular genotypes has been shown in *Pseudotsuga menziesii* (Mirb.) Franco (Radwan 1972).

Woolly aphids are common on ponderosa pine in the Central Rocky Mountains. We scored all trees in the

Boulder Canyon population for presence or absence of woolly aphids by sampling 40 branches per tree. Trees with aphids ($N = 51$) were compared with trees without aphids ($N = 166$) in the frequencies of the most common allele and of all other alleles pooled together. Significant differences were found at the F. E. locus. The frequencies of allele 2 was 0.500 in trees with aphids, and 0.651 in noninfested trees ($X^2 = 6.82$, $P < 0.01$). These data suggest that patterns of animal attacks on trees are not random. Similar results have been obtained in ponderosa pine by Edmunds and Alstadt (1978) in a study of the black pine leaf scale, *Nuculapsis californica*, and by Sturgeon (1979) in a study of the bark beetle, *Dendroctonus brevicomis*.

Synthesis

The genetic clustering reported in the species cited in this report presents a different picture from the one observed in some other conifers studied to date. For example, no subdivision has been detected in *Pinus rigida* Mill. (Guries and Ledig 1977) or *Pinus taeda* L.⁵ Rudin and others (1977) studying *Pinus sylvestris* L. and Tigerstedt (1973) studying *Picea abies* (L.) Karst. show that in these species the subdivision is not very pronounced. These differences may result, in part, from study techniques: our study at Boulder Canyon involves mapping of all individuals established within a given area, with a sample size of more than 200 individuals; the studies cited earlier typically involved selected individuals and sample sizes of 60 to 70 trees. More importantly, the differences probably reflect real differences in biological features of the species concerned. Ponderosa pine in the Rocky Mountains often grows in all-aged stands which are composed of patches of relatively uniform size and age (Schubert 1974). In species where such mosaics do not occur, and where stands are primarily single-aged and single-sized, less opportunity for such differentiation exists. This may be particularly true in species such as *Pinus rigida* Mill. that depend on fire for their establishment, so that extensive stands get started under similar conditions.

Because heterogeneity of the physical environment is limited within a site, causes of the heterogeneity must be sought elsewhere. We believe that clusters are probably composed of genetically related individuals, given the short dispersal distances of both tree seeds (Levin and Kerster 1974) and viable pollen (Coles and Fowler 1976, Muller 1977). Differentiation among clusters may also occur if different clusters, which have different age structures (*fig. 1*), come from genetically-different parents, are exposed to different selective pressures in their early years, or both. Comparisons between major age classes in ponderosa pine (Beckman 1977) have shown major age-associated genetic differences within some populations.

Animal predation may also contribute to the genetic structure of the Boulder Canyon population. Both deer and woolly aphids show preferences for some phenotypes, and it is likely that other animals with more destructive habits could practice similar discrimination.

Table 2—Comparisons between individuals with low and high relative cone productions in a population of ponderosa pine, at the Boulder Canyon population

Locus	Allele	Cone production		Chi square	Significance
		Low F $\pm$ S.E.	High F $\pm$ S.E.		
PER	1	0.07 $\pm$ 0.026	0.01 $\pm$ 0.010	5.5	$P < 0.07$
	2	.71 $\pm$.045	.70 $\pm$.046		
	3	.221 $\pm$.041	.29 $\pm$.045		
GLU	1	.66 $\pm$.047	.74 $\pm$.044	1.6	$P < .50$
	2	.34 $\pm$.047	.26 $\pm$.044		
F.E.	1	.20 $\pm$.040	.31 $\pm$.046	6.6	$P < .05$
	2	.74 $\pm$.044	.57 $\pm$.050		
	3	.06 $\pm$.024	.12 $\pm$.035		
PHI	1	.01 $\pm$.010	.01 $\pm$.010	1.4	$P < .50$
	2	.98 $\pm$.014	.99 $\pm$.010		
	3	.01 $\pm$.010	.00		
C.E.	1	.67 $\pm$.047	.82 $\pm$.038	5.9	$P < .05$
	2	.05 $\pm$.044	.03 $\pm$.017		
	3	.28 $\pm$.045	.15 $\pm$.035		
PGM-1	1	.03 $\pm$.017	.08 $\pm$.027	2.5	$P < .50$
	2	.97 $\pm$.017	.92 $\pm$.027		
PGM-2	1	.09 $\pm$.029	.06 $\pm$.024	2.2	$P < .50$
	2	.90 $\pm$.030	.94 $\pm$.024		
	3	.01 $\pm$.010	.00		

The extent of differentiation in allele frequencies observed between clusters within the Boulder Canyon ($\bar{D} = 0.015$) and Glacier Lake populations ($\bar{D} = 0.022$) are similar in magnitude to the differentiation between the Boulder Canyon and Glacier Lake populations ($\bar{D} = 0.030$). These populations are separated by more than 1000 m in elevation and several kilometers in their linear distance. This suggests that a substantial proportion of total genetic variation can be found within populations of ponderosa pine and supports the findings of Madsen and Blake (1977) and O'Malley and others (1979). Similar results for other conifers have been obtained by Yeh (1981), Lines and Mitchell (1966), and Fins (1979). However, this observation cannot be applicable to all classes of genetic variation, otherwise we would not have witnessed the repeated disasters produced by planting seeds of far-away sources in a given area; nevertheless, it probably applies to at least some of the variability that forest geneticists are intent upon capturing.

CONCLUSIONS

The populations of ponderosa pine we have studied show a great deal of intra-population differentiation which is not obviously associated with heterogeneity of the physical environment. Similar results in other herbaceous and woody species suggest that such differentiation is a common feature of the genetic structure of plant populations.

The highly localized heterogeneity we have demonstrated, along with the genetic differences observed between heavy and light cone bearers have practical implications. Whenever seed collections or regeneration after logging depend on a few trees with large numbers of cones, the genetic base of the following generations may be severely restricted. Genetic analyses, provenance studies, and seed orchards cannot produce accurate pictures of total genetic variation if they depend only on progenies from trees producing many seeds in a given year.

Acknowledgments

We thank Joseph Beckman, Diane Bowman, Karen Jerger, Jan Logan, Mary Jo Tesitor, Dorothea Slater, and Charles Nilon for technical assistance. The research is supported by grants BMS 75-14050 and DEB 78-16798 from the National Science Foundation, and by grants from the University of Colorado.

LITERATURE CITED

- Allard, R. W.
1960. *Principles of plant breeding*. 485 p. Wiley, New York.
- Barber, H. N.
1965. *Selection in natural populations*. *Heredity* 20:551-572.
- Beckman, J. S.
1977. *Adaptive peroxidase allozyme differentiation between colonizing and established populations of Pinus ponderosa Laws. on the Shanahan Mesa, Boulder, Colorado*. M.A. thesis. Univ. Colorado, Boulder, Colo.
- Benson, L., E. A. Phillips, P. A. Wilder, and others.
1967. *Evolutionary sorting of characters in a hybrid swarm. I. Direction of slope*. *Amer. J. Bot.* 54:1017-1026.
- Christiansen, F. B., and O. Frydenberg.
1976. *Selection component analysis of natural polymorphisms using mother-offspring samples of successive cohorts*. In *Population genetics and evolution*. S. Karlin and E. Nevo, eds. p. 277-302. Academic Press, New York.
- Clegg, M. T., A. L. Kahler, and R. W. Allard.
1978. *Estimation of life cycle components of selection in an experimental plant population*. *Genetics* 89:765-792.
- Coles, J. F., and D. P. Fowler.
1976. *Inbreeding in neighboring trees in two white spruce populations*. *Silvae Genetica* 25:29-34.
- Darwin, Charles.
1859. *On the origin of species by means of natural selection*. 502 p. Murray, London.
- Dorman, K. W.
1976. *The genetics and breeding of southern pines*. U.S. Dep. Agric., Agric. Handb. 471, 407 p.
- Eanes, W. F., P. M. Gaffney, R. K. Koehn, and C. M. Simon.
1977. *A study of sexual selection in natural populations of the milkweed beetle Tetraopes tetraophthalmus*. In *Measuring selection in natural populations*. F. B. Christiansen and T. M. Fendel, eds. Springer-Verlag, New York.
- Edmunds, G. F., and D. N. Alstadt.
1978. *Coevolution in insect herbivores and conifers*. *Science* 199:941-945.
- Epling, C., and T. Dobzhansky.
1942. *Genetics of natural populations. VI. Microgeographical races in Linanthus parryae*. *Genetics* 27:317-332.
- Feeny, P. P.
1975. *Biochemical coevolution between plants and their insect herbivores*. In *Coevolution of animals and plants*. L. E. Gilbert and P. H. Raven, eds. p. 3-18. Univ. Texas Press, Austin, Texas.
- Fins, L.
1965. *Genetic architecture of giant sequoia*. Ph.D. Dissertation. Univ. California, Berkeley. 255 p.
- Fowler, D. P.
1965. *Effects of inbreeding in red pine, Pinus resinosa Ait. III. Factors affecting natural selfing*. *Silvae Genetica* 14:37-46.
- Gan, Y. Y., F. W. Robertson, P. S. Ashton.
1977. *Genetic variation in wild populations of rain-forest trees*. *Nature* 269:323-324.
- Grant, M. C., and J. B. Mitton.
1977. *Genetic differentiation among growth forms on Engelmann spruce and subalpine fir at tree line*. *Arct. Alp. Res.* 9:259-263.
- Guries, R. P., and F. T. Ledig.
1977. *Analysis of population structure from allozyme frequencies*. *Proc. Southern Forest Tree Improv. Conf.* 14:246-253.
- Guries, Raymond P., and F. Thomas Ledig.
1981. *Genetic structure of populations and differentiation in forest trees*. In *Proceedings of symposium on isozymes of North American forest trees and forest insects*, July 27, 1979, Berkeley, Calif. Gen. Tech. Rep. PSW-48, p. 42-47. Pacific Southwest Forest and Range Exp. Stn., Forest Serv., U.S. Dep. Agric., Berkeley, Calif.
- Harper, J. L.
1977. *Population biology of plants*. 892 p. Academic Press, New York.
- Hermann, R. H., and D. P. Lavender.
1967. *Early growth of Douglas-fir from various altitudes and aspects in southern Oregon*. *Silvae Genetica* 16:143-151.
- Kafton, D.
1977. *Isozyme variability and reproductive phenology of Monterey cypress*. Ph.D. Dissertation. Univ. California, Berkeley. 157 p.
- Langlet, O.
1971. *Two hundred years genecology*. *Taxon* 20:653-722.
- Ledig, F. T.
1974. *An analysis of methods for the selection of trees from wild stands*. *Forest Sci.* 20:2-18.
- Levin, D. A.
1977. *The organization of genetic variability in Phlox drummondii*. *Evolution* 31:477-484.

- Levin, D. A., and H. W. Kerster.
1974. **Gene flow in seed plants.** *Evol. Biol.*, p. 139-220.
- Libby, W. J., R. F. Stettler, and F. W. Seitz.
1969. **Forest genetics and forest-tree breeding.** *Ann. Rev. Genet.* 3:469-494.
- Lines, R., and R. F. Mitchell.
1966. **Differences in phenology of Sitka spruce provenances.** Forest Research Report. p. 173-184, G. B. Forestry Commission.
- Linhart, Y. B., J. B. Mitton, D. M. Bowman, K. B. Sturgeon, and J. L. Hamrick.
1979a. **Genetic aspects of fertility differentials in ponderosa pine.** *Genet. Res.* 33:237-242.
- Linhart, Y. B., J. B. Mitton, D. M. Bowman, K. B. Sturgeon, and J. L. Hamrick.
1979b. **Some genetic consequences of differential reproduction in populations of forest trees.** In IUFRO, Flowering and seed development in forest trees. F. T. Bonner, ed. p.41-50. Southern Forest Exp. Stn. Starkville, Miss. 380 p.
- Lundkvist, K., and D. Rudin.
1977. **Genetic variation in 11 populations of *Picea abies* in Sweden as determined by isozyme analysis.** *Hereditas* 85:67-74.
- Madsen, J. L., and G. M. Blake.
1977. **Ecological genetics of ponderosa pine in the Northern Rocky Mountains.** *Silvae Genetica* 26:1-8.
- Mitton, J. B., Y. B. Linhart, J. L. Hamrick, and J. Beckman.
1977. **Population differentiation and mating system in ponderosa pine of Colorado Front Range.** *Theor. Appl. Genet.* 51:5-14.
- Mitton, J. B., Y. B. Linhart, K. B. Sturgeon, and J. L. Hamrick.
1979. **Allozyme polymorphisms detected in mature needle tissues of ponderosa pine.** *J. Heredity* 70:86-89.
- Muller, G.
1977. **Cross-fertilization in a conifer stand inferred from enzyme gene-markers in seeds.** *Silvae Genetica* 26:223-226.
- O'Malley, D. M., F. W. Allendorf, and G. M. Blake.
1979. **Inheritance of isozyme variation and heterogeneity in *Pinus ponderosa*.** *Biochem. Genet.* 17:233-250.
- Nei, M.
1972. **Genetic distance between populations.** *Amer. Nat.* 106:283-292.
- Radwan, M. A.
1972. **Differences between Douglas-fir genotypes in relation to browsing preference by black-tailed deer.** *Can. J. For. Res.* 2:250-255.
- Rehfeldt, G. E.
1978. **The genetic structure of a population of Douglas-fir (*Pseudotsuga menziesii* var. *glauca*) as reflected by its wind-pollinated progenies.** *Silvae Genetica* 27:49-52.
- Rhoades, D. F., and R. G. Cates.
1976. **Toward a general theory of plant antiherbivore chemistry.** *Recent Adv. Phytochem.* 10:168-213.
- Rudin, D., G. Eriksson, and M. Rasmuson.
1977. **Inbreeding in a seed trees stand of *Pinus sylvestris* L. in northern Sweden. A study by the aid of the isozyme technique.** Inst. for Skogsgenetik Res. Note 25, Swedish University of Agricultural Sciences, Umeå, Sweden.
- Schaal, B. A.
1975. **Population structure and local differentiation in *Liatris cylindracea*.** *Amer. Nat.* 109:511-528.
- Schaal, B. A., and D. A. Levin.
1976. **The demographic genetics of *Liatris cylindracea* Michx.** *Amer. Nat.* 110:191-206.
- Schubert, G. H.
1974. **Silviculture of southwestern ponderosa pine: The status of our knowledge.** USDA Forest Serv. Res. Paper. RM-123, 71 p. Rocky Mountain Forest and Range Exp. Stn., Fort Collins, Colo.
- Shaw, D. V., and R. W. Allard.
1981. **Analysis of mating system parameters and population structure in Douglas-fir using single-locus and multilocus methods.** In Proceedings of symposium on isozymes of North American forest trees and insects, July 27, 1979, Berkeley, Calif. Gen. Tech. Rep. PSW-48, p. 18-22. Pacific Southwest Forest and Range Exp. Stn., Forest Serv., U.S. Dep. Agric., Berkeley, Calif.
- Stern, K., and L. Roche.
1974. **Genetics of forest ecosystems.** 330 p., Springer-Verlag, New York.
- Sturgeon, K. B.
1979. **Monoterpene variation in ponderosa pine xylem resin related to western pine beetle predation.** *Evolution* 33:803-814.
- Tigerstedt, P. M. A.
1973. **Studies on isozyme variation in marginal and central populations of *Picea abies*.** *Hereditas* 75:47-60.
- Workman, P. L., and J. D. Niswander.
1970. **Population studies on southwestern Indian tribes. II. Local differentiation in the Papago.** *Amer. J. Human Genet.* 22:24-49.
- Yeh, Francis C.
1981. **Analyses of gene diversity in some species of conifers.** In Proceedings of symposium on isozymes of North American forest trees and insects, July 27, 1979, Berkeley, Calif. Gen. Tech. Rep. PSW-48, p. 48-52. Pacific Southwest Forest and Range Exp. Stn., Forest Serv., U.S. Dep. Agric., Berkeley, Calif.

Applying Isozyme Analyses in Tree-Breeding Programs¹

W. T. Adams²

Abstract: Four examples illustrate the potential for practical use of isozyme analyses in applied breeding programs. These include identifying parent trees and clones, seed sources, and parentage of controlled crosses, and evaluating the effectiveness of different procedures involving open-pollination to produce seed of specific crosses. The improved ability to assess the true identity of parents, seed stocks, and controlled crosses justifies the added expense resulting from the use of isozyme analyses in applied tree-breeding programs.

Some papers in this symposium, as well as others previously published (Bergmann 1978, Guries and Ledig 1977, Lundkvist and Rudin 1977, Mitton and others 1977, Muller 1977, Rudin and others 1974), demonstrate the use of isozymes in studying genetic structure and variation patterns in populations of forest trees. Recently, isozyme investigations also have been started in seed orchards³ (Rudin and Lindgren 1977). Information from studies in both natural and seed orchard populations no doubt will help to increase the efficiency of future tree-improvement programs.

In addition to their research value in forest genetics, isozymes have more immediate application as genetic markers in tree-breeding programs. By providing examples primarily from my own work, I illustrate several possible applications. These examples demonstrate that isozyme analyses can be used by tree breeders to determine and maintain the identity of individuals, crosses, and breeding populations with greater accuracy than was possible previously.

IDENTIFYING PARENT TREES AND CLONES

If isozymes are to be effective genetic markers in breeding programs (Rudin and Lindgren 1977), isozyme variation among parent trees or orchard clones must be large. This is expected in most forest tree species because an effort is usually made to incorporate a broad base of genotypes in breeding populations. The large isozyme variation found among clones in one pitch pine (*Pinus rigida* Mill.) breeding orchard and three loblolly pine (*Pinus taeda* L.) seed orchards³ (Hunter 1977) supports this expectation (table 1). As many as nine alleles were present

at any one isozyme locus, and the average percentage of heterozygous loci per clone ranged from 15 to 33 percent.

For clonal identification it is preferable to have many alleles per locus, each of about equal frequency. However, although isozyme loci in these four orchards (table 1) averaged about three alleles, one allele was usually in high frequency. Isozyme variation among these clones, therefore, was far from "ideal." Nevertheless, enough variation among clones existed so that most could be individually identified by using the sampled isozymes as genetic markers. On the basis of the 11 polymorphic loci scored in each clone, 28 unique genotypes were recognized among the 32 clones in the pitch pine orchard; similarly, 22 twelve-locus genotypes and 25 eleven-locus genotypes were recognized, respectively, among the 23 and 27 clones in the two South Carolina loblolly orchards. All 27 clones in the North Carolina loblolly orchard could be uniquely identified on the basis of their twelve-locus genotypes. Complete resolution of clones in the first three orchards could be achieved if additional loci were evaluated. These results indicate that isozyme gene markers can be useful for maintaining the identity of ortets and their ramets in a breeding program.

Hunter (1977), for example, has used isozymes to detect labeling errors in seed orchard trees. Six of the 22 clones with more than one ramet in the North Carolina loblolly orchard had at least one mislabeled ramet, discovered because the genotypes of the mislabeled ramets differed from those of the other ramets representing a clone.

The genotype of each orchard clone was determined from seed samples on the basis of the composition of allozymes (allelic isozymes) observed in megagametophytes. When this method is used, a genotype will be incorrectly identified only if all megagametophytes sampled from a heterozygote carry the same allozyme allele; with N megagametophytes, this probability is $(1/2)^{N-1}$ for a single locus. For a sample of only 10 seeds, therefore, the probability of incorrectly identifying a clone's genotype at any one locus is less than 0.2 percent. Assuming all loci are independent, the probability of error in identifying a 12-locus genotype is less than 2.5 percent.

In general, extracting enzymes from seed tissues is easier than from vegetative tissues (for example, needles, buds, or bark) of seedlings or older trees. However, analyzing

¹ Presented at the Symposium on Isozymes of North American Forest Trees and Forest Insects, July 27, 1979, Berkeley, Calif. (Paper FRL 1389, Forest Research Laboratory, Oregon State University, Corvallis, Oreg.)

² Assistant Professor of Forest Genetics, Oregon State University, Corvallis, Oreg.

³ Adams, W.T., and R.J. Joly. Allozyme studies in loblolly pine seed orchards: clonal variation and frequency of progeny due to self-fertilization. (In press.)

Table 1—*Variation at isozyme loci polymorphic among clones in one pitch pine breeding orchard and three loblolly pine seed orchards*

Orchard ¹	Number		Alleles per locus		Most frequent allele at each locus		Heterozygous loci per clone		
	Clones	Loci	Most	Mean	Range	Mean	Range	Mean	Mean
							<i>Number</i>	<i>Percent</i>	
Pitch pine	32	11	6	2.9	0.98 to 0.70	0.91	0 to 4	1.7	15
Loblolly pine ²									
South Carolina-LSG	23	12	6	3.0	.98 to .46	.82	1 to 5	3.0	25
Loblolly pine ²									
South Carolina-HSG	27	11	5	3.0	.98 to .44	.81	0 to 6	2.9	26
Loblolly pine ³									
North Carolina	27	12	9	3.3	.96 to .26	.73	1 to 8	4.0	33

¹The two South Carolina orchards are located at the same site. The low specific gravity (LSG) orchard consists of clones selected for wood of low specific gravity; the high specific gravity (HSG) orchard consists of clones selected for wood of high specific gravity.

²Data from Adams and Joly. (In press.)

³Data from Hunter (1977).

isozymes from vegetative tissues is necessary if genotypes of juvenile and nonflowering mature individuals are to be identified. A set of isozyme loci expressing clearly in both seed and vegetative tissues would be of most benefit; several examples of such loci have been reported (Bartels 1971, Lundkvist 1975, Mitton and others 1977, Rudin 1975).

SEED CERTIFICATION

Interest in seed certification will intensify as demand for genetically improved seed increases (Rudolph 1974). Because considerable isozyme variation has been observed among orchard clones, the seed of individual clones or that from specific crosses can be distinguished from the seed of most other clones or crosses. Isozymes can be used, therefore, to certify the identity of specific seed lots.

The expected frequencies of isozymes in the seed of the two South Carolina loblolly pine orchards were determined by weighting each clonal genotype according to the number of its ramets and assuming random mating (*table 2*).³ Genetic makeup of seed lots from these orchards was expected to be similar because the clones came from parent trees in the same breeding zone and, except for wood density, were selected for the same qualities. This was generally borne out—at most loci, the most frequent allele in one orchard was also the most frequent in the other. Differences between orchards were also observed, however; some alleles had dissimilar frequencies, and some occurred in one orchard but not in the other. Therefore, one could readily determine from which orchard a particular seed lot was derived.

In several species, evidence is accumulating to indicate that variation in isozyme frequencies between provenances can be expected (Bergmann 1978, Lundkvist and Rudin 1977, Mitton and others 1977, Rudin and others 1974). Thus, even greater differences in isozymes may occur among seed lots from orchards composed of clones from various geographical areas. Similarly, isozymes could certify the geographical origin of seed collected from

natural stands (Bergmann 1972). Changes in isozyme frequency may be strongly correlated with changes in environment (Bergmann 1978, Grant and Mitton 1977, Mitton and others 1977), in which case isozymes can be helpful in defining breeding and seed zones (Feret and Bergmann 1976).

Table 2—*Expected allelic frequencies in the seed crops of two South Carolina loblolly pine seed orchards*

Locus	Allele ²	Orchards		Locus	Allele	Orchard	
		LSG	HSG			LSG	HSG
PGI1	1	0.015	0.015	LAP2	1	0.417	0.503
	2	.985	.985		2	.583	.450
					3	—	.047
GOT1	1	.156	.007	PGI2	1	.047	—
	1L	— ³	.025		2	.822	.930
	2	.184	.362		3	.108	.057
AP2	3	.660	.606	GOT2	4	.023	.013
	1	.015	.052		1	.168	.098
	2	.889	.938		2	.007	—
MDH2	3	.096	.010	6PGD	3	.829	.902
	1	.905	.993		1	.025	—
	2	.095	.007		2	.397	.407
LAP1					3	.053	.015
	1	.065	—		4	.015	.085
	2	.836	.862		5	.503	.481
GDH	3	.048	.045	PGM1	5L	.007	—
	N	.051	.093		6	—	.012
	1	—	.038		1	.944	.978
	2	.939	.787		2	.056	.022
	3	.013	—				
	4	—	.010	PGM2	1	.028	—
	5	.048	.165		2	.972	1.00

¹ LSG is low specific gravity; HSG is high specific gravity. (From Adams and Joly. In press.)

²Alleles are numbered in ascending order from the fastest to the slowest migrating isozymes. N is a null allele; L is an allele coding a very lightly staining band.

³A dash (—) means no allele present.

DETERMINING VALIDITY OF CONTROLLED CROSSES

Tree-breeding programs usually require many controlled crosses, which generally are assumed to have been made without error. This assumption, however, has not been tested because no accurate method has been available. Using isozyme markers to test the validity of controlled crosses is relatively simple (if sizable isozyme variability exists among clones or parent trees): first, parental isozyme genotypes are determined, and second, their progeny are analyzed to see whether their isozyme makeup is as expected.

To demonstrate this application, we conducted the following study. Seeds were supplied from 30 two-parent loblolly pine crosses made by three different organizations (designated A, B, C) in the North Carolina State-Industry Cooperative Tree Improvement Program (*table 3*). Additional seed lots from all the clones involved as male parents and from clones of the female parents used in one-half of the 30 crosses also were obtained. Megagametophytes of at least 10 seeds from each parental seed lot were analyzed and the isozyme genotypes of the clones inferred at six loci. Embryo and megagametophyte tissues from 20 to 30 seeds of each of the 30 controlled crosses were then assayed at the same six loci. By analyzing both the haploid and diploid seed tissues, it was possible to determine the allelic composition of the ovule and pollen gametes forming the embryo. Because the allelic composition of the ovule is the same as that in the megagametophyte, the remaining allele (after accounting for the allele contributed by the ovule) must be from the pollen gamete. The pollen pools (that is, pollen gametes effective in fertilizing viable seeds) of the crosses, therefore, can be compared with those expected on the basis of each pollen parent's isozyme genotype; similar comparisons can be made for the ovule pools of the 15

crosses for which independent information on the genotypes of the female parents is available.

Evidence of error in female identity was found in only one (cross 18 X 19) of the 30 crosses. Seed from clone 18 had apparently been mixed with seed from one or more other clones, because five of the 20 seeds sampled contained alleles in megagametophytes that could not have come from clone 18. At least some of the contaminated seed could have come from clone 19; this clone has alleles at two loci not found in any of the other sampled clones, and both of these alleles were found in the contaminated seed. The allelic composition of megagametophytes was as expected in the remaining 14 crosses for which the isozyme genotypes of the maternal parents were known, and no more than two alleles per locus were observed among the megagametophytes of the other crosses sampled.

Pollen contamination seemed to be a real problem in organization C (*table 3*). Although the allelic compositions of the pollen pools were as expected in all but one of the 20 crosses made by organizations A and B, 8 out of 10 crosses for organization C contained alleles that could not have come from the supposed pollen parent. Apparently, procedural problems in pollen handling, bagging, or making the crosses resulted in considerable pollen contamination.

Caution is needed in comparing the validity of crosses made by different organizations when using the technique of isozyme markers. The ability to detect errors is somewhat dependent upon parents carrying alleles at one or more marker loci that are relatively infrequent among other sources of pollen. If the variation among clones is generally large, however, the results from different organizations should be roughly comparable. The extent of contamination is underestimated with this technique because contaminants having the same allelic compositions as those expected would remain undetected. Nevertheless, isozyme gene markers will probably uncover most errors in seed-handling and control-crossing procedures.

Table 3—Number of detected pollen contaminants (d.c.)¹ observed in the pollen pools of seeds from 10 controlled two-parent crosses of loblolly pine from each of three organizations

Organization A			Organization B			Organization C		
Cross		Observed pollen gametes	Cross		Observed pollen gametes	Cross		Observed pollen gametes
♀ ♂	d.c.	Total	♀ ♂	d.c.	Total	♀ ♂	d.c.	Total
1 X 2	0	29	15 X 16	0	29	² 31 X 32	5	30
² 3 X 2	0	30	17 X 16	0	28	² 33 X 32	7	30
² 4 X 5	0	30	² 18 X 19	0	20	² 34 X 31	0	30
6 X 5	0	27	20 X 21	0	22	² 35 X 31	1	30
7 X 8	0	28	15 X 22	0	30	² 36 X 35	2	30
9 X 8	0	28	23 X 18	0	30	² 3 X 35	0	30
10 X 11	0	28	24 X 25	0	29	² 3 X 37	2	30
12 X 11	0	26	² 26 X 27	0	30	² 38 X 39	23	30
13 X 14	0	30	28 X 27	12	30	² 40 X 41	16	30
7 X 14	0	30	² 29 X 30	0	21	² 30 X 39	14	30

¹ Detected pollen contaminants (d.c.) are pollen gametes with an allele at one or more of six isozyme loci *not* carried by the supposed pollen parent involved in the cross.

² Those clones used as female parents whose genotypes at six isozyme loci were determined from megagametophytes of an independent seed sample (see text for details).

EFFECTIVENESS OF OPEN-POLLINATION TO PRODUCE SEED OF SPECIFIC CROSSES

To avoid costly control-pollinations, alternate methods for producing hybrid seed or seed of outstanding specific combinations have been sought. Among the possibilities are wind-pollinated two-clone (or two-species) seed orchards (Hyun 1976, Wright 1976) or supplemental mass pollination (Little and Trew 1977, Wakeley and others 1965, Woessner and Franklin 1973). If the isozyme makeup of parents is sufficiently different, crosses between and within parents can be distinguished. Isozymes, therefore, can help to determine the effectiveness of such alternatives.

The isozyme approach was used to estimate the proportion of hybrid seed obtained when unbagged female flowers of pitch pine were mass pollinated with pollen mixes from loblolly pine parents.⁴ An isozyme gene marker was found in high frequency in the pollen mix but was rare in any pitch pine source. On the basis of the proportion of pollen pool gametes of mass-pollinated clones carrying this marker, the proportion of hybrid seed was estimated to range from 2 to 42 percent (averaging 16 percent) for six clones. A nearly identical mean percentage (17 percent) of hybrids was found for the progeny of these clones by scoring 1-year-old seedlings in the nursery into hybrid and nonhybrid classes on the basis of their stem and foliage characteristics.⁵ Because the isozyme technique determines the extent of hybrid seed set before nursery establishment, only the seed of clones with high hybrid seed percentages need be sown.

DISCUSSION AND CONCLUSIONS

The preceding examples are but a few of the many possible uses of isozymes in applied tree-breeding programs. Although these applications can be invaluable to the tree breeder, their practical implementation will depend on a cost-benefit analysis. The ability to better determine and maintain the true identity of breeding materials can only be assessed relative to the probability of errors under current breeding technology. As seen in the controlled crosses, this may vary widely among organizations. Even if error in the handling of breeding materials is now relatively small, its effect on reduction of genetic gains is likely to be magnified in advanced generation breeding. Maintaining the identity of breeding materials, therefore, will become increasingly important in the future.

Although costs of isozyme analyses will vary substantially among laboratories, it is helpful to provide some

rough estimates. For instance, I have found that processing from 100 to 150 seed tissue extracts by using horizontal starch gel electrophoresis (Conkle 1972) on 1 working day is about maximum for a full-time technician with a half-time assistant to help prepare plants, wash dishes, and do other required tasks. A laboratory with this capability can be equipped for less than \$15,000 (including large glass door refrigerator, power supplies, incubator, balance, pH meter, plexiglas gel forms and buffer trays, and miscellaneous glassware and supplies). This estimate assumes a standard laboratory with piped-in distilled water.

For the enzymes normally analyzed, our daily cost of chemicals and supplies averages about \$40. With 1 day's run, we can assess the genetic makeup of 9 to 14 parent trees at 10 to 15 loci or analyze from 25 to 30 progeny each from 1.5 to 2.5 controlled crosses for the same number of loci. We can genetically identify a clone for \$3 to \$5 and determine the validity of a controlled cross for \$15 to \$25. Although these figures do not include labor, utilities, or depreciation on equipment, at even double the costs, they seem fairly reasonable compared with other expenses normally incurred in a tree-improvement program (Porterfield 1974).

I am not suggesting that it is economically feasible or prudent for most individual organizations to establish their own electrophoresis laboratories at this time. The potential benefits of isozyme analyses in breeding programs, however, seem great enough for research units of larger organizations and tree-improvement cooperatives to consider establishing such laboratories soon. Because of the long generation times faced by most tree breeders, we can hardly afford not to seriously consider isozyme analyses for increasing breeding efficiency.

LITERATURE CITED

- Bartels, Horst.
1971. **Genetic control of multiple esterases from needles and macrogametophytes of *Picea abies***. *Planta* 99:283-289.
- Bergmann, Fritz.
1972. **Experiments on possibilities for genetic certification of forest seed**. In Proc., IUFRO and SABRO joint meeting [Oct. 5-20, 1972, Tokyo, Japan], A-6(V):1-5.
- Bergmann, F.
1978. **The allelic distribution at an acid phosphatase locus in Norway spruce (*Picea abies*) along similar climatic gradients**. *Theor. Appl. Genet.* 52:57-64.
- Conkle, M. Thompson.
1972. **Analyzing genetic diversity in conifers ... isozyme resolution by starch gel electrophoresis**. USDA Forest Serv. Res. Note PSW-264, 5 p. Pacific Southwest Forest and Range Exp. Stn., Berkeley, Calif.
- Feret, P. P., and F. Bergmann.
1976. **Gel electrophoresis of proteins and enzymes**. In *Modern methods in forest genetics*. p. 49-77. J. P. Miksche, ed. Springer-Verlag, New York.
- Grant, Michael C., and Jeffrey B. Mitton.
1977. **Genetic differentiation among growth forms of Engelmann spruce and subalpine fir at tree line**. *Arct. Alp. Res.* 9:259-263.

⁴Joly, R.J., and W.T. Adams. Allozyme estimation of the effectiveness of mass-artificial pollination for the production of pitch x loblolly pine hybrid seed. Manuscript in preparation.

⁵I.F. Trew, Westvaco Corporation, unpublished data, on file at Westvaco Corporation, Crozet, Virginia.

- Guries, R. P., and F. T. Ledig.
1977. **Analysis of population structure from allozyme frequencies.** *In* Proc., 14th Southern Forest Tree Improv. Conf. [June 14-16, 1977, Gainesville, Fla.], p. 246-253.
- Hunter, Serena Colvin.
1977. **An electrophoretic analysis of isoenzyme variation in a Piedmont loblolly pine seed orchard.** M.S. Thesis, North Carolina State Univ., Raleigh. 48 p.
- Hyun, S.K.
1976. **Interspecific hybridization in pines with special reference to *Pinus rigida* x *taeda*.** *Silvae Genetica* 25:188-191.
- Little, S., and I.F. Trew.
1977. **Progress report on testing pitch x loblolly pine hybrids and on providing hybrid seed for mass plantings.** *In* Proc., 24th North-eastern Forest Tree Improv. Conf. [July 26-29, 1976, College Park, Md.], p. 14-28.
- Lundkvist, Kenneth.
1975. **Inheritance of acid phosphatase isozymes in *Picea abies*.** *Hereditas* 79:221-226.
- Lundkvist, Kenneth, and Dag Rudin.
1977. **Genetic variation in eleven populations of *Picea abies* as determined by isozyme analysis.** *Hereditas* 85:67-74.
- Mitton, J. B., Y. B. Linhart, J. L. Hamrick, and J. S. Beckman.
1977. **Observations on the genetic structure and mating system of ponderosa pine in the Colorado Front Range.** *Theor. Appl. Genet.* 51:5-13.
- Muller, G.
1977. **Short note: cross-fertilization in a conifer stand inferred from enzyme gene-markers in seeds.** *Silvae Genetica* 26:223-226.
- Porterfield, Richard L.
1974. **Predicted and potential gains from tree improvement programs — a goal programming analysis of program efficiency.** Tech. Rep. 52, 112 p. School of Forest Resour., North Carolina State Univ., Raleigh.
- Rudin, Dag.
1975. **Inheritance of glutamate-oxalate-transaminases (GOT) from needles and endosperms of *Pinus sylvestris*.** *Hereditas* 80:296-300.
- Rudin, Dag, Gosta Eriksson, Inger Ekberg, and Marianne Rasmusson.
1974. **Studies of allele frequencies and inbreeding in Scots pine populations by the aid of the isozyme technique.** *Silvae Genetica* 23:10-13.
- Rudin, Dag, and Dag Lindgren.
1977. **Isozyme studies in seed orchards.** *Studia Forestalia Suecica* Nr 139, 23 p. Swedish Coll. Forestry, Stockholm.
- Rudolf, Paul O.
1974. **Tree-seed marketing controls.** *In* Seeds of woody plants in the United States. p. 153-166. U.S. Dep. Agric., Agric. Handb. 450.
- Wakeley, Philip C., O. O. Wells, and T. E. Campbell.
1965. **Mass production of shortleaf x slash pine hybrids by pollinating unbagged female flowers.** *In* Joint Proc., 2d Genet. Workshop Soc. Amer. For. and 7th Lake States Forest Tree Improv. Conf. [Oct. 21-23, 1965, Hickory Corners, Mich.], p. 78-79.
- Woessner, R. A. and E. C. Franklin.
1973. **Continued reliance on wind-pollinated southern pine seed orchards—is it reasonable?** *In* Proc., 12th Southern Forest Tree Improv. Conf. [June 12-13, 1973, Baton Rouge, La.], p. 64-73.
- Wright, Jonathan W.
1976. **Introduction to forest genetics.** 463 p. Academic Press., Inc., New York.



The Forest Service of the U.S. Department of Agriculture

- ... Conducts forest and range research at more than 75 locations from Puerto Rico to Alaska and Hawaii.
- ... Participates with all State forestry agencies in cooperative programs to protect and improve the Nation's 395 million acres of State, local, and private forest lands.
- ... Manages and protects the 187-million-acre National Forest System for sustained yield of its many products and services.

The Pacific Southwest Forest and Range Experiment Station

- ... Represents the research branch of the Forest Service in California, Hawaii, and the western Pacific.

Conkle, M. Thompson, technical coordinator.

1981. **Proceedings of the symposium on isozymes of North American forest trees and forest insects, July 27, 1979, Berkeley, California.** Gen. Tech. Rep. PSW-48, 64 p., illus. Pacific Southwest Forest and Range Exp. Stn., Forest Serv., U.S. Dep. Agric., Berkeley, Calif.

These 10 symposium papers discuss gene resource management, basic genetics, genetic variation between and within tree species, genetic variability and growth, comparisons of tree life history characteristics, genetic variation in forest insects, breeding systems, and applied uses of isozymes in breeding programs.

Retrieval Terms: forest tree breeding, isozymes, genetic variation, North America, symposia

