

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 ± 0.0370	Yeh and O'Malley (1979)
interior variety	11	21	.6800	.1747 ± .0422	Yeh (In preparation) ²
<i>Picea sitchensis</i> (Bong.) Carr.	10	24	.5130	.1467 ± .0400	Yeh and El-Kassaby (1979)
<i>Pinus contorta</i> ssp. <i>latifolia</i> (Engelm.)	10	25	.5914	.1544 ± .0365	Yeh (In preparation) ³
Critchfield	9	25	.5867	.1601 ± .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.

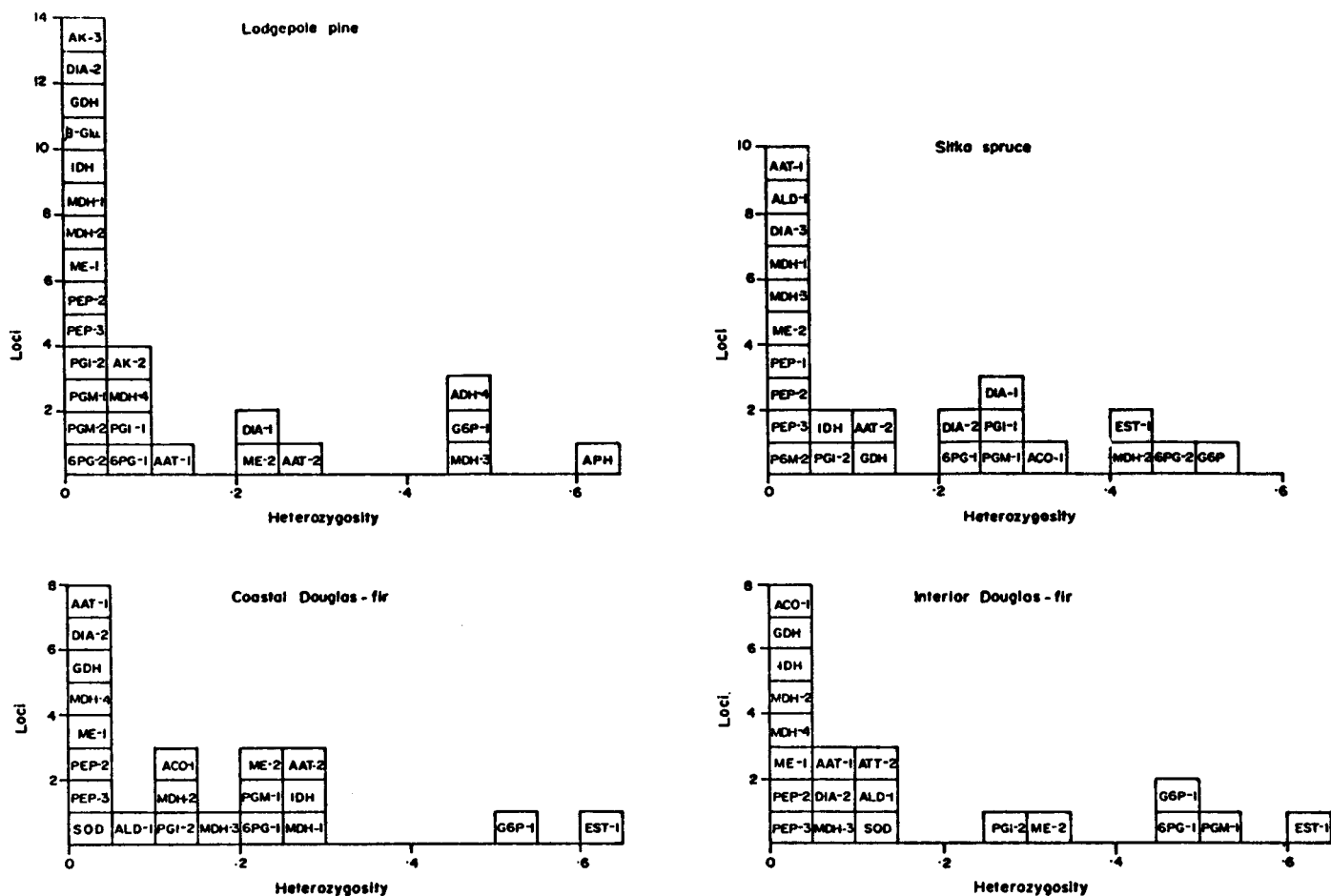


Figure 1—Distribution of heterozygosities for alleles determining electrophoretic variants in some species of conifers.

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 fine: 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.