REPRINTED FROM:

Proceedings 17th Meeting of the
Canadian Tree Improvement
Association, Part 2 (1980)
(Pages 109 - 117)

AMOUNT AND DISTRIBUTION OF ISOZYME VARIATION IN VARIOUS CONIFER SPECIES

M. Thompson Conkle

Pacific Southwest Forest and Range Experiment Station
Forest Service, U.S. Department of Agriculture,
P.O. Box 245
Berkeley, California, 94701 U.S.A.

ABSTRACT

Estimation of the relative amount and the geographic distribution of genetically controlled variation is a central topic of tree resource management. Biochemical data from the analysis of forest tree enzyme variants provides a direct and precise measure of allele frequencies of tree genes.

The amount of genetic variation in several conifers; Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco), loblolly pine (*Pinus taeda* L.), sugar pine (*P. lambertiana* Dougl.), knobcone pine *P. attenuata* Lemm.), lodgepole pine (*P. contorta* var. *latifolia* Engelm.), and Jeffrey pine (*P. Jeffreyi* Grove and Balfour) is assessed by determining the number of alleles per locus and species heterozygosities. Both measures indicate that native conifers are genetically variable and species differ in the kind and amount of variation they contain. Comparisons with herbaceous plants rank conifers among the most variable plants thus far studied.

The distribution of genetic variation measured by analysing stands and geographic samples is an active area of research. Published studies and current unpublished work leads to the conclusion that geographic trends can be identified that support the hypothesis that subsamples are significantly differentiated with respect to allele frequencies but the major amount of variation resides within the subsamples. Studies using isozyme data are providing data on breeding systems and comparisons of life history characteristics are showing consistent patterns which characterize the genetic strategies of forest trees.

RÉSUMÉ

L'évaluation de la qualité relative et de la distribution géographique de la variation faisant l'objet de contrôle génétique sont des problèmes de base dans l'organisation de l'exploitation des ressources en arbres. Les données biochimiques obtenues d'après l'analyse des variations d'enzymes des arbres de forêt donnent une indication directe et précise de la fréquence des allèles dans les gènes des arbres.

Le taux de variation génétique de plusieurs conifères; le sapin Douglas (*Pseudotsuga menziesii* (Mirb.) Franco), le pin à l'encens (*Pinus taeda* L.), le pin gigantesque (*P. lambertiana* Dougl.), le pin knobcone (*P. attenuata* Lemm.), le pin lodgepole (*P. contorta* var. *latifolia* Englem.), et le pin de Jeffrey (*P. Jeffreyi* Greve and Balfour) est trouvé en déterminant le nombre d'allèles par locus et les espèces hétérozygotes. Ces deux mesures indiquent que les conifères indigènes sont génétiquement variables et que les espèces diffèrent par la nature et le taux de variation qu'elles contiennent. Des études comparatives avec les plantes herbacées rangent les conifères parmi les végétaux les plus variables qui aient été étudiés jusqu'à présent.

La distribution de variation génétique mesurée sur des plants d'arbres et des échantillons géographiques représente une branche active de la recherche. Des études publiées et des travaux en cours non publiés nous poussent à conclure que l'on peut identifier des tendances géographiques qui affirment l'hypothèse selon laquelle les échantillons secondaires sont largement différenciés en ce qui concerne les fréquences d'allèles, mais que le plus grand taux de variation se trouve à l'intérieur des échantillons secondaires. Des recherches basées sur l'isozyme apportent des informations sur les systèmes de reproduction. Des comparaisons sur la nature du cycle de la vie présentent des modèles stables qui caractérisent les stratégies génétiques des arbres de forêt.

INTRODUCTION

A current challenge in forest research is to estimate the amount and distribution of genetic variation between and within taxa. This information is applied in assessing how significant improvement in the commercial products of forests can be realized while maintaining sufficient genetic variability to cope with uncertain future environments.

If a scale for measuring relative amounts of genetic variation was developed with natural populations representing maximum and cultivated varieties representing lesser values, most native North American forest trees would probably rank near the top. Cultivated varieties and introduced species of trees have not displaced native vegetation and most native populations have not shrunk to the extent that tree breeders are confronted with a severe reduction of genetic variability. Cautions from breeders working with animals and crop plants that have long histories of domestication, however, are influencing tree improvement plans. Many, perhaps most, domesticated species and varieties have severely restricted genetic resources (Harlin 1975). Breeders are actively searching for usable genetic variability among native populations (Harlin 1976, Brown et al. 1978). Forest tree breeders need to identify the amount and distribution of genetic variation in managed species to develop efficient tree improvement strategies. Isozyme analyses play an important role in the description of genetic variability.

My current research effort is to identify allelic forms of isozymes and characterize the genetic variability within six conifer species. This research, when added to estimates of heterozygosity in isozymes for other conifers, suggests that trees are among the most heterozygous of higher organisms analysed (Hamrick et al. 1980). Current reports on several species further suggest that most genetic variability is within geographic samples and a small proportion (about 7 percent) of total variation is between geographic samples.

THE PROCEDURES

Open-pollinated seed from individual parent trees germinates to the stage where the radicle just emerges beyond the seed coat and is then refrigerated until analysed.

As a minimum, six gametophytes from seed of each parent tree are individually macerated and electrophoresed. With this sample size, homozygotes are identified without error and true heterozygotes are expected to be misclassified in less than 3 percent of the samples. Electrophoresis and staining procedures are similar to those described by Fowler and Morris (1977), and Guries and Ledig (1978). Phenotypes of enzyme bands including migration distances and staining intensities are the basis for identifying allelic forms and inferring parent tree genotypes.

Seed materials came from a variety of sources (Table 1). Diploid tree genotypes were inferred for individual trees in the samples and allele frequencies for loci were assessed on the basis of two times the number of parent trees.

The techniques for obtaining enzyme bands is similar for all conifers, but the analysis of the species included in this study spans several years of research. Different numbers of loci per species resulted from improvements in laboratory procedures. Some bias toward more variable loci may be a feature of species with few numbers of loci, but all enzymes with clear bands are included in this survey.

Two measures of genetic variability are examined for each species. The average number of alleles per locus estimates the allelic differentiation within a species. A few loci in some species are monomorphic, but characteristically, from two to a maximum of nine different alleles are identified for each locus.

The second measure of variability estimates the average number of heterozygous loci per parent tree. Average hetereozygosity is calculated from species allele frequencies for each locus by subtracting the sum of the squared allele frequencies from 1.0. This heterozygosity value estimates the number of heterozygotes expected under Hardy-Weinberg conditions. For species in this study, the estimated numbers of heterozygotes approximates the actual count of heterozygotes. The choice of using computed rather than observed values is to make the data

Table 1. Species samples for evaluating the amount of allelic variation in isozyme loci.

Species and Location	Parent Trees 1/ in the Sample
Douglas-fir, *Pseudotsuga menziesii* (Mirb.) Franco	
Natural population, Oregon	152
Loblolly pine, *Pinus taeda* L.	
Natural population, North Carolina	146
Superior trees	90
Sugar pine, *P. lambertiana* Dougl.	
California and Oregon	58
Jeffrey pine, *P. Jeffreyi* Grev. & Balf.	
Four natural stands, California	75
Lodgepole pine, *P. contorta* Dougl. ssp. *murrayana*	
Natural stand, California	40
Knobcone pine, *P. attenuata* Lemm.	
Ten geographic areas, Oregon and California	49

1/ Allele frequencies are 2 X these numbers.

compatible with other literature. The number of heterozygotes per locus tends to be large when allele frequencies have intermediate values and when alleles per locus are numerous.

RESULTS

Species differed in both measures of genetic variability (Table 2). Douglas fir (*Pseudotsuga menziesii* (Mirb.) Franco) and loblolly pine (*Pinus taeda* L.) averaged almost four alleles per enzyme locus. Sugar, pine (*P. lambertiana* Dougl., Jeffrey pine (*P. Jeffreyi* Grev. & Balf., and Lodgepole pine (*P. contorta* Dougl.) averaged slightly less than three alleles per locus, and knobcone pine (*P. attenuata* Lemm.) had only about two. Confidence intervals for the means indicate that the three groups

Table 2. Isozyme variability in several coniferous species.

Species	Loci Analysed	Alleles per Locus ($\bar{X} \pm$ S.D.)	Heterozygosity per Locus ($\bar{X} \pm$ S.D.)
Douglas-fir	19	4.00 ± .22	0.33 ± .03
Loblolly pine			
Natural stand	11	3.73 ± .32	.34 ± .04
Select trees	30	3.87 ± .13	.26 ± .01
Sugar pine	20	2.85 ± .18	.26 ± .03
Jeffrey pine	44	2.85 ± .33	.26 ± .05
Lodgepole pine	37	2.78 ± .08	.19 ± .01
Knobcone pine	22	2.09 ± .11	.14 ± .02

of species differ significantly from one another.

Mean heterozygosities per locus followed the trends set by the average number of alleles per locus. Douglas-fir and the natural stand of loblolly pine had trees that on the average were heterozygous at about one-third of the loci in the sample. The superior tree sample of loblolly pine had a lower estimate for heterozygosity than the natural stand (Table 2), but when the same genes are compared, the values are nearly equal. Additional loci analysed recently in our laboratory had several allelic variants per locus, but often had one allele in high frequency.

The select loblolly pines had the same values for heterozygosity per locus as did sugar and Jeffrey pines. These three species were heterozygous for about one-fourth of their loci. Lower values were observed for lodgepole pine (0.19) and knobcone pine (0.14).

DISCUSSION

Contrasts between conifer species for the amount of isozyme variation each possesses are heightened by the finding that red pine (*Pinus resinosa* Ait.) is estimated to have an average heterozygosity value of zero (Fowler and Morris 1977). Nine loci were studied for red

pine. In other conifers, several loci which were among those sampled in red pine, had many alleles and high estimates of heterozygosity. Because several of the red pine seed samples were from widely separated geographic locations, the lack of variation was not the result of restricted sampling. It is noteworthy that red pine has high self-compatibility and low levels of phenotypic variability in growth traits. The zero value for red pine heterozygosity is the low benchmark.

The values reported here for alleles per locus and heterozygosity per locus for Douglas-fir and loblolly pine are among the highest obtained for conifers. A similarly high value was reported for bristlecone pine (*P. aristata* Engelm.): 2.35 alleles per locus and heterozygosity of 0.36 (Hiebert 1977). The highest values for conifers, therefore, approach four alleles per locus and average heterozygosities in excess of one-third of the loci per individual.

CONCLUSIONS

Broad-scale comparisons of species and taxa are outdated as new data become available and should be interpreted cautiously because comparisons are rarely made on the basis of the same subset of loci. But inclusion of large samples of loci improves average values for species. Hamrick et al. (1979, 1980) summarize plant species information currently available. On the average, genetic variation in plants is roughly equivalent to that in invertebrates but significantly more than in vertebrate species. Trees average significantly more variation than herbaceous plants. Plant species, with the greatest number of alleles per locus and the highest values for heterozygosity, have large geographic distributions, high fecundities, outcrossing as the mode of reproduction, long generation times, and are wind-pollinated.

Total genetic variation within a taxon can be partitioned to within and between subsamples. Recent studies of variation in conifers show that most of the genetic variation is within subsamples. Pitch pine (*P. rigida* Mill.), had 99 percent of total variation within populations (Guries and Ledig 1977) and only one percent between populations. Yeh (1980) reports similar large amounts of variation within populations of Douglas-fir, Sitka spruce, and lodgepole pine (ssp. *latifolia*) from British Columbia, Canada. From 92 to 97 percent of the total variation is within populations of these three species. Brown and Moran (1980) summarized 17 studies of forest trees and found that wind-pollinated species averaged 93 percent of total variation within subpopulations, 7 percent between populations.

These low levels of variation among subpopulations strengthen the inferences drawn from the species data reported in this study. Estimates of the amount of genetic variation of single populations are expected to compare favorably with samples representing species collections. Further work on forest trees is expected to assign tree species on a heterozygosity scale that ranges from near zero heterozygosity to the most heterozygous organisms studied to date.

It seems reasonable to assume that variation in enzymes is a direct measure of overall genetic variation. Isozymes have proved their usefulness for estimating relative amounts of genetic variability within taxonomic units and for comparing the degree of similarity between taxa. Allelic differences in enzymes, however, measure genetic variation in primary gene products. These primary gene products are many steps removed and represent only a small sample of the genes that contribute to the expression of a tree's phenotype. The challenge in research is to establish a relationship, if it exists, between enzyme data and phenotypically valuable traits in forest trees.

REFERENCES

Brown, A.H.D., E. Nevo, D. Zohary and O. Dagan. 1978. Genetic variation in natural populations of wild barley (*Hordeum spontaneum*). Genetica 49:97-108.

Brown, A.H.D. and G.F. Moran. 1980. 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, California, (In press).

Fowler, D.P. and R.W. Morris. 1977. Genetic diversity in red pine: evidence for low genetic heterozygosity. Can. J. Forest. Res. 7(2): 343-347.

Guries, R.P. and F.T. Ledig. 1977. Analysis of population structure from allozyme frequencies. In: Proceedings of the Fourteenth Southern Forest Tree Impr. Conf., June 14-16, 1977, Gainesville, Fla. Southern Forest Tree Comm. Publ. 36, p. 246-253.

Guries, R.P. and F.T. Ledig. 1978. Inheritance of some polymorphic isoenzymes in pitch pine (*Pinus rigida* Mill.). Heredity 40:27-32.

Hamrick, J.L., Y.B. Linhart and J.B. Mitton. 1979. Relationships between life history characteristics and electrophoretically-detectable genetic variation in plants. Annu. Rev. Ecol. Syst. 10:173-200.

Hamrick, J.L., J.B. Mitton and Y.B. Linhart. 1980. Levels of genetic variation in trees: The influence of life history characteristics. In Proceedings of the Symposium on Isozymes of North American Forest Trees and Forest Insects, July 27, 1979, Berkeley, California, (In press).

Harlin, J.R. 1975. Our vanishing genetic resources. Science 188:618-621.

Harlin, J.R. 1976. Genetic resources in wild relatives of crops. Crop Sc. 16:329-333.

Hiebert, R.D. 1977. The population biology of bristlecone pine (*Pinus aristata*) in the eastern Great Basin. Ph.D. Dissertation, Univ.

Kansas, Lawrence, 82 p.

Yeh, F.C. 1980. The analysis of gene diversity in several forest tree species. In: Proceedings of the Symposium on the Isozymes of North American Forest Tree and Forest Insects, July 27, 1979, Berkeley, California, (In press).

DISCUSSION

Dr. E.K. Morgenstern. In the last slides you showed that a low amount of variation had been found in Norway spruce but a very large amount of variation in Monterey pine - two species which have range differences of the order of 1 000. Could you explain?

Dr. M.T. Conkle. Your observation should account for variation within the two species. The total isozyme variation for Norway spruce from Sweden was estimated to be three times larger than for California Monterey pine, 41 percent vs 15 percent. The illustration, which you referred to, divided total variation into two classes; among populations and within populations. Monterey pine had significant variation among populations (15 percent), while Norway spruce populations were less distinct (3 percent of total variation). The contrast between these two species suggests a relationship between variation in isozymes and the size and isolation of populations within species. Norway spruce is widely distributed with large population numbers and few restrictions to gene flow. Monterey pine has small populations separated by distances of 50-100 kilometres. The information for Norway spruce was from a 1979 Hereditas article by Kenneth Lundkvist. Data for California populations of Monterey pine were from a seminar by Gavin Moran.

Dr. E.K. Morgenstern. I guess in your Table 1 you have a low number of samples of Norway spruce. I wonder if you could explain this.

Dr. M.T. Conkle. Data for forest trees are just becoming available. I selected studies that reported ten or more loci to gain accuracy in the average estimates. Other studies of Norway spruce, however, give values similar to those in the illustration. Researchers, I might add, are working toward standardized isozyme analysis so results from different laboratories will be comparable.

Dr. D.P. Fowler. The question in Norway spruce is a bit bothersome. I suggest it makes very big differences if your three or four samples are from Poland, let us say, compared to a broader range-wide sample.

Dr. M.T. Conkle. Isozyme studies of various plants and animals commonly report large proportions of total variation within geographic samples and small proportions among samples. Thus, I expect variation within stands from Poland would be minor, similar to the findings of the Swedish study. Comparing Polish and Swedish samples, however, should

produce larger differences.

Forest geneticists are accustomed to significant differences among provenances in growth trials and these differences are often greater than 3 percent. But, the comparison of isozyme and growth results may not account for the compounding of growth differences over the life of a trial. Would growth differences within field trials average about 3 percent annual rates when discounted over several years? Francis Yeh's research addresses the comparison of isozyme and growth variability; we should watch his results as they become available.

Dr. R.B. Hall. This population versus within stand versus between stands selection seems to be very important and I am wondering how much extrapolation you can make between particular enzyme systems which may not be under strong selection pressures versus something like photoperiod response which would be under very strong selection pressures.

Dr. M.T. Conkle. People are actively researching your question. Some isozyme variation probably represents noise (random variation) while other isozyme variation results from selection. Isozyme alleles differ in heat deactivation and temperature optima, and the biochemical function of most enzymes is known. Some isozyme allele frequencies may reflect selection as strongly as photoperiod response.

Dr. F.C. Yeh. I would like to add a little bit to the analysis. We have further subdivided the analysis into subpopulation aspects. Like we looked at within populations. This is confounded because it included family variation within populations. And that component comes to about one-third of the variation. Therefore, family differences within populations account for one-third of the differences that you see.