# FOLIAR ISOZYME VARIATION IN TWENTY SEVEN PROVENANCES OF <u>PINUS</u> <u>STROBUS</u> L.: GENETIC DVERSITY AND POPULATION STRUCTURE<sup>1</sup>

by Jang Bal Ryu, Research Assistant, and Robert T. Eckert, Assistant Professor of Forest Resources and Genetics, Institute of Natural and Environmental Resources, University of New Hampshire, Durham, NH 03824.

## ABSTRACT

Genetic structure of <u>Pinus strobus</u> L. has been investigated by starch gel electrophoresis. Eight isozymes coded by 12 loci were used for this study. Interprovenancial allele differentiation amounted to eight percent of the total variation. Frequencies of PGM22 and average heterozygosity are correlated with latitude. Cluster analysis using frequencies of six alleles, chosen for their high contributions to common variance, showed four clusters of provenances. Some ecotypic variation was indicated in the southern Appalachians and in the northern part of the species range.

## INTRODUCTION

Assessments of genetic variation within and among populations are important for deciding strategy and tactics of tree improvement. The recent use of gel electrophoresis in isozyme studies has aided in assessing genetic structure of some conifer species (Guries and Ledig 1982, Linhart

The genetic structure of natural populations is the result of a dynamic process involving the opposing forces of gene flow and selection: the former homogenizes but the latter differentiates. For forest trees, particularly the wind-pollinated conifers, gene flow has been viewed as an overpowering force and a great degree of differentiation is not expected within stands, unless gene flow is restricted. On the other hand, natural selection is viewed as a strong force operating in heterogeneous environments. Tree species having a wide natural range or living in heterogeneous environments may be subjected to relatively stronger selection pressures (Mitton et al. 1977).

This paper reports an analysis of genetic structure of 27 provenances of <u>P. strobus</u> based on eight foliar isozymes coded by 12 loci. Since samples were collected from a provenance test plantation, they may not exactly represent original natural populations. The study trees are progeny of a limited number of trees, and selection, if any, has operated for one

This study was carried out in partial fulfillment of the senior author's Ph.D dissertation and was supported by McIntire-Stennis grant Number 11 of the New Hampshire Agricultural Experiment Station. Scientific Contribution Number 1203 from the New Hampshire Agricultural Experiment Station.

generation. Therefore, the term provenance is used instead of population. Cones for the provenance test were collected from approximately ten average trees on five acres per provenance. It is assumed that the genetic structure of provenances does not depart far from that of the original populations.

#### MATERIALS AND METHODS

Foliage samples were collected from a provenance test plantation at Alfred, Maine in the winter of 1980. A total of 675 samples, 25 trees from each of 27 provenances (Table 1) were collected. Procedures of electrophoresis, staining methods, band patterns, genotype and allele frequencies, are available elsewhere (Ryu 1982). Locus designations remain tentative until confirmed by control crosses, however isozymes in this report are observed in expected patterns and co-electrophorese to compatible Rf positions with megagametophyte allozymes reported earlier (Eckert, <u>et al</u>. 1981).

Allelic variation within provenance was quantified by measuring the proportion of polymorphic loci, average number of alleles per locus, and average heterozygosity (Lewontin 1967). Genetic differentiation among provenances was investigated by two methods. The first is partitioning of genetic variation within and among provenances (Nei 1973). The second procedure is genetic distance among provenances (Nei 1972). Genetic distance was compared with geographic distance among provenances)

Relationships among arcsine transformed allele frequencies and provenance geographic data were investigated by Pearson correlation coefficients. Only the most common allele (frequency less than .99 at each locus was used, because other allele(s) at each locus were strongly correlated with the most common allele (-.07 to -1.00, mean -.69). In reporting our results, when more than one zone was recognized for an enzyme, the faster migrating zone was designated as locus 1 and the next zone as locus 2, etc. This convention was also followed for allele designation at a locus.

Cluster analysis of 27 provenances was carried out based on allele frequencies at six loci. The six alleles were selected using alpha factor analysis with varimax rotation on the most common allele at ten polymorphic loci. The factor analytic approach was used to reduce the number of isozyme allele frequencies to those contributing most to common variance among variables.

Calculations for these measures and statistics were carried out using the Statistical Package for the Social Sciences (SPSS) or Fortran programs written for specific problems. The cluster analysis was carried out by a packaged program of Biomedical Computer Programs P-series (BMDP2M). Calculation formulae for allele diversity and genetic distance are listed in Appendix 1.

NO.	Location		Latitude	Longitude
1	Union	Georgia	34.8N	84.Ow
2	Transylvania	North Carolina	35.2	82.6
3	Greene	Tennessee	36.0	82.8
4	Garrett	Maryland	39.7	78.8
5	Greenbrier	West Virginia	38.0	80.5
6	Monroe	Pennsylvania	41.1	75.4
8	Clearfield	Pennsylvania	41.0	78.5
9	Clearfield	Pennsylvania	40.8	78.5
10	Ulster	New York	42.0	74.0
11	Ulster	New York	41.8	74.3
12	Franklin	New York	44.4	74.3
13	Worcester	Massachusetts	42.5	72.3
14	Penobscot	Maine	44.9	68.6
15	Allamakee	Iowa	43.3	91.5
16	Ashland	Ohio	40.8	82.3
18	Forest	Wisconsin	45.5	88.5
19	Cass	Minnesota	47.5	94.5
20	Lunenburg	Nova Scotia	44.4	64.6
21	Sunbury	New Brunswick	46.0	66.3
22	Quebec	Quebec	47.5	72.0
23	Pontiac	Quebec	47.5	77.0
24	Norfolk	Ontario	42.7	80.5
25	Algoma	Ontario	46.2	82.6
27	Carroll	New Hampshire	43.8	71.4
28	Lake	Minnesota	48.1	91.3
29	Houghton	Michigan	44.3	84.8
30	Pulaski	Virginia	36.9	81.0

Table 1. Twenty-sevenprovenances of Pinus strobus.

#### RESULTS AND DISCUSSION

## <u>Measures of Variation within Provenance</u>

The mean value of proportion of polymorphic loci (P), number of alleles per locus (A) and average heterozygosity (H) based on 12 loci were .69, 2.01 and .236 respectively. P,A, and H values of .53, 2.06 and .175 were calculatead from data of Eckert, <u>et al</u>. (1981), based on 17 loci in megagametophyte tissue from 35 clones of this species. It should be noted that values of P, A, and H depend heavily on the choice of loci, occurrence of rare alleles, number of samples per provenance, and areas sampled, in addition to possible tissue differences. The importance of examining a large number of loci should be noted in order to obtain a reliable index.

Hamrick <u>et al</u>. (1981) compared levels of genetic variation and life history characteristics of 20 conifer species and concluded that species of ater successional stages, mesic habitat types, with open cones and southern or western species have more genetic variation than species with alternate combinations of characteristics. The mean values of P, A and H from the 20 conifer species were .68, 2.29 and .207. Our values, based on 12 loci detected in foliage, are reasonably similar to those reported for other conifer species. <u>P. strobus</u> fits all of the categories for high variation except natural range distribution. So the considerable genetic variation in <u>strobus</u> is not unexpected.

## <u>Allelic Diversity</u>

Partitioning genetic variation into within and among provenances, and allelic differentiation resulted in values of allele differentiation (GST) ranging from .272 at ME1 to .000 at fixed loci. LAP1, GOT3, PER2, PGM2, and MDH1 were highly diverse loci, indicated by large total diversity values (Table 2).

Locus	Total Diversity	Within Provenance	Among Provenance	Allele Differentiation
	(HT)	(HS)	(DST)	(GST = DST/HT)
LAP1	.6589	. 5986	. 0603	.092
LAPI LAP2	.1876	.1715	.0161	.092
LAP3	.0478	.0449	.0029	.061
дот3	. 5000	. 4735	. 0265	.053
SDH1	.1044	.0951	.0093	.089
4E1	.2300	.1674	.0626	.272
PER2	.4548	.3987	.0598	.130
F-EST1	.0519	.0493	.0026	.050
PGM1	.0000	.0000	.0000	.000
PGM2	.3880	.3637	.0243	.063
MDH1	.5727	.4723	.1004	.175
MDH2	.0000	.0000	.0000	.000
Mean	.2663	.2326	.0304	.080

Tble 2 Analsis of allele diversit amon 27 rovenances of E strobusa

<sup>a/</sup>genic diversity of Nei (1973) based on most common allele (<.99)

The mean value of GST was .080, thus eight percent of the genetic variation over all provenances was due to I interprovenancial allele differentiation. This value Is similar to the average allele differentiation of 23 forest tree species summarized by Brown and Moran (1981), who found that wind-pollinated conifers, on the average, had seven percent of total variation among populations and 93 percent within populations. These comparisons may also be affected by the loci studied and the populations sampled. Broad-leaved trees showed greater differentiation than conifers (Brown and Moran 1981). Inbreeding plant species showed even higher differentiation among populations, though total variation (HT) was less than outcrossing plant species (Brown 1979).

#### <u>Genetic Distance</u>

Genetic differentiation among provenances was further investigated using Nei's (1972) genetic distance. The genetic distance between two provenances varied (Appendix 2): the highest value (.139) was found between provenance 2 (NC) and 3 (TN), and the lowest value (.005) between provenance 13 (Mass.) and 20 (Nova Scotia), and between provenance 15 (Iowa) and 27 (NH).

The correlation coefficient for a specific provenance was calculated from genetic and geographic distances from that provenance to all the other 26 provenances (Table 3). If genetic differentiation is largely due to isolation by distance, then the genetic distance and geographic distance are expected to be positively correlated. The correlation coefficient between the two distances over all 351 possible combinations ((27 x 27 -27)/2) was statistically significant, though the coefficient was only .11 (Table 3). The proportion of variation in genetic distance explained by geographic distance is generally very low, with the exception of two NY provenances (10, 12). In general, there is no relationship between geographic and genetic distance in these data thus the hypothesis that genetic differentiation is due to isolation by distance is not supported.

A significant correlation between genetic distance and geographic distance was found in <u>Pseudotsuga menziesii</u> from British Columbia (Yeh and O'Malley 1980). Similar results were reported by Yang <u>et al</u>. (1977) for the same species, and in <u>Picea abies</u> in Sweden (Lundkvist and Rudin 1977). Only weak correlation between the two distances was found in eleven populations of P. <u>rigida</u> (Guries and Ledig 1982). Linhart <u>et al</u>. (1981) reported no correlation between the two distances in <u>P. ponderosa</u>, where substantial genetic differentiation was detected among clusters within a small area. Two herbaceous selfing species in Israel, <u>Hordeum spontaneum</u> and <u>Avena barbata</u>, also showed no correlation between genetic distance and geographic distance (Nevo gt .L. 1979, Kahler et al. 1980).

## Patterns of Variation Related to Environment

A few significant correlations were found with latitude, while longitude correlated only with LAP32 frequencies at the five percent level (Table 4). Correlation between latitude and PGM22 was significant at the .1 percent level. LAP22 was correlated with latitude at the 1.0 percent level, whereas LAP32 and GOT32 were correlated at the five percent level. These correlations suggest a general cl 1 nal pattern of variation, which may be of use in tree improvement. Isozyme frequencies of LAP, PGM, and ME are significantly correlated with relative height growth of eastern white pine in provenance test plantings located at mid to southern locations (Eckert and Ryu 1982).

Studies of the relationship between individual isozyme frequency and geographic or climatic variables in forest trees are rare. PER and PGM1 in <u>P. ponderosa</u> differentiated with elevation (Mitton <u>et al</u> 1980). Clinal variation of allele frequency along altitudinal gradients has often been reported (Lundkvist 1979, Yang et al.1977, Yeh and O'Malley 1980). Increasing latitude and elevation resulted in increases in AP allele frequency in <u>Picea abies</u> (Bergmann 1978). Allele frequencies of MDH and 6-PGD in P. <u>taeda</u> were correlated with temperature, longitude and annual

Prove.	Mean Genetic Distance	Mean Geographic Distance	Correlation Coefficient	Coefficient of Determination <sup>d∕</sup>		
1 <b>GA</b>	.036	1,249	195	.038		
2 NC	.099	1,163	.138	.019		
3 TN	.052	1,100	104	.011		
4 MD	.032	793	.035	.001		
5 WVA	.032	896	273a	.045		
6 Penn.	.026	793	.138	.019		
8 Penn.	.035	740	.123	.015		
9 Penn.	.024	745	.079	.006		
10 NY	.040	811	.457c	.209		
11 NY	.035	810	.180	.032		
12 NY	.030	786	.516c	.266		
13 Mass.	.029	905	.313a	.098		
14 Maine	.027	1,120	.185	.034		
15 Iowa	.030	1,261	299a	.089		
16 Ohio	.039	809	.006	.000		
18 Wis.	.038	1,174	.046	.002		
1 <b>9 Minn.</b>	.043	1,572	016	.000		
20 NS	.022	1,442	.213	.045		
21 NB	.036	1,329	.421b	.177		
22 Que.	.035	977	.029	.001		
23 Que	.063	963	.117	.014		
24 Ont.	.034	750	076	.006		
25 Ont.	.028	923	.127	.016		
27 NH	.032	954	.300a	.090		
28 Minn.	.040	1,390	.044	.002		
29 Mich.	.035	933	091	.008		
30 VA	.031	974	113	.013		
	Val	ues from all 351	L combination			
	.037	1,013	.110b	.012		

Table 3. Correlations between genetic and geographic distances for al provenances.

a, b and c: **mignificant** at the 10, 5 and 1 percent level, respectively.

d: Measures the proportion of variation in genetic distance determined by variation in geographic distance.

precipitation (Florence and Rink 1980). Clinal variations in allele frequencies are often claimed to be caused by selection (Bergmann 1978). However, the possibility of processes other than natural selection causing the clinal variation can not be excluded (Kimura and Maruyama 1971). Table 4. Correlation coefficients among isozyme frequencies, average heterozygosity (HET), latitude and longitude of the provenances.

Isozyme	Latitude	Longitude
LAPiN LAP22 LAP32 GOT32 SDH12 ME12 PER22 F-EST1 PGM22 MDH11 Average heterozygosity	03 <sub>b</sub> 48 42 <sup>a</sup> .36 <sup>a</sup> .30 04 .12 01 <sub>b</sub> .48	.01 21 .3 a 02 .25 .10 23 .01 02 11 16

a, b and c: significant at the 5, 1 and .1 percent level, respectively.

Average heterozygosity (Het) showed a positive correlation with latitude (p)=.01) (Table 4).

ten loci showed negative correlations with latitude; those at SDH and F-EST were significant at the five percent level. Yang <u>et al</u>. (1977) reported low heterozygosity in northern sources in <u>Pseudotsuga menziesii</u>. Only three loci were used for Yang's study however.

# Factor and Cluster Analysis

Arcsine transformed frequencies of the most common allele at ten polymorphic loci were alpha factored with varimax rotation of factors to simple structure. Six loci were selected, which had loadings greater than 0.5 on one factor and low loadings on the remaining factors, for cluster analysis (Table 5). Such a procedure eliminates alleles which do not contribute strongly to the simple factor structure.

cluster analysis of 27 provenances based on mean arcsine frequencies of LAP22, LAP32, GOT32, SDH2, PER22, and PGM22 resulted in four clusters (Fig. 1). Interestingly, an east to west clustering seems to appear in the data moving from Cluster I to III (Fig. 1). Cluster I included four provenances (6, 11, 14 and 20) and covered Nova Scotia, approximately onethird of the New England area and the northeastern part of Pennsylvania. Cluster I may represent a maritime ecotype. Cluster 11 contained eight provenances, west of cluster 1 and approximately east of the Appalachian ridges and may represent populations adopted to growth in the Appalachian highlands. One exception was provenance 28 (Lake, Minn.) in cluster II. Cluster III included 11 provenances; all provenances west of cluster 11, except provenances 23 (Pontiac, Quebec) and 28 (Lake, Minn.). Cluster III ocated on Appalachian plateaus and plains west of the Appalachians. Cluster IV contained four provenances, three provenances in the southern Appalachians and provenance 23. Provenances in cluster IV were different

from provenances in other clusters and were substantially different from each other.

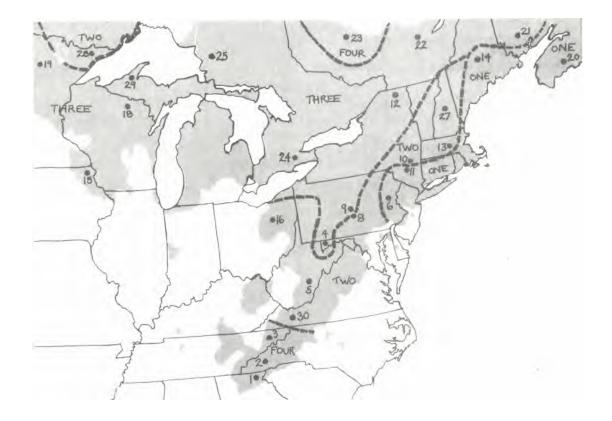
	Factor							
Allele	1	2	3	4				
LAPIN LAP22 LAP32 GOT32 SDH12 ME12 PER22 F-EST12 PGM22 MDH11	016 .089 046 .748a .926a .216 .083 .495 317 .022	.065 .108 094 .310 098 .404 .961a .123 251 069	.006 .338 .942a .187 151 .060 141 392 .028 017	.349 .565a .115 1 .001 .055 .053 305 .697a 044				
Covariance	1.82	1.30	1.24	1.05				

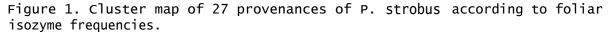
Table 5. Varimax rotated alpha factor of ten foliar isozyme alleles

a. Indicates high loading on the factor. These alleles were chosen for cluster analysis.

The three southern provenances in cluster IV may be refugia from glaciation. Because of complex topography of the Appalachians, gene flow among the three provenances may have been restricted resulting in the three provenances having different genetic composition from each other and from other provenances. Growth performance (Demeritt and Kettlewood 1975, Wright 1970), and other morphological and physiological studies (Mergen 1963) also suggest ecotypic variation around these three provenances. The interpretation of the high differentiation of provenance 23 is not readily discernable. However, an extremely cold climate north of the provenance (Wyman and Flint 1967) may have affected high genetic differentiation in this provenance.

Clustering of provenances in an east-west plane using these isozyme variables may indicate that "ecotypes" related to factors other than latitudinally related variables such as temperature, number of frost-free days, etc. may be detected with this set of isozymes. Recall, however, that frequencies of LAP22, LAP32, GOT32, and PGM22 were correlated with latitude, and these variables were selected according to factor analysis for inclusion in cluster analysis. One would expect the major clusters to fall out along latitude, however this did not seem to be the case except for cluster IV. Cluster analysis based on morphological, or physiological variables, in addition to isozyme variables may clarify interpretations.





### CONCLUSIONS

Estimates of genetic diversity based on 12 fol iar isozyme loci in eastern white pine are close to average figures reported for conifers in other isozyme studies. Intraprovenancial variation accounts for approximately 8 percent of the total genetic variation observed in our 27 provenance rangewide sample. Genetic and geographic distance are generally correlated. However, the hypothesis that genetic differentiation is due to separation by distance is not supportable. Correlations of four alleles and average heterozygosity with latitude may be useful in tree improvement programs if these variables are correlated with economic traits.

Multivariate data exploration of isozyme data resulted in four provenance clusters, three of which may be representative of populations adapted to differing geographic and climatic conditions. Ecotypic variation reported for three southern provenances and one northern provenance based on growth performance and/or morphological variables in earlier studies was also observed : or the same provenances using foliar isozyme data.

<sup>1</sup> Geographic distances were measured on the Classic map of the United States (1:5,000,000), Hammond incorporated, Maplewood, NJ.

### ACKNOWLEDGEMENT

The authors acknowledge excellent critical reviews by Y.T. Kiang, University of New Hampshire and Robert D. Westfall, Ph.D., Davis, CA.

## Literature Cited

- Bergmann, F. 1978. The allelic distribution at an acid phosphatase locus in Norway spruce (<u>Picea abies</u>) along similar climatic gradients. Theor. Appl. Genet. 52: 57-64.
- Braun, E. L. 1964. Deciduous forests of eastern North America. The Blakiston Co. 596pp.
- Brown, A. H. D. 1978. Isozymes, plant population genetic structure and genetic conservation. Theor. Appl. Genet. 52: 145-157.
- Brown, A. H. D. and G. F. Moran 1981. Isozymes and the genetic resources of forest trees. Proc. symposium on isozymes of North American forest trees and forest insects. 1-10.
- Demeritt, M. E. Jr. and H. C. Kettlewood 1975. Eastern white pine seed source variation in the northeastern United States: 16-year results. Proc. Lake States Forest Tree imp. Conf. 12: 80-87.
- Eckert, R. T. and J. B. Ryu 1982. Foliar isozyme variation in 27 provenances of <u>Pinus</u> <u>strobus</u> L.: Relationships with growth and some multivariate considerations. Proc. 7th North American Forest Biology Workshop, Lexington, Ky. in press.
- Eckert, R. T., R. J. Joly and D. B. Neale 1981. Genetics of isozyme variants and linkage relationships among allozyme loci in 35 eastern white pine clones. Can. J. For. Res. 11: 573-579.
- Florence, Z. and G. Rink 1980. Geographic patterns of al lozymic variation in loblolly pine. Proc. Southern Forest Tree imp. Conf. 15: 33-41.
- Guries, R. P. and F. T. Ledig 1982. Genetic diversity and population structure in pitch pine (<u>Pinus</u> <u>rigida</u> Mill.). Evolution 36:387-402.
- Hamrick, J. L., J. B. Mitten and Y. B. Linhart 1981. Levels of genetic variation in trees: influence c r life history characteristics. Proc. symposium on Isozymes of Korth American forest trees and forest insects. 35-40.
- Kehler, A. L., R. W. Allard, M. Krzakowa, C. F. Wehrhahm and E. Nevo 1980. Associations between isozyme phenotypes and environment in the slender wild oat (Avena barbata) in israel. Theor. Appl. Genet. 56: 31-47.
- Kimura, M. and T. Maruyama 1971. Patterns of neutral polymorphism in a geographically structured population. Genet. Res. 18: 125-131.
- Lewontin, R. C. 1967. An estimate of average heterozygosity in man. Amer. J. Human Genetics 19: 681-685.

Linhart, Y. B., J. B. Mitton, K. B. Sturgeon and M. L. Davis 1981. Genetic variation in space and time in a population of ponderosa pine. Heredity 46: 407-426.

- Lundkvist, K. and D. Rudin 1977. Genetic variation in eleven populations of <u>Picea abies</u> as determined by isozyme analysis. Hereditas 85: 67-74.
- Mergen, F. 1963. Ecotypic variation in <u>Pinus</u> <u>strobus</u> L. Ecology 44: 716-727.
- Mitton, J. B., K. B. Sturgeon and M. L. Davis 1980. Genetic differentiation in ponderosa pine along a steep elevational transect. Silvae Genet 29: 100-103.
- 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.
- 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 Colorado front range. Theor. Appl. Genet. 51: 5-13.
- Muller, E. H. 1965. Quaternary geology of New York. in The Quaternary of the United States. 99-112. Princeton Univ. Press. 922pp.
- Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Nat. Acad. Sci. 70: 3321-3323.
- Nei, M. 1972. Genetic distance between populations. The Amer. naturalist 106: 283-292.
- Nevo, E., D. Zohary, A. H. D. Brown and M. Haber 1979, Genetic diversity and environmental associations of wild barley, <u>Hordeum</u> <u>spontaneum</u>, in israel. Evolution 33: 815-833.
- Ryu, J. B. 1982. Genetic structure of <u>Pinus strobus</u> L. based on foliar isozymes from 27 provenances. Unpublished Ph.D. dissertation. University of New Hampshire. Durham, N.H. 03824 133pp.
- Wright, J. W. 1970. Genetics of eastern white pine. USDA For. Serv. Res. Paper W0-9. 16pp.
- Wyman, D. and H. L. Flint 1967. Plant hardiness zone maps. Arnoldia 27: 53-56.
- 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 Genet. 26: 10-18.
- Yeh, F. Ch. -H. and D. O'Malley 1980. Enzyme variation in natural populations of douglas-fir, <u>Pseudotsuga menziesii</u> (Mirb.) Franco, from British Columbia. I. Genetic variation patterns in coastal populations. Silvae Genet. 29: 83-92.

Appendix I. Calculation formulae for genetic differentiation among provenances.

Allele diversity (Nei 1973). Ι. HT<sup>=</sup> 1-(  $x^{2}ik + xik \times jk)/s^{2}$ , k i=j k HS= 1-( x2ik)/s, DST= HT-HS, GST= DST/HT, where xik : frequency of the kth allele in the ith provenance, and s : number of provenance. 2. Genetic distance (Nei 1972).  $j_x = Xi^2$ ,  $jy = Yi^2$ ,  $j_{xy} = XiYi$ ,  $l^= J_{xy}/J_x J_y$ ,  $D^= -log_e l$ , where X, Y : provenance, xi, Yi : allele frequency of provenance X and Y,  $j_{\boldsymbol{x}},\,\boldsymbol{j}_{\boldsymbol{Y}}$  : probability of identity of two alleles in provenance X and Y,  $j_{x \mathsf{Y}} \mathbb{I}$  probability of identity of an allele from provenance X and an allele from provenance Y,  $J_{xy}$ ,  $J_x$  and  $J_y$ : arithmetic means of  $j_{xy}$ ,  $j_x$ , and  $j_y$ , respectively, and genetic distance. D

Prov			18	1 <b>9</b>	20	21	22	23	24	25	27	28	29	30
	16		.011	.064	.013	.050	.010	.091	.039	.036	.045	.012	.052	.020
	TO	18		.054	.033	.010	.010	.077	.045	.028	.043		.040	.047
	1		19		.036	.046	.048	.055	.026	.034	.042	.054 .024	.024	.008
	-	2		20		.018	.019	.042	.012	.011	.009	.024	.019	.042
2	.064		3		21		.038	.055	.036	.011	.019	.030	.010	.029
3	.053	.139		4		22		.091	.044	.026	.043	.080	.076	.060
4	.035	.108	.041		5		23	2.4	.032	.061	.058 .024	.080	.029	.026
5	.037	.084	.061	.019		6	•	24	25	.030	.024	.032	.013	.028
6	.026	.098	.040	.017	.013		8	0	25	27	.010	.059	.015	.022
8	.034	.084	.027	.022	.030	.037	001	9	10	21	28	.055	.055	.026
9	.021	.081	.040	.020	.019	.013	.021	010	TO	11	20	29		.036
10	.049	.093	.079	.028	.025	.021	.049	.018	.042	11	12	23		
11	.048	.106	.022	.020	.022	.027	.013	.024 .013	.042	.021		13		
12	.032	.109	.051	.017	.015	.015	.026 .038	.013	.025	.036	.019		14	
13	.029	.090	.057	.041	.025	.016	.038	.010	.027	.035	.022	.020		15
14	.016	.091	.041	.018	.026	.007	.037	.013	.028	.043	.020	.007	.016	
15	.024	.098	.056	.039	.026	.012	.044	.032	.061	.029	.042	.042	.029	.046
16	.038	.110	.021	.025	.053	.040	.017	.029	.062	.023	.036	.035	.034	.041
18	.046	.133	.029	.026	.040	.035 .022	.013	.025	.043	.055	.041	.043	.022	.036
19	.030	.085	.086	.025	.019 .030	.022	.020	.010	.036	.026	.024	.005	.010	.008
20	.016	.081	.031	.027	.030	.022	.020	.017	.026	.033	.018	.010	.032	.016
21	.050	.130	.063	.037 .013	.023	.022	.011	.021	.048	.018	.027	.044	.028	.040
22	.040	.121	.022	.013	.033	.054	.075	.052	.049	.065	.050	.035	.060	.047
23	.061	.071	.126	.072	.040	.026	.046	.018	.025	.054	.030	.020	.016	.022
24	.019	.068	.077	.033	.022	.010	.031	.013	.023	.027	.011	.013	.014	.012
25	.032	.125	.043	.018	.022	.017	.041	.011	.031	.046	.026	.007	.019	.005
27	.022	.086	.032	.047	.036	.034	.016	, 035	.061	.012	.040	.049	.034	.056
28	.051	.115	.020	.019	.030	.014	.047	.011	.031	.046	.026	.021	.017	.015
29 30	.033 .021	.127 .082	.026	.032	.035	.016	.034	.023	.041	.029	.030	.021	.012	.021

Appendix 2. Geneticdistances among 27 provenances of strobus.