Study on the Genetic Structure in *Pinus massoniana* (Masson Pine) Populations

By Q. Q. HUANG[1]) and Y. X. ZHANG[2])

(Received 23rd September 1999)

Summary

The result of isozyme analyses in the six natural populations of Masson pine (*Pinus massoniana* LAMB.) showed that the level of genetic variation in the population is relatively high. The proportion of polymorphic loci is 76.2%, the average number of alleles per locus is 2.39, the effective number of alleles per locus is 1.62, and average expected heterozygosity is 0.2730. On the other hand, the genetic differentiation between the populations is very low (G_{ST}=0.017, D=0.011 ±0.005); about 98% of the total genetic variation comes from the individuals within the populations. The genetic variation between the different populations is about 2%, and mainly originates from one third of the gene loci. The population sampled is almost a single panmictic unit, with most loci are in HARDY-WEINBERG equilibrium, some degrees of inbreeding in some loci.

Key words: Pinus massoniana, genetic structure, population.

Introduction

Since 1970, the gel electrophoresis technique of allozymes has been extensively used to study the genetic diversity of populations for forest trees, such as *Pinus ponderosa* (LINHART et al., 1981), *P. contorta* (DANCIR and YEH, 1983; WHEELER et al., 1982), *P. banksiana* (FINS and LIBBY, 1982), *P. sylvestris* (YAZDANI et al., 1985), *P. rigida* MILL. (GURIES and LEDING, 1982), *P. radiata* (PLESSAS and STRAUSS, 1986), *Abies mariesii* MAST. (SUYAMA, 1992), White spruce (FURNIER et al., 1991). The study on genetic structure of *Pinus massoniana* populations only in EST (esterases: EC 3.1.1.2), MDH (malate dehydrogenase: EC 1.1.37), POD (peroxidase: E.C. 1.11.1.7) was also reported by S. GE et al. (1988). Masson pine is one of the main woody species in South China, and it has different patterns because of random genetic drift, mutation migration, and natural selection in a long evolutionary process. Therefore, it is important that a great deal of information on the genetic structure of population of masson pine provide some help to guide the effects made to ensure gene conservation of the species.

Materials and Methods

Sampling

Masson pine is one of the main woody species in South China (thirteen provinces). Guizhou province is one of main areas for *Pinus massoniana*. As natural forest stands had been hard wrecked, a few natural populations in masson pine survive around distributing area, Guizhou province so is. The material obtained for this study was from six natural population of *Pinus massoniana* at Duyun, Suiyang and Jianhe counties, Guizhou province. The age of trees was on the average 25 years. A sample plot was established in each population, and height, diameter at breast height, and mid-height were measured. Respectively 21 and 49 trees were selected in each population. Cones were harvested on these trees and seeds extracted and stored at –20°C until analysis.

Electrophoresis

The seeds were soaked in running water at 4°C for 48 hours, and then incubated on moistened filter paper in petri dishes at 4°C for 3 weeks. Haploid megagametophytes were excised individually from stratified seeds for analysis by polyacrylamide gel electrophoresis on vertical slab gels. The electrophoresis was performed following the procedures of DAVIS (1964) and ORNSTEIN (1964) with slight modifications. Methods for the extraction and staining of enzymes were slightly modified from those described by CONKLE et al. (1982) and SHIRAISHI (1988). Thirteen enzymes systems were analyzed in this study. These enzyme systems were: PGM(E.C.2.7.5.1), GDH(E.C.1.4.1.2), MDH(E.C.1.1.37), SHDH(E.C.1.1.1.25), 6PGD(E.C.1.1.1.44), G6PD(E.C.1.1.1.49), DIA(E.C.1.6.4.3), GOT(E.C.2.6.1.1), EST(E.C.3.1.1), AMY(E.C.3.2.1), LAP(E.C.3.4.1.1), AAP (E.C.3.4.11.1), PGI(E.C.5.3.1.9). The inheritance of isozymes for each enzyme systems had been investigated by HUANG et al. (1994).

Quantitative analysis

Genetic diversity based on allele frequencies at each locus were obtained by the following eight criteria.

(1) Proportion of polymorphic loci (P)
(2) Number of alleles per locus (Na)
(3) Effective number of alleles per locus (Ne)
(4) Observed heterozygosity (Ho)
(5) Expected heterozygosity (He)
(6) WRIGHT's fixation index (F)
(7) Coefficient of gene differentiation (G_{ST})
(8) NEI's genetic distance (D)

The chi-square statistic was used to detect significant differences in allele frequencies among populations for each locus:

$$\chi^2 = 2NG_{ST}(a-1) \qquad df = (a-1)(n-1)$$

Where N is the total sample size, a is the number of alleles at the locus, and n is the number of populations. A significant χ^2 value indicates heterogeneity in allele frequencies among populations. Likewise, deviations of F_{IT} and F_{IS} from zero were tested by chi-square using the following formula:

$$\chi^2 = F^2N(a-1),\ df = a(a-1)/2$$

These two analyses reveal whether differences in the frequency of genotypes among populations are due to differences in allele frequencies alone or if they can also be attributed to non-random association into individuals.

Results

Gene diversity

Thirteen enzyme systems were analyzed and a total of 84 alleles at 28 loci were identified. Among these loci, 24 were

---

[1]) Institute of Environmental Applied Chemistry, Guangzhou Normal University, Jiefang north Rd., Guihuagang East 1, Guandong, Guazhou, China 510405

[2]) Guizhou Forest Research Institute, 287 Fuyuan South Rd., Guiyang, Guizhou, China

polymorphic, and thus the proportion of polymorphic loci is 76.2% (*Table 2*). The 4 monomorphic loci (Got-1, Shd-1, Mdh-4 and Lap) were fixed in all the six populations. Besides, 9 out of 24 polymorphic loci were fixed in some of the populations. But 15 loci more than half of total loci, were segregating in every population. That is particularly the case of Pgm, Aap-1, Est-1, Est-2, Est-3, Amy and G6p.

The number of alleles and the effective number of alleles per locus were the two other indices used to measure genetic variation. The results showed that the average number of alleles at a locus (Na) was 2.39, 18 loci had either three or four alleles, and 9 loci possessed more than 4 alleles. Est-2, Est-3 loci had the highest number of alleles (6 alleles). The average effective number of alleles per locus was 1.57, with 13 loci (46.4%) having an effective number of alleles greater than two. The Pgm, Aap-1 and Amy loci had an effective number of alleles of 4. The effective number of alleles per locus was much lower than the actual number of alleles, especially for the loci Aap-2, Dia-3, Dia-4, Est-2 and Est-4. This means that in the populations there were a lot of genes in which the frequencies were

Table 1. – Allele frequencies at 28 loci in 6 populations of *P. massoniana*.

Population	*Pgm*				*Pgi-1*			*Pgi-2*			*Got-1*	*Got-2*			*Got-3*	
	a	b	c	d	a	b	c	a	b	c	a	a	b	c	a	b
Yul	0.469	0.234	0.156	0.141	0.156	0.803	0.031	0.156	0.813	0.031	1.000	0.016	0.953	0.031	0.938	0.063
Yun I	0.333	0.167	0.333	0.167	0.000	0.976	0.024	0.000	0.976	0.024	1.000	0.000	1.000	0.000	0.952	0.048
Yun II	0.381	0.071	0.301	0.167	0.024	0.976	0.000	0.024	0.976	0.000	1.000	0.000	1.000	0.000	0.881	0.119
Jianhe	0.429	0.286	0.071	0.214	0.000	1.000	0.000	0.000	1.000	0.000	1.000	0.000	1.000	0.000	0.976	0.024
Yu 2	0.396	0.271	0.000	0.333	0.063	0.938	0.000	0.063	0.930	0.000	1.000	0.021	0.958	0.021	0.917	0.083
Suiyang	0.337	0.276	0.102	0.286	0.031	0.969	0.000	0.031	0.969	0.000	1.000	0.061	0.898	0.041	0.929	0.071
Average	0.387	0.229	0.158	0.226	0.051	0.940	0.009	0.051	0.940	0.009	1.000	0.024	0.955	0.021	0.932	0.068
X^1	33.56**				27.01**			27.01**				5.38			0	

Population	*shd-1*	*shd-2*				*Aap-1*				*Aap-2*				*Aap-3*		
	a	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c
Yul	1.000	0.016	0.203	0.109	0.672	0.016	0.484	0.297	0.203	0.000	0.000	1.000	0.000	0.000	1.000	0.000
Yun I	1.000	0.000	0.095	0.214	0.690	0.214	0.310	0.310	0.167	0.000	0.024	0.976	0.000	0.000	1.000	0.000
Yun II	1.000	0.000	0.190	0.214	0.595	0.167	0.429	0.095	0.310	0.071	0.000	0.929	0.000	0.024	0.976	0.000
Jianhe	1.000	0.000	0.000	0.238	0.762	0.238	0.262	0.214	0.286	0.048	0.000	0.952	0.000	0.071	0.929	0.000
Yu 2	1.000	0.042	0.104	0.354	0.500	0.250	0.271	0.208	0.271	0.000	0.000	0.958	0.042	0.000	0.968	0.042
Suiyang	1.000	0.000	0.102	0.316	0.582	0.071	0.214	0.276	0.439	0.000	0.000	0.990	0.010	0.000	0.990	0.010
Average	1.000	0.009	0.119	0.247	0.625	0.137	0.318	0.244	0.361	0.015	0.003	0.973	0.009	0.012	0.979	0.009
X^1		17.54				24.90				7.96				9.74		

Population	*Mdh1*		*Mdh2*		*Mdh-3*			*Mdh-4*	*Dia-1*			*Dia-2*		*Dia-3*			
	a	b	a	b	a	b	c	a	a	b	c	a	b	a	b	c	d
Yul	0.875	0.125	0.641	0.359	0.969	0.016	0.016	1.000	0.203	0.656	0.141	0.281	0.719	0.016	0.984	0.000	0.000
Yun I	0.714	0.286	0.619	0.381	0.952	0.048	0.000	1.000	0.095	0.786	0.119	0.143	0.857	0.000	0.929	0.024	0.048
Yun II	0.810	0.190	0.524	0.476	0.929	0.071	0.000	1.000	0.225	0.550	0.225	0.143	0.857	0.024	0.905	0.000	0.071
Jianhe	0.810	0.190	0.810	0.190	0.881	0.119	0.000	1.000	0.214	0.667	0.119	0.238	0.762	0.000	1.000	0.000	0.000
Yu 2	0.760	0.250	0.583	0.417	0.958	0.042	0.000	1.000	0.237	0.711	0.053	0.146	0.854	0.00	1.000	0.000	0.000
Suiyang	0.847	0.153	0.714	0.286	0.969	0.031	0.000	1.000	0.068	0.852	0.080	0.133	0.867	0.041	0.959	0.000	0.000
Average	0.813	0.188	0.658	0.342	0.949	0.048	0.003	1.000	0.159	0.723	0.118	0.179	0.821	0.018	0.964	0.003	0.015
X^1	1.68		8.74		1.28				10.95			0.67		2.82			

Population	*Est-1*			*Est-2*						*Est-3*						*Lap-1*
	a	b	o	a	b	c	d	e	o	a	b	c	d	e	o	a
Yul	0.266	0.563	0.172	0.000	0.517	0.033	0.017	0.000	0.433	0.016	0.047	0.750	0.156	0.000	0.031	1.000
Yun I	0.119	0.667	0.214	0.000	0.762	0.000	0.000	0.000	0.238	0.000	0.000	0.619	0.071	0.000	0.310	1.000
Yun II	0.167	0.548	0.286	0.000	0.833	0.024	0.000	0.119	0.024	0.000	0.071	0.452	0.310	0.000	0.167	1.000
Jianhe	0.381	0.333	0.286	0.000	0.450	0.025	0.075	0.100	0.350	0.000	0.024	0.500	0.286	0.024	0.167	1.000
Yu 2	0.250	0.375	0.375	0.000	0.523	0.023	0.000	0.273	0.182	0.021	0.042	0.396	0.271	0.000	0.271	1.000
Suiyang	0.173	0.449	0.378	0.042	0.448	0.125	0.010	0.063	0.313	0.000	0.051	0.439	0.235	0.020	0.255	1.000
Average	0.220	0.485	0.295	0.012	0.562	0.052	0.015	0.083	0.275	0.006	0.042	0.524	0.220	0.009	0.199	1.000
X^1	20.70*			117.94**						53.42**						

(continuation *table 1*)

Population	*Dia-4*				*Amy*					*G6p*			*Gdh*		
	a	*b*	*c*	*o*	*a*	*b*	*c*	*d*	*e*	*a*	*b*	*c*	*a*	*b*	*o*
Yul	0.047	0.953	0.000	0.000	0.203	0.266	0.375	0.156	0.000	0.375	0.625	0.000	0.156	0.844	0.000
Yun I	0.000	1.000	0.000	0.000	0.333	0.119	0.452	0.048	0.048	0.262	0.690	0.048	0.048	0.929	0.024
Yun II	0.119	0.881	0.000	0.000	0.071	0.500	0.262	0.167	0.000	0.286	0.667	0.048	0.071	0.929	0.000
Jianhe	0.167	0.833	0.000	0.000	0.095	0.286	0.548	0.024	0.048	0.381	0.571	0.048	0.071	0.905	0.024
Yu 2	0.146	0.813	0.021	0.021	0.104	0.313	0.417	0.125	0.042	0.333	0.646	0.021	0.021	0.979	0.000
Suiyang	0.071	0.888	0.020	0.020	0.112	0.214	0.582	0.071	0.020	0.367	0.602	0.031	0.041	0.939	0.020
Average	0.086	0.896	0.009	0.009	0.149	0.271	0.458	0.098	0.024	0.342	0.628	0.030	0.068	0.920	0.012
X^1	33.16**				47.85**					0			8.13		

Population	*6Pg-1*		*6Pg-2*	
	a	b	a	b
Yul	0.109	0.891	0.047	0.953
Yun I	0.238	0.762	0.000	1.000
Yun II	0.175	0.825	0.000	1.000
Jianhe	0.000	1.000	0.000	1.000
Yu 2	0.146	0.854	0.021	0.979
Suiyang	0.153	0.847	0.000	1.000
Average	0.138	0.862	0.012	0.988
X^1	9.21		0	

*) Significant at 5% level

**) Significant at 1% level

low and not contributing much to the population genetic variation. However, many loci were in strong genetic differentiation. In addition, at the species level expected heterozygosity (He) was 0.273. On average, 76.2% of the loci were polymorphic within populations, with population values ranging from 64.3% (in Jianhe) to 82.1% (in Yu2 and Suiyang). Within populations, there was an average of 2.39 alleles per locus.

Genetic structure

Significant differences in allele frequencies were found among populations for 8 polymorphic loci; with no significant differences detected for the remaining 14 (*Table 1*). The highest D_{ST} and G_{ST} values were 0.040, 0.070 respectively. Mean gene diversity among populations (D_{ST}-value) and the mean coefficient of genetic differentiation (G_{ST}-value) were 0.008, 0.017 respectively (*Table 3*), which indicated that only about 2% of the genetic variation was between populations, while about 98% of the species' total variation is occurred within populations.

Analyses of NEI's genetic distance between following populations showed that the mean D-value was low (0.001). Yu2 and Suiyang populations were very similar (D = 0.003). Greater differences (D-value=0.010) were found between Yun I and Yun II, despite the two study populations were located 10 km from each other, in other words, its means that there is a significant negative correlation between genetic distance and geographic distance among populations. Dendrogram from clustering based on NEI's genetic distance for 6 natural populations is given by *figure 1*.

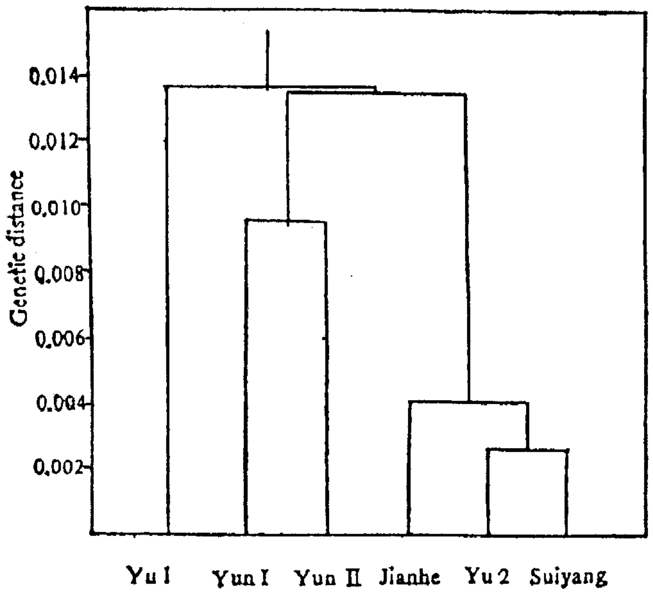

Figure 1. – Dendrogram from clastering based on NEI's genetic distance for the six populations of *P. massoniana*.

Mean observed and expected heterozygosities did not differ significantly for the 5 populations excepted Yul, that is also the case for the mean F_{IS}-value (0.0671) and F_{IT}-value (0.0932) which are not significant different from zero. This showed that most loci are in HARDY-WEINBERG equilibrium. Indeed, the Chi-square analysis of F_{IS}-values indicated significant overall deviations from zero for Got-3, Shd-2, Mdh-3, Dia-1, Dia-3, Est-2 and Amy, and excess homozygosity at these loci. Similar analysis of F_{IT} estimates showed the same results. These results correspond with *P. rigiada, Oenothera grandis* and *Desmodium undiflorm* already published by other authors. But

Table 2. – Total and mean of member of alleles frequency of the most Common alleles, and number of fixed population at 28 loci in 6 populations of *P. massoniana.*

Loci	Number of alleles				Frequence of the most common alleles		No. of fixed populations
	Nai	$\bar{N}ai$	*Nei*	$\bar{N}ei$	$\bar{X}i$	Range	
Pgm	4	3.82	4	3.83	0.232	0.071–0.469	0
Pgi-1	3	2.00	1	1.00	0.945	0.813–1.000	1
Pgi-2	3	2.00	1	1.00	0.945	0.813–1.000	1
Got-1	1	1.00	1	1.00	1.000	1.000	6
Got-2	3	2.00	1	1.00	0.968	0.898–1.000	3
Got-3	2	2.00	1	1.00	0.923	0.881–0.976	0
Shd-1	1	1.00	1	1.00	1.000	1.000	6
Shd-2	4	3.17	3	2.17	0.643	0.500–0.762	0
Aap-1	4	4.00	4	3.50	0.280	0.095–0.484	0
Aap-2	4	1.83	1	1.00	0.968	0.929–1.000	1
Aap-3	3	1.67	1	1.00	0.975	0.929–1.000	2
Mdh-1	2	2.00	2	1.33	0.801	0.714–0.857	0
Mdh-2	2	2.00	2	1.83	0.500	0.190–0.810	0
Mdh-3	3	2.17	1	1.00	0.943	0.881–0.969	0
Mdh-4	1	1.00	1	1.00	1.000	1.000	6
Dia-1	3	3.00	2	1.83	0.704	0.550–0.852	0
Dia-2	2	2.00	2	1.33	0.819	0.719–0.867	0
Dia-3	4	2.00	1	1.00	0.963	0.905–1.000	2
Dia-4	4	2.50	1	1.00	0.895	0.813–1.000	1
Est-1	3	3.00	3	2.50	0.333	0.119–0.667	0
Est-2	6	4.17	3	2.33	0.423	0.024–0.833	0
Est-3	6	4.50	3	2.67	0.322	0.031–0.750	0
Amy	5	4.67	4	3.00	0.291	0.071–0.548	0
G6p	3	2.83	2	2.00	0.484	0.262–0.690	0
6pg-1	2	1.83	2	1.17	0.863	0.752–1.000	1
6pg-2	2	1.33	1	1.00	0.989	0.953–1.000	4
Lap-1	1	1.00	1	1.00	1.000	1.000	6
Gdh	3	2.50	1	1.00	0.921	0.844–0.972	0

some loci show an exceeding and insufficient heterozygosity. Therefore we should pay attention to this in future breeding.

Discussion and Conclusion

Many published papers showed that a high amount of genetic diversity exists in conifer species. HAMRICK and GOUT (1989) reported that the percentage of polymorphic loci (P), average number of alleles per locus (Na), and expected heterozygosity (He) for 55 gymnosperms were 70.9%, 2.35, 0.173 respectively. S. GE et al. (1988) also reported similar results (P = 61.5%, Na = 2.26, H = 0.206) for 25 conifer species. S. GE (1985) studied genetic variations for Got, Mdh, Gdh and Ldh in *P. massoniana,* and founded following values: P = 64.5%, Na = 1.65, He = 0,216, compared with these results, the percentage of polymorphic loci (p = 76.2%), average number of alleles per loci (Na = 2.39), effective number of alleles per loci (Ne = 1.62) and expected heterozygosity (He = 0.273) for *P. massoniana* in this study were somewhat higher.

Lower observed G_{ST}- value in *P. massoniana,* however, is the partitioning of genetic variation among populations. In general, conifer have a G_{ST}-value around 6.8% (HAMRICK and GOUT, 1989). Higher G_{ST}-value were reported in some pine species, for example, *P. pungens* (0.135, GIBSON and HAMRICK, 1991), *P. halepensis* MILL. (0.300, SCHILLER et al., 1986). Compared with these wind-pollinated, outcrossed, and long lived, woody species, the G_{ST}-values (0.0172) found here is the lowest one, and suggest a very small genetic differentiation among populations of *P. massoniana.*

Continuously distributed populations can be substructured into smaller individual breeding units as a result of random genetic drift, mutation, and differential natural selection in a long evolutionary process. On the other hand, high rates of gene flow via pollen and seeds, result in little spatial genetic differentiation of populations. This phenomenon has been shown in *P. pungens* (GIBSON and HAMRICK, 1991), *P. rigida* (GURIES and LEDIG, 1982), *P. contorta* (WHEELER and GURIES, 1982), *P. ponderosa* (HAMRICK et al., 1989). When factors act to restrict gene flow, the opportunity for genetic divergence decreases. For wind-pollinated, highly outcrossed species such as pines, geographic isolation of populations would be an

Table 3. – Estimates of genetic diversity parameters for 28 loci surveyed in 6 *P. massoniana* populations.

Loci	H_T	H_S	D_{ST}	G_{ST}	F_{IT}	F_{IS}	F_{ST}
Pgm	0.7245	0.7031	0.0241	0.0333	0.0369	0.0075	0.0295
Pgi-1	0.1046	0.1004	0.0042	0.0402	0.1032	0.0066	0.0402
Pgi-2	0.1046	0.1004	0.0042	0.0402	0.1032	0.0066	0.0402
Got-1	0.0000	0.0000	0.0000	0.0000	–	–	–
Got-2	0.0622	0.0617	0.0005	0.0080	0.0322	0.0024	0.0080
Got-3	0.1277	0.1277	0.0000	0.0000	0.2638**	0.2638**	0.0000
Shd-1	0.0000	0.0000	0.0000	0.0000	–	–	–
Shd-2	0.5237	0.5195	0.0092	0.0174	0.1125**	0.0968**	0.0174
Aap-1	0.7368	0.7186	0.0182	0.0247	0.0801	0.0568	0.0247
Aap-2	0.0636	0.0631	0.0005	0.0079	–0.0189	–0.0269	0.0079
Aap-3	0.0482	0.0475	0.0007	0.0145	–0.0166	–0.0316	0.0145
Mdh-1	0.3199	0.3188	0.0017	0.0050	0.0178	0.0144	0.0034
Mdh-2	0.4574	0.4455	0.0119	0.0260	0.0942	0.0070	0.0260
Mdh-3	0.1081	0.0179	0.0002	0.0019	0.0962*	0.0945*	0.0019
Mdh-4	0.0000	0.0000	0.0000	0.0000	–	–	–
Dia-1	0.4615	0.4540	0.0275	0.0163	0.3417**	0.3308**	0.0163
Dia-2	0.2970	0.2964	0.0006	0.0020	0.1313	0.1296	0.0020
Dia-3	0.0727	0.0725	0.0002	0.0028	0.4938**	0.4924**	0.0028
Dia-4	0.1913	0.1850	0.0063	0.0329	0.0240	–0.0240	0.0329
Est-1	0.6303	0.6109	0.0194	0.0308	–0.0435	–0.0625	0.0308
Est-2	0.5794	0.5387	0.0407	0.0702	0.4525**	0.3109**	0.0702
Est-3	0.6346	0.6144	0.0202	0.0318	0.0462	0.0148	0.0318
Amy	0.6959	0.6709	0.0247	0.0356	0.1765**	0.1462*	0.0355
G6p	0.4919	0.4919	0.0000	0.0000	0.0876	0.0876	0.0000
Gdh	0.1482	0.1464	0.0018	0.0121	0.0216	0.0096	0.0121
6pg-1	0.2370	0.2305	0.0065	0.0274	–0.0181	–0.0469	0.0274
6pg-2	0.0225	0.0225	0.0000	0.0000	–0.0088	–0.0088	0.0000
Lap-1	0.0000	0.0000	0.0000	0.0000	–	–	–
Average	0.2803	0.2732	0.0080	0.0172	0.0932	0.0671	0.0170

*) Significant at 5% level

**) Significant at 1% level

effective factor to inhibit gene flow (HAMRICK, 1987). In *P. rigida,* whose natural distribution has been reduced to three isolated mainland populations and two island populations, 16.2% of its variation is found among populations (MORAN et al., 1988). *P. halepensis* MILL., a species distributed in discrete populations around the Mediterranean Sea, 30.0% of its variation is found among populations (SCHILLER et al., 1986). For both species, limited gene flow due to population isolation is probably the main factor contributing to interpopulational differentiation. Masson pine is distributed along a continuous mainland population in southern China. However, indirect estimate of Nm based on G_{ST}-value (14.45) was higher than the values obtained for other pine species (HAMRICK, 1987; GIBSON and HAMRICK, 1991; SCHILLER et al., 1986). At a species level, *P. massoniana* maintains the high amount of genetic diversity, and has been the little genetic differentiation (G_{ST} = 0.0172) among populations. High rate of gene flow rates, which via pollen or seeds may be explanatory note to this phenomenon.

Generally, the correlation between genetic and geographic distances among populations tend to be more genetically similar. But the lower coefficient between genetic and geographic distances among population would occurs because of high gene flow, mutation, migration, natural selection and geographic isolation in small area. For wind-pollinated, highly outcrossed species such as pines, through high gene flow rates via pollen or seeds, lower genetic distance is frequently found among populations separated by far geographic distance. On the contrary, higher genetic distance may be found among the near populations because of mutation, natural selection and special geographic isolation. Yu1 and Yu2, Yun I and Yun II were located in the same area respectively, and relatively high genetic distances (D = 0.013, 0.009) have been shown between them. This phenomenon may be explained by natural selection, mutation and limited gene flow due to the special population isolation. Yu2 and Suiyang were separated by far geographic distance, and lower genetic distance (D = 0.003) has been found between the two populations, it means mutation, natural selection and special geographic isolation are probably the main factors contributing to genetic differentiation.

References

DANCIK, B. P. and YEH, F. C.: Allozyme variability and evolution of lodgepol pine (*Pinus contorta* var. Lafifolia) and Jack pine (*P. banksiana*) in Alberta. Can. J. Genet. Cytol. **25**: 57–64 (1983). — DAVIS,

B. J.: Disk Electrophoresis II: Method and application to human serum proteins. Ann. N. Y. Acad. **121**: 404–427 (1964). — FINS, L. and LIBBY, W. J.: Population Variation in Sequoiadendron: Seed and Seedling Studies, Vegetative Propagation, and Isozyme Variation. Silvae Genet. **31**: 102–110 (1982). — FURNIER, G. R., MOHN, M. S. C. A. and CLYDE, M. A.: Geographic Patterns of Variation in Allozymes and Height Growth in White Spuce. Can. J. For. Res. **21**: 707–711 (1991). — GIBSON, J. P. and HAMRICK, J. L.: Genetic Diversity and Structure in *Pinus pungens* (table Mountain Pine) Population. Can. J. For. Res. **21**: 635–642 (1991). — GULLBERG, U. et al.: Allozymes Variation in Scots Pine (*Pinus sylvestris* L.) in Sweden. Silvae Genet. **34**: 193–201 (1985). — GURIES, R. P. and LEDIG, F. T.: Genetic Diversity and Population Structure in Pitch Pine (*Pinus rigida* MILL.). Evolution (Kawrence, Kans.) **36**: 378–420 (1982). — HAMRICK, J. L.: Gene Flow and Genetic Variation in Plant Populations in Differentiation Patterns in Higher Plants. Edited by K. URBANSKA. Academic Press, New York. Pp. 53–67 (1987). — HAMRICK, J. L. and GODT, J. W.: Allozyme Diversity in Plant Species. In: Plant Population Genetic, Breeding and Gene Resources. Edited by A. D. H. BROWN, M. T. CLEGG, A. L. KAHLER and B. S. WIER: Sinauer Associates Inc., Sunder Land, Massachusetts. Pp. 43–63 (1989). — HUANG, Q. Q. et al.: Genetic Control of Isozyme Variation in Masson Pine, *Pinus massoniana* LAMB. Silvae Genet. **43**: 285–291 (1994). — KIMURA, M.: Evolutionary Rate at the Molecular Level. Nature (London) **217**: 624–626 (1968). — O'MALLEY, D. M.: Inheritance of Isozyme Variation and Heterozygosity in *Pinus pondersa*. Biochemical Geneticas **17**: 233–250 (1979). — ORNSTEIN, L.: Disk Electrophoresis-II: Background and Theory. Ann. Acad. Sci. **121**: 321–349 (1964). — PLESSAS, M. E. and STRAUSS, S. H.: Allozyme Differentiation among Populations, Stands, and Cohorts in Monterey Pine. Can. J. For. Res. **16**: 115–164 (1986). — SCHILLER, G. et al.: Isozyme Variation among Nature Stand and Plantations of Aleppo Pine in Israel. Isr. J. Bot. **35**: 161–174 (1986). — S. GE et al.: The Genetic Variation in Natural stands of Masson pine (Chinese). Scientic Silvae Sinicae **24**(4): 399–408 (1988). — SHIRAISHI, S.: Inheritance of Isozyme Variation in Japanese Black Pine, *Pinus thunbergii* parl. Silvae Genet. **37**: 93–100 (1988). — SUYAMA: Inheritance of Isozyme Variation in Natural Populations of *Abies mariesii* MAST. Thesis for Doctorate in University of Tsukuba (1992). — WHEELER, N. C. and GURIES, R. P.: Populations Structure, Genetic Diversity and Morphological Variation in *Pinus contorta* DOUGL. Can. J. For. Res. **12**: 595–606 (1982). — YAZDANI, R., MUONA, O., RUDIN, D. and SZMIDT, A. E.: Genetic Structure of a *Pinus sylvestris* L., Seed-tree Standard Naturally Regenerated Understory. For. Sci. **31**: 430–436 (1985).

Optimising Early Selection Using Longitudinal Data

By L. A. APIOLAZA1)2)3), D. J. GARRICK1) and R. D. BURDON2)

(Received 17th January 2000)

Abstract

This study analysed the use of longitudinal data, i.e. repeated assessment of the same individuals at different ages, in the context of early selection. Autoregressive relationships, banded correlations and unstructured ('unsmoothed') matrices were used to model the additive genetic covariance matrix ($\mathbf{G}_0$) for 10 total height measurements of a *Pinus radiata* open-pollinated progeny test. We examined the effects on response to selection of inferred covariance structure, mass versus combined selection, one or multiple assessments, and two breeding-delay intervals. End results are expressed as predicted average gain per year. The patterns of predicted response to selection vary widely between inferred covariance structures. Considering the autoregressive model (based on logarithm of age ratios between assessments) as an example, the effect of combining information from relatives on response to selection is more important (16% to 41% extra gain) than using extra measurements (2% to 25%), when predicting individual breeding values, although the economics of extra gain vs extra assessment costs must be carefully analysed. It is expected that using multiple assessments could be advisable for datasets with lower genetic autocorrelations. An approximate comparison across covariance models showed the autoregressive model to exhibit the best ability to produce 'correct' selections as well as the highest predicted response to selection.

Key words: longitudinal data, early selection, covariance structures, tree model, *Pinus radiata*.

Introduction

Trees included in breeding programs are often evaluated in progeny trials to predict their genetic value. Results from testing determine the participation of the trees in the breeding population, as well as their use as parents of future plantations (ZOBEL and TALBERT, 1984; WHITE, 1987). The breeding objective includes tree characteristics at harvest age (e.g. volume and wood density); however, progeny tests are assessed at one or more early ages, often less than half the rotation age. The problem of early selection arises with the existence of less than perfect genetic and phenotypic correlations between performance at early assessments and performance at harvest age.

Extending the testing period increases accuracy of selection, i.e. the correlation between predicted and real breeding values, but also increases financial costs and time delays to achieve gain. Optimising response to selection for a given objective involves finding the appropriate combination of accuracy and evaluation time. Traditionally, this has been achieved by calculating gain for different selection ages using the formula for correlated response to mass or index selection, which includes the heritability at early and mature ages and genetic correlation between the ages (SEARLE, 1965; FALCONER and MACKAY, 1996). The selection age that maximises either a biological criterion (response per year, e.g. LAMBETH, 1980) or an economic criterion (net present value, e.g. NEWMAN and WILLIAMS, 1991) may be chosen.

---

1) Institute of Veterinary, Animal and Biomedical Sciences, Massey University, Palmerston North, New Zealand.

2) New Zealand Forest Research Institute, Private Bag 3020, Rotorua, New Zealand.

3) Current address: CRC for Sustainable Production Forestry, School of Plant Science, University of Tasmania, GPO Box 252-55, Hobart, Tasmania 7001, Australia.