Genotype X Environment Interaction and Genotypic Stability in Loblolly pine*

IV. Correlation studies

By F. Owino[1])

(Received July / September 1976)

Introduction

Genetic gain to be realized from a selection program depends on three factors (1) intensity of selection. (2) magnitude of additive genetic variance and (3) the accuracy of selection or the ability to select superior genotypes. The intensity of selection depends on the proportion of individuals saved and Burdon and Shelbourne (1971) have discussed the limits to stretching gains by manipulating this component. Additive genetic variance measures the breeding value of genotypes and there is very little that a breeder can do to change this component. The accuracy of selection is measured by the correlation between the selected individual's breeding value and the criterion upon which selection is based. This is practically the only factor that the breeder can alter to increase gains from his selection.

The criterion upon which selection is based can be (a) individual phenotype, (b) mean of individual's progeny, (c) performance of the same individual tested in different environments, and (d) pedigree estimates of the individual; *i.e.*, the performance of its relatives. All these criteria are currently used by tree breeders. Thus, selection in the wild stands corresponds to (a) abuve while progeny in several environments combines (b) and (c) above. Clonal tests, as proposed by Libby (1964) corresponds to (c) while criterion (d) is becoming more and more important in advanced generation breeding.

In recent tree breeding literature, there has been considerable discussion about the relative merits of aids to selection such as progeny testing, clonal testing, and the use of pedigree information in the form of combined family and individual selection indices (Libby 1969, Namkoong *et al.* 1966) Shelbourne 1969, Burdon and Shelbourne 1971, van Buijtenen 1972, Zobel *et al.* 1972). These authors have, however, based their arguments on genetic gain predictions according to selection theory. One of the several assumptions on which this theory rests is that there is no genotype X environment interaction. Thus, the merits of progeny testing has been assessed mainly in terms of the added gains to the rogued clonal or seedling seed orchard. When genotpye X environment interaction proves important, then progeny testing becomes even more beneficial. That is, it then fulfills two important functions:

(i) to provide information which can form the basis for genetic thinning of seed orchards and hence extra genetic gains.

(ii) when tested in several locations and years, to reduce the environmental "noise" that would otherwise be associated with measurement of the selection unit.

Progeny tests have a third function which is not relevant to the present discussion; if based on properly designed mating schemes, they form the genetic resource for advanced selection and breeding.

The experiment reported here was set up in an attempt to study the merits of progeny testing in several locations relative to testing in one location. In addition, the merits of using a common genetic check are assesed. The specific objectives of the study were:

1. to measure the repeatability of family means across environments
2. to measure the relative efficiency of progeny testing in p locations as opposed to testing in one location.
3. to compare observed family means with the most probable family means after adjustments for effects of genotype X environments interaction.
4. to evaluate the merits of using a common genetic check as has been generally adopted in tree breeding.

Genetic Background

If P is the phenotypic measure and A is the breeding value, then the accuracy of selection is, by definition, the correlation between A and P (r_{AP}). In the simple case of mass selection $r_{AP} = h$ where h is the square root of the narrow sense heritability. The response to selection in this case is $R = r_{AP}.i.\ \sigma_A = h.i.\sigma_A$ where i = intensity of selection in standard units and σ_A = the square root of the additive genetic variance.

If the breeder uses the family mean at a location as his selection unit and if this family mean is the average performance over n replications with k trees to a plot, then progeny testing in p locations increases the response to selection by increasing h as follows:

$$h_p = \sqrt{\frac{pnk}{1 + (pnk - 1)r_f}}\ .h$$

where r_f is the repeatability of family means across the environments and h_p is the square root of the heritability based on progeny testing in p locations. It can be readily seen that progeny testing increases the heritability and hence the genetic gain by a factor

$$\sqrt{\frac{pnk}{1 + (pnk - 1)r_f}}.$$

Since, in tree breeding progeny testing, the combination of pnk is usually large, the critical factor in progeny test efficiency is the repeatability of family means with the efficiency being greater the lower the repeatability.

The relative efficiency of selection (Q) based on progeny test at p locations relative to that at one location has the relationship

$$Q = 100\sqrt{\frac{pnk}{1 + (pnk - 1)r_f}}$$

---

* This study was carried out in partial fulfilment of the requirements for the Ph. D. degree, North Carolina State University, 1975. Part V of this paper was already published in Silvae Genetica 26, 4 (1977).

[1]) Lecturer, Eotany Department, University of Nairobi, Kenya.

A common genetic check used in the North Carolina State Cooperative Tree Improvement Programme consists of the half-diallel of crosses among the parents used as tester males i.e. a mixture of six crosses among tester males used to weight family means in space and time.

LUSH (1943) applied these concepts to practical breeding situations where the object is for the breeder to determine the "most probable" worth of the family ($\hat{P}_f$) devoid of the effects of genotype × environment interaction. According to his formulation,

$$\hat{P}_f = \bar{P} + [\frac{pnk(r_f)}{1 + (pkn - 1)r_f}(P_f - \bar{P})]$$

where P_f is the average of the individual's phenotype and $\bar{P}$ is the average phenotypic value of the population of origin. The quantity $(P_f - \bar{P})$ is the selection differential *i.e.* the difference between the selected individual and the mean of the population from which it is selected.

There are four different methods of progeny testing that tree breeders use. The first method, and one that is rarely used, is progeny testing at a single location. Secondly, one can test all the families in the same p locations. When there is enough seed for p values of 3 to 5, this is the optimum method. However, often there is difficulty with effecting all the required pollination either because of pollen scarcity or excessive work load. It is therefore common practice to test different families in different locations or years. When a common genetic check is included to adjust for environmental differences, then this consitutes the third progeny test method. When no such genetic check is used, this consitutes the fourth progeny test method.

Materials and Methods

To avoid in-balance, only three test locations were used for correlation studies: Halifax Co. N.C., Bullock Co. Ga., and Murray Co. Ga. Height growth and fusiform rust score were analysed for four sets of families — Weyerhaeuser open-pollinated progeny, Hoerner-Waldorf within orchard crosses, Weyerhaeuser × Continental Can (La.) crosses, and Hoerner-Waldorf × Westvaco crosses.

Statistical model and data analysis

The model used for variance component and intra-class correlation estimation was as follows:

$$Y_{ijk} = mu + g_i + l_j + (gl)_{ij} + b_{ijk}$$

where Y_{ijk} = mean performance of the i^{th} set of families in the k^{th} replicate in the j^{th} location.
mu = overall mean of all families over all locations.
g_i = average effect of the i^{th} set of families.
$(gl)_{ij}$ = interaction effect of the i^{th} set of families and the j^{th} location.
l_j = average effect of the j^{th} location.
b_{ijk} = effect of the k^{th} replication.

For balanced data and assuming a random statistical model, the appropriate expectation mean squares in the analysis of variance are as shown in *table 1*. Expressions for calculation of intra-class correlations together with their standard errors for four different methods of progeny testing have been formulated by NORDSKOG and KEMPTHORNE (1960). Their formulations have been modified to forms that are applicable to the experimental tree breeding situation as presented in *appendix 1*. An adjustment in genotypic variances necessary to remove the contribution due to differences in genotypic variance from location to location is shown in *appendix 2*.

Table 1. — Form of analysis of variance on plot mean basis used in the calculation of repeatability of family mean.

Source	DF	Expected mean squares[1])
Family set	g — 1	$\sigma^2_{w/k} + \sigma^2_p + n\sigma^2_{gl} + nl\sigma^2_g$
Location	l — 1	$\sigma^2_{w/k} + \sigma^2_p + n\sigma^2_{gl} + ng\sigma^2_l$
Family × location	(g — 1) (l — 1)	$\sigma^2_{w/k} + \sigma^2_p + n\sigma^2_{gl}$
Error	gl (n — 1)	$\sigma^2_{w/k} + \sigma^2_p$

[1]) σ^2_w = within plot variance.
σ^2_p = plot to plot variance.
σ^2_{gl} = variance due to family × location interaction.
σ^2_g = variance due to family differences.
σ^2_l = variance due to location differences.
n = number of replication.
k = harmonic mean of the number of trees per plot.

Appendix 1. — Formulations for the calculation of repeatability of family means (r_f) and their standard errors ($S.E._{r_f}$) for four methods of progeny testing.

Test 1.	$r_f = \frac{\sigma^2_g{}^{1)}_l + \sigma^2_{gl}}{\sigma^2_g + \sigma^2_{gl} + (\sigma^2_{w/k} + \sigma^2_p)/n}$
Test 2.	$r_f = \frac{\sigma^2_g}{\sigma^2_g + \sigma^2_{gl}/p + [(\sigma^2_{w/k}\sigma^2_{gl})/p]}$
Test 3.	$r_f = \frac{\sigma^2_g}{\sigma^2_g + 2[(\sigma^2_{w/k} + \sigma^2_p)/pn + \sigma^2_{gl}/p]}$
Test 4.	$r_f = \frac{\sigma^2_g}{\sigma^2_g + (\sigma^2_{gl} + \sigma^2_l)/p + (\sigma^2_{w/k} + \sigma^2_p)/pn}$

[1]) The same as given in table 1.

Appendix 2. — Adjustment for genotype × environment interaction variance and the calculation of their corresponding standard errors.

The estimated genotype × environment interaction variance can be due to two different phenomena — low genetic correlation from location or differences in genotypic variances from location to location. Thus,

$$\sigma^2_{gl} = \sum_{i \neq j} [(\sigma_i - \sigma_j)^2 + 2\sigma_i\sigma_j (1 - p_{ij})]/p(p - 1)$$

where σ_i = square root of the genotypic variance for the genotypes in location i.
σ_j = square root of the genotypic variance for the genotypes in location j.
p_{ji} = genetic correlation of genotypes in locations i and j.
p = number of locations.

The contribution from differences in genotypic variance was estimated as follows (SCHULTZ and BERNARD, 1967):

$$\text{Var } (V_{gl}) = \sum_{i \neq j} (\sigma_i - \sigma_j)^2 p\,(p - 1)$$

The adjusted interaction variance was therefore calculated as follows: —

$$\sigma^2_{gl\ (adjusted)} = \sigma^2_{gl} - \text{Var } (V_{gl})$$

The standard error of a component (S.E.)

This was calculated as $\text{S.E. } (\sigma^2_i) = \sqrt{\frac{2}{c^2}\sum \frac{(M \cdot S_i)^2}{d \cdot f_i + 2}}$ where c = the coefficient of the corresponding component, MS_i = mean square involved in the computation of the component and df_i = degrees of freedom corresponding to the i^{th} mean square.

Table 2. — Components of variance estimates for height growth and fusiform rust score.

Component[1]) of variance	Sets of families			
	Weyerhaeuser O.P.	Hoerner-Waldorf × Hoerner-Waldorf	Weyerhaeuser × Continental Can (La.)	Hoerner-Waldorf × Westvaco
		Height growth		
σ^2_g	.64 ± .979	.34 ± .948	1.08 ± 1.476	.56 ± .596
σ^2_l	6.30 ± 6.724	1.15 ± 1.715	9.11 ± 6.554	45.16 ± 24.444
σ^2_{gl}	.63 ± .472	.53 ± .645	1.30 ± .715	.53 ± .283
$(\sigma^2_{w/k} + \sigma^2_p)$	.32 ± .205	2.02 ± .673	.42 ± .101	.28 ± .050
		Rust score		
σ^2_g	.009 ± .0003	.093 ± .118	Negative	.027 ± .042
σ^2_l	.228 ± .052	.465 ± .456	.003 ± .012	.669 ± .387
σ^2_{gl}	.001 ± .000	.094 ± .079	.020 ± .009	.021 ± .028
$(\sigma^2_{w/k} + \sigma^2_p)$	.051 ± .0002	.072 ± .031	.007 ± .003	.021 ± .004

[1]) σ^2_w = within plot variance.
σ^2_p = plot to plot variance.
σ^2_{gl} = variance due to family × location interaction.
σ^2_g = variance due to family differences.
σ^2_l = variance due to location differences.
n = number of replications.
k = harmonic mean of the number of trees per plot.

Table 3. — Estimates of repeatability of family means (intra-class correlation) for four sets of families and for four different methods of progeny testing.

Progeny test method[1])	Sets of families				
	Weyerhaeuser O.P	Hoerner-Waldorf Crosses	Weyerhaeuser × Louisiana	Hoerner-Waldorf × Westvaco	Mean
		Height growth			
Test 1	.91	.56	.94	.92	.83
Test 2	.72 ± .13	.36 ± .21	.75 ± 12	.84 ± .04	.67
Test 3	.56	.22	.60	.72	.52
Test 4	.21	.22	.29	.07	.20
		Rust score			
Test 1	.37	.89	—	.87	.71
Test 2	.60 ± .19	.61 ± .19	—	.84 ± .04	.68
Test 3	.43	.44	—	.73	.53
Test 4	.10	.24	—	.19	.18

[1]) Test 1 → single location progeny test.
Test 2 → progeny test of same families in the same three locations.
Test 3 → progeny test of different families in three different locations but with a common genetic check.
[1]) Test 4 → progeny test of different families in three different locations without the use of a genetic check.

Result and Discussion

The component of variance due to genotype × environment interaction for height growth was as large as or even larger than that due to genotypic differences for all the four sets of families after appropriate adjustments *(table 2)*. Although the standard errors of these estimates are very large, the family and familiy × environment sources of variation were significant at the .05 probability level for the four sets of families. For rust score, the genotype × environment component was found to be much less than that for family differences.

Repeatabilities of family means for four different methods of progeny testing are given in *table 3*. For height growth there is a big difference in repeatability between testing all families in the same environments and testing different families in different environments without the use of a common genetic check. This relationship applies to all sets of families except for Hoerner-Waldorf within orchard crosses. The use of a genetic check markedly increases the repeatabilities relative to testing different genotypes in different locations without a genetic check.

The efficiencies of progeny testing in three locations relative to testing in one location for three methods of progeny testing are given in *table 4*. Extrapolation for efficiencies of progeny testing in five locations are also included in the table. It can be seen from *tables 3* and *4* that selection efficiency increases as the repeatability of family means decreases. This is expected because, if there is a high correlation between family means in different environments (high r_f), then only one location can provide most of the information. The results of extrapolation to five locations indicate that selection efficiency does not increase very much by increasing the number of locations above three.

The differences between the observed family means and the most probable family worth based on a 5% selection differential is shown graphically in *figure 1*. The 5% selection differential is used here because it is in the range of

Table 4. — Efficiency of selection for height growth based on progeny testing at 3 and 5 locations relative to that base on progeny testing single location.

Table[1])	Progeny test at:	
	3 locations	5 locations
Test 1	100	100
Test 2	121.72	121.89
Test 3	137.67	138.07
Test 4	216.83	219.47

[1]) Same as in table 3.

what is usually achieved in the wild stands selection of loblolly pine (PORTERFIELD 1974). For an eight-year-old loblolly progeny test, this differential can roughly be achieved by selecting the top 1/3 of the families (MATZIRIS 1974). Using an average repeatability of .67, the observed family means and the most probable family worths fall very close to one another. This means that the effects of genotype × environment interaction are not large enough to lower selection efficiency when progeny testing is carried out in three locations.

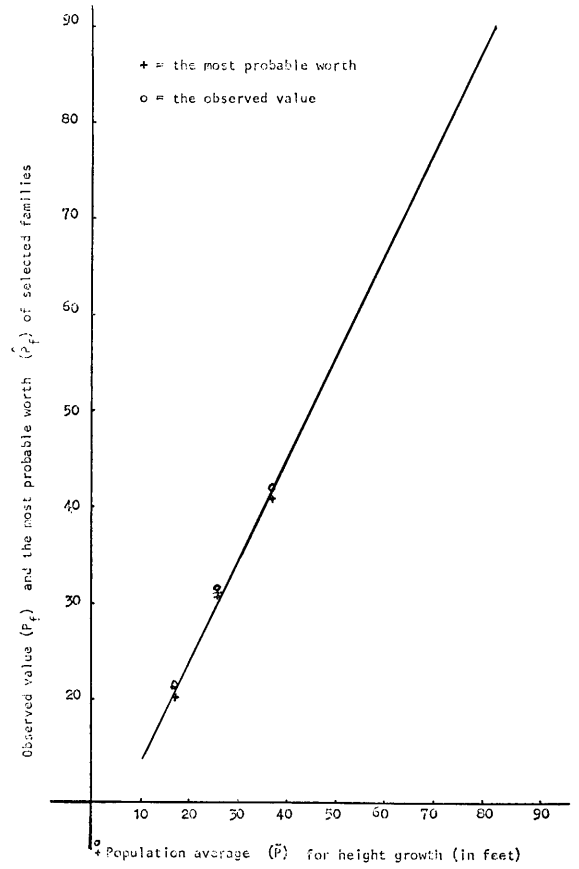

Figure 1. — Graphical representation of the observed and 'most probable worth' of families based on 5% selection differential.

Acknowledgements

The author expresses gratitude to Professor B. J. ZOBEL for his guidance and encouragement during the course of this study. Special thanks go to members of the North Carolina State University — Industry Tree Improvement Co-operative Program and to fellow graduate students who provided valuable assistance at various stages of the study.

Summary

Correlation studies indicated that families of loblolly pine, when progeny tested in the same three sites, show repeatabilities of family means across environments in the neighbourhood of .72. However, when the same families are tested at different locations without a genetic check, repeatability of family means was found to be .21. Finally, use of genetic check to link families tested at three different sites raised repeatability of family means to about .56. Results also indicated that progeny testing at three locations can result in 20% superiority in selection efficiency compared to tests at single locations.

Key words: *Pinus taeda*, progeny testing, genotype × environment interaction, repeatability of family means, "observed" and "most probable" worth of genotypes.

Zusammenfassung

Der realisierte genetische Gewinn hängt neben Selektionsintensität und Größe der additiven genetischen Varianz wesentlich auch von der Effizienz ab, überlegene Genotypen zu selektionieren, wobei letztere stark von der Größe der Genotyp-Umwelt Interaktion bestimmt wird. Die Untersuchung verschiedener Methoden der Nachkommenschaftsprüfung von 8jährigen Familien von *Pinus taeda* zeigen für das Merkmal Baumhöhe, daß die Selektionseffizienz bei vorhandener G × E bei Prüfung der Nachkommen auf 3 Standorten um 20% größer ist, als wenn nur an einem Ort geprüft würde.

Literature Cited

BURDON, R. D. and SHELBOURNE, C. J. A.: Breeding populations for recurrent selection: conflicts and possible solutions. N.Z. Jour. For. Sci. 1 (2): 174—193 (1971). — LIBBY, W. J.: Seedling versus vegetative orchards. Paper to FAO-N.C. State University Forest Tree Improvement Training Centre p. 306—316 (1969). — LUSH, J. L.: Animal breeding plans. Iowa State College Press 443 pp. (1943). — MATZIRIS, D. I.: Predicted versus realized gain in loblolly pine *(Pinus taeda)* improvement. Unpublished Ph. D. dissertation, Department of Forestry, North Carolina State University at Raleigh (1974). — NAMKOONG, G., SNYDER, E. B. and STONECYPHER, R. W.: Heritability and gain concepts for evaluating breeding system such as seedling orchards. Silvae Genet. 15: 76—84 (1966). — NORDSKOG, A. W. and KEMPTHORNE, O.: Importance of genotype × environment interactions in random sample poultry tests. Biometrical Genetics. Pergamon Press. p. 159—168 (1960). — POTERFIELD, L. R.: Predicted and potential gains from tree improvement programs. A goal programming analysis. Ph. D. dissertation, Department of Forestry, Yale University, New Haven (1973). — SHELBOURNE, C. J. A.: Tree Breeding methods. Forest Res. Inst., New Zealand Forest Service, Tech. Paper No. 55 (1969). — VAN BUIJTENEN, J. P.: Efficiency of mating designs for second generation selection. Proc. Meeting of the working Party on Progeny Testing, IUFRO, Macon, Ga. p. 103—126 (1972). — ZOBEL, B. J., WEIR, R. J. and JETT, J. B.: Breeding methods to produce progeny for advanced-generation selection and to evaluate trees. Can. Jour. For. Research 2 (3): 339—345 (1972).

Addenda: BECKER, W. A.: Manual of quantitative genetics. Washington State University Press (1975). — SCHULTZ, W. M. and BERNARD, R. L.: Genotype × environment interactions in the regional testing of soybean strains. Crop Sci. 7: 125—130 (1967).

Addendum to Appendix 1: For the case when the same sets of families are tested in the same locations (i.e. Test 2) the standard errors of the repeatability of family means (S.E.r_f) are calculated as follows (BECKER, 1975).

$$S.E.r_f = \sqrt{\frac{2(1-r_f)^2\,[1+(l-1)r_f]^2}{l(l-1)\,(N-1)}}$$

Where l = the number of locations i.e. 3 in this case
N = the number of families i.e. 10 in this case
r_f = the repeatability of family means.