Estimates of Genetic Parameters and Prediction of Breeding Values for Growth in *Eucalyptus globulus* Combining Clonal and Full-Sib Progeny Information

By J. A. Araújo[1]), R. Sousa[1]), L. Lemos[1]) and N. M. G. Borralho[2])[3])

(Received 1st July 1996)

Abstract

Analysis of height growth from clonal and progeny trials from Soporcel's *Eucalyptus globulus* breeding program gave regional heritabilities of 0.28 ± 0.03 and 0.19 ± 0.04 for the north-west and south of Portugal, respectively. Results also showed that cloning the parents used in the crossing program, even when based on a few ramets per clone, improved considerably the accuracy of genetic parameter estimates and efficiency of backward selections. The benefits of cloning the parents for selection efficiency seem justifiable even when large scale vegetative propagation is not feasible.

Key words: BLUP, REML, individual model, heritability, genetic correlations.

FDC: 165.3; 165.4; 165.441; 232.11; 176.1 *Eucalyptus globulus.*

Introduction

The first breeding populations combine 2 generations: the base parents, either selected in native stands or in unimproved plantations, and their progeny, established in well replicated progeny trials. Unfortunately, in most cases the only information available for selection is the progeny's performance. Measurements on the base parents are not available, or in the case of diameter or height in native stands of unknown age, they are too unreliable to be incorporated in the analysis. This is unfortunate because parent-offspring relationships result in estimates with lower error variances than methods based on intraclass correlation amongst collateral relatives (*e.g.* Hill, 1990). They are also more robust against assumptions about the intraclass correlation in open-pollinated families (Borralho and Kanowski, 1994), not a trivial problem for many species, and they improve the accuracy of breeding value predictions (*e.g.* Nicholas, 1987).

A way to record the performance of the parents is to propagate a few vegetative propagules, either by tissue culture or cuttings, from rejuvenated tissue. Such approach depends to some extent on the magnitude of propagation effects associated with cuttings, althought simulation seems to suggest that biases in genetic parameters are likely to remain small (Borralho and Kanowski, 1995). Propagating mature trees is achievable for most tropical and temperate hardwood species, using coppice or epicormic shoots. In this paper we describe such a case, developed for *Eucalyptus globulus* by Soporcel, a major forest and pulp company in Portugal. Between 1985 and 1989, cuttings were taken from the coppice of plus-trees selected in genetically unimproved stands in Portugal and the rooted cuttings were then established in clonal trials. At the same time, some of these plus-trees were used in a crossing program and their offspring established in progeny trials.

The study has 2 main objectives: (1) to apply a REML analysis to both the clonal and progeny data to obtain accurate estimates of heritability and across region genetic correlations for early growth of *E. globulus* in Portugal; and, (2) to investigate the importance of parental information based on clonal trials in the accuracy of breeding values estimates using BLUP techniques.

Material and Methods

Genetic Material

The clones used were selected in genetically unimproved stands of *E. globulus* growing in central and north Portugal. The population includes 152 base parents, 150 of which were successfully propagated by stem cuttings and later established in clonal trials. The families in the progeny trials were derived from controlled and open pollinated crosses between 10 base parents, 8 of which were also represented in clonal trials. The crossing scheme is a very sparse diallel. Due to the limited number of crosses successfully completed in that year, the trial was supplemented with open pollinated seed from some of the parents (see *Table 1*).

Table 1. – Crossing scheme used in trials 12 and 13.

Females	Males					
	1035	1049	1113	1197	1301	OP
1035		X	X	X	X	X
1041	X	X		X	X	X
1043		X		X	X	X
1052		X				
1053		X				X
1197					X	X
1300	X	X		X		X
1301		X				

Trials and Measurements

The clonal and progeny trials were established by Soporcel, and are part of the company's breeding program. Trials are located on a range of conditions in Portugal, but were grouped in 2 main geographic regions: the Northwest, generally of high productivity, and the South, with drier conditions, hence less productive (*Table 2*). Height was considered to represent the same trait within a region, and a different trait between regions. All trials were established as randomised complete blocks, with each family represented once (or occasionally twice) in each replicate by a 5-tree row plots, and clones represented by a single ramet in each replicate. Height was measured on all trees between 2 to 4 years.

---

[1]) Soporcel, Centro de Investigação Florestal, Apartado 15, 2065 Alcoentre, Portugal

[2]) Cooperative Research Centre for Temperate Hardwood Forestry, Dept. of Plant Science, University Tasmania, GPO Box 252C, Hobart, Tasmania 7001, Australia

[3]) To whom correspondance should be addressed

Table 2. – Details of the clonal and progeny trial used in this study. Trials were grouped in 2 regions, the North West (NW) and the South (S) in Portugal. Also given are the total number of clones in each clonal trial and how many of them were also included as parents in progeny trial 12 and 13.

Trial No.	Type	Region	No. Blocks	No. Clones/ parents	No. With progeny
2	Clonal	NW	10	27	6
4	Clonal	NW	10	24	3
5	Clonal	NW	10	65	1
7	Clonal	NW	10	51	1
8	Clonal	NW	10	50	2
9	Clonal	NW	16	57	2
10	Clonal	S	16	39	2
11	Clonal	NW	15	53	2
12	Progeny	NW	5	10	10
13	Progeny	S	5	10	10
14	Clonal	S	10	81	1

Data Transformation

Because trials were established and measured at different years and on sites of different productivity, means and variances differed. In order to combine measurements from different trials in a single analysis, data from each trial was initially standardised to a variance of one.

Statistical Analysis

The analysis used an individual tree model, represented in matrix notation by:

$$y = \mu + Xb + Za + e \qquad (1)$$

where y is a vector of observations for height, μ is the overall mean, b is the vector of effects due to blocks (assumed fixed), a is a vector of additive genetic effects of both parents and progeny and e is the vector of residual variances. X and Z are known design matrices relating observations to effects. The analysis assumes all random effects to be normally distributed with mean zero and (co)variances as follow:

$$Var\begin{bmatrix} y \\ a \\ e \end{bmatrix} = \begin{bmatrix} R+ZGZ' & ZG' & R \\ GZ' & G & 0 \\ R & 0 & R \end{bmatrix} \qquad (2)$$

where $G = Diag\ \{A\ x\ T_A\}$ is the genetic variance-covariance matrix, A is the numerator relationship matrix between trees, T_A the matrix of additive genetic covariance between height in the 2 regions (see Meyer, 1989), and R is the matrix of variance-covariance of the residuals. Residual covariances between the 2 heights (in NW and S region) was assumed to be zero. Height measurements taken on different ramets of the same clone were treated as repeated measurements of the same trait with zero residual covariance. Estimates of covariance components using the model described in equation 1, were obtained by restricted maximum likelihood (REML), using K. Meyer's DFREML programs (Meyer, 1991). Sampling errors for heritabilities were approximated by fitting a cubic function to the likelihood surface (Meyer, 1991). Sampling errors for the genetic correlations were not available from DFREML and were approximated using Robertson's (1959) equation:

$$\sigma_{r_G} = (1 - r_G^2)\sqrt{\frac{\sigma_{h_x^2}\sigma_{h_y^2}}{h_x^2 h_y^2}} \qquad (3)$$

Best Linear Unbiased Predictions of expected breeding values (EBV) and corresponding prediction error variances (PEV) were based on the same model using the variance ratios estimated from the data. The accuracy of breeding value estimation ($r_{a\hat{a}}$), or the correlation between true breeding values (TBV) and expected breeding values (EBV), were following Henderson (1984) as:

$$r_{a\hat{a}} = \sqrt{1 - \frac{PEV}{\sigma_a^2}} \qquad (4)$$

where $PEV = G - C_{22}$, and G is the genetic variance-covariance matrix and C_{22} is a g-inverse of the mixed model equations referring to the breeding values.

Results and Discussion

Genetic Parameters

Genetic parameters for height in the 2 regions are listed in *table 3*. Heritabilities using the combined clonal and progeny data were $h^2 = 0.28$ across the more productive sites in the North West of Portugal (NW), and $h^2 = 0.19$ in the drier, less productive sites in the South (S). These estimates are on the lower end of the range reported by Borralho *et al.* (1992a) based on open-pollinated material from plantations or seed orchards in Portugal. Genetic correlation between the 2 regions was nevertheless high ($r_G = 0.83$) in agreement with previous G x E studies for the same species, by Kube *et al.* (1993) and Borralho *et al.* (1995) in Australia, and Borralho *et al.* (1992b) in Portugal.

Table 3. – Genetic parameters based on progeny data (full-sib progeny trials) only and combining base parent (clonal trials) and progeny data (full-sib progeny trials) for 2 to 5 year old height in the northwest (NW) and south (S) of Portugal.

	Progeny	Parents + Progeny
Heritability		
NW	0.52 ± 0.37	0.28 ± 0.03
S	0.13 ± 0.11	0.19 ± 0.04
Genetic Correlation	0.33 ± 0.52 †	0.84 ± 0.03 †

† Calculated from Robertson (1959)

It is apparent from *table 3* that the accuracy of the genetic parameters changed markedly when the data from clonal trials was included, with the error variances of the estimates reduced 3 and 10 fold, for the S and NW regions, respectively. The much larger heritability calculated using progeny data only, for the NW sites ($h^2 = 0.52 \pm 0.37$) was in fact not significantly different to the estimate obtained for the same region using the combined dataset ($h^2 = 0.28$).

The relatively high error variances of the estimates based on the progeny data only is due to the reduced number of families and parents represented and low progeny sizes (of between 5 and 25 progeny per family at each site). Published standard errors for heritability estimates of *E. globulus*, based on open-pollinated material, ranged between 0.03 and 0.08 (Borralho *et al.*, 1992a, 1995), but they were obtained from a much larger array of families. The standard errors of estimates based on the combined analysis (0.03 and 0.04) were of the same magnitude as those studies, despite being based on a much smaller set of parents and progeny sizes, and demonstrate well the impact on the accuracy of the estimates of including parental information in the analysis.

A key distinction between the two datasets is the fact that heritability estimates using clonal information covers a wide range of sites, whereas the progeny data is based on a single-site heritability, with one trial in the NW and another in the South. Thus heritability estimates from the combined analysis are likely to be lower, as a result of genotype by environment interactions. This was apparent in the NW, where there was a marked drop from the single-trial heritability of 0.52 (using progeny data only), to 0.28 (using combined across-site data), although the 2 values were not significantly different. In the South, the heritability estimates were very similar and did not differ significantly between datasets. The fact that there were no statistically significant changes in the heritabilities from the progeny to the combined multiple trial situation, suggest that genotype-site interactions should be unimportant. The genetic correlation between height in the 2 regions was accordingly very high ($r_G = 0.84 \pm 0.03$). Interestingly, the genetic correlation between the 2 progeny trials, thus using progeny data only, was much lower, ($r_G = 0.33 \pm 0.52$), although the accuracy of the latter estimate was poor.

Breeding Values

The improvement in the accuracy of genetic parameter estimates, discussed in the previous section, has important advantages, as the use of wrong genetic parameters is known to result in biased estimates of breeding values (*e.g.* see White and Hodge, 1990 for a discussion). Another advantage is the reduction in the prediction error variance (denoted PEV) of the breeding value estimates themselves. The PEV is a useful statistic because it is related with the accuracy of selection, that is the correlation between the predicted breeding value and the true but unknown breeding value [denoted here $r_{a\hat{a}}$, or in White and Hodge (1990) terminology, the $corr(g, \hat{g})$], and although PEV's are independent of the expected breeding values, they provide a useful measure of the risk associated with the selections (Meuwissen and Woolliams, 1994). The accuracy of both parent and progeny's breeding values in Soporcel's breeding population, using the progeny data with and without the clonal data, are summarized in *table 4*.

Table 4. – Average accuracy of EBV, given as correlation between true and estimated breeding values ($r_{a\hat{a}}$), for progeny and base trees when different information was used.

	$r_{a\hat{a}}$	
	Using all data	**Using only progeny data**
Progeny		
from trial 12	0.75	0.73
from trial 13	0.71	0.69
Parents		
cloned with progeny	0.98	0.78
cloned without progeny	0.92	–
not cloned with progeny	0.93	0.91

As expected, parent trees in clonal trials and with offspring in progeny trials had the highest accuracy of breeding value estimates, with $r_{a\hat{a}} = 0.98$. Hence backward selection of these parent trees would be close to 100% accurate. Trees only presented in clonal trials or alternatively parent trees which have only been progeny tested but not clonal tested, also had a high accuracy, with $r_{a\hat{a}} = 0.92$ and 0.93, respectively. Compared with backward selections, accuracy of selecting trees in the progeny trials was lower, with $r_{a\hat{a}} = 0.75$ and 0.71 for trials 12 and 13, respectively. The impact of adding parental information on the selection of these progeny was moderate, with accuracies improving from 73% to 75% in trial 12 and from 69% to 71% in trial 13. Approximate selection accuracy for parents and progeny with and without being clonally tested are given in *table 4*.

In many programs it is common to keep a separate breeding and deployment population, the latter seen as an elite group from the trees previously selected for breeding. As a consequence, results from the clonal trials (when available) are poorly linked and generally do not contribute information to the selections carried out in the breeding population. Results from this paper suggest there are important advantages for selection efficiency in cloning the parents and combining them in the same analysis. The establishment of clonal trials of (at least some) parents results in considerably improved estimates of genetic parameters and an increase in the accuracy of breeding values estimation in the breeding population. Although propagation of parents is generally difficult, small scale cloning programs, such as the one described here for *E. globulus* for a species known to be difficult to clone (Wilson, 1993), can be executed.

Cloning the progeny from the breeding population has been shown to be an effective way to improve accuracy of selection (Shaw and Hood, 1985; Shelbourne, 1991; Mullin and Park, 1992). While this may be a formidable task when applied to the whole progeny population (of several thousand individuals), the cloning of parent trees (usually only a few hundred and based on a few ramets per clone) could be more attainable. The results found here for Soporcel's *E. globulus* population suggests that there was a considerable improvement on the estimates of genetic parameters (thus reducing bias in the subsequent estimation steps), allowed for a better representation of genotypes across the range of sites, and improved considerably the selection accuracy of parental breeding value predictions, the latter being particularly relevant if backward selection is a favoured option for deployment purposes.

As programs move toward the second and third generations of breeding, the relative importance of cloning the parents is likely to be reduced, as more information becomes available (the phenotype of the parents and the parents' sibs, for example) but their importance will not disappear altogether. Cloning, which can be seen as a way to have repeated measurements, will remain a valuable strategy to improve selection accuracy in poorly inherited (low repeatability) traits, such as growth. Furthermore, cloning parents and testing them with their progeny (preferably in the same trials) will provide a useful links between different generations of breeding. This will be useful in separating the effects of genetic and silviculture on increasing productivity.

Although parental cloning may seem an interesting alternative to Shelbourne's (1991) breeding population cloning, there is one important distinction. Cloning the parents will increase mostly the accuracy of the parents EBV (which is important for backward selection), but the improvement in accuracy in the progeny's EBV comes only from an increase in pedigree information, thus, benefit from cloning the parents is a result of an increase in accuracy of between family selection. Cloning the progeny, on the other hand, can improve the accuracy of within family selection (Shelbourne, 1991), but has a relatively small impact on the accuracy of genetic parameters or parental selection. The 2 strategies seem therefore to be complementary in their benefits to selection programs.

Acknowledgments

We thank Soporcel for their support and for making the data available, and BRAD POTTS and TONY SHELBOURNE for their comments on the manuscript.

Literature

BORRALHO, N. M. G., ALMEIDA, I. A. and COTTERILL, P. P: Genetic control of growth of young *Eucalyptus globulus* clones in Portugal. Silvae Genet. **41**: 100–104 (1992b). — BORRALHO, N. M. G., JARVIS, S. F. and POTTS, B. M.: Multivariate restricted maximum likelihood estimation of genetic parameters for diameter at 3 to 5 years in the *Eucalyptus globulus* base population in Australia. In: “Eucalypt Plantations: Improving Fibre Yield and Quality”. (Eds. B. M. POTTS, N. M. G. BORRALHO, J. B. REID, R. N. CROMER, W. N. TIBBITS and C. A. RAYMOND). pp. 230–232. Proc. CRC-IUFRO Conf., Hobart, 19 to 24 Feb. (1995). — BORRALHO, N. M. G. and KANOWSKI, P. J.: Correspondence of performance between genetically related clones and seedlings. Can. J. For. Res. **25**: 500–506 (1994). — BORRALHO, N. M. G., KANOWSKI, P. J. and COTTERILL, P. P.: Genetic control of growth of *Eucalyptus globulus* in Portugal. I. Genetic and phenotypic parameters. Silvae Genet. **41**: 39–45 (1992a). — HENDERSON, C. R.: Applications of Linear Models in Animal Breeding. Univ. of Guelph. 423 pp. (1984). — HILL, W. G.: Considerations in the design of animal breeding experiments. In: Advances in Statistical Methods for Genetic Improvement of Livestock, (D. GIANOLA and K. HAMMOND, eds.). pp. 59–76. Springer-Verlag, Berlin (1990). — KUBE, P. D., MATHESON, A. C., RAYMOND, C. A. and BOOMSMA, D. B.: Provenance variation in *Eucalyptus globulus*. In: “Eucalypt Plantations: Improving Fibre Yield and Quality”. (Eds. B. M. POTTS, N. M. G. BORRALHO, J. B. REID, R. N. CROMER, W. N. TIBBITS and C. A. RAYMOND). pp. 245–246. Proc. CRC-IUFRO Conf., Hobart, 19 to 24 Feb. (1995). — MEUWISSEN, T. H. E. and WOOLLIAMS, J. A.: Response versus risk in breeding schemes. Proc. 5th World Congress on Genetics Applied to Livestock Production, Univ. Guelph **18**: 236–243 (1994). — MEYER, K.: Restricted maximum likelihood to estimate variance components for animal models with several random effects using a derivative-free algorithm. Genet., Sel., Evol. **21**: 317–340 (1989). — MEYER, K.: DFREML Version 2.0 – Programs to Estimate Variance Components by Restricted Maximum Likelihood Using a Derivative-Free Algorithm. User Notes. Animal Genetics and Breeding Unit, University of New England, Armidale, N. S. W. Mimeo (1991). — MULLIN, T. J. and PARK, Y. S.: Estimating genetic gains from alternative breeding strategies for clonal forestry. Can. J. For. Res. **22**: 14–23 (1994). — NICHOLAS, F. W.: Veterinary Genetics. Oxford Science Publications (1987). — ROBERTSON, A.: The sampling variance of genetic correlation coefficient. Biometrics **15**: 234–249 (1959). — SHAW, D. V. and HOOD, J. V.: Maximising gain per effort by using clonal replicates in genetic tests. Theor. Appl. Genet. **71**: 392–399 (1985). — SHELBOURNE, C. J. A.: Genetic gains from different kinds of breeding population and seed or plant production population. In: “Intensive Forestry: The Role of Eucalypts”, held in Durban, South Africa, in September 1991. South African Forestry Journal **160**: 49–65 (1991). — WHITE, T. L. and HODGE, G. R.: Predicting Breeding Values with Applications in Forest Tree Improvement. Kluwer Academic Publishers, Dordrecht (1990). — WILSON, P. J.: Propagation characteristics of *Eucalyptus globulus* LABILL. ssp *globulus* stem cuttings in relation to their original position in the parent shoot. J. Hort. Sci. **68**: 715–724 (1993).

Stand Density Influences Outcrossing Rate and Growth of Open-Pollinated Families of *Eucalyptus globulus*

By C. M. HARDNER, R. E. VAILLANCOURT and B. M. POTTS

Department of Plant Science and the Cooperative Research Centre for Temperate Hardwood Forestry,
University of Tasmania, GPO Box 252-255, Hobart, Tasmania, Australia

(Received 12th July 1996)

Summary

Using isozymes markers, we test BORRALHO and POTTS' (1996) hypothesis that poorer growth of open-pollinated (OP) families of *Eucalyptus globulus* grown from seed collected from less dense stands is a consequence of lower outcrossing rates. A significant relationship between stand density class and outcrossing rate was found, with outcrossing rates near one for families from closed forests. In support of BORRALHO and POTTS' (1996) hypothesis, lower parental breeding values were generally associated with lower outcrossing rates. This is no doubt due to a higher proportion of less vigorous selfs within an OP family depressing mean growth relative to families where inbred progeny were virtually absent. It is suggested that the precision of predicting breeding values with OP progenies may be improved by adjusting family performance for individual outcrossing rate.

Key words: Inbreeding depression, breeding values, selfing rate, mating system.

FDC: 232.11; 165.41; 176.1 *Eucalyptus globulus.*

Introduction

In a recent paper, BORRALHO and POTTS (1996) suggested that the growth of native forest open-pollinated (OP) families of *Eucalyptus globulus* was significantly affected by the density of the stand from which the seed was collected. They argued that differences in stand density would influence outcrossing rates and that accounting for this would improve heritability estimates and the prediction of breeding values. The density of the stand surrounding each parent was classified on a four point scale with 1 representing isolated parent trees (i.e. no observable neighbours) to 4 being for parent sampled from closed continuous forests. Classification was based on photographs taken at the time of the original seed collection because it was impractical to relocate all 596 parent trees used in the analysis. Families from parents classified as isolated were on average significantly smaller at 4 years than families of parents from closed forests. BORRALHO and POTTS (1996) suggested this trend could result from selection for faster growth in closed forests, but was more likely due to inbreeding depression caused by reduced outcrossing rates at lower stand densities.

Eucalypts are pollinated by a variety of generalist insects and animals (GRIFFIN, 1980). There is ample opportunity for self pollination and at least some seed is set following controlled self pollination (GRIFFIN *et al.*, 1987; HARDNER and POTTS. 1995). Estimates of outcrossing rates for individuals trees vary