Genetic Mapping of Quantitative Trait Loci Controlling Growth and Wood Quality Traits in *Eucalyptus grandis* Using a Maternal Half-Sib Family and RAPD Markers

Dario Grattapaglia,*,† Fernando L. G. Bertolucci,‡ Ricardo Penchel‡ and Ronald R. Sederoff†

**Plant Genetics Laboratory, Centro Nacional de Pesquisa de Recursos Genéticos e Biotecnologia, Empresa Brasileira de Pesquisa Agropecuária, Brasília D.F. 70879-970, Brazil, †Forest Biotechnology Group, North Carolina State University, Raleigh, North Carolina 27695 and ‡Gerência de Tecnologia, Aracruz Celulose S.A., Aracruz ES 291907-000, Brazil*

Manuscript received August 24, 1995
Accepted for publication August 12, 1996

ABSTRACT

Quantitative trait loci (QTL) mapping of forest productivity traits was performed using an open pollinated half-sib family of *Eucalyptus grandis*. For volume growth, a sequential QTL mapping approach was applied using bulk segregant analysis (BSA), selective genotyping (SG) and cosegregation analysis (CSA). Despite the low heritability of this trait and the heterogeneous genetic background employed for mapping, BSA detected one putative QTL and SG two out of the three later found by CSA. The three putative QTL for volume growth were found to control 13.7% of the phenotypic variation, corresponding to an estimated 43.7% of the genetic variation. For wood specific gravity five QTL were identified controlling 24.7% of the phenotypic variation corresponding to 49% of the genetic variation. Overlapping QTL for CBH, WSG and percentage dry weight of bark were observed. A significant case of digenic epistasis was found, involving unlinked QTL for volume. Our results demonstrate the applicability of the within half-sib design for QTL mapping in forest trees and indicate the existence of major genes involved in the expression of economically important traits related to forest productivity in *Eucalyptus grandis*. These findings have important implications for marker-assisted tree breeding.

THE major obstacle in forest tree improvement is the time necessary to complete a breeding generation. Current practice has relied almost exclusively on the analysis of phenotypes at rotation age. For most traits of commercial value, early selection of individual trees only becomes efficient half-way through rotation, even for fast-growing species of Eucalyptus. Tree breeding is made even more difficult by the changes that occur during the transition from juvenility to maturity. Wood properties change during growth and maturation. Wood specific gravity is only adequately expressed at the phenotypic level after the tree has produced several growth rings, and early height growth is often a poor predictor of volume at rotation age. Methods to improve the accuracy of early selection at the individual level would be of considerable value to increase the genetic gain per unit time. In that respect, a promising method is the direct identification of genotypes using a diagnostic system based on molecular markers cosegregating with the traits of interest.

Results from quantitative trait loci (QTL) mapping studies in recent years support the existence of a few major genes controlling large proportions of the total variation in a range of quantitatively inherited traits in crop plants (reviewed by STUBER 1992; DUDLEY 1993). Recent reports on the genetic architecture of quantitatively inherited traits in forest trees, *i.e.*, the number, chromosomal location and magnitude of effect of QTL controlling these traits, also indicate that major genes exist in trees such as *Pinus taeda* (GROOVER *et al.* 1995), *Populus trichocarpa x deltoides* (BRADSHAW and STETTLER 1995), and *Eucalyptus grandis* and *E. urophylla* (GRATTAPAGLIA *et al.* 1995).

QTL mapping in trees has been carried out using both three-generation outbred pedigrees as well as two-generation pedigrees involving crosses between highly heterozygous parents (pseudo-testcross strategy). No study to date has explored the use of a half-sib family, the most commonly available pedigree in tree breeding. Such an approach has been successfully used in domestic animals to detect QTL for growth and production traits (*e.g.*, GELDERMANN *et al.* 1985; BEEVER *et al.* 1990; CLAMP *et al.* 1992; GEORGES *et al.* 1995). If a parent tree is heterozygous at a marker locus, its half-sibs can be partitioned into two groups, those that received the chromosomal segment marked by one allele and those that received the homologous segment with the alternative marker allele. If the tree is also heterozygous at a closely linked QTL, the groups of half-sibs will also differ with respect to the quantitative trait.

In this study we present a QTL mapping analysis, using a maternal open pollinated half-sib family of *E. grandis* at rotation age (6.5 years). The objectives of the present study were as follows: (1) attempt to locate

Corresponding author: Dario Grattapaglia, Laboratório de Genetica de Plantas, CENARGEN-EMBRAPA, C.P. 02372, Brasília D.F. 70879-970, Brazil. E-mail: dario@cenargen.embrapa.br

genomic regions controlling growth and wood quality traits at rotation age in fast growing Eucalyptus, (2) test the adequacy of the within half-sib approach for QTL mapping in a forest tree, (3) apply a sequential approach of QTL mapping in a half-sib by analyzing first the extreme phenotypes by bulk segregant analysis (BSA), selective genotyping (SG) and then the whole segregating progeny by cosegregation analysis (CSA).

MATERIALS AND METHODS

Plant material: The experimental material consisted of an open pollinated maternal half-sib family of an elite genotype of *E. grandis* (Coffs Harbor, Australia). Seeds were obtained from a managed seed orchard at Aracruz Celulose S.A., Brazil where the *E. grandis* genotype is used as the only female clone, and 25 *E. urophylla* clones are used as pollinators. The ratio of pollinators to maternal clonal ramets was 3:1 and the maternal crowns were surrounded by pollinator trees. Pollination in Eucalyptus is predominantly entomophilous. Cross pollination was enhanced by the establishment of bee cages in the orchard, and a 800-m strip of native forest separated the orchard from other Eucalyptus stands. The maternal *E. grandis* clone was characterized as highly self-incompatible by controlled self-pollination. Therefore the seeds collected in the orchard are derived from outcrossing events with very high probability. Furthermore, seedlings derived from self-pollination (estimated at <3%) are typically identified and rogued in the nursery stage (IKEMORI and CAMPINHOS 1983). In spite of this we still found 3.6% putative selfed individuals using a combination of four maternal codominant markers. These individuals were eliminated from the experiment and substituted by true half-sibs.

Hundreds of hectares of commercial production plantations have been established with this seed lot at Aracruz Celulose. We selected and delimited a square area of ~1 hectare in an homogeneous forest stand planted in 1986 at a spacing of 3 m between rows × 2 m between plants. At time of trait evaluation, the stand was at harvest age (6.5 years). Even terrain, uniform soil type and minimal number of missing trees were the main factors considered in the selection of the experimental area.

Traits measured: The following traits of commercial value were measured: (1) circumference at breast height (1.3 m) (CBH), estimated individual tree heritability of 0.31 (REZENDE and BERTOLUCCI 1993); (2) wood specific gravity (WSG), individual tree heritability estimated between 0.5 and 0.7 (F. L. G. BERTOLUCCI, unpublished results); (3) percentage dry weight of bark (%BARK) (no estimate of heritability available) and (4) cellulose pulp yield (%PULP), broad sense heritability of 0.2 (DEMUNER and BERTOLUCCI 1992). CBH was measured on a total of 1085 standing trees that were then felled for leaf sample collection for DNA analysis. Height growth was measured on 50 felled trees to estimate height × diameter correlation. A 4–5-cm-thick wood disk at breast height was sampled from 400 individuals for measurement of wood quality traits. WSG determinations were made for the entire wood disk by a gravitometric method. The wet weight of the water saturated wood disk was measured. Then, disk volume was determined as the ratio between the weight of the displaced water and the specific gravity of water at 26° (996.6 kg/m^3). Dry weight of the disk was obtained following oven drying at 105° for a minimum of 48 hr. WSG was calculated as the ratio between the dry weight (kg) and the volume of the disk (m^3). Throughout this procedure measurements were made separately for the bark and the solid wood so that WSG with and without bark were obtained. These were found to be highly correlated so that only WSG without bark was considered in the study. Percentage bark was calculated as the ratio between the dry weight of bark and the total dry weight of the disk.

For a sample of 164 individuals, a micropulping technique developed by Aracruz Celulose S.A. was used to estimate percentage cellulose pulp yield. Briefly, a sample consisting of 15 g of uniformly prepared wood chips (1.0–1.2 × 5–11 × 30–50 mm) were taken from the center part of the dried wood disks. Wood chip samples were subject to alkaline cooking concentrations corresponding to 160–180 kg of active alkali (expressed as NaOH) per ton of dry wood chip, *i.e.*, 16–18% active alkali. Cooking was performed at 170° during 2 hr into rotatory mini-digestors at a 5:1 ratio of white licor to wood chip. Following digestion, the dark licor was rinsed off, and the wood fibers in suspension were disintegrated in a blender until they became pulp. The pulp was then washed, filtered under vacuum and dried at 105° for 48 hr. Pulp yield was calculated as the ratio between the dry weight of pulp and the initial weight of wood chips.

RAPD marker genotyping: Leaf samples from a total of 1085 individuals were collected. Random amplified polymorphic DNA (RAPD) assay conditions, marker identification and scoring are described elsewhere (GRATTAPAGLIA and SEDEROFF 1994). A linkage map of RAPD markers had been constructed for the common maternal parent using the pseudo-testcross strategy so that the RAPD markers were previously selected to be present and heterozygous in the *E. grandis* genotype and absent, homozygous null in *E. urophylla* pollinators. We screened the 25 pollinator trees for the RAPD markers used in this study and none of them showed the amplified fragments. Thus, the RAPD markers used uniquely identified the maternal gametic contribution to the half-sibs.

Bulk segregant analysis (BSA): The DNA pooling technique originally proposed by ARNHEIM *et al.* (1985) and later adopted by MICHELMORE *et al.* (1991) was employed as a first step to identify genomic regions controlling CBH. Individuals that were 1.7 σ_p (phenotypic standard deviation) above and below the mean for CBH were selected to compose the bulked DNA samples. Four bulks of 10 individuals each, from each tail of the distribution, were constructed by mixing equimolar amounts of genomic DNA. The 40 individuals in each tail were ranked phenotypically and sequentially assigned to the four bulks. A total of eight bulks, four "low" and four "high" together with the maternal genomic DNA and a pooled DNA sample of six paternal individuals as controls, were screened with a total of 130 framework markers covering an estimated 90% of the genome (GRATTAPAGLIA and SEDEROFF 1994).

Selective genotyping (SG): From the same population of 1085 individuals, 96 progeny with extreme phenotypes for CBH (48 in each tail, *i.e.*, 4.4% of the total mapping population) were chosen for selective genotyping. The two groups included the 80 individuals used to compose the bulks in the BSA step. These individuals were genotyped for a total of 52 evenly spaced markers in the *E. grandis* map covering an estimated 75% of the genome and amplified with 33 RAPD primers. Two analytical approaches were used to test for putative QTL: a *t*-test for difference of marker allele means, *i.e.*, selective genotyping (LANDER and BOTSTEIN 1989) and a *z*-test for difference in marker allele frequencies between the two extreme groups, the so-called trait-based analysis (LEBOWITZ *et al.* 1987). Significant differences ($P \leq 0.01$) were taken as an indication of marker-trait associations.

Cosegregation analysis (CSA): A total of 300 random individuals for which data was available for CBH, WSG and %BARK were genotyped for 77 RAPD marker loci. A subset of these individuals ($n = 164$) had also been evaluated for %PULP. A genetic map was calculated from the genotypic

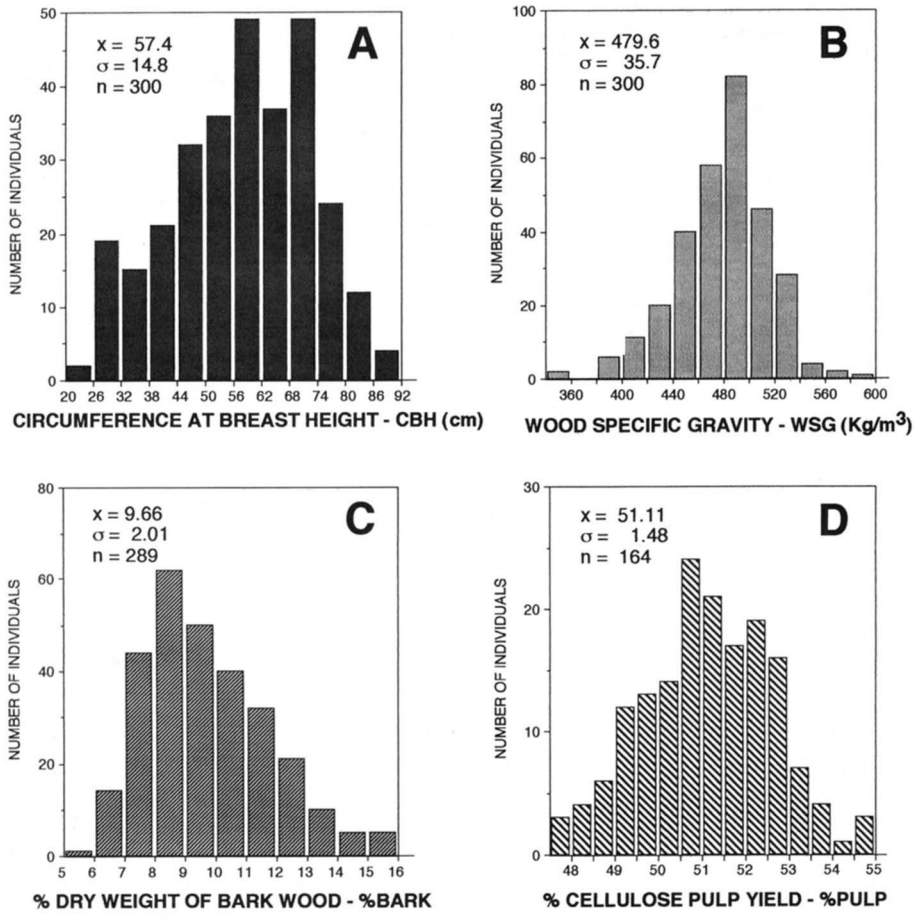

FIGURE 1.—Frequency distributions for growth and wood quality traits mapped in the half-sib family. Mean (x), standard deviation (σ) and sample size used in the cosegregation QTL analysis are indicated.

data *de novo* and checked for consistency with the previously reported maps (GRATTAPAGLIA and SEDEROFF 1994; GRATTAPAGLIA *et al.* 1995).

QTL mapping analysis was performed by linear regression using GLM (SAS 1988) and by interval mapping implemented by MAPMAKER-QTL (LANDER and BOTSTEIN 1989) under a backcross model. A nominal significance level for each individual test of $P \leq 0.01$ or a LOD score threshold of 1.9 were used to declare the presence of a putative QTL. With this threshold an overall false positive rate of ~5% was expected, given that a sparse map with an average of seven markers per chromosome at 30-cM intervals was used. This low threshold was adopted to ensure that any QTL with small but significant effect contributing to the traits were detected. For each LOD peak, the 1.0 LOD support intervals were determined. For all detected QTL, the percentage of variance explained as estimated by interval mapping was also reported. When linked QTL with no overlapping 1.0 LOD support intervals were detected, the locus with highest LOD score was fixed both for position and effect and the chromosome scanned again for the linked effect.

Multiple QTL models were evaluated by interval mapping with MAPMAKER-QTL and by multivariate linear regression using GLM (SAS 1988). To test for digenic epistasis, multivariate regression analysis was used including all the main effects and two-way interactions of unlinked markers. Significant effects were tested using Type III Sum of Squares to account for cases of missing cells (KNAPP *et al.* 1992).

RESULTS

Quantitative traits: The frequency distributions of phenotypes for three out of the four traits examined showed an approximately normally distributed continuous variation (Figure 1) as tested by the Shapiro-Wilk statistic calculated using PROC UNIVARIATE (SAS 1988). Only %BARK showed a significant departure from normality with a right-skewed distribution. Phenotypic correlations were estimated among traits within the family studied. Significant ($P \leq 0.01$) correlations were observed between CBH and WSG ($r = 0.28$), CBH and %BARK ($r = -0.67$) and WSG and %BARK ($r = -0.24$). As expected, phenotypic correlation between CBH and height growth was high ($r = 0.95$). Results presented for CBH should also be regarded as valid for height growth and therefore could be interpreted as representative of the relevant commercial trait, *i.e.*, final volume growth of the tree.

Linkage map construction: To perform cosegregation analysis of all traits on all RAPD markers, we reconstructed a genetic linkage map for the *E. grandis* maternal genotype using its open pollinated half-sib family. The linkage grouping obtained in this experiment was essentially the same as the one obtained previously from the full-sib pedigree with $n = 122$ (GRATTAPAGLIA *et al.* 1995), confirming the breakage of linkage group 11 into two pieces that merged with linkage groups 9 and 13, and the merger of groups 8 and 12. However, in this experiment we observed a merger of group 6 and group 9/11. Although the recombination fraction between the two groups was high

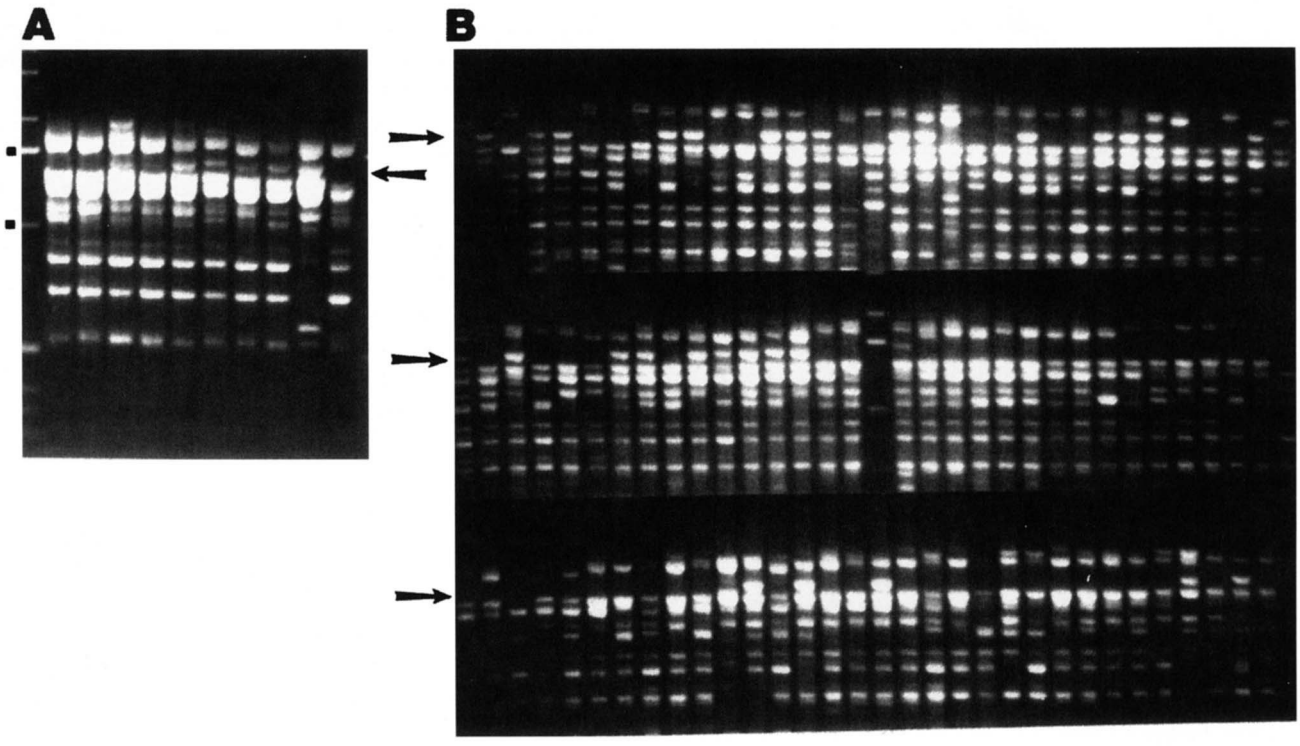

FIGURE 2.—QTL mapping for CBH using extreme phenotypes. Multiple bulk segregant analysis. (A) From left to right: 1 kb size standard, four "low" CBH bulked DNA samples, four "high" CBH bulked DNA samples; polymorphic marker U10_1500 between the two groups of bulks is indicated by arrow. (B) Selective genotyping: gel showing the same marker U10_1500 (indicated by arrows) genotyped for individual samples. From left to right, top to bottom, 1 kb size standard, 48 individuals from the high extreme of CBH distribution (>1.7 sp), 1 kb size standard, 48 individuals from the low extreme of CBH distribution (<1.7 sp), 1 kb size standard. Dots along the side indicate from top to bottom standard fragments of size 1600 and 1018 bp.

(0.38), the LOD score was >5.0. With this merger, we achieved 11 linkage groups, corresponding to the expected number based on the haploid number of chromosomes in *E. grandis* ($n = 11$).

The order of framework markers on the map is well conserved when compared to our previous maps of the *E. grandis* genotype. When we increased sample size from 62 to 122 meioses we observed 7% of order changes, most of these changes being simple order switches of adjacent markers (GRATTAPAGLIA *et al.* 1995). In this study we increased the sample size to 300 and observed 14% of order changes, the great majority being again of simple order switches of pairs of adjacent markers.

Sequential approach to QTL mapping: From a total of 130 markers screened in the BSA experiment, only one marker showed a distinctive polymorphic pattern between the high and low CBH bulks. The polymorphism was obtained with marker U10_1500, present in all the high CBH bulks and absent in all the low bulks (Figure 2A). For 10 other markers, inconsistent BSA polymorphisms were observed, when, within each group of bulks, the marker was present or absent in

TABLE 1

Results of selective genotyping and trait-based analysis for CBH

Marker	Linkage group	Marker allele means[a] Means ± SD +	Marker allele means[a] Means ± SD −	*P* value (*t* test)	Marker allele frequencies[b] High CBH tail	Marker allele frequencies[b] Low CBH tail	*P* value (*z* test)
U20_1000	2	61.3 ± 29.7	48.4 ± 30.1	0.03	0.60	0.39	0.044
R4_1300	5	50.7 ± 30.3	67.3 ± 29.9	0.01	0.42	0.72	0.004
U10_1500	7	75.1 ± 22.4	46.1 ± 29.4	0.000	0.50	0.10	0.000
A11_980	7	48.2 ± 29.9	71.9 ± 25.1	0.000	0.58	0.92	0.000
K9_1660	7	62.1 ± 29.3	47.9 ± 30.2	0.02	0.61	0.37	0.025

[a] Selective genotyping. Means for the alternative RAPD marker alleles, (+) presence of the RAPD band; (−) absence of band.
[b] Trait-based analysis. Frequencies of the (+) presence of the RAPD marker allele.

TABLE 2

Summary of significant marker-trait associations ($P \leq 0.01$) detected by linear regression using GLM (SAS)

Marker	Linkage group	CBH	WSG	% BARK
K10_835	2	0.002	0.01	
X1_1450	5	0.001	0.002	0.01
R4_1300			0.000	0.001
K1_1000			0.000	0.008
M6_961			0.000	0.007
Y16_1500			0.004	0.003
Y17_1500		0.000	0.002	
V7_1200			0.000	
Y15_760	6/9		0.007	
Y15_650				0.008
P10_530				0.004
N15_1079	7	0.007	0.004	
A11_980			0.000	
R20_1150	14			0.01

three out of four bulks. Replicate RAPD analyses with these markers indicated that these inconsistencies were due to sampling effect in the bulk composition and not to failure of the PCR assay.

Out of 52 evenly spaced markers screened in the selective genotyping experiment, four displayed a significant difference both in marker allele means and in marker allele frequency of the RAPD fragment between the two extreme groups. These markers were as follows: R4_1300 on linkage group 5, and markers K9_1660, U10_1500 and A11_980 on group 7. We interpreted these markers as being putatively associated to QTL controlling volume growth. Marker U20_1000 on linkage group 2 was only significant at the 0.05 level (Table 1).

Cosegregation analyses were performed and results are presented for untransformed phenotypic data. For %BARK an analysis was carried out on normalized (log transformed) data, however results were virtually the same as for the raw untransformed data. Using linear regression a total of 23 out of 308 (four traits × 77 markers) F tests were found to be significant ($P \leq 0.01$) (Table 2). Several of the significant markers associated with trait expression were linked on the same linkage group. Counting only genetically independent associations, *i.e.*, one per linkage group per trait, the linear regression analysis detected a total of 10 QTL regions: three for CBH, four for WSG and three for %BARK. No QTL was detected for %PULP as the highest P value found was 0.03 on group 6/9. Overlapping QTL for CBH, WSG and %BARK were observed on group 5 where either one or more genetic loci with pleiotropic effect or a cluster of linked genes control all three traits. Overlapping of QTL positions was also observed for CBH and WSG on groups 2 and 7 and for WSG and %BARK on group 6/9 (Figure 3).

The results of the interval mapping analysis using MAPMAKER-QTL agreed with the results of linear regression. However putative QTL for CBH on group 7 and for %BARK on group 14 were detected by linear regression but were not significant by interval mapping. Therefore a total of eight unlinked QTL were detected by interval mapping using MAPMAKER-QTL (Table 3). No QTL was detected for %PULP in the interval mapping analysis. However a LOD score of 1.5 was detected on group 6/9 in the interval U20_900–Z18_900.

Tests for linked QTL effects for CBH, WSG and %BARK were carried out on group 5 using MAPMAKER-QTL. For WSG, upon fixing the QTL in the first interval (X1_1450–R4_1300), a raise in LOD = 1.82 was observed at the position of the second putative QTL (interval Y17_1500–V7_1200). When the second QTL was fixed, a raise in LOD = 2.1 was observed at the first QTL. This result suggests two linked QTL effects for WSG on group 5. For CBH and %BARK however, no evidence for linked QTL were found.

Estimates of the proportion of phenotypic variation explained by each QTL were obtained with MAPMAKER-QTL (Table 4). For all traits, individual QTL explained between 3 and 10% of the phenotypic variation. Least square means of the alternative QTL genotypes and their associated standard deviations were also estimated (Table 3). For all the putative QTL detected, variances of the alternative QTL genotype classes were generally equal and close in value across QTL within traits. Estimates of the standardized difference in phenotypic mean trait value between the two alternative QTL genotypes (δ) ranged from 0.19 to 0.51 phenotypic standard deviations (σ_p) and most of the values were ~0.4.

Comparison of BSA, SG and CSA: BSA detected one clear polymorphic marker, U10_1500 on linkage group 7. With SG, significant differences both in marker allele means and frequencies between the high and low groups were also observed for a set of adjacent markers (U10_1500, A11_980 and K9_1660) corroborating the BSA result (Figure 2B). When CSA was applied, marker U10_1500 was not significantly associated with CBH. However, by linear regression, an adjacent genomic region on group 7 involving marker N15_1079 was found to be significant ($P = 0.007$) (Table 2). Furthermore, a genomic region involving markers A11_980 and K9_1660 was deemed significant for WSG (LOD 3.7) (Table 3). This result could be explained by the fact that CBH and WSG are positively correlated in this family, and WSG is a trait of higher heritability.

BSA for CBH did not detect the putative QTL on group 5 that was found by SG (marker R4_1300) and later found to be the one of largest effect by CSA. In our study, a multiple BSA strategy was adopted, which is significantly more stringent than the conventional BSA with only one bulked sample per phenotypic extreme. Only when all four bulks were polymorphic, a

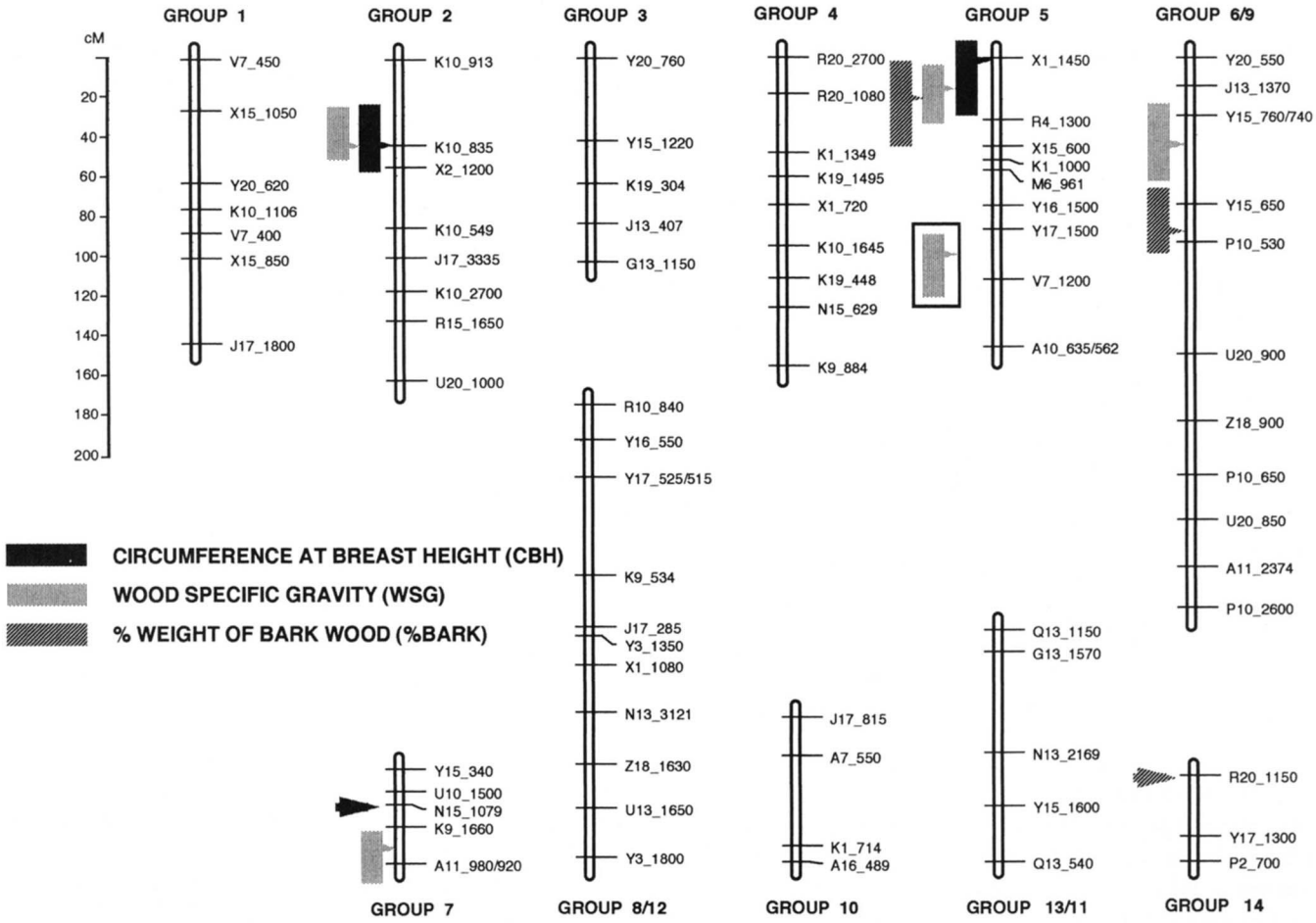

FIGURE 3.—QTL map of productivity traits for the *E. grandis* maternal genotype. Linkage maps of RAPD markers were constructed using MAPMAKER (LOD 5.0; $\theta = 0.30$) and markers were ordered with log-likelihood support $\geq$ 1000:1. Bars to the left of linkage groups correspond to the 1.0 LOD support intervals for the location of the QTL. Arrows from bars indicate the most likely position estimated with MAPMAKER-QTL. Arrows without bars for CBH on group 7 and %BARK on group 14 correspond to QTL detected only by linear regression. Boxed bar on group 5 corresponds to putative linked QTL effect detected with MAPMAKER-QTL.

putative linked marker was declared. This stringent procedure controls for false positives (*i.e.*, markers that appear polymorphic between bulks but are not linked to the trait expression), however it apparently results in a reduction of detection power. We assayed, *a posteriori*, individual samples that composed our bulks for markers later found to be in QTL regions (data not shown). We observed that one individual in 10 carrying the RAPD marker was enough to "contaminate" an otherwise null genotype bulk resulting in a negative BSA screening result. For example, marker Y17_1500 on group 5 was among the markers that showed ambiguous BSA polymorphisms (three out of four bulks being polymorphic) and was later determined to be linked to a major QTL region affecting CBH, WSG and %BARK by CSA.

Multi QTL analysis: Multilocus models by interval mapping and linear regression provided estimates of the total proportion of phenotypic variance explained by the joint action of the mapped QTL (Table 4). Over 10% of the phenotypic variance could be accounted for by the QTL mapped for CBH and %BARK and 22% for WSG. When testing for significant epistatic effects, with the exception of CBH, no significant two-way interactions were detected at $P \leq 0.01$. For CBH a highly significant positive interaction was detected between QTL on groups 5 (marker Y17_1500) and 7 (marker N15_1079) ($P = 0.0097$). This result indicates that a two-locus model including markers Y17_1500, N15_1079 and their interaction could account for 15% of the phenotypic variance in CBH. The nonadditive epistatic term was responsible for an almost 50% increase in the phenotypic variance explained by the model.

DISCUSSION

QTL mapping using extreme phenotypes: Our BSA experiment was successful in detecting a QTL for volume growth, a trait of low heritability, in a heterogeneous genetic background such as a half-sib family of an undomesticated forest tree. It is shown however, that BSA might miss important QTL due to sampling effects in the bulk composition and the sensitivity of the PCR for the specific RAPD marker assayed in the region.

TABLE 3

QTL summary for the *Eucalyptus grandis* maternal parent as determined by interval mapping analysis using MAPMAKER-QTL

Trait[a]	Linkage group	Marker interval	QTL[b] position	LOD peak	Support[c] interval	% variation[d] explained	Means ± SD[e] +	Means ± SD[e] −	δ[f]
CBH	2	K10_835–X2_1200	0.0	1.9	20.0–K10_835–X2_1200	3.9	58.9 ± 1.2	55.0 ± 1.3	0.26
	5	X1_1450–R4_1300	2.0	3.1	OFF END–X1_1450–26.0	6.6	60.0 ± 1.1	53.8 ± 1.3	0.42
WSG	2	K10_835–X2_1200	0.0	1.9	14.0–K10_835–6.0	3.4	473.9 ± 3.8	486.0 ± 4.9	0.34
	5	X1_1450–R4_1300	20.0	4.7	28.0–R4_1300–2.0	10.2	470.9 ± 3.1	486.6 ± 3.1	0.44
	5	Y17_1500–V7_1200	12.0	4.2	22.0–V7_1200–12.0	8.4	485.2 ± 2.9	471.6 ± 3.2	0.38
	6/9	Y15_760–Y16_650	14.0	3.1	4.0–Y15_760–30.0	9.1	472.6 ± 3.0	487.4 ± 3.5	0.41
	7	K9_1660–A11_980	10.0	3.7	16.0–A11_980–OFF END	6.1	483.3 ± 2.9	476.4 ± 3.1	0.19
% BARK	5	X1_1450–R4_1300	20.0	3.0	28.0–R4_1300–X15_600	7.2	10.19 ± 0.18	9.36 ± 0.18	0.41
	6/9	Y15_650–P10_530	16.0	3.3	6.0–Y15_650–P10_530–3.0	9.2	10.17 ± 0.16	9.23 ± 0.19	0.47

Listed are the locations and magnitudes of effect of QTL controlling growth and wood quality traits.

[a] CBH, circumference at breast height; WSG, wood specific gravity; % BARK, percentage dry weight of bark.

[b] Most likely QTL position corresponding to LOD peak, as estimated by MAPMAKER-QTL; cM distance from leftmost marker of interval.

[c] Interval over which the position of the QTL is at most 10 times less likely than the most likely position estimated by MAPMAKER-QTL; from left to right: cM distance from the left, marker segment and cM distance to the right; OFF END, off the end of linkage group.

[d] Percentage of the phenotypic variation explained, as estimated by MAPMAKER-QTL.

[e] Least square estimates of genotype means for the alternative RAPD marker-linked QTL alleles; (+) presence of the RAPD band; (−) absence of band.

[f] Difference between alternative QTL genotypes expressed in phenotypic standard deviations.

If a very large sample size is already available and the phenotype determination is quick and inexpensive, BSA followed by genotyping a random sample of individuals might be a very useful way to quickly identify one or some of the genomic regions controlling a quantitative trait of interest in a half-sib family. However, BSA is certainly not adequate to identify all the major genetic factors controlling the trait. A less stringent screening than the one adopted in this study might be useful in this respect. However, the large number of individual assays necessary to screen the false positives (*i.e.*, markers that appear polymorphic between bulks but are not linked to the trait expression) might overrule the intended advantage of BSA as a short cut for QTL mapping. Our experimental results agree with a recent theoretical assessment of DNA pooling strategies for mapping QTL where it is suggested that these might be successful in tagging only QTL of very large effect (WANG and PATERSON 1994).

Two out of the three genomic regions later detected by CSA as harboring QTL for CBH were identified by selectively genotyping a fraction of ~4.4% of the individuals at each extreme of the trait distribution. The two QTL detected by SG were the ones detected with higher significance level later in the CSA. It is important to point out, however, that marker U20_1000 in the vicinty of the third genomic region on linkage group 2, later identified as having a QTL, displayed a P value = 0.03 in the SG analysis (Table 1). Therefore, at a more relaxed threshold (α = 0.05), all three QTL found by CSA were in fact detected by SG. Selective genotyping was applied exclusively to CBH where trait determination is easy and an almost unlimited sample size was available for the study. We did not apply SG to wood quality traits. WSG and %PULP determinations are relatively costly, and an enlarged sample size could not be justified for SG. Furthermore new extreme phenotype groups would have to be composed and genotyped reducing the efficiency of the SG approach.

Within half-sib QTL mapping in forest trees: We have shown that the linkage disequilibrium within a half-sib family allows one to identify genomic regions

TABLE 4

Summary of the half-sib QTL analysis for growth and wood quality traits

	CBH	WSG	% BARK
No. of putative QTL detected			
Linear regression	3	4	3
Interval mapping (MAPMAKER-QTL)z	2	5	2
% Total phenotypic variance explained			
Linear regression	10.9	21.2	11.6
MAPMAKER-QTL	13.7	24.7	12.6

with significant effects on quantitatively inherited traits of relevance to forest productivity in Eucalyptus. This is the first report on the use of a half-sib pedigree for QTL mapping in forest trees.

A half-sib approach for QTL mapping in animals has been investigated by simulations (SOLLER and GENIZI 1978; HALEY 1991), and experimental results in cattle (*e.g.,* GELDERMANN *et al.* 1985; BEEVER *et al.* 1990; GEORGES *et al.* 1995) demonstrated the detection of significant marker-associated quantitative effects for a range of productivity traits. HALEY (1991) estimated that half-sib mapping will only have sufficient power to detect genes of relatively large effects. Using Haley's expression, at $\alpha = 0.001$, for a maximum map distance between a marker and QTL of 10 cM, our sample size ($n = 300$) would achieve a power of ~0.5 to detect QTL with $\delta = 0.5$ (δ corresponds to the standardized phenotypic difference between half-sibs of the two marker genotype classes) and power ≥ 0.95 for QTL with $\delta = 1.0$. For the majority of the QTL detected in our experiment, the standardized difference of the phenotypic means between marker genotype classes were between 0.2 and 0.5 (Table 3). It is unlikely therefore that QTL with effects $\delta \geq 1.0$ were missed in our analysis, however we might have missed approximately half of the QTL with effects $\delta = 0.5$, and evidently the majority of those with $\delta \leq 0.5$. The number of mapped QTL in this study should therefore be regarded as minimal.

To increase power for QTL detection in a half-sib experiment, a "granddaughter" design has been proposed where progeny tests of half-sib sires from a single elite sire are evaluated to decrease the error variance of quantitative trait evaluation (WELLER *et al.* 1990) This design is particularly attractive for animals where very large half-sib families from a single sire are difficult to obtain, and highly specialized pedigrees involving multiple generations are available. In forest trees, on the other hand, large half-sib families are readily available and three generation pedigrees are difficult to obtain.

Recently, we carried out QTL mapping using a full-sib family from a cross between heterozygous individuals using the pseudo-testcross strategy (GRATTAPAGLIA *et al.* 1995). With that strategy, the QTL variation contributed by both parents can be examined, while in a half-sib family only the variation contributed by the common parent is accounted for. Furthermore, detecting QTL, particularly for low heritability traits, could require more experimental effort in half-sib families. However, as pointed out by O'MALLEY and MCKEAND (1994) in the population breeding approach generally used in tree breeding programs, QTL identified in a full-sib family are not defined with respect to the additive genetic variance of the breeding population. Although in a full sib family one looks at QTL variation both in the maternal and paternal parents, we compare the effect of the two alleles from one parent, averaged over only two other alleles, those that come from the other parent. Half-sib families sample more of the genetic variation of the population as the effect of the QTL alleles segregating in the common maternal parent is measured with respect to the whole breeding population. In other words, one compares the effect of the two maternal alleles averaged over a much larger sample of alleles at that same QTL coming in the pollen pool. QTL identified in half-sib families can thus in principle be directly related to breeding value, additive genetic variation and thus the theory and practice of population based tree breeding programs.

Full-sib and half-sib approaches for QTL mapping depend on the heterozygosity of the QTL and linked marker. While the heterozygosity of the marker should not represent a limitation with most marker technologies, no prior information is available on the heterozygosity of the QTL. Genetically heterogeneous undomesticated populations such as those of forest trees are likely to have multiple alleles at QTL. Therefore it is expected that in trees, QTL should be segregating and could in principle be detected in almost any cross. The critical issue for detection becomes the distribution of allelic effects, *i.e.,* the relative magnitude of effects of the segregating alleles and their interactions. Evidently as populations undergo selection, QTL detection is expected to become more difficult as the frequency of favorable alleles with comparable effects increase. However in genetically heterogeneous populations we expect to find not only multiple alleles but also multiple QTL. Thus opportunity will exist to uncover important discrete genetic factors controlling traits of interest for at least several generations.

Genetic architecture of quantitative traits in Eucalyptus: To our knowledge this is the first report of QTL mapping for productivity related traits at rotation age in Eucalyptus. Our findings shed some light on the architecture of quantitative traits in forest trees and may have some important implications for the planning of future QTL mapping experiments and for marker assisted tree breeding. Given the magnitude of the phenotypic variation explained by the joint action of the putative QTL mapped, these results indicate that, genes or groups of closely linked genes with relatively large phenotypic effects ($\delta \sim 0.5$) are involved in the control of the quantitative traits investigated.

In a study of dairy cattle, BOVENHUIS and WELLER (1994) estimated that a QTL for butterfat accounted for additive genetic variance equivalent to 3.6% of the phenotypic variation. The heritability of butterfat was 12.9% and this QTL effect represented ~28% of the additive genetic variance estimated for the whole population. In our study, we can arrive at such an assessment based on estimates of heritability for the same traits, in similar genetic material, age and environment. For CBH, individual tree heritability of 0.31 (REZENDE and BERTOLUCCI 1993), the two QTL for CBH controlling

13.7% of the phenotypic variation, could account for 44% of the additive genetic variation. For WSG, heritability has been estimated between 0.5 and 0.7 (F. L. G. BERTOLUCCI, unpublished results), therefore the four QTL mapped could explain between 35 and 49% of the additive genetic variation.

By fitting multiple QTL models and their two-way interactions we detected a significant case of digenic epistasis between unlinked QTL that increased the total proportion of the phenotypic variance explained by the detected QTL. QTL mapping studies offer a valuable tool to investigate this source of nonadditive genetic variation. Epistatic interactions among QTL might prove of considerable importance in the architecture of quantitative traits and the advancement of selection in breeding populations of Eucalyptus. At the location of the major QTL involved in the interaction (on group 5), QTL for CBH, WSG and %BARK were also mapped. Perhaps a major gene located in that region is involved not only in the control of all the growth traits analyzed, but also affects the final trait expression by epistatic interactions with other genes.

Marker assisted selection (MAS) in Eucalyptus: Two steps are generally necessary for the implementation of MAS: location of QTL and their manipulation with genetic markers in subsequent generations of selection and recombination. With a half-sib approach used in this study the first step can be contemplated by a retrospective QTL analysis, *i.e.*, using an existing half-sib family. Existing tree genetic trials typically consist of several full-sib families that are too small for QTL mapping. However in diallel mating designs, half-sib families can frequently be composed by a collection of full-sib families related by a common parent tree. Furthermore, like in this study, large half-sib families are sometimes available in operational family block plantings. In the single-tree QTL approach adopted, marker/trait associations are established at the individual level, and therefore substantial linkage disequilibria is expected to be maintained. Close linkages established between markers and QTL could be followed for several subsequent generations of breeding. After QTL for individual trees have been detected in relatively large experiments, the number of markers genotyped in subsequent generations derived from those trees could be reduced as only those particular marker segments containing the QTL of interest would be tracked. Progeny sizes could then vary depending on the number of genomic regions targeted at selection to increase the probability of recovering genotypes with the correct QTL allele profiles.

The contemplation of marker assisted breeding should be done on a case-by-case basis (WILLIAMS and NEALE 1992). In the context of Eucalyptus breeding the prospects are positive. Hybrid breeding combined with clonal propagation of selected individuals is increasingly being used. Large amounts of linkage disequilibrium are generated by hybridization and substantial amounts of nonadditive genetic variation can be captured by vegetative propagation. These are favorable conditions for MAS (WELLER and FERNANDO 1991; STRAUSS *et al.* 1992). Individual trees of specialized breeding populations could be efficiently QTL mapped by analyzing the performance of their offspring, both full-sibs and half-sibs. From a cross involving one or two QTL mapped trees, efficiency of within-family selection for superior individuals could be significantly improved by the incorporation of marker information. Marker information would also be very useful for selecting specific pairs of parents that contain complementary QTL for subsequent breeding.

In our experiment, three QTL detected for CBH accounted for an estimated 44% of the additive genetic variance. When the proportion of the additive genetic variance explained by the marker loci exceeds the heritability of the character, selection on the markers alone is more efficient than selection on the individual phenotype (SMITH 1967). Plotting our results on the efficiency curves presented by LANDE and THOMPSON (1990), individual tree selection within-family for CBH based on the marker intervals detected could have an efficiency of ~1.3 times that of phenotypic selection alone.

Besides increasing accuracy of selection at the individual level for low heritability traits, efficiency could be significantly improved by decreasing the generation interval. Furthermore, indirect selection for traits difficult to score could also benefit from marker assisted breeding. In these situations not only low heritability traits could be considered but also high heritability traits. WSG is a trait of high heritability. However even in fast growing Eucalyptus, WSG requires a few years to reach full expression. Combined early selection for CBH and WSG using markers would allow a significant increase in selection intensity within families at a much younger age and save considerable time in the establishment of field trials of the individuals tested as clones.

With the advent of more efficient marker technologies and the demonstration of the existence of major genes for quantitative traits in trees, the prospects for the use of MAS in advanced generation breeding of forest trees are promising, contrary to earlier evaluations (STRAUSS *et al.* 1992). However expectations should not be overstated until data are accumulated on realized comparative gains from marker assisted selection *vs.* conventional phenotypic selection.

We thank our colleagues at North Carolina State University (NCSU), in particular JOSÉ X. CHAPARRO, DAVID O'MALLEY, BEN H. LIU, BRUCE WEIR and ZHAO-BANG ZENG. D.G. had a fellowship from CAPES (Brazilian Ministry of Education) and a research grant from PADCT/SBIO/FINEP (Brazilian Ministry of Science and Technology). We thank the NCSU Forest Biotechnology Industrial Associates Consortium and the USDA Competitive Grants Program (USDA award 91-37300-n6431) for their support. We also acknowledge the insightful reviews of three anonymous referees.

LITERATURE CITED

ARNHEIM, N., C. STARNGE and H. ERLICH, 1985 Use of pooled DNA samples to detect linkage disequilibrium of polymorphic restriction fragments and human disease: studies of the *HLA* class II loci. Proc. Natl. Acad. Sci. USA **82:** 6970–6974.

BEEVER, J. E., P. D. GEORGE, R. L. FERNANDO, C. J. STORMONT and H. A. LEWIN, 1990 Associations between genetic markers and growth and carcas traits in a paternal half-sib family of Angus catle. J. Anim. Sci. **68:** 337–344.

BOVENHUIS, H., and J. I. WELLER, 1994 Mapping and analysis of dairy cattle quantitative trait loci by maximum likelihood methodology using milk protein genes as genetic markers. Genetics **137:** 267–80.

BRADSHAW, H. D. JR., and R. F. STETTLER, 1995 Molecular genetics of growth and development in populus. IV. Mapping QTL with large effects on growth, form and phenology traits in a forest tree. Genetics **139:** 963–973.

CLAMP, P. A., J. E. BEEVER, R. L. FERNANDO, D. G. MCLAREN and L. B. SCHOOK, 1992 Detection of linkage between genetic markers and genes that affect growth and carcass traits in pigs. J. Anim. Sci. **70:** 2695–2706.

DEMUNER, B. J., and F. L. G. BERTOLUCCI , 1992 Tree breeding: a new approach through evaluation of genetic and phenotypic parameters of eucalypt wood and pulp. Proceedings of the 26th Congress of the Brazilian Pulp and Paper Industries, pp. 411–423, Sao Paulo.

DUDLEY, J. W., 1993 Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. Crop Sci. **33:** 660–668.

GELDERMANN, H., U. PIEPER and B. ROTH, 1985 Effects of marked chromosome sections on milk performance in cattle. Theor. Appl. Genet. **70:** 138–146.

GEORGES, M., D. NIELSEN, M. MACKINNON, A. MISHRA, R. OKIMOTO *et al.*, 1995 Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. Genetics **139:** 907-920.

GRATTAPAGLIA, D., and R. R. SEDEROFF, 1994 Genetic linkage maps of *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross mapping strategy and RAPD markers. Genetics **137:** 1121–1137.

GRATTAPAGLIA, D., F. L. G. BERTOLUCCI and R. R. SEDEROFF, 1995 Genetic mapping of QTLs controlling vegetative propagation in *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross mapping strategy and RAPD markers. Theor. Appl. Gen. **90:** 933–947.

GROOVER, A., M. DEVEY, T. FIDDLER, J. LEE, R. MEGRAW *et al.*, 1994 Identification of quantitative trait loci influencing wood specific gravity in an outbred pedigree of loblolly pine. Genetics **138:** 1293–1300.

HALEY, C. S., 1991 Use of DNA fingerprints for the detection of major genes for quantitative traits in domestic species. Anim. Genetics **22:** 259–277.

IKEMORI, Y., and E. CAMPINHOS, JR., 1983 Produção de sementes de *Eucalyptus grandis* × *Eucalyptus urophylla* por polinização aberta: resultados preliminares. Silvicultura **8:** 306–308.

KNAPP, S. J., W. C. BRIDGES and B. H. LIU, 1992 Mapping quantitative trait loci using nonsimultaneous and simultaneous estimators and hypothesis tests, pp. 209–238 in *Plant Genomes: Methods for Genetic and Physical Mapping,* edited by J. BECKMANN and T. C. OSBORN. Kluwer Academic Publishers, Dordrecht, The Netherlands.

LANDE, R., and R. THOMPSON, 1990 Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics **124:** 743–756.

LANDER, E. S., and D. BOTSTEIN, 1989 Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics **121:** 185–199.

LEBOWITZ, R. J., M. SOLLER and J. S. BECKMANN, 1987 Trait-based analyses for the detection of linkage between marker loci and quantitative trait loci in crosses between inbred lines. Theor. Appl. Genet. **73:** 556–562.

MICHELMORE, R. W., I. PARAN and R. V. KESSELI, 1991 Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. USA **88:** 9828–9832.

O'MALLEY, D. M., and S. E. MCKEAND, 1994 Marker assisted selection for breeding value in forest trees. Forest Genetics **1(4):** 231–242.

REZENDE, G. D. S. P., and F. L. G. BERTOLUCCI, 1993 Uso da seleção combinada na determinacão da eficiência da seleção precoce em progênies de meios-irmãos de *Eucalyptus urophylla.* Proceedings of the 1st Panamerican Forestry Congress, pp. 158–162, Curitiba.

SAS, 1988 *SAS/STAT User's Guide, version 6.03.* SAS Institute, Cary, NC.

SMITH, C., 1967 Improvement of metric traits through specific genetic loci. Anim. Prod. **9:** 349–358.

SOLLER, M., and A. GENIZI, 1978 The efficiency of experimental designs for the detection of linkage between a marker locus and a locus affecting a quantitative trait in segregating populations. Biometrics **34:** 47–55.

STRAUSS, S. H., R. LANDE and G. NAMKOONG, 1992 Limitations of molecular-marker aided selection in forest tree breeding. Can. J. For. Res. **22:** 1050–1061.

STUBER, C. W., 1992 Biochemical and molecular markers in plant breeding, pp. 37–61 in *Plant Breeding Reviews,* Vol. 9, edited by J. W. DUDLEY, A. R. HALLAUER and M. RYDER. John Wiley & Sons, Inc., New York.

WANG, G.-L., and A. H. PATERSON, 1994 Assessment of DNA pooling strategies for mapping of QTLs. Theor. Appl. Genet. **88:** 355–361.

WELLER, J. I., and R. L. FERNANDO, 1991 Strategies for the Improvement of animal production using marker-assisted selection, pp. 305–328 in *Gene-mapping Techniques and Applications,* edited by L. B. SCHOOK, H. A. LEWIN and D. G. MCLAREN. Marcel Dekker Inc., New York.

WELLER, J. I., Y. KASHI and M. SOLLER, 1990 Power of "daughter" and "granddaughter" designs for mapping of quantitative traits in dairy cattle using genetic markers. J. Dairy Sci.**73:** 2525–2532.

WILLIAMS, C. G., and D. B. NEALE, 1992 Conifer wood quality and marker-aided selection: a case study. Can. J. For. Res. **22:** 1009–1017.

Communicating editor: P. D. KEIGHTLEY