Quantitative trait dissection analysis in Eucalytus using RAPD markers: 2. Linkage disequilibrium in a factorial design between E. urophylle and E. Grandis.

Daniel Verhaegen, Christophe Plomion, Mireille Poitel, Paulo Costa, Antoine Kremer

QUANTITATIVE TRAIT DISSECTION ANALYSIS IN *EUCALYPTUS* USING RAPD MARKERS: 2. LINKAGE DISEQUILIBRIUM IN A FACTORIAL DESIGN BETWEEN *E. UROPHYLLA* AND *E. GRANDIS*

Daniel Verhaegen[1], Christophe Plomion[2], Mireille Poitel[1], Paulo Costa[2] & Antoine Kremer[2]

[1)] CIRAD Forêt, Baillarguet BP 5035, F–34032 Montpellier Cedex 1, France
[2)] INRA, Station de Recherches Forestières, Laboratoire de Génétique et Amélioration des Arbres Forestiers, BP 45, F–33610 Cestas, France

Received January 15, 1997; accepted January 15, 1998

ABSTRACT

A 13 × 13 factorial design between *E. urophylla* and *E. grandis,* comprising 87 full-sib families, was used to assess the relationships between RAPD marker frequency classes obtained from parental genotypes and the interspecific additive mean (IAM) of the hybrid progeny. For any marker showing a significant association, the cumulative number of the " present band " allele in the parents was significantly correlated either positively or negatively, with the IAM of the traits studied: *i.e.* volume, stem taper and wood quality. We discuss the potential origin of such correlations in terms of linkage disequilibrium between QTL allele and marker allele. We also examine the possible use of such information, firstly in order to select the parents for further generations of breeding, and secondly in order to choose the hybrid families in which QTAs of specific value could be detected and used to identify the best trees to be vegetatively propagated for the production of clonal variety.

Key words: *Eucalyptus*, factorial design, QTL, marker-assisted selection, linkage disequilibrium

INTRODUCTION

Forest trees are allogamous undomesticated organisms characterized by a high genetic heterogeneity (HAMRICK *et al.* 1992). The association between alleles at marker loci and at quantitative trait loci (QTL, GELDERMANN, 1975) is likely to be in linkage equilibrium in the populations (reviewed by STRAUSS *et al.* 1992), due to high levels of heterozygosity and because of recombination. Therefore, marker-assisted selection (MAS) might be questionable for forest trees. Conversely, in autogamous crop plants or species treated as such for breeding purposes, it is expected that the association between a marker allele and a linked QTL allele (QTA) would be consistent over many individuals of a breeding population due to linkage disequilibrium.

However, relying on the RAPD (random amplified polymorphic DNA) technology (WILLIAMS *et al.* 1990), GRATTAPAGLIA and SEDEROFF (1994) presented a solution to deal with linkage equilibrium in forest trees. They showed that it was possible to construct individual tree maps in a very short period of time, allowing the determination of location, number and the effect of QTLs within virtually every two-generation pedigree in a breeding program. In the mapping strategy they developed ("the two-way pseudo-testcross mapping strategy"), two parental maps are first constructed, with RAPD segregating in the 1:1 ratio within a hybrid family, and a QTL analysis carried out independently for both parents of the cross under the backcross model. In such full-sib progeny of heterozygous parents, the analysis of marker-trait association leads to the detection of "specific" QTLs (LEONARDS-SCHIPPERS *et al.*, 1994; GRATTAPAGLIA *et al.* 1995, VERHAEGEN *et al.* 1997). QTLs detected in such a narrow genetic background can readily be used to identify the ideotype to be propagated for clonal variety production (GRATTAPAGLIA *et al.* 1995; O'MALLEY, 1996, VERHAEGEN *et al.* 1997). Such a marker-assisted clonal propagation is valuable for those species where vegetative multiplication is efficient (*e.g.* eucalyptus, poplar, willow, cryptomeria, larch, radiata pine, acacia, gmelina, ayous). However, in order to use such QTL for breeding purposes, *i.e.* to select parents for subsequent generations of breeding, the value of favorable QTAs need to be evaluated over a wider genetic background. An analytical method aimed at estimating the breeding value of such RAPD-QTA in forest tree populations has been proposed by PLOMION and DUREL (1996). It is based on the analysis of the half-sib families, involving both parents of the hybrid progeny. Alternatively, ARCADE *et al.* (1996) have used factorial design in larch to detect QTL of general value. Their approach

consisted on genotyping the parents with RAPD markers, measuring quantitative traits on the full-sib progenies, and investigating the statistical relationship between expected RAPD marker "present band" allele frequency in the F_1, and family performances. MURANTY (1996) also proposed the use of diverse experimental designs involving FS families as an efficient strategy to detect QTL in outbred species. Her approach involved marker genotyping and trait measurement on F_1 progeny. However the number of progeny to be genotyped could be rapidely prohibitive. When few traits are analyzed, MURANTY *et al.* (1997) considered a selective genotyping approach for the location and estimation of the effect of a quantitative trait locus.

Here, we present experimental data on *Eucalyptus* where the significance of marker-trait associations was determined using a factorial design of the breeding program. Our objectives were twofold: (i) to determine to what extent markers (whether linked or not to putative QTLs detected in a single full-sib family) could affect the expression of wood density, volume of the bole and form of the trunk, in a larger genetic background, and (ii) to use molecular marker data to predict the interspecific additive mean of hybrid families. The eucalyptus breeding program is based on hybrids of *E. urophylla* × *E. grandis* which have demonstrated a high potential for industrial plantation in the Congo (VIGNERON 1991). Today, commercial plantations of *Eucalyptus* in this country cover 45,000 hectares and the principal product is round wood for export to pulp mills.

MATERIALS AND METHODS

Genetic materials and factorial design analysis

A factorial design (R90-11, Table 1) consisting of 13 *E. urophylla* as female parents and 13 *E. grandis* as male parents was established in 1990. Mother trees were selected in eight open pollinated families from the Monte Lewotobi provenance in the Island of Flores, and in three open pollinated families from the Monte Egon provenance on the Island of Timor, of the Sonda Archipelago. Pollen used for crosses was collected from seven provenances in the northern part of the natural area near Atherton, Queensland (Australia). This design was incomplete since the ratio of full cells to total cells was 51%, *i.e.* 87 interspecific F_1 progeny were available (Table 1). The design was established in 32 incomplete blocks with a variable number of replicates (1 to 4 among crosses, Table 1). The experimental unit was a square plot of 4*4=16 hybrids. The traits were: (i) pilodyn pin penetration depth in the trunk (PIL) measured with a penetrometer at 18 months; this trait is commonly used to evaluate wood density in the eucalyptus breeding program, (ii) vigor measured at 18, 26 and 38 months by means of the volume of the bole (VIG); the volume was calculated from height and circumference, by considering the trunk as a cone surperimposed on a cyclinder, (iii) height/diameter ratio (HDR) also measured at 18, 26 and 38 months. 38 months corresponded to actual selection age. Individual values were adjusted for the block effect with the software package OPEP (BARADAT 1989) and then used in the following fixed model:

$$Y_{ijk} = \mu + M_i + F_j + M^*F_{ij} + \varepsilon_{ijk}$$

were Y_{ijk} is the phenotypic value of individual k^{th} of the crossing between i^{th} male and j^{th} female; μ is the overall mean; M_i is the effect of i^{th} male; F_j is the effect of j^{th} female and M^*F_{ij} is the interaction effect between i^{th} male and j^{th} female; and ε_{ijk} is the within-replicate residual.

The factorial mating design allowed estimation of some genetic variance components: additive and dominance variances. We assumed that epistatic variance was negligible and that the inbreeding coefficient of the parents was zero. In this paper we only consider the mean estimated for the hybrid population. The dominance effects for the studied traits were rather low, and decreased with age (BOUVET & VIGNERON 1995). This parameter was not considered in this study but many details relating to the prediction of the specific combining ability have been reported by BARRIL *et al.* (1997a, 1997b). We did not investigate the marker-GCA associations at parental level because sample sizes were low (13 trees in each species). Rather, we studied the association between the interspecific additive mean ($IAM_{ij} = \mu + M_i + F_j$) and the frequency classes of the RAPD " present band " allele in hybrid families.

RAPD assay

DNA was extracted from dried leaves of the 26 parents and used for RAPD assay as described by VERHAEGEN and PLOMION (1996). These parents were genotyped by two sets of markers:

- •26 RAPD loci located in the two maps of *E. urophylla* and *E. grandis* (see markers in bold type in figure 1 of VERHAEGEN *et al.* 1997), either linked or not with putative QTLs. Amplified products were subjected to electrophoresis on 1.8% agarose gels cast in 0.5× TBE and run in 0.5× TBE at $4V \cdot cm^{-1}$ for 5 hours.
- 415 polymorphic RAPD fragments. These markers were not located in the parental maps. They were

Table 1. Factorial design with clone nomenclature. Trees follows by the same number of asterisk are half-sibs. (Numbers (1,2,3,4) report to the number of replicates of 16 trees.

E. urophylla	*Eucalyptus grandis*												
	9–30	9–33	9–36	9–37	9–38	9–39	9–40	9–41	9–42	9–43	9–44	9–45	9–46
14–128			4	4	2		4	4	4	2	1		1
14–130		1	4	2	4				2	4			3
14–132*			4	4			4	4		4			2
14–133**	4		4	2		4	4	3	4		4		
14–135	1				4	4			4				2
14–136*	2	2	4	4	4		4	4	4	4	4		1
14–137**					1			4	4				
14–138	4			2					4				
14–139***		3	1		4	4	1	4	4	4	4		
14–140***				2			4		4				
14–141	3	3	2	2	1		4	4	2		1	1	
14–144			4	4	4	4	4	4	4	4	4	4	4
14–145					1	4							

Table 2. Marker genotypes of the female and male parents, and expected frequency of the RAPD "present band"allele (+) in hybrid families originating from different parental genotype combinations. Squared cells are not distinguished with dominant markers and represent three possible classes (I, II, III) to which the hybrid progeny can be assigned.

		male parent		
	PHENOTYPE GENOTYPE	band present [+]		band absent [–]
		+ +	+ –	– –
female parent				
PHENOTYPE	GENOTYPE	(I)		(II)
band present [+]	++	++ 100 %	½ ++ : ½ + – 75 %	+ – 50 %
	+ –	½ ++ : ½ +– 75 %	¼ ++ : ½ + – : ¼– – 50 %	½ + – : ½ – – 25 %
		(II)		(III)
band absent [–]	– –	+ – 50 %	½ + – : ½ – – 25 %	– – 0 %

used as an independant data set for validation purpose. Amplified products were loaded in 8% acrylamide gels, and run in 0.5× TBE at 18.5V·cm^{-1} for 2 hours.

All the reactions were repeated twice and only reproducible bands were considered in this study. The negative films were scanned and electronic images analyzed with the BIOIMAGE imaging system (Bio Image, Ann Arbor, MI) using the Whole Band Analyzer. A spreadsheet with 1 and 0 coding for the presence and absence of a RAPD fragment, respectively, was automatically produced using the LANE MATCH" option under the MATCHING RESULTS" command, and later used for analysis.

Marker-trait association

All the RAPD markers were scored by the presence or the absence of a specific amplification product in the 26 parents. The relationships between the molecular

data and the interspecific additive mean (IAM) in the hybrid population (87 data points) was evaluated using one-way ANOVA, with the "present band" allele frequency classes for each marker taken as a factor. These frequencies were deduced from the RAPD genotypes of the parents (Table 2). Basically, the 87 full-sib families were grouped into either 2 or 3 frequency classes, depending on whether the marker was polymorphic in one or both parental species, respectively. It is obvious that the dominant nature of the RAPD markers led to a certain imprecision in the estimation of the "present band" allele frequency. Families belonging to class I (see Table 2) were composed of hybrid progenies with " present band " allele frequency with values of 100%, 75% or 50%. Families belonging to class II (see Table 2) were composed of hybrid progenies with "present band " allele frequency with values of 50% or 25%. Only families produced from homozygous null parents could be classified unambiguously as 0% (class III in Table 2).

All parents of the factorial design were represented but not equally. Indeed, the number of FS families and the number of F1 individuals within each FS family could vary (Table 2). In addition sample size within each genotypic class could also vary. In order to take into account this unbalanced experiment, a significance threshold for marker class-IAM association was established by using a randomization test, implemented in OPEP, as follows: (i) the data were permuted by scrambling the relationship between the 87 IAM and the marker classes. This created data where the null hypothesis (no association) was true; (ii) one-way ANOVA with permuted data was performed with the marker class as a factor; (iii) these two steps were repeated 1000 times; (iv) finally we chose a threshold that would be exceeded by only 1% of all permutations.

After this selection procedure, significant RAPD markers were used in a multiple regression analysis using the Splus software (BECKER *et al.* 1992) in order to predict the IAM of any family within the factorial design. Basically each characteristic was treated as a dependent variable and the various RAPD markers as independent variables.

RESULTS AND DISCUSSION

QTA of "specific" value vs. QTA of "general" value

The association between RAPD polymorphisms and the IAM was evaluated at 18 months for PIL and at 18, 26 and 38 months for VIG and HDR. The 26 parents of the factorial design were genotyped with 26 markers of known location in the maps reported by VERHAEGEN *et al.* (1997). Of these, 13 of them were associated with a QTL while 13 were not, as indicated in the third column of Table 3. Marker class-IAM association was investigated using one-way ANOVA. Some markers presented significant effects (based on the permutations test) of the "present band" allele frequency classes on the IAM (Table 3).

The most important point was that for markers segregating in both species, the three classes could be ordered with the mean value of the progeny belonging to either class I or class III being the best or the worst, and the mean value of the progeny belonging to class II being intermediate. An example of such these result is given for VIG38 in Figure 1. The fact that the three classes could be ordered, *i.e.* that a significant marker class-IAM could be detected from "present band" allele frequency changes, agrees with the existence of QTLs of "general combining" value.

Out of the 13 RAPD markers associated with a quantitative trait in the full-sib (FS) progeny, 4 were associated with an identical trait at an identical age in the factorial design (FD): i.e., for *E. grandis* X12_633 (VIG38), L08_343 (HDR26), B01_576 and U20_1358 (PIL18). These results agree with the existence of a linkage disequilibrium between these markers and QTL alleles in the studied hybrid population. However linkage disequilibrium was not a general rule. Indeed, 5 markers were not associated with any traits: i.e., for *E. grandis* D03_618, N14_1588, A09_1192, Q05_525, F04_796. Also, 2 markers were associated with the same trait but at different age: *i.e.*, for *E. urophylla* Z09_808, for *E. grandis* R15_625. Other markers were associated with completely different traits: *e.g.* in *E. urophylla*, Z03_925 associated with VIG38 in the FS was associated with PIL18 in the FD. A very interesting example of association between marker allele and QTA in the population was shown by two tightly linked markers in LG1 for *E. grandis:* L08_343 and X12_633. These markers were associated with several QTLs (PIL18,26, HDR26,38, and VIG26,38) in the FS family (Fig.1 of VERHAEGEN *et al.* 1997). However, whilst the first marker showed a strong relationship with HDR26 and no association with VIG38 in the factorial design, the second marker showed strictly opposite associations, i.e. it was strongly associated with VIG38 and did not show any association with HDR26. Indeed, because of the linkage equilibrium, the physical association between favorable alleles at a marker and at a QTL could have been broken even between closely linked markers. As pointed out by STRAUSS *et al.* (1992), given the large effective population sizes (Ne>1000) in forest trees, linkage disequilibrium due to physical linkage (CROW & KIMURA 1970) would be expected only at recombination frequencies (θ) below

Table 3. RAPD: interspecific additive mean associations for markers located in the *E. urophylla* (A) and *E. grandis* (B) genetic maps. Linkage group numbers refer to VERHAEGEN and PLOMION (1996). It is indicated where a RAPD marker was linked with a QTL detected in the full-sib family used by VERHAEGEN *et al.* (1997) (QTL link.). Shown are the values of the Fisher test and in parentheses the associated F-value of the 1% probability level determined by bootstrap as described in the Material and Method section. ns: non significant. The adjusted R-squared (adj-R^2) were determined by multiple regression analyses carried out with all significant markers. Abbreviations are: HDR - height diameter ratio; VIG - vigor and PIL - pilodyn pin penetration depth, measured at 18, 26 and 38 months.

RAPD marker	Linkage group	QTL link	DR18	VIG18	HDR26	VIG26	HDR38	VIG38	PIL18
A									
K10_528	1	/	ns	6.1 (6.0)	ns	9.9 (4.7)	ns	10.6 (4.8)	ns
Z03_925	1	VIG38	ns	ns	6.3 (6.4)	ns	ns	ns	9.8 (4.7)
Y13_869	2	/	ns	ns	ns	11.3 (6.4)	ns	17.1 (4.9)	ns
K10_771	3	/	ns	ns	ns	ns	ns	ns	ns
Z09_808	3	HDR18	ns	ns	ns	ns	6.3 (5.0)	ns	ns
D02_1050	4	/	ns	ns	ns	ns	ns	ns	ns
A10_1304	4	/	ns	ns	ns	ns	ns	ns	ns
I4_1209	6	/	ns	ns	9.4 (4.4)	ns	5.6 (5.1)	ns	ns
G14_841	6	/	ns	ns	11.1 (5.2)	ns	ns	ns	ns
Z12_730	7	/	ns	ns	ns	10.9 (5.6)	ns	15.8 (6.6)	16.6 (6.3)
X19_562	11	PIL26,38 HDR18,26	10.2 (4.4)	7.6 (5.2)	ns	6.4 (4.9)	ns	ns	ns
B									
X12_633	1	VIG 26,38 HDR 26,38 PIL 18,26	ns	ns	ns	ns	ns	5.3 (5.0)	ns
L08_343	1	VIG 26,38 HDR 26, 38 PIL 18, 26	ns	ns	7.9 (5.5)	ns	ns	ns	ns
D03 _618	2	VIG 38 HDR 38	ns	ns	ns	ns	ns	ns	ns
M05_1572	3	/	ns	ns	ns	ns	ns	ns	ns
R15_625	4	PIL 38	ns	7.8 (4.3)	ns	11.4 (8.0)	ns	11.1 (4.7)	7.5 (5.1)
U20_1358	5	PIL 18	ns	7.4 (5.6)	15.4 (6.0)	ns	9.1 (5.0)	ns	11.4 (5.0)
Y05_1299	5	/	ns	ns	ns	ns	ns	ns	ns
N14_1588	5	VIG 26 HDR 26	ns	ns	ns	ns	ns	ns	ns
A09_1192	7	PIL 18	ns	ns	ns	ns	ns	ns	ns
Q05_525	8	HDR 18, 26	ns	ns	ns	ns	ns	ns	ns
R12_1654	8	/	ns	8.3 (6.7)	ns	14.3 (7.4)	ns	11.3 (6.8)	ns
R04_1844	8	/	ns	26.9 (7.6)	ns	19.2 (7.6)	ns	11.6 (4.8)	ns
B01_576	11	PIL 18	ns	ns	ns	ns	ns	ns	15.4 (6.0)
F04_796	2	VIG 26	ns	ns	ns	ns	ns	ns	ns
A09_619	11	/	ns	ns	6.4 (5.3)	ns	ns	ns	ns
adj R^2 (P<0.001)			0.19	0.46	0.56	0.59	0.38	0.52	0.77

0.025cM (θ <Ne/4, HILL & ROBERTSON 1968), which is below the level of our map (VERHAEGEN & PLOMION 1996) and any published genetic map. Another possible explanation could be that these two markers had different allelic status in the parental species and therefore presented different " present band " allele frequency in the hybrid progeny (see table 1), leading to different patterns of associations.

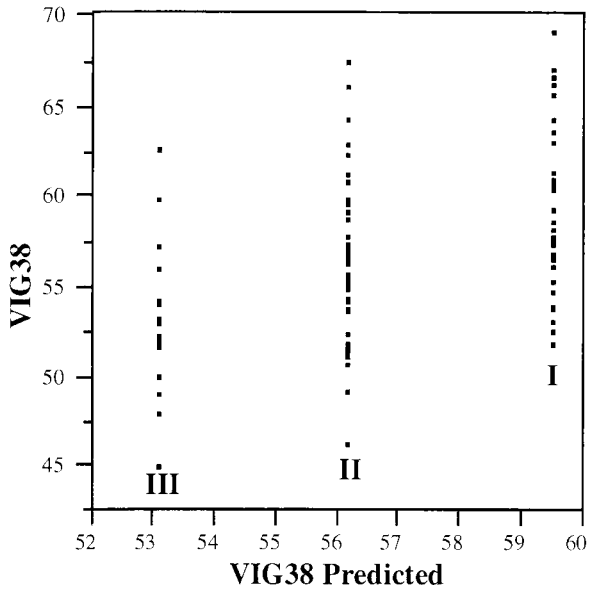
Out of the 13 RAPD markers that were not associated with any trait in the FS; 4 markers did not show any association with any trait in the factorial design: i.e. for *E. urophylla* K10_771, A10_1304, and for *E.*

Table 4. Same as for Table 3 on a restricted data set.

RAPD marker	Linkage group	QTL link	DR18	VIG18	HDR26	VIG26	HDR38	VIG38	PIL18
A									
K10_528	1	/	ns	8.7 (6.2)	ns	12.6 (8.3)	ns	13.3 (6.4)	ns
Z03_925	1	VIG38	ns	ns	5.1 (5.0)	ns	ns	ns	10.5 (6.2)
Y13_869	2	/	ns	ns	ns	19.5 (6.8)	ns	21.5 (6.8)	ns
K10_771	3	/	ns	ns	ns	ns	ns	ns	ns
Z09_808	3	HDR18	ns	ns	ns	ns	9.8 (4.8)	ns	ns
D02_1050	4	/	ns	ns	ns	ns	ns	ns	ns
A10_1304	4	/	ns	ns	ns	ns	ns	ns	ns
I4_1209	6	/	ns	ns	13.4 (4.4)	ns	5.0 (2.5)	ns	ns
G14_841	6	/	ns	ns	8.7 (4.7)	ns	ns	ns	ns
Z12_730	7	/	ns	ns	ns	14.3 (7.2)	ns	20.4 (6.6)	14.4 (5.9)
X19_562	11	PIL26,38 HDR18,26	8.4 (5.4)	6.9 (4.0)	ns	6.2 (4.8)	ns	ns	ns
B									
X12_633	1	VIG 26,38 HDR 26,38 PIL 18,26	ns	ns	ns	ns	ns	5.3 (5.0)	ns
L08_343	1	VIG 26,38 HDR 26, 38 PIL 18, 26	ns	ns	6.6 (6.3)	ns	ns	ns	ns
D03 _618	2	VIG 38 HDR 38	ns	ns	ns	ns	ns	ns	ns
M05_1572	3	/	ns	ns	ns	ns	ns	ns	ns
R15_625	4	PIL 38	ns	7.6 (5.2)	ns	9.4 (5.3)	ns	8.4 (4.8)	**ns**
U20_1358	5	PIL 18	ns	5.7 (4.7)	15.7 (5.3)	ns	9.8 (4.5)	ns	11.0 (5.0)
Y05_1299	5	/	ns	ns	ns	ns	ns	ns	ns
N14_1588	5	VIG 26 HDR 26	ns	ns	ns	ns	ns	ns	ns
A09_1192	7	PIL 18	ns	ns	ns	ns	ns	ns	ns
Q05_525	8	HDR 18, 26	ns	ns	ns	ns	ns	ns	ns
R12_1654	8	/	ns	8.9 (7.4)	ns	12.9 (7.0)	ns	9.4 (7.9)	ns
R04_1844	8	/	ns	26.4 (7.4)	ns	15.9 (5.8)	ns	9.4 (6.8)	ns
B01_576	11	PIL 18	ns	ns	ns	ns	ns	ns	17.4 (4.7)
F04_796	2	VIG 26	ns	ns	ns	ns	ns	ns	ns
A09_619	11	/	ns	ns	9.5 (7.2)	ns	ns	ns	ns
adj R^2 ($P<0.001$)			0.19	0.52	0.60	0.71	0.49	0.56	0.80

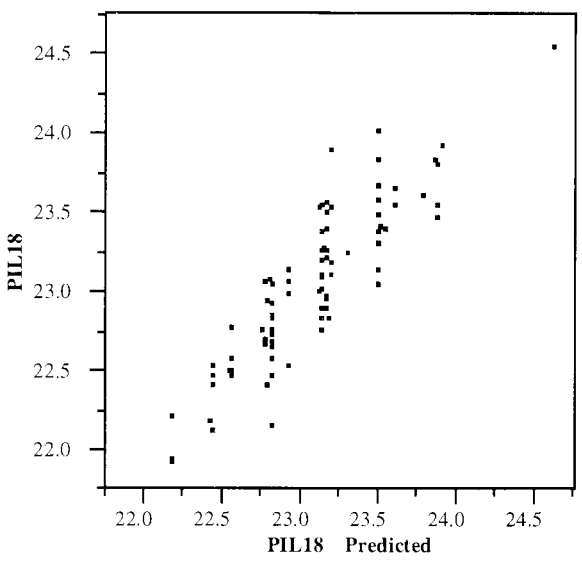
grandis Y05_1299, M05_1572; whereas 9 presented highly significant associations with the IAM for at least one trait, e.g. for *E. urophylla* K10_528 in LG1 and Y13_869 in LG2. Some of them presented interesting features: e.g. for *E. urophylla* I14_1209 and G14_841 located at both end of LG6 (Fig.1 of VERHAEGEN *et al.* 1997), as well as for *E. grandis* R12_1654 and R04_-1844 closely linked in LG8, presented a similar pattern of association since they were linked with the same traits (Table 3). This demonstrated again the existence of linkage disequilibrium between marker allele and QTA in the studied population without necessarily any strong physical linkage.

Such disequilibrium was not expected for an undomesticated outbreed species like eucalyptus. The lack of linkage disequilibrium between marker loci and loci involved in the variation of a quantitative character for a non-domesticated allogamous species such as forest trees, is well documented in population genetics studies (reviewed by STRAUSS *et al.* 1992). The origin of the observed linkage disequilibrium in our study is



Source	DF	SS	MS	F ratio	Prob>F
Model	2	470.8	235.4	11	0.0001
Error	84	1788.9	21.3		
C Total	86	2259.7			

Figure 1. Analysis of variance among RAPD "present band" frequency classes (I, II, III, see table 1) at marker R15_625 for VIG 38.



Effect test

Source	DF	SS	F Ratio	Prob>F
Z03_925	2	1.16	9.14	0.0003
Z12_730	1	2.24	35.18	0.0000
R15_625	2	2.94	23.11	0.0000
U20_1358	2	2.18	17.15	0.0000
B01_576	2	1.83	14.35	0.0000

ANOVA

Source	DF	SS	MS	F ratio	Prob>F
Model	9	16.72	1.86	29.20	0.0000
Error	77	4.90	0.06		
C Total	86	21.62			

Figure 2. Multiple regression on Interspecific Additive Mean (IAM) using the 5 RAPD markers associated with the pilodyn pin penetration depth measured at 18 months (PIL 18). The adjusted determination coefficient of this analysis is 0.77 ($P<0.001$).

unknown and whether these associations would hold for the whole breeding population is still an open question that need to be tested. We can only suggest that this linkage disequilibrium may originate, in part, from a physical linkage, which means that linkage between QTA and allele at marker loci has been preserved throughout many generations.

This disequilibrium could well be a simple consequence of the narrow genetic base of one or both sets of parents, 3 out of the 13 mother trees being half-sibs (Table 1). We therefore perfomed the same analysis on a restricted data set (76 families) were 3 half-sib mother trees (accession 14-140, 14-137 and 14-132 in table 1) were discarded. The same significant associations were detected (Table 4) but for one case (PIL18 against RAPD marker R15-625).

Alternatively, linkage disequilibrium might originate from selection on epistatic interaction (LEWONTIN 1974), or genetic drift (HILL & ROBERTSON 1968; OHTA & KIMURA 1969) caused by the geographical origin of both species. Indeed, the two selected populations of *Eucalyptus* used for the breeding program in the Congo originated from two particular natural areas. The *E. urophylla* trees were sampled from a very small area on the Flores Island, while the *E. grandis* trees were sampled near Atherton (Australia) in the north of the natural range.

The eucalyptus breeding program involves crossings between highly heterozygous individuals, where allele at marker and QTL loci are likely to be in linkage equilibrium within each species. Interspecific hybridization is known to generate linkage disequilibrium (STRAUSS *et al.* 1992) and therefore provided opportunities to detect associations between QTL and markers. Interestingly other authors dealing with interspecific crosses (European crossed by Japanese larch) have also detected stronger linkage disequilibrium than expected in intraspecific crosses (ARCADE *et al.* 1996).

Stability of marker class-IAM associations across ages

VIG and HDR were measured at three different ages in the factorial design. This made it possible to investigate the stability of marker class-IAM associations across ages. The results (Table 3) showed that some markers were consistently associated with a particular trait: e.g. K10_528 (LG 1, *E. urophylla),* R12_1654, R04_1844 (LG8, *E. grandis*) and R15_625 (LG4, *E. grandis*) with VIG. Other markers were significantly associated at two ages or only one, *e.g.* I14_1209 (*E. urophylla*, LG6) and U20_1358 (*E.grandis*, LG5) for HDR26 and HDR38. Selection age of eucalyptus trees occurs at 38 months. This result shows that molecular markers might be a useful tool to reduce this time lag and therefore increase genetic gain per unit of time.

Marker-assisted selection of parental trees and hybrid families

In an advanced tree breeding program using molecular marker technology, it will be important to choose those trees that combine both favorable QTAs of "general" value and superior phenotypes. In the particular case of the reciprocal recurrent selection scheme (RRS) developped for *Eucalyptus* (VIGNERON 1991), selection of parental trees to be crossed within each species is based upon their additive effects, determined in factorial designs involving controlled crossings between both species.

For markers segregating in both species and therefore presenting three frequency classes for the " present band" allele (either absent in both parent, present in one parent and absent in the other, or else present in both parents), we observed that the cumulative number of "present band" allele in the parents was correlated either positively or negatively with the IAM. QTL detected by such analysis may be of importance for the RRS scheme. Indeed, they could help with the identification of parents containing QTAs useful for improving performances of hybrid families, in subsequent generation of breeding.

Markers showing significant "present band" frequency classes-IAM associations were used in a multiple linear regression analysis in order to predict the performances of hybrid families (Table 3). Adjusted R-squared were significant ($P<0.001$) and ranged from 0.20 to 0.77. For VIG38, a trait that is used as an early selection criteria in the breeding program, the 7 significant markers explained 52% of the IAM variation. For PIL18, the 5 significant markers accounted for 77% of the IAM variation (Fig. 2). Narrow sense heritability of this trait was 0.85 (BOUVET 1995). In order to validate this approach in the prediction of IAM, an independent data set of 415 markers was used; from which we only report results for VG38 and PIL18. A total of 28 and 20 RAPD markers were significantly associated with VIG38 and PIL18 ($P<0.01$ determined by permutation tests), respectively. The exhaustive method of the stepwise function of the Splus software was used to find the best model for VIG38 and PIL18, including the same number of markers; *i.e.* 7 and 5 explanatory markers for these two traits respectively. For VIG38 and PIL18, adjusted R^2 were 90% and 74%, with Cp Mallow's coefficient taking values of 8 and 6, respectively.

Both data sets demonstrated the predictive power of the multiple regression procedure. This method could be used either as an initial screening step for the identification of the best existing hybrid families or for selecting the best parental combination, and eventually for producing hybrid progeny of great value in which specific QTLs could be mapped. It follows that these QTLs could be used to detect ideotypes that combine several favorable QTAs, which could then be vegetatively propagated for the production of clonal varieties (VERHAEGEN *et al.* 1997).

ACKNOWLEDGMENTS

This work was supported by CIRAD-Forêt and INRA, for the duration of the Ph.D of DV at the INRA Station de Recherches Forestières de Bordeaux-Cestas. We wish to thank J. M. Bouvet for collecting leaves in the Congo and for measuring the traits. Thanks are also due to the comments of the referees.

REFERENCES

ARCADE, A., FAIVRE-RAMPANT, P., LE GUERROUE, B., PAQUES, L.E. & PRAT, D. 1996: Quantitative traits and genetic markers: analysis of a factorial mating design in larch. *In*: "Somatic Cell Genetics and Molecular Genetics of Trees". Eds MR Ahuja, W Boerjan, D Neale. Kluwer Academic Publishers, Dordrecht, The Netherlands, 211–216.

BARADAT, PH. 1989: Amélioration génétique des arbres forestiers. Eléments méthodologiques. INRA, Laboratoire de Génétique et d'amélioration des arbres forestiers.

BARRIL, C. P., VERHAEGEN, D., VIGNERON, P., BOUVET, J.M. & KREMER, A. 1997a: Stucture of the specific ability between between two species of *Eucalyptus*. I. RAPD data. *Theor Appl Genet* **94**:796–803.

BARRIL, C. P., VERHAEGEN, D., VIGNERON, P., BOUVET, J.M. & KREMER, A. 1997b: Stucture of the specific ability between two species of Eucalyptus. II. A clustering approach and a multiplicative model. *Theor. Appl. Genet.* **94**:804–809.

BECKERT, R.A., CHAMBERS, J.M. & WILKS, A.R. 1992: The new S Language: a Programming Environment for Data

Analysis and Graphics. Wadsworth and Brooks; Cole Advanced Books and Software, Pacific Grove, California.

BOUVET, J.-M. 1995: Evolution de la variabilité avec l'âge et corrélation juvénile-adulte dans des populations d'Eucalyptus. Thèse de l'Institut National Agronomique Paris-Grignon, France, 236 p.

BOUVET, J.–M. & VIGNERON, P. 1995: Age trends in variances and heritabilities in *Eucalyptus* factorial mating designs. *Silvae Genet.* **44**:206–216.

CROW, J.F. & KIMURA, M. 1970: An Introduction to Population Genetics. Harper and Row, New York.

GELDERMANN, H. 1975: Investigations on inheritance of quantitative characters in animals by gene markers. I. Methods. *Theor. Appl. Genet.* **70**:138–146.

GRATTAPAGLIA, D. & SEDEROFF, R. 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., BERTOLUCCI, F. L. G. & SEDEROFF, R. 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. Genet.* **90**:933–947.

HAMRICK, J. L., GODT, M. J. W. & SCHERMAN-BRAYLES.S. T. 1992: Factors influencing levels of genetic diversity in woody plant species. *New Forest* **6**:95–124.

HILL, W. G. & ROBERTSON, A. 1968: Linkage disequilibrium in finite populations. *Theor. Appl. Genet.* **38**:226–231.

LEONARDS-SCHIPPERS C., GIEFFERS W., SCHÄFER-PREGL R., RITTER, E., KNAPP, S. J., SALAMINI, F. & GEBHARDT, C. 1994: Quantitative resistance to *Phytophtora infestans* in potato: a case study for QTL mapping in an allogamous plant species. *Genetics* **137**:67–77.

LEWONTIN, R. C. 1974: The Genetic Basis of Evolutionary Change. Columbia Univ Press, New York.

MURANTY, H. 1996: Power of tests for quantitative trait loci detection using full-sib families in different schemes. *Heredity* **76**:156–165.

MURANTY, H. & GOFFINET, B. 1997: Selective genotyping for location for location and estimation of the effect of a quantitative trait locus. *Biometrics* **53**:629-643.

O'MALLEY, D. M. 1996: Complex trait dissection in forest trees using molecular markers. *In*: The impact of plant molecular genetics (BWS Sobral Ed). Birkhäuser, Boston, USA, p 49–70.

OHTA T. & KIMURA M. 1969 Linkage disequilibrium due to random genetic drift. *Genet. Res.* **13**:47–55.

PLOMION, C. & DUREL, C.-E. 1996: Estimation of the average effects of specific alleles detected by the pseudo-testcross QTL mapping strategy. *Genet. Sel. Evol.* **28**:223–235.

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

VERHAEGEN, D. & PLOMION, C. 1996: Genetic mapping in *Eucalyptus urophylla* and *E. grandis* using RAPD markers. *Genome* **39**:1051–1061.

VERHAEGEN, D., PLOMION, C., GION, J. M., POITEL, M., COSTA, P. & KREMER, A. 1997: Quantitative trait dissection analysis in *Eucalyptus* using RAPD markers: 1 – Detection of QTLs in an interspecific hybrid progeny, stability of QTL expression across different ages. *Theor. Appl. Genet.* **95**:597-608.

VIGNERON, P. 1991: Création et amélioration de variétés d'hybrides d'*Eucalyptus* au Congo. *In*: Proc IUFRO Symp "Intensive Forestry: the role of Eucalyptus" Durban, South Africa, p. 345–360.

WILLIAMS, J. G. K., KUBELIK, A. R., LIVAK, K. J. & RAFALSKI, J. A. 1990: DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nuc. Ac. Res.* **18**:6531–6535.