Theor Appl Genet (1998) 97: 816–827 

**R. P. V. Brondani · C. Brondani**
**R. Tarchini · D. Grattapaglia**

# Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*

Received: 16 March 1998 / Accepted: 22 March 1998

**Abstract** We report on the development, genetic characterization and linkage mapping of a battery of SSR (simple sequence repeat) loci in *Eucalyptus grandis* and *E. urophylla*. This study reveals the abundance of SSRs in *Eucalyptus*, the very high information content of these markers for mapping and individual identification, and demonstrates the feasibility of constructing a comprehensive microsatellite-based linkage map for *Eucalyptus*. Primer sequence for a set of 20 highly informative EMBRA (*Eucalyptus* microsatellites from Brazil) loci are made available together with their map position and estimates of the expected heterozygosity and allele size range in these two species. Using genomic library enrichment and anchored-PCR screening prior to sequencing, the efficiency of SSR marker locus development was 63% from sequencing data to operationally useful SSR loci. Absolute transportability between the two species and very high levels of allelic variability and expected heterozygosity (H) were seen at all SSR loci surveyed. The number of alleles per locus ranged from 9 to 26 with an average of 16.3 $\pm$ 4.8. The average H of 15 loci was 0.86 $\pm$ 0.04, 0.83 $\pm$ 0.08 and 0.89 $\pm$ 0.04, respectively, for *E. urophylla*, *E. grandis* and the combined two-species estimate. In the mapping analysis 16 out of 20 marker loci segregated in a fully informative configuration, allowing the determination of synteny of six homologous linkage groups between the two species. The availability of transportable, multiallelic, PCR-based co-dominant SSR loci represents a dramatic improvement in our ability to carry out detailed population genetic analysis and to search, understand, and manipulate allelic variation at QTLs (quantitative trait loci) in species of *Eucalyptus*.

**Key words** Simple sequence repeats (SSRs) · Microsatellite · Linkage map · *Eucalyptus grandis* · *E. urophylla*

Communicated by P. M. A. Tigerstedt

R. P. V. Brondani · C. Brondani · D. Grattapaglia (✉)
Laboratório de Genética de Plantas,
CENARGEN-EMBRAPA C.P. 02372 Brasília D.F. 70849-970, Brazil
and
Departamento de Biologia Celular
Universidade de Brasíla UnB-Brasília D.F., Brazil
Tel: +55-61-340 3515
Fax: +55-61-3403624
E-mail: dario@cenargen.embrapa.br

R. Tarchini
Biotechnology Research, Agricultural Products Department,
E.I. du Pont de Nemours and Co (Inc.) – POB 80402,
Experimental Station, Wilmington DE, 19880-0402, USA

## Introduction

The development of high-resolution linkage maps based on sets of highly informative simple sequence repeats (SSR) or microsatellite-based markers has revolutionized mammalian genetic analysis (e.g., Dib et al. 1996; Dietrich et al. 1996). The same trend is expected for plants, since the occurrence, properties and map position of SSR loci for several crop and tree species have been reported in a number of studies in recent years, both from library and electronic database screening (reviewed by Powell et al. 1996).

Microsatellites are very attractive to plant geneticists as they combine several features of the ultimate marker: (1) they are typically co-dominant and multiallelic with an expected heterozygosity frequently above 0.7; (2) they are highly polymorphic allowing precise discrimination even of closely related individuals; (3) they are abundant and uniformly dispersed in plant genomes; (4) they can be efficiently analyzed by a rapid and simple PCR assay; (5) SSR marker genotyping can be semi-automated in multiplex assays using fluorescence-labeled primers; and (6) marker information based on primer sequences, may be easily shared

between laboratories improving cooperative efforts in research and development. Furthermore, microsatellite markers are transportable across pedigrees, populations, and frequently across genetically related species of the same genus.

One of the first reports describing the presence of microsatellites in plant genomes was in forest trees (Conditt and Hubbell 1991). Since then, the occurrence and characteristics of nuclear microsatellites has been reported in several forest-tree species such as *Pinus radiata* (Smith and Devey 1994), *Quercus macrocarpa* (Dow et al. 1995), *Pinus sylvestris* (Kostia et al. 1995), *Picea sitchensis* (van de Ven 1996), *Pinus strobus* (Echt et al. 1996), *Pithecellobium elegans* (Chase et al. 1996) and *Swietenia humilis* (White and Powell 1997). All these studies, however, resulted in the development of only a small number of markers, typically less than ten. In *Eucalyptus nitens* the development of four microsatellite markers was reported, but surprisingly the primer sequences were not published (Byrne et al. 1996). With the exception of two SSR-based markers mapped in *Pinus radiata* (Devey et al. 1996), the generation of linkage information for SSR-based markers in trees has not been contemplated to-date.

The development of SSR-based markers has become increasingly accessible in recent years mainly due to novel library enrichment strategies and rapid fluorescence-based automatic sequencing technologies (reviewed by Powell et al. 1996). Particularly for the genetic analysis of highly heterozygous forest-tree species, the availability of linkage maps based on transportable PCR-based co-dominant multiallelic markers will become increasingly crucial as one moves from descriptive QTL mapping to effective use of QTL information for selective breeding. Applications of SSR-based markers in tree breeding will range from genotype identification and genetic management of breeding populations, QTL-allele tagging, comparative QTL mapping and gene-flow studies for transgenic tree deployment. We expect to use SSR-based markers not as the sole marker system, but rather in combination with RAPD and AFLP technologies depending on the objectives of the study.

As part of our program to incorporate genomic genetics in *Eucalyptus* breeding and germplasm management, we are currently developing a large battery of informative microsatellite markers. The objectives of the present study were: (1) to establish the presence, abundance and characteristics of two classes of dinucleotides repeats (AG/TC and AC/TG) in two *Eucalyptus* species; (2) to develop and characterize the genetic information content of an initial set of SSR marker loci based on such repeats; and (3) to map this set of SSR loci on existing RAPD framework linkage maps to demonstrate the feasibility of constructing a comprehensive reference microsatellite-based linkage map for *Eucalyptus*.

## Materials and methods

### Genetic material

Two selected genotypes of *E. grandis* (Coffs Harbour provenance) and *E. urophylla* (Timor provenance) from the germplasm collection of EMBRAPA were used to develop the libraries enriched for $(AG)_n$ and $(AC)_n$ dinucleotides. Genetic mapping of the SSR loci was carried out on a population of 94 $F_1$ individuals derived from a cross between *E. grandis* G44 and *E. urophylla* URO28. Genomic DNA was extracted from parents and progeny individuals as described earlier (Grattapaglia and Sederoff 1994).

### Southern analysis with SSR probes

DNA from a single individual of *E. grandis* was digested with five different restriction enzymes, *Hin*fI, *Mse*I, *Tsp*509I, *Rsa*I and *Sau*3AI, according to the manufacturer's instructions, electrophoresed in 1.5 % agarose gels and blotted onto a nylon membrane. Oligonucleotide probes, complementary to simple-sequence repeats of poly (dA-dC) and poly (dA-dG) were labeled with $^{32}P$. Membranes were pre-hybridized and hybridized to the probe in the same solution at the appropriate temperature, followed by stringent washes and exposure to autoradiography.

### SSR-enriched genomic libraries construction

Genomic library construction was carried out following previously developed and optimized protocols (Rafalski et al. 1996; Taramino and Tingey 1996; White and Powell 1997). Genomic DNA was digested with *Mse*I, and fragments in the range of 280–600 bp were recovered on a DEAE-cellulose/NA-45 membrane via electrophoresis. Fragments were ligated to adaptors at the *Mse*I restriction site. Positive clones containing SSRs were selected by hybridization with biotinilated-oligonucleotides complementary to the repetitive sequence AG/CT or AC/GT, and recovered by magnetic beads linked to streptavidine. Fragments were amplified by PCR and cloned into phage Lambda ZapII (Stratagene).

### Selection and sequencing of positive clones and primer design

Lambda Zap libraries were screened by lifting plaques onto nylon membranes and hybridizing them to random-primed $^{32}P$-labeled poly (dA-dC) or poly (dA-dG) probes, respectively, for $(AC)_n$ or $(AG)_n$ repeats. Plaques containing the respective SSR were identified by autoradiography. Positive clones were transferred to 96-well plates for plasmid excision. The presence of the SSR and its position within the cloned insert was determined by an anchored-PCR screening strategy (Taylor et al. 1992; Rafalski et al. 1996). Subsequent agarose-gel analysis revealed clones containing SSR inserts and the direction within the vector from which they were to be sequenced. After in vivo excision of the pBluescript plasmid from Lambda Zap and the preparation of plasmid DNA (Wizard Minipreps, Promega), double-stranded miniprep DNAs were sequenced on an Applied Biosystems 377 instrument using dye-terminator fluorescent chemistry (Applied Biosystems Incorporated). Primer pairs complementary to sequences flanking the repeat element were designed using a modification (M. Hanafey, unpublished software) of the Primer program (Lincoln et al. 1991). To reduce problems with spurious banding patterns generated during amplification, to allow the use of all primer pairs under the same PCR conditions, and to develop single-reaction multiplex systems, two stringent criteria in primer design were employed: (1) a primer Tm range from 52°C to

58°C; and (2) a maximum of 3°C difference in Tm between primers in a pair. Primers were synthesized by Operon Technologies Incorporated (Alameda, California).

### PCR amplification of SSR loci

Microsatellite-marker amplification was performed in 96-well V bottom plates in a 13-µl reaction volume containing 0.3 µM of each primer, 1 unit of *Taq* DNA polymerase, 0.2 mM of each dNTP, 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, DMSO (50%), and 7.5 ng of template DNA. Amplifications were performed using a MJ Research PT-100 thermal controller with the following conditions: 96°C for 2 min, then 29 cycles of 94°C for 1 min, 56°C for 1 min and 72°C for 1 min; and a final elongation step at 72°C for 7 min. Inheritance, segregation and mapping was carried out in a 3.5% Metaphor agarose (FMC Bioproducts) gel containing 0.1 µg/ml of ethidium bromide in 1 × TBE buffer (89 mM Tris-borate, 2 mM EDTA pH 8.3). Submarine electrophoresis was carried out at 120 V for 2 h in custom-made gel boxes that contained 96 samples per gel.

### Inheritance, segregation and mapping of SSR marker loci

A progeny set of 94 $F_1$ individuals used earlier to construct linkage maps based on RAPD markers (Grattapaglia and Sederoff 1994) was employed for inheritance, segregation and mapping analyses of the SSR loci. SSR marker loci were identified by the acronym EMBRA, which stands for *Eucalyptus* microsatellite from Brazil, followed by a sequential number. SSR markers were localized on existing framework linkage groups of RAPD markers constructed at a likelihood support of 1000 : 1. Multipoint linkage analysis was carried out using MapMaker™ (Lander et al. 1987) version 2.0 for Mac. As the RAPD framework maps were constructed under a pseudo-testcross design, linkage analysis of the SSR markers was done for each parental data set separately. Since the MapMaker™ model assumes that all markers are in coupling phase it does not recognize linkages for markers in repulsion. To allow for the detection of linkage of SSR markers to RAPD markers in any linkage phase, the data set to be entered into MapMaker™ had to be duplicated and re-coded. Therefore, for each SSR locus the two alleles from a heterozygous parent were coded 1 and 2 and vice versa. A co-segregation analysis of these data was then carried out with the RAPD-marker data set for that same parent. To map the same SSR locus on the second parental linkage map, the other two alleles from the second heterozygous parent were again coded 1 and 2, and vice versa, and a co-segregation analysis of these data was then carried out with the RAPD-marker data set for the second parent. Thus, when the same SSR locus was identified in both genetic maps it allowed the prompt alignment of linkage groups of the two *Eucalyptus* species. Markers could be mapped on both parental maps when they were in a fully informative configuration, i.e., both parents heterozygous and at least three different alleles segregating in the mating configuration so that four different progeny genotypes could be identified. When only one of the parents was heterozygous, the SSR marker could only be mapped on one of the maps, namely the one for the heterozygous parent.

### Polymorphism analysis of SSR loci

A subset of 15 SSR loci was characterized for genetic information content by genotyping 32 randomly chosen adult individual trees from a germplasm bank with no history of selection, 16 from *E. grandis* and 16 from *E. urophylla*. PCR products were separated on 6% denaturing polyacrylamide gels in 7 M urea and 1 × TBE buffer, visualized by silver staining (Bassam et al. 1991) and sized by comparison to a 10-bp ladder (BRL) standard. The number and frequency of alleles observed in each species, the observed heterozygosity and the expected heterozygosity estimates (H), defined as $H = 1 - \Sigma(p_1)^2$, where $p_1$ is the frequency of the $i^{th}$ allele in the population sample studied, were determined for the group of 15 loci characterized. Average expected heterozygosity as defined by $H_{av} = \Sigma H_n/n$, where $n$ is the total number of loci, was also estimated. Separate analyses were carried out for the two *Eucalyptus* species and estimates were also obtained for the two species combined.

## Results

### Abundance of microsatellite in the *Eucalyptus* genome

Digestion of *E. grandis* DNA with different restriction enzymes revealed that *Mse*I produced the most adequate digestion profile for library construction, with a range of fragments between 200 and 800 bp. Genome Southern screening with poly (dA-dC) and poly (dA-dG) dinucleotide repeat probes revealed that in *Eucalyptus*, as in several other plant species surveyed to-date, AG repeats are more frequent than AC repeats. Furthermore, both $(AG)_n$ and $(AC)_n$ repeats appear to be well dispersed throughout the genome as an intense hybridization signal was seen in a wide range of fragment size.

### SSR-enriched libraries

A total of seven libraries were constructed, five enriched for AG, three for *E. grandis* and two for *E. urophylla*, and two for AC, one for each species. Libraries had an average insert size of 500 bp, therefore facilitating later single-pass sequencing. Plaque hybridizations showed that the enrichment step yielded between 10 and 40% of clones containing a SSR, the percentage varying with the library. Out of 960 positive clones screened by anchored-PCR, 714 had a SSR and of these, 490 had an SSR of adequate size and position in the cloned insert. In other words the efficiency was 74% for the recovery of clones containing a SSR and 51% for the recovery of clones with a SSR that was positioned far away from both cloning sites so that enough DNA on both sides of the SSR would be later available for primer design. A total of 207 anchored-PCR selected $(AG)_n$ clones were sequenced, resulting in 180 useful sequences (87%) for designing primers. The remainder included 12 (6%) redundant sequences and 15 sequences that would only allow the design of one primer requiring reverse sequencing (7%) to become usable. No primers were designed for $(AC)_n$ clones.

### Sequence characterization of SSRs

SSRs can be classified as perfect, imperfect or compound (Weber 1990). Sequencing of clones containing

$(AG)_n$ repeats showed a distinct predominance (81%) of perfect repeats (no interruption in the dinucleotide repeat sequence). Clones containing $(AC)_n$ showed approximately the same proportion of perfect and compound repeats. The compound sequences consisted of classes of di, tri- and tetra-nucleotide repeats alternating with $(AC)_n$ repeats. Imperfect repeat sequences were relatively rare (around 15%) for both types of repeat (Fig. 1 A). Microsatellite length was defined by the longer run of uninterrupted repeats. Although a greater number of $(AG)_n$-containing clones (207) were sequenced when compared to the number of $(AC)_n$ clones (80), the majority of sequences, both for AC and AG repeats, had a length range between 11 and 30 dinucleotide repeat units. The $(AG)_n$ repeats, however, were more evenly distributed in size with a length ranging from 6 to 45 units (Fig. 1B). The average number of dinucleotide repeats was relatively similar for the two motifs, $23.3 \pm 7.1$ for $(AG)_n$ and $19 \pm 4.8$ for $(AC)_n$.

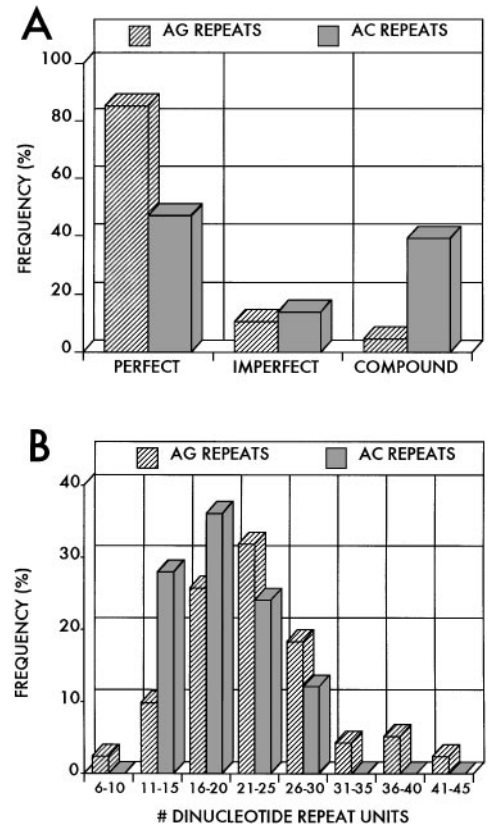


**Fig. 1A, B** Sequence characteristics of AC and AG SSRs in *Eucalyptus*. **A** Frequency distribution of different classes of repeats; **B** length distribution in the number of dinucleotide repeat units. Results based on sequencing data from 207 AG clones and 80 AC clones

## Primer screening, inheritance and mapping of SSR loci

Using a single PCR program, from the 180 pairs of primers designed and tested, 63 (35%) pairs amplified very clean and easily interpretable PCR products; 67 (37%) amplified clearly interpretable PCR products but also showed some non-specific amplification of secondary bands; and 28% either did not amplify any product or resulted in a smear of bands that could not be interpreted. For this last group of primer pairs further optimization of the reaction conditions could be pursued. From the 63 primer pairs that amplified clearly interpretable products, 20 were selected at random to verify the inheritance and segregation of the locus in the $F_1$ progeny set and map their position (Table 1). These pairs amplified markers that showed Mendelian inheritance, segregated in the mapping population, and could easily be interpreted in Metaphor agarose gels (Fig. 2). Of the 20 SSR loci tested, 19 of them were heterozygous and therefore segregated in the female parent (*E. grandis*), 17 were heterozygous and segregated in *E. urophylla*, and 16 segregated in both parents. The observed segregation data in agarose was checked in polyacrylamide gels, confirming the non-informative homozygous genotypes for locus EMBRA7 in *E. grandis* and loci EMBRA2, EMBRA8 and EMBRA20 in *E. urophylla.*

## Linkage mapping of SSR marker loci

Although parents were homozygous for some loci, all 20 randomly chosen SSR marker loci were polymorphic between the two parents and there was no locus where both parents were homozygous. Therefore all 20 SSR loci could be readily mapped onto existing RAPD linkage maps for the two parents constructed earlier by a pseudo-testcross strategy (Grattapaglia and Sederoff 1994). With the exception of locus EMBRA9, with a LOD score of 2.44, LOD scores for map position were high and similar for the *E. grandis* and *E. urophylla* maps (Fig. 3). Sixteen of these marker loci (80%) were heterozygous in both parents and segregated in a fully informative fashion with three or four different and clearly interpretable segregating alleles in agarose gels, generating four distinguishable genotypes in the $F_1$ progeny (Fig. 2). Because of the mating configuration, these marker loci were mapped on both parental maps allowing the determination of synteny between homologous linkage groups in the two species (Fig. 3). Of the remaining four loci, one of them was mapped only on the *E. urophylla* map (EMBRA7) and three only on the *E. grandis* map (EMBRA2, EMBRA8 and EMBRA20). With no exception, SSR loci that mapped on the same linkage group in one species also did so in the other species, further validating the fact that such loci are homologous SSR loci in the two species. SSR marker loci were placed on a total of 9 out of the 11 existing linkage groups ($n = 11$ in *Eucalyptus*). Locus order was the same in both species maps with

**Table 1** SSR locus denomination, type of repeat, forward and reverse primer sequences, primer annealing temperatures, and the expected allele size from the sequencing data for 20 SSR loci in *Eucalyptus*

| SSR locus | Repeat | Primer sequences | Annealing temp. (°C) | Size (bp) |
|---|---|---|---|---|
| EMBRA1 | $(AG)_{33}$ | 5′-gAT AgA ACT TTC CTA TTT gAT Cg-3′<br>5′-gTA ggA TTT gAT gTC TgC AA-3′ | 56 | 127 |
| EMBRA2 | $(AG)_{15}$ | 5′- CgT gAC ACC Agg ACA TTA C-3′<br>5′-ACA AAT gCA AAT TCA AAT gA-3′ | 56 | 121 |
| EMBRA3 | $(AG)_{19}$ | 5′-gAT Cgg ATT ggA ggA gAC-3′<br>5′-AAT TCA ATT CAT CCA AAg C-3′ | 56 | 123 |
| EMBRA4 | $(AG)_{23}$ | 5′-ATA CAA TgA TTT gAA Agg gg-3′<br>5′-gAg TTg TTT gTT TTg TCg AA-3′ | 56 | 64 |
| EMBRA5 | $(AG)_{22}$ | 5′-ATg CTg gTC CAA CTA AgA TT-3′<br>5′-TgA gCC TAA AAg CCC AAC-3′ | 56 | 88 |
| EMBRA6 | $(AG)_{19}$ | 5′-AgA gAA TTg CTC TTC ATg gA-3′<br>5′-gAA AAg TCT gCA AAg TCT gC-3′ | 56 | 98 |
| EMBRA7 | $(AG)_{15}$ | 5′-CAC ACC gTg TCA gTT AgC-3′<br>5′-AAT AAg gAg gAT TCC ATg g-3′ | 56 | 115 |
| EMBRA8 | $(AG)_{21}$ | 5′-CAC AAC TAA AAA TCA AAA CCC-3′<br>5′-AAA gAg CAg ATT ATT ACA gAA gC-3′ | 56 | 127 |
| EMBRA9 | $(AGA)_3(AG)_{28}$ | 5′-AgT gAg AgA gAT ATT CgC gT-3′<br>5′-CCA ATA CAA TCA TCA ATC CA-3′ | 56 | 94 |
| EMBRA10 | $(CCT)_3(AG)_{14}$ | 5′-gTA AAg ACA TAg TgA AgA CATTCC-3′<br>5′-AgA CAg TAC gTT CTC TAg CTC-3′ | 56 | 95 |
| EMBRA11 | $(AG)_4GG(AG)_{13}$ | 5′-gCT TAg AAT TTg CCT AAA CC-3′<br>5′-gTA AAA TCC ATg ggC AAg-3′ | 56 | 97 |
| EMBRA12 | $(AG)_{22}$ | 5′-Agg ATT TgT ggg gCA AgT-3′<br>5′-gTT CCC CAT TTT CAT gTC C-3′ | 56 | 98 |
| EMBRA13 | $(AG)_{27}$ | 5′-ATT TCC CTA ggT TTg ACA Tg-3′<br>5′-TCC AAC ATC TTA CTC AAC CA-3′ | 56 | 130 |
| EMBRA14 | $(AG)_8AAC(AG)_{25}$ | 5′-gCC TCA AAC CAA TTC AAA T-3′<br>5′-CAT gAT TCA TCC CAC TCC TC-3′ | 56 | 109 |
| EMBRA15 | $(AG)_{21}$ | 5′-TTT gTT ggA TgA ggA CTT-3′<br>5′-CAA CAT gTT CTC CgA AAA g-3′ | 56 | 66 |
| EMBRA16 | $(AG)_{21}$ | 5′-CAA CgT TCC CCT TTC TTC-3′<br>5′-ATg TTA ggC CAA ACC CAg-3′ | 56 | 98 |
| EMBRA17 | $(AG)_{18}$ | 5′-Agg ATA CTC gTg AgA gAA gC-3′<br>5′-gTA gAT CTg TTC TgC ATg TTg-3′ | 56 | 184 |
| EMBRA18 | $(AG)_3GG(AG)_{19}$ | 5′-CAg CTA ggA TgT TAg ACT Tgg-3′<br>5′-gCA CAC CTA gAA TTT TCA AAC TA-3′ | 56 | 87 |
| EMBRA19 | $(AG)_{23}$ | 5′-gAC ggT TgA TTT CCT gAT T-3′<br>5′-gTg gTg CTC CTC TCC TCT-3′ | 56 | 124 |
| EMBRA20 | $(AG)_{19}$ | 5′-gTg AgT ggg TAT CCA TCg-3′<br>5′-gCT ggA ACT ggT CTT gAg-3′ | 56 | 97 |

the exception of two of the 16 loci. Linkage-group numbering followed the number established earlier for *E. grandis* (Grattapaglia et al. 1995), as this species will be used as the reference species for the construction of a comprehensive linkage map.

## Polymorphism analysis of SSR loci

Absolute transportability of the SSR loci between the two species and very high levels of allelic variation were observed (Fig. 4). A similar mean number of alleles at a SSR locus was found in the two species, 11.8 for *E. urophylla* and 11 for *E. grandis*. On average 40% of the alleles were shared by the two species. Totalling across species, the number of alleles varied from a minimum of nine to a maximum of 26 with a mean of $16.3 \pm 4.8$. The allele sizes varied between 65 and 193 bp. Observed heterozygosity values were in general lower than the expected values estimated from allele frequencies. The average observed heterozygosity of 15 loci was $0.59 \pm 0.14$, $0.56 \pm 0.13$ and $0.57 \pm 0.12$, respectively, for *E. urophylla*, *E. grandis* and the combined two-species estimate (Table 2). Frequencies of all alleles at all loci were less than 0.3 with a relatively homogeneous distribution (data not shown but available upon request) resulting in very high expected heterozygosities. In *E. urophylla* all but one locus, and in *E. grandis* all but three loci, had values above 0.8. Combined estimates for the two species yielded nine loci with a heterozygosity above 0.9. The average expected heterozygosity of 15 loci was $0.86 \pm 0.04$, $0.83 \pm 0.08$ and

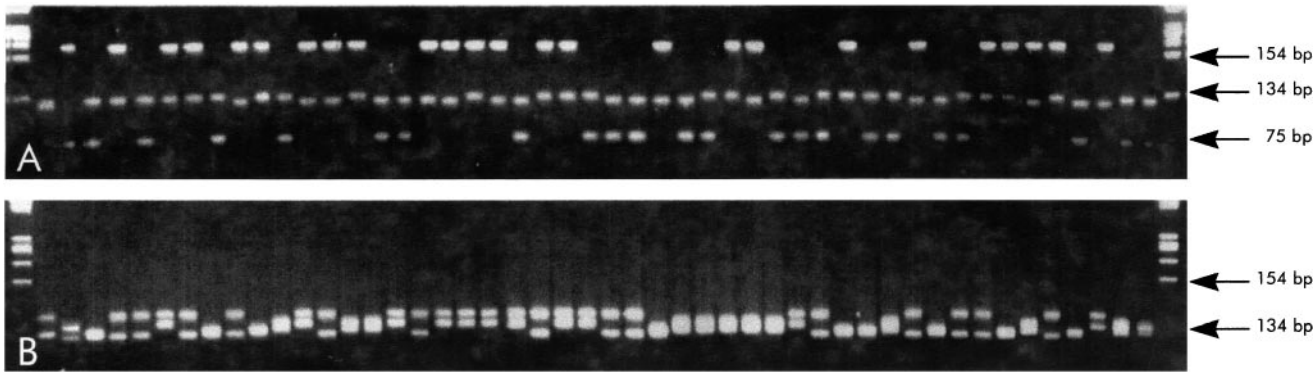


**Fig. 2A, B** Inheritance and segregation of fully informative SSR loci in *Eucalyptus*. Agarose-gel resolution and detection by ethidium-bromide staining of SSR loci EMBRA4 (panel **A**) and EMBRA 10 (panel **B**): *lanes 1 and 50* are 1 Kb ladder (BRL) size standards with the sizes of some fragments indicated in base pairs; *lanes 2 and 3* are the two parents, *E. grandis* and *E. urophylla*, followed by 46 $F_1$ progeny individuals

$0.89 \pm 0.04$, respectively, for *E. urophylla*, *E. grandis* and the combined two-species estimate (Table 2).

## Discussion

This study reveals the abundance of SSRs in the *Eucalyptus* genome, determines the very high information content of these markers for mapping and individual identification in two species of the genus, and demonstrates the feasibility of constructing a comprehensive microsatellite-based linkage map for *Eucalyptus*. Primer sequences for a set of 20 highly informative SSR marker loci are made available together with their map position and estimates of expected heterozygosity in two of the most-widely planted eucalypt species in the world.

### Characteristics of microsatellites in *Eucalyptus*

The frequency of each class of SSR is highly variable among plant species (Wang et al. 1994). Our results indicate that both AG and AC repeats appear to be well-dispersed throughout the genome of *E. grandis* and *E. uraphylla*, although AG repeats were seen to be more abundant than AC repeats. Based on a limited sample of 13 SSR-containing clones, Byrne et al. (1996) also reported that AG repeats were more abundant in *E. nitens*. These results demonstrate that in *Eucalyptus*, as in the majority of other plant species surveyed to-date (reviewed by Powell et al. 1996), AG repeats are more abundant than AC repeats, as distinct from the reports showing that AC repeats are more frequent in mammalian genomes (e.g., Weissenbach et al. 1992). In contrast to *Eucalyptus*, a recent survey of dinucleotide repeat frequency in pine trees revealed a greater abundance of AC than AG repeats (Echt and May-Marquardt 1997). No precise estimate of the frequency of AG and AC SSRs in the *Eucalyptus* genome can be made from our experimental work as the development of the microsatellite loci was based on highly enriched libraries. The high efficiency of recovery of SSR-containing clones from genomic libraries constructed with different restriction enzymes and the occurrence of two different classes of repeats suggests, however, that a large number of dinucleotide repeats is available in the *Eucalyptus* genome for marker development.

The average number of AG repeats and the predominance of perfect repeats for this motif in *Eucalyptus* was very similar to that found in rice by Chen et al. (1997). For the AC repeats, an almost equal proportion of perfect and compound or imperfect repeats was present, and the number of repeated motifs was on average four-times higher for AG sequences than for AC in *Eucalyptus*. Similar results were reported in corn, where despite differences in the structure of the AG and AC repeats, the number of alleles and the expected heterozygosity were not significantly different between the two motifs (Taramino and Tingey 1996). In this study, no primer pairs were tested for AC microsatellite markers, therefore no direct comparison of the level of polymorphism can be made between AG and AC microsatellite loci. No positive relationship was seen in *Eucalyptus* between the number of tandemly repeated AG motifs and the level of polymorphism for this first set of 20 SSR loci evaluated in a sample of 32 individuals. Although studies in humans (Weber 1990), mice (Love et al. 1990) and plants (Bell and Ecker 1994; Saghai-Maroof et al. 1994) revealed the existence of a similar relationship, other reports did not find that such a correlation existed in humans (Valdes et al. 1993) and *Brassica* (Szwec-McFadden et al. 1996), or else described significant exceptions to this rule in corn (Taramino and Tingey 1996). The development of a larger number of SSR loci based both on AG and AC repeats and the evaluation of allelic diversity in larger

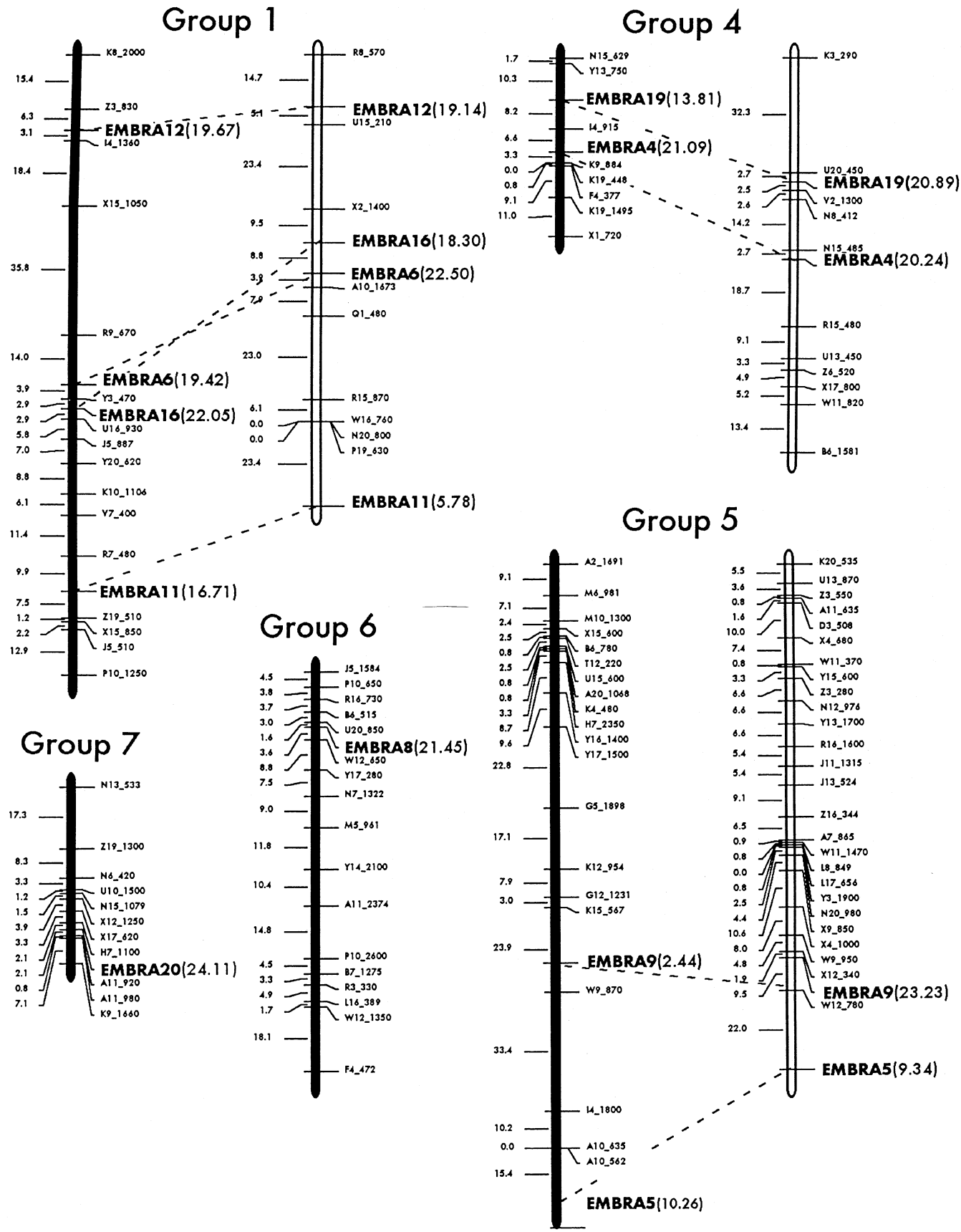


**Fig. 3** For legend see page 823

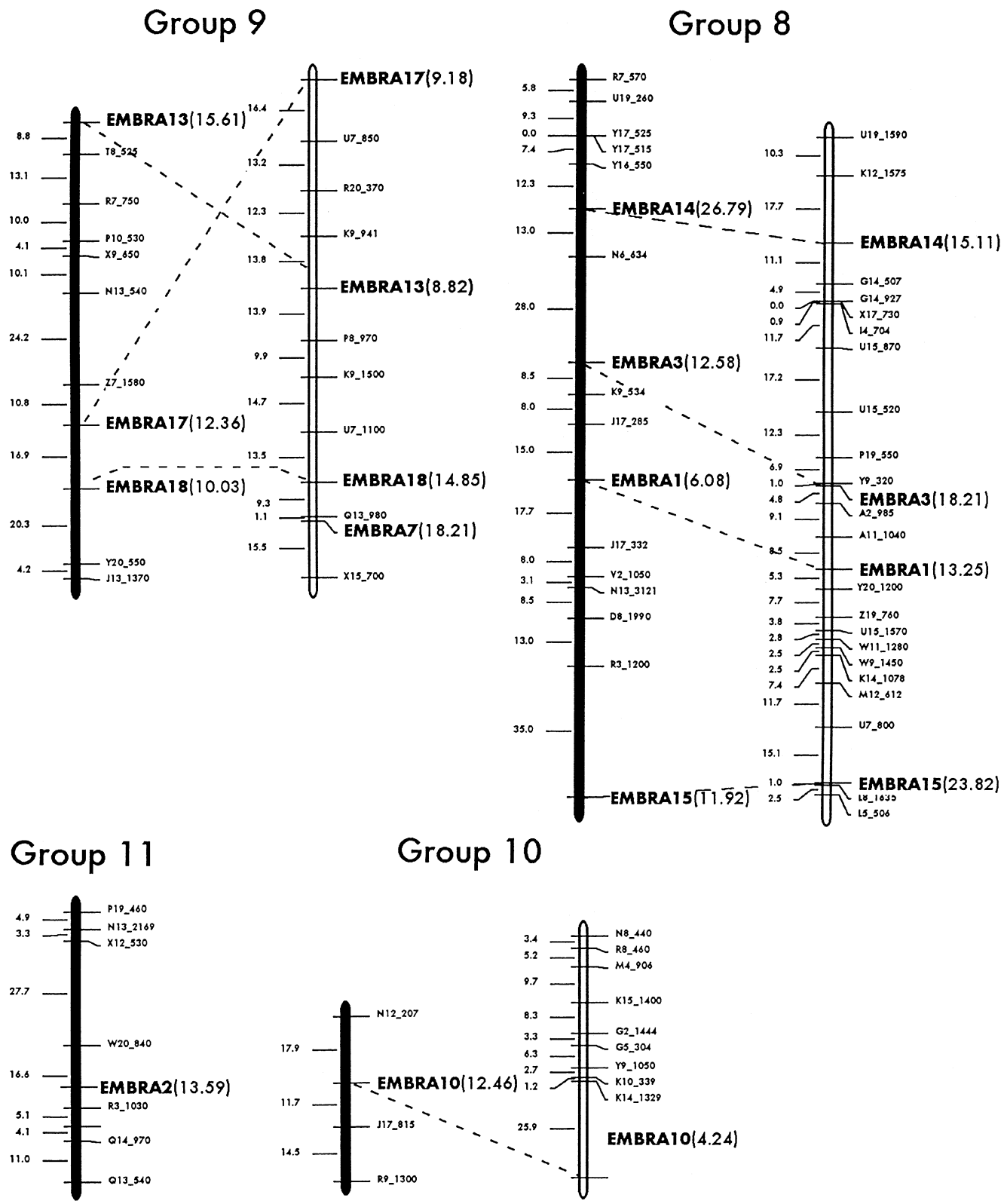


**Fig. 3** Linkage mapping of SSR markers on RAPD maps. Twenty SSR loci have been mapped on RAPD framework linkage groups of *E. grandis* and *E. urophylla*. Sixteen of them were mapped on both species maps, allowing the determination of synteny of six homologous linkage groups between the two species. *E. grandis*: *black-filled* linkage groups; *E. urophylla*: *white-filled* linkage groups; LOD scores are in parenthesis after SSR locus denomination

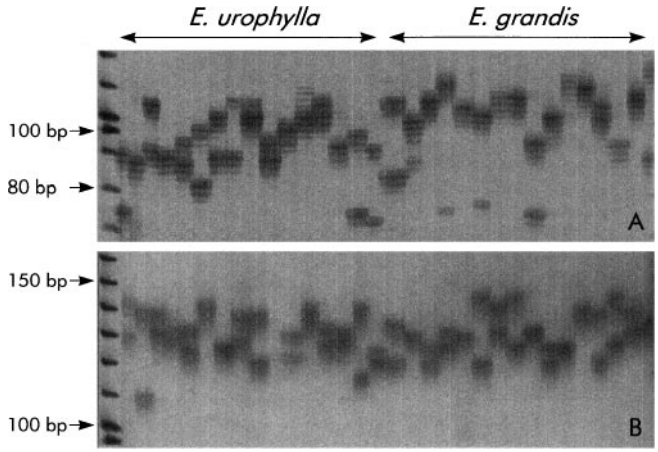


**Fig. 4A, B** Allelic variation at *Eucalyptus* SSR loci. Silver-stained denaturing polyacrylamide-gel images of SSR loci EMBRA4 (panel **A**) and EMBRA 12 (panel **B**) analyzed in 16 individuals of *E. urophylla* and 16 of *E. grandis*. The *first lane* is a 10-bp ladder (BRL) size standard with the sizes of some fragments indicated in base pairs

population sizes is needed to better elucidate these issues in *Eucalyptus*.

## Efficiency of SSR marker development in *Eucalyptus*

We have used a very efficient method to generate highly polymorphic microsatellite-based markers. Following the enrichment step, between 10 and 40% of the cloned fragments contained a SSR. Of these, 51% were selected for sequencing based on anchored-PCR screening step, yielding 87% of clones with non-redundant sequences that allowed primer design. Seventy two percent of the primer pairs yielded clearly interpretable PCR products under a single set of PCR conditions, and 100% of the tested primer pairs yielded informative loci in the pedigree analyzed. The overall efficiency, considering that an average of 20% of the Lambda clones contained a SSR, was therefore 8.8% ($0.2 \times 0.51 \times 0.87$) of the total recombinant Lambda genomic clones yielding primer sequences and 6.3% ($0.2 \times 0.51 \times 0.87 \times 0.72$) yielding informative SSR loci. Using non-enriched libraries and colony hybridization screening the reported efficiencies in microsatellite development were respectively 0.21% of total recombinant clones yielding polymorphic markers in humans (Weissenbach et al. 1992), 0.24% of total recombinant clones yielding primer sequences in swine (Roher et al. 1994), and 0.12% yielding informative SSR loci in bovine (Bishop et al. 1994). Our efficiency was approximately 37-times higher for the yield of primer sequences and 30–50-times higher for the yield of informative loci. Evidently as the number of SSR loci which are developed increases this efficiency will tend to come down as a result of an increased number of redundant sequences. This can be overcome, however, by the

**Table 2** Map position, allele size range, allele numbers, and observed and expected heterozygosities of 15 SSR loci in *E. grandis* and *E. urophylla*

| SSR Locus | Linkage group | Allele size range (bp) | # Alleles *E. urophylla* | # Alleles *E. grandis* | # Shared alleles | Total # alleles | $H_{obs.}$ *E. urophylla* | $H_{exp.}$ *E. urophylla* | $H_{obs.}$ *E. grandis* | $H_{exp.}$ *E. grandis* | $H_{obs.}$ combined | $H_{exp.}$ combined |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| EMBRA1 | 8 | 100–145 | 14 | 12 | 11 | 15 | 0.62 | 0.91 | 0.6 | 0.89 | 0.61 | 0.92 |
| EMBRA3 | 8 | 115–193 | 11 | 15 | 9 | 17 | 0.53 | 0.90 | 0.46 | 0.90 | 0.5 | 0.92 |
| EMBRA4 | 4 | 72–178 | 13 | 17 | 8 | 22 | 0.53 | 0.90 | 0.37 | 0.92 | 0.45 | 0.93 |
| EMBRA5 | 5 | 115–160 | 18 | 16 | 8 | 26 | 0.8 | 0.93 | 0.66 | 0.93 | 0.73 | 0.95 |
| EMBRA6 | 1 | 120–170 | 10 | 11 | 5 | 16 | 0.5 | 0.87 | 0.56 | 0.84 | 0.53 | 0.90 |
| EMBRA10 | 10 | 110–155 | 12 | 10 | 4 | 18 | 0.68 | 0.85 | 0.37 | 0.87 | 0.52 | 0.91 |
| EMBRA11 | 1 | 123–165 | 14 | 9 | 5 | 18 | 0.46 | 0.89 | 0.46 | 0.79 | 0.46 | 0.91 |
| EMBRA12 | 1 | 105–145 | 12 | 11 | 7 | 16 | 0.8 | 0.88 | 0.62 | 0.87 | 0.70 | 0.91 |
| EMBRA13 | 9 | 75–120 | 10 | 7 | 6 | 11 | 0.87 | 0.86 | 0.81 | 0.60 | 0.84 | 0.78 |
| EMBRA14 | 8 | 120–150 | 8 | 7 | 6 | 9 | 0.53 | 0.85 | 0.53 | 0.80 | 0.53 | 0.87 |
| EMBRA15 | 8 | 90–125 | 16 | 14 | 6 | 24 | 0.62 | 0.92 | 0.73 | 0.91 | 0.67 | 0.94 |
| EMBRA16 | 1 | 110–165 | 14 | 9 | 8 | 15 | 0.56 | 0.87 | 0.62 | 0.81 | 0.59 | 0.86 |
| EMBRA17 | 9 | 120–170 | 5 | 10 | 4 | 11 | 0.33 | 0.74 | 0.35 | 0.86 | 0.34 | 0.87 |
| EMBRA18 | 9 | 70–110 | 10 | 10 | 7 | 13 | 0.43 | 0.81 | 0.5 | 0.79 | 0.46 | 0.85 |
| EMBRA19 | 4 | 65–145 | 10 | 7 | 4 | 13 | 0.64 | 0.81 | 0.76 | 0.77 | 0.70 | 0.84 |
| | | Mean ± SE | 11.8 ± 3.2 | 11 ± 3.2 | 6.5 ± 2.0 | 16.3 ± 4.8 | 0.59 ± 0.14 | 0.86 ± 0.04 | 0.56 ± 0.13 | 0.83 ± 0.08 | 0.57 ± 0.12 | 0.89 ± 0.04 |

construction of libraries using different restriction enzymes which would provide largely independent sources of microsatellite markers (Chen et al. 1997).

Few estimates of the efficiency of SSR development in plants are available in the literature. In *Brassica* out of 141 clones sequenced, primer pairs could be designed for 21 (15%) and, of these, 13 amplified polymorphisms in a test array, i.e., an overall efficiency of 9.2% from sequencing to the production of operational SSR loci (Szewc-McFadden et al. 1996). In rice, where libraries were not enriched but an anchored-PCR presequencing screen was used, 47% of the sequenced AG clones generated discrete PCR products, i.e., approximately one out of two sequences yielded a SSR locus (Chen et al. 1997). In this work, the efficiency obtained from sequencing data to the formation of operationally useful SSR loci was 63% ($0.87 \times 0.72$), i.e., almost two out of three sequences yielded a polymorphic SSR locus with Mendelian inheritance. Optimization of PCR conditions could be used to increase this efficiency; however, we believe that it is preferable to use the same PCR conditions and either sequence new clones or design new sets of primers to allow high-throughput genotyping.

## Variability at SSR loci for *Eucalyptus* genotype identification

Very high levels of allelic variability were seen at all 15 microsatellite loci surveyed, inspite of the relatively limited sample size ($n = 32$) used to derive estimates of the number of alleles. The smallest number of alleles seen in a sample of 16 individuals of *E. urophylla* was five. However, the relatively homogeneous distribution of allele frequencies still resulted in an expected heterozygosity of 0.74. The range of variation in the number of alleles found at 15 loci in a single species (5–18 in *E. urophylla* and 7–17 in *E. grandis*) are very similar to the range observed, 5–16, in a sample of 20 individuals of *E. nitens* at four loci (Byrne et al. 1996). The range and mean number of alleles found in *Eucalyptus* is on average two to three times larger than that found for several crop plants, including *Brassica* (Szwec-Mcfadden et al. 1996), tomato, rice, wheat and maize (reviewed by Powell et al. 1996). Allelic variability for *Eucalyptus* is close to that seen for soybean and barley which in several reports have shown a very high number of alleles at SSR loci with a range of 11 to 26 for soybean and up to 37 alleles for barley (Morgante et al. 1994; Saghai-Maroof et al. 1994).

Observed heterozygosities at the loci surveyed were in general lower than those expected. This was also the case in *E. nitens* (Byrne et al. 1996). Inbreeding due to a certain proportion of selfing in *Eucalyptus* could potentially play a role in the reduced observed heterozygosity. This is typical in *Eucalyptus* mating-system studies when seedlings are used (House and Bell 1994) because selfed individuals survive at younger ages but are then gradually eliminated by inbreeding depression (Griffin and Cotterill 1989). In our study, however, this explanation is unlikely because the 32 trees surveyed were genetically unrelated adult individuals sampled from a large germplasm bank with no history of selection. The difference between observed and expected heterozygosities is more likely to be due to the limited sample size used to estimate allele frequencies, which would prevent the observation of several classes of heterozygotes, particularly those involving rare alleles.

Due to the very high observed heterozygosity, all 32 individual trees could be uniquely fingerprinted with a single SSR locus, namely EMBRA5, the one that showed the largest number of alleles and the highest expected heterozygosity. The high degree of multiallelism and the very clear and simple co-dominant Mendelian inheritance of the set of microsatellites developed here provide an extremely powerful system for the unique identification of *Eucalyptus* individuals for fingerprinting purposes and parentage testing. We are currently characterizing a subset of these 15 loci in a very large sample sizes of 200 individuals where we have estimated probabilities of identity in the order of $10^{-10}$ for as few as six loci combined (Kirst et al., in preparation). Furthermore, we are optimizing multiplexed, multicolor fluorescent detection of alleles combined with automated sizing with internal size standards to increase precision in allele-size calling. This high-throughput approach is routinely employed in linkage mapping and parentage identification in humans (e.g., Ziegle et al. 1992) and has proven extremely powerful for *Brassica* (Kresovich et al. 1995) and soybean identification (Diwan and Cregan 1997).

The battery of SSR marker loci developed in this work should be very useful to the vast majority of *Eucalyptus* researchers. Our results show an absolute transportability of the SSR loci from *E. grandis* to *E. urophylla* and vice versa. Of the 20 SSR loci, 12 were derived from *E. grandis* and eight from *E. urophylla* genomic libraries. These two species belong to the same subgenus (*Symphyomyrtus*) that includes practically all of the species used in intensive eucalypt-based production forestry both in tropical and temperate regions of the world. Conservation of four SSR loci across species within *Symphyomyrtus* was also found by Byrne et al. (1996) between *E. nitens* and three other species, *E. grandis* being one of them. The very high rate of SSR marker-loci conservation between the most important *Eucalyptus* species in the world is a definite advantage. It will allow rapid spreading of the use of these markers in *Eucalyptus* genetics research, facilitating information exchange and a comparison among different laboratories particularly in respect of genetic mapping data (see below).

There are several envisaged applications of a set of highly discriminating SSR marker loci in *Eucalyptus* population genetics, breeding, and operational forestry.

Precise multilocus estimates of outcrossing rates in natural populations, seed orchards, and breeding groups could easily be generated with a small battery of such SSR loci. The build-up of inbreeding could be precisely monitored in incomplete pedigreed breeding populations. Quality control of controlled crosses could be quickly carried out, and the high power of exclusion provided by SSR markers would allow paternity identification of selected half-sib progenies produced by open pollination or by pollen mix. This could be a very useful tool for breeding programs conducted mainly by open pollination to reconstruct specific superior individuals or full-sib families by controlled crosses. The high probability of finding rare alleles in particular individuals would allow for the accurate estimation of the rate and distance of gene flow in forest stands. This information would provide fundamental subsidies for the establishment of forestry management procedures to prevent the escape of transgenes in transgenic *Eucalyptus*.

## Towards a comprehensive linkage map of SSR loci for *Eucalyptus*

All 20 SSR loci were genetically informative between the two parents of the cross and could therefore be mapped with very high confidence on a framework map of RAPD markers. A remarkable results was that 16 out of 20 markers segregated in fully informative fashion, i.e, with at least three different alleles segregating in the mapping population, allowing the determination of synteny of six homologous linkage groups of *E. grandis* and *E. urophylla*. Integration of the five remaining linkage groups should follow with the analysis of further SSR loci. Byrne et al. (1996) reported that the average number of alleles detected at SSR marker loci is nearly twice the average number found at RFLP loci in *Eucalyptus*. Not surprisingly, the proportion of fully informative SSR loci found in this study (80%) is significantly higher than the proportion of fully informative RFLP markers found segregating in radiata pine (19%) (Devey et al. 1996) and in a wide interpopulational cross of *E. nitens* (22%) (Byrne et al. 1995). Evidently besides the higher number of alleles expected at SSR loci, the interspecific nature of the pedigree used for linkage mapping in this study also increases the probability of detecting fully informative markers.

Similar to human genetics studies in the pre-microsatellite era, a key obstacle to a more precise analysis of linkage and the detection of QTLs in *Eucalyptus* has been the outbred nature of the pedigrees available and the limited polymorphism information content of the molecular–marker classes used to-date to construct linkage maps. The availability of microsatellite maps will certainly represent a dramatic improvement in our ability to carry out high-quality mapping experiments and implement marker-assisted selection. This work demonstrates that the construction of a comprehensive high-density linkage map of SSR loci should be achievable for *Eucalyptus*. Given the very broad genetic base of most *Eucalyptus* breeding programs, and heterozygosities in the 0.7–0.9 range for SSR loci we anticipate that essentially every family segregating for the trait of interest should provide marker-QTL linkage information. Transportability of SSR loci among pedigrees, not only within the same species but also across species of the same subgeneus, should allow the determination of QTL synteny and facilitate a directed search for new allelic variation at known QTLs within and among species. Finally, as most QTLs in *Eucalyptus* and in other genetically heterogeneous forest trees are likely to be multiallelic, only multiallelic markers will permit tracking, understanding and the adequate manipulating of allelic variation at QTLs in marker-assisted breeding.

**Acknowledgements** We gratefully acknowledge our friends Dr. Scott Tingey, Dr. Antoni Rafalski and Dr. Julie Vogel of Dupont who all contributed to make this work possible. We also thank the Brazilian Ministry of Science and Technology – PADCT-FINEP Competitive Grants in Biotechnology to D.G.; the ministry of Education CAPES Doctoral fellowships to R.P.V.B. and C.B.; the Brazilian National Research Council-CNPq Research fellowship to D.G., and Aracruz Celulose S.A. for providing the mapping population.

## References

Bassam BJ, Caetano-Anolles G, Gresshoff PM (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal Biochem 196:80–83

Bell CJ, Ecker JR (1994) Assignment of 30 microsatellite loci to the linkage groups of *Arabidopsis*. Genomics 19:137–144

Bishop MD, Kappes SM, Keele JW, Stone RT, Sunden SLF, Hawkins GA, Toldo SS, Fries R, Grosz MD, Yoo J, Beattie GW (1994) A genetic linkage map for cattle. Genetics 136:619–639

Byrne M, Murrell JC, Allen B, Moran, G (1995) An integrated linkage genetic linkage map for eucalyptus using RFLP, RAPD and isozyme markers. Theor Appl Genet 91:869–875

Byrne M, Marques-Garcia MI, Uren T, Smith DS, Moran GF (1996) Conservation and genetic diversity of microsatellite loci in the geneus *Eucalyptus*. Aust J Bot 44:331–341

Chase MR, Kesseli R, Bawa KS (1996) Microsatellite markers for population and conservation genetics of tropical trees. Am J Bot 83:51–57

Chen X, Temnykh S, Xu Y, Cho YG, McCouch SR (1997) Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L.). Theor Appl Genet 95:553–567

Conditt R, Hubbell SP (1991) Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. Genome 34:66–71

Devey ME, Bell JC, Smith DN, Neale DB, Moran GF (1996) A genetic linkage map for *Pinus radiata* based on RFLP, RAPD, and microsatellite markers. Theor Appl Genet 92:673–679

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazdan L, Schoun F, Lathrop M, Gyapay G, Morrissette J, Weissenbach J (1996) A comprehensive gentic map of the human genome based on 5264 microsatellites. Nature 380:152–154

Dietrich WF, Miller JC, Steen RG, Merchant M, Damron-Boles D, Husain Z, Dredge R, Daly MJ, Ingalls KA, O'Connor TJ, Evans CA, DeAngelis MM, Levinson DM, Kryglyak L, Goodman N,

Copeland NG, Jenkis NA, Hawkins TL, Stein L, Page DC, Lander ES (1996) A comprehensive genetic map of the mouse genome. Nature 380 : 149–152

Diwan N, Cregan PB (1997) Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. Theor Appl Genet 95 : 723–733

Dow BD, Ashley MV, Howe HF (1995) Characterization of highly variable (Ga/CT) *n* microsatellites in the bur oak, *Quercus macrocarpa*. Theor Appl Genet 91 : 137–141

Echt CS, May-Marquardt P (1997) Survey of microsatellite DNA in pine. Genome 40 : 9–17

Echt CS, May-Marquardt, Hseih M, Zahorchak R (1996) Characterization of microsatellite markers in eastern white pine. Genome 39 : 1102–1108

Grattapaglia D, Sederoff RR (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. Genetics 137 : 1121–1137

Grattapaglia D, Bertolucci FL, Sederoff RR (1995) Genetic mapping of quantitative trait loci controlling vegetative propagation in *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross mapping strategy and RAPD markers. Theor Appl Genet 90 : 933–947

Griffin AR, Cotteril PP (1988) Genetic variation in growth of outcrossed, selfed and open-pollinated progenies of *Eucalyptus regnans* and some implications for breeding strategy. Silvae Genet 37 : 124–131

House APN, Bell JC (1994) Isozyme variation and mating system in *Eucalyptus urophylla* S.T. Blake. Silvae Genet 43 : 167–176

Kostia S, Varvio SL, Vakkari P, Pulkkinen P (1995) Microsatellite sequences in a conifer, *Pinus sylvestris*. Genome 38 : 1244–1248

Kresovich S, Szewc-McFadden AK, Bliek SM, McFerson JR (1995) Abundance and characterization of simple sequence repeat (SSR) loci isolated from a size-fractionated genomic library of *Brassica napus* L. (rapessed). Theor Appl Genet 91 : 206–211

Lander ES, Green P, Abrahamson J, Baarlow A, Daly MJ, Lincoln SE, Newburg L (1987) MapMaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1 : 174–181

Lincoln SE, Daly MJ, Lander ES (1991) "Primer" software, available from E. Lander, Whitehead Institute, Cambridge, Massachusets

Love JM, Knight AM, McAleer MA, Todd JA (1990) Toward construction of a high-resolution map of the mouse genome using PCR-analysed microsatellites. Nucleic Acids Res 18 : 4123–4130

Morgante M, Rafalski A, Biddle P, Tingey S, Olivieri AM (1994) Genetic mapping and variability of seven soybean simple sequence repeat loci. Genome 37 : 763–769

Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends Plant Sci 1 : 209–245

Rafalski JA, Vogel JM, Morgante M, Powell W, Andre C, Tingey SV (1996) Generating and using DNA markers in plants. In: Birren B, Lai E (eds) Analysis of non-mammalian genomes – a practical guide. Academic Press, New York, pp 75–134

Roher GA, Alexander LJ, Keele JW, Smith TP, Beattie CW (1994) A microsatellite linkage map of the porcine genome. Genetics 136 : 231–245

Saghai-Maroof MA, Biyashev RM, Yang GP, Zhang Q, Allard RW (1994) Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations and population dynamics. Proc Natl Acad Sci USA 91 : 5466–5470

Smith DN, Devey ME (1994) Occurrence and inheritance of microsatellites in *Pinus radiata*. Genome 37 : 977–983

Szwec-McFadden AK, Kresovich S, Bliek SM, Mitchell SE, McFerson JR (1996) Identification of polymorphic, conserved simple sequence repeats (SSRs) in cultivated *Brassica* species. Theor Appl Genet 93 : 534–538

Taramino G, Tingey S (1996) Simple sequence repeats for germplasm analysis and mapping in maize. Genome 39 : 277–287

Taylor GR, Haward S, Noble JS, Murday V (1992) Isolation and sequencing of CA/GT repeat microsatellites from hromosomal libraries without subcloning. Anal Biochem 200 : 125–129

Valdes AM, Slatkin M, Freimer NB (1993) Allele frequencies of microsatellite loci: the stepwise mutation model revised. Genetics 133 : 737–749

Ven WTG vad de, McNicol RJ (1996) Microsatellites as DNA markers in Sitka spruce. Theor Appl Genet 93 : 613–617

Wang Z, Weber JL, Zhong G, Tanksley SD (1994) Survey of plant short tandem DNA repeats. Theor Appl Genet 88 : 1–6

Weber JL (1990) Informativeness of human $(dC\text{-}dA)_n$ $(dG\text{-}dT)_n$ polymorphisms. Genomics 7 : 524–530

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, Lathrop M (1992) A second-generation linkage map of the human genome. Nature 359 : 794–801

White G, Powell W (1997) Isolation and characterization of microsatellite loci in *Swietenia humilis* (Meliaceae): an endangered tropical hardwood species. Mol Ecol 6 : 851–860

Ziegle JS, Su Y, Corcoran KP, Nie L, Mayrand PE, Hoff LB, McBride LJ, Kronick MN, Diehl SR (1992) Application of automated DNA-sizing technology for genotyping microsatellite loci. Genomics 14 : 1026–1031