Theor Appl Genet (2000) 101:1282–1290

P.A. Butcher · S. Decroocq · Y. Gray · G.F. Moran

Development, inheritance and cross-species amplification of microsatellite markers from *Acacia mangium*

Received: 30 December 1999 / Accepted: 10 May 2000

Abstract Microsatellite markers were developed in *Acacia mangium* Willd. to provide highly variable co-dominant markers for linkage mapping and studies of the breeding system. After an enrichment procedure 40% of colonies contained microsatellites in contrast with less than 1% from a non-enriched library. The majority of microsatellite sequences were AC repeats. Co-dominant segregation of alleles in two full-sib crosses of *A. mangium* was demonstrated at 33 microsatellite loci. The markers were highly variable relative to restriction fragment lengths polymorphisms (RFLPs). In the two pedigrees 53% of microsatellite loci were fully informative compared with 15% of RFLPs. Based on alleles detected among four parental genotypes, the microsatellites consisting of dinucleotide repeats were more polymorphic than those with tri- and tetra-nucleotide repeats. The microsatellite markers were not as transferable across species in the genus *Acacia* as RFLPs. Two thirds of the primers developed in *A. mangium* (subgenus Phyllodineae, section Juliflorae) amplified DNA from other species within the same section but failed to amplify in species from the subgenus Acacia. The availability of multiallelic, PCR-based, co-dominant microsatellite loci makes possible efficient studies of gene flow and breeding systems in *A. mangium*, a species with low allozyme variation.

Keywords *Acacia mangium* · Microsatellites · Simple sequence repeats · Genome mapping

Introduction

Acacia mangium Willd. is a tropical tree species which occurs naturally in Australia, New Guinea and islands of the Moluccas. Its fast growth rates and adaptability to a wide range of sites have led to it becoming the most-widely planted species in South-east Asia for pulp and paper production. Early domestication of the species was from seed production stands which were established from a narrow genetic base (Butcher et al. 1996) and which had inferior growth rates to several natural populations (Harwood and Williams 1992). More recently breeding programs have been directed towards broadening the genetic base of breeding populations for recurrent selection and developing clonal breeding programs. The application of such strategies has opened opportunities to further improve the efficiency of breeding programs using marker-assisted selection.

A genetic linkage map is being constructed for *A. mangium* to locate quantitative trait loci (QTLs) controlling wood quality and disease resistance traits which could be used for marker-assisted selection. The development of restriction fragment length polymorphisms (RFLPs) for *A. mangium* revealed variation which was adequate for linkage mapping; however, the proportion of fully informative loci (i.e. at least three alleles and segregation on both sides of a cross) was low (15%) (Butcher et al. 2000). Fully informative markers are more likely to be variable in other mapping pedigrees and are therefore particularly useful for producing consensus maps from different pedigrees and/or different species. By combining male and female segregations, they are also useful for deducing the mode of gene action of QTLs. Comparative studies of RFLPs and microsatellite loci, also termed simple sequence repeats (SSRs), in forest trees have shown that microsatellite loci are more variable (Byrne et al. 1996; Devey et al. 1999). Both types of marker are co-dominant and appear to be randomly distributed throughout tree genomes (Brondani et al. 1998; Devey et al. 1999). Microsatellite loci are assayed by the polymerase chain reaction (PCR) and therefore require less DNA than RFLPs. This, together with the ability to semi-automate their analysis, makes them particularly useful for screening large numbers of individuals in breeding populations.

Communicated by M.A. Saghai Maroof

P.A. Butcher (✉) · S. Decroocq · Y. Gray · G.F. Moran
CSIRO Forestry and Forest Products, PO Box E4008, Kingston, ACT 2604, Australia
e-mail: Penny.Butcher@ffp.csiro.au
Tel.: +61-2-6281 8289, Fax. +61-2-6281 8233

The relative genomic abundance of the different microsatellite repeat motifs will determine which motifs are most suitable for marker-development efforts. In mammals the (AC) motif is the most common dinucleotide repeat (Beckmann and Weber 1992), and (CAG) and (AAT) the most common trinucleotide repeats (Stallings 1994). In contrast, (AT) and (AG) repeats appear to be more abundant in plant genomes than (AC) repeats (Lagercrantz et al. 1993). Evidence suggests tri- and tetranucleotide repeats are less abundant in plant genomes (Lagercrantz et al. 1993; Ma et al. 1996). However, the loci from these repeat classes may be easier to score because of a lower incidence of stutter bands and larger unit allelic-size differences (Hague and Litt 1993; Kijas et al. 1995). To examine the factors affecting the efficiency of microsatellite development in *A. mangium*, the abundance and variability of microsatellites from libraries enriched for dinucleotide and trinucleotide repeats were compared with that from a *Sau*3AI library.

The informativeness of microsatellite loci for linkage mapping in *A. mangium* was ascertained by comparing the segregation of alleles at microsatellite and RFLP loci in two mapping pedigrees. The amplification of microsatellite primers developed in *A. mangium* was also tested in a range of other acacias to explore their potential application in comparative studies of genome organisation.

Materials and methods

Standard library construction and screening

Genomic DNA was extracted from *A. mangium* phyllodes (female mapping parent, cross B) following the procedures in Butcher et al. (1998). Size-selected (250–600 bp) genomic-DNA digested with *Sau*3A1 was ligated into *Bam*H1, digested and de-phosphorylated pUC18 (Pharmacia), and used to transform competent *Escherischia coli* DH10B (Life Technologies). The library was screened by colony hybridisation using 150–200 ng each of ^{32}P-labelled $(GA)_{10}$, $(CA)_{10}$ and $(CAC)_{15}$ oligonucleotides in 6 x SSPE, 5×Denhardt's reagent (Sambrook et al. 1989), 1% sodium dodecyl suphate (SDS) with 3×20 min post-hybridisation washes, using 0.5×SSC and 1% SDS, at 65°C. Positive colonies were re-plated and a second round of hybridisation carried out to confirm their positive status. Plasmid DNA from individual colonies was prepared using a Promega Wizard kit (Madison, Wis., USA).

Construction and screening of libraries enriched for microsatellites

To improve the efficiency of microsatellite development, the enrichment procedure of Edwards et al. (1996) was used with some modification. Two sets of membranes were prepared, one to enrich for dinucleotide repeats, another to enrich for tri- and tetranucleotide repeat sequences. Fifty nanograms of each microsatellite motif, [$(GA)_{10}$, $(CA)_{10}$] (Pharmacia), [$(CGC)_{15}$, $(AGC)_{15}$, $(CTC)_{15}$, $(CAC)_{15}$, $(AAC)_{15}$, $(ACAT)_{12}$, $(CCTT)_{12}$, $(ACGT)_{12}$, $(ACAG)_{12}$ and $(GCAT)_{12}$] (Life Technologies), were added to 40 ul of 0.4 mM NaOH and spotted onto a 0.5 cm^2 Hybond N$^+$ membrane (Amersham, Ill., USA), dried for 2 h at 80°C, boiled in 1% SDS to remove weakly bound DNA and pre-hybridised for 24 h in 6×SSPE, 5×Denhardt's reagent and 1% SDS at 65°C. Genomic DNA was digested with *Rsa*1, ligated with a *Mlu*I adaptor (Edwards et al. 1996) and purified using BRESA-CLEAN (Bresatek Ltd.) according to the manufacturer's directions. Enrichment for microsatellites was carried out using 100 ng of adaptor-ligated, denatured genomic DNA and 4 µg of the 21-mer oligonucleotide (5´CTCTTGCTTACGCGTGGACTA3´), and hybridised in a single tube containing both membranes for 24 h at 65°C. Membranes were then washed in 2×SSC, 0.1% SDS at room temperature for 10 min, at 65°C for 10 min, then 3-times in 0.5×SSC, 0.1% SDS for 10 min at 65°C. Bound-DNA was eluted by boiling membranes in 200 µl of sterile distilled water for 5 min. One microlitre of eluted-DNA was amplified by PCR with 0.8 µM of the 21-mer *Mlu*I primer in 25 µl of 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, and 1 U of Taq Polymerase (Life Technologies) using a Perkin-Elmer 9600 thermocycler. The PCR products were restricted with *Mlu*I and cloned into a PjVI vector (Edwards et al. 1996) prior to transformation into DH10B (Life Technologies). The library of microsatellite-enriched clones was screened by colony hybridisation using di-, tri-, and tetra-nucleotide motif probes, as described for the standard library. One round of screening was carried out for the dinucleotide enriched library and two rounds for the tri- and tetra-nucleotide-enriched library to confirm the positive status of colonies.

Sequencing and primer design

Insert DNA was amplified in 25-µl PCR reactions using 400 nM of the 21-mer *Mlu*I primer, and 2 µl run on a 2% agarose gel to determine insert size by comparison with a 100-bp ladder (Gibco). Single PCR products larger than 400 bp were cleaned using spin columns (QIAgen), and sequenced by the dideoxy chain-termination method using a fluorescent dye–primer cycle sequencing kit (Applied Biosystems Inc.) and an ABI 377 Sequencer. If necessary, clones containing a microsatellite were re-sequenced with reverse primers to obtain reliable flanking sequence. Redundant clones were identified using the program AMPLIFY (Engels 1992). Where the microsatellite repeat region was larger than 20 bp and and flanking sequences were unique, primers were designed using the program PRIMER Version 0.5 (Lincoln et al. 1991) and synthesised by Life Technologies Inc.

PCR-amplification of microsatellites and fragment analysis

Fifty nanograms of template genomic DNA was amplified by PCR with 0.2 µM of each primer and 0.2 mg ml^{-1} of bovine serum albumin in 25-µl reactions, using conditions described previously. Annealing temperatures were varied to optimise the amplification of each primer pair. Amplification cycles were: 94°C for 2 min; 30 cycles of (94°C for 30 s, 55°C, 60°C or 65°C for 30 s and 72°C for 60 s); 72°C for 10 min. The PCR products were separated on 20-cm 8% non-denaturing polyacrylamide gels run in 1×TBE, with cooling, at 350 V for 4–9 h, depending on the size of the amplification products, then stained in ethidium bromide (0.5 µl ml^{-1} H_2O) for 30 min. *Hpa*II-digested pUC19 was used as a molecular-size marker.

Following optimisation of PCR conditions for each microsatellite locus, dye-labelled reverse primers were synthesised for semi-automated analysis. PCR reactions were scaled down to 10 µl using 20 ng of template DNA and dye-labelled primers (Life Technologies); all other conditions remaining the same. PCR products were separated by capillary electrophoresis on an ABI PRISM 310 Genetic Analyzer and fragment-sizes determined using GENESCAN and GENOTYPER fragment analysis software (Applied Biosystems).

Inheritance of microsatellite loci

The segregation of alleles at 33 microsatellite loci was compared with expected Mendelian ratios by a χ^2 goodness-of-fit analysis. Segregation ratios were calculated for two 2-generation outcrossed pedigrees of *A. mangium*, described in Butcher et al. (2000). The progeny arrays consisted of 108 and 123 individuals in crosses A and B respectively.

Polymorphism and cross-species amplification of microsatellite loci

To compare the relative informativeness of microsatellite and RFLP loci in the two *A. mangium* mapping pedigrees, the number of alleles was determined at 33 microsatellite loci and compared who that at 344 RFLP loci (Butcher et al. 2000). The degree of length-polymorphism was also compared between microsatellite sequences consisting of di-, tri- and tetra-nucleotide repeats.

The amplification of *A. mangium* microsatellite primers was examined in three unrelated individuals from each of 14 acacia species representing the subgenera Acacia and Phyllodineae. Individuals of *Pararchidendron pruinosum, Archidendron lucji* and *Archidendron ramiflorum* were also screened to determine whether primers would amplify in species from other genera. Twelve primer pairs were amplified separately with DNA from each individual, using the same PCR conditions and methods of analysis as for *A. mangium*.

Results

Library screening

The number of clones sequenced, primers designed, and primers which amplified a single locus in *A. mangium*, are summarised in Table 1 for the *Sau*3AI and enriched libraries. Less than 1% of colonies were SSR-positive in the *Sau*3A1 library, compared with 40% of colonies positive for dinucleotide repeat motifs and 17% of colonies positive for trinucleotide repeat motifs in the enriched libraries. The insert size ranged from 200 to 1400 bp with an average of 680 bp.

A higher proportion of sequenced clones from the dinucleotide-enriched library resulted in useable microsatellite loci than from the *Sau*3A1 library or the trinucleotide-enriched library (Table 1). Clone sequences were discarded prior to primer design for several reasons which included there being insufficient flanking sequence to design primers, no microsatellite present and duplicate or multiple clone sequences. The relatively low number of tri- and tetra-nucleotide microsatellite sequences was due to a high proportion of redundant clones (30%). This may reflect the lower frequency of tri- and tetra- nucleotide repeats in non-coding regions of plant genomes (Wang et al. 1994).

Sequence characterisation

Of the 75 sequences from the enriched libraries for which primers were designed (Table 1) 50 were compound, i.e. contained two or more different repeat motifs. Only one of the six microsatellites from the *Sau*3A1 library was compound. Edwards et al. (1996) also reported that a significant proportion of clones from enriched libraries contained more than one microsatellite sequence.

The most common dinucleotide repeat motif was $(AC)_n$ (70%), followed by $(AT)_n$ (25%) and $(AG)_n$ (5%). (Microsatellite motifs were named using the convention of alphabetical ordering, see Echt and May-Marquardt 1997). For sequences with tri- and tetra-nucleotide repeats the most abundant motif was $(ACC)_n$, (28%). Other motifs in order of frequency were $(AAC)_n$ (14%); $(AAG)_n$, $(AAT)_n$ and $(CTC)_n$ (8%); $(AGT)_n$ and $(AGG)_n$ (6%); $(ACG)_n$, $(AGC)_n$, $(ATG)_n$, $(ACAT)_n$, $(AAAC)_n$, $(ACTC)_n$ and $(AATG)_n$ (3%). The average number of (AC) repeat units (16.6±0.9) was greater than (AG) repeat units (12.5±1.9). The maximum repeat number found for $(AC)_n$ was 93 and for $(AG)_n$ was 15. Trinucleotide microsatellite sequences ranged from 3 to 11 repeat units with an average of 7.5±0.6. For compound microsatellite sequences, the calculation of repeat length was based on the longest uninterrupted repeat. This was determined by evidence that length variations preferentially affect the longest repeat in compound and interrupted microsatellite sequences (Loridon et al. 1998).

Polymorphism of microsatellite loci

Forty three microsatellite primer pairs were screened across four first-generation individuals of the two mapping pedigrees. Thirty three loci were polymorphic and primer sequences for these loci are listed in Table 2. Eight loci were invariant in both mapping pedigrees and two had null alleles. Of the 33 variable loci, three or more alleles were detected in the mapping parents at 26 loci (78%) compared with 28% of the RFLP loci (Fig. 1). Sixteen percent of the loci were fully informative in both crosses, 38% and 53% of loci were fully informative in crosses A and B respectively.

Polymorphism, as measured by the number of alleles detected in the four mapping parents, was correlated with microsatellite repeat length (r=0.59; P<0.001 determined from a one-way ANOVA). Microsatellite sequences consisting of only dinucleotide repeats were more polymorphic (mean of 4.5 alleles) than those with tri- and tetra-nucleotide repeats (mean of 2.3 alleles) (P<0.001 from a two-sample t-test assuming unequal variance). Only two alleles were detected among the four mapping parents for seven of the nine microsatellite loci containing tri- or tetra-nucleotide repeats. In contrast, more than two alleles were detected among parental samples with all microsatellite loci consisting of

Table 1 The number of clones sequenced from *Sau*3A1 and enriched libraries which contained microsatellites (SSRs), were suitable for primer design and amplified a single locus in *A. mangium*

Library	Clones sequenced	Contain SSR	Primers designed	Amplified
*Sau*3AI	35	11 (31%)	6	4
Di-enriched	115	101 (88%)	60	32
Tri- and tetra-enriched	73	54 (74%)	15	7
Total	223	166 (74%)	81	43

Table 2 Primer sequences of 33 microsatellite loci and length variants detected among four *A. mangium* mapping parents

Name	Primer sequences (5′–3′)	Allele sizes (bp)
Dinucleotide microsatellites		
Am008	CCACCCGTTACCCATTTATG CCGTGATTGACTCTCAGCG	106, 104, 96, 94, 88
Am012	TGAGTCGATCGCTTAGCTTG TCCCGTTATTATGCCAAAGTG	156, 154, 150
Am014	GTACTAACGTTGCTATATGAGAAAGG CTGGTTGTTCGCTTATATGG	154, 150
Am018	CACGGCTGTTATTTCCTTCG GGAAAGAGGTGTGACAGAGGAC	144, 137, 135, 129
Am030	GAGGTAATATTTTGAATTCCTTGAAC GGTGTATACCTCTTTCCTGTGG	121, 107, 96, 89
Am041	TAGGCTAATGGTCATATTCCTAG AGAGATAGGGGTACACACTAAAAAAC	147, 143, 137, 129, 123, 116, 114
Am136	CCCATTGCCGTTTCTTTG GCATTTCCCTTGGAACAGTC	125, 121, 113, 111
Am164	ACCCGGACGTATAGAAATAAATACA CGTGGAGGCAAGCAATATC	220, 160, 156, 139, 128, 98, 81, 74
Am173	TTGGATGTCAAGATTTTACGG CATTAGGCCACGTTTTGATAG	103, 101, 97, 94
Am326	GGACCAAACTTATGCAACACC GCATCAATGTACTAAACCATTTCC	249, 243, 241, 230, 218
Am341	CCATTCGAGCATCCTAAGAG CGTATGGCTGAGCTACTTAATCA	130, 128, 126, 124
Am352	CCTCATGTCCTTGAATGTCAC GACTAACCCACAAGGAAGAGTTAC	129, 127
Am367	CGCAACTCCATCTGATTTACTG TTATGTTGGGTTAATACGCTAACTG	112, 111, 99, 96
Am368	CAC AAG GAA CTG AGC AAT GG TCT TGC CTA GGT AGA TTT TGA CC	100, 96, 94, 92, 91
Am384	AGA CTT CAT AAA TAA GAT GGA AGA GG ATG CCA AAT TTT CTT ATT GGA G	197, 195
Am387	TGATACAAGGGAAGACAGAGTGG CCAACTCAAAACCTGACAACG	114, 112, 104
Am389	AATCCTTCCGAAAGTTATACATGG GCACTTGTAAGTCGGAACTGC	220, 218, 216
Am391	ATGGACAGTTAATTAGTGATCTGATG GGACTAACCCAGATTCCACC	86, 82, 74
Am396	CGGATGAGAAGTGGCATTAGG GCCATCCAGAGTTGAAATTAGC	99, 73, 69
Am400	AAGCACTTCTAGATTGGCAGC CAATAATCTAGAAACTACGGGACTTG	89, 83, 81
Am424	AATACATGGAAGAGGATGAGATG ATTGCATTTCATTTGTTGCC	206, 187, 185, 180, 176, 174
Am429	CCTTCTTCTCTCATCTACCAAACC CCCACATCATCACTCACAACT	180, 178, 174, 170
Am435	ACCCTTTATTTCTCACACGGA ACAGAAGAAGATGCAAAGAAGG	152, 143, 141, 139
Am436	ATGGATCTTGTCCTTATCTTGA GGGCCAATTTGAGTTTGGAA	246, 244, 242, 240
Am460	CACTAATTGCTCACACATTCCA ATTCATAGCCTCTCCCTTCAG	140, 138
Am463	AACATCCCTGCCAATAAACA TCAACTCCGTAGGACCAAAG	94, 92, 90
Am465	TGGGTATCACTTCCACCATT AGGCTGCTTCTTTGTGCAGG	180, 176, 174, 170, 162, 158, 150
Am477	GGTCATGGAAGAGGAGAATG TCTTATACTAGGCTCACAACACTC	101, 99, 91
Am484	ACATTATTTCCCTTCAGTCTATATG GATTTGCTTTATCGTTCTTCTG	115, 113, 109, 105
Am503	GTATGAGTTCCAGTCCTACCATCA CAGTCCGGTTTTTGCTGTCA	170, 168, 162
Trinucleotide microsatellites		
Am502	CAAATGGCCAAGTTACGACTG TTCTGGTAATCCAAACTTATGTGG	128, 122
Am522	ATGAGTTTGACCCGACTTGA TGCGTAACACGATTATCCCT	132, 121
Am770	CAGAGGTGGCAGATGATGTC AAGCCTTTAGTTGGGCGTTC	95, 93

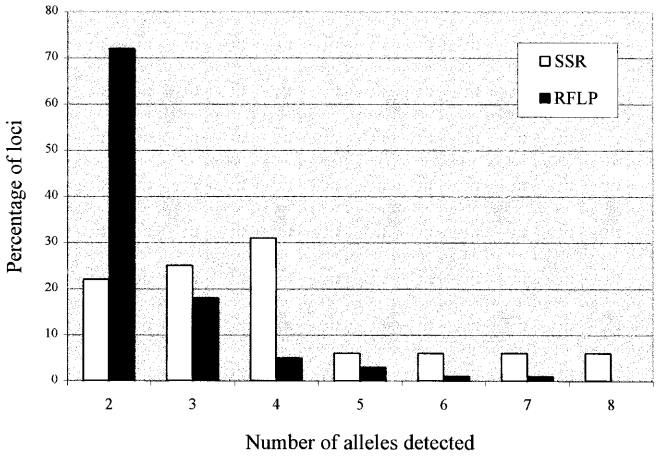

Fig. 1 Comparison of the variability of microsatellites (33 loci) and RFLPs (344 loci) based on the number of alleles detected in four first-generation mapping individuals

only dinucleotide repeats. Forty five percent of microsatellite loci with tri- and tetra- nucleotide motifs in their sequence were polymorphic in both pedigrees compared with 74% of microsatellites with only dinucleotide repeats.

Inheritance of microsatellite loci

Segregation in the two mapping pedigrees was assessed at 33 microsatellite loci (Table 3). Twenty one loci were variable in both crosses, three were variable in cross A only, and nine in cross B only. Segregation ratios were distorted at two loci in cross A and one locus in cross B, within the 5% probability expected due to chance alone. The loci with distorted segregation in cross A segregated according to the expected Mendelian ratios in cross B (Table 3). Slightly more loci segregated on the female side of the pedigrees than the male; 79% and 83% of loci segregated on the female side of crosses A and B, respectively, compared with 62% and 80% on the male side.

Cross-species amplification of *A. mangium* microsatellite loci

Cross-species amplification was scored as positive only when sharp peaks greater than 80 bp were produced in the three individuals from each species. In *Acacia*[1] one-third of the microsatellite loci amplified in all species in the subgenus Phyllodineae, and two-thirds amplified within the section Juliflorae (Table 4). Only one microsatellite locus amplified in the subgenus Acacia, but surprisingly, one-quarter of the microsatellite primers amplified fragments in the genera *Archidendron* and *Pararchidendron*. The alleles were not sequenced so it is not possible to say whether loci are homologous.

[1] The genus *Acacia* belongs to the tribe Acacieae within the legume subfamily Mimosoideae (Maslin and Stirton 1997). It includes three subgenera: Acacia; Phyllodineae; and Aculeiferum. *A. mangium* is classified in the section Juliflorae within the subgenus Phyllodineae. This subgenus includes seven sections (namely Botrycephalae, Pulchellae, Alatae, Lycopodiifoliae, Phyllodineae, Plurinerves and Juliflorae). The genera *Archidendron* and *Pararchidendron* are in the tribe Ingeae (Leguminosae: Mimosoideae)

Discussion

Variability of microsatellite loci

The variability of microsatellite loci in the two mapping pedigrees, based on the number of alleles detected in the first-generation individuals (Fig. 1) and the proportion of fully informative loci, was more than three-times higher than previously detected with RFLPs (Butcher et al. 2000). This is consistent with data from a survey of 20 unrelated individuals from natural populations of *A. mangium* which revealed levels of genetic diversity detected at five microsatellite loci (A=6.6; H_T=0.704) that were more than three-times higher than previously detected in the same individuals with 58 RFLP loci (A=2.1; H_T=0.205) (DeCroocq et al. 1997). Microsatellite variation was 30-times higher than reported in an allozyme survey which screened a larger number of individuals and populations (Moran et al. 1989).

The level of polymorphism detected amongst the four mapping parents was positively correlated with the number of contiguous repeats in the microsatellite sequence. Similar results have been reported in humans (Weber et al. 1990) and *Arabidopsis* (Loridon et al. 1998). In *A. mangium* the trinucleotide microsatellite loci had fewer repeat units (mean of seven contiguous repeats) than the dinucleotides (mean of 18 contiguous repeats) and significantly less variation was detected at these loci. A similar trend was observed in an enriched library from *Melaleuca alternifolia* where dinucleotide microsatellite loci had a higher average number of repeats (22) than trinucleotides (8) and were more polymorphic (Rossetto et al. 1999).

Sequence characterisation

The most abundant dinucleotide motif in *A. mangium* was $(AC)_n$. This contrasts with surveys in plants that have shown the most frequently occurring dinucleotide repeats are $(AT)_n$, with $(AG)_n$ and $(AC)_n$ as second and third most frequent (Condit and Hubbell 1992; Lagercrantz et al. 1993; Morgante and Olivieri 1993; Wang et al. 1994; Echt and May-Marquardt 1997). The (AT) repeat is self-complementary and therefore is difficult to screen for by colony hybridisation. As a result, these repeats were only detected in the *A. mangium* libraries in compound microsatellite sequences. Surveys of microsatellite sequences in forest trees have generally found the $(AG)_n$ motif to be more abundant than $(AC)_n$, for example, *Quercus macrocarpa* [83% $(AG)_n$] (Dow et al. 1995), *Eucalyptus nitens* (Byrne et al. 1996) *Eucalyptus*

Table 3 Segregation ratios at 33 microsatellite loci in two outcrossed pedigrees (cross A and B) of *A. mangium* and probability (Pr) of Mendelian inheritance based on χ^2 analysis

Locus	Library	X	Repeat motif(s)	Segregation	Expected	χ^2	Pr
Am030	*Sau*3A1	A	$(AT)_9(GT)_{15}$	56:55	1:1	0.01	0.92
		B		67:58	1:1	0.65	0.42
Am041	*Sau*3A1	A	$(GT)_{36}$	24:29:28:30	1:1:1:1	0.75	0.96
		B		52:73	1:1	3.53	0.06
Am136	*Sau*3A1	A	$(CT)_{20}$	61:49	1:1	1.31	0.25
		B		30:66:29	1:2:1	0.41	0.82
Am164	*Sau*3A1	A	$(TG)_{93}$	29:29:26:27	1:1:1:1	0.24	0.97
		B		39:28:26:32	1:1:1:1	3.16	0.21
Am008	enriched	A	$(TG)_{14}$	31:34:26:20	1:1:1:1	4.06	0.26
		B		39:32:31:23	1:1:1:1	4.12	0.13
Am012	enriched	A	$(TC)_{15}(AC)_7$-$(AC)_{10}AG$	58:53	1:1	0.23	0.64
		B		63:62	1:1	0.01	0.93
Am014	enriched	A	$(ATAC)_3(AC)_{26}(AT)_3(GTAT)_2\ (AT)_3AC$	Invariant			
		B		31:64:30	1:2:1	0.07	0.96
Am018	enriched	A	$(AC)_{14}$	62:49	1:1	1.52	0.22
		B		55:69	1:1	1.58	0.21
Am173	enriched	A	$(AC)_{18}$	Invariant			
		B		32:29:35:29	1:1:1:1	0.79	0.73
Am326	enriched	A	$(CA)_{20}$	31:29:26:25	1:1:1:1	0.82	0.29
		B		40:31:30:24	1:1:1:1	4.18	0.85
Am341	enriched	A	$(CA)_{12}(TA)_2$	28:55:28	1:2:1	0.01	0.99
		B		37:32:31:25	1:1:1:1	2.33	0.58
Am352	enriched	A	$(TTC)_2TA(AC)_{14}$	Invariant			
		B		68:56	1:1	1.16	0.28
Am367	enriched	A	$(A)_7G(A)_6GG(A)_{14}(CA)_{14}$	56:55	1:1	0.01	0.92
		B		34:31:30:30	1:1:1:1	0.34	0.86
Am368	enriched	A	$(A)_9(CA)_{13}$	58:53	1:1	0.23	0.64
		B		39:31:28:27	1:1:1:1	2.84	0.29
Am384	enriched	A	$(TG)_{11}$-$(TG)_6\ T(TG)_5$-$(TG)_6(AG)_8$-$(TG)_4(GAA)_3$	60:51	1:1	0.73	0.39
		B		Invariant			
Am387	enriched	A	$(AT)_2(GT)_2(AT)_2(GT)_{17}$-$(TA)_8$	57:54	1:1	0.08	0.78
		B		25:36:40:24	1:1:1:1	6.1	0.11
Am389	enriched	A	$(AT)_3(GT)_9GC(GT)_2$	Invariant			
		B		44:34:24:23	1:1:1:1	9.30	0.02
Am391	enriched	A	$(AT)_2(TG)_{13}$	Invariant			
		B		37:32:28:27	1:1:1:1	2.00	0.66
Am396	enriched	A	$(AC)_{16}(AT)_6$	60:51	1:1	0.73	0.39
		B		73:52	1:1	3.53	0.06
Am400	enriched	A	$(GCA)_2(GT)_{11}$	30:28:27:26	1:1:1:1	0.32	0.96
		B		67:58	1:1	0.65	0.42
Am424	enriched	A	$(GT)_2\ (TG)_2$,$A(GT)_4$,$GA(GT)_9$	60:51	1:1	0.73	0.39
		B		39:32:30:24	1:1:1:1	3.67	0.37
Am429	enriched	A	$(GT)_{16}$	31:30:26:24	1:1:1:1	1.18	0.76
		B		Invariant			
Am435	enriched	A	$(CT)_5$,$CC(CT)_2(CA)_9(CACT)_3$	43:24:24:20	1:1:1:1	11.56	0.01
		B		65:60	1:1	0.20	0.66
Am436	enriched	A	$(TG)_{14}$,T	61:50	1:1	1.09	0.30
		B		39:31:28:27	1:1:1:1	2.84	0.24
Am460	enriched	A	$(AT)_4$,$(GT)_{12}$,$(GGAAT)_2$,$(GT)_2$	59:52	1:1	0.44	0.51
		B		67:58	1:1	0.65	0.42
Am463	enriched	A	$(CA)_{11}$	59:52	1:1	0.44	0.51
		B		63:62	1:1	0.01	0.93
Am465	enriched	A	$(AC)_{23}$	41:20:28:22	1:1:1:1	9.68	0.02
		B		34:39:31:21	1:1:1:1	5.53	0.34
Am477	enriched	A	$(GT)_{13}$	Invariant			
		B		35:34:30:26	1:1:1:1	1.62	0.50
Am484	enriched	A	$(TA)_2$,TG,$(TC)_6$,$(TG)_{10}$	Null allele			
		B		41:36:25:23	1:1:1:1	7.19	0.08
Am503	enriched	A	$(TG)_{13}$,$(TG)_8$	Invariant			
		B		34:33:31:26	1:1:1:1	1.23	0.58
Am502	enriched	A	$(TTC)_3$-$(GGA)_8$,$AGA(GGA)_2$	58:53	1:1	0.23	0.64
		B		Invariant			
Am522	enriched	A	$(AAT)_3$-$(AGT)_2(GGT)_7$	26:54:31	1:2:1	0.53	0.77
		B		60:65	1:1	0.20	0.66
Am770	enriched	A	$CTC(CAC)_5CGC(CAC)_3$	Invariant			
		B		30:71:24	1:2:1	2.89	0.24

Table 4 Cross-species amplification of microsatellite primers developed in *A. mangium* in species representing six sections of the subgenus Phyllodineae, the subgenus Acacia and the closely related genera *Pararchidendron* and *Archidendron*. Numbers of alleles detected in three unrelated individuals are shown for all species except *A. mangium* (four individuals), *P. pruinosum*, *A. lucji* and *A. ramiflorum* (one individual from each species)

Species	Am 030	Am 041	Am 136	Am 164	Am 014	Am 018	Am 173	Am 387	Am 465	Am 502	Am 522	Am 770
Subgenus Phyllodineae												
Section Juliflorae												
A. mangium	4	7	4	8	2	4	4	3	7	2	2	2
A. holosericea	5	–	3	–	–	2	–	1	4	1	1	1
A. neurocarpa	2	–	2	–	–	1	–	1	2	1	2	–
A. aneura	2	–	2	–	–	3	–	4	3	1	2	2
Section Plurinerves												
A. melanoxylon	1	–	–	–	–	–	–	–	6	2	–	2
A. flavescens	1	–	–	–	–	–	–	1	4	2	–	1
Section Botrycephylae												
A. dealbata	1	–	–	–	–	–	–	1	2	3	–	1
A. mearnsii	2	–	–	–	–	–	–	1	3	3	–	1
Section Phyllodineae												
A. victoriae	2	–	–	–	–	–	–	–	4	3	–	2
A. falciformis	1	–	–	–	–	–	–	–	4	4	–	2
Section Lycopodiifoliae												
A. lycopodiifoliae	1	–	–	–	–	–	–	–	3	3	–	2
Section Alatae												
A. alata	1	–	–	–	–	–	–	–	1	1	–	1
Subgenus Acacia												
A. bidwillii	–	–	–	–	–	–	–	–	3	–	–	–
A. nilotica	–	–	–	–	–	–	–	–	–	–	–	–
A. boliviana	–	–	–	–	–	–	–	–	–	–	–	–
Other genera												
P. pruinosum	–	–	–	–	–	–	2	1	–	1	–	–
A. lucji	1	–	–	–	–	–	1	–	–	1	–	–
A. ramiflorum	1	–	–	–	–	–	1	–	–	–	–	–

grandis and *Eucalyptus urophylla* (Brondani et al. 1998), *Populus tremuloides* (Dayanandan et al. 1998), *Shorea curtisii* (Ujino et al. 1998), *Melaleuca alternifolia* [43% $(AG)_n$] (Rossetto et al. 1999) and *Grevillea macleayana* (England et al. 1999). Of the conifers only *Larix ssp*. and *Pinus strobus* had more $(AC)_n$ than $(AG)_n$ micosatellite sequences (Echt and May-Marquardt 1997).

The factors responsible for differences in the abundance of microsatellite motifs among plant genera are not well understood (Echt and May-Marquardt 1997); however, for hybridisation surveys they may simply reflect differences in cloning, enrichment or screening procedures. There is evidence that the repeat type and number is influenced by the restriction enzyme used to size-fractionate the genome when constructing the library (Hamilton and Fleischer 1999). However, in *A. mangium* the AC motif was most common in both the *Sau*3AI and *Rsa*I libraries. Condit and Hubbell (1992) found that smaller inserts gave rise to higher estimates for $(AC)_n$ abundance and suggested that poly-AC sites are not randomly or uniformly spread through the genome of tropical trees but are highly clustered in certain regions. They observed that 10–20% of phage inserts that carried an AC site also carried an AG site, indicating a clustering of repeat regions. Non-random distribution and the clustering of AC repeats has also been proposed in the *Elymus caninus* genome (Sun et al. 1998) and in the wheat genome (Ma et al. 1996). Enrichment procedures which produce a high proportion of compound microsatellites (see Edwards et al. 1996) may therefore also increase the probability of detecting AC repeats.

To-date, there have been few comparative studies of the lengths of different microsatellite motifs in tree species. In *A. mangium* the most-common motif also had the highest average repeat length. The same trend was reported in a *Q. macrocarpa*-enriched library while in *E. grandis/E. urophylla* there was no significant difference in the number of dinucleotide repeats for the $(AG)_n$ and $(AC)_n$ motifs. In *A. mangium* the longest and most-polymorphic microsatellite had 93 AC repeats. This contrasts with database screening of the soybean genome (Leguminosae) which revealed no $(AC)_n$ microsatellites with

more than eight repeats, leading Akkaya et al. (1992) to suggest that AC sequences would not offer sufficient polymorphism to permit their use as genetic markers in soybean and perhaps in other plants.

Efficiency of microsatellite marker development in *Acacia*

The efficiency of the enrichment procedure for dinucleotide repeats was similar to that reported by Brondani et al. (1998) in *Eucalyptus* hybridising with biotinilated-oligonucleotides (10–40%). Enrichment for trinucleotide repeats in *A. mangium* was less efficient due to the high proportion of redundant clones. Given that trinucleotide microsatellites were no easier to score when using dye-labelled primers and capillary electrophoresis, the extra effort in their development in *A. mangium* may not be warranted. While the number of trinucleotide primers tested was limited, there was some evidence to suggest that they were less-variable in *A. mangium* but amplified across a broader range of species. Further sampling is required to test this hypothesis.

Cross-species amplification

The proportion of *A. mangium* microsatellite primers which amplified alleles in other species in the genus *Acacia* was relatively low compared with reports from other genera, which suggests that regions flanking the microsatellites are not highly conserved (in nucleotide sequence or relative position). Peakall et al. (1998) reviewed cross-species amplification of microsatellite loci in plants, citing moderate to complete primer conservation (50–100%) within genera, not all of which were informative (20–100% polymorphic). The relatively low cross-species amplification in *Acacia* may reflect the degree of evolutionary divergence among the different sections and subgenera. Approximately one-third of microsatellite primers amplified alleles in all species sampled in the subgenus Phyllodineae. This subgenus, which includes over 900 species, is largely restricted to Australia. Only one microsatellite locus was amplified in the subgenus Acacia. This subgenus has a pantropical distribution, centred in Africa and South America. The Phyllodineae is thought to have evolved independently from the subgenus Acacia (Maslin and Stirton 1997) and was treated by Pedley (1986) as a distinct genus.

More loci amplified in species within the section Juliflorae than in other sections of the subgenus. While uncertainties remain concerning the evolutionary history of the genus *Acacia* and the hierarchical arrangement of its species (Brain and Maslin 1996), the pattern of cross-species amplification was consistent with a general trend in plants of decreasing amplification and decreasing polymorphism with increasing evolutionary distance from the source species. Whitton et al. (1997) reported a decline in amplification success with increasing divergence within the Asteraceae. Similar results were reported in *Brassica* (Lagercrantz et al. 1993), *Eucalyptus sensu lato* (Byrne et al. 1996) and in the Leguminosae (Dayanandan et al. 1997; Peakall et al. 1998). Primer sites for microsatellite sequences in *Pithecellobium elegans* were conserved only among taxa within the tribe Ingeae (Leguminosae: Mimosoideae) and none amplified in *Acacia collincii* (Dayanandan et al 1997). In *Quercus* (Fagaceae) the proportion of amplified fragments, as well as the detected polymorphism, decreased with increasing evolutionary distance (Steinkellner et al. 1997).

The exception to this trend was the amplification of alleles in *Archidendron* and *Pararchidendron*. Fragments were in a similar size range to those amplified in *A. mangium*; however, without sequencing it is not possible to determine whether the amplified loci are homologous. Evidence from cross-species comparisons in other genera suggests that the variation in allele sizes among species can reflect variation in the flanking regions, in addition to variation within the repeat sequence (Steinkellner et al. 1997; Peakall et al. 1998; Karhu et al. 2000). This, together with evidence of size homoplasy among, and even within, species (Grimaldi and Crouau-Roy 1997), means that it is not possible to make inferences concerning evolutionary relationships based on microsatellites without sequencing.

While the potential for the transfer of microsatellite markers outside the section Juliflorae (235 species, Brain and Maslin 1996) is limited, this section includes *Acacia auriculiformis, Acacia crassicarpa, Acacia peregrina,* in addition to *A. mangium,* which are the most commercially important tropical acacias and therefore the most-likely candidates for genetic mapping and gene-isolation efforts. The low cross-species amplification of microsatellite loci contrasts with the high level of transferability of RFLP loci across the genus (Butcher et al. 2000).

The development of microsatellite markers for *A. mangium* has provided highly informative markers for linkage mapping. The higher variability of these markers compared with RFLPs and allozymes increases the power to genetically discriminate between individuals. The microsatellite markers developed in *A. mangium* could be used to estimate pollen contamination in seed orchards, rates of inbreeding and the success of supplemental mass pollination, to determine mating patterns within seed orchards and to identify superior male parents based on the performance of progeny. The relatively low conservation of microsatellite loci across species, together with evidence from other genera of size homoplasy in source and non-source species, indicates they may, however, be of limited use in population and evolutionary studies in other acacias.

Acknowledgements Financial support for this research was provided by CSIRO's Australian Tree Seed Centre (ATSC). The PjVI vector for the enrichment procedure was kindly provided by K. Edwards, Long Ashton Research Station, University of Bristol. Les Pedley (Brisbane Herbarium), Greg Leach (Darwin Botanical Gardens), Australian National Botanic Gardens and ATSC provided leaf material of *Acacia* species. Thanks to Tom Uren, Jeff Glaubitz and Mike Devey for comments on the manuscript.

References

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphism of simple sequence repeat DNA in soybean. Genetics 132:1131–1139

Beckmann JS, Weber JL (1992) Survey of human and rat microsatellites. Genomics 12: 627–631

Brain P, Maslin BR (1996) A serological investigation of the classification of *Acacia* subgenus Phyllodineae (Leguminosae: Mimosoideae). Biochem Syst Ecol 24:379–392

Brondani RPV, Brondani C, Tarchini R, Grattapaglia D (1998) Development, characterisation and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. Theor Appl Genet 97:816–827

Butcher PA, Moran GF, Perkins, HD (1996) Genetic resources and domestication of *Acacia mangium*. In: Dieters MJ, Matheson AC, Nikles DG, Harwood CE, Walker SM (eds) Tree improvement for sustainable tropical forestry. Proc QFRI-IUFRO Conf, Caloundra, 27 Oct–1 Nov 1996, pp 467–471

Butcher PA, Moran GF, Perkins HD (1998) RFLP diversity in the nuclear genome of *Acacia mangium*. Heredity 81:205–213

Butcher P, Moran GF, Bell R (2000) Genetic linkage mapping in *Acacia mangium.* 1. Evaluation of restriction endonucleases, inheritance of RFLP loci and their conservation across species. Theor Appl Genet 100:576–583

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

Condit R, Hubbell SP (1992) Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. Genome 34:66–71

Dayanandan S, Bawa KS, Kesseli R (1997) Conservation of microsatellites among tropical trees (Leguminosae). Am J Bot 84:1658–1663

Dayanandan S, Rajora OP, Bawa KS (1998) Isolation and characterisation of microsatellites in trembling aspen (*Populus tremuloides*). Theor Appl Genet 96:950–956

DeCroocq S, Butcher PA, Moran GF (1997) Genetic diversity and conservation of microsatellite loci in acacias. Poster presented at Plant and Animal Genome V Conference, San Diego, Jan 12–17, 1997

Devey ME, Sewell MM, Uren TL, Neale DB (1999) Comparative mapping in loblolly and radiata pine using RFLP and microsatellite markers. Theor Appl Genet 656–662

Dow BD, Ashley MV, Howe HF (1995) Characterisation 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

Edwards KJ, Barker JHA, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques 20:758–760

Engels WR (1992) Amplify. Version 1.0. Department of Genetics, University of Wisconsin, Madison, Wisconsin

England PR, Ayre KJ, Whelan RJ (1999) Microsatellites in the Australian shrub *Grevillea macleayana* (Proteaceae). Mol Ecol 8:689–690

Grimaldi M-C, Crouau-Roy B (1997) Microsatellite allelic homoplasy due to variable flanking sequences. J Mol Evol 44:336–340

Hague XY, Litt M (1993) A study of the origin of ‘shadow bands’ seen when typing dinucleotide repeat polymorphisms by the PCR. Hum Mol Genet 2:411–415

Hamilton MB, Fleischer RC (1999) Cloned microsatellite repeats differ between 4-base restriction endonucleases. J Hered 90: 561–563

Harwood CE, Williams ER (1992) A review of provenance variation in growth of *Acacia mangium*. In: Carron LT, Aken KM (eds) Breeding technologies for tropical acacias. ACIAR Proc No. 37, ACIAR, Canberra, pp 22–30

Karhu A, Dieterich J-H, Savolainen O (2000) Rapid expansion of microsatellite loci in pines. Mol Biol Evol 17:259–265

Kijas JMH, Fowler JCS, Thomas MR (1995) An evaluation of sequence-tagged microsatellite-site markers for genetic analysis within *Citrus* and related species. Genome 38:349–355

Lagercrantz U, Ellegren H, Andersson L (1993) The abundance of various polymorphic microsatellite motifs differs betweeen plants and vertebrates. Nucleic Acids Res 21:1111–1115

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

Loridon K, Cournoyer B, Goubely C, Depeiges A, Picard G (1998) Length polymorphism and allele structure of trinucleotide microsatellites in natural accessions of *Arabidopsis thaliana*. Theor Appl Genet 97:591–604

Ma ZQ, Roder M, Sorrells ME (1996) Frequencies and sequence characteristics of di- tri- and tetra-nucleotide microsatellites in wheat. Genome 39:123–130

Maslin BR, Stirton CH (1997) Generic and infrageneric classification in *Acacia* (Leguminosae: Mimosoideae): a list of critical species on which to build a comparative data set. Bull Int Group for the Study of Mimosoideae 20:22–44

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 3:175–182

Moran GF, Muona O, Bell JC (1989) *Acacia mangium*: a tropical forest tree of the coastal lowlands with low genetic diversity. Evolution 43:231–235

Peakall, R., Gilmore S, Keys W, Morgant M, Rafalski A (1998) Cross-species amplification of soybean (*Glycine max*) simple-sequence-repeats (SSRs) within the genus and other legume genera. Mol Biol Evol 15:1275–1287

Pedley L (1986) Derivation and dispersal of *Acacia* (Leguminosae), with particular reference to Australia, and the recognition of Senegalia and Racosperma. Bot J Linn Soc 92:219–254

Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends Plant Sci 7:215–222

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) Non-mammalian genomic analysis. A practical guide. Academic Press, New York, pp 75–134

Rossetto M, McLauchlan A, Harriss FC, Henry RJ, Baverstock PR, Lee LS, Maguire TL, Edwards KJ (1999) Abundance and polymorphism of microsatellite markers in the tea tree (*Melaleuca alternifolia*, Myrtaceae). Theor Appl Genet 98:1091–1098

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. a laboratory manual, 2nd edn. Cold Spring Harbour Laboratory, Cold Spring Harbor, New York

Sim BL (1984) The genetic base of *Acacia mangium* Willd. in Sabah. In: Barnes RD, Gibson GL (eds) Provenance and genetic improvement strategies in tropical forest trees. Commonwealth Forestry Institute, Oxford and Forest Research Centre, Harare, pp 597–603

Stallings RL (1994) Distribution of trinucleotide microsatellites in different categories of mammalian genomic sequence: implications for human genetic diseases. Genomics 21:116–121

Steinkellner H, Fluch S, Turetschek E, Lexer C, Streiff R, Kremer A, Burg K, Glossl J (1997) Identification and characterization of $(GA/CT)_n$-microsatellite loci from *Quercus petraea*. Plant Mol Biol 33:1093–1096

Sun G-L, Salomon B, Bothmer RV (1998) Characterisation and analysis of microsatellite loci in *Elymus caninus* (Triticeae: Poaceae). Theor Appl Genet 96:676–682

Ujino T, Kawahara T, Tsumura Y, Nagamitsu T, Yoshimaru H, Ratnam W (1998) Development and polymorphism of simple sequence repeat DNA markers for *Shorea curtisii* and other Dipterocarpaceae species. Heredity 81:422–428

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-dA)*n* (dG-dT)*n* polymorphisms. Genomics 7:524–530

Whitton J, Rieseberg LH, Ungerer MC (1997) Microsatellite loci are not conserved across the Asteraceae. Mol Biol Evol 14: 204–209